BANCA D'ITALIA
EUROSISTEMA

Fiscal Policy and Growth

Seminari e convegni
Workshops and Conferences

July 2013

number 13

Earlier versions of the papers in this volume were presented at the Banca d'Italia workshop on "Fiscal Policy and Growth" *held in Perugia, S.A.Di.Ba., on 29-31 March, 2012. We wish to thank the staff of S.A.Di.Ba. for their assistance in all aspects of the organization of the workshop.*

Previous Public Finance workshops have addressed the following topics:

"Indicators of Structural Budget Balances" (1998)

"Fiscal Sustainability" (2000)

"Fiscal Rules" (2001)

"The Impact of Fiscal Policy" (2002)

"Tax Policy" (2003)

"Public Debt" (2004)

"Public Expenditure" (2005)

"Fiscal Indicators" (2006)

"Fiscal Policy: Current Issues and Challenges" (2007)

"Fiscal Sustainability: Analytical Developments and Emerging Policy Issues" (2008)

"Pension Reform, Fiscal Policy and Economic Performance" (2009)

"Fiscal Policy: Lessons from the Crisis" (2010)

"Rules and Institutions for Sound Fiscal Policy after the Crisis" (2011)

The proceedings may be requested from:

Banca d'Italia
Biblioteca
Via Nazionale, 91
00184 Roma
Italy

They are also available at the Banca d'Italia website:

http://www.bancaditalia.it/studiricerche/convegni/atti;internal&action=_setlanguage.action?LANGUAGE=en
and:
http://www.bancaditalia.it/pubblicazioni/seminari_convegni;internal&action=_setlanguage.action?LANGUAGE=en

Printed by the Printing Office of Banca d'Italia, Rome, November 2012.

CONTENTS

Session 2

GOVERNMENT BUDGETS AND POTENTIAL GROWTH

Session 3

TAXATION, REGULATION AND PUBLIC SERVICES

Session 4

POLICIES TO PROMOTE SUSTAINABLE GROWTH

Contributors

António Afonso *(ECB)*
Nikola Altiparmakov *(Fiscal Council Republic of Serbia)*
Milojko Arsić *(University of Belgrade)*
Jan Babecký *(National Bank of the Czech Republic)*
Stefan Bach *(DIW Berlin)*
Ray Barrell *(Brunel University)*
Salvador Barrios *(European Commission)*
Anja Baum *(University of Cambridge)*
Péter Benczúr *(Bank of Hungary)*
Carine Bouthevillain *(Banque de France)*
Adi Brender *(Central Bank of Israel)*
Francesco Caprioli *(Banca d'Italia)*
Cristina Checherita-Westphal *(European Central Bank)*
Sarah Ciaglia *(Ministry of Finance, Germany)*
Sergio Clavijo *(ANIF Colombia)*
Jérôme Creel *(OFCE/Sciences Po, France and ESCP Europe)*
Maria De los Ángeles Mignon *(Instituto de Economía y Finanzas, Universidad Nacional de Córdoba, Argentina)*
Francesco Di Comite *(European Commission)*
Gilles Dufrénot *(Banque de France)*
Werner Ebert *(Ministry of Finance, Germany)*
Balázs Égert *(OECD)*
Glenn Follette *(Federal Reserve Board)*
Daniele Franco *(Banca d'Italia)*
Ángel Gavilán *(Banco de España)*
Gabriele Giudice *(European Commission)*
Arlind Gjokuta *(Bank of Albania)*
Francesco Grigoli *(IMF)*
Fuad Hasanov *(IMF)*
David A. Heald *(University of Aberdeen)*
Pablo Hernández De Cos *(Banco de España)*
Dawn Holland *(National Institute of Economic and Social Research)*
Paul Hubert *(OFCE/Sciences Po, France)*
Ian Hurst *(National Institute of Economic and Social Research)*
John Janssen *(New Zealand Treasury)*
Juan F. Jimeno *(Banco de España)*

Gábor Kátay *(Bank of Hungary)*
Áron Kiss *(Bank of Hungary)*
Walpurga Köhler-Töglhofer *(Österreichische Nationalbank)*
Manmohan Kumar *(IMF)*
Julia Lendvai *(European Commission)*
Eduardo Ley *(The World Bank)*
Yngve Lindh *(Ministry of Finance, Sweden)*
Byron Lutz *(Federal Reserve Board)*
Maria Rosaria Marino *(Banca d'Italia)*
André Martínez Fritscher *(Banco de México)*
Sandro Momigliano *(Banca d'Italia)*
Gilles Mourre *(European Commission* and *Free University of Brussels)*
Nivedita Mukherji *(IMF)*
Jorge Onrubia-Fernández *(Universidad Complutense de Madrid)*
Hans Pitlik *(WIFO Austria)*
Jonathan Pycroft *(European Commission)*
Olivér M. Rácz *(Bank of Hungary)*
Agustín Ramello De la Vega *(Instituto de Economía y Finanzas, Universidad Nacional de Córdoba, Argentina)*
Ernesto Rezk *(Instituto de Economía y Finanzas, Universidad Nacional de Córdoba, Argentina)*
Carolina Rodríguez Zamora *(Banco de México)*
Juan A. Rojas *(Banco de España)*
Marzia Romanelli *(Banca d'Italia)*
Philipp Rother *(European Central Bank)*
A. Jesús Sánchez-Fuentes *(Universidad Complutense de Madrid)*
Francesco Saraceno *(OFCE/Sciences Po, France)*
Bert Saveyn *(European Commission)*
Margit Schratzenstaller *(WIFO Austria)*
Gerti Shijaku *(Bank of Albania)*
Douglas Sutherland *(OECD)*
Martino Tasso *(Banca d'Italia)*
Ingrid Toming *(European Commission)*
João Tovar Jalles *(ECB)*
Alessandro Turrini *(European Commission)*
Sergey Vlasov *(Bank of Russia)*
Jaejoon Woo *(IMF)*
Pietro Tommasino (Banca d'Italia)

FOREWORD

Daniele Franco[*]

This volume brings together the papers presented at the 14th Banca d'Italia Public Finance Workshop, held in Perugia from 29 to 31 March 2012.

The workshop focused on the issue of Fiscal Policy and Growth. In recent years this issue has come again to the fore. The global economy continues to recover, but growth remains uneven across countries. While it is strong in emerging markets, growth has been relatively weak in several advanced countries, with employment lagging. This adds to the challenge of fiscal adjustment following unprecedented government action to support the economy during the crisis.

In the short/medium term, the appropriate phasing out of fiscal stimulus measures will be crucial to sustaining the fiscal consolidation effort. Over a longer horizon, as fiscal space remains tight in industrial countries owing to unfavourable demographics and high public debts, the composition of government budgets will be a key policy variable. The quality of regulation and effectiveness in the provision of public goods and services will also play an important role. The need to ensure the sustainability of economic development impinges on an array of public policies and issues, ranging from environment protection to income redistribution, from the design of fiscal frameworks to the creation of an innovation-friendly economic context.

The workshop provided an overview of the recent theoretical and empirical work on the link between public policies and economic growth. It examined the short-term effects of fiscal policy, the size of short-term multipliers of individual budget items, comparing alternative strategies to increase government revenue and decrease spending. It also examined the link between the structure of government budgets and long-term growth, considering which tax reforms can be growth-enhancing and how the level of taxation and the relative weights of income, consumption and wealth taxes affect potential growth. It assessed the contribution to growth of public spending for the formation and preservation of human and physical capital, as well as the role of regulation in facilitating business activity. It examined how reforms in the management of such public services as education, health care and justice can improve their quality and augment potential economic growth. Finally, it evaluated the design of fiscal frameworks that can increase economic stability and the policies that can foster innovation and social cohesion.

Banca d'Italia is grateful to the institutions that contributed to the success of the initiative, to the experts who provided research papers and to all who came to Perugia to take part in the discussion.

This volume extends the analysis of fiscal policy issues carried out in the previous workshops, which were devoted to *Indicators of Structural Budget Balances* (1998), *Fiscal Sustainability* (2000), *Fiscal Rules* (2001), *The Impact of Fiscal Policy* (2002), *Tax Policy* (2003), *Public Debt* (2004), *Public Expenditure* (2005), *Fiscal Indicators* (2006), *Fiscal Policy: Current Issues and Challenges* (2007), *Fiscal Sustainability: Analytical Developments and Emerging Policy Issues* (2008), *Pension Reform, Fiscal Policy and Economic Performance* (2009), *Fiscal Policy: Lessons from the Crisis* (2010) and *Rules and Institutions for Sound Fiscal Policy after the Crisis* (2011).

[*] Banca d'Italia, Economic Research and International Relations Area.

INTRODUCTION

Maria Rosaria Marino,[*] *Martino Tasso*[*] *and Pietro Tommasino*[*]

The papers included in this volume provide an overview of recent work on the issue of fiscal policy and growth.

This is, of course, a vast and multi-faceted topic and the various parts of the conference where only able to discuss some of the aspects involved.

A first issue is the link between fiscal policy and short-term growth. This boils down to the hotly debated topic of the role of discretionary policy actions for macroeconomic management and, more technically, to the size of "fiscal multipliers". In the pre-crisis years, a certain consensus emerged, according to which monetary policy was to be preferred as a countercyclical tool to fiscal policy, mainly due to the long lags and to the not-completely-understood channels through which fiscal stimulus hits the economy. The recent crisis, however, has shaken this assumption. Indeed, conventional monetary policy in some countries reached its limits, while the length and the depth of the recession made the issue of implementation lags less pressing. In such a juncture, research on fiscal multipliers was also breathed new life into. While we still lack a new consensus, most economists would probably agree that multipliers are very much context- and time-specific. They are a function of the fundamentals of the economy (for example, the state of public finances) and of the business cycle.

A second issue is the relationship between fiscal choices and long-run growth. This issue has an extensive tradition in public finance studies, even if the recent crisis shifted the emphasis on the effects of fiscal policy at business-cycle frequencies. Fiscal variables influence long-run performance through several channels. First, sound public finances are a crucial element of a stable and predictable macroeconomic environment, which, in turn, is a precondition for growth. Second, expenditure and taxation policies influence individual behaviour: sometimes they are useful in correcting market failures, sometimes they induce distortions and discourage labour supply, therefore putting growth at risk.

Furthermore, as in modern economies the public sector is one of the biggest service provider, its efficiency and effectiveness are crucial issues, also in light of tighter budget constraints.

Finally, it is important to remark that governments have an impact on the economy (both in the short and long run) not only through the budget but also through regulation.

The papers presented at the Workshop were divided among four sessions, which correspond to the sections of the present volume. Session 1 examines the short-term impact of fiscal policy; Session 2 is about the link between government budgets and potential growth; Session 3 concentrates on taxation, regulation and public services; and in Session 4 policies to promote sustainable growth are discussed.

1 The short-term impact of fiscal policy

The contributions in Section 1 deal with the impact of fiscal policy on the macro-economy. The papers use a variety of methods, and touch both positive and normative ones. They differ both with respect to the dependent variable of interest and for the way in which fiscal action is measured. The first paper uses a theory-based economic model to assess the size of fiscal

[*] Banca d'Italia, Structural Economic Analysis Department, Public Finance Division.

multipliers as a function of the features of the economy. The second focuses on the impact of discretionary fiscal action on aggregate consumption, merging original data on the size and composition of policy measures with previous estimates of fiscal multipliers. The third paper is also empirical, but focuses on the impact of fiscal policy on unemployment and job market flows. The fourth paper has a more normative focus and discusses optimal fiscal policy in a context in which the economy's trends are summarized by a relatively parsimonious VAR. The fifth paper discusses a more subtle issue: are the effects of fiscal policy stronger in recessions than in booms? To answer the question, it relies on relatively new time series models, estimated on Italian data. The last paper of the section discusses in depth a single case study, that of Latvia during the recent crisis, and highlights the reasons for the success of the Latvian fiscal consolidation.

Ray Barrell, Down Holland and Ian Hurst use the National Institute Global Econometric Model (NiGEM) to assess the size of fiscal multipliers in different countries under different scenarios. NiGEM is an open-economy macro-econometric model which incorporates forward-looking aspects. Among its key features are: a consumption equation which depends on future income streams and takes borrowing constraints into account; rational expectations in financial markets; and flexible exchange rates. The NiGEM features several fiscal variables: it distinguishes between corporate and personal income taxes and between direct and indirect taxes; on the spending side, it differentiates between public investment, public consumption, and transfers. To ensure public sector solvency, NiGEM assumes that the income tax rate rises (resp. decreases) if the deficit-to-GDP ratio is above (resp. below) an exogenously-given target level. The authors use the model (which is calibrated in part and partly estimated) to assess the effects of a reduction in deficit by 1 per cent of GDP in 18 advanced economies. It turns out that multipliers tend to be smaller in more open economies, in smaller countries, and if the short-run elasticity of consumption is smaller. Moreover, multipliers for government consumption tend to be larger than those of other budget items. If the fiscal action is temporary (*i.e.*, it is done keeping fixed the target level of the deficit) multipliers are greater than if the action is permanent (*i.e.*, if the target level of the deficit is permanently reduced by 1 per cent of GDP as well).

Glenn Follette and Byron Lutz aim at measuring the impact on aggregate demand of countercyclical fiscal policy in the US during the post-WWII period. To do this, as a first step the authors identify discretionary policy actions using a variety of public sources and documents (the so-called "narrative approach"). Contrary to other narrative studies, they do not drop the measures which were not explicitly motivated by countercyclical reasons. For each policy episode, Follette and Lutz assess the size of the change induced with respect to previous legislation, keeping track of which budgetary item was used. It turns out that discretionary fiscal policy was mostly counter-cyclical. As a second step, Follette and Lutz multiply the size of each policy change for an estimate of the associated multiplier, taken from the existing literature and from the FRB/US macro model (different budgetary items are assigned different multipliers). In this way, they are able to show that, on average, a 1 percentage point increase in the deficit for two years boosts demand by 0.4 percent of GDP in the first year and 0.6 percent of GDP in the second year. This figure is relatively small, mainly because countercyclical policy in the US has been mainly pursued through tax cuts, which have relatively small multipliers.

Alessandro Turrini estimates the impact of fiscal consolidation on short- and long-run unemployment, job creation, and job destruction in 13 EU countries for the years 1978-2009. To build its explanatory variable, he uses both the "narrative approach" (as do Follette and Lutz) resorting to a recent database of consolidation episodes built by the IMF, and a more standard cyclically-adjusted deficit measure. In this latter case, the fiscal consolidation measure is instrumented and only episodes in which the change in the parameter is above 0.5 percentage points of GDP are considered in the second stage. Results show that fiscal consolidations have a significant impact on cyclical unemployment. However, the effect is not large (about

0.1 percentage points), mostly due to expenditure-based consolidations. Results are mostly robust if one substitutes narrative measures of fiscal action to cyclically-adjusted measures. To shed light on the interaction between fiscal policy and labour market regulation, Turrini also runs his regressions separately for countries with high and low employment protection legislation (EPL), *i.e.*, countries with an average value of the OECD Employment Protection Indicator above (resp. below) the EU27 median. It turns out that the effect on unemployment is not different in the two sub-samples. However, while in low-EPL countries the effect operates through an increase in the job destruction rate, in high-EPL countries it operates through a decrease in job creation. Since a reduced job-finding rate corresponds to a longer average duration of unemployment spells, in high-EPL countries fiscal policy shocks also tend to raise the share of long-term unemployment.

Francesco Caprioli and Sandro Momigliano use Vector Auto Regression (VAR) models to capture the influence of the state of the economy on the effects of public expenditures shocks. Their focus is on the Italian economy over the period 1982-2011. They use quarterly fiscal variables recorded on a cash basis, and identify exogenous fiscal shocks using the methodology originally developed by Blanchard and Perotti (2002). In addition to variables which are standard in the literature (private GDP, inflation, interest rates, net revenue and government consumption), the authors include in their VARs government debt and foreign demand. To take into account the influence of the state of the economy on the effects of public expenditure shocks, the authors follow three approaches. First, they estimate a standard structural VAR model over the two sub-samples identified as "recessions" and "expansions" by the Italian statistical authorities. Second, they estimate an Endogenous Threshold VAR (ETVAR) using alternatively lagged private GDP growth and the output gap as business cycle indicators. With this approach, whether an economy is in a recession or in a boom is determined endogenously. Finally, the paper considers a Smooth Transition VAR (STVAR), where the probability of transition between booms and recessions is a continuous function of a business cycle indicator. The main results are the following. Without distinguishing across regimes, the response of private GDP to an expenditure shock is positive, hump-shaped and highly significant for approximately two years. The median value of the expenditure multiplier is equal to 1.04 on impact and reaches its peak (1.8) after three years. Furthermore, when the split-sample methodology is adopted, it emerges, as expected, that the effect of public expenditure shocks is larger in recessions than in booms. However, this difference is no more statistically significant when the ETVAR model is used. The difference becomes even less clear-cut when the STVAR model is estimated. These mixed results may be due to the limited size of the two sub-samples, and/or to the fact that most recessions in the sample are quite mild.

Francesco Di Comite, Gabriele Giudice, Julia Lendvai and Ingrid Toming discuss the severe fiscal consolidation engineered in Latvia in the period 2009-11 and its effects on growth. They notice that the size of the fiscal effort as measured in the government reports (as against an unchanged-policies scenario) is huge, and much bigger than if measured by the change in the cyclically-adjusted primary balance. They argue that this is mainly attributable to shortcomings of the latter indicator which, as is well known, does not take into account the composition in the drop of GDP matters for the effects of the cycle on revenues, nor considers fully the cyclicality of expenditures. They go on to describe the composition of the manoeuvre, which was mostly expenditure-based. Concerning the timing of consolidation, they argue that it was quite frontloaded, as the bulk of the adjustment was legislated and entered into force already in the second half of 2009. After the description of the size, timing and composition of the Latvian fiscal consolidation, the authors provide DSGE-based estimates of what should have been the effect of the package implemented in Latvia. They use a quite standard open-economy new-Keynesian model. Interestingly, it turns out that the model implies a much sharper and long-lasting contraction than the one observed in the data. According to the authors, this discrepancy means that non-Keynesian mechanisms were at play. This is suggested by the fact that consumer confidence indicators increased, and government bond yields decreased, soon after the main part of the package was implemented, therefore closely mimicking the recovery in output. From a policy perspective, the Latvian experience suggests that a quick, sizable and expenditure-based consolidation might be the best option when facing an unsustainable macroeconomic and fiscal status quo.

Jan Babecký comments on the first two papers. He argues that both contributions highlight the large variation in reported multipliers, be it across countries (the first paper) or over time (the second one). Difficulties in obtaining reliable estimates of fiscal effects obviously limit the role of economic analysis as guidance to policy. According to Babecký, a possible way out would be to explore this variation employing the methods of quantitative review of literature (Meta-Regression Analysis). Apart from understanding the reasons behind differences in the estimate of multipliers, Meta-Regression Analysis can also help identifying the "best-practice" specification. He also suggests, as further directions for future research, topics such as the role of debt sustainability expectations, the impact of consolidation on risk premia, and fiscal stress testing.

Adi Brender discusses the papers by Momigliano and Caprioli. Concerning the former contribution, he highlights four potential directions for improvement. First, the authors should tell if their framework ensures that the inter-temporal budget constraint of the government is always satisfied. Second, they could consider a more flexible model which allows for a change in the relationship between fiscal policy and interest rates when going from the pre-Euro to the Euro period. Finally, it should be investigated whether the response of short-term interest rates reflects only monetary policy, or also the sentiment of investors and whether there is an asymmetry between the consequences of negative and positive fiscal shocks.

Walpurga Köhler-Töglhofer reviews the papers by Di Comite, Giudice, Lendvai and Toming and by Arpaia and Turrini. As discussed above, Di Comite *et al.* argue that non-keynesian effects of fiscal consolidations are more likely to emerge if the fiscal retrenchment is large, frontloaded, and expenditure-based. Köhler-Töglhofer suggests to add to the list of relevant factors ownership, commitment, and fairness. Moreover, she questions the assertion that the Latvian adjustment was relatively painless, given that from 2007 to 2010 Latvia's GDP fell by about 25 per cent. Concerning the paper by Arpaia and Turrini, she argues that, given the absence of a fully-fledged theoretical model, one should be very careful in drawing policy implications from the authors' empirical findings. Moreover, she urges to consider in their analysis not only the employment protection legislation regime, but also other institutional features of the labour markets, such as unemployment benefits.

2 Government budgets and potential growth

Session 2 contains a series of studies on the relationship between government policies and economic outputs. The first two papers are econometric studies on the topic of the impact of public debt on growth, while the following three combine theory with data to illustrate different aspects of the implications of public policies for the potential growth of the economy. The session is concluded by a case-study on the effects of budget policies in Albania.

The paper by Anja Baum, Cristina Checherita-Westphal and Philipp Rother deals with the highly-debated topic of the relationship between government debt and growth. The authors focus on the euro area and on the 1990-2010 period. The study shows that additional debt has a positive effect on growth only if the stock of debt is relatively low, that is, when it is below 67 per cent of GDP. When the incidence of debt over product is very high (above 95 per cent), additional government debt is instead associated with negative growth. Moreover, Baum *et al.* claim that their results can be explained by the increased pressure on long-term interest rates which is usually associated with high levels of public debt. The authors conduct a series of robustness checks, which, for the most part, support their main findings: they expand their baseline specification with more covariates, including the 1980-1989 period and controlling for possible endogeneity problems and for the effect of excluding outliers. Unlike other studies in this particular area of public finance, this paper uses a novel dynamic threshold panel methodology by extending previous econometric work to the case of panel data and applying it to this topic for the first time.

Manmohan Kumar and Jaejoon Woo study the relation between initial public debt and subsequent growth rates in a sample of advanced and emerging economies over the 1970-2008 period. The authors use a variety of reduced form econometric techniques, such as pooled OLS, between estimator, fixed effects, and system GMM. In the main specification, they find that a 10 percentage point increase in the initial debt-to-GDP ratio is associated with a reduction of real per capita growth of about 0.2 percentage points per year. The authors find that their results are broadly robust to a series of specification tests. Using a model in which initial debt is interacted with dummies for ranges of initial debt, they find some evidence of a non-linear effect on the growth rate of the economy: debt-to-GDP ratios above 90 per cent are more strongly associated with subsequent slower growth. Finally, Kumar and Woo explore the determinants of the negative relationship between debt and growth: using a growth-accounting exercise, they find that higher initial public debt is associated with significantly weaker labor productivity growth, mainly due to a reduction in domestic investments.

Maria Rosaria Marino, Marzia Romanelli and Martino Tasso present some preliminary results from an ongoing project aimed at building a dynamic model of family labor supply to be used to conduct policy-relevant analysis for the Italian economy. In this version of the model, agents decide about female labor supply and asset accumulation, taking into account the main features of the Italian tax-and-benefit system. The model, which is estimated using both longitudinal and cross-sectional data for the 2004-10 period, replicates quite closely what is observed in the data in terms of female labor supply. This version of the study presents preliminary results for a small set of policy experiments: an increase in households' non-labor income decreases the overall poverty rates but lowers the incentives of married women to participate in the labor market. On the contrary, policies aimed at increasing the return of the hours worked (such as an increase in the amounts of individual work-related tax credits) have positive effects on both dimensions. The authors conclude with an agenda for ongoing and future work.

Nivedita Mukherji and Fuad Hasanov build a two-sector model of production to analyze the effect of public policies on informality: alongside the informal sector, the authors model formal firms which can produce either formal or informal goods. This paper finds an unconventional result: higher taxes may actually reduce informality. When both tax evasion and the payment of bribes by informal firms are allowed, higher taxes on formal firms increase the relative price of formal goods; as a consequence, there would be a reallocation of producers towards their production. Unlike previous literature on the relation between taxes and informality (which mainly focuses on reallocations between formal and informal sectors), this paper shows how public policies may impact the distribution of production within formal sectors too. The authors complement their theoretical analysis with a set of empirical tests based on cross-country data and focus on testing whether public policies which are responsible for promoting informality are associated with an increase in net revenue as well. They show that it is possible to increase net revenue by having higher taxes, a larger government, and a higher number of regulations. Almost all of these factors are found to be associated with an increased informality as well.

António Afonso and João Jalles develop a growth model that accounts for the role of the government: in it, the higher the level of resources devoted to financing the public sector, the lower the optimal level of private consumption and output per worker. Afonso and Jalles test the implications of the model using an unbalanced panel data of 108 countries for the period

1970-2008. They find that government size is negatively related to growth and that institutional quality (measured with a set of different proxies) has a positive effect on GDP per capita (this effect is weaker for countries of Scandinavian legal origin). Government consumption is found to be negatively associated with growth in all the groups of countries considered in the study. Moreover, the negative effect of government size on growth turns out to be even stronger in those countries with lower levels of institutional quality. These findings are robust to a set of checks.

Gerti Shijaku and Arlind Gjokuta analyze the relationship between fiscal policy and growth in the case of Albania. To this end, they derive some testable implications from an endogenous growth model for a small open economy. Their paper is divided in two main parts: in the first one, the authors review the Albanian fiscal policy in the last decades, while in the second they turn to their methodology and to the analysis of their econometric estimates. They find that economic growth is more responsive to changes in revenues than to changes in expenditures. Moreover, the authors find that the responsiveness of the growth rate to fiscal reforms varies according to the type of tax which is involved: one percentage point increase in the revenue collected by distortionary taxes would slow the growth rate by about 0.64 percentage points, whereas if the same resources were collected through less distortionary taxes, the growth rate would decrease by just 0.13 percentage points. Similarly, productive expenditures have a much stronger impact on growth than non-productive ones. The authors derive some policy advices from their study: in particular, should a tax increase be required, operating over indirect taxes would be preferable.

John Janssen comments on the first paper of the session, *i.e.* the work by Baum *et al.* on the relationship between government debt and growth in the Euro Area. Janssen appreciates the methodology which is used to estimate the threshold in the debt ratio over which additional debt becomes detrimental for economic growth. He points out, though, that the source of the increase of public debt may play a role too: additional debt incurred to finance productive investments may be quite different from that incurred to finance consumption. Janssen comments on the study by Kumar and Woo as well. While he appreciates both the methodology used and the robustness checks conducted by the authors of the paper, he suggests a series of possible extensions. In particular, he calls for a sensitivity analysis in which the sample period is extended over the years affected by the global financial crisis. He also suggests that a research on whether the maturity structure of debt plays a role could be an interesting research topic. Janssen concludes his comments on these two studies on the debt-to-growth relation by illustrating the role played by public debt in the New Zealand government balance sheet.

Finally, two papers in this session which deal with the topic of the effects of taxation (Marino *et al.* and Mukherji and Hasanov) are commented by Gilles Mourre. He starts by recognizing the importance of the fiscal system in explaining both labor force participation and informality. He then advises the authors of the first paper to explore the possibility of incorporating into the model a few additional features such as a more heterogeneous utility function and a more detailed treatment of childcare services and of the fixed costs of work. Finally, he observes that the model by Mukherji and Hasanov assumes perfect labor mobility, which he judges to be quite a strong hypothesis. As regards their empirical analysis, he underlines that the number of observations may be low and suggests a series of possible solutions. Mourre suggests to the authors of this paper to check whether their results are still valid in a sample of euro-area countries, and to use other variables rather than the top marginal personal income tax rates to measure tax pressure.

3 Taxation, regulation and public services

While all the papers discussed in the third session deal with the economic impact of the structure of the public sector, their approaches and methodologies vary. The session is opened by a theoretical study on the cost of government inefficiencies. It is followed by a descriptive study on the structure and size of public sectors and by a reduced-form cross-country econometric paper on service regulation. The fourth paper, on the topic of the efficiency costs of different forms of taxation, derives its results from a computational general-equilibrium model. The fifth paper of the session raises the concern that the rankings of countries by living standards would change quite dramatically if government inefficiencies could be quantified. The second-to-last paper of this session is a static labor supply model, while the last one is an *ex post* difference-in-difference study of the effects on public health of an institutional reform which took place in Mexico in 1997.

The work by Jorge Onrubia-Fernández and Jesús Sánchez-Fuentes deals with the topic of the cost of public-sector inefficiencies. The possible budgetary savings related with the improvement of the productive efficiencies of the government can indeed constitute an alternative fiscal policy tool. These savings can be sizeable: the OECD estimates that the gradual adoption of best practices in primary and secondary education could save resources for around 0.5 per cent of GDP, while improvements in the health sector could potentially lead to savings of the order of 2 per cent of GDP. This paper provides a theoretical framework to quantify the social welfare changes which derive from variations of public-sector performances: the authors of this paper derive one measure from the cost function and one from the production side of the economy. The authors claim that their approach could be adapted to be used in a variety of empirical applications with the aim of monitoring the performance of the government.

The composition of public expenditure and the tax structure can both influence the growth level of a country. The study of Hans Pitlik and Margit Schratzenstaller deals with this issue by analyzing the growth-friendliness of fiscal and regulatory structures in a cross-section of EU and highly developed OECD countries. On the basis of the indications of economic theory and previous empirical studies, the authors divide public expenditures between "productive" and "non-productive" ones: core public services, infrastructure spending, and expenditures for health and education services are among the former, redistribution, culture, religion, and interests payments among the latter. The authors then rank countries on the basis of the share of expenditures in 2004-08 which are considered productive: Korea, New Zealand, and Ireland are on top of the list, while Austria, Greece, and Germany are considered to have the least productive expenditure mix. A similar approach is used to rank countries with respect to several indicators related to their tax structure, to their approach to business regulation, and with respect to an average indicator about the "growth-friendliness" of their policies (in this case, New Zealand, and Korea still lead the ranking, while Germany, Italy, and Greece get the lowest marks).

Guglielmo Barone and Federico Cingano study the effect of service regulation on the performance of downstream manufacturing activities, using a panel data of OECD countries in the 1996-2002 period. They examine whether countries with a lower level of service regulation in 1996 saw faster value-added productivity and export growth in those manufacturing industries which used services more intensively. The authors rely on OECD indicators for anti-competitive regulatory frameworks for the energy sector, telecommunications, transportation, and professional services. They find that service regulation plays a non-negligible role in explaining subsequent growth in the manufacturing sector: in particular, this study suggests that strongest gains from deregulation would come from removing restrictions to price setting in professional services and from unbundling power generation from distribution in the energy sector.

Salvador Barrios, Jonathan Pycroft, and Bert Saveyn study the relative magnitude of the economic distortions imposed by different kinds of taxes. Given the need for fiscal consolidation in many European countries, the authors contrast the marginal cost of public funds (MCF, that is the costs imposed on the economy by levying an extra euro in tax revenue) of labor and environmental taxation ("green taxes"). Using a computational general-equilibrium model for the EU (the so-called GEM-E3), they find that the distortions provoked by a labor tax hike are larger than those related to green taxes: on average, raising 1 euro by taxing labor would result in a loss of efficiency of about 90 cents; on the other hand, this loss is quantified at around 8 cents in the case of environmental taxes. Even though these figures vary across countries, in any member state the MCF of labor taxes is higher than that of green taxes. The authors then study the robustness of their findings and show that, once spillovers effects between countries are accounted for, the advantage of green taxes over labor ones is weaker. They also find that their results are sensitive to different assumptions on the flexibility of the labor market. The authors leave the topic of the progressivity of the different forms of taxation to future work.

National account systems equate output to input costs when evaluating the value of government production. Since many countries are affected by large inefficiencies in the production of government services, Francesco Grigoli and Eduardo Ley argue that this method can lead to severe distortions when countries are compared in terms of GDP per capita. Purging GDP from the resources which are wasted in the provision of government services, the authors are able to build an alternative measure of living standards. For this, they rely on a series of pre-existing empirical studies which quantify "waste" in public sectors in several countries: in a cross-section of 24 countries, this loss is about 4.1 per cent of GDP on average, but it displays large variability. The results of the study indicate that the rankings of countries by standard of living would change dramatically if this correction were to be taken into account.

The paper of Péter Benczúr, Gábor Kátay, Áron Kiss and Olivér M. Rácz estimates the effect of income tax and welfare transfers on the labor force participation in Hungary. The authors use pooled cross-sectional data for the 1998-2008 period taken from the Hungarian Household Budget Survey. They define as "gains-to-work" the algebraic sum of lost welfare benefits and acquired salary which comes with the transition from unemployment to full-time work. This study finds that participation probabilities are strongly influenced by this variable: a 10 per cent increase in the gains-to-work increases the probability of being active by 2.9 per cent. Moreover, the strength of this effect is heterogeneous in the population: people around retirement age, married women, and women at child-bearing age exhibit larger elasticities. Finally, the authors use their model to simulate the effects of the 2012 Hungarian reform of the tax and benefit system, which eliminated an employee tax credit, cut the tax rate below a certain income threshold, and raised social security contributions by one percentage point. The results are heterogeneous in the population, but this study finds that this reform will have a slightly negative effect on aggregate activity (a decrease of about 1 percent).

André Martínez Fritscher and Carolina Rodríguez Zamora analyze the effects of the 1997 health sector reform in Mexico which transferred both financial resources and responsibilities from the central government to the states. This paper therefore falls within the large literature on fiscal decentralization. The authors exploit the specific features of the reform to evaluate its impact on several indicators of health status. The findings of the study indicate that the decentralization reform did not increase the overall efficiency of the provision of health services. Moreover, states which received higher transfers after the reform did not perform significantly better that the others. The authors suggest that these results could be explained by the sudden implementation of the reform on one hand, and on the lack of a funding allocation mechanism which encouraged the adoption of best practices by the states on the other. The authors conclude by arguing that, on the basis of their study, a successful decentralization is based on some requirements, such as revenue collection decentralization, and an improvement of transparency at the state level.

Stefan Bach comments on three papers presented in this section. He first analyzes the work by Onrubia-Fernández and Sánchez-Fuentes on the cost of public-sector inefficiencies. Bach recognizes the importance of this topic, given the potential for both budgetary savings and indirect welfare gains. He calls for an extension of the study to the area of pure public goods and their financing and he points out the difficulty of obtaining reliable data for the empirical analysis of this important topic. Bach then turns to comment the work by Grigoli and Ley: in this case as well, he stresses the need for better and more detailed data. He concludes by analyzing the study by Martínez Fritscher and Rodríguez Zamora on the impact of the 1997 Mexican reform on public health outcomes. His comments revolve around two main points: first, it would be interesting to study the impact of the reform on long-term indicators as well, and, second, the reasons for the ineffectiveness of the reform should be addressed in greater detail.

Sergio Clavijo comments on both the work by Barone and Cingano and the one by Pitlik and Schratzenstaller. As regards the former, he encourages the authors to extend their research in two directions: the study of a possible unbundling in the health sector between the provision of insurance and the provision of health services and the extension of the dataset to a sample of developing countries. He agrees with the main conclusion of the paper about the possible harmful consequences on the growth of an economy of added regulation in services, with the caveat that this result should be applied with caution to other industries. As for the second paper, Clavijo suggests the adoption of a theoretical model to better justify the categorization of expenditures between "productive" and "non-productive"; he also calls for the use of effective tax burden indices rather than marginal tax rates to rank countries on the basis of the distortions due to the tax system. Finally, he asks the authors for a deeper analysis of the relationship between the score assigned to different countries in 2008 and their subsequent different reaction to the most recent financial crisis.

Commenting the paper by Benczúr *et al.*, Yngve Lindh recognizes the importance of the design of taxes and transfers for labor force participation and thus for growth. He recalls that recent reforms which increased the "gains-to-work" in Sweden were found to have a positive effect on labor supply there too. He asks the authors for a more detailed description of the reforms to the tax and benefit system which took place in Hungary in the period analyzed and suggests to possibly consider the effects of the reforms on labor demand as well. Lindh lists a series of very detailed comments on the paper by Savey *et al.* too. In particular, he asks for a more accurate description of the model (in particular, about the mechanisms beyond the spillover effects), calling for an extension of the study to other kinds of taxes.

4 Policies to promote sustainable growth

The papers presented in Session 4 focus on how reforms and fiscal policies might promote sustainable growth. The contributions are very different from each other: from empirical to theoretical, plus three case studies concerning Argentina, Spain and Serbia.

The paper by Douglas Sutherland focuses on the implications of reducing debt levels for growth in the short and in the long term. Overall, the link between economic growth and the post-crisis debt overhang is complicated. On the one hand, high debt seems to be associated with lower growth. On the other, however, fiscal consolidation may weaken growth both in the near term and over a longer horizon. Realistically, debt problems are so serious in many countries that consolidation has the potential to strongly hamper growth. In the short run, consolidation may weaken demand and monetary policy may not be able to compensate for such effects for some time to come. This argues for the necessity of phasing in consolidation. Appropriate and clear fiscal objectives together with institutions that ensure accountability may help to preserve credibility in the process. However, to maintain it, it may also be necessary to take some action up-front, in which case instruments with small short-term multipliers may be given some weight. This may involve some political economy risks, in that this may skew consolidation because of inappropriate instruments. Slow consolidation may also entail a price insofar as it involves a higher debt and thereby higher interest rates. In the longer run, the effects of consolidation on growth will depend on the choice of instruments. Some instruments are available that will have limited detrimental impact on growth and enter in little or no conflict with other policy objectives. Notably, increasing spending efficiency, reforming unsustainable pension systems, putting prices on environmental externalities and maximising the benefits of structural reforms could make sizeable contributions to consolidation. In addition, reviewing tax and benefit systems within a wider horizon could help determine how policy objectives could be achieved at a lower cost and where, instead, support is less justified.

The essay presented by Ernesto Rezk, Maria De los Ángeles Mignon and Agustín Ramello De la Vega aims at assessing, using an Augmented Solow Model, the impact on GDP of the investment in education in Argentina. To this end, the author uses a proxy for the propensity to invest in human capital accumulation consisting in the percentage of the working age population enrolled in secondary school. In connection to this, one of the main contributions of the paper is having improved three aspects of the standard model: (a) finding a better representation for the average propensity to invest in human capital; (b) missing components, such as the opportunity costs incurred by parents and students, are added to all government and educational levels' budgetary expenditures; (c) a methodology has been developed for the measurement of the stock of human capital such that the variable can be used in a second stage of the analysis in place of human capital accumulation rate. Given the econometric problems caused by the variables' non-stationarity, the author discarded the usual estimation procedures and used alternative approaches, such as cointegration and the error correction model, including lags and dummies. Results point to the existence of long-run equilibrium relations among variables; the coefficients showed the expected signs and were, in all cases, significantly different from zero. Moreover, although the formation of human capital grew substantially during the analysed period, there didn't seem to exist a clear relationship between the characteristics and effectiveness of spending programmes and the needs of the country's productive technological matrix. As for the link between human capital formation and economic growth, the author shows that either human capital did not help enhancing Argentine growth or its effect was negligible. Rezk finds one of the possible reasons for that in a design of public policies in this field inefficient and ineffective to obtain an adequate contribution of human capital to GDP.

The paper by Ángel Gavilán, Pablo Hernández de Cos, Juan F. Jimeno and Juan A. Rojas concentrates on Spain and uses a large overlapping generations model of a small open economy featuring imperfect competition in the labour and product markets to understand which were the main determinants of the large expansionary phase experienced in that country from the mid-1990s until the arrival of the global financial crisis in 2007-08, what role fiscal policy and structural reforms could have played to avoid the build-up of large external imbalance over this period, and how these policies could affect the recovery of economic activity after the crisis. The authors find that falling interest rates and demographic changes were the main drivers of the Spanish expansionary phase and that, over this period, a tighter fiscal policy or structural reforms designed to foster competition in the labour and product markets could not have avoided the build-up of a large external imbalance. Concerning the macroeconomic aspects, the model is able to reproduce the trade-off faced by tighter fiscal policies after the crisis, *i.e.*, they may reduce output losses induced by the crisis in the medium term, but at the expense of (mild) output losses in the years immediately after the crisis. On the contrary, structural reforms do not face this trade-off and may contribute to reduce output losses in the short and medium term, while inducing a positive long-run effect on the level of output.

The paper by Carine Bouthevillain and Gilles Dufrénot argues against the widespread view at the European level that, in order to get out of the economic depression while maintaining the sustainability of public finances, the EU countries should implement common fiscal policies. The authors argue that higher growth rate in the EU cannot be achieved with the same fiscal mix in all member states and this view is based on quantile estimates showing heterogeneous reactions across EU countries. They claim it is important to distinguish between member states which were part of the EU since the beginning and the emerging countries which entered the EU in the early 2000's. Social security spending, direct taxes, welfare and sovereign expenditure and human capital expenditure have strikingly different effects on different countries' real GDP growth rate. An increase in human capital spending is growth-enhancing in industrialized EU countries, while welfare and sovereign expenditure play a more important role in fostering growth in emerging economies. Direct taxes exert a much more detrimental impact in countries that are growing rapidly than in those that experience a slow growth. When the growth rate is considered in per-capita terms, indirect taxes appear to exert an asymmetric effect on EU countries: they are harmful in the low-growth countries, but not inconsistent with stronger growth dynamics in countries that grow rapidly. Direct taxes are growth-enhancing if an economy has either a slow or fast growth rate and are neutral at moderate growth rates. The authors suggest that an implication of these results is that, analyzing growth-friendly fiscal policies, it may not be helpful to use average fiscal multipliers. It is necessary, instead, to consider the different growth impacts normal times and in times of crisis and to acknowledge the different ways in which the same policies can affect growth rates in different countries. This rules out the use of a single fiscal/growth model for all EU countries.

The paper by Jérôme Creel, Paul Hubert and Francesco Saraceno evaluate the macroeconomic impact of three different fiscal rules that have been, will, or might be implemented in Europe: a balanced (at 0.5 per cent of GDP) structural budget and constant debt reduction rule established by the Treaty on Stability, Coordination and Governance in the Economic and Monetary Union (known as Fiscal Compact), the 3 per cent ceiling on public deficit and an investment rule in the vein of the United Kingdom golden rule of public finances. The authors simulate a small-scale New Keynesian model with both forward and backward expectations and the calibration draws on the existing literature and on the 2011 values of public finance data for four euro-area countries which are taken as representative of the different types of eurozone member states. The authors focus on two different scenarios: the first involves assessing the path followed by the four economies under each fiscal rule during the fiscal consolidation from 2011 debt and deficit levels, towards the Maastricht steady state. The second assesses the impact of demand and supply shocks affecting the economy at the steady state. Results are manifold. First, abiding by the rules produces in all cases a short-run recession, even in a country with a small fiscal multiplier and a low initial public debt like the Netherlands. Second, during a consolidation phase, the investment rule performs better than the other rules, *i.e.*, the recession is milder and shorter, thus leading to a substantially lower average output loss over a 20-year horizon. Third, if the economy is hit by a demand or supply shock at the steady state, none of the rules emerges as superior in coping with them. Fourth, the Treaty on Stability, Coordination and Governance in the Economic and Monetary Union, with its constant debt reduction rule, generally imposes large costs to the economy, while not necessarily improving public finances' sustainability.

Balázs Égert recognizes that governments and central banks of developed countries swiftly reacted to the 2007-08 financial and economic crisis by implementing substantial fiscal and monetary policy easing, coupled with state aid to the troubled financial sector. These actions helped contain the Great Recession, but pro-cyclical discretionary fiscal expansion and the banking sector bail-outs led to an unprecedented rise in public debt-to-GDP ratios. Against this backdrop, a number of papers found that an excessively high public debt-to-GDP ratio hampers economic activity. In particular, Reinhart and Rogoff (2010) showed that there was a tipping point at 90 per cent of GDP, *i.e.*, economic growth slows down sharply if the debt-to-GDP ratio exceeds this level. The paper aims to check the robustness of the 90 per cent threshold, using a subset of a variant of the Reinhart-Rogoff dataset. The author estimates a bivariate relationship between growth and debt (and lagged debt) in a two-regime threshold model for a variety of thresholds. A robustness check of the threshold is also performed by jack-knifing the sample, *i.e.*, dropping one country from the sample at a time. Égert finds that the threshold may be different from 90 per cent and it varies a lot depending on the inclusion in the model of the contemporaneous or the lagged-debt variable. Furthermore, the negative impact of debt on growth is sensitive to outlier observations.

Nikola Altiparmakov and Milojko Arsić's paper starts by acknowledging that in the last two couple of decades the tax systems around the world have changed in response to the rapid globalization, which introduced international mobility of capital, goods and services and, to a lesser extent, labour. This caused a worldwide reduction in custom duties, corporate income taxes and tax wedges on labour and the reduction in tax rates was especially stark in emerging European countries, which experienced a fierce income tax competition in order to attract foreign investors (the so-called "race to bottom" phenomenon). Faced with reduced revenues, EU countries increasingly relied on consumption taxation, in particular VAT. But shifting the burden from income to consumption taxation is, in practice, challenging due to political considerations and the common belief that VAT is a regressive tax causing adverse distributional effects on poor households. However, recent research has unambiguously shown that much of the estimated extremely regressive incidence of consumption taxes against annual income originates from measurement errors inherent in expenditure surveys and that the theoretical basis for assessing the VAT incidence against annual income instead of annual expenditures or lifetime income is rather weak. Recent empirical estimates in EU member states, based on the lifetime tax incidence approach, reveal, on the contrary, a slightly progressive VAT incidence. A micro-simulation analysis of Serbian expenditure survey data conducted by the authors yields similar conclusions. However, the authors stress that this result is driven by some specific features of many emerging European countries (e.g., Poland, Romania and Serbia) compared to developed European ones. In particular, a significant presence of own-source small farming production and associated in-kind consumption, which enhances the progressivity of VAT systems, and the significant presence of a shadow economy and the evasion of direct income taxes, which suggests that household expenditure is a more meaningful indicator of the living standard and ability to pay taxes than the registered income. Overall, the authors conclude that common beliefs of regressive VAT taxation are overstated and poorly founded in economic reality of emerging European countries.

Werner Ebert and Sarah Ciaglia note that in the context of the current EMU debate on austerity and stimulus, the papers by Bouthevillain and Dufrénot and by Gavilán, Hernández de Cos, Jimeno and Rojas address important questions. In particular, the papers stress that, as fiscal policy is the only policy area in which instruments affect growth in different ways, the question on how heterogeneous growth patterns in the euro area can be shaped by fiscal policy measures compared to structural reforms is topical. Historical experience with fiscal policy measures shows that a "one-size-fits-all" approach does not work well, particularly in a common-currency area. To disaggregate and to be more country-specific in order to derive practical policy conclusions is wiser and this is done in both papers in two different ways: Bouthevillain and Dufrénot disaggregate public expenditures and revenues and select different growth periods; Gavilán *et al.* follow a country-specific long-term approach including open-economy and external-imbalances variables. The first paper concentrates on fiscal policy and growth and inquires whether a common fiscal policy would enhance or reduce growth in a similar way across countries; the second paper focuses on structural policies with a specific view on macroeconomic imbalances and growth and tries to explain how external imbalances evolved in Spain and in the euro area. Ebert and Ciaglia offer the authors suggestions for future extensions of the papers and advices to improve data. Finally, Ebert tries to draw from the papers lessons for strengthening the governance in the euro area. He argues that the approach of Bouthevillain and Dufrénot calls for a renewed agenda on the quality of public finances which should be integrated in Europe 2020 and the SGP, whereas the approach in the paper by Gavilán *et al.* could help to analyse the links between the Macroeconomic Imbalance Procedure and fiscal policy observation under the SGP. In particular, while currently no "one-size-fits-all" approach for EU member states' fiscal policies is possible or desired, alternative measures could be devised.

David Heald finds the Égert paper very interesting and admits that it almost convinced him that the Reinhart and Rogoff results that very high debt ratios are damaging are not robust. The paper makes patent its counter-intuitive result that, beyond 90 per cent, the effects on growth become less negative or neutral, but what is not clear is to what extent that is due to the particular data or econometric techniques used. In particular, concerning the data, Heald mentions that the Égert analysis is based on two time series: a longer one that looks at central government debt and a shorter one that looks at general government debt. The results are not substantially different, but it is important to choose which data to look at (central government, general government or public sector) in order to avoid arbitrage mechanisms. Another point is that net debt misses important information, such us pension liabilities. Finally, there is remarkable neglect of the assets side of the public sector balance sheet. In accruals-based government financial reports, the focus is on the net assets figure or, in national accounts, on the net worth figure. It is common knowledge that data often are not very good but, thinking about what kind of policy response there should be to particular levels of government debt, it is important to look at both sides of the balance sheet. Heald also reviews the paper by Jérôme Creel, Paul Hubert and Francesco Saraceno. He raises two main points: firstly, if the modelling assumptions determine the results and to what extent the judgements behind the model building prejudge the results that are going to be achieved; secondly, considering the criterion applied here, what official modelling has been done within the European Commission or elsewhere. The Saraceno results favour the old UK-style golden rule (where investment is outside the golden rule) rather than the new European Fiscal Compact. The point that Heald makes is that alternative modelling, which can be defended on technical economic grounds, might generate different results and that sometimes economic and political judgements can be obscured by modelling complexities. Commenting the paper by Rezk *et al.*, Heald finds the theoretical part on how human capital might influence growth very helpful and informative. As for the empirical part, the discussant concentrates on the issue of finding good data for human capital. He emphasises the importance of good social statistics, as well as reliable economic statistics, arguing that if there appears to be a complete separation between the social data and the economic data, it is possible to concentrate on securing good economic data. But, as soon as one starts arguing that human capital development is important in a growth context, it is essential to emphasise good social statistics and making sure that national statistics and social statistics do not themselves become a casualty of fiscal consolidation.

In discussing the paper by Douglas Sutherland, Sergey Vlasov notes that in the calculation of what has already been done or is under way, as well as what should be done in terms of adjustment in the long run, Greece, Portugal, Spain and Ireland – the countries more at risk with possible debt crisis in Euro area – not only have the largest cumulative fiscal tightening between the deficit trough and 2012, but also have the most modest adjustment need up to 2050, under the condition of bringing down gross financial liabilities to 50 per cent of GDP (with the only exception of Ireland). Then the discussant poses a few questions about how large the risk is that in 2012 the reporting figures would not correspond to those planned, how much the estimates are correlated with the low sovereign ratings given to these countries by international rating agencies and, finally, if there is a preliminary estimation on 2012 supporting authors' calculations. Vlasov argues that the debt overhang can be worked off in two ways: by a primary-balance tightening and by using the real-growth and real-interest rate effects. The authors analyze a wide range of possible instruments of fiscal consolidation and quantify their contribution to primary-balance tightening for each country. On this point the discussant casts some doubts related to: (i) the use of the OECD average as a target value for a set of instruments, as countries' peculiarities have to be taken into account too; (ii) the fact that the level of discontent among the population as a result of possible employees' layoffs, social spending cuts and increase in so-called "sin" taxes, and the way pension reforms should be carried out have been disregarded. Some other criticisms stem from the absence of (i) a discussion on how the primary balance might be substantially improved through the operation of automatic stabilizers and of (ii) an estimate of the effects of fiscal consolidation on GDP growth rates for OECD and/or individual countries. Commenting the paper by Nikola Altiparmakov and Milojko Arsić, Vlasov suggests to the authors to present their proposals of modifying the VAT system in Serbia as a way to offer special consumption incentives, boosting economic growth and improving fiscal sustainability. As for the methods, he argues to recur not only to the abolition of the reduced rate or to the elimination of certain exemptions from VAT, but to other forms of tax relief as well. This latter possibility would have the advantage of allowing to pursue specific goals on social ground (*i.e.*, increase fertility rate).

Session 1

THE SHORT-TERM IMPACT OF FISCAL POLICY

FISCAL MULTIPLIERS AND FISCAL CONSOLIDATIONS

Ray Barrell,[*] Dawn Holland[**] and Ian Hurst[**]

In this paper we look at fiscal multipliers in 18 OECD economies. The prospects for fiscal consolidation depend up the problems the country may face with its debt stock, the political will to deal with these problems and on the costs of consolidation. These costs are a function of the impacts of fiscal policy on the economy. Our analysis is based on a series of simulations using the National Institute Global Econometric Model, NiGEM. We first discuss the NiGEM model, as our results depend upon our description of the world. We then go on to decompose some of the factors that might affect our results. We consider the differences between temporary and permanent shifts in fiscal policy, the impact of an interest rate response, the role of expectations and the sensitivity to liquidity constrained consumers. Multipliers are time and state dependent. They are smaller the more open the economy and they appear to have been falling over time. They depend on the offsetting feedbacks in the economy, and in particular on the offsetting reactions of interest rates. A tighter fiscal policy will allow short term interest rates to be lower now and in the future if there is no change to the monetary target, and hence long term interest rates will be lower now, and the exchange rate will fall. Equity prices will rise and forward looking wage bargainers will change their behaviour. Each of these helps offset the contractionary effects of fiscal consolidation.

Introduction

This paper assesses various fiscal consolidation aspects for 18 OECD economies. The prospects for fiscal consolidation depend upon the problems a country may face with its debt stock, the political will to deal with these problems and on the costs of consolidation. These costs are a function of the impacts of fiscal policy on the economy. The analysis is based on a series of simulations using the National Institute Global Econometric Model, NiGEM. The NiGEM model will be discussed first, as the results depend upon the model properties. The key features of the model are that it is estimated and has a common structure across the 18 countries. If the results differ across countries it will be because they are different. Some of these differences, such as the openness of the economy, are important. They change over time and they are not related to estimation. Others, such as the speed of response to changes in income, do depend upon how the model was estimated. Although the model is estimated it has a strong role for expectations, and it is also flexible, as it can be run under different models of expectations formation, depending upon the thought experiment being undertaken.

Then the factors that might affect the results will be decomposed, for instance, by looking at temporary and permanent shifts in fiscal policy. In each case the first year multipliers will be presented. In the first year taxes will be raised or spending cut so that *ex ante* the deficit would improve by 1 per cent of GDP. Government consumption on goods and services and government transfers to individuals (mainly benefits and state pensions) will be changed, as well as income tax and indirect taxes. In the latter two the tax rate will be changed, and this has implications elsewhere

[*] Brunel University.

[**] National Institute of Economic and Social Research.

Corresponding author: Dawn Holland – National Institute of Economic and Social Research – 2, Dean Trench Street, Smith Square – London SW1P 3HE, United Kingdom. E-mail: dholland@niesr.ac.uk, Tel. +44 (20) 7654 1921.

The authors would like to acknowledge the financial support of the OECD towards the production of this paper, which is largely drawn from a broader study published as OECD Economics Department Working Paper No. 933, available from www.oecd.org/eco/workingpapers

in the economy. Each experiment is undertaken with the same set of assumptions, which will be discussed. The effects of government investment or corporate taxes will not be investigated. Government investment and corporate tax receipts are generally a small proportion of the economy, and a 1 per cent of GDP change to either would be a large proportionate change. In a temporary shock, the impact of a shift in government investment would be the same as a government consumption shock of the same magnitude. A long run shock to either government investment or the corporate tax rate would change the real equilibrium of the economy.

When undertaking experiments it is important to be able to dissect the contributing factors. These will be decomposed by removing them or changing them one at a time. Models such as NiGEM have to run with a monetary and a fiscal feedback rule and they use rational expectations. The rules and the assumptions about expectations affect outturns. The impacts of the assumptions will be investigated, looking at the role of forward looking bond and exchange rate markets, forward looking equities, forward looking wage bargainers and forward looking consumers. It is possible to run NiGEM with some or all of these, the effects on the multipliers will be investigated. Multipliers are time and state dependent. As we showed in Barrell, Fic and Liadze (2009), they are smaller the more open the economy and they appear to have been falling over time. They depend on the offsetting feedbacks in the economy, and in particular on the offsetting reactions of interest rates. A tighter fiscal policy will allow short-term interest rates to be lower now and in the future if there is no change to the monetary policy target, and hence long-term interest rates will be lower now. And the exchange rate will fall. Equity prices will rise and forward looking wage bargainers will change their behaviour. Each of these helps offset the contractionary effects of fiscal consolidation. It is also possible that the timing of fiscal consolidation and type of rule applied may affect outcomes. If fiscal policy is expected to be tightened in the future then long rates will fall now, increasing the offset, and perhaps even inducing a short-term expansion of output. Expansionary fiscal contractions are exceptionally rare, however.

The NiGEM model

The National Institute's global econometric model (NiGEM) can be used in a number of ways, from a backward looking structural model to a version that has similar long-run properties as the dynamic stochastic general equilibrium models used by institutions such as the Bank of England.[1] GDP (Y) is determined in the long run by supply factors, and the economy is open and has perfect capital mobility. The production function has a constant elasticity of substitution between factor inputs, where output depends on capital (K) and on labour services (L), which is a combination of the number of persons in work and the average hours of those persons. Technical progress ($tech$) is assumed to be labour augmenting and independent of the policy innovations considered here.

$$Y = \gamma(\delta(K)^{-\rho} + (1-\delta)(Le^{\lambda_L tech})^{-\rho})^{-1/\rho} \tag{1}$$

In general, forward looking behaviour in production is assumed and because of "time to build" issues investment depends on expected trend output four years ahead and the forward looking user cost of capital. However, the capital stock does not adjust instantly, as there are costs involved in doing so that are represented by estimated speeds of adjustment. The equilibrium level of unemployment is the outcome of the bargaining process in the labour market, as discussed in

1 The Bank of England Quarterly model is discussed in Harrison *et al*. (2005). NiGEM is discussed in Barrell, Holland and Hurst (2007), Barrell, Hurst and Mitchell (2007) and in other papers at www.niesr.ac.uk. NiGEM does not impose maximising equilibrium conditions in the same way as Dynamic Stochastic General Equilibrium models, but has the same steady-state equilibrium properties.

Barrell and Dury (2003), and the speed of adjustment depends on (rational) expectations of future inflation unless backward oriented learning is used. Financial markets normally follow arbitrage conditions and they are forward looking. The exchange rate, the long-term interest rate and the equity price will all "jump" in response to news about future events. Fiscal policy making involves gradually adjusting direct taxes to maintain the deficit on target, but it is assumed that taxes have no direct effect on labour supply decisions. Monetary policy making involves targeting inflation with an integral control from the price level, as discussed in Barrell, Hall and Hurst (2006) and inflation settles at its target in all simulations. Some of the key features of the model that determine the outturns of the simulation studies are detailed further below.

Consumer behaviour

As Barrell and Davis (2007) show, both the level of total asset based wealth $(\ln(TAW)$ or $\ln(NW+HW))$ and changes in financial $(d\ln(NW))$ and especially housing wealth $(d\ln(HW))$ will affect consumption (C).[2] Their estimates suggest that the impact of changes in housing wealth have five times the impact of changes in financial wealth in the short run, although long-run effects are the same. Barrell and Davis (2007) also show that adjustment to the long-run equilibrium shows some inertia as well. Al Eyd and Barrell (2005) discuss borrowing constraints, and investigate the role of changes in the number of borrowing constrained households. It is common to associate the severity of borrowing constraints with the coefficient on changes in current real incomes $(d\ln(RPDI))$ in the equilibrium correction equation for consumption. These coefficients are important in evaluating impact multipliers, and may increase during a severe economic downturn. One can write the equation for $d\ln(C)$ as:

$$\begin{aligned} d\ln(C_t) &= \lambda\{\ln(C_{t-1}) - [a + b_0\ln(TAW_{t-1}) + (1-b_0)\ln(RPDI_{t-1})]\} \\ &\quad + b_1 d\ln(RPDI_t) + b_2 d\ln(NW_t) + b_3 d\ln(HW_t) \end{aligned} \qquad (2)$$

where the long-run relationship between $\ln(C)$ and $\ln(RPDI)$ and $\ln(TAW)$ determine the equilibrium savings rate, and this relationship forms the long-run attractor in an equilibrium correction relationship. The logarithmic approximation is explained in Barrell and Davis (2007).

Operating in forward-looking consumption mode, consumers react to the present discounted value of their future income streams, which is approximated by total human wealth (TW), although borrowing constraints may limit their consumption to their personal disposable income in the short run. Total human wealth is defined as:

$$TW_t = Y_t - T_t + TW_{t+1}/((1+rr_t)(1+my_t)) \qquad (3)$$

Y is real income, T are real taxes, and the subscript $t+1$ indicates an expected variable which is discounted by the real interest rate rr_t and by the myopia premium of consumers, my_t. The equation represents an infinite forward recursion, and permanent income is the sustainable flow from this stock.

Prices

Consumer prices (CED) are modelled as a dynamic weighted average of unit costs of production and import prices, adjusted by the indirect tax rate. A policy shift that changes the indirect tax rate, therefore, has a direct impact on the price level. Unit costs of production (UTC)

2 Throughout d is the change operator and $\ln$ is the natural logarithm.

are derived from the cost minimization problem around the underlying production function, given by:

$$\text{Minimize } C = WL + rK \tag{4}$$

$$\text{s.t. } Y = \gamma(\delta(K)^{-\rho} + (1-\delta)(Le^{\lambda_L tech})^{-\rho})^{-1/\rho} \tag{5}$$

where the factors of production L and K are associated with factor prices W (wages) and r (user cost of capital).

The first order conditions of the cost minimisation problem give the optimal input ratio, which can be substituted into the production function to derive the cost minimising levels of factor inputs to produce a given level of output. It is assumed that firms operate on their factor demand curves, at least in the long run, which leads to the following expression for marginal costs:

$$\ln(MC) = \theta_1 + \ln(W) - (1+\rho)\ln\left(\frac{Y}{L}\right) + \rho\lambda_L tech \tag{6}$$

where:

$$\theta_1 = \rho\ln(\gamma) - \ln(1-\delta) \tag{7}$$

Marginal costs are treated as a shadow price, whereas observed basic prices (P) incorporate an endogenous mark-up, which is modelled as a function of the output gap.

Government sector

In order to evaluate multipliers a reasonably disaggregated description of both spending and tax receipts is needed. Corporate $(CTAX)$ and personal (TAX) direct taxes and indirect taxes $(MTAX)$ on spending are modelled, along with government spending on investment (GI) and on current consumption (GC), and transfers $(TRAN)$ and government interest payments (GIP) are separately identified. Each source of taxes has an equation applying a tax rate to a tax base (profits, personal incomes or consumption). As a default, government spending on investment and consumption are rising in line with trend output in the long run, with delayed adjustment to changes in the trend. They are re-valued in line with the consumers' expenditure deflator (CED). Government interest payments are driven by a perpetual inventory of accumulated debts. Transfers to individuals are composed of three elements, with those for the inactive of working age and the retired depending upon observed replacement rates. Spending less receipts gives the budget deficit (BUD), which adds to the debt stock.

$$\text{BUD} = \text{CED}*(\text{GC}+\text{GI})+\text{TRAN}+\text{GIP}-\text{TAX}-\text{CTAX}-\text{MTAX} \tag{8}$$

It has to be considered how the government deficit (BUD) is financed. Either money (M) or bond financing $(DEBT)$ are allowed:

$$BUD = d(M) + d(DEBT) \tag{9}$$

and rearranging gives:

$$DEBT = DEBT_{t-1} + BUD - d(M) \tag{10}$$

In all policy analyses a tax rule is used to ensure that governments remain solvent in the long run. The default rule is applied to the personal direct tax rate, which is adjusted endogenously to bring the government deficit into line with a specified target. This ensures that the deficit and debt stock return to sustainable levels after a shock. A debt stock target can also be implemented and this is discussed below. The income tax rate $(TAXR)$ equation is of the form:

$$TAXR = \text{f(target debt or deficit ratio} - \text{actual debt or deficit ratio)} \tag{11}$$

If the government budget deficit is above the target, (*e.g.*, 3 per cent of GDP and the target is 1 per cent) then the income tax rate is increased.

Monetary policy

Interest rates are set by the monetary authority in relation to a targeting regime, where policy interest rates are set in relation to a rule that is normally forward looking. We distinguish two types of rules, those that target only inflation and those that target the price level or a nominal variable such as GDP or the money stock. During the "great moderation" era central bankers and many economists became convinced that they had changed the world they lived in by adopting simple feedback rules for monetary policy in combination with rules for fiscal policy that kept debt in bounds. The simple feedback rule was based on the Taylor Rule (TR) that suggests that when inflation increases the central bank should increase the interest rate more than in proportion to the rise in inflation, and hence the real interest rate would rise and help choke off demand. In a forward looking world it is possible to improve on this principal. If agents see the central bank as fully credible, then the announcement of a price level target (PLT), rather than just an inflation target, will stabilise fluctuations in output and in inflation. A price level targeting central bank will loosen policy more rapidly as it has to get the price level back to target. The converse will be true in a boom. These two feedback rules are shown in equation (12) below, with int being the intervention rate, ssr being the steady state (endogenous) real interest rate, og being the output gap, inf and $inft$ being the inflation rate and the target, and P and PT being the price level and the price level target.

$$int_t = a_0 + a_1 ssr_t + a_2 og_t + a_3 (inf_{t+1} - inft) + a_4 (P_t - PT_t) \tag{12}$$

In a Taylor Rule a_0 is zero, a_1 is 1.0, a_2 is 0.5, a_3 is 1.5 and a_4 is zero, whilst in a PLT regime $a_{(1)}$ is zero, $a_{(2)}$ is also zero, and $a_{(3)}$ is set to 0.7 and $a_{(4)}$ to 0.4. The PLT rule has the advantage of working only on observables. The same is true of a two pillar strategy as embraced by the ECB. The bank responds to deviations of inflation from target and also deviations of a nominal aggregate (*NOM*) – the money stock for instance – as described in equation:

$$int_t = b_0 + b_1 (inf_{t+1} - inft) + b_2 (NOM_t - NOMT_t) \tag{13}$$

Forward looking financial markets

A deflationary shock such as a fiscal tightening will have a weaker interest rate response under a Taylor Rule than under price level targeting, and both may be weaker than a two pillar rule. If actors know the rule is in place then they will form expectations of the future path of short rates, and this will cause the current long rate to change, along with the exchange rate and the equity price. Forward looking long rates (LR) should be related to expected future short-term rates:

$$(1 + LR_t) = \prod_{j=1}^{T} (1 + int_{t+j})^{1/T} \tag{14}$$

Forward looking equity prices (EQP) are related to future profits (PR) in a forward recursion where $eprem$ is the equity premium:

$$EQP_t = PR_t + \frac{EQP_{t+1}}{(1 + int_t)(1 + eprem_t)} \tag{15}$$

The exchange rate depends on the expected future path of interest rates and the exchange rate risk premia, solving an uncovered interest parity condition, so that the expected change in the exchange rate is given by the difference in the interest earned on assets held in local and foreign currencies:

$$e_t = e_{t+1}\left(\frac{1+\text{int}^*_t}{1+\text{int}_t}\right)\left(1+rp_t\right) \tag{16}$$

where e_t is the bilateral exchange rate at time t (defined as domestic currency per unit of foreign currency), int_t is the short-term nominal interest rate at home set in line with a policy rule, int_t^* is the interest rate abroad and rp_t is the exchange rate risk premium.

Fiscal multipliers

NiGEM is an estimated and calibrated model with a supply side and rational expectations, but is does not go as far in this direction as modern DSGE models which are theory based, but fail in their description of the world. In a model such as ours multipliers are small. They average around 0.3 or less, as can be seen from Tables 1 and 2 below. Even then these estimates probably exceed the multipliers that one would see with any actual consolidation programme, because for some actions implementation speed is faster in the model than in the world. If one allows for more gradual implementation, this would reduce average multipliers to below 0.2. This matters in particular when comparing multipliers for taxes and benefits to those for spending. Taxes or benefits can be cut by 1 per cent of GDP relatively easily both in the model and in the world. Multipliers in response to income tax and benefit adjustments are small, as a part of the decline in personal sector income is offset by a temporary adjustment in the savings rate. As one can see from the tables, multipliers appear larger for cuts in real government spending. This is in part because of the assumption that such cuts can be implemented immediately, and this is certainly not the case. It is also in part because government consumption is part of the income identity and hence when they are cut (and reduce the number of people employed or goods and services bought) measured real output falls. If one were to reduce government spending by as much, but do it through wage reductions, then the impact on real GDP would be much less, and the second round effects of the shock would effectively be the same as an increases in taxes.

In order to determine the effects of an *ex ante* change in fiscal policy one has to avoid offsetting or reinforcing policy effects, but the model must otherwise be allowed to run. In each of our simulations in this section we make the following assumptions:

- Policy reactions are turned off for the first year:
  - The central bank does not change the short-term interest rate for a year, whatever the shock. It then follows a targeting regime that stabilises either the inflation rate or the price level.
  - The government does not target the deficit for the first year. The model has a feedback rule which adjusts the direct tax rate in relation to the gap between actual and target deficits. This is switched off for a year.
  - Government investment is fixed at the baseline for a year and does not respond to long-term factors in the first year. The same, where this is appropriate, is true for government consumption.
  - Other tax rates and all benefit replacement rates are held constant throughout the simulation period.
- Markets work and all quantities and prices can react and there are no exogenous variables in the model, with the exceptions of policy targets, labour supply and risk premia:
  - Financial markets look forward and are assumed to follow arbitrage paths, and expectations for those paths are outturn consistent.

  - Long-term government bond rates are the forward convolution of future short- term policy rates plus an exogenous premium.
  - Long-term real interest rates are the forward convolution of future short-term real policy rates plus an exogenous risk premium made up of the bond premium plus private sector risks.
  - Equity prices are the discounted value of future profits, where the discount factor is the market interest rate plus the exogenous equity premium.
  - Exchange rates "jump" when future interest rates change and they follow the arbitrage path given by nominal interest rates.
- Labour markets are described by an exogenous labour supply, a labour demand equation and by a wage equation based on search theory, where the bargain depends on backward and forward looking inflation expectations.
- Capital stocks adjust slowly towards that associated with expected capacity output four years ahead, which in turn depends upon a forward looking user cost of capital. Expectations are rational and factor demands and capacity output are based on a CES production function.
- Consumers respond to their forward looking financial wealth, but are not fully forward looking.

In the next sections the implications of several of these default assumptions will be tested.

Table 1 reports the estimates of the first year multipliers for 18 OECD countries, under the default assumptions described above, for a 1 per cent (*ex ante*) GDP rise in taxes or cut in spending that is reversed after one year. The multipliers for cuts in government consumption spending and spending on benefits are reported, as well as for rises in indirect taxes and direct (personal) taxes. Simulations are run one country at a time, so there are no spillovers across countries in the reported multipliers. Generally multipliers peak in the first year and then decline, and the *ex post* improvement in government revenues will normally be less than 1 per cent of GDP as tax bases change. Some of the effects of the impulse will be offset by declines in interest rates. Both short and long rates should fall, but the former may be trapped at the lower bound at present. This will have a limited impact on our results as long rates are forward looking and can move even when current short rates are restrained by the zero bound. In NiGEM, investment behaviour is mainly influenced by long real rates through the user cost of capital, and these are free to fall in response to the temporary fiscal tightening.

The multipliers reported in Table 1, illustrate some of the key differences across fiscal instruments, and also highlight important differences across countries. Government consumption spending multipliers tend to be larger than tax or benefit multipliers, as a fraction of any disposable income change is absorbed through a temporary adjustment to savings. However we should bear in mind the caveat mentioned above that it is not necessarily feasible to cut the provision of government goods and services at short notice.

Country size is an import distinguishing factor across country multipliers, as the long term fall in real interest rates that is produced by consolidations that is reflected in current long term real interest rates is an international phenomenon. When capital moves freely between countries, real interest rates are determined largely by the balance between global saving and global investment, and large countries such as the United States have much more impact than small ones such as Greece. In addition the initial interest rate response will be smaller in countries in EMU because the ECB responds to euro area inflation.

Multipliers tend to be smaller in more open economies, because the more open an economy is the more of a shock will spread into other countries through imports, and small open economies such as Belgium have small multipliers. Another structuring factor is the degree of dependence of

Table 1

First-year Multipliers from 1 Percent of GDP Temporary Innovations

Country	Government spending		Taxes	
	Consumption	Benefits	Indirect	Direct
Australia	–0.82	–0.27	–0.25	–0.22
Austria	–0.53	–0.17	–0.09	–0.13
Belgium	–0.17	–0.04	–0.05	–0.03
Canada	–0.53	–0.16	–0.05	–0.12
Denmark	–0.53	–0.10	–0.06	–0.04
Finland	–0.64	–0.14	–0.09	–0.08
France	–0.65	–0.32	–0.09	–0.27
Germany	–0.48	–0.29	–0.09	–0.27
Greece	–1.07	–0.44	–0.22	–0.32
Ireland	–0.33	–0.09	–0.07	–0.08
Italy	–0.62	–0.17	–0.07	–0.12
Japan	–1.27	–0.65	–0.34	–0.57
Netherlands	–0.53	–0.19	–0.07	–0.16
Portugal	–0.68	–0.15	–0.08	–0.11
Sweden	–0.39	–0.14	–0.06	–0.16
Spain	–0.71	–0.15	–0.17	–0.09
United Kingdom	–0.74	–0.22	–0.16	–0.15
United States	–1.12	–0.35	–0.35	–0.25

Note: No shift in the budget target. Experiments conducted in one country at a time.

consumption on current income. This is often related to liquidity constraints, with a higher current income elasticity more common in financially unliberalised economies such as Greece than in Belgium or the United States. Finally the speed of response of the economy depends in part on the flexibility of the labour market and the speed at which policies, such as a rise in VAT feed into prices.

Barrell, Holland and Hurst (2012) compare the temporary government consumption spending and direct tax multipliers from Table 1 to some of the key factors determining the differences in the magnitude of multipliers across countries: country size, import penetration and the estimated short-term income elasticity of consumption. This identifies a strong correlation between country size and the tax and spending multipliers, suggesting that the larger the economy the bigger the multiplier. The large economy impact on world interest rates must be more than offset by other features of large economies, such as the tendency to be less open to imports than the smaller economies, as the interest rate change in response to a temporary shock is very small. Import penetration has a very strong correlation with the impact multipliers, suggesting that more open economies tend to have smaller multipliers, both in response to spending cuts and tax rises.

Figure 1 illustrates the strength of this correlation with the temporary spending on goods and services multiplier.

The short-term income elasticity of consumption has little relationship with the first year government consumption multipliers, but shows a 50 per cent correlation with income tax multipliers, which feed directly into personal income. This relationship is illustrated in Figure 2. The indirect tax multiplier will also depend upon the speed at which real wealth effects reduce consumption. An indirect tax increase reduces real wealth, and as it affects consumption in the long run, it affects the multipliers.

A permanent fiscal consolidation also involves changing the budget deficit target. The reported multipliers in Table 2 are derived from the shocks applied in Table 1, but with the cut in spending or increase in taxes being permanent and also the deficit target is shifted by 1 per cent of GDP. This changes the shape of the multiplier, as income taxes will rise in all scenarios from the second year of the simulation to cover any shortfall in the 1 per cent of GDP consolidation, and long-term interest rates will fall by more than for a temporary consolidation. The impact of tax increases in the second year varies across shocks, depending on the degree of shortfall in the *ex post* budget improvement compared to the *ex ante* estimates.

In general, permanent multipliers should be smaller than temporary ones, as the impact of the fiscal contraction on long rates will be larger, and the fall in long rates will induce increases in asset prices and in investment.[3] Country size plays a much more direct role in determining the offset on a permanent consolidation relative to a temporary one than in determining the size of the multiplier itself. Figure 3 plots the ratio of permanent to temporary multipliers in response to an innovation in government consumption. There is a 60 per cent correlation between these ratios and economy size, measure as GDP in prices and PPPs of 2005. Larger countries, such as the United States, which has an important role in determining global interest rates, sees a much bigger decline in the magnitude of the multiplier when the consolidation is permanent, compared to small EMU countries such as Finland, where monetary policy is not independent. The five countries with the largest differences between temporary and permanent multipliers all have independent monetary policies and hence a fiscal contraction will induce a larger decline in long rates and in the exchange rate than is observable in the countries within EMU.

US fiscal multipliers under different monetary policy reactions

The fiscal multipliers reported in Tables 1 and 2 above are based on the series of assumptions detailed in the previous section. However, multipliers are not immutable, and in the next two sections the implications of some of these assumptions will be assessed, and the impact on the estimated multipliers from adopting an alternative set of assumptions reported. In this section the focus is on the choice of the monetary policy response to a fiscal consolidation. We use the United States as an example, but similar results can be expected in other large advanced economies.

Under the default assumptions, nominal short-term interest rates are initially fixed for one year. Thereafter, the monetary authority is assumed to follow the standard feedback rule, which applies a combined target to both inflation and a nominal aggregate. If one allows interest rates to respond immediately, the monetary authority will cut interest rates in the first year to offset part of the contractionary impact of the fiscal consolidation. This reduces the fiscal multiplier slightly in

[3] The impact of the consolidation on risk premia is not taken into account. These are largely absent currently for large countries such as the United States, the United Kingdom, France and Germany. For small countries such as Greece, Ireland and Portugal this is important.

Figure 1

Temporary Spending Multiplier and Import Penetration

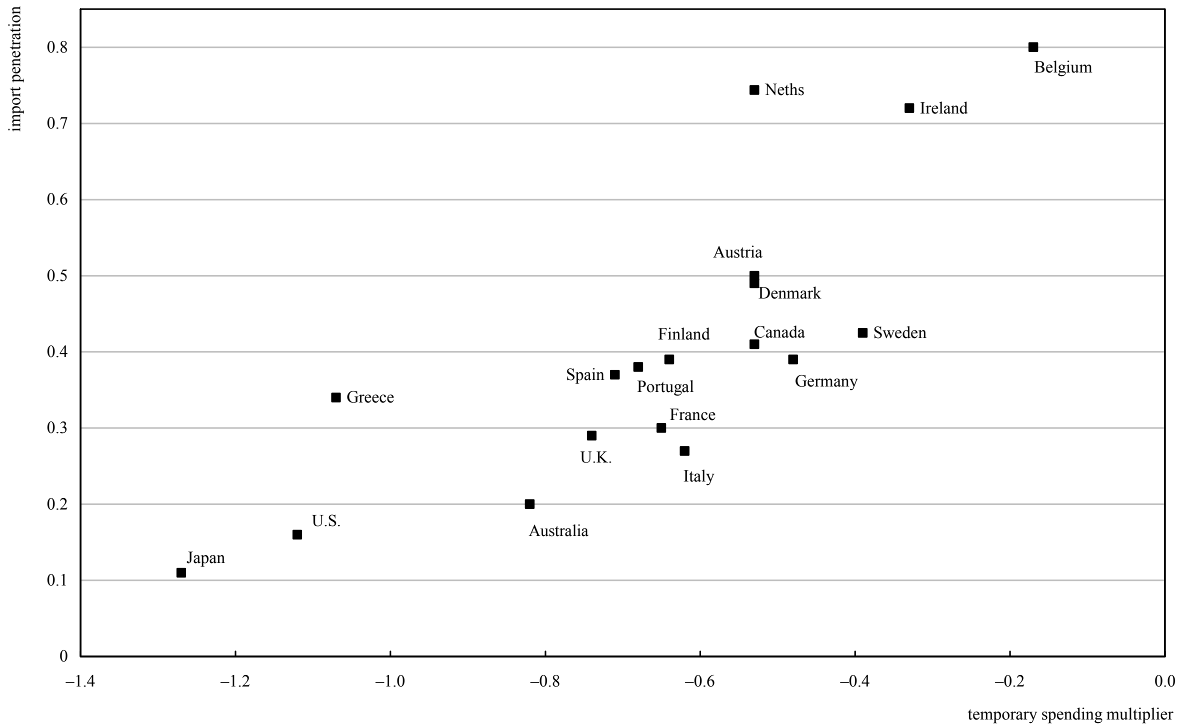


Figure 2

Temporary Tax Multiplier and Income Elasticity of Consumption

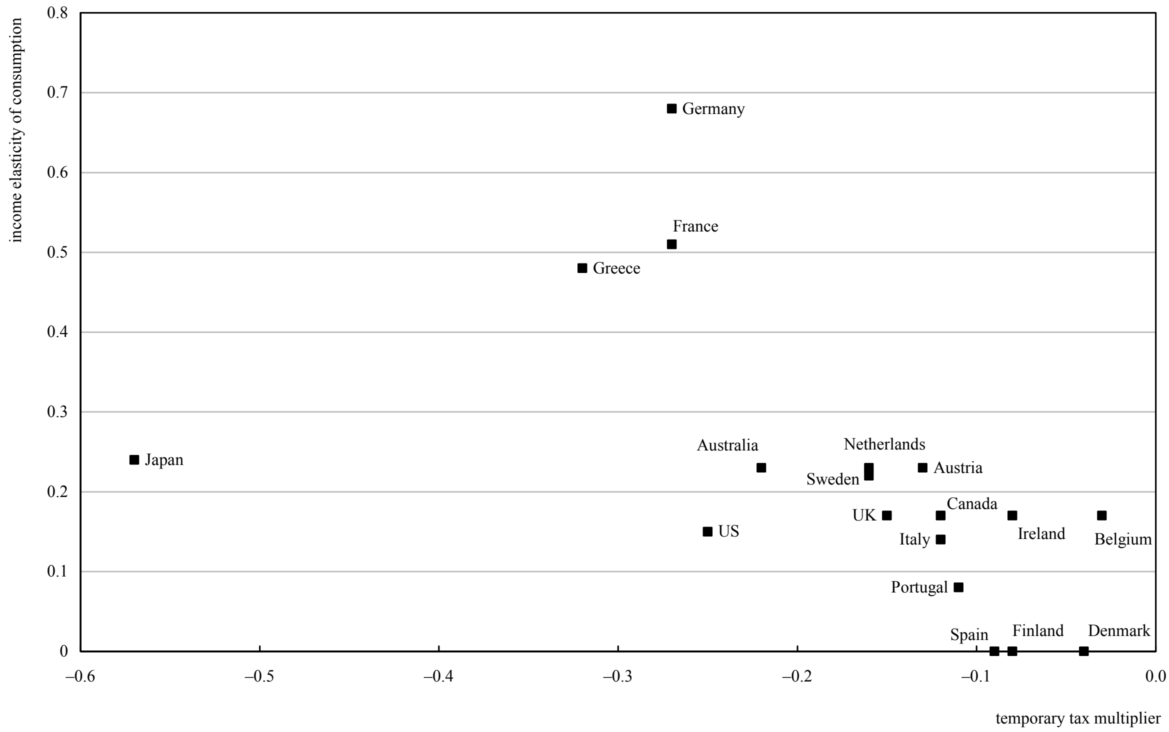


Table 2

First-year Multipliers from 1 Percent of GDP Permanent Consolidation

Country	Government Spending		Taxes	
	Consumption	Benefits	Indirect	Direct
Australia	–0.61	–0.17	–0.32	–0.12
Austria	–0.55	–0.18	–0.05	–0.13
Belgium	–0.16	–0.04	–0.02	–0.03
Canada	–0.43	–0.13	–0.10	–0.08
Denmark	–0.54	–0.10	–0.02	–0.05
Finland	–0.67	–0.16	–0.05	–0.10
France	–0.65	–0.33	–0.11	–0.26
Germany	–0.46	–0.29	–0.12	–0.25
Greece	–1.02	–0.44	–0.29	–0.37
Ireland	–0.33	–0.11	–0.06	–0.08
Italy	–0.62	–0.17	–0.06	–0.12
Japan	–1.15	–0.58	–0.43	–0.48
Netherlands	–0.51	–0.19	–0.05	–0.15
Portugal	–0.70	–0.17	–0.06	–0.12
Sweden	–0.40	–0.17	–0.05	–0.13
Spain	–0.74	–0.17	–0.16	–0.12
United Kingdom	–0.55	–0.14	–0.14	–0.08
United States	–0.90	–0.25	–0.27	–0.16

Note: Budget target shifted by 1 percent of GDP. Simulations conducted in one country at a time.

Figure 3

Ratio of Permanent to Temporary Government Consumption Multipliers

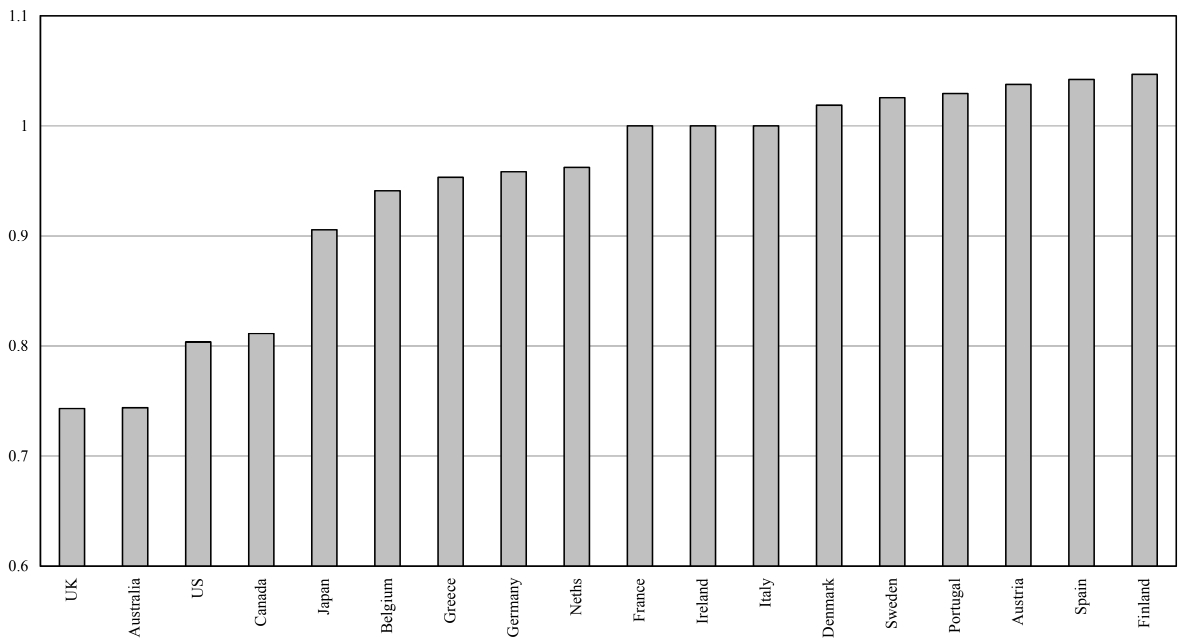
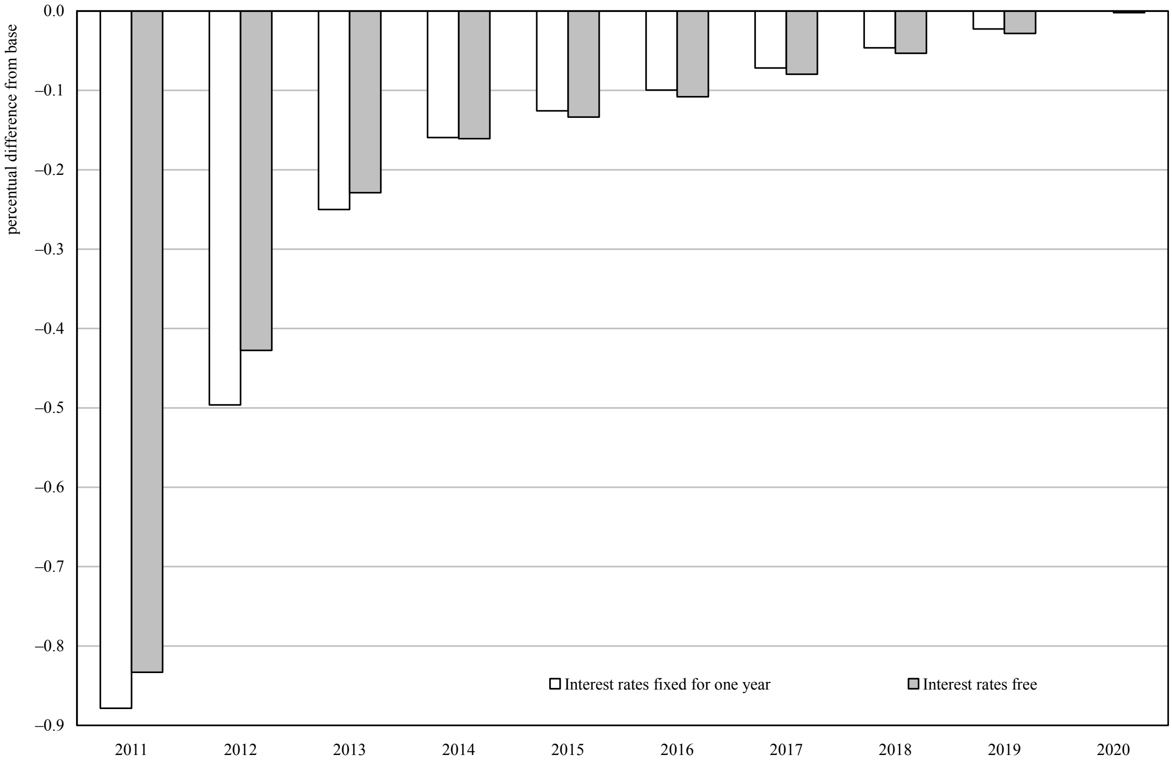


Figure 4

Impact on US GDP of 1 Percent Permanent Spending-based Consolidation

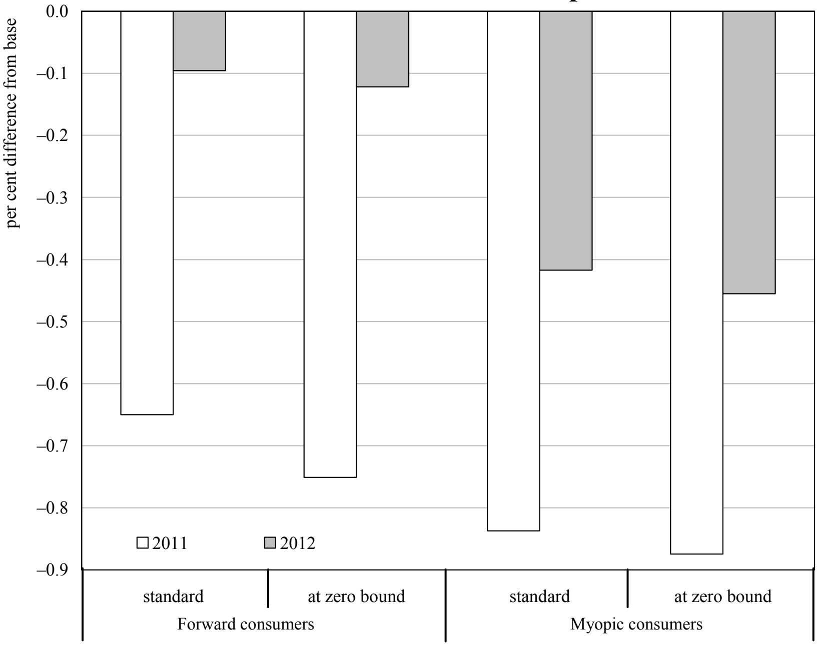


the first three years, as illustrated in Figure 3, but raises it slightly in subsequent years, so that the net cumulative impact of this speed of interest rate response is negligible.

It may of course be the case that monetary policy cannot react immediately because interest rates are at zero. In the baseline in mid 2011 interest rates in the US start to rise from the very low level seen since 2009, and hence a cut is possible. However, this January 2011 baseline included a significant increase in oil prices which would raise inflation in the United States and induce an interest rate response. Hence that baseline cannot be used to evaluate the importance of a zero lower bound, but it is possible to construct a counterfactual history where this is possible by removing the oil price shock and creating a new baseline. If we undertake this simulation then interest rates in the United States would be trapped at 0.001 until the first quarter of 2012, and hence one can evaluate the role of the zero bound over this period. The fiscal consolidation was simulated on the standard base and the counterfactual base with forward looking consumers and with myopic consumers. Forward looking consumers (discussed below) take the net present value of their future incomes and spend in relation to this. In a normal baseline a fiscal consolidation reduces interest rates in the short term and hence consumption rises as a result. At the zero bound interest rates cannot fall (for at least five quarters in our experiment) and hence consumption does not absorb as much of the shock and output falls by 0.1 percentage points more than in the normal case with forward looking consumers. In NiGEM myopic consumers are less influenced by short-term interest rates and investment decisions depend upon the user cost of capital. Hence the zero bound raises the multiplier by less if consumers are myopic, as can be seen from Figure 5. In general, the lower bound is not very important, but the longer it is expected to last the greater the effect on the consolidation multiplier.

Figure 5

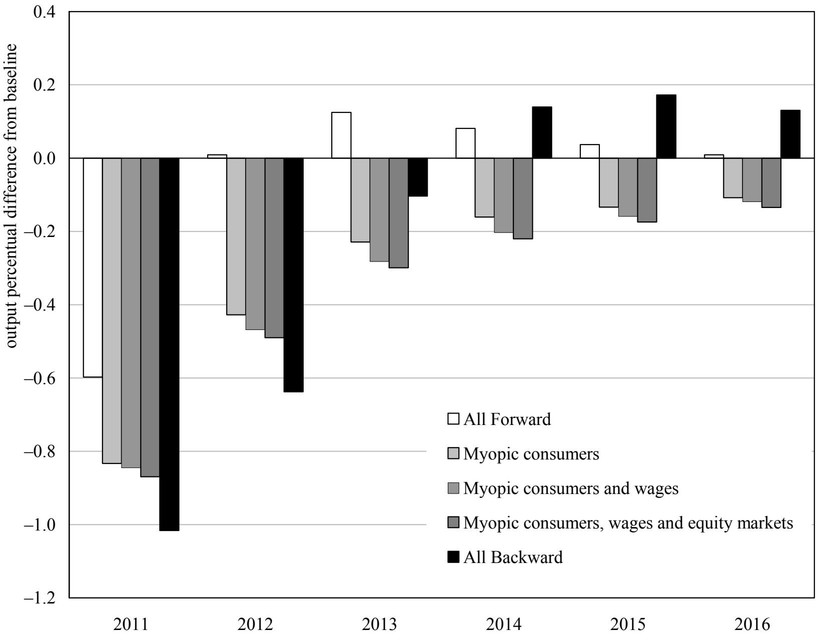
Impact of the Zero Lower Bound on Interest Rates on the US Consolidation Multiplier



Note: Forward consumers use forward-looking model-consistent expectation whereas myopic consumers are backward looking.

Figure 6

Fiscal Multipliers in the US Under Different Forms of Expectations



Note: the figure shows the deviations from baseline following a permanent fiscal shock.

Fiscal multipliers and expectations

Perhaps the most important set of assumptions affecting the size of the multiplier concern the role of expectations. In the standard set of simulations, the assumption is made that financial markets are forward looking. Long-term interest rates, equity price and exchange rates follow a forward looking arbitrage path, which is consistent with the simulation outturns. Wage setting is also partly forward looking, with wage settlements driven by a weighted average of current and expected inflation. Consumers are assumed to be myopic, but respond to their forward looking financial wealth, albeit rather slowly.

In this section some of these assumptions are relaxed in order to assess their impact on the estimated fiscal multipliers. Figure 6 shows the US multiplier in response to a permanent spending consolidation under the default assumptions (labelled as myopic consumers in the figure) and compares this to a range of alternative sets of assumptions regarding expectations. If one turns labour markets and equities backward looking so that they do not depend upon expectations about the future then the multiplier path is little affected. This is illustrated by the lines labelled myopic consumers and wage setters, and myopic consumers, wage setters and equity markets in the figure. The size of the multiplier is marginally larger under these assumptions, but not significantly so. The shock still operates with a monetary feedback rule and slower growth will reduce inflation and hence interest rates in the future will be lower. This will cause the forward looking exchange rate to jump down and forward looking long rates to do the same. If one turns long rates backward looking and fixes the exchange rate in the first period (and thereafter in this experiment), the multiplier in response to the consolidation programme in the US increases to over one in the first year. This is labelled "All backward" in the figure. Short term interest rates still fall and if one did not allow this to happen then the multiplier would be marginally larger still.

One can also move in the other direction and assume consumers are forward looking and react to the expected value of their future incomes. As taxes will be lower in the future and hence the net present value of incomes is higher, consumption is initially higher with forward looking consumers than it is with myopic ones. There is estimated inertia in the adjustment to the long run even with our forward looking consumption equations. Reducing the mark up would shrink the multiplier further from the –0.6 in the chart, but it would still be negative. However, as the myopia premium shrinks to zero the model comes close to be fully Ricardian in that future tax liabilities are more fully taken into account.

Fiscal multipliers and liquidity constraints

In the presence of perfect capital markets and forward-looking consumers with perfect foresight, households will smooth their consumption path over time, and consumer spending will be largely invariant to the state of the economy or temporary fiscal innovations. However, some fraction of the population at any given time is liquidity constrained with little or no access to borrowing, so that their current consumption is largely restrained by their current income. The share of the population that is liquidity constrained will affect the short-term income elasticity of consumption, given by parameter b_1 from equation (2), which we reproduce below:

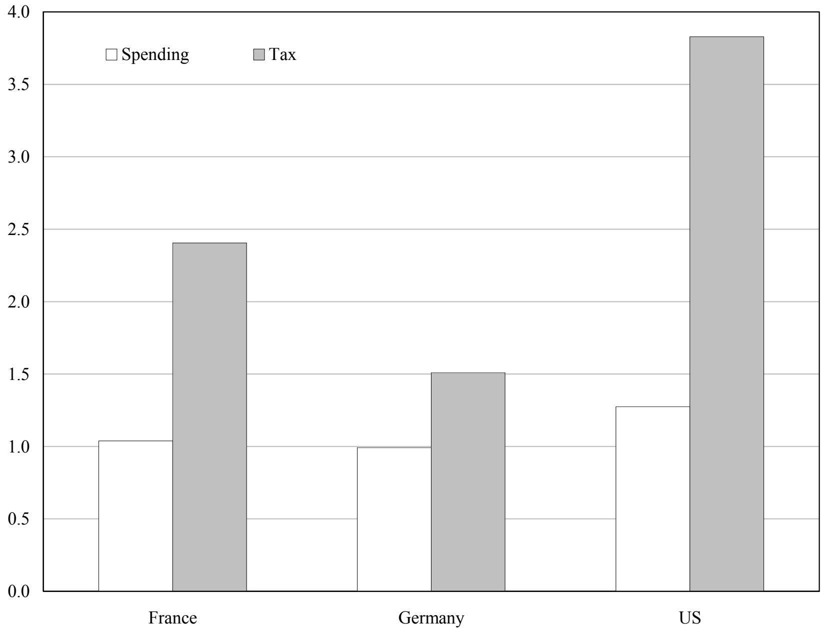
$$\begin{aligned} d\ln(C_t) = \lambda\{\ln(C_{t-1}) - [a + b_0\ln(TAW_{t-1}) + (1-b_0)\ln(RPDI_{t-1})]\} \\ + b_1 d\ln(RPDI_t) + b_2 d\ln(NW_t) + b_3 d\ln(HW_t) \end{aligned} \tag{2}$$

Cross-country differences in the average short-term income elasticity of consumption have a strong correlation with the tax multipliers, as illustrated in Figure 2. However, access to credit is dependent both on credit history and on current income, and so is necessarily sensitive to the state of the economy. As unemployment rises, a greater share of the population will be unable to access credit at reasonable rates of interest – at precisely the moment when they are in need of borrowing to smooth their consumption path. This means that consumption is likely to be cyclical, and that b_1 is likely to be time varying and dependent on the position in the cycle. Following a banking crisis the effects can be expected to be particularly acute, as banks tighten lending criteria, as discussed by Barrell, Fic and Liadze (2009). This also suggests that fiscal multipliers are dependent on the state of the economy – especially tax innovation multipliers – and this is consistent with recent studies such as Delong and Summers (2012) and Auerbach and Gorodnichenko (2012).

In order to assess the sensitivity of fiscal multipliers to the magnitude of the liquidity constraints parameter, b_1, we compare our standard multiplier for a 1 per cent of GDP innovation to government consumption and income tax to one where the liquidity constraints parameter is increased by 0.5. The ratio of the multipliers is illustrated in Figure 7. The spending multipliers are not affected dramatically – although the effects in the US are somewhat stronger than in France or Germany. The tax multipliers, on the other hand, are significantly increased when liquidity

Figure 7

Ratio of Multiplier with Heightened Liquidity Constraints to Baseline Multiplier



constraints are high – by nearly 4-fold in the US, 240 per cent in France and 150 per cent in Germany. This will significantly narrow or eliminate the gap between spending and tax multipliers during a downturn. This suggests that there may be little scope to apply a balanced-budget stimulus through an adjustment to policy instruments, especially during a banking crisis-driven recession.

Conclusion

In general in most countries fiscal policy multipliers are small, but are negative when fiscal policy is tightened. These effects are likely to be magnified during a recession, especially when banking systems are impaired. Tighter fiscal policy reduces growth in the short run in almost all circumstances, but a lower debt stock reduces pressures on real interest rates and hence in the longer term can raise sustainable output. This effect is larger for larger countries, and there are noticeable spillovers through real interest rates from policies in the United States (or from the euro area as a whole). If fiscal policy were to be noticeably tightened in the United States and Japan, as it should be, this could boost activity in the euro area as lower long-term real interest rates may well stimulate demand.

REFERENCES

Al-Eyd, A. and R. Barrell (2005), "Estimating Tax and Benefit Multipliers in Europe", *Economic Modelling*, Vol. 22.

Auerbach, A.J. and Y. Gorodnichenko (2012), "Fiscal Multipliers in Recession and Expansion", mimeo.

Barrell, R., D. Holland and A.I. Hurst (2007), "Correcting US Imbalances", NIESR, Discussion Paper, No. 290.

——————(2012), "Fiscal Consolidation Part 2: Fiscal Multipliers and Fiscal Consolidations", OECD, Economics Department, Working Paper, No. 933.

Barrell, R. and E.P. Davis (2007), "Financial Liberalisation, Consumption and Wealth Effects in Seven OECD Countries", *Scottish Journal of Political Economy*, Vol. 54, No. 2.

Barrell, R. and K. Dury (2003), "Asymmetric Labour Markets in a Converging Europe: Do Differences Matter?", *National Institute Economic Review*, No. 183.

Barrell, R., A.I. Hurst and J. Mitchell (2007), "Uncertainty Bounds for Cyclically-adjusted Budget Balances", in M. Larch and L.N. Martins (eds.), *Fiscal Indicators,* European Commission, Brussels.

Barrell, R., S.G. Hall and A.I. Hurst (2006), "Evaluating Policy Feedback Rules Using the Joint Density Function of a Stochastic Model", *Economics Letters*, Vol. 93, No. 1.

Barrell, R., T. Fic and I. Liadze (2009), "Fiscal Policy Effectiveness in the Banking Crisis", *National Institute Economic Review,* No. 207.

Delong, J.B. and L.H. Summers (2012), "Fiscal Policy in a Depressed Economy", mimeo.

Harrison, R., K. Nikolov, M. Quinn, G. Ramsay, A. Scott and R. Thomas (2005), *The Bank of England Quarterly Model*, Bank of England, London.

FISCAL MULTIPLIERS: HOW MUCH BANG FOR THE BUCK?

Glenn Follette[*] *and Byron Lutz*[*]

The U.S. federal government has consistently conducted an expansionary fiscal policy during the period following business cycle peaks (during the downturn and early in the recovery). However, the selected policies have frequently included actions that have relatively low direct multipliers. This study examines the historical record to gauge the effectiveness of fiscal policy – in terms of the timing, size and composition of the policy response – in stimulating demand. We use a narrative method to identify the policy choices following each post-war recession and draw on econometric evidence from the literature and from FRB/US macro model to estimate the boost to aggregate demand. We find that the direct multipliers are frequently well below 1 owing to a reliance on tax cuts.

Introduction

Considerable attention and debate has centered on the fiscal policy actions undertaken in response to the economic turmoil following the recent financial crisis. The United States implemented a substantial counter-cyclical policy by augmenting the automatic stabilizers – the boost in spending and reduction in tax payments that occur endogenously during an economic downturn – with a variety of discretionary tax and spending programs through the American Recovery and Reinvestment Act (ARRA) and other actions. This counter-cyclical response is not unusual; previous work by Follette and Lutz (2010) demonstrated that discretionary fiscal policy has typically been expansionary following business cycle peaks. This paper examines the issue more closely by detailing the *types* of policy actions taken in response to recessions and evaluating their impact on government budgets and on aggregate demand. The approach of the paper is as follows. We begin by outlining the issues and our methodology. We then turn to describing the discretionary policy actions in response to each of the post-World War II recessions and estimate their impact on the federal budget deficit. Next, we calculate the *direct* impulse to aggregate demand from these actions to gauge the "bang for the buck" (but do not consider follow-on, or total multiplier, effects). We do so using parameter values from the econometric literature on the response of consumers, businesses and subnational governments to federal government taxation and spending. Next we examine the role of automatic stabilizers in stabilizing demand. Our analysis suggests that the support to aggregate demand from automatic stabilizers is modest, leaving a potential role for active fiscal policy. Finally, we offer some concluding comments.

Methodological notes on measuring discretionary fiscal policies

We identify discretionary fiscal policy actions by using a narrative approach, similar to that pursued by Romer and Romer's (2009) analysis of tax policies. Our focus is on the federal government policies, where most significant counter-cyclical policy actions occur.[1] Accordingly, we use a variety of sources including the *Treasury Annual Report*, *Monthly Treasury Statement*, Congressional Budget Office documents, Joint Committee on Taxation budget estimates of

---

[*] Federal Reserve Board.

The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the research staff or the Board of Governors.

[1] In Follette and Lutz (2010) we document the small size and pro-cyclical movement of state and local fiscal actions.

proposed legislation, BEA's detailed tables on the National Income and Product Accounts, and *Congressional Quarterly* to estimate the budget effects of fiscal policy actions.

Unlike Romer and Romer, we are concerned with all of the budget actions that occur during and following recessions, regardless of their motivation. Romer and Romer's work focused on those tax actions that were not counter-cyclical to estimate total tax multipliers because those actions were least likely to be correlated with other impulses on aggregate demand. Our inclusive examination is consistent with the observations of Perotti (2012); when he examined budget consolidation efforts, the results were sensitive to whether all fiscal actions were included. Perotti notes that fiscal actions are often in response to other fiscal policies and thus looking at a subset will be biased. For this reason we would not want to exclude exogenous defense spending decisions, for example, because the counter-cyclical actions the government takes will be conditioned on these defense spending actions. While Romer and Romer's methodology of excluding some fiscal actions is proper and innovative for estimating total multipliers, it would not be appropriate for our goal of assessing the overall magnitude and effect of policy during and following recessions. Focusing on only explicit counter-cyclical actions could mischaracterize the government's response in many instances because the response is in part conditioned on the knowledge of the other fiscal policies. Therefore, focusing solely on the explicit counter-cyclical policies would not be helpful.[2] Nonetheless, we do attempt to decompose the change in policy around recessions into an explicit counter-cyclical component (*i.e.*, stimulus) and a non-counter-cyclical component.

To implement our narrative approach we have to define what constitutes a policy action and how to measure its size and timing. For purchases, we define the discretionary policy action to be equal to the real change in purchases over the period. Thus, no change in policy would be zero real growth in consumption and investment. Our definition is useful for examining short-run changes in policies and their effects on aggregate demand relative to zero growth – *i.e.*, to answer the question of whether fiscal policy is contributing to an increase aggregate demand. It would not be appropriate for longer run analysis, or evaluating whether fiscal policy is pushing demand above or below trend growth; in such cases alternative measure such as real purchases as a share of GDP, or as a share of potential GDP, would be more appropriate. In addition, although policies are formally set in nominal terms through the annual appropriations process, we are assuming that policy makers' decisions are based on the underlying real quantities. Finally the timing of the action is equated with the actual increase in spending and not when the decision is made.[3] For taxes and transfers we use the effect on revenues or outlays of changes in law and not movements automatically triggered by changes in economic activity. The timing of the policy change is set equal to the change in actual collections, rather than the time of enactment. The size of the policy action is based relative to prior law, except when prior law assumes the expiration of a tax. One potential flaw in our measure of using prior law is that it is sensitive to whether the baseline law includes inflation indexation. For example, during the high inflation 1970s the taxes rose as a share of GDP owing to bracket creep that was not fully offset by legislated tax cuts. With our measure, tax policy looks to be loose, when it was actually somewhat restrictive. For grants in aid, we use the change in real grants disbursed for non-Medicaid grants and the changes in laws for Medicaid grants, which move mostly endogenously.

2 For example, the Bush Administration did little explicit counter-cyclical policy in 2001 and 2002, perhaps owing to the fact that its tax policy was already counter-cyclical.

3 For example, a permanent increase in defense spending of 1 percent would increase actual purchases by roughly 0.6 per cent in the first year owing to time to build and other implementation lags. Our measure would yield an estimate of 0.6 per cent for the policy action.

Discretionary fiscal policy actions

This section provides a brief narration of the discretionary fiscal policy enacted during and following each post-war recession. The total effects on the deficit for each episode are summarized on Table 1. The total effect is also decomposed into three pieces: defense, stimulus (*i.e.*, policy actions undertaken explicitly for counter-cyclical reasons) and other.

1953. According to the National Bureau of Economic Research (NBER), the 1953 recession spanned from the third quarter of 1953 through the second quarter of 1954. The economy was overheating at the time of the recession, with the unemployment rate at 2-1/2 per cent at the business cycle peak. With the outbreak of the Korean War in June 1950, defense spending rose rapidly driving down the unemployment rate. Truman offset a portion of the increased demand by raising taxes and implementing wage/price and other controls on private demand. The 1950 personal and corporate income tax increases were permanent, whereas increases in these taxes in 1951 were temporary and slated to end in 1954. Balancing the budget was an important goal of the Truman and Eisenhower administrations (beginning January 1953), but the Truman tax increases were insufficient to keep the budget in balance. When the war ended in the summer of 1953 defense spending began to fall rapidly and the economy moved into recession. The Eisenhower administration let most of the temporary taxes expire as scheduled and enacted some additional tax cuts. But these were smaller than the declines in defense spending and the budget moved into surplus. Accordingly, as shown in Table 1, discretionary policy actions were pro-cyclical on net.

1957. The next business cycle peak was August 1957 and the trough was reached in April 1958. Again, the economy was rising briskly until the peak, with the unemployment rate falling to 4 per cent. It rose to 7-1/2 per cent over the recession and then declined to 5-1/2 per cent during the first year of the recovery. The Eisenhower administration was still more concerned with keeping the budget near balance than using counter-cyclical policy. Nonetheless, discretionary policies were mildly expansionary owing to increases in non-defense purchases that outstripped defense cuts and tax increases. The increase in non-defense purchases reflected policy decisions, such as the interstate highway program (enacted in 1956), that were taken before the recession.

1960. The business cycle peak was April 1960 and the trough was reached in February 1961. Real GDP rose 5 per cent over the four quarters ending 1960Q1 and then fell 1 per cent over the succeeding four quarters during which time unemployment rose from 5.1 per cent to peak at 7 per cent in 1960Q2. Fiscal policy was somewhat pro-cyclical during this period – tight during the recession and loose during the expansion – and there was little explicit counter-cyclical policy. The key policy changes included an increase in the social security tax rate in 1960Q1 just before the peak, a cut to corporate taxes in 1962, and increased real defense purchases owing to foreign entanglements. A small increase in unemployment benefits was enacted to provide additional weeks of unemployment insurance to those exhausting their benefits in 1961, after the trough.[4] In each subsequent recession extended UI benefits would be granted temporarily.

1970. The next business cycle peaked in December 1969 and the trough was reached in November 1970. The recession likely reflected, at least in part, a tightening in monetary policy to attack rising inflation. The Nixon administration also responded with wage/price controls during this period. The economy limped into this recession expanding only 2 per cent over the year prior to the peak and the unemployment rate was drifting up, although it was quite low at 3.6 per cent in 1969Q4. During 1970 the unemployment rate rose to 5.8 per cent and real output essentially moved sideways. During the recession real federal purchases were falling owing to a reduction in

4 Unemployment benefits typically run out after 26 weeks. The legislation added up to 13 additional weeks of unemployment compensation and cost $1 billion. It was financed by increased UI taxes in 1962 and 1963.

Table 1

Budget and Economic Effects of Discretionary Fiscal Actions
(percent of GDP)

Date	Policy	Recession	*t*+1	*t*+2	Recession	*t*+1	*t*+2
		Effect on Budget Deficit			Aggregate Demand Effect		
2008Q1	Discretionary actions	1.3	2.5	3.5	0.7	1.6	2.6
	Defense	0.3	0.6	0.7	0.3	0.6	0.7
	Stimulus	0.9	1.8	2.5	0.3	0.8	1.7
	Other	0.1	0.2	0.2	0.1	0.2	0.2
2001Q1	Discretionary actions	0.9	1.9	3.0	0.6	1.5	2.4
	Defense	0.2	0.4	0.8	0.2	0.4	0.8
	Stimulus	0.4	0.2	0.7	0.2	0.1	0.3
	Other	0.3	1.3	1.6	0.2	0.9	1.2
1990Q3	Discretionary actions	–0.1	–0.4	–0.5	0.0	–0.3	–0.4
	Defense	0.0	–0.4	–0.6	0.0	–0.4	–0.6
	Stimulus	0.0	0.1	0.2	0.0	0.1	0.2
	Other	–0.1	–0.1	–0.1	0.1	0.0	0.1
1981Q3	Discretionary actions	0.2	1.9	2.7	0.3	1.4	2.3
	Defense	0.5	1.0	1.4	0.5	1.0	1.4
	Stimulus	0.0	0.1	0.1	0.0	0.1	0.1
	Other	–0.3	0.7	1.2	–0.2	0.2	0.8
1973Q4	Discretionary actions	–0.6	0.9	0.9	–0.4	0.2	0.7
	Defense	–0.2	–0.2	–0.2	–0.2	–0.2	–0.2
	Stimulus	0.0	1.1	1.0	0.0	0.4	0.6
	Other	–0.5	0.0	0.1	–0.2	0.0	0.3
1970Q1	Discretionary actions	0.5	0.4	1.4	–0.6	–0.7	–0.3
	Defense	–1.1	–2.2	–2.8	–1.1	–2.2	–2.8
	Stimulus	0.0	0.1	0.3	0.0	0.1	0.3
	Other	1.5	2.6	3.8	0.5	1.4	2.2
1960Q2	Discretionary actions	–0.6	0.5	1.5	–0.4	0.5	1.3
	Defense	–0.1	0.7	1.1	–0.1	0.7	1.1
	Stimulus	0.0	0.1	0.0	0.0	0.1	0.0
	Other	–0.5	–0.3	0.3	–0.4	–0.3	0.2
1957Q3	Discretionary actions	0.1	0.5	0.2	0.1	0.4	0.3
	Defense	0.2	0.0	–0.3	0.2	0.0	–0.3
	Stimulus	0.0	0.0	0.0	0.0	0.0	0.0
	Other	–0.1	0.5	0.5	–0.2	0.4	0.6
1953Q3	Discretionary actions	–0.3	–2.8	–3.3	–0.7	–3.6	–4.1
	Defense	–1.0	–4.0	–4.6	–1.0	–4.0	–4.6
	Stimulus	0.0	0.0	0.0	0.0	0.0	0.0
	Other	0.7	1.2	1.3	0.2	0.4	0.5

Recession is the first 4 quarters following business cycle peak (beginning in the quarter indicated), *t*+1 is following 4 quarters, and *t*+2 is next 4 quarters.

Vietnam-related defense spending. The Nixon Administration raised transfers significantly in 1970Q2 and cut personal and corporate income taxes, in part, by letting Johnson's temporary tax surcharges expire. In January 1971 Nixon's State of the Union message promoted an expansionary budget to help stimulate the economy – by increasing transfers and grants to state and local governments. The goal was a full employment budget balance at 4 per cent unemployment. From a budgetary perspective these increases were larger than the cuts in defense spending. We have only identified the extra UI benefits as stimulus because the other parts of the Nixon program were permanent. Overall, policy actions boosted the deficit, but reduced aggregate demand owing to their timing and composition.

1973. Economic policy making during the 1973-75 recession was complicated by the combination of high inflation – reflecting in part sharp rises in oil and commodity prices – and a long and deep recession. The economy peaked in November 1973 and reached bottom in March 1975. The economy fell particularly steeply during the second half of the period. The unemployment rate rose from 4.8 to nearly 9 per cent. Fiscal policy was somewhat restrictive in 1974, in part owing to declining defense spending, but also because of an increase in social security taxes that was only partly offset by increased social security benefits.[5] In 1974, inflation was seen as a more urgent problem (Stein, *Presidential Economics*, p. 212) and the Whip Inflation Now program was unveiled. But when the economy began to drop quickly in the fall of 1974, policy makers shifted to stimulative fiscal policy. First extended unemployment benefits were proposed, along with expenditure restraint and taxes to pay for the benefits. However, after the severity of the recession became apparent a substantial stimulus program was enacted in February 1975 that included permanent and temporary tax cuts. Personal income tax cuts averaged 1¼ per cent of GDP in calendar 1975, with most of it delivered as a rebate in the second quarter – boosting disposable personal income by more than 3 per cent of GDP in that quarter. The economy began to expand in Q2. We have assigned the permanent portion of the 1975 tax cut as part of stimulus program as well as the rebate and the extended unemployment benefits.[6]

1981. High inflation and the subsequent tight monetary policies led to a sharp recession that lasted from July 1981 to November 1982 and the unemployment rate rose from 7.4 to 10.7 per cent. Real government purchases were on an upswing at that point as real defense spending began to climb under Carter and accelerated under Reagan. Significant tax cuts for individuals (phased in over 1981Q3, 1982 Q2, and 1983Q3) and corporations, including investment incentives, were enacted in August 1981. These were partly offset by tax increases in 1982 and 1983 in response to rising deficits and to cuts in transfers and grants. Deficits rose and fiscal policy was expansionary.

1990. The 1990 recession followed a period of restrictive monetary policy and a spike in oil prices caused by the invasion of Kuwait. The economy peaked in July 1990 and moved sideways for a few months before declining 1 percent through March 1991. The unemployment rate rose from 5.3 to 6.6 per cent by the trough and then continued to drift up over the next year and a half owing to the shallowness of the recovery. Fiscal policy was contractionary owing to the enactment of the 1990 Budget Enforcement Act which raised taxes, cut entitlements, and capped expenditures on federal purchases. In addition, defense purchases were put on a downward path reflecting the end of the Cold War, but this was offset temporarily – during the recession year – by expenditures related to the first Gulf War. Increased non-defense purchases, from budget decisions made before the 1990 Budget Enforcement Act moderated the degree of restraint.

2001. After a long expansion, culminating in the dot.com boom and bust, the economy peaked in March 2001 and then fell into a mild recession which reached the trough in November. GDP

[5] In addition, policy was tight because of rising bracket creep which is not captured in our measures of fiscal policy.

[6] We did not include any of the 1970 actions in stimulus because the goal was set forth as balancing the high-employment budget. In this case the goal was boosting economic activity and the actions were taken in response to the severity of the recession.

declined ¾ per cent over the period and the unemployment rate rose modestly from 4 to 5½ per cent. The unemployment rate continued to rise during the meager recovery, reaching 6¼ per cent in 2003. Fiscal policy was expansionary owing to tax cuts that had been planned before the economy weakened and increased spending on defense (in response to the terrorist attacks of 9-11) and domestic initiatives (e.g., No Child Left Behind and several expansions of Medicare benefits). The 2001 tax act was originally conceived as a phased-in reduction of income and estate taxes beginning in 2001 and the 2001 portion of the tax cut was increased and a rebate on 2000 taxes was added in response to the weak economic outlook. Additional stimulus provisions were enacted in 2002 (partial expensing extended unemployment benefits) and 2003 (accelerating provisions of the 2001 act, cutting taxes on dividends and temporary boost to grants) owing to the subpar recovery. We designate the 2001 rebate and the 2002 and 2003 actions as stimulus.

2008. The Economy peaked in December 2007 with an unemployment rate of 4.8 per cent and output fell 5 percent over the next six quarters until the trough was reached in June 2009 with the unemployment rate up to 9.3 per cent. Despite the large decline, the economy only slowly recovered and the unemployment rate continued to move up to 10 per cent by the end of the 2009 before drifting down. Against this backdrop, several stimulus actions were taken. First in 2008 a temporary income tax cut was enacted and subsequently unemployment benefits were augmented.[7] In February 2009 ARRA was enacted which included temporary tax cuts and increases in transfers, aid to state and local governments, and federal purchases. Subsequently, in 2010 some of these programs were extended, and in 2011 and 2012 a payroll tax cut was put in place.

Budget effects of policy decisions

From this narrative, one can readily see that substantial counter-cyclical policy actions enacted explicitly for counter-cyclical reasons were only taken twice, towards the end of the 1973-75 recession and in 2008-9. In addition, policy moved in a decidedly counter-cyclical direction in 1982 and 2001 owing to campaign promises. By contrast, policies were pro-cyclical in 1990 and 1953.

Defense purchases have been an important component of fiscal actions during many of the cycles. Although these purchases are often considered to be exogenous, the manner in which they are financed is a crucial determinant of the overall stance of fiscal policy. Thus, it is instructive to examine defense purchases as a separate category.

A number of conclusions can be drawn by comparing and contrasting the post-war recessions. As revealed on Table 1:

- As noted above, of the nine recessions, only the 1973 and 2008 recessions had large explicit counter-cyclical policies either during the recession or shortly thereafter. And only in 2008 was counter-cyclical policy put in place during the initial year of the recession.
- In three of the recessions defense spending was falling rapidly (1953, 1970, 1990) and in four it was rising quickly (1960, 1981, 2001, and 2008).
- Between 1953 and 1973 there is a negative correlation between defense spending and other policies during the recession and early recovery period, consistent with a strong balanced budget motive, while after 1973 there is a strong positive correlation.

[7] In the summer and fall of 2008 financial markets were addressed by nationalizing the government sponsored enterprises (GSEs) Fannie Mae and Freddie Mac and providing liquidity and support to banks and the auto sector through TARP – these actions are not included in our fiscal measures. The Troubled Asset Relief Program (TARP) has been virtually budget neutral – the government made money on loans to banks which were offset by losses associated with AIG and GM. The bailout of Fannie Mae and Freddie Mac may cost the government about 1½ percent of GDP. The economic effects of these programs while possibly substantial, are extremely difficult to quantify. Blinder and Zandi (2011) have tried to do so.

- In part because of the change in the correlation between defense and other policies, fiscal policies has generally have been much more expansionary, in terms of the size of the deficit, since the mid-1970s than in the earlier period.
- The largest fiscal programs were put in place following the 1981, 2001, and 2008 recessions.
  - The 1981 and 2008 recessions were the deepest in the post-war period and thus would be natural to have larger responses. By contrast, the 2001 recession was quite mild in terms of the loss of GDP or the peak unemployment rate, yet the fiscal response was very strong.[8]
  - That said, defense spending was an important component in all three episodes and tax cuts had ben preannounced in two of them. Thus, the size and timing of the 1981 and 2001 policies may have been somewhat fortuitous.
- Policy has tended to be relatively modest in the recession year, with much larger actions in the second (t+1) and third years (t+2). The delayed timing may reflect the recognition lags, as well as a general reluctance to pursue counter-cyclical policy until a recession is shown to be substantial.

Aggregate demand effects of discretionary policies

Romer and Romer (2009) argue persuasively that omitted variable bias issues make it exceedingly difficult to estimate the effectiveness of counter-cyclical policies. Accordingly, we estimate the aggregate demand effects by looking to research on macro consumption and investment functions as well as research based on panel studies to choose parameter values for responses to tax and transfer policies.[9] We estimate aggregate demand effects by summing the changes in real government purchases, plus the induced consumption from tax and transfer policies based on an estimated consumption function, plus increased investment from changes in taxes and subsidies, and add in an assumed response by state and local governments to changes in discretionary grants-in-aid. By design these estimates only include the direct effect and not the follow-on multiplier effects. Thus, the change in aggregate demand from an increase in real purchases is 1.0 because there are no leakages from imports, crowding out from higher interest rates, or second round multiplier effects. While these may be important the focus here is on the impact effect due to the composition of policies chosen. The 2008-10 fiscal policies and estimated economic effect are described in detail as a guide to the procedure.

2008-10 Stimulus policies[10]

The federal government enacted two pieces of stimulus legislation in 2008. First, a temporary tax cut of $100 billion (0.7 per cent of GDP) was delivered in Q2 and Q3, along with a one-year 50 per cent partial expensing provision ($40 billion in 2008, but only $10 billion over ten years). Second, temporary extended unemployment benefits were put in place the third quarter. These benefits were then enlarged in November and the program's duration was extended several times (including by the ARRA legislation in 2009). The benefits were initially equal to 0.1 per cent of GDP in 2008Q3 and grew to 0.6 per cent of GDP by 2010Q2. These actions probably had only a small effect on aggregate demand in 2008. Empirical investigations by Shapiro and Slemrod (2009), Parker, Souleles, Johnson and McClelland (2011), Sahm, Shapiro, and Slemrod (2010) and

8 Judging the size of the stimulus relative to the size of the shock is complicated by the fact that greater stimulus will reduce the output gap and thus generate a smaller *ex post*-measured shock.

9 Many of the studies do not offer quarterly timing of the demand effects and we therefore judgmentally set the quarterly timing.

10 The 2011 stimulus policies are not included because they fall outside our three year window.

others suggest that that 25 to 50 per cent of a temporary tax rebate is spent within 2 quarters of receipt and the rest is saved (we assume 40 per cent is spent, 25 per cent in the first quarter and 15 per cent in the following one). Work by House and Shapiro (2008) and Cohen and Cummins (2006) suggest that partial expensing has little impact on investment. By contrast, much of the increase in unemployment benefits probably was spent – we assume 85 per cent, in part, because it is targeted to those with significant income losses.[11] As a result, as shown in Figure 1, only a small portion of the stimulus was spent in 2008, with the ratio of increased demand to increased budget deficit, interpretable as an "aggregate MPC", of only 0.35 – largely because most of the tax rebate was saved, but also because the lack of stimulus from the partial expensing provisions.

In early 2009 the American Recovery and Reinvestment Act (ARRA) was passed. It included personal and corporate income tax cuts, grants to state and local governments, increases in transfer payments, and a

Figure 1

Stimulus Policies and Aggregate Demand, 2008-11

(percent of GDP)

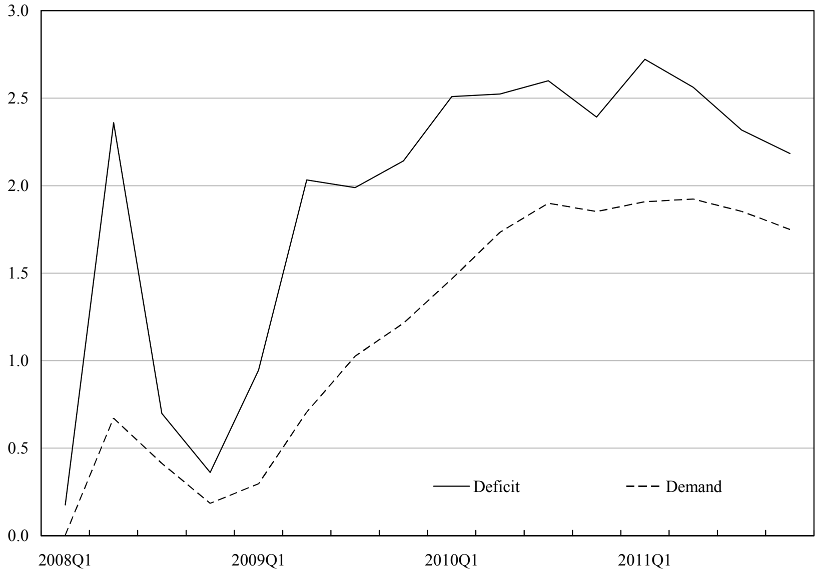
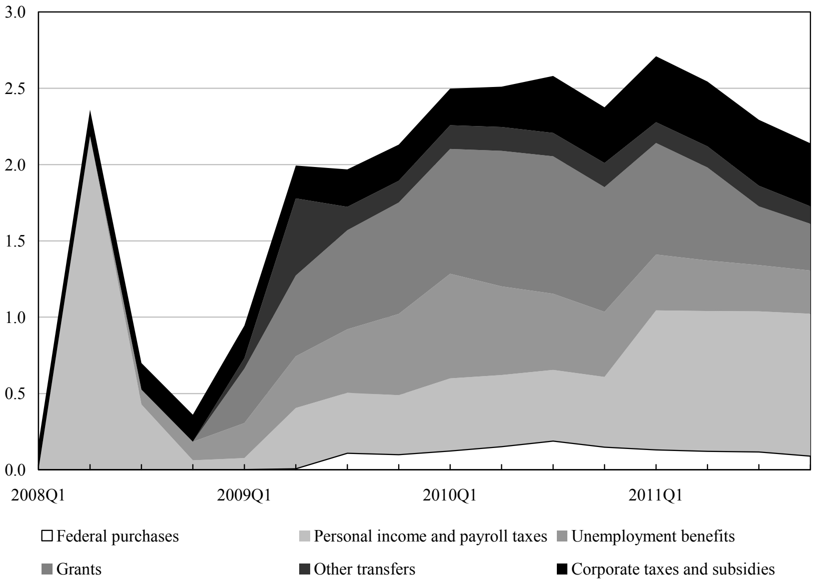


Figure 2

Composition of Stimulus Policies, 2008-11



[11] Aggregate consumption functions typically indicate that consumption out of transfers is higher than that out of other income. We base our demand effects on the consumption function used in the FRB/US structural model, with an MPC of 0.85 over 8 quarters, versus 0.7 for other income. See Brayton and Tinsley (1996). Note, that these benefits are targeted to those with unemployment durations longer than 26 weeks.

small increase in federal purchases of goods and services. The program – excluding the routine extension of AMT and extension of UI benefits (which we included in the 2008 actions) – totaled about $700 billion, or 5 per cent of GDP, that was expected to be largely spent out by the government over several years. Subsequently the grants programs were extended and at the end of 2010, with the economy still weak, the expiring $60 billion personal tax cut included in ARRA was replaced by a $100 billion payroll tax cut. The table below sketches out the effect of ARRA and other stimulus legislation on major budget aggregates and the quarterly pattern is displayed in Figure 2.

Composition of 2008-11 Stimulus Program
(effect on budget, percent of GDP)

Year	Personal and Payroll Taxes	Unemployment Benefits	Other Transfers	Grants	Federal Purchases	Corporate Taxes and Subsidies
2008	0.67	0.05	0.00	0.00	0.00	0.17
2009	0.32	0.38	0.22	0.57	0.05	0.23
2010	0.47	0.55	0.16	0.86	0.15	0.31
2011	0.92	0.32	0.13	0.51	0.11	0.43

An important feature of the stimulus program was the significant use of temporary grants to state and local governments. Empirical work on the effect of state and local grants is not dispositive. Early research, such as Gramlich (1969), suggested that an increase in grants is spent by the government and in the area that for which the grant was made – and dubbed the flypaper effect. More recent work, such as that by Brian Knight (2002), suggests that increased grants sometimes result in lower taxes. Interesting work by Suarez and Wingender (2010), Shoag (2012) and others on state fiscal multipliers are consistent with the grants being spent and spent quickly.[12] Moreover, in the current episode the temporary grants are of roughly the same magnitude – but smaller – than the cyclical shortfall in revenue. With state and local governments restricted by their constitutions to run budgets that are close to balance it would probably be optimal to spend the extra grants to prevent *temporary* swings in provision of state and local services – namely education and health. Accordingly, consistent with the new state-level fiscal multiplier literature, we assume that the increased grants are spent out over the four quarters following receipt.

The net result of the effect of all stimulus actions on aggregate demand is the pattern shown in Figure 1: the direct boost to aggregate demand is consistently below that of the effect on the budget, but the ratio of the two – the "aggregate MPC" – moves towards 1 over time, rising from 0.35 in 2008 to 0.7 in 2010. This reflects, in part, the phased-in response of consumers and state and local governments to the tax cuts, transfers, and grants.

[12] By contrast, Cogan and Taylor (2011) argue that state and localities saved the extra grants. However, their regression analysis rests on a levels regression using non-stationary variables. Moreover, their hypothesis suggests that state and local budgets would be flush with funds, however, state budget balances are quite low by historical standards (see National Association of State Budget Officers, 2012) and state and local deficits as measured in the NIPA are exceptionally large.

Budget and Demand Effects from 2008-10 Fiscal Policies
(percent of GDP)

Year	Stimulus		All Discretionary Policies		Memo: Change in Demand per 1 ppt Deficit	
	Budget	Demand	Budget	Demand	Stimulus	Total
Recession (2008)	0.9	0.3	1.3	0.7	0.35	0.54
t+1 (2009)	1.8	0.8	2.5	1.6	0.45	0.62
t+2 (2010)	2.5	1.7	3.5	2.6	0.67	0.76

GDP effects are annual average (year over year) to be comparable to the budget effects.

Budget and Demand Effects from 1973-75 Fiscal Policies
(percent of GDP)

Year	Stimulus		All Discretionary Policies		Memo: Change in Demand per 1 ppt Deficit	
	Budget	Demand	Budget	Demand	Stimulus	Total
Recession (1973Q4-74Q3)	0.0	0.0	–0.6	–0.4	n.a.	0.64
t+1 (1974Q4-75Q3)	1.1	0.4	0.9	0.2	0.36	0.23
t+2 (1973Q4-76Q3)	1.0	0.3	0.9	0.6	0.63	0.69

Other discretionary fiscal policies were being implemented in addition to the stimulus actions, namely expanding defense commitments. Taken together, the increase in the deficit was 1.3 per cent of GDP in 2008, rising to 3.5 per cent by 2010, with the direct boost to aggregate demand estimated to be .7 per cent of GDP in 2008 and 2.6 per cent of GDP in 2010. Figure 3 compares the effects of all discretionary policies to stimulus policies.

1973-75 stimulus policies

The other period of active counter-cyclical fiscal policy was in response to the 1973-75 downturn. As described earlier, the stimulus was implemented through a one-time rebate and what became permanent tax cuts, and extended unemployment benefits, Figure 4. Given the low MPC that are estimated for rebates, and the slow adjustment by consumers to permanent tax cuts, the stimulus was rather modest. Moreover, these policies were against a backdrop of a downturn in defense spending. In sum, fiscal policy was not very stimulative and did not turn stimulative until the recession was ending, and the bang for the buck was initially muted.[13]

[13] The stimulus package was implemented at the business cycle trough and thus some argue that the package was unnecessary. But, the recovery may have begun at this point because of the additional boost to aggregate demand created by fiscal policy, either directly or through shifts in expectations.

The 1981 and 2001 recessions

Policy in these two periods was similar. In the first year, tax cuts and increased defense spending were enacted to carry out campaign promises with little explicit regard to the cyclical position of the economy. In subsequent years the policy actions of the two periods diverged a bit. After the initial bout of tax cuts, tax increases were enacted in 1982 and 1983 in response to the budget deficits, while in 2002 and 2003 additional tax cuts and spending programs were enacted, with some of the 2002 and 2003 tax cuts explicitly implemented as part of a stimulus program. The heavy use of tax cuts, 70 per cent of the discretionary increase in the deficit in 1981-83 and 40 per cent in the 2001-03 period would normally imply that the demand effects of the discretionary policies will be somewhat muted, but the rapid increases in defense purchases boosted the "bang for the buck".[14] Indeed, the ratio of the increase in aggregate demand to the increase in the deficit was about 0.8 in the both episodes.

Figure 3

Discretionary Policies and Demand, 2008-11

(percent of GDP)

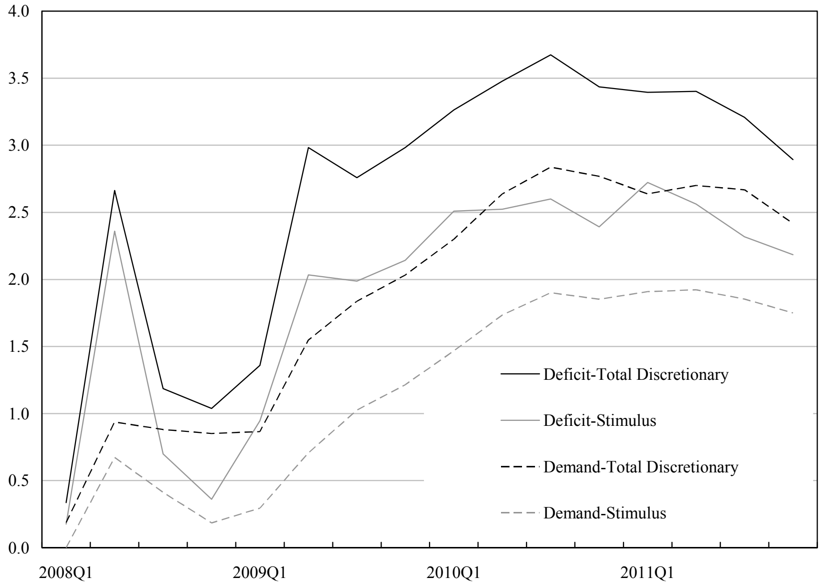
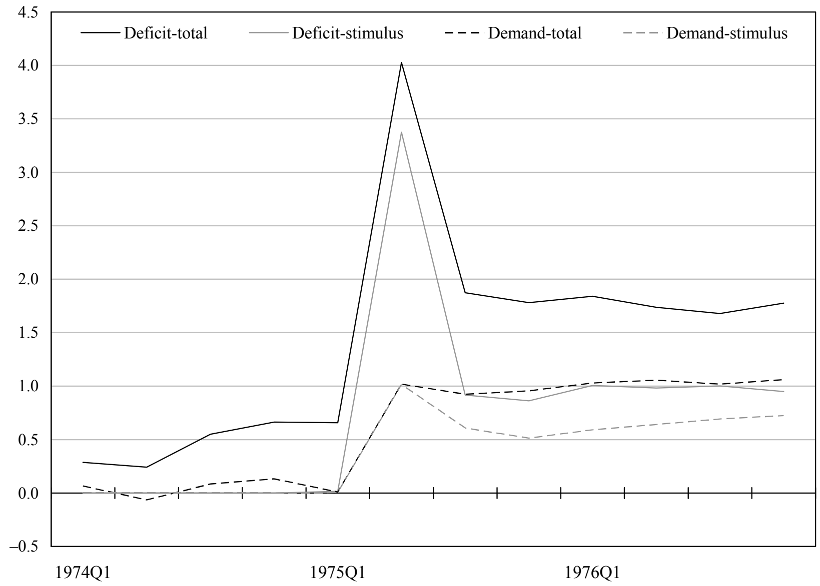


Figure 4

Discretionary Policies and Demand: 1974-76



---

[14] The Reagan Administration also cut some spending programs, particularly grants. Of the change in the deficit, 60 per cent was from increased defense spending, 70 per cent was from tax cuts and –30 per cent was from grants and other spending. Over the 2001-3 period, defense purchases contributed 25 per cent, non-defense purchases 10 per cent, and grants 15 per cent of the increase in the deficit owing to discretionary policy actions, with tax cuts the remainder.

Table 2

Direct Fiscal Multipliers
(percent of GDP)

Year	Total				Excluding Defense			
	Recession	*t*=1	*t*=2	3-Year Average	Recession	*t*=1	*t*=2	3-Year Average
2008	0.5	0.6	0.8	0.7	0.4	0.5	0.7	0.6
2001	0.7	0.8	0.8	0.8	0.6	0.7	0.7	0.7
1990	–0.5	0.7	0.7	0.6	–1.0	0.5	0.7	0.7
1981	1.6	0.7	0.8	0.8	0.6	0.4	0.7	0.6
1973	0.6	0.2	0.8	0.4	0.4	0.4	0.8	0.6
1970	–1.2	–1.7	–0.2	–0.7	0.3	0.6	0.6	0.5
1960	0.8	1.0	0.9	1.0	0.8	1.0	0.6	1.0
1957	0.6	0.8	1.7	1.0	1.3	0.8	1.3	1.0
1953	2.3	1.3	1.2	1.3	0.4	0.4	0.4	0.4

The multipliers are calculated as the ratio of the demand effect in the current period (from both current and previous policy actions), divided by current policy actions. No follow-on multiplier effects are included.

Looking at fiscal policy actions during the two years following the business cycle peak, we see that discretionary fiscal policies excluding defense spending pack little power. On average the ratio of the direct boost to demand relative to a sustained increase in the deficit is only about 0.4 in the recession year and about 0.6 in year *t*+1. The increase is largely accounted for by the lagged response of consumption to tax cuts and increased transfers. This is similar to the effectiveness of the stimulus programs put in place in 2008-10.

Indirect aggregate demand effects

The estimates above are only of the direct effects of fiscal policy actions, effectively translating fiscal actions into aggregate demand shocks. Table 2 reports the fiscal demand shocks as a ratio to the budget effects, which we label as the direct fiscal multiplier. The total effect of fiscal policy on the economy depends critically on the stance of monetary policy. Coenen *et al.* (2012) examine fiscal policy simulations using the structural models used by the IMF, Federal Reserve and other organizations. Their estimates suggest that if monetary policy is accommodative then the total fiscal multiplier would be in the range of 1.2 to 2.2 times that of the direct effect after two years. By contrast, with monetary policy not accommodative, then the multiplier falls to a range of 0.7 to 0.9 of the direct effect, figures below unity owing to crowing out of domestic demand and net exports.

Automatic stabilizers

This section considers the effect of the automatic stabilizers on aggregate demand to provide a point of comparison to the magnitudes of the discretionary actions taken by policymakers. See

Table 3

Recessions, Automatic Stabilizers, and Aggregate Demand

Cycle	Cyclical Deficit			Induced Demand			Unemployment Rate			GDP Gap		
	***t*–1**	**Recession**	***t*+1**	***t*–1**	**Recession**	***t*+1**	***t*–1**	**Recession**	***t*+1**	***t*–1**	**Recession**	***t*+1**
2008Q1	0.0	0.8	2.4	–0.1	0.2	1.1	4.6	5.8	9.3	–0.2	–2.6	–7.3
2001Q1	–1.1	–0.1	0.6	–0.6	–0.3	0.1	4.0	4.7	5.8	2.5	0.0	–1.5
1990Q3	–0.4	0.6	1.3	–0.2	0.1	0.6	5.3	6.3	7.3	1.0	–1.7	–3.0
1981Q3	0.7	1.2	2.2	0.2	0.6	1.4	7.5	8.5	10.3	–1.7	–4.2	–7.0
1973Q4	–0.8	–0.2	1.5	–0.2	–0.3	0.6	5.0	5.2	8.1	2.6	0.1	–4.1
1970Q1	–1.2	0.0	0.4	–0.8	–0.4	0.2	3.5	5.0	5.9	3.5	–0.2	–0.6
1960Q2	–0.2	0.4	0.6	–0.2	–0.1	0.1	5.3	6.0	6.4	0.4	–1.6	–1.8
1957Q3	–0.6	0.1	0.3	–0.2	–0.1	0.1	4.1	5.7	6.2	2.6	–0.8	–1.1
1953Q3	–1.3	–0.4	0.0	–0.4	–0.3	–0.1	2.8	4.4	5.1	5.7	1.8	0.7
Average	–0.5	0.3	1.0	–0.3	–0.1	0.4	4.7	5.7	7.1	1.8	–1.0	–2.9

Cyclical Deficit measures the cyclical effect of the economy on the budget surplus as a percent of potential GDP. A positive sign indicates that the actual surplus is higher than the high-employment surplus.
Induced Demand measures the contribution to the level of GDP, as a percent of GDP, from the demand induced by the cyclical swing in transfers and taxes.
Unemployment rate is the average unemployment rate for the period.
GDP gap measures the difference between actual and potential GDP as a percent of potential GDP.
Recession designates the four quarters following the business cycle peak, beginning in quarter indicated.
t–1 designates the four quarters before the "recession" year.
t +1 designates the four quarters following the "recession" year. For 1973, 1981 and 2008 recessions it includes some recessionary quarters.
The brief 1980 recession is omitted because much of the post recession period overlaps with the 1981 recession.

Follette and Lutz (2010) for details describing the estimation procedure for the budget effects of the automatic stabilizers. The methodology for calculating the aggregate demand effect arising out of these changes in government activity is discussed below.

The automatic stabilizers are primarily composed of personal and corporate income taxes, social insurance taxes, and unemployment benefits. Most of the budgetary effect is on the tax side of the ledger: We estimate that for every 1 percentage point swing in cyclical GDP there is a 0.35 percentage point increase in the federal deficit with 0.3 percentage point coming from taxes. The aggregate demand effects are a bit less unbalanced because the marginal propensity to consume (MPC) out of cyclical corporate taxes is probably tiny and that of transfers tends to be larger than that of personal taxes. Table 3 provides our estimate of the budget effects and aggregate demand effects of the automatic stabilizers in the year of the recession and the following year. By comparison we also show the depth of the recession. A key take-away is that the offset to the weakness in aggregate demand that is provided by the automatic stabilizers is modest. This is largely a consequence of our assumption that the marginal propensity to consume (MPC) for these policies is small initially, in line with responses by consumers to income in general. Given that this income is temporary, the small MPC is probably appropriate, but given that it is delivered in a targeted fashion to those with income losses, the actual MPC may be higher than assumed, particularly to the extent that consumers are liquidity constrained. A second observation is that the amount of support to aggregate demand from the automatic stabilizers is frequently much less than that provided by discretionary actions.

Conclusion

Fiscal policy has frequently been stimulative during recessions and early during the recovery. Much of the stimulus has come from policies that were put in place for non-counter-cyclical reasons, notably defense spending and structural changes to the tax system. Owing to the composition of policies chosen, on average, a 1 percentage point increase in the deficit for two years is estimated to boost demand by 0.4 percent of GDP in the first year and 0.6 per cent of GDP in the second year. This was the case for the 2008-10 stimulus program, for example. One reason for the low direct multipliers is that it is difficult to increase federal purchases quickly, and federal transfer programs and grants programs that may have relatively high multipliers are generally small and slow to implement. Accordingly, tax cuts have been an important component of stimulus programs, but they are not particularly effective.

Further work may be fruitful in two areas: improved measurement of the fiscal policy changes, and estimating the aggregate demand response. With regard to measuring policy changes, more attention can be given to precise timing (when the policy is announced versus when it is implemented), size (initial estimates by budget agencies versus *ex post* values), and defining the baseline. With regards to the latter, the U.S. has shifted to inflation-indexed tax and benefit systems in the 1970s and 1980s. As noted earlier, in the 1970s some tax policies are scored as tax cuts even when they allow effective tax rates to rise due to bracket creep. Moreover, the failure to adjust the tax code during episodes of high inflation should arguably be scored as a tax increase. Our survey of the empirical literature with regards to the to the demand effects of policy actions indicates that the direct effect on consumption from tax and transfer changes is better understood than the changes in state and local spending to federal aid, or the changes in investment to temporary tax credits (partial expensing, first time home buyers credits, etc). Better understanding of the state and local government response to temporary increases in aid would particularly useful.

REFERENCES

Agarwal, S., C. Liu and N.S. Souleles (2007), "The Reaction of Consumer Spending and Debt to Tax Rebates – Evidence from Consumer Credit Data", Federal Reserve Bank of Chicago, Working Paper, No. 2007-10.

Blinder, A. and M. Zandi (2010), "How the Great Recession Was Brought to an End", unpublished paper, Moody's Analytics, available at: economy.com/mark-zandi

Brayton, F. and P. Tinsley (1996), "A Guide to FRB/US: A Macroeconomic Model of the United States", Federal Reserve Board, Washington (D.C.).

Coenen, G., C.J. Erceg, C. Freedman, D. Furceri, M. Kumhof, R. Lalonde, D. Laxton *et al.* (2012), "Effects of Fiscal Stimulus in Structural Models", *American Economic Journal: Macroeconomics*, Vol. 4, No. 1, pp. 22-68.

Cohen, D. and J. Cummins (2006), "A Retrospective Evaluation of the Effects of Temporary Partial Expensing", Board of Governors of the Federal Reserve System (U.S.) Finance and Economics Discussion Series, No. 2006-19.

Cogan, J.F. and J.B. Taylor (2011), "What the Government Purchases Multiplier Actually Multiplied in the 2009 Stimulus Package", unpublished paper, Stanford University, January.

DeLong, J.B. and L.H. Summers (2012), "Fiscal Policy in a Depressed Economy", Brookings Papers on Economic Activity, Spring.

Follette, G. and B.F. Lutz (2011), "Fiscal Policy in the United States: Automatic Stabilizers, Discretionary Fiscal Policy Actions, and the Economy", in Banca d'Italia (ed.), *Fiscal Policy: Lessons from the Crisis*, Banca d'Italia.

Gramlich, E. (1969), "State and Local Governments and their Budget Constraint", *International Economic Review*, Vol. 10, pp. 163-82, June.

House, C.L. and M.D. Shapiro (2008), "Temporary Investment Tax Incentives: Theory with Evidence from Bonus Depreciation", *American Economic Review*, Vol. 98, No. 3, pp. 737-76, June.

Knight, B. (2002), "Endogenous Federal Grants and Crowd-Out of State Government Spending: Theory and Evidence from the Federal Highway Aid Program", *American Economic Review*, Vol. 92, pp. 71-92, March.

National Association of State Budget Officers (2012), "Fiscal Survey of the States", Fall.

Parker, J.A., N.S. Souleles, D.S. Johnson and Robert McClelland (2011), "Consumer Spending and the Economic Stimulus Payments of 2008", NBER, Working Paper, No. 16684.

Perotti, R. (2012), "The 'Austerity Myth': Gain Without Pain?", in A. Alesina and F. Giavazzi (eds.), *Fiscal Policy after the Financial Crisis*, University of Chicago Press.

Ramey, V.A. (2011), "Identifying Government Spending Shocks: It's All in the Timing", *Quarterly Journal of Economics*, No. 126, February.

Romer, C. (2011), "What Do We Know About the Effects of Fiscal Policy? Separating Evidence from Ideology", lecture, Hamilton College, November.

Romer, C. and D. Romer (2009), "A Narrative Analysis of Postwar Tax Changes", University of California, Working Paper, Berkeley (CA).

——————(2010), "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks", *American Economic Review*, No. 100, June.

Sahm, C.R., M.D. Shapiro and J. Slemrod (2010), "Household Response to the 2008 Tax Rebates: Survey Evidence and Aggregate Implications", in J.R. Brown (ed.), *Tax Policy and the Economy*, MIT Press, Cambridge (MA).

——————(2011), "Check in the Mail or More in the Paycheck? Does the Effectiveness of Fiscal Stimulus Depend on How It Is Delivered?", *American Economic Journal: Economic Policy*, forthcoming.

Shapiro, M.D. and J. Slemrod (2009), "Did the 2008 Tax Rebates Stimulate Spending?", *American Economic Review,* Papers and Proceedings, No. 99, May.

Shoag, D. (2010), "The Impact of Government Spending Shocks: Evidence on the Multiplier from State Pension Plan Returns", unpublished paper, Harvard University.

Stein, H. (1994), *Presidential Economics*, American Enterprise Institute, Washington (D.C.).

Suárez Serrato, J.C. and P. Wingender (2010), "Estimating Local Fiscal Multipliers", unpublished paper, University of California, Berkeley (CA), November.

FISCAL CONSOLIDATION IN REFORMED AND UNREFORMED LABOUR MARKETS

Alessandro Turrini[*]

This paper estimates the impact of fiscal consolidation on unemployment and job market flows across EU countries using a recent database of consolidation episodes built on the basis of a "narrative" approach (Devries et al.*, 2011). Results show that fiscal consolidation does have a significant impact on cyclical unemployment, although not large. As expected, the impact of fiscal policy shocks on job separation rates is much stronger in low-EPL countries, while high-EPL countries suffer from a stronger reduction in the rate at which new jobs are created. Since a reduced job-finding rate corresponds to a longer average duration of unemployment spells, fiscal policy shocks also tend to raise the share of long-term unemployment in high-EPL countries. Results are broadly confirmed when using "top-down" fiscal consolidation measures based on adjusting budgetary data for the cycle.*

1 Introduction

Since the outburst of the of the 2008 financial crisis, Europe is witnessing a worrying upsurge in unemployment and an unprecedented degree of dispersion of unemployment rates. The implementation of major and protracted fiscal consolidation strategies in such a context, and without prospects of a stable worldwide recovery, has stimulated debate on the growth and employment impact of consolidation measures, with implications for the coordination of timing and modalities of budgetary adjustment across EU countries (e.g., Corsetti, 2012).

Despite these concerns, a number of EU countries not only have recently put in place ambitious fiscal consolidation plans, but have also at the same time carried out major labour market reforms. In particular, the notoriously rigid and hard-to-reform Employment Protection legislation (EPL) systems of Southern European countries have been profoundly shaken with a view to stimulate job creation and tackle the problem of labour market segmentation at a juncture where severe budgetary cuts to reassure markets and put public finances on a sustainable footing where necessary.

Against this background, this paper aims at addressing a number of questions: to what extent continued fiscal consolidation across Europe would impact on unemployment? Which type of consolidation, expenditure or revenue-based, would be most employment-friendly? Does the impact of fiscal consolidation on unemployment come mostly from the job destruction side or does job creation play a relevant role as well? How do employment protection reforms interact with fiscal consolidation in determining unemployment and labour market flows? Are budgetary cuts more harmful when dismissals are less costly?

The analysis presented in this paper builds on various streams of existing literature. The literature on large episodes of fiscal consolidation focuses on the possible expansionary effects linked to the forward-looking behaviour of agents (e.g., Giavazzi and Pagano, 1990; Alesina *et al.*, 2002) and on the effectiveness of these episodes in durably improving the state of public finances (e.g., Alesina and Ardagna, 1998). Another stream of literature focuses on the estimation of fiscal

---

[*] DG ECFIN, European Commission.

I would like to thank David Leigh for guidance on the data and Alfonso Arpaia, Marco Buti and Lucio Pench for helpful discussions and suggestions.

The views expressed in this paper are those of the author and do not correspond necessarily with those of the European Commission.

multipliers. Most of the empirical literature based on structural VARs identifies fiscal shocks from a-priori information on the impact of the cycle on revenues and expenditures and generally find significantly positive multipliers, but seldom larger than one (e.g., Blanchard and Perotti, 2002; Perotti, 2005).

Analyses based on a "narrative", "action-based" approach to the identification of fiscal shocks, which requires a bottom-up computation of discretionary fiscal measures reported in official documents, also estimate significantly positive multipliers, but values are often large, well above unity (e.g., Romer and Romer, 2010; Guajardo *et al.*, 2011). Most empirical analyses on the impact of fiscal policy focus on output.

A few analyses look at the unemployment and labour market impact. Monacelli *et al.* (2010) develop a structural VAR for the US and estimate a negative and significant impact of government spending on unemployment and job creation, while job destruction falls.

The aim of this paper is to fill gaps in the existing literature in two main respects. First, it presents estimates of the impact of fiscal policy on unemployment and job market flows on EU countries: evidence is scarce for these countries. Second, it aims at shedding light on the interaction between fiscal consolidation and labour market regulation in driving labour market developments.

The baseline measure of fiscal consolidation used in the analysis is the action-based fiscal consolidation variable constructed in Devries *et al.* (2011), which present the double advantage of not including cyclical elements and being largely exogenous. As a countercheck, a "top-down" fiscal consolidation variable based on the cyclical adjustment of budgetary data is also used. The impact of fiscal consolidation is assessed on cyclical unemployment, on job separation and finding rates (hazard rates), and on the share of long-term unemployment. In light of limited sample size, econometric analysis spans the whole available panel of data for EU countries, but separate analysis is carried out for countries with a high vs. low degree of employment regulation.

Results confirm the finding that fiscal consolidation, notably government expenditure cuts have a significant impact on unemployment, although not large, and that this impact comes both from an increase in job destruction and a reduction in job creation. Interestingly, this unemployment impact does not differ much between high or low-EPL countries. There are considerable differences instead for what concerns job market flows, with fiscal consolidation in high-EPL countries having a less strong impact on job destruction but also leading to a more pronounced reduction in job finding rates.

The remainder of the paper is organised as follows. In the next section the data and the empirical strategy are illustrated. Section 3 presents results. The last section concludes with remarks on policy implications and suggestions for further analysis.

2 Data and empirical strategy

2.1 Data

The analysis focuses on EU countries and spans the 1980-2010 period, although lack of data availability for some countries and variables restricts the sample.

The baseline measure of fiscal consolidation is the "action-based" variable constructed in Devries *et al.* (2011). Data are collected over the period 1978-2009 for 17 OECD countries, 13 of which are EU countries (Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Portugal, Spain, Sweden, the UK). This action-based consolidation variable contains bottom-up estimates of the amount of measures taken by the government during years where the overall objective of fiscal policy, as reported in official statements and documents, was that of reducing the deficit and improving the state of public finances. If in a given year, in a given country, fiscal policy resulted in a reduction of the budget deficit and the reduction of the deficit, the variable reports the estimated amount of discretionary measures, separately for revenues and expenditures. In all other cases, the variable is set to zero, *i.e.*, there is no consolidation, either because the fiscal stance was expansionary or because fiscal contraction was mainly aimed at keeping under control domestic demand or at other purposes different than budgetary correction.

These "action-based" measures have a double advantage. First, they are not affected by the economic cycle, the reason being that their construction follows a bottom-up approach, *i.e.*, the amount of measures is computed by summing up estimates contained in official documents, so that cyclical movements in the budget are kept out from the start. Second, these consolidation measures are unlikely to imply risks of reverse causation because only the fiscal adjustment episodes ex-ante driven by the objective to adjust the budget are considered.

The analysis is complemented with the use of "top-down" fiscal consolidation measures. To this purpose, data on the change in the primary structural balance, structural revenues, primary structural expenditures from the DG ECFIN AMECO database are used, which are available for all EU countries (starting from 1995 only for countries having acceded the EU in 2004 or afterwards). Budgetary data are purged from the impact of the cycle and, for years after 2002, from one-off measures.

To address the issue of reverse causation, these top-down fiscal policy measures are instrumented using the variables normally used in the estimation of fiscal policy determinants by means of "fiscal reaction functions" (e.g., Bohn, 1998; Galí and Perotti, 2003). These variables are the own lag of the dependent variable, the lagged output gap, the lagged government/GDP ratio (the source for all instruments is the DG ECFIN AMECO database).

With a view to limit the analysis only to consolidation episodes all observations where the change in the instrumented fiscal balance is less than 0.5 per cent of GDP are set to zero. Hence, as in the case of the action-based fiscal variable measure, also this variable reports measures only in periods of fiscal consolidation that are unlikely to be related to the reaction of fiscal authorities to unemployment. The 0.5 per cent cut-off value for the instrumented change in the structural balance nets out minor consolidation episodes and permits to isolate a roughly equal number of consolidation episodes as those identified with the action-based approach over the sample period for the 13 EU countries for which data are available for both measures (120 action-based consolidation period, 117 top-down consolidation periods). The action-based and the top-down consolidation measures also exhibit a roughly similar average (respectively, 1.2 per cent of GDP and 0.8 per cent of GDP, respectively) and a rather high (0.38), statistically significantly rank correlation.

As for unemployment, the baseline variable used is the cyclical unemployment, as obtained from the difference between the overall unemployment rate and the NAWRU (source: AMECO database). The data are available for all EU27, but only starting from 1995 for countries that acceded the EU in 2004 or after. By dealing with cyclical unemployment, the risk of panel non stationarity is reduced, so that the complications linked to panel cointegration analysis are avoided. The underlying assumption is that, any impact of consolidation on unemployment is mostly arising from variations in cyclical unemployment.

Regarding data on job separation and job finding rates (hazard rates), the have been constructed as described in Arpaia and Curci (2010), following the methodology proposed by Shimer (2007). Data on job flows are available for all EU27 countries but for shorter time series compared with cyclical unemployment (going back to 1997 at the earliest). Data on the share of long-term unemployment on overall unemployment are taken from Eurostat, are available for all EU27 countries, and are available starting from 1992 at the earliest.

Figure 1 displays *prima facie* evidence of the link between cyclical unemployment and fiscal consolidation. Figure 1a reports for each country the action-based fiscal variable and cyclical unemployment figures. It appears that cyclical unemployment was quite often relatively high during the periods where fiscal consolidations took place. Figure 1b confirms this finding in a scatterplot that exhibits a positive, although weak relation between consolidation and cyclical unemployment across the panel. Of course, this prima-facie evidence does not imply causation but is suggestive of a possible link running from fiscal policy to unemployment outcomes.

2.2 Empirical strategy

The baseline regression framework used in the analysis of cyclical unemployment is as follows:

$$u_{i,t} = \alpha u_{i,t-1} + \beta u_{i,t-2} + \gamma FC_{i,t} + \theta_i + \eta_t + \varepsilon_{i,t} \quad (1)$$

where i, t denote country and year respectively, u is cyclical unemployment, FC is a consolidation variable, θ and η are, respectively, country and year fixed effects, while ε is a standard white noise error.

The specification amounts to an augmented AR2 model, which is motivated in light of broadly regular oscillations of cyclical unemployment around the mean (zero) over large samples.

In (1), the use of the simultaneous fiscal policy variable is justified in the case of action-based variables due to low risk of endogeneity and associated reverse causation problems. The top-down fiscal policy variables are instead instrumented to address the simultaneity issue.

The modelling of the impact of fiscal policy on other labour market variables is analogous to (1) except that, for the case of job market flows (hazard rates) and share of long-term unemployment, the second autoregressive term is dropped (being largely insignificant).

Equation (1) is estimated by means of panel fixed effect estimation (least square dummy variables) with robust standard errors for the case of action-based consolidation measures. For top-down measures, estimation is performed in two stages: first, the instrumenting regressions are run and the prediction obtained is "trimmed" in such a way to set to zero all observations corresponding to improvement in the instrumented primary structural balance below 0.5 per cent of GDP; second, panel regressions are run using the instrumented and trimmed consolidation variable.

With a view to shed light on the interaction between fiscal policy and labour market regulation, regressions are run separately for high and low EPL countries. The break down of countries is perfomed in the most straightforward way: countries with high (low) EPL are assumed to be those with an average value over the sample period of the OECD overall EPL indicator above the median of such averages across the whole panel of EU27 countries.

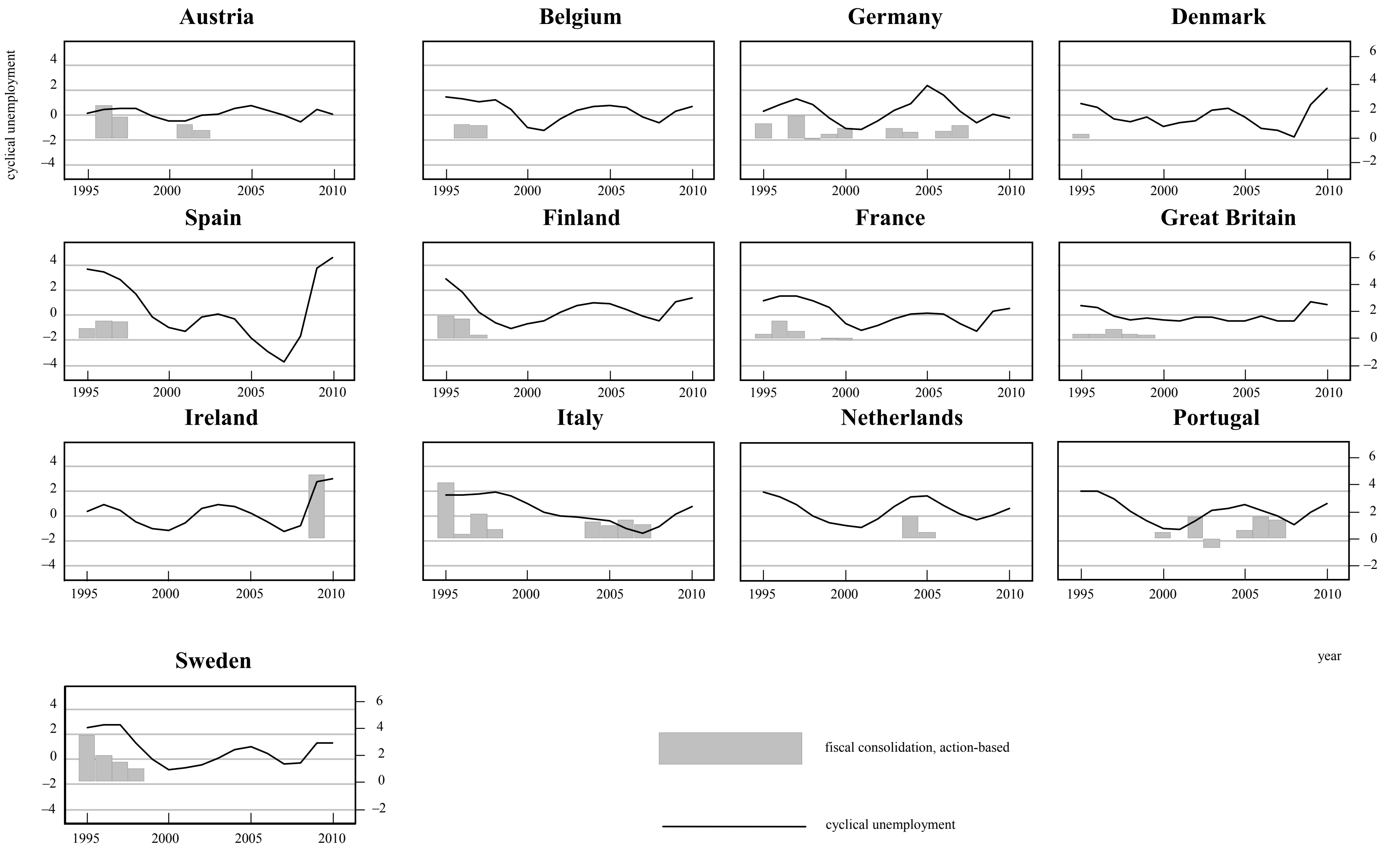
3 Results

3.1 Fiscal consolidation and unemployment

Table 1 reports results concerning the estimated impact of fiscal policy (action-based) on cyclical unemployment. The unemployment impact multiplier of the overall budgetary consolidation variable is positive but not large, amounting to less than 1/10 of a percentage point of unemployment for each GDP point of consolidation. While the impact of government revenue is non-significant, that of government expenditure is negative and higher in absolute value and of a higher order of significance that that for the overall budget balance.

Figure 1a

Cyclical Unemployment and Fiscal Consolidations (Action-based), 13 EU Countries, 1995-2009



Due to the auto-regressive process of unemployment, the peak multiplier is above the impact multiplier, as the adjustment of unemployment to the fiscal shock takes time. As shown in Figure 2, the peak effect materializes after one year (reaching almost 0.1 per cent for the overall budget and about –0.18 for expenditure cuts) and decays to zero after about 5 years. Afterwards, cyclical unemployment tends gradually to revert to pre-shock levels due to its stationarity properties. The overall impact of fiscal policy on unemployment has to take into account the sum of effects (overall multiplier).

The unemployment impact of fiscal consolidation is similar if measured according to top-down variables and notwithstanding the sample used in this case comprises a larger number of countries (Table 2).

The impact of consolidation takes similar values also if measured on the overall unemployment rate rather than on cyclical unemployment (Table 3), with the exception of revenues, whose coefficient is in this case positive, even if non-significant. This result reassures for what concerns the use of cyclical unemployment as baseline variable, and

Figure 1b

Relation Between Consolidation and Cyclical Unemployment

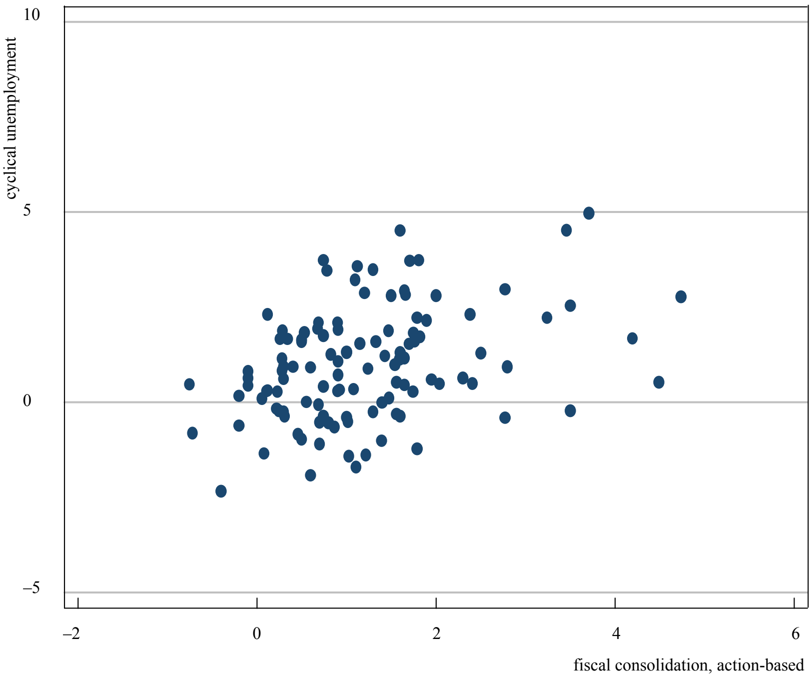


Figure 2

Fiscal Consolidation Impact on Cyclical Unemployment. Impulse Response Function

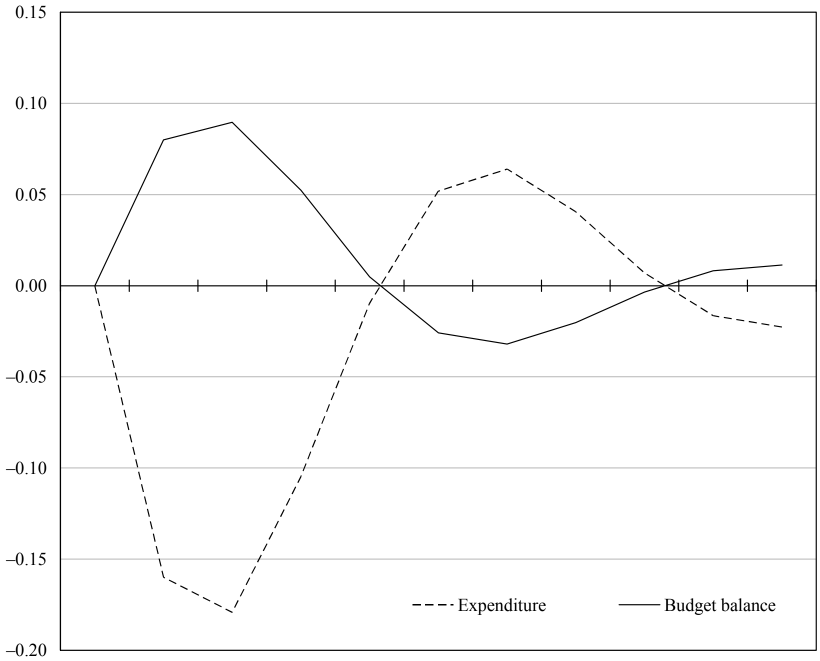


Table 1

Impact of Consolidation on Cyclical Unemployment,
Action-based Fiscal Policy Variables – 13 Countries EU, 1980-2009

Dependent Variable: Cyclical Unemployment	**(1) Budget Balance, Action-based**	**(3) Revenue, Action-based**	**(3) Expenditure, Action-based**
Explanatory variables:			
Cyclical unemployment (1 lag)	1.206 *(17.87)***	1.22 *(18.09)***	1.194 *(17.57)***
Cyclical unemployment (2 lags)	–0.609 *(7.50)***	–0.611 *(7.46)***	–0.607 *(7.56)***
Fiscal policy variable	0.08 *(1.65)*+	0.018 *(0.20)*	–0.16 *(2.35)**
Constant	0.307 *(1.58)*	–0.236 *(2.30)**	0.303 *(1.58)*
Observations	353	353	353
Number of countries	13	13	13
R^2	0.86	0.86	0.87

+, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method. fixed effect panel OLS, standard errors robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda:
Budget balance, action-based: year-on-year change in government budget balance associated with fiscal consolidation measures (source: Devries *et al.*, 2011).
Revenue, action-based: year-on-year change in government revenues associated with fiscal consolidation measures on the revenue side (source: Devries *et al.*, 2011).
Expenditure, action-based: year-on-year change in government expenditure associated with fiscal consolidation measures on the expenditure side (source: Devries *et al.*, 2011).

indicates that most of the effect of fiscal policy on unemployment falls on the cyclical component of the unemployment, with relatively minor implications for the NAWRU.[1]

Turning to the impact of fiscal policy on job market flows (Tables 4-7), it turns out that, in line with expectations, fiscal consolidation has a positive and significant impact on separation rates. Action-based consolidation measures have all significant coefficients, while in the case of top-down measures the coefficient of revenues lacks significance. Results are also broadly in line with expectation for what concerns job finding rates. In this case regression coefficients do not reach significance levels but the signs of the coefficients of all variables indicate a negative impact of consolidation on job finding rates, irrespective how consolidation is measured. Moreover, *t*-statistics take all values between 1 and 1.5, not far from cut off values for statistical significance at 10 per cent level.

Results concerning the impact of fiscal consolidation on the share of long-term unemployment do not lend themselves to an obvious interpretation. While the impact appears to be largely insignificant using action-based variables, top-down consolidation variables yield a

1 These conclusions are, however, to be taken with caution in light of the risk of inconsistent estimates in Table 3 arising from the likely non-stationarity of the unemployment rate, revealed, *inter alia*, by the high first-order auto-regressive coefficient.

Table 2

Impact of Consolidation on Cyclical Unemployment,
"Top-down" Fiscal Policy Variables – EU27, 1980-2010

Dependent Variable: Cyclical Unemployment	**(1)** Change in Structural Balance	**(2)** Change in Structural Revenue	**(3)** Change in Structural Primary Expenditure
Explanatory variables			
Cyclical unemployment (1 lag)	1.098 *(15.41)***	1.094 *(15.47)***	1.063 *(15.01)***
Cyclical unemployment (2 lags)	–0.491 *(6.11)***	–0.485 *(6.11)***	–0.458 *(5.64)***
Fiscal policy variable	0.142 *(1.61)*	–0.037 *(0.88)*	–0.138 *(2.28)**
Constant	–0.355 *(2.96)***	–0.289 *(2.38)**	–0.31 *(2.43)**
Observations	546	547	548
Number of countries	27	27	27
R^2	0.75	0.75	0.74

+, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method: fixed effect panel instrumental variables (the fiscal policy variable is instrumented with its own lag, the lagged output gap, and the lagged government debt/GDP ratio). Standard errors are robust with respect to heteroschedasticity and non-independence within country clusters.

Legenda:
Change in structural balance =year-on-year change in cyclically-adjusted government budget balance, information on one-off measures netted out when available (source: ECFIN AMECO database)
Change in structural revenue = year-on-year change in cyclically-adjusted government revenues, information on one-off measures netted out when available (source: ECFIN AMECO database).
Change in structural expenditure= year-on-year change in cyclically-adjusted government primary expenditure, information on one-off measures netted out when available (source: ECFIN AMECO database).
Consolidation episodes: change in instrumented structural balance > 0.5 % GDP.

significant negative impact for revenue increases, while the effect of expenditure cuts is positive. *A priori*, there is no clear expectation on the impact effect of fiscal consolidation on the share of long-term unemployment. On the one hand, since fiscal policy retrenchment implies more job dismissals, the increase of unemployment inflows would lead to a reduction of the share of long-term unemployment. On the other hand, the reduction of job finding rates linked to fiscal consolidation would play in the opposite sense: longer spells into unemployment for those already jobless, and a consequent in crease the long-term unemployment share. In light of these opposite effects, I find no surprising that results are non-significant or ambiguous in this case.

3.2 *The role of employment regulations*

The next step in the analysis aims at estimating separately the impact of consolidation on unemployment for high and low EPL countries, with a view to assess the interplay between the unemployment effects of fiscal policy and the role of labour market regulations.

Table 3

**Impact of Consolidation on Unemployment,
Action-based Fiscal Policy Variables – 13 Countries EU, 1980-2009**

Dependent Variable: Unemployment Rate	(1) Budget Balance, Action-based	(3) Revenue, Action-based	(3) Expenditure, Action-based
Explanatory variables:			
Unemployment (1 lag)	1.459 *(19.94)***	1.481 *(20.14)***	1.457 *(20.27)***
Unemployment (2 lags)	–0.589 *(7.80)***	–0.603 *(7.90)***	–0.589 *(7.92)***
Fiscal policy variable	0.129 *(1.87)*+	0.145 *(1.20)*	–0.179 *(1.65)*+
Constant	0.97 *(5.05)***	0.928 *(4.73)***	0.987 *(5.11)***
Observations	353	353	353
Number of countries	13	13	13
R^2	0.93	0.93	0.93

+, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method. fixed effect panel OLS, standard errors robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda:
Budget balance, action-based: year-on-year change in government budget balance associated with fiscal consolidation measures (source: Devries *et al.*, 2011).
Revenue, action-based: year-on-year change in government revenues associated with fiscal consolidation measures on the revenue side (source: Devries *et al.*, 2011).
Expenditure, action-based: year-on-year change in government expenditure associated with fiscal consolidation measures on the expenditure side (source: Devries *et al.*, 2011).

Table 4

**Impact of Consolidations on Job Separation Rates,
Action-based Fiscal Policy Variables – 13 EU, 1997-2009**

Dependent Variable: Job Separation Rates	(1) Budget Balance Action-based	(2) Revenue Action-based	(3) Expenditure Action-based
Explanatory variables:			
Job separation rate (1 lag)	0.778 *(8.24)***	0.783 *(8.31)***	0.776 *(8.16)***
Fiscal policy variable	0.03 *(2.54)**	0.054 *(2.39)**	–0.046 *(1.81)*+
Constant	–0.141 *(0.57)*	–0.145 *(0.59)*	–0.14 *(0.57)*
Observations	115	115	115
Number of countries	13	13	13
R^2	0.72	0.72	0.72

+, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method: fixed effect panel OLS, standard errors robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: See footnotes to Table 1.

Table 5

**Impact of Discretionary Fiscal Policy on Job Separation Rates,
"Top-down" Fiscal Policy Variables – EU27, 1997-2010**

Dependent Variable: Job Separation Rates	**(1)**	**(2)**	**(3)**
	Change in Structural Balance	**Change in Structural Revenue**	**Change in Structural Primary Expenditure**
Explanatory variables:			
Job separation rate (1 lag)	0.78	0.782	0.783
	*(11.27)***	*(11.23)***	*(11.34)***
Fiscal policy variable	0.041	–0.021	–0.014
	*(2.44)**	*(0.80)*	*(0.94)*
Constant	0.119	0.152	0.139
	(1.19)	*(1.54)*	*(1.46)*
Observations	225	225	225
Number of countries	27	27	27
R^2	0.68	0.67	0.67

+, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method: fixed effect panel instrumental variables (the fiscal policy variable is instrumented with its own lag, the lagged output gap, and the lagged government debt/GDP ratio). Standard errors are robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: See footnotes to Table 2.

Table 6

**Impact of Consolidations on Job Finding Rates
Action-based Fiscal Policy Variables – 13 EU, 1997-2009**

Dependent Variable: Job Finding Rates	**(1)**	**(2)**	**(3)**
	Budget Balance Action-based	**Revenue Action-based**	**Expenditure Action-based**
Explanatory variables:			
Job finding rate (1 lag)	0.718	0.718	0.72
	*(5.80)***	*(5.79)***	*(5.85)***
Fiscal policy variable	–0.305	–0.516	0.523
	(1.45)	*(1.57)*	*(1.46)*
Constant	3.645	3.646	3.631
	(1.30)	*(1.30)*	*(1.30)*
Observations	115	115	115
Number of countries	13	13	13
R^2	0.59	0.59	0.59

+, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method: fixed effect panel OLS, standard errors robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: See footnotes to Table 1.

Table 7

Impact of Consolidations on Job Finding Rates
"Top-down" Fiscal Policy Variables – EU27, 1997-2010

Dependent Variable: Job Finding Rates	(1) Change in Structural Balance	(2) Change in Structural Revenue	(3) Change in Structural Primary Expenditure
Explanatory variables:			
Job finding rate (1 lag)	0.665	0.666	0.661
	(6.31)**	*(6.24)***	*(6.22)***
Fiscal policy variable	–0.243	–0.262	0.148
	(1.26)	*(1.00)*	*(1.13)*
Constant	3.271	3.382	3.142
	(3.80)**	*(4.04)***	*(3.64)***
Observations	229	229	229
Number of countries	27	27	27
R^2	0.52	0.52	0.52

$^{+}$, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method. Columns: fixed effect panel instrumental variables (the fiscal policy variable is instrumented with its own lag, the lagged output gap, and the lagged government debt/GDP ratio). Standard errors are robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: See footnotes to Table 2.

Table 8

Impact of Consolidations on the Share of Long-term Unemployment,
Action-based Fiscal Policy Variables – 13 EU Countries, 1992-2009

Dependent Variable: Long-term Unemployment Share	(1) Budget Balance Action-based	(2) Revenue Action-based	(3) Expenditure Action-based
Explanatory variables:			
Long-term unemployment share (1 lag)	0.798	0.798	0.798
	(20.56)**	*(20.58)***	*(20.63)***
Fiscal policy variable	0.037	–0.001	0.107
	(0.10)	*(0.00)*	*(0.18)*
Constant	10.193	10.247	10.175
	(5.10)**	*(5.12)***	*(5.20)***
Observations	206	206	206
Number of countries	13	13	13
R^2	0.86	0.86	0.86

$^{+}$, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method: fixed effect panel OLS, standard errors robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: See footnotes to Table 1.

Table 9

Impact of Consolidations on the Share of Long-term Unemployment, "Top-down" Fiscal Policy Variables – EU27, 1992-2010

Dependent Variable: Long-term Unemployment Share	(1) Change in Structural Balance	(2) Change in Structural Revenue	(3) Change in Structural Primary Expenditure
Explanatory variables:			
Long-term unemployment share (1 lag)	0.686 *(16.50)***	0.689 *(16.38)***	0.684 *(16.37)***
Fiscal policy variable	–0.707 *(1.34)*	–0.783 *(1.95)*+	–0.382 *(1.06)*
Constant	15.485 *(6.85)***	16.558 *(7.95)***	16.333 *(7.71)***
Observations	368	368	368
Number of countries	27	27	27
R^2	0.7	0.7	0.7

+, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method. Columns: fixed effect panel instrumental variables (the fiscal policy variable is instrumented with its own lag, the lagged output gap, and the lagged government debt/GDP ratio). Standard errors are robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: See footnotes to Table 2.

Tables 10 and 11 report results for the impact on cyclical unemployment. When running the analysis separately for high-EPL and low-EPL countries, it is found that fiscal consolidations have a somehow larger effect in regulated labour markets, even though, most probably in light of the reduction in sample size, the estimated fiscal policy effect is not anymore significant when the sample is split according to EPL.

The result that fiscal consolidation is not less harmful in more regulated labour markets runs against the intuition. The explanation could lie in the different behaviour of job creation and job destruction. It is well-known from existing theory and evidence that strict EPL is associated with lower exit rates from unemployment but also with a lower probability for the unemployed to find a new job (Mortensen and Pissarides, 1994; Gomez-Salvador *et al.*, 2004). It could be the case that in high-EPL countries fiscal policy shocks destroy less jobs but also lead to a stronger reduction in the rate at which new jobs are created, with a possibly overall strong effect on cyclical unemployment.

The estimation of the impact of fiscal consolidation on job market flows separately for high and low-EPL countries supports the above hypothesis. As shown in Tables 12 and 13, job separation rates rise significantly with fiscal retrenchments only in low-EPL countries. The result is particularly neat using action-based consolidation measures: discretionary changes in the overall budget balance, government revenue, government expenditure are all insignificant in high-EPL countries while they are largely significant and with the expected sign in low-EPL countries. Conversely, job separation rates appear to react mostly in high-EPL countries (Tables 14 and 15). The change in the overall balance leads to a significant reduction in job finding rates only in high-EPL countries, irrespective of the measurement of fiscal policy. The estimates using the action-based variable reveal that this is mostly the outcome of a different reaction of job finding rates to expenditure cuts: only in high-EPL countries the reduction of government expenditure and the associated fall in aggregate demand leads to a significant impact on hiring and job finding rates.

Table 10

Impact of Consolidations on Cyclical Unemployment by EPL Strictness, Action-based Fiscal Policy Variables – 13 EU Countries, 1980-2009

Dependent Variable: Cyclical Unemployment	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Budget Balance, Action-based		**Revenue, Action-based**		**Expenditure, Action-based**	
	Low EPL	**High EPL**	**Low EPL**	**High EPL**	**Low EPL**	**High EPL**
Explanatory variables:						
Cyclical unemployment (1 lag)	1.206 *(19.70)***	1.216 *(14.40)***	1.224 *(20.16)***	1.22 *(14.27)***	1.19 *(19.39)***	1.198 *(14.27)***
Cyclical unemployment (2 lags)	–0.62 *(10.27)***	–0.614 *(7.26)***	–0.628 *(10.33)***	–0.601 *(6.98)***	–0.619 *(10.37)***	–0.603 *(7.24)***
Fiscal policy variable	0.069 *(1.61)*	0.127 *(1.53)*	0.008 *(0.10)*	–0.014 *(0.09)*	–0.148 *(2.34)**	–0.273 *(2.31)**
Constant	0.427 *(2.13)**	–0.072 *(0.28)*	0.475 *(2.35)**	–0.049 *(0.19)*	0.427 *(2.17)**	–0.276 *(1.08)*
Observations	196	157	196	157	196	157
Number of countries	7	6	7	6	7	6
R^2	0.86	0.89	0.86	0.89	0.87	0.89

$^{+}$, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method: fixed effect panel OLS, standard errors robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: Fiscal variables: see footnote to Table 1. The grouping of countries with respect to the OECD overall EPL indicator is built on the basis of the median country-specific average value of the indicator over the sample period.

Table 11

Impact of Consolidations on Cyclical Unemployment by EPL Strictness, "Top-down" Fiscal Policy Variables – 21 EU Countries, 1980-2010

Dependent Variable: Cyclical Unemployment	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
	Change in Structural Balance		**Change in Structural Revenues**		**Change in Structural Primary Expenditures**	
	Low EPL	**High EPL**	**Low EPL**	**High EPL**	**Low EPL**	**High EPL**
Explanatory variables:						
Cyclical unemployment (1 lag)	1.168 *(18.32)***	1.211 *(10.17)***	1.163 *(18.42)***	1.193 *(10.28)***	1.162 *(18.34)***	1.199 *(10.38)***
Cyclical unemployment (2 lags)	–0.595 *(9.53)***	–0.59 *(4.21)***	–0.596 *(9.60)***	–0.576 *(4.21)***	–0.598 *(9.77)***	–0.591 *(4.31)***
Fiscal policy variable	0.04 *(0.59)*	0.076 *(0.88)*	–0.077 *(1.17)*	–0.081 *(1.44)*	–0.116 *(2.26)**	–0.133 *(2.07)**
Constant	0.518 *(3.37)***	–0.28 *(2.11)**	0.541 *(3.43)***	–0.223 *(1.56)*	0.61 *(2.98)***	–0.216 *(1.53)*
Observations	243	233	243	233	243	233
Number of countries	11	10	11	10	11	10
R^2	0.83	0.82	0.83	0.82	0.83	0.82

$^{+}$, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method. Columns: fixed effect panel instrumental variables (the fiscal policy variable is instrumented with its own lag, the lagged output gap, and the lagged government debt/GDP ratio). Standard errors are robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: Fiscal variables: see footnote to Table 2. The grouping of countries with respect to the OECD overall EPL indicator is built on the basis of the median country-specific average value of the indicator over the sample period.

Table 12

Impact of Consolidations on Job Separation Rates, Distinguishing by EPL Strictness, Action-based Fiscal Policy Variables – 13 EU Countries, 1997-2009

Dependent Variable: Job Separation Rates	(1)	(2)	(3)	(4)	(5)	(6)
	Budget Balance Action-based		Revenue Action-based		Expenditure Action-based	
	Low EPL	High EPL	Low EPL	High EPL	Low EPL	High EPL
Explanatory variables:						
Separation rate (1 lag)	0.233	0.825	0.268	0.827	0.262	0.825
	(1.07)	*(6.65)***	*(1.28)*	*(6.66)***	*(1.20)*	*(6.81)***
Fiscal policy variable	0.065	–0.027	0.12	–0.023	–0.105	0.088
	*(3.75)***	*(0.48)*	*(3.82)***	*(0.27)*	*(3.17)***	*(0.80)*
Constant	0.641	0.058	0.62	0.059	0.618	0.057
	*(3.97)***	*(0.65)*	*(3.96)***	*(0.64)*	*(3.78)***	*(0.65)*
Observations	63	52	63	52	63	52
Number of countries	7	6	7	6	7	6
R^2	0.61	0.83	0.6	0.83	0.59	0.83

+, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method: fixed effect panel OLS, standard errors robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: Fiscal variables: see footnote to Table 1. The grouping of countries with respect to the OECD overall EPL indicator is built on the basis of the median country-specific average value of the indicator over the sample period.

Table 13

Impact of Consolidations on Job Separation Rates, Distinguishing by EPL Strictness, "Top-down" Fiscal Policy Variables – 21 EU Countries, 1997-2009

Dependent Variable: Job Separation Rates	(1)	(2)	(3)	(4)	(5)	(6)
	Change in Structural Balance		Change in Structural Revenue		Change in Structural Expenditure	
	Low EPL	High EPL	Low EPL	High EPL	Low EPL	High EPL
Explanatory variables:						
Separation rate (1 lag)	0.652	0.776	0.705	0.761	0.694	0.78
	*(6.51)***	*(7.56)***	*(6.84)***	*(7.64)***	*(6.82)***	*(8.08)***
Fiscal policy variable	0.06	–0.026	0.013	–0.064	–0.02	–0.011
	*(3.54)***	*(0.48)*	*(0.43)*	*(1.27)*	*(0.92)*	*(0.48)*
Constant	0.15	–0.142	0.123	0.145	0.123	–0.157
	(1.68)+	*(0.56)*	*(1.31)*	*(1.25)*	*(1.36)*	*(0.63)*
Observations	102	83	102	83	102	83
Number of countries	11	10	11	10	11	10
R^2	0.67	0.75	0.65	0.76	0.65	0.75

+, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method: fixed effect panel instrumental variables (the fiscal policy variable is instrumented with its own lag, the lagged output gap, and the lagged government debt/GDP ratio). Standard errors are robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: Fiscal variables: see footnote to Table 2. The grouping of countries with respect to the OECD overall EPL indicator is built on the basis of the median country-specific average value of the indicator over the sample period.

Table 14

Impact of Consolidations on Job Finding Rates, Distinguishing by EPL Strictness, Action-based Fiscal Policy Variables – 13 EU Countries, 1997-2009

Dependent Variable: Job Finding Rates	(1)	(2)	(3)	(4)	(5)	(6)
	Budget Balance Action-based		Revenue Action-based		Expenditure Action-based	
	Low EPL	High EPL	Low EPL	High EPL	Low EPL	High EPL
Explanatory variables:						
Finding rate (1 lag)	0.837	0.659	0.841	0.671	0.835	0.655
	*(4.61)***	*(4.34)***	*(4.69)***	*(4.36)***	*(4.61)***	*(4.40)***
Fiscal policy variable	–0.146	–1.663	–0.173	–1.761	0.338	4.292
	(0.67)	*(1.96)*+	*(0.44)*	*(1.52)*	*(0.82)*	*(2.26)**
Constant	0.189	3.315	0.088	3.285	0.231	3.73
	(0.07)	*(2.45)**	*(0.03)*	*(2.36)**	*(0.09)*	*(1.35)*
Observations	63	52	63	52	63	52
Number of countries	7	6	7	6	7	6
R^2	0.58	0.68	0.58	0.67	0.58	0.68

+, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method: fixed effect panel OLS, standard errors robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: Fiscal variables: see footnote to Table 1. The grouping of countries with respect to the OECD overall EPL indicator is built on the basis of the median country-specific average value of the indicator over the sample period.

Table 15

Impact of Consolidations on Job Finding Rates, Distinguishing by EPL Strictness, "Top-down" Fiscal Policy Variables – 21 EU Countries, 1997-2010

Dependent Variable: Job Finding Rates	(1)	(2)	(3)	(4)	(5)	(6)
	Change in Structural Balance		Change in Structural Revenue		Change in Structural Expenditure	
	Low EPL	High EPL	Low EPL	High EPL	Low EPL	High EPL
Explanatory variables:						
Finding rate (1 lag)	0.825	0.618	0.83	0.639	0.829	0.636
	*(5.86)***	*(3.94)***	*(5.93)***	*(4.01)***	*(5.95)***	*(4.04)***
Fiscal policy variable	0.112	–1.286	–0.095	–0.257	0.135	0.419
	(0.56)	*(1.86)*+	*(0.22)*	*(0.48)*	*(0.59)*	*(1.28)*
Constant	0.123	4.473	0.064	3.512	1.636	2.908
	(0.08)	*(1.57)*	*(0.04)*	*(2.38)**	*(1.21)*	*(2.24)**
Observations	102	85	102	85	102	85
Number of countries	11	10	11	10	11	10
R^2	0.59	0.58	0.59	0.56	0.59	0.56

+, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method: fixed effect panel instrumental variables (the fiscal policy variable is instrumented with its own lag, the lagged output gap, and the lagged government debt/GDP ratio). Standard errors are robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: Fiscal variables: see footnote to Table 2. The grouping of countries with respect to the OECD overall EPL indicator is built on the basis of the median country-specific average value of the indicator over the sample period.

Table 16

Impact of Consolidations on the Share of Long-term Unemployment, Distinguishing by EPL Strictness EU27, Action-based Fiscal Policy Variables – 13 EU Countries, 1992-2009

Dependent Variable: Long-term Unemployment Share	(1)	(2)	(3)	(4)	(5)	(6)
	Budget Balance Action-based		Revenue Action-based		Expenditure Action-based	
	Low EPL	High EPL	Low EPL	High EPL	Low EPL	High EPL
Explanatory variables:						
Long-term unemployment share (1 lag)	0.758 *(12.08)***	0.789 *(13.94)***	0.779 *(13.28)***	0.811 *(14.56)***	0.746 *(11.47)***	0.78 *(13.46)***
Fiscal policy variable	–0.512 *(1.14)*	1.422 *(1.80)*$^{+}$	–0.528 *(0.69)*	1.249 *(1.01)*	1.098 *(1.35)*	–2.262 *(1.81)*$^{+}$
Constant	8.839 *(4.23)***	6.978 *(2.28)**	3.59 *(1.78)*$^{+}$	12.345 *(4.30)***	9.189 *(4.29)***	7.451 *(2.38)**
Observations	110	96	110	96	110	96
Number of countries	7	6	7	6	7	6
R^2	0.89	0.86	0.89	0.85	0.89	0.86

$^{+}$, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method: fixed effect panel OLS, standard errors robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: Fiscal variables: see footnote to Table 1. The grouping of countries with respect to the OECD overall EPL indicator is built on the basis of the median country-specific average value of the indicator over the sample period.

The fact that job market flows react differently to fiscal consolidation according to the EPL regime helps disentangling the impact of fiscal retrenchment on the share of long-term unemployment. Since a reduced job finding rate corresponds to a longer average duration of unemployment spells, one would expect that fiscal policy shocks also tend to raise the share of long-term unemployment in high-EPL countries. The evidence reported in Tables 16 and 17 supports this expectation. While, as discussed above, over the whole available sample fiscal consolidation does not exhibit a significant relation with the share of long-term unemployment, when separating countries according to EPL, a pattern emerges: the effect is more strongly positive in high-EPL countries.

4 Concluding remarks

Overall, the evidence confirms that fiscal consolidation does have a significant impact on cyclical unemployment, which peaks after one year and gradually fades away. Results indicate however that the impact of budgetary consolidation is rather moderate (less than 0.1 per cent of additional cyclical unemployment at peak for each GDP point of budgetary cuts) and significant only for measures on the expenditure side.

Results also show that while fiscal consolidation in regulated labour markets is not necessarily less harmful in terms of unemployment, there are well-grounded reasons to expect it to be more worrying in terms of unemployment composition, being high EPL associated with a stronger reduction in job creation and a higher incidence of long-term unemployment. In these

Table 17

Impact of Consolidations on the Share of Long-term Unemployment, Distinguishing by EPL Strictness, "Top-down" Fiscal Policy Variables – 21 EU Countries, 1992-2010

Dependent Variable: Long-term Unemployment Share	(1)	(2)	(3)	(4)	(5)	(6)
	Change in Structural Balance		Change in Structural Revenue		Change in Structural Expenditure	
	Low EPL	High EPL	Low EPL	High EPL	Low EPL	High EPL
Explanatory variables:						
Long-term unemployment share (1 lag)	0.71	0.695	0.723	0.707	0.721	0.691
	*(14.30)***	*(9.13)***	*(14.06)***	*(9.31)***	*(13.85)***	*(9.31)***
Fiscal policy variable	–1.365	–0.307	–0.262	–0.945	0.096	–0.412
	(1.93)+	*(0.38)*	*(0.44)*	*(1.93)*+	*(0.21)*	*(0.65)*
Constant	17.036	15.756	16.046	15.132	15.966	16.167
	*(5.20)***	*(4.17)***	*(5.01)***	*(4.02)***	*(4.78)***	*(5.02)***
Observations	155	153	155	153	155	153
Number of countries	11	10	11	10	11	10
R^2	0.82	0.68	0.81	0.68	0.81	0.68

+, **, ** denote statistical significance at the 10, 5, 1 per cent level respectively. *T*-tests are reported in square brackets.
Specification. All regressions include country and year fixed effects.
Estimation method: fixed effect panel instrumental variables (the fiscal policy variable is instrumented with its own lag, the lagged output gap, and the lagged government debt/GDP ratio). Standard errors are robust with respect to heteroschedasticity and non-independence within country clusters.
Legenda: Fiscal variables: see footnote to Table 2. The grouping of countries with respect to the OECD overall EPL indicator is built on the basis of the median country-specific average value of the indicator over the sample period.

respects, the findings bode well for the strategy recently followed by some EU countries and support the view that in the current juncture tackling the challenges facing the euro area requires a multi-pillar approach comprising both fiscal consolidation and courageous structural reforms (Buti and Padoan, 2012).

The findings in this paper have also implications for the feasibility of structural reforms during austerity periods. Although it is well-known that certain labour market reforms may be hard to square with fiscal consolidation because of their electoral (e.g., Buti *et al.*, 2010) or budgetary costs (e.g., Deroose and Turrini, 2005), governments with a strong mandate to bring public finances on a sustainable footing while taking courageous measures to improve to capacity of the economy to create jobs may be able to carry out austerity measures and reform employment protection at the same time.

Further analysis on this topic seems deserved, not only to further check robustness of results with respect to the measurement of fiscal policy, the specification of empirical equations, and the definition of the sample, but also to better qualify results in terms of which EPL policy settings matter most in driving results.

REFERENCES

Alesina, A. and S. Ardagna, (1998), "Tales of Fiscal Adjustment", *Economic Policy*, No. 27, pp. 489-45.

Alesina, A., S. Ardagna, R. Perotti and F. Schiantarelli (2002), "Fiscal Policy, Profits, and Investment", *American Economic Review*, No. 92, pp. 571-89.

Arpaia, A. and N. Curci (2010), "EU Labour Market Behaviour During the Great Recession", European Economy, Economic Paper, No. 405.

Blanchard, O. and R. Perotti (2002), "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output", *Quarterly Journal of Economics*, No. 117, pp. 1329-68.

Boeri, T. (2011), "Institutional Reforms and Dualism in European Labour Markets", *Handbook of Labour Economics*, Elsevier, Amsterdam, pp. 1173-236.

Bohn, H. (1998), "The Behaviour of Us Public Debt and Deficits", *Quarterly Journal of Economics*, No. 113, pp. 949-63.

Buti, M., A. Turrini, P. Van den Noord and P. Biroli (2010), "Reforms and Re-elections in OECD Countries", *Economic Policy*, No. 25, pp. 61-116.

Buti, M. and P.C. Padoan (2012), "From Vicious to Virtuous: A Five-point Plan for Euro-zone Restoration", CEPR, Policy Insight, No. 61.

Corsetti, G. (2012), "Has Austerity Gone Too Far?", Vox.org, 2 April.

Deroose, S. and A. Turrini (2005), "The Short-term Budgetary Implications of Structural Reforms: Evidence from a Panel of EU Countries", CEPR, Discussion Paper, No. 5217.

Devries, P., J. Guajardo, D. Leigh and A. Pescatori (2011), "A New Action-based Dataset of Fiscal Consolidation", IMF, Working Paper, No. WP/11/128.

Galí, J. and R. Perotti (2003), "Fiscal Policy and Monetary Integration in Europe", *Economic Policy*, Vol. 18, No. 37, pp. 533-72.

Gómez-Salvador, R., J. Messina and G. Vallanti (2004), "Gross Job Flows and Institutions in Europe", *Labour Economics,* Vol. 11, pp. 469-85.

Guajardo, J., D. Leigh and A. Pescatori (2011), "Expansionary Austerity: New International Evidence", IMF, Working Paper, No. WP/11/158.

IMF (2010), *World Economic Outlook*, October, IMF, Washington (D.C.).

Monacelli, T., R. Perotti and A. Trigari (2010), "Unemployment Fiscal Multipliers", *Journal of Monetary Economic*s, Vol. 97, No. 5, pp. 531-53.

Mortensen, D. and C. Pissarides (1994), "Job Creation and Job Destruction in the Theory of Unemployment", *The Review of Economic Studies,* Vol. 61, No. 3, pp. 397-415.

Perotti, R. (2005), "Estimating The Effects of Fiscal Policy in OECD Countries", proceedings, Federal Reserve Bank of San Francisco.

Romer, C.D. and D.H. Romer (2010) "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks", *American Economic Review*, Vol. 100, No. 3, pp. 763-801.

Shimer, D. (2007), "Reassessing the Ins and Outs of Unemployment", NBER, Working Paper, No. 13421.

THE MACROECONOMIC EFFECTS OF EXPENDITURE SHOCKS DURING GOOD AND BAD TIMES

Francesco Caprioli and Sandro Momigliano**

We study how the effects of expenditure shock on economic activity are influenced by the state of the economy on the basis of various autoregressive models and indicators of cyclical conditions. For Italy over the period 1982-2011 we find some, but not conclusive, evidence that expenditure multipliers tend to be higher in recessions than in expansions.

1 Introduction

The large stimulus packages implemented by governments in most advanced countries to contrast the global recession that begun in mid-2008 stimulated a large debate (see Corsetti *et al.*, 2010; Romer and Romer, 2010) and brought renewed attention to the old question of the usefulness of fiscal policy to smooth cyclical fluctuations. More recently, a similar debate stemmed from fiscal consolidation policies and focused on the size of fiscal multipliers (IMF, 2102).

The theoretical literature provides limited guidance on these issues, as the qualitative effects of fiscal policy are model-dependent (see Cogan *et al.*, 2009); the empirical evidence is still not conclusive either, although it suggests that fiscal expansions generally boost private consumption and output.[1]

It has been often pointed out that the effects of fiscal policy may depend on the state of the economy (e.g., Parker, 2011; IMF, 2012), but there is still little empirical research trying to assess how the size of fiscal multipliers varies over the cycle. Indeed, most of the existing empirical literature uses linear models which, by construction, are unable to capture any dependence of fiscal multipliers on the level of aggregate demand.

In this paper we contribute to the debate by estimating for the Italian economy some threshold VARs, which allow to analyse the influence of the state of the economy on the effects of expenditure shocks.

Our starting point is the Structural VAR (SVAR) employed in Caprioli and Momigliano (2011). Fiscal shocks are identified using the methodology developed by Blanchard and Perotti (2002), which delivers relatively efficient estimates in small samples as recently stressed by Chahrour *et al.* (2010).[2] The model includes two additional variables – government debt and foreign demand – with respect to the standard model found in the literature (5 variables: private

* Bank of Italy.
E-mail: francesco.caprioli@bancaditalia.it and sandro.momigliano@bancaditalia.it

We are indebted to Stefano Neri for encouragement and comments. Any remaining errors are our own. The views in this paper are exclusively the responsibility of the authors and should not be interpreted as those of the Bank of Italy.

1 See Coenen *et al.* (2010). The two main empirical approaches that attempt to assess the effects of fiscal policy have specific limits. Reliable and non-interpolated quarterly fiscal data over a sufficiently long period of time, a prerequisite for the VAR approach, exist only for a few countries. The "narrative" approach (*i.e.*, Ramey and Shapiro, 1997, and Edelberg *et al.*, 1999) is resource-intensive and intrinsically subjective, making it almost impossible to apply across countries.

2 Other approaches most commonly used to identify structural shocks are the sign restrictions on impulse responses (see Mountford and Uhlig, 2002), the dummy variable one (see, e.g., Romer and Romer, 2010) and the Choleski ordering one, see, e.g., Fatás and Mihov, 2001. The literature about the effects of fiscal policy using Vector Autoregression is large and to offer a comprehensive survey goes beyond the scope of this paper. See Blanchard and Perotti (2002), Perotti (2004), Fatás and Mihov (2001), Mountford and Uhlig (2002), Giordano *et al.* (2008), Ramey and Shapiro (1997), Edelberg *et al.* (1999), and Burriel *et al.* (2010) among many others.

GDP, inflation, interest rates, net revenue and government consumption). The inclusion of debt is important because it allows to take into account its influence on the fiscal authorities' decisions, particularly important in the case of Italy.[3] The inclusion of foreign demand is warranted by its strong influence on economic activity, Italy being a small open economy.

To take into account the state of the economy we first estimate the SVAR described above splitting the sample into "recessions" and "expansions" on the basis of the official chronology of the Italian economy published by Isco/Isae/Istat (Altissimo *et al.*, 1999; Istat, 2011), which identifies recessions following the methodology used for the US by the National Bureau of Economic Research.

Secondly, to assess the robustness of our results to different ways of identifying recessions and also to overcome the problem arising from the limited number of observations for such regime in the official chronology we estimate an Endogenous Threshold VAR (Fazzari *et al.*, 2012). This approach allows to endogenously identify recessions and expansions, based on an indicator of cyclical conditions. As in the previous analysis, the sample is split and two sets of parameters are estimated. We apply this method to two alternative indicators of cyclical conditions: private GDP growth in the previous year and the output gap.

Finally, we apply the Smooth Transition VAR model proposed by Auerbach and Gorodnichenko (2012) to the same two indicators. The model allows a gradual transition between recessions and expansion: in each quarter, the parameters of the model are a linear combination of two sets of values (corresponding to each regime), weighted by the degree of being in each regime.

The main results of this paper can be summarized as follows.

In the SVAR analysis which doesn't take into account the state of the economy, the response of private GDP to an expenditure shock is positive, hump-shaped and highly significant for approximately two years. The median value of the expenditure multiplier is equal to 1.04 on impact and reaches its peak (1.8) after three years.

When we split the sample in "expansion" and "recession" regimes, either on the basis of the official chronology or applying the ETVAR approach to our cyclical indicators, for both regimes we obtain impulse response functions (IRFs) broadly similar to those estimated for the full sample but generally less precise. Under the recession regime, the response of private GDP is larger and more prompt and the cumulative multiplier is stronger on average in the first year. However, the confidence bands of the estimate are relatively large, and the difference in the median value of the multiplier across regimes is not statistically significant.

Estimates are generally less precise and results become ambiguous when we apply the STVAR model. Depending on the cyclical indicator used, in recessions the median value of the expenditure multiplier is constantly higher or constantly lower.

In conclusion, our empirical investigation shows some weak evidence that expenditure multipliers are higher in recessions than in expansions. While this result is influenced by the limited size of the sub-samples, it seems to suggest that the differences in the multipliers have not been extremely large, at least in the period under examination.

The paper proceeds as follows. Section 2 describes the data. In Section 3 we outline the specification of the VAR, ETVAR and STVAR models and our identification strategy. In Section 4

[3] Other researchers have included public debt in a SVAR exercise examining fiscal multipliers. We broadly follow the methodology of Favero and Giavazzi (2007), who add a deterministic equation linking debt dynamics to the government budget balance. Chung and Leeper (2007) employ a conceptually similar approach. Creel *et al.* (2005) include public debt as an additional variable. This second approach allows the analysis of the effects of direct shocks on government debt. This, however, comes at the cost of estimating a higher number of parameters than actually needed, as the government budget constraint is disregarded.

we analyze the effects of government consumption shocks without distinguishing between states of the economy. In Section 5 we discuss the results when we distinguish between the two regimes (expansions and recessions). We conclude with Section 6.

2 Data and variables

We extend up to 2011:2 the database of quarterly cash fiscal data used in Caprioli and Momigliano (2011) on the basis of the Italian Ministry for the Economy and Finance Quarterly Report and the general government borrowing requirement published by the Bank of Italy. The specification includes seven variables: private GDP (*i.e.*, total GDP net of government consumption, y_t); the inflation rate (π_t) based on the private GDP deflator; the nominal interest rate on government debt (i_t); government consumption (g_t); net taxes (t_t); the debt-to-GDP ratio (d_t); and foreign demand (f_t).

As in Caprioli and Momigliano (2011), we include GDP net of government consumption instead of total GDP. This choice stems from the fact that cash government consumption has a different quarterly profile from the corresponding national accounts aggregate, which complicates somewhat the interpretation of the effects on total GDP of a shock to (cash) government consumption, as it cannot be assumed (contrary to the case of national accounts fiscal data) to have a one-to-one impact on aggregate demand. Moreover, excluding the government component of aggregate demand from total GDP allows us to answer directly the most relevant policy question, that is how the private sector reacts to a fiscal shock.

We construct the interest rate on government debt as a weighted average of the yield on short-term and on long-term government debt, where the weight is given by the share of debt obligations with maturity shorter than one year. Government consumption is the sum of government spending on goods and services and government wages. Net taxes are computed by subtracting government consumption, interest payments and investment from the borrowing requirement; therefore this variable includes monetary transfers as well as revenue.[4]

All variables, apart from inflation, interest rate and the debt-to-GDP ratio, are log-transformed, converted in real terms using the private GDP deflator and seasonally adjusted using the TRAMO-SEATS procedure.

To identify expansions and recessions, we use the following indicators: i) the official chronology of the Italian economy, based on the National Bureau of Economic Research, produced by Istat-Isae-Isco (cfr. Altissimo *et al.*, 1999); ii) past 4-quarters private GDP growth; iii) the output gap, computed on the basis of the Hodrick Prescott filter (as commonly found in the literature, λ is set equal to 1600).

3 The models and the identification strategy

3.1 The SVAR

The reduced-form VAR is specified in level (as shown by Sims *et al.*, 1990), in large samples it is possible to ignore the cointegrating vector) and can be written as follows:

---

[4] We exclude public investment from our benchmark specification (as in Giordano *et al.*, 2008), because we are not confident enough about the quality of the data. Results do not qualitatively change as a result of adding investment to either government consumption or net revenue, as shown in Subsection 4.2.

$$X_t = \sum_{i=1}^{k_1} C_i \; X_{t-i} + \sum_{i=1}^{k_2} \gamma_i \; d_{t-i} + \sum_{i=0}^{k_3} \delta_i \; \log(f_{t-i}) + U_t \tag{1}$$

where:

$$X_t = \begin{bmatrix} \log(y_t) \\ \pi_t \\ i_t \\ \log(t_t) \\ \log(g_t) \end{bmatrix} \tag{2}$$

k_1, k_2 and k_3 are the number of lags for the variables included in the VAR, for the debt-to-GDP ratio and for the foreign demand variable respectively.

U_t is the vector of reduced-form residuals. k_1, k_2 and k_3 are set to the minimum number of lags that delivers serially uncorrelated reduced-form residuals. In particular, they are set equal to 2, 1 and 1 respectively. A constant and a deterministic linear trend are included. According to equation (1), past values of the debt-to-GDP ratio influence the current values of macroeconomic variables, which conversely influence the current value of the debt-to-GDP ratio according to the following law of motion:

$$d_t = \frac{1 + R_t}{(1 + \pi_t)\left(\dfrac{y_t}{y_{t-1}}\right)} d_{t-1} + \frac{g_t - t_t}{y_t} \tag{3}$$

where:

$$R_t = \sum_{j=0}^{N} \frac{i_{t-j}}{N} \tag{4}$$

Equation (3) represents the period-by-period government budget constraint, expressed as a ratio to total GDP. Changes in the interest rate on government debt i_t only gradually affect its average cost R_t in equation (4); we set $N = 20$, as 5 years is approximately the financial duration of the debt at the end of our sample.

Compared with Favero and Giavazzi (2007), we add equation (4) and include in equation (1) the actual yield at issuance instead of the average cost of servicing public debt. We do so to identify more precisely the reaction of financial markets to the state of the public finances. In fact, the yield at issuance responds immediately to investors' sentiments, while the average cost adjusts with a relatively long delay, depending on the maturity structure of government obligations. Moreover, the yield at issuance is more directly relevant for investment decisions in the private sector.

We assume that, while current and past values of foreign demand affect the current values of macroeconomic and fiscal variables, the reverse is not true. This assumption seems appropriate as Italy is a relatively small open economy. As a measure of foreign demand, we follow Busetti *et al.* (2011), who compute the demand of Italian goods from abroad as:

$$f_t = \sum_{j=1}^{N} M_{j,t} \; \overline{q_j} \tag{5}$$

where $M_{j,t}$ corresponds to the total imports of goods by country j in volume at time t weighted by $\overline{q}_j$, the average ratio over the period 1999-2001 between Italian exports towards country j and total Italian exports. Busetti *et al.* (2011) construct this index for commercial partners both belonging to the Euro area and outside the EU. As a measure of global foreign demand, we consider the sum of the two indices.[5]

3.2 *Test for nonlinearities*

As the ETVAR and the STVAR models are non-linear multivariate system of equations, we test whether our data are consistent with the existence of non-linearities related to cyclical conditions, on the basis of equation (10) below.

$$X_t = \sum_{i=1}^{k_1} C_i X_{t-i} + \sum_{i=1}^{k_2} \gamma_i d_{t-i} + \sum_{i=0}^{k_3} \delta_i \log(f_{t-i}) + \sum_{i=1}^{k_1} \Lambda_i X_{t-i} z_t + U_t \tag{10}$$

In the equation, z_t is an indicator of cyclical conditions. If the true model is linear, then the coefficients in Λ are jointly insignificant. The likelihood ratio test, performed for each of our three indicators of cyclical conditions, rejects the null hypothesis that $\Lambda = 0$, giving support to the hypothesis that non-linearities are an important feature of the data.

3.3 *The ETVAR and STVAR models*

In the ET-VAR model, equation (1) is substituted by equations (6) and (7), while equations (2)-(5) remain unchanged.

$$\begin{aligned} X_t =& (1-F(z_{t-1}))[\sum_{i=1}^{k_1^E} C_i^E X_{t-i} + \sum_{i=1}^{k_2} \gamma_i^E d_{t-i} + \sum_{i=0}^{k_3} \delta_i^E log(f_{t-i}) + U_t^E] + \\ &+ F(z_{t-1})[\sum_{i=1}^{k_1^R} C_i^R X_{t-i} + \sum_{i=1}^{k_2} \gamma_i^R d_{t-i} + \sum_{i=0}^{k_3} \delta_i^R log(f_{t-i}) + U_t^R] \end{aligned} \tag{6}$$

Equation (7) states that the economic system is described by two piecewise linear models with different sets of coefficients. The recessionary and expansionary regimes are identified by the transition indicator function $F(z_{t-1})$, defined by:

$$F(z_{t-1}) = I(z_{t-1} \leq r) = \begin{cases} 1, & if z_{t-1} \leq r \\ 0 & \forall t >> r \end{cases} \tag{7}$$

where z_{t-1} is the threshold variable and r is the threshold value. The threshold variable is lagged to avoid contemporaneous feedback effects from the model to the probability. The model, which allows different lags for the autoregressive part across regimes and regime-specific covariance matrices, is estimated in two steps. First, for a given value of the threshold r, the regime-specific coefficients and covariance matrices are estimated by OLS using observations from each regime; second, the threshold value is estimated by minimizing the conditional likelihood over a grid of values, namely:

[5] As a robustness check, we use also the world trade series obtained from IMF International Financial Statistics. The use of this series to measure foreign demand does not change results.

$$\hat{r} = arg\min_{r\epsilon R}\left(\frac{T_R}{2}log|\hat{\Omega^R}_t| + \frac{T_E}{2}log|\hat{\Omega^E}_t|\right) \tag{8}$$

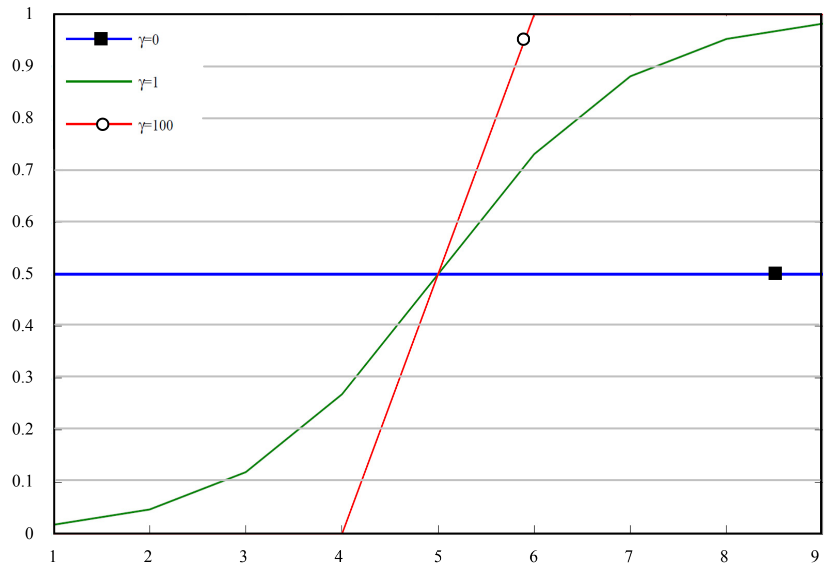
In the STVAR model the transition function takes a more general form, given by:

$$F(z_{t-1}) = \frac{exp(-\gamma z_{t-1})}{1+exp(-\gamma z_{t-1})}, \gamma > 0 \tag{9}$$

which allows a smoother transition across the two regimes than what prescribed by equation (7). The weighting function $F(z_{t-1})$, the probability to be in a recession, depends on the business cycle indicator z_{t-1}; as it is imposed $\gamma>0$ in equation (9), the lower the z_{t-1}, the higher $F(zt-1)$, as shown in Figure 1, for different values of the γ parameter. As for the case of the threshold value in ETVAR, in the STVAR the γ parameter is estimated by minimizing the likelihood function.

Figure 1

1–*F*(*z*): Probability To Be in an Expansion



3.4 *Identification strategy*

The identification strategy is identical for the three models described above. The only difference is that in the ETVAR the procedure is applied to the residuals of each regime.

Reduced-form residuals associated with the fiscal variables, u_t^g and u_t^t can be written as linear combinations of the structural fiscal shocks and of the reduced-form residuals of the other variables in the VAR:

$$u_t^g = \alpha_y^g\, u_t^y + \alpha_\pi^g\, u_t^\pi + \alpha_i^g\, u_t^i + \beta_t^g\, \varepsilon_t^t + \varepsilon_t^g \tag{7}$$

$$u_t^t = \alpha_y^t\, u_t^y + \alpha_\pi^t\, u_t^\pi + \alpha_i^t\, u_t^i + \beta_g^t\, \varepsilon_t^g + \varepsilon_t^t \tag{8}$$

The α coefficients contain both the automatic elasticity and the discretionary change to the macro variables innovations, while the β coefficients measure the response of the fiscal variables to a structural shock. To estimate the α and β coefficients in equations (7)-(8) we follow the approach in Blanchard and Perotti (2002). First, we assume that, within a quarter, the discretionary change of fiscal variables to innovations in the macro variables is zero. Using quarterly data, this assumption can be justified on the ground of decision lags in fiscal policy-making which last longer than three months. Secondly, we estimate the α in equations (7)-(8) using external information on the elasticities of government consumption and taxes to output, inflation and interest rate. Following Giordano *et al.* (2008) (Appendix B therein) in this paper, we set $\alpha_{\pi}^{g} = -0.9$, $\alpha_{y}^{t} = 0.3$, $\alpha_{\pi}^{t} = -0.4$ and all the other α equal to zero. In addition, we assume that government consumption does not contemporaneously adjust to revenues, *i.e.*, we set β_{t}^{g} equal to zero. Consequently, we estimate β_{g}^{t} from equation (8) using OLS. We verify that even sizeable changes in these parameters do not significantly affect our results.

Finally, we estimate the coefficients relating the reduced-form macro variables residuals to the fiscal ones by instrumental variables, using as instruments for u_{t}^{g} and u_{t}^{t} their corresponding structural shocks, uncorrelated by definition.

It is important to notice that the identification strategy for structural shocks does not depend on the presence of the debt-to-GDP ratio, as the latter follows a deterministic law of motion. In other words, equation (3) holds as an identity and therefore it does not add any shock to the ones already included in the VAR model specified in equation (1).

A problem with the fiscal shocks identified using the VAR approach is that they may be anticipated by economic agents, owing to the delay between the announcement of fiscal measures and their actual implementation. In order to check for this possibility, we run Granger causality tests between the fiscal shocks estimated with the benchmark model and survey expectations about future policy actions and macro variables. The results do not support the hypothesis that fiscal shocks were anticipated.[6]

4 The effects of government consumption shocks in a SVAR model

Figure 2 shows the response of the fiscal and macroeconomic variables to an exogenous shock (equal to 1 per cent of private GDP) to government consumption. In each panel the solid line represents the median response, while the dashed lines represent two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution of the responses at each horizon, as commonly done in the literature.[7]

Concerning the reaction of fiscal variables, two points are worth mentioning. The first is that the government consumption shock is largely short-lived, being equal to 0.1 per cent of private GDP already after four quarters. The second is that the higher public consumption is rapidly financed by higher revenues, which increase already in the first quarter, remain broadly constant at 0.2 per cent of GDP for two years and then slowly decrease. The rise in net revenue, ensuring that the initial surge in the debt is fully absorbed within three years, reflects their direct stabilizing discretionary reaction to the debt and, to a lesser extent, to the increase in private GDP (see below).

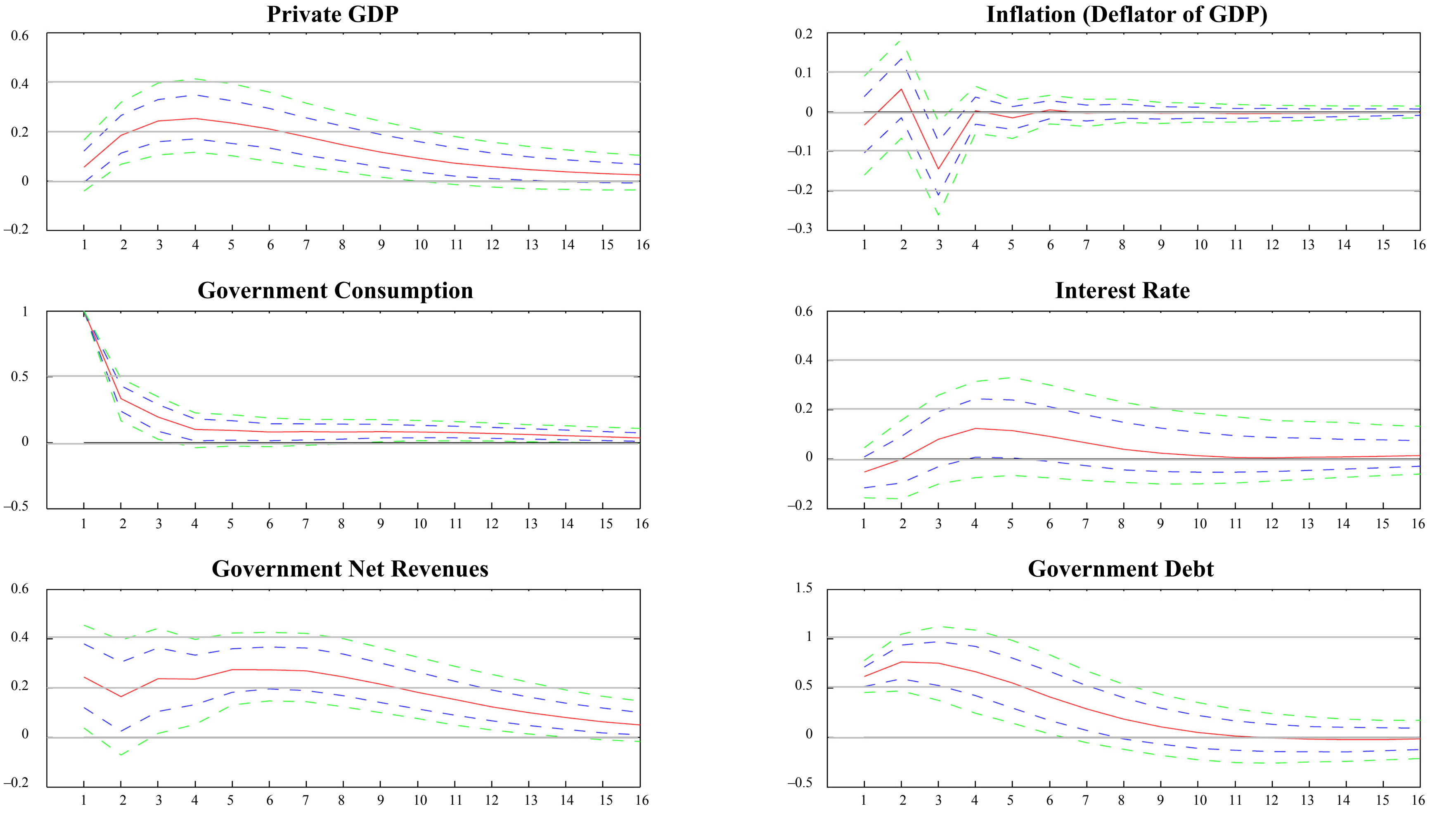
---

[6] As for survey expectations, we use the Consensus mean forecasts of *i*) the annual growth rate of real GDP, private consumption, gross fixed investment, industrial production, consumer and producer prices, *ii*) unemployment rate (as a percentage of the labor force), current account and state sector budget balance, and *iii*) three-month euro-area interest rate and 10-year Italian government bond yield. Following Ramney (2008) and Kirchner *et al.* (2010), the fiscal shocks at time t are regressed on a constant, its own lag and the previous forecasts made in period $t-1$ for period t.

[7] We compute confidence bands for IRF by bootstrapping. After estimating equation (1), we obtain fitted residuals $\hat{u}_1, \ldots, \hat{u}_T$ normally distributed with zero mean and covariance matrix Ω. We draw errors from this distribution to simulate the system of equations (1)-(5) L times. For each draw we compute the IRF as described in the previous footnote. Finally, we collect the α^{th} and $1-\alpha^{th}$ percentile across the L draws. In the simulation we set L=1000.

Figure 2

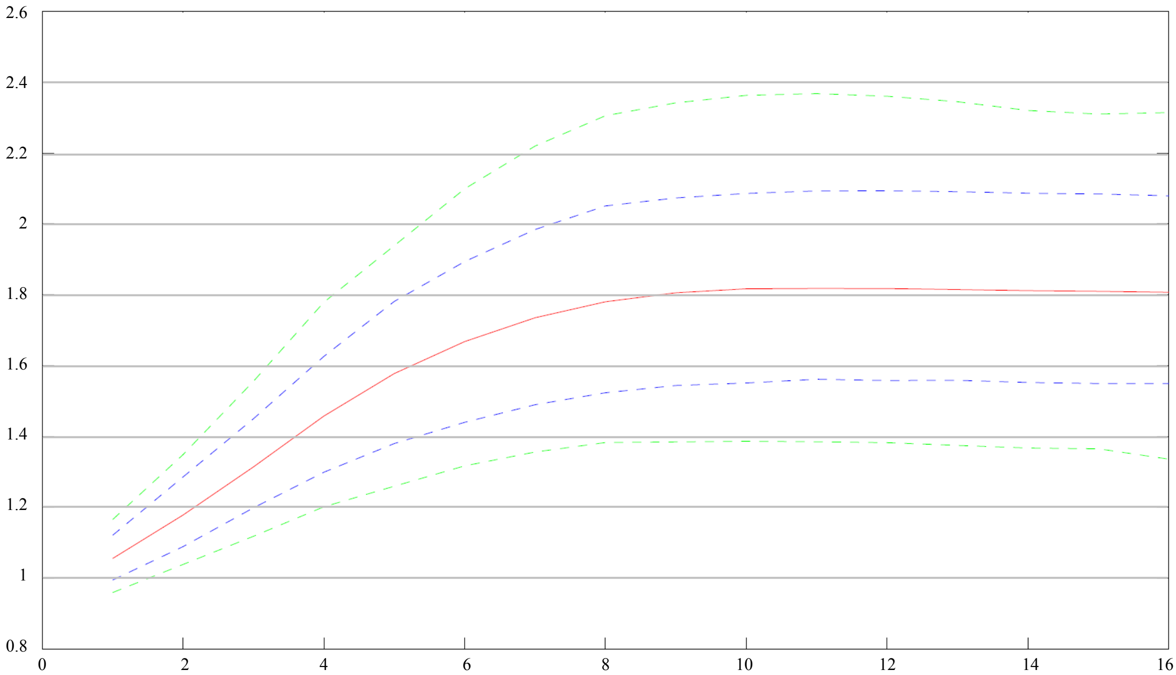
Impulse Responses to a Positive Government Consumption Shock Equal to 1 Per Cent of Private GDP: SVAR Model



The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution. Responses, except for inflation and interest rate, are deviations from the baseline and expressed in percentage points of private GDP. Inflation and interest rate responses are deviations from the baseline and expressed in percentage points.

Figure 3

Cumulative Multiplier of Government Consumption on GDP: SVAR Model



The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution.

After a shock to public consumption, the response of private GDP is positive and highly significant for approximately two years. The peak, reached at the fourth quarter, is equal to 0.25 per cent of GDP. Positive and significant effects of government consumption shocks on economic activity represent a relatively common result of the VAR literature (e.g., Giordano *et al.*, 2008; Perotti, 2004; Mountford and Uhlig, 2002; and Neri, 2001). The output response to government consumption reflects the low persistence of the shock. To make our results more directly comparable with analyses which focus on total GDP (instead of its private component) and analyze shocks with a different persistence, we compute the cumulative multiplier (*i.e.*, the ratio of the cumulative change in total GDP to the cumulative change in total government consumption)[8] charted in Figure 3. The median value is equal to 1.04 on impact, reaches its peak (1.8) after three years and remains roughly constant thereafter. The confidence bands are relatively narrow compared with similar studies, with the 95th and the 5th percentiles of the distribution remaining above 1.3 and below 2.4 after the fifth quarter.

The median value for the long-run fiscal multiplier lies in the upper part of the wide range of estimates provided by the empirical literature. As shown in Spilinbergo *et al.* (2009), the relatively high value of the multiplier may be due to the debt-stabilizing reaction of fiscal variables. The transitory nature of the government consumption shock, rapidly compensated by higher revenues, and the small – and delayed – increase in interest rates do not pose a threat to the sustainability of the Italian public debt, notwithstanding its high level, making any precautionary savings by

8 Following Giordano *et al.* (2008), we compute total GDP in this context by adding the cash-based government consumption included in the model to private GDP.

households unnecessary. The response of private GDP is robust across alternative specifications of the model.[9]

The reaction of inflation to a government consumption shock is not statistically significant. This is in line with the analyses of Marcellino (2006), King and Plosser (1985) and Henry *et al.* (2004). The response of interest rates is relatively small, hump-shaped and never statistically significant. The existence of a positive relationship between interest rates and the level of government debt can be found in many empirical studies (see Bernheim, 1987, 1989; Gale and Orzag, 2002; Miller and Russek, 1996; and Engen and Hubbard, 2004).[10]

4.1 The role of government debt and foreign demand

The model includes two additional variables – government debt and foreign demand – with respect to the standard model found in the literature (5 variables: private GDP, inflation, interest rates, net revenue and government consumption). The inclusion of debt is important because it allows to better understand the fiscal framework associated with the shock. In particular, the reaction of fiscal variables – namely, government spending and net revenue – to changes in public debt can be analyzed.[11] Empirical evidence (see Bohn, 2007; Trehan and Walsh, 1991; Hamilton and Flavin, 1986; and Golinelli and Momigliano, 2008) suggests that this feedback effect is generally important. In the case of a high-debt country like Italy, the influence of debt on the fiscal authorities' decisions is likely to be particularly large.[12] The inclusion of foreign demand is warranted by its strong influence on economic activity, Italy being a small open economy. As it can be safely assumed that foreign demand, measured by world demand, is not significantly influenced by Italian macro or fiscal variables, its inclusion in the VAR comes at a relatively small cost in terms of additional parameters to be estimated.

The left and right panels of Figure 4 show the impact of including public debt and/or foreign demand in the model respectively on the median response of private GDP and on the accuracy of this estimate, measured by the distance between the 95th and the 5th percentiles of the distribution.

[9] As robustness checks, we considered the following model specifications in which: i) we include the interest rate only on debt obligations with a maturity shorter than one yea; ii) we use the gross yield on debt obligations with a maturity longer than three years; iii) the specification of the VAR includes a quadratic trend instead of a linear one; iv) we include government investment in our definition of government consumption; v) net revenues come first when identifying the shocks (in the benchmark model, government consumption is ordered first); v*i*) the reduced-form residuals of fiscal variables depend explicitly on the level of government debt; and v*ii*) the average financial duration is set equal to two years instead of its end-of-sample value (five years). We do not report these robustness checks, as estimates stay almost unchanged with respect to the benchmark specification. The results obtained with these alternative specifications confirm the hump-shaped pattern of private GDP and, apart from the "quadratic trend" specification for the quarters 6-10, they are well within the upper (95th percentile) and lower (5th percentile) bands of the GDP response in the benchmark specification. In the case of the "quadratic trend" specification, the lower impact on private GDP largely reflects the shorter persistence of the expenditure shock. The cumulative multiplier is very close to that for the benchmark specification.

[10] The results for inflation and interest rates are also robust across the alternative specifications described in the previous footnote.

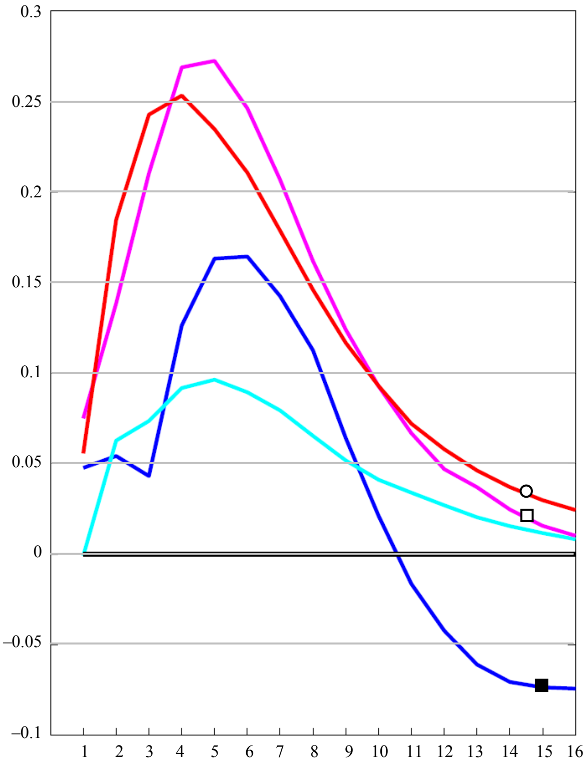
[11] Recent research suggests that, depending on whether or not an expenditure shock is reabsorbed in the medium-long term, fiscal multipliers may have different values (see Corsetti *et al.*, 2009; and Ilzetzki *et al.*, 2009).

[12] Other researchers have included public debt in a SVAR exercise examining fiscal multipliers. We broadly follow the methodology of Favero and Giavazzi (2007), who add a deterministic equation linking debt dynamics to the government budget balance. Chung and Leeper (2007) employ a conceptually similar approach. Creel *et al.* (2005) include public debt as an additional variable. This second approach allows the analysis of the effects of direct shocks on government debt. This, however, comes at the cost of estimating a higher number of parameters than actually needed, as the government budget constraint is disregarded.

Figure 4

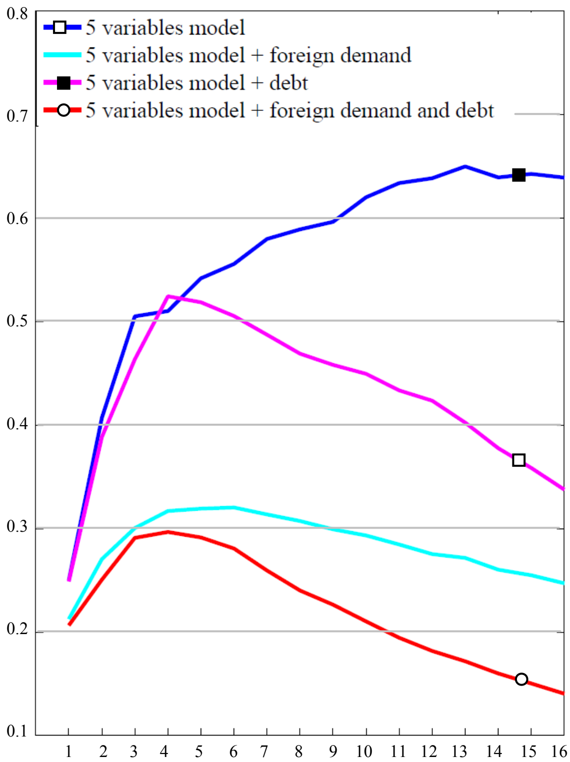
GDP and Shocks to Government Consumption

Effects on Private GDP of a Shock to Government Consumption
(median values; percent of private GDP)



SVAR model and alternative models which exclude debt and/or foreign demand.

Size of Confidence Bands of the Estimates of the Effects on Private GDP of a Shock to Government Consumption
(difference between the 95th and 5th percentiles of the distribution of the private GDP responses; percent of private GDP)



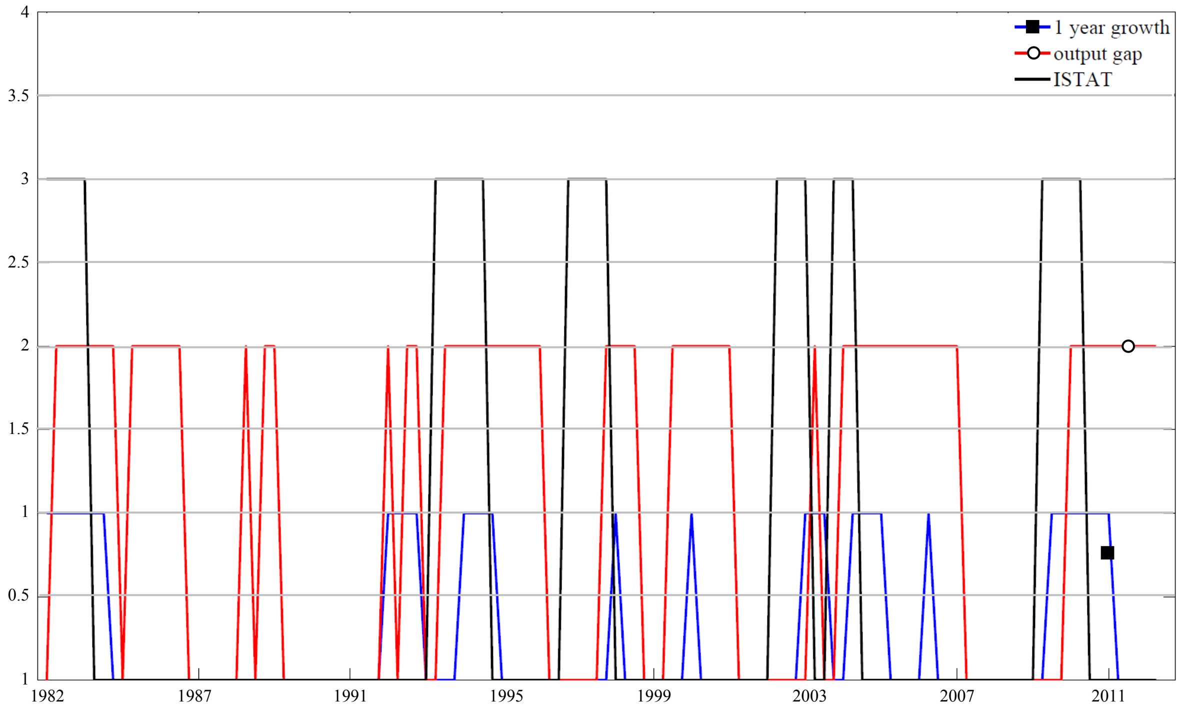
SVAR model and alternative models which exclude debt and/or foreign demand. Benchmark specification and alternative models which exclude debt and/or foreign demand.

Compared with a five-variable model that excludes both public debt and foreign demand, adding public debt determines a stronger (twice larger on average in the first two years) and longer lasting response of private GDP to a consumption shock (left panel). These results give support to the argument of Favero and Giavazzi (2007) that omitting debt in the model can result in biased estimates of the effects on GDP of fiscal shocks. The authors stressed the need to take into account the reactions of fiscal variables to changes in debt. In our case, these reactions would dampen the effects on output. On the contrary, we find a larger effect on private GDP, which comes from allowing a direct influence of debt on output.[13] Adding also the foreign demand (so as to reach our benchmark specification) does not instead have a sizeable effect on the response of private GDP.

[13] Another possible explanation for the greater response of private GDP could be that the inclusion of debt led to a better identification of the exogenous fiscal shocks (as the endogenous reactions of fiscal variables to changes in debt were excluded). However, we compared estimated fiscal shocks obtained with and without debt and differences were negligible.

Figure 5

Comparing Recession Dates According to the ISTAT Chronology and Alternative Models



Compared with a five-variable model that excludes both public debt and foreign demand, adding public debt determines a very large improvement in the precision of estimates: the confidence band of the response shrinks almost to a third, on average (right panel of Figure 4). This is not a surprise, given its major influence on Italian macroeconomic developments. Adding the debt also improves the accuracy of the estimates further, but to a lesser extent.

5 Distinguishing across states of the economy: the effects of government consumption shocks

Compared to the analysis presented in Section 4, here we estimate the effects of expenditure shocks distinguishing between states of the economy. There is no consensus in the literature on the most appropriate indicator for the business cycle. The official chronology of the Italian economy (Altissimo et al, 1999; Istat, 2011) which follows the methodology followed in the US by the National Bureau of Economic Research, identifies 28 quarters (out of 118) as "recessions" in the period 1982-2011. We label the other periods as "expansions". As a dichotomy variable cannot be used in the ETVAR and STVAR models, we also employ the following two alternative indicators to measure cyclical conditions: 4-quarters private GDP growth; the output gap, computed on the basis of the HP filter.

Figure 5 shows the recession periods based on the official chronology and the recession periods estimated by the ETVAR model using, respectively, the alternative indicators just mentioned. There are sizeable differences between the three estimates of "recessions". The ETVAR-based recessions are generally shifted forward with respect to the official chronology. Also, with ETVAR, the sample is more evenly split between the two regimes.

Nevertheless, Figure 5 allows to readily identify the four major recessions in our sample: the one at the beginning of the eighties, triggered by the second oil shock; the one at the beginning of the nineties, determined by the financial crisis; the strong slowdown in the initial years of the last decade and, finally, the last episode, influenced by the Lehman Bros' collapse.

5.1 Identifying recessions on the basis of the official chronology of the Italian economy

Figures 6 and 7 show the response of the fiscal and macroeconomic variables to a government consumption shock in the recession periods identified by the official chronology and in the other periods ("expansions"), respectively. As in the IRFs previously discussed, the shock is equal to 1 per cent of private GDP and the solid line represents the median response, while the dashed lines represent the 5th, 16th, 84th and 95th percentiles of the distribution of the responses. To facilitate comparison, in the Figure 6 we include also the median responses found in the expansion regime (solid line with squares).

When constructing impulse responses for a given regime, we assume that the state of the economy when the shock occurs does not change; in particular, we ignore any feedback effect from the fiscal shock to the type of regime. As this assumption becomes stronger the more we extend the time horizon of our analysis, we narrow it to 8 quarters in these IRFs and in the following ones.

The results are relatively close to those described in Section 4: in both regimes, the response of private GDP is positive and hump-shaped; revenues show a positive reaction making the initial surge in the debt to be gradually absorbed. However, under the recession regime, the response of private GDP is more prompt (the peak effect of 0.4 per cent is reached in the second quarter) and stronger on average in the first year. The response fades faster than in the expansion regime, but this is due to the fact that in the latter regime the expenditure shock is more persistent. Interest rates are higher in the expansions, but values are not significant in both regimes.

Figure 8 compares the response of private GDP and the cumulative multiplier in the two regimes. In recessions, the median value of the multiplier is constantly higher than in expansions; however, due to the very large confidence bands in the recession regime, the two bands largely overlap (note that the graph includes only one standard deviation bands), indicating that the difference between the two estimates is not statistically significant.

The very large size of the confidence bands estimated in the recession regime reflects the limited number of observations for this regime in the official chronology. To try to overcome this problem, we also run the SVAR adding to each recession the semester following it, as it is likely that substantial slack remains at the start of a recovery (in this way the number of observations for recessions increases to 40). This change improves the precision of estimates (in particular, the effects on private GDP in recession becomes statistically significant for 4 quarters) but it does not eliminate the overlap between the confidence bands of the estimates of the multipliers (Figure 9).

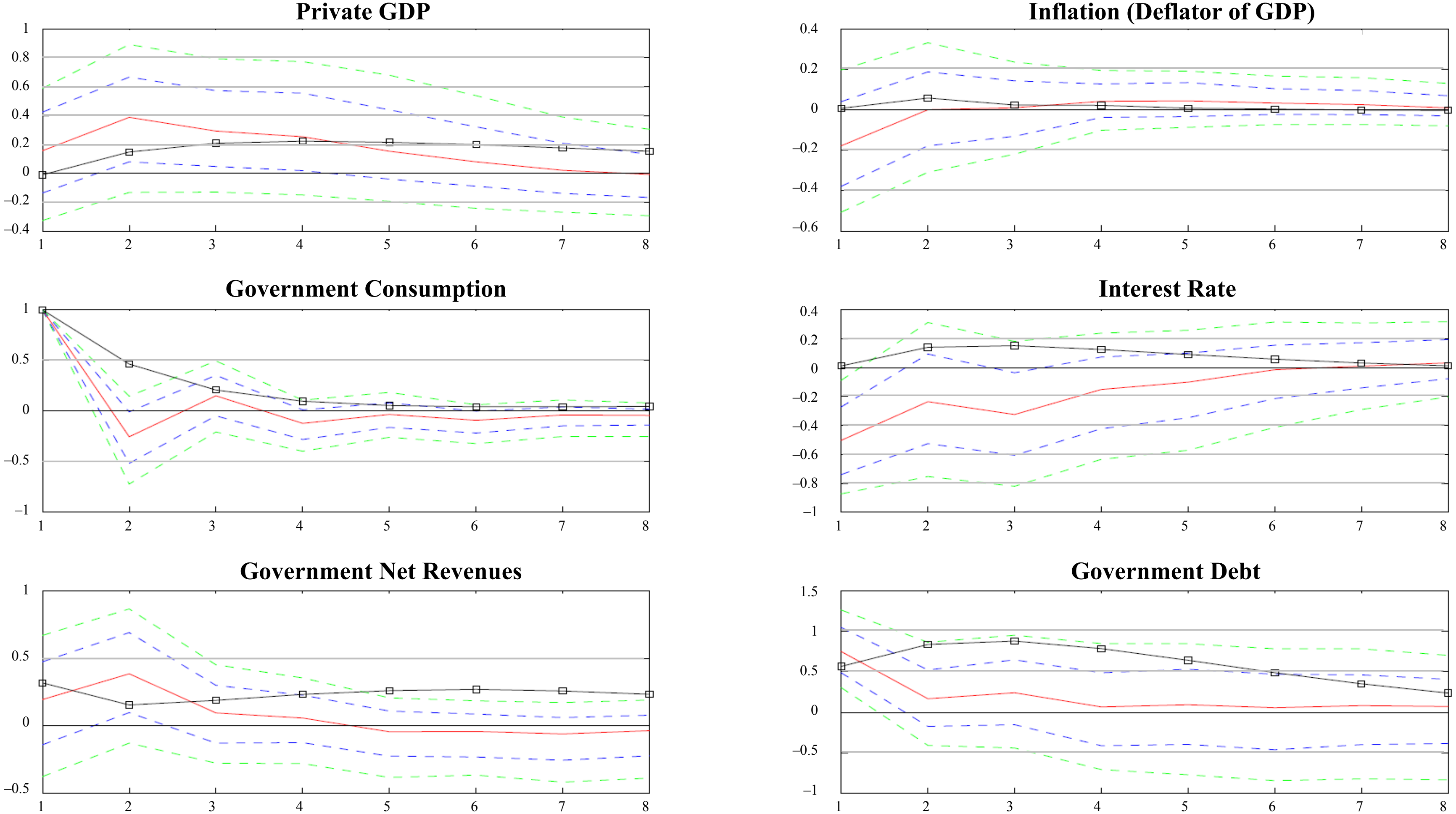
5.2 Identifying recessions on the basis of an ETVAR

While in the previous analysis recessions and expansions were identified outside the model, in this section the two regimes are endogenously identified in the ETVAR model on the basis of an indicator of cyclical conditions. We replicate the analysis on the basis of our two alternative indicators; we show in Appendix individual IRFs (Figures A1-A6); the expenditure multipliers are reported in Figure 10.

Results shown in Figure 10 broadly confirm the analysis based on the official chronology: under the recession regime, the response of private GDP is more prompt and stronger, determining

Figure 6

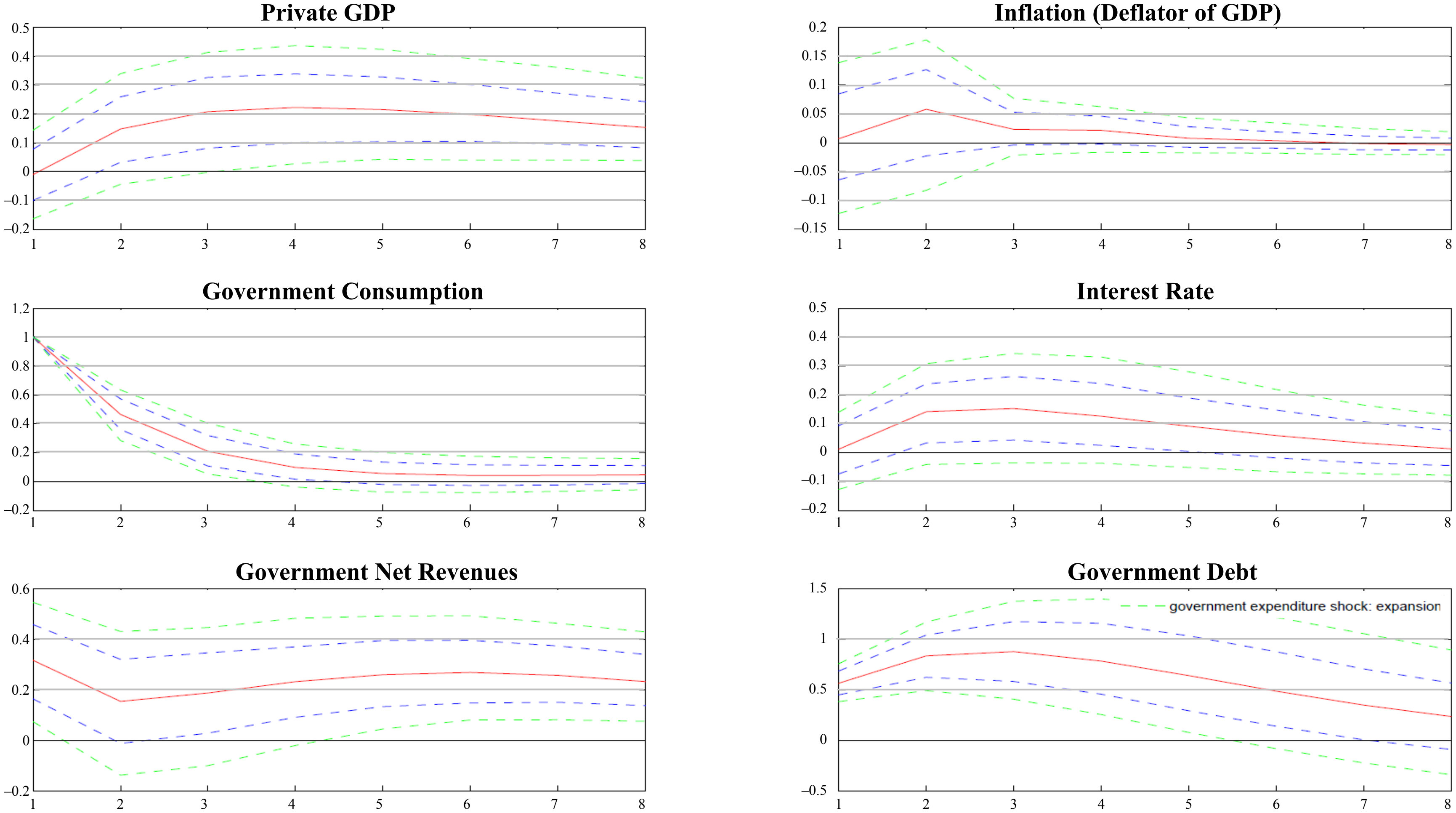
Impulse Responses in Recession to a Positive Government Consumption Shock Equal to 1 Percent of Private GDP



SVAR model using the ISTAT chronology. The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution. The solid curve with bullets represents the median response in expansion. Responses, except for inflation and interest rate, are deviations from the baseline and expressed in percentage points of private GDP. Inflation and interest rate responses are deviations from the baseline and expressed in percentage points.

Figure 7

Impulse Responses in Expansion to a Positive Government Consumption Shock Equal to 1 Percent of Private GDP



SVAR model using the ISTAT chronology. The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution. Responses, except for inflation and interest rate, are deviations from the baseline and expressed in percentage points of private GDP. Inflation and interest rate responses are deviations from the baseline and expressed in percentage points.

a higher fiscal multiplier. Compared to the analysis based on the official chronology, the estimates for the recession regime (which includes 32, 45 observations respectively for the two indicators) are more accurate, but there is still a large overlap between the confidence bands of the fiscal multipliers. The response of interest rates is broadly similar, and not significant, in the two regimes.

5.3 Identifying recessions on the basis of an STVAR

In the previous two sections, the sample was split into two regimes (recessions and expansions) and two sets of parameters were estimated. The STVAR model allows instead a smooth transition between the two regimes; in each quarter the parameters of the model are a linear combination of two sets of values (corresponding to each regime), weighted by the degree of being in each regime (Auerbach and Gorodnichenko, 2012).

In order to produce IRFs with this approach (and also to identify fiscal multipliers), we need to select two benchmarks, representative for recessions and expansions. We select the benchmarks so to leave outside them

Figure 8

Fiscal Multipliers in Recession (—o—) and Expansion (—■—) Using the ISTAT Chronology

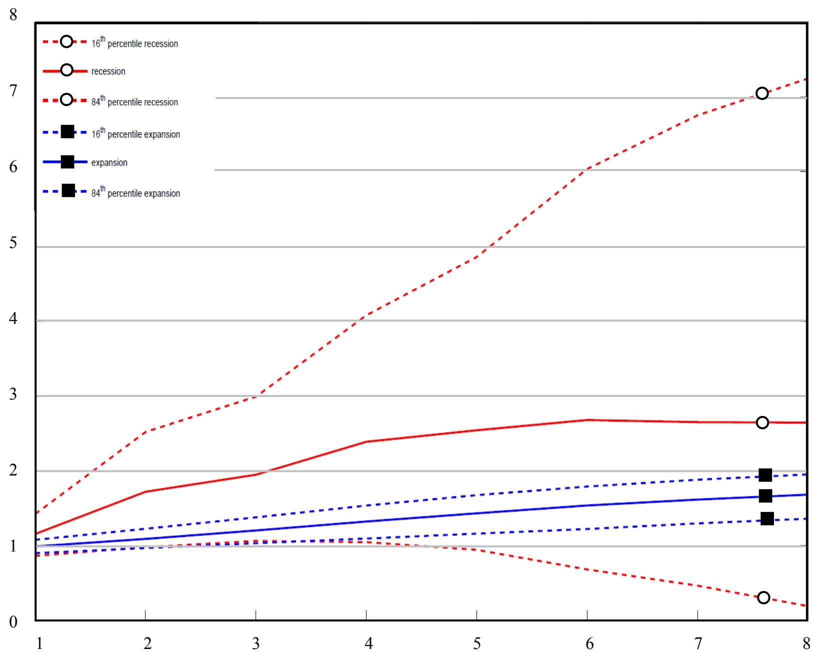


Figure 9

Fiscal Multipliers in Recession (—o—) and Expansion (—■—) Using the ISTAT Chronology Adding Two Quarters at the End of Each Recession Date

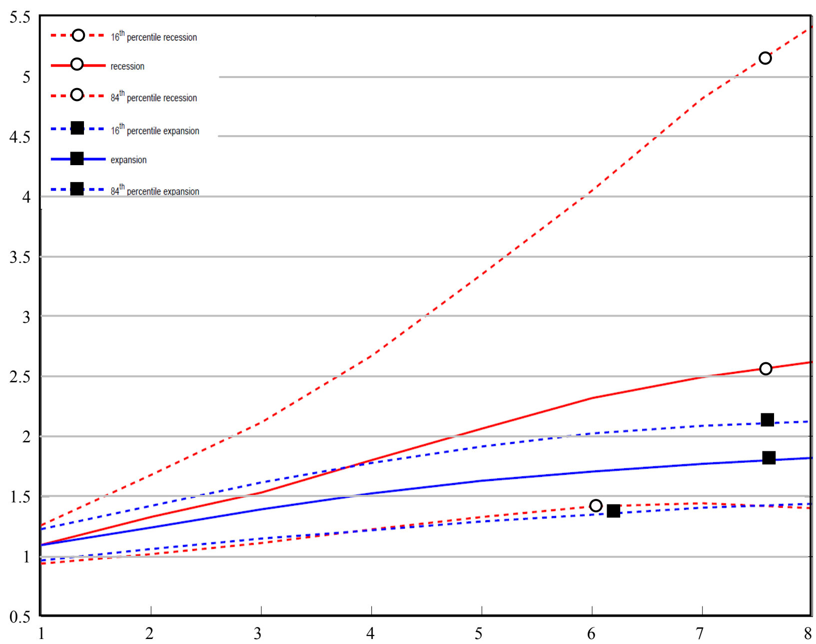
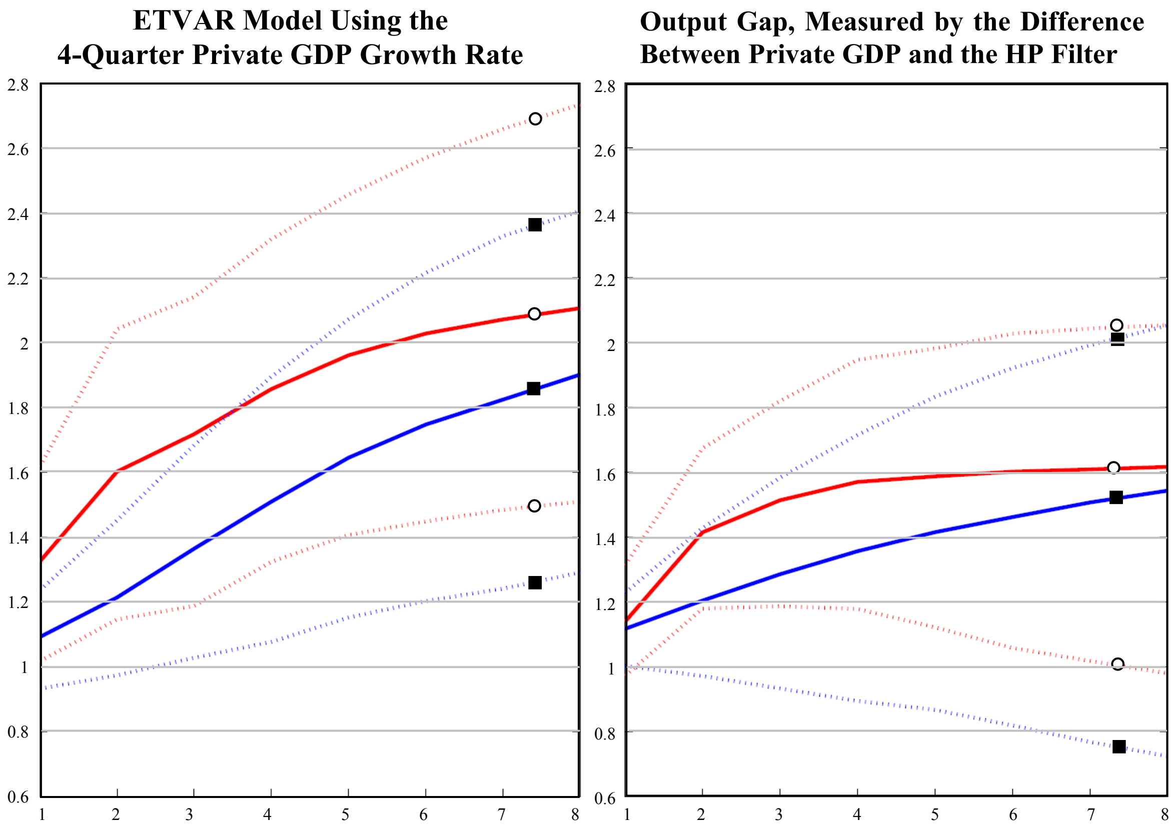


Figure 10

Fiscal Multipliers in Recession (—o—) and Expansion (—■—)



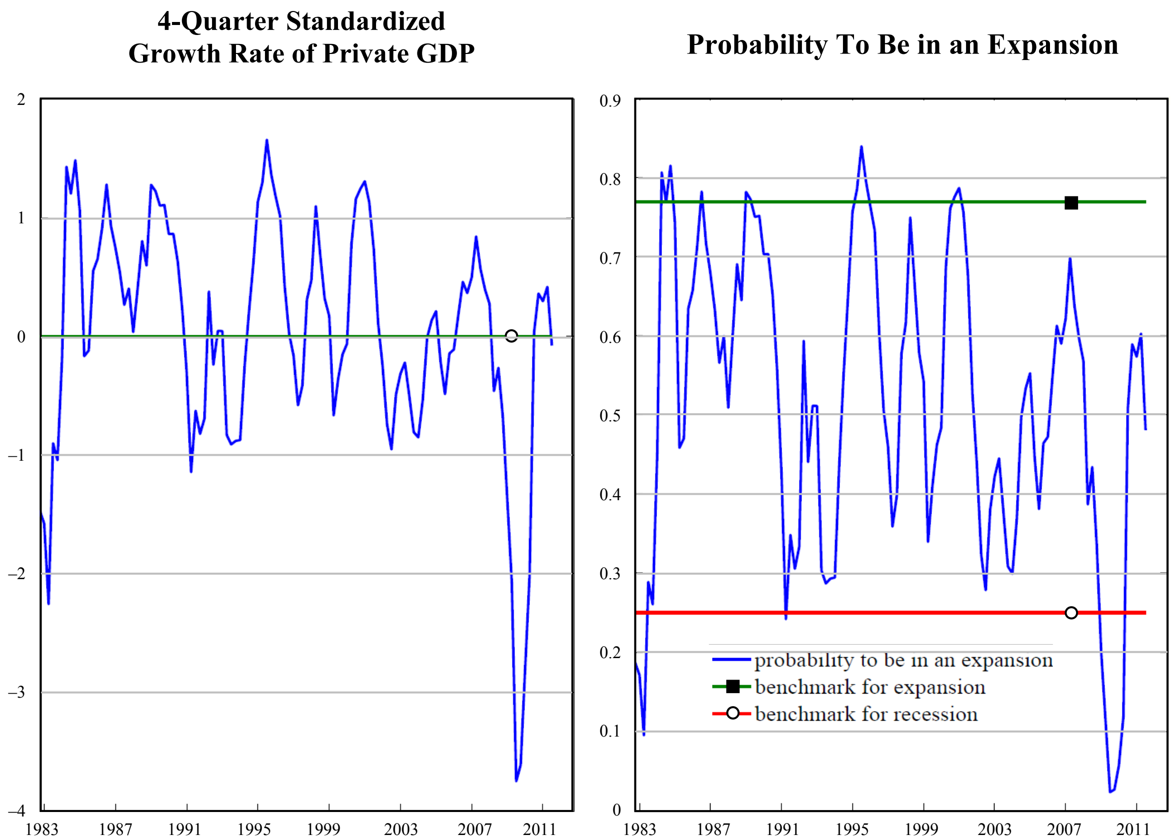
20 per cent of the observations. For the 4-quarter moving average of the private GDP annualized growth rate, these benchmarks corresponds to, respectively, –2.5 and 1.5. The left panel of Figure 11 shows the 4-quarter average of the private GDP annualized growth rate; the right panel shows the probability to be in an expansion (1–*Fz*; see Section 2) and the selected benchmarks.

We report in Appendix the IRFs, based on the STVAR model for the two indicators of cyclical conditions (Figures A7-A12). The estimates of the expenditure multipliers are reported in Figure 12. Results are far less clear-cut than in the previous analysis. Depending on the cyclical indicator used, in recessions the expenditure multiplier is constantly higher (output gap or constantly lower (4-quarters private GDP growth). Estimates are generally less precise than those obtained with the ETVAR model, as shown by the larger confidence bands compared to those in Figure 10.

8 Conclusions and future research

In this paper we study how the effects of expenditure shock on economic activity are influenced by the state of the economy on the basis of various autoregressive models and indicators of cyclical conditions. We rely on quarterly cash-basis fiscal data for the Italian economy covering the period 1982:1-2011:2.

Figure 11



The main results can be summarized as follows.

Independently of the method that we use and whether we distinguish or not between states of the economy, the expenditure shocks that we estimate tend to be largely transitory and revenues show a positive reaction, making the initial surge in the debt to be gradually absorbed.

In the analysis which doesn't take into account the state of the economy, the response of private GDP to an expenditure shock is positive and highly significant for approximately two years. The government consumption multiplier (1.04 on impact and 1.8 at the peak) lies in the upper part of the wide range of estimates provided by the empirical literature.[14]

When we split the sample in "expansion" and "recession" regimes, either on the basis of the official chronology or applying the ETVAR approach to our two cyclical indicators, for both regimes we obtain IRFs broadly similar to those estimated for the full sample but generally less precise.

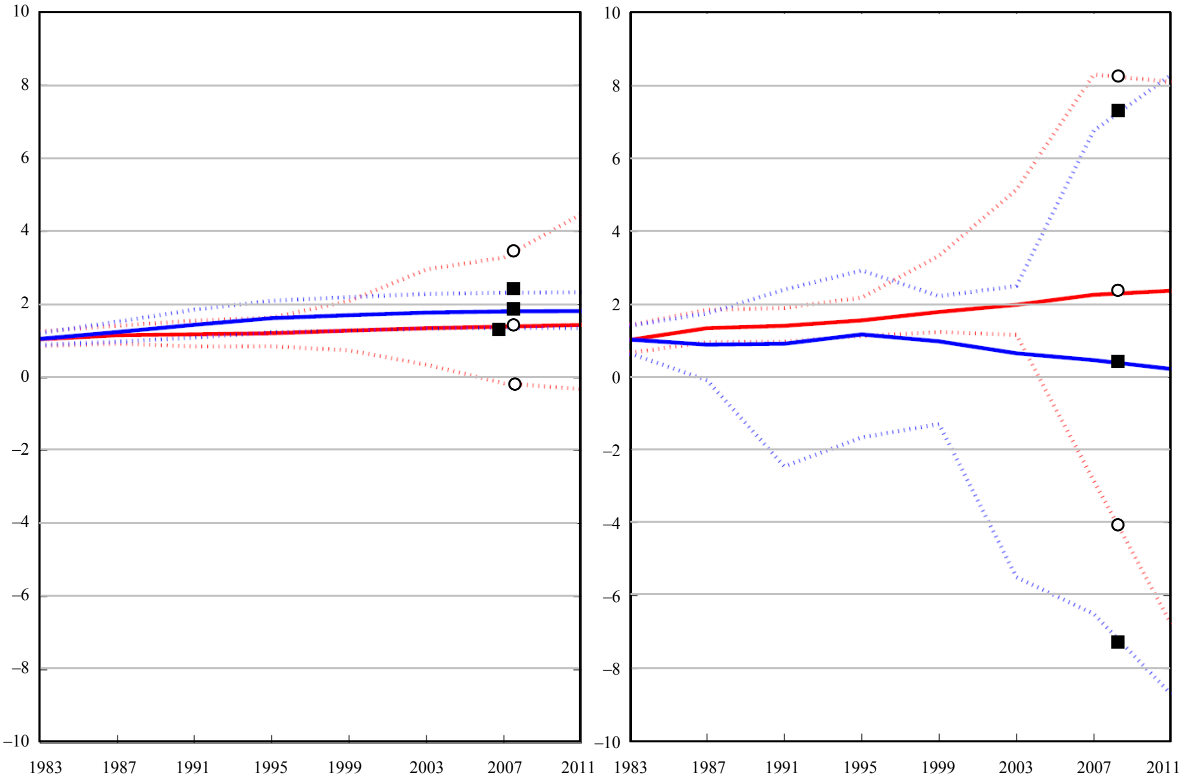
Under the recession regime, the response of private GDP is larger and more prompt; the cumulative multiplier is also stronger. However, the confidence bands of the estimates are relatively large, and the difference in the median across regimes is not statistically significant.

[14] This may be due to the debt-stabilizing reaction of revenues, in line with the idea that the effects of fiscal stimulus on economic activity depend positively on the soundness of fiscal policy (see, e.g., Corsetti *et al.*, 2009). The transitory nature of the shocks that we observe and their small size may also have a bearing on the value of the multiplier.

Figure 12

Fiscal Multipliers in Recession (—○—) and Expansion (—■—)

STVAR Model Using the 4-Quarters Private GDP Growth Rate | **Output Gap Measured by the Difference Between Private GDP and the HP Filter**



Median results are ambiguous when we apply the STVAR model: depending on the cyclical indicator used, in recession the median value of the expenditure multiplier is constantly higher or constantly lower. This result is associated to estimates that are less precise than those obtained with the ETVAR approach. This is somewhat unexpected, as this method is deemed to use more efficiently the information of the sample; a possible explanation is that weighting function $F(z_{t-1})$, in the presence of highly imperfect indicators of cyclical conditions, represents a sort of unwarranted straightjacket for the data.

Results give a very weak support to the idea that higher multipliers may be due, at least partly, to a different behaviour of interest rates.

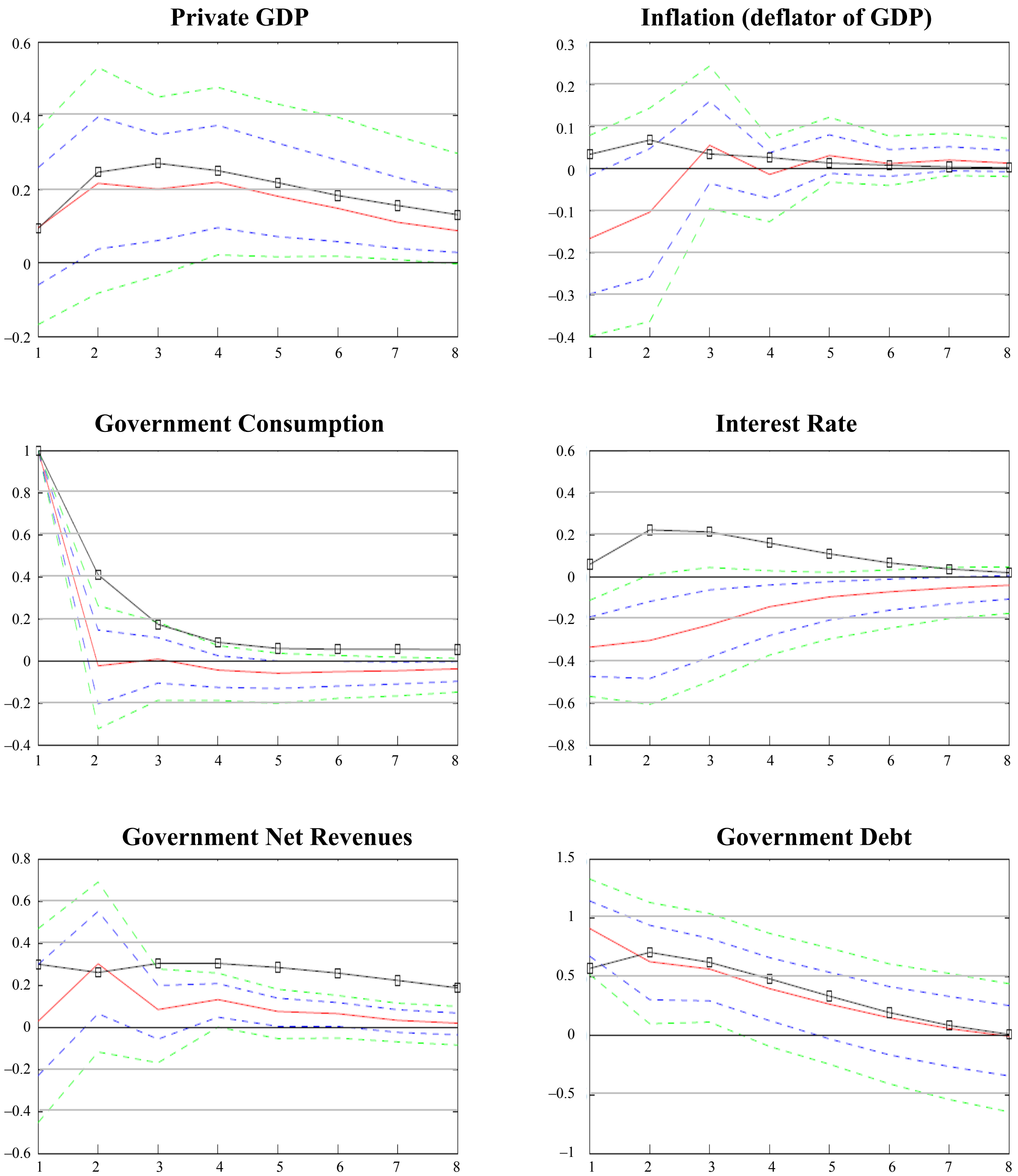
In conclusion, our empirical investigation shows some weak evidence that expenditure multipliers tend to be higher in recessions than in expansions. While this result is influenced by the limited size of the sub-samples, it seems to suggest that the differences in the multipliers have not been extremely large, at least in the period under examination.

Finally, our empirical analysis could be strengthened along at least two lines. First, the assumption that the initial state of the economy does not change when constructing impulse responses for a given regime should be relaxed. Second, a more thorough selection and discussion of business cycle indicators should be conducted.

APPENDIX

Figure 13

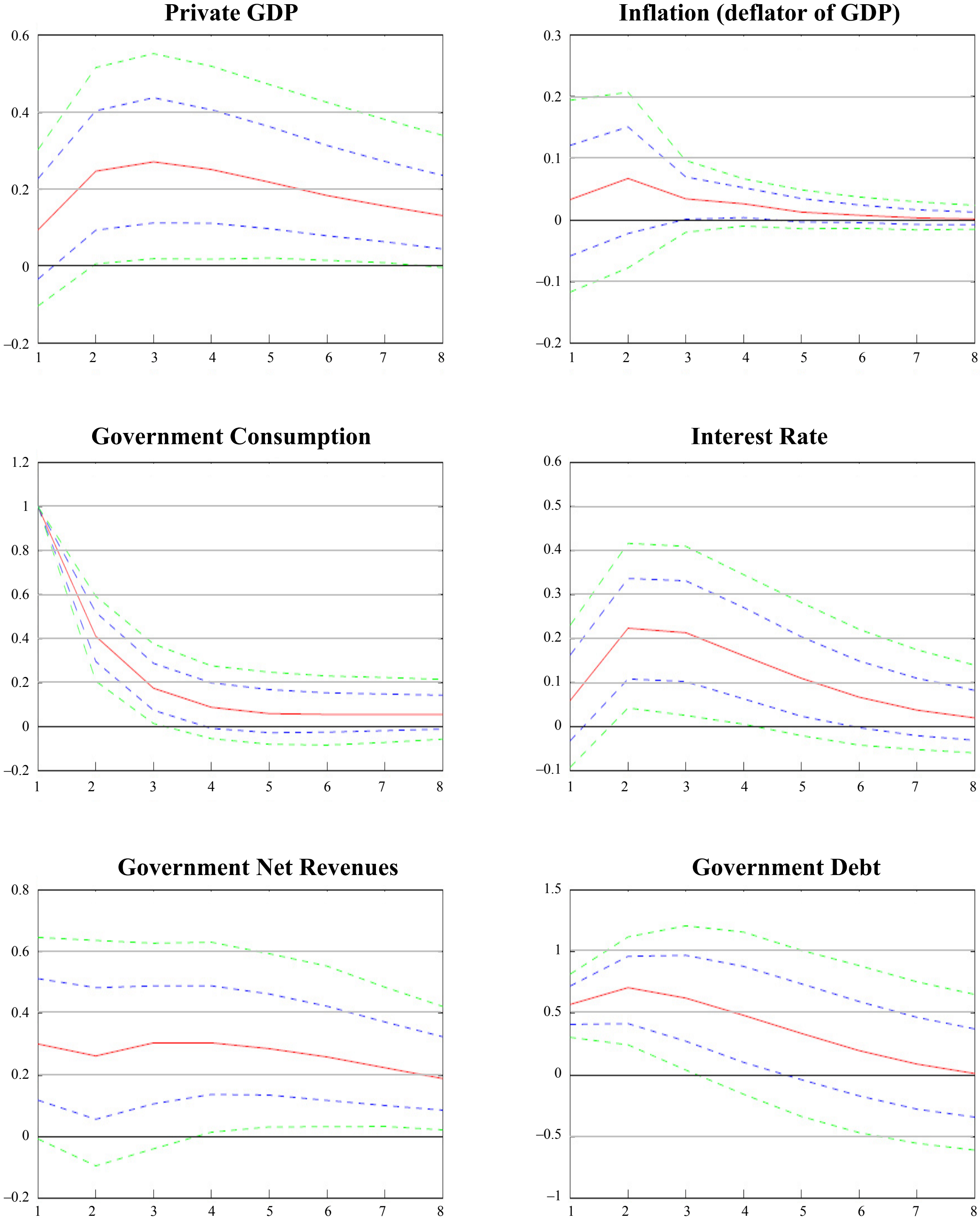
Impulse Responses in Recession to a Positive Government Consumption Shock Equal to 1 Percent of Private GDP



SVAR model using the ISTAT chronology adding two quarters at the end of each recession date. The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution. The solid curve with bullets represents the median response in expansion. Responses, except for inflation and interest rate, are deviations from the baseline and expressed in percentage points of private GDP. Inflation and interest rate responses are deviations from the baseline and expressed in percentage points.

Figure 14

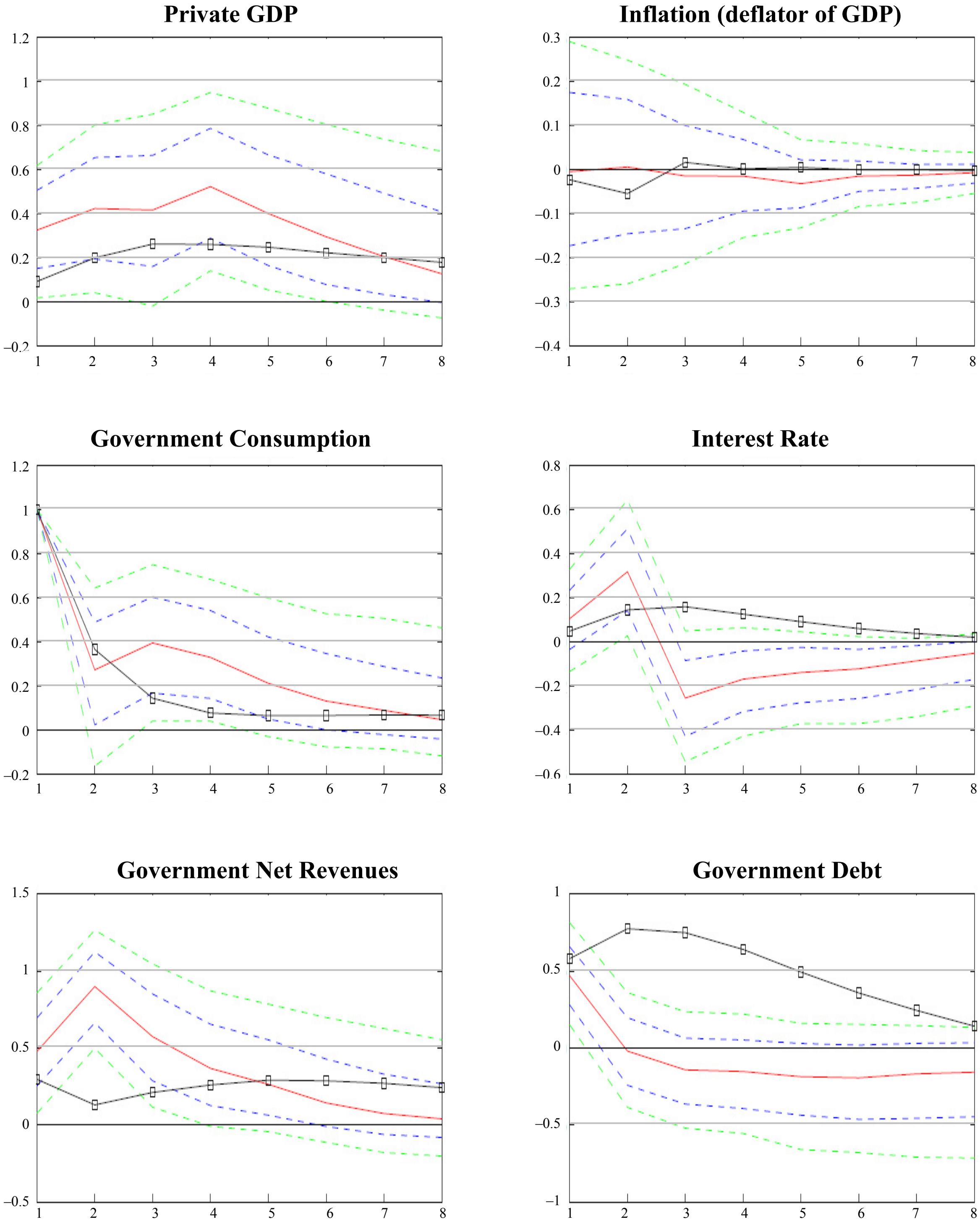
Impulse Responses in Expansion to a Positive Government Consumption Shock Equal to 1 Percent of Private GDP



SVAR model using the ISTAT chronology adding two quarters at the end of each recession date. The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution. Responses, except for inflation and interest rate, are deviations from the baseline and expressed in percentage points of private GDP. Inflation and interest rate responses are deviations from the baseline and expressed in percentage points.

Figure 15

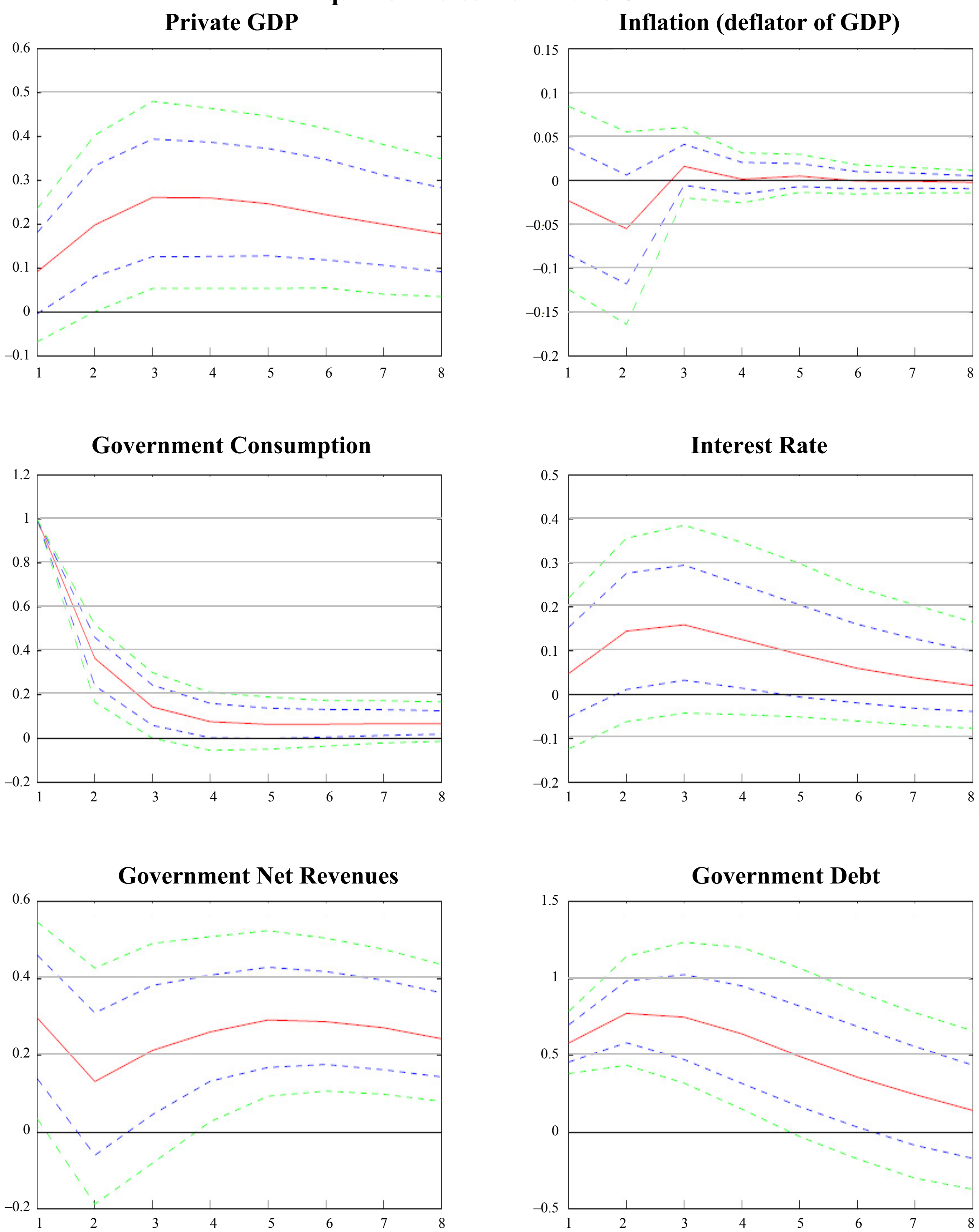
Impulse Responses in Recession to a Positive Government Consumption Shock Equal to 1 Percent of Private GDP



ETVAR model with 4-quarters private GDP growth rate. The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution. The solid curve with bullets represents the median response in expansion. Responses, except for inflation and interest rate, are deviations from the baseline and expressed in percentage points of private GDP. Inflation and interest rate responses are deviations from the baseline and expressed in percentage points.

Figure 16

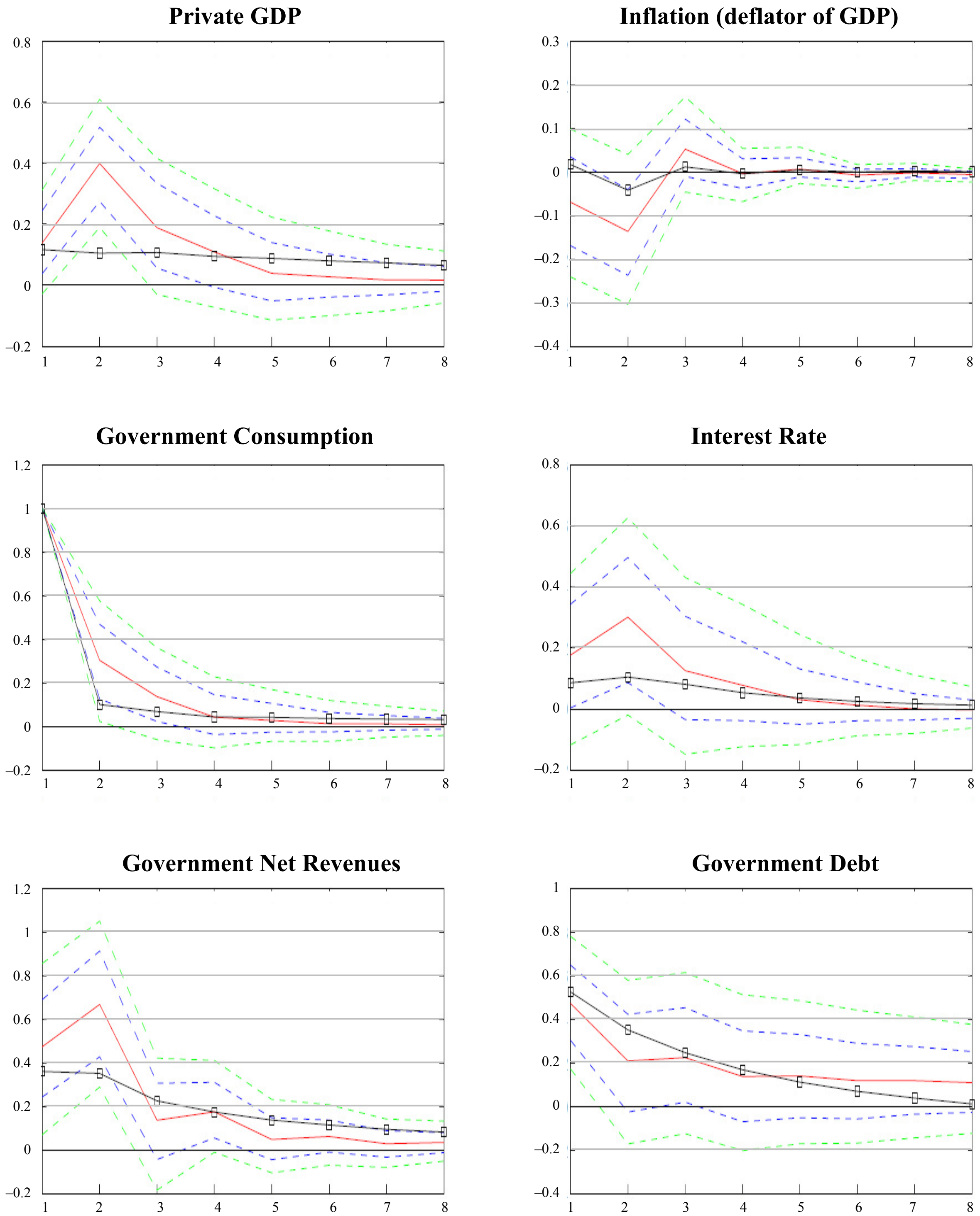
Impulse Responses in Expansion to a Positive Government Consumption Shock Equal to 1 Percent of Private GDP



ETVAR model with 4-quarters private GDP growth rate. The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution. Responses, except for inflation and interest rate, are deviations from the baseline and expressed in percentage points of private GDP. Inflation and interest rate responses are deviations from the baseline and expressed in percentage points.

Figure 17

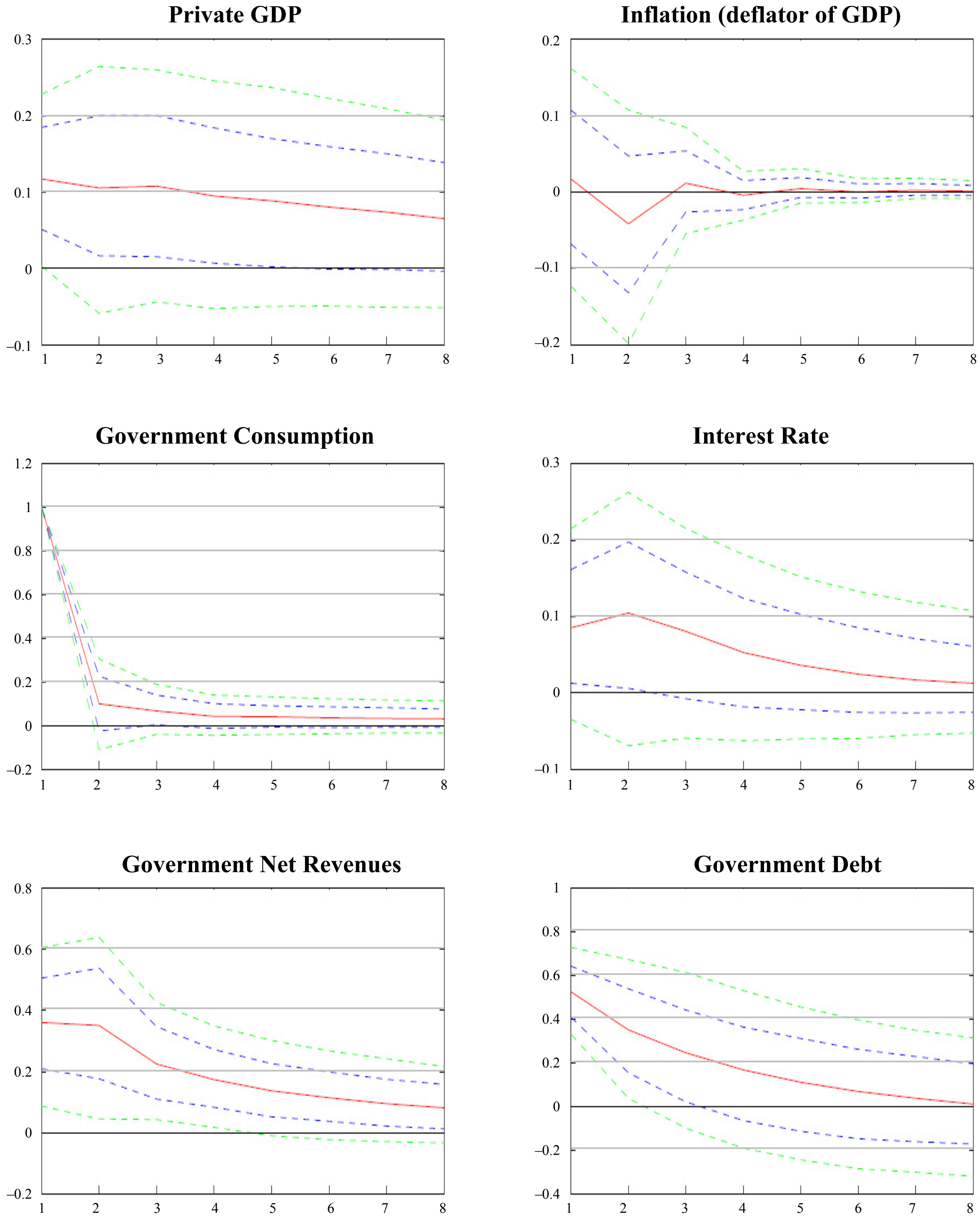
Impulse Responses in Recession to a Positive Government Consumption Shock Equal to 1 Percent of Private GDP



ETVAR model with output gap measured by the difference between private GDP and the HP filter. The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution. The solid curve with bullets represents the median response in expansion. Responses, except for inflation and interest rate, are deviations from the baseline and expressed in percentage points of private GDP. Inflation and interest rate responses are deviations from the baseline and expressed in percentage points.

Figure 18

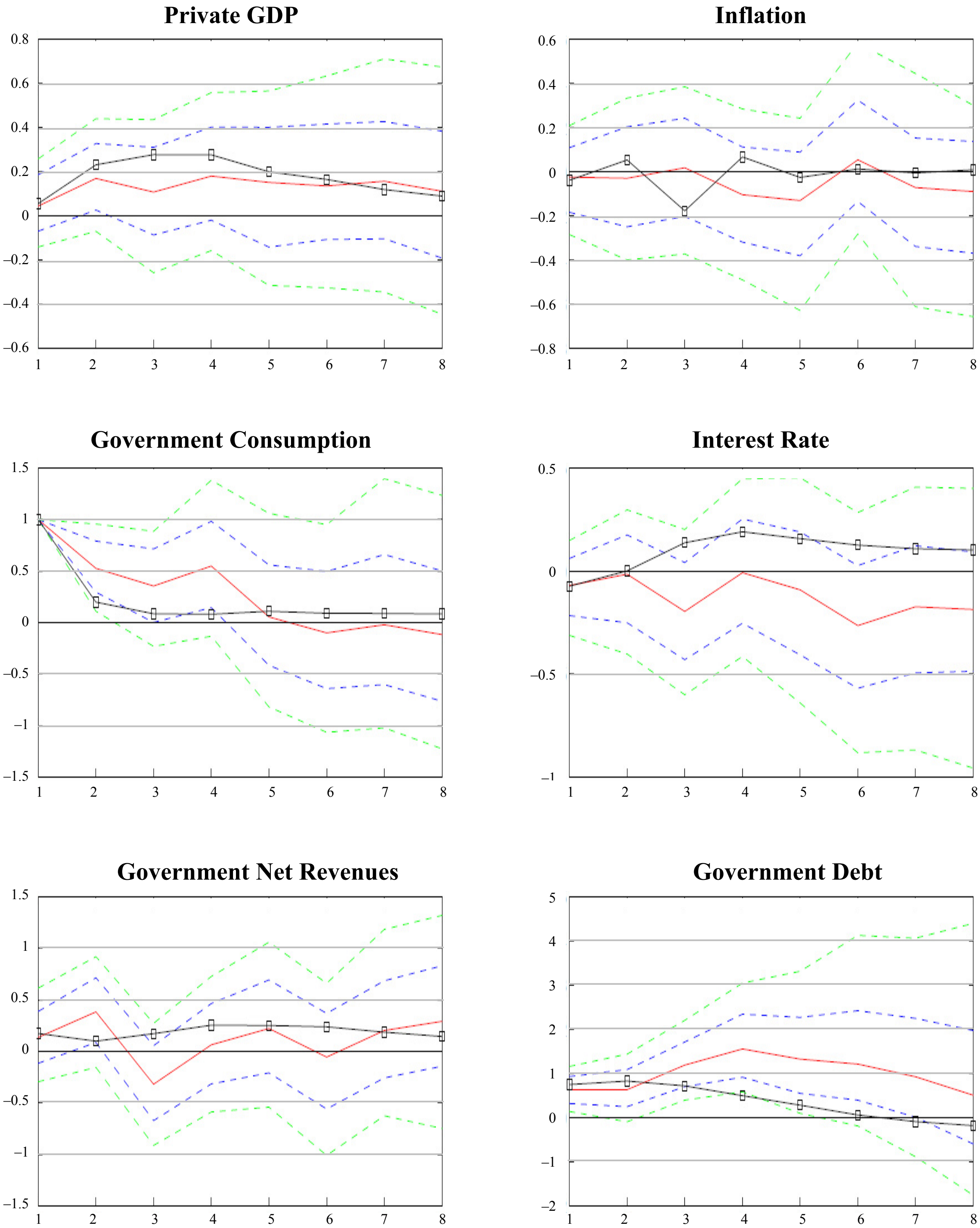
Impulse Responses in Expansion to a Positive Government Consumption Shock Equal to 1 Percent of Private GDP



ETVAR model with output gap measured by the difference between private GDP and the HP filter. The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution. Responses, except for inflation and interest rate, are deviations from the baseline and expressed in percentage points of private GDP. Inflation and interest rate responses are deviations from the baseline and expressed in percentage points.

Figure 19

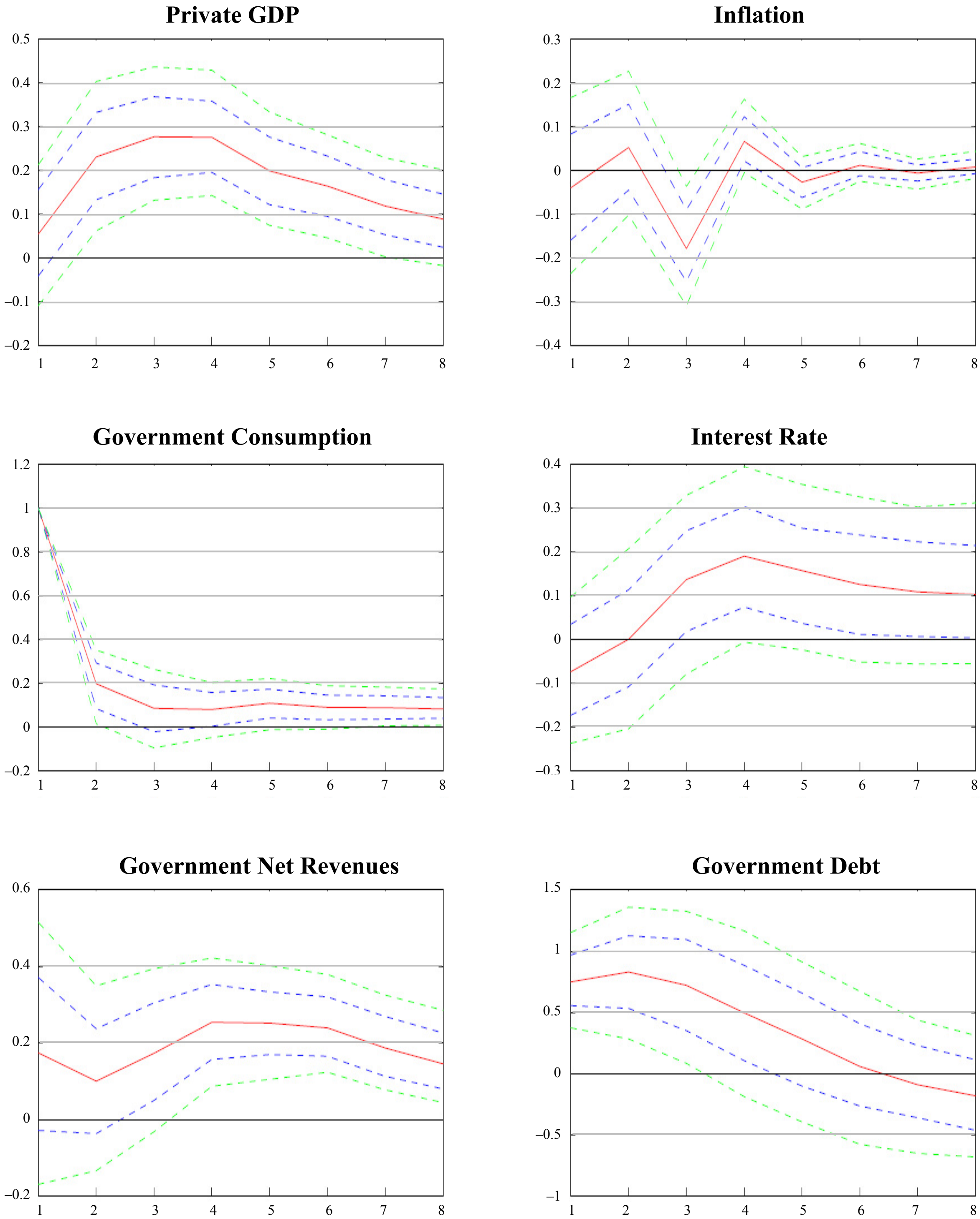
Impulse Responses in Recession to a Positive Government Consumption Shock Equal to 1 Percent of Private GDP



STVAR model with the 4 quarters private GDP growth rate. The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution. The solid curve with bullets represents the median response in expansion. Responses, except for inflation and interest rate, are deviations from the baseline and expressed in percentage points of private GDP. Inflation and interest rate responses are deviations from the baseline and expressed in percentage points.

Figure 20

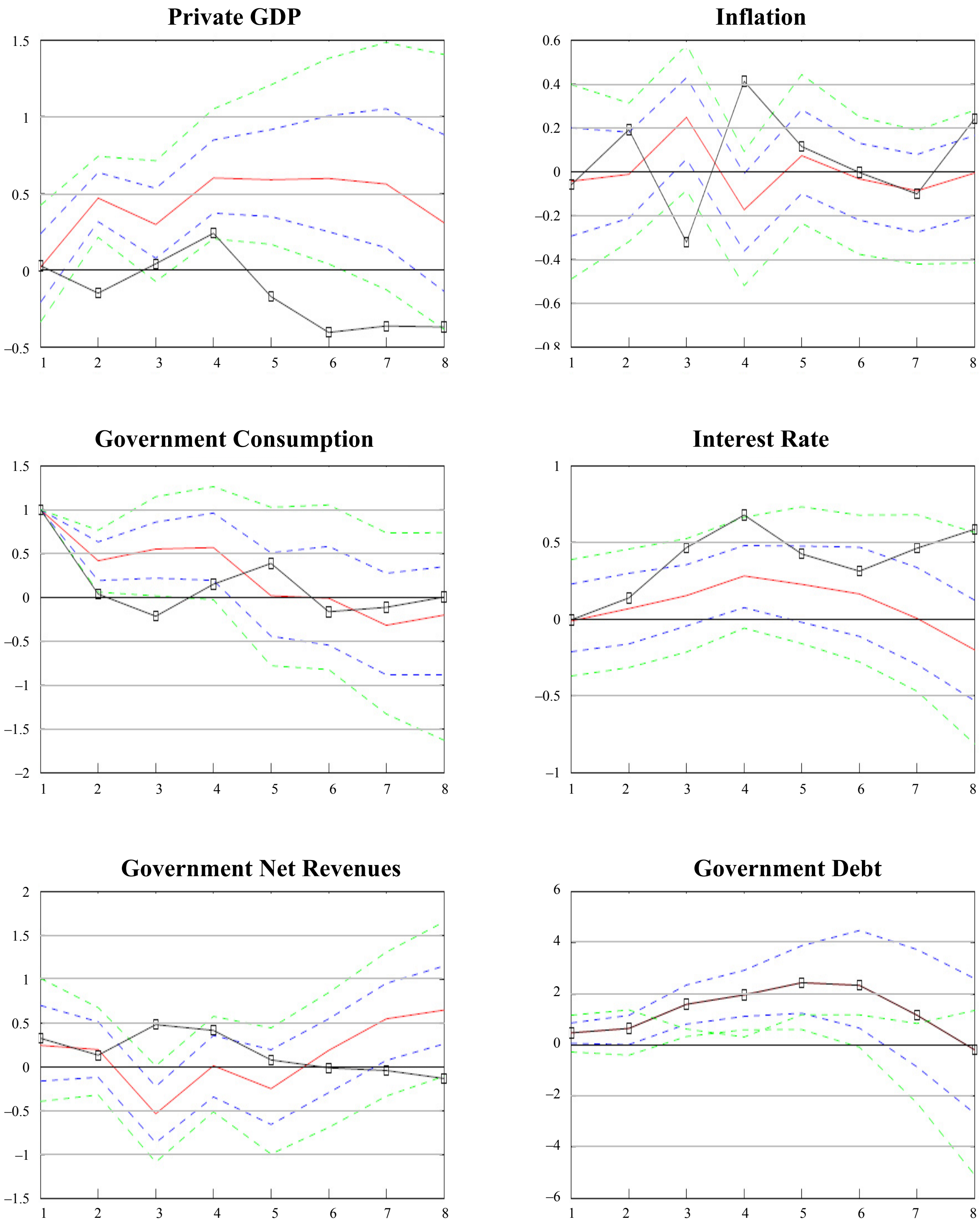
Impulse Responses in Expansion to a Positive Government Consumption Shock Equal to 1 Percent of Private GDP



STVAR model with the 4 quarters private GDP growth rate. The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution. Responses, except for inflation and interest rate, are deviations from the baseline and expressed in percentage points of private GDP. Inflation and interest rate responses are deviations from the baseline and expressed in percentage points.

Figure 21

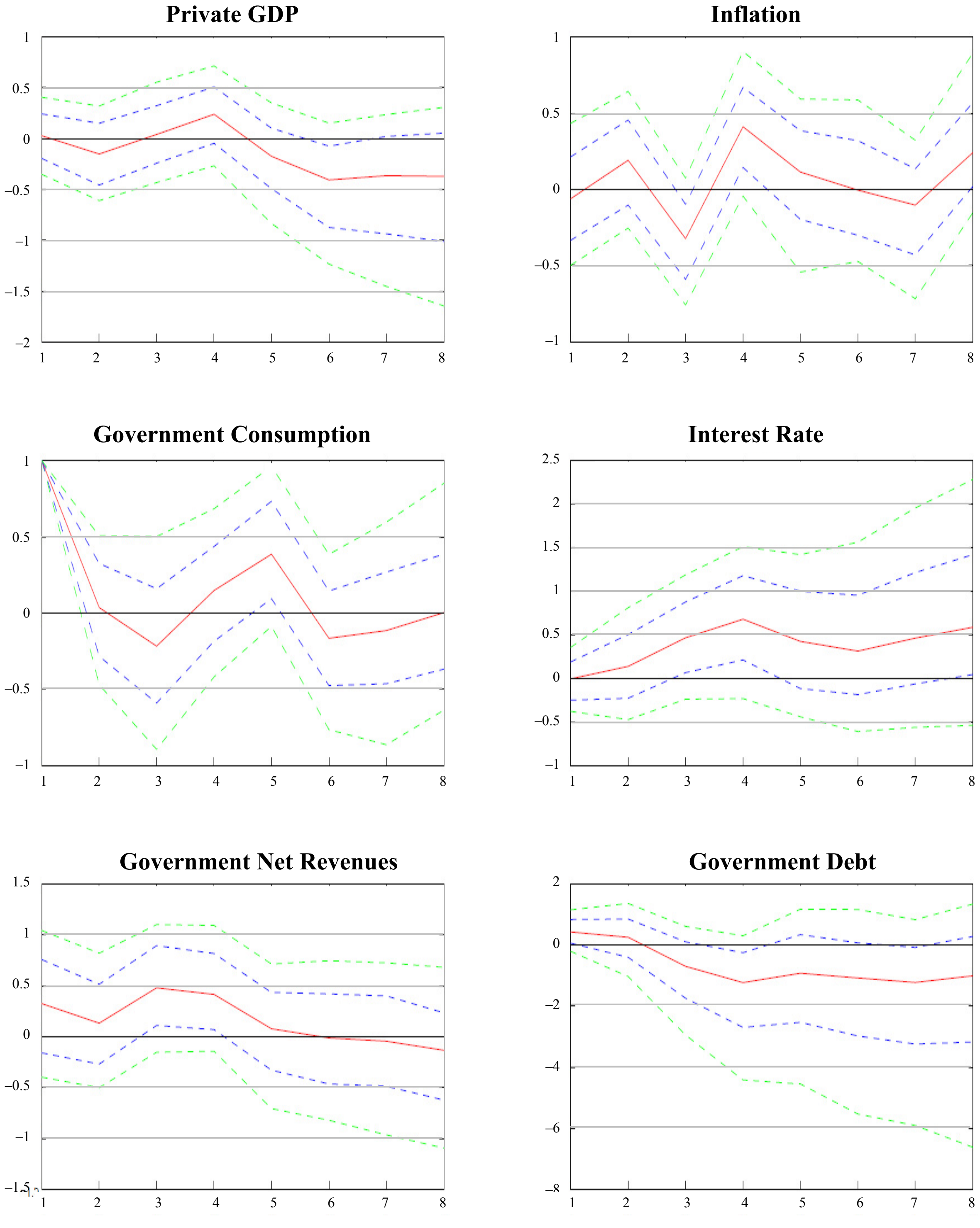
Impulse Responses in Recession to a Positive Government Consumption Shock Equal to 1 Percent of Private GDP



STVAR model with the output gap measured by the difference between private GDP and the HP filter. The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution. The solid curve with bullets represents the median response in expansion. Responses, except for inflation and interest rate, are deviations from the baseline and expressed in percentage points of private GDP. Inflation and interest rate responses are deviations from the baseline and expressed in percentage points.

Figure 22

Impulse Responses in Expansion to a Positive Government Consumption Shock Equal to 1 Percent of Private GDP



STVAR model with the output gap measured by the difference between private GDP and the HP filter. The curves represent the median and two sets of lower and upper bands, corresponding to the 5th, 16th, 84th and 95th percentiles of the distribution. Responses, except for inflation and interest rate, are deviations from the baseline and expressed in percentage points of private GDP. Inflation and interest rate responses are deviations from the baseline and expressed in percentage points.

REFERENCES

Altissimo, F., D.J. Marchetti and G.P. Oneto (2000), "The Italian Business Cycle: Coincident and Leading Indicators and Some Stylized Facts", Temi di Discussione, No. 377, Banca d'Italia.

Auerbach, A. and Y. Gorodnichenko (2012), "Measuring the Output Responses to Fiscal Policy", *American Economic Journal: Economic Policy*, forthcoming.

Baum, A. and G. Koester (2011), "The Impact of Fiscal Policy on Economic Activity Over the Business Cycle – Evidence from a Threshold VAR Analysis", Deutsche Bundesbank, Discussion Paper, No. 03/2011.

Baum, A., M. Poplawski-Ribeiro and A. Weber (2012), "Fiscal Multipliers and the State of the Economy", IMF, Working Paper, forthcoming.

Bernheim, B.D. (1987), "Ricardian Equivalence: An Evaluation of Theory and Evidence", in *NBER Macroeconomic Annual*, MIT Press, Cambridge (MA), pp. 263-304.

Blanchard, O. and R. Perotti (2002), "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output", *Quarterly Journal of Economics*, No. 117.

Bohn, H. (2007), "Are Stationarity and Cointegration Restrictions Really Necessary for the Intertemporal Budget Constraint?", *Journal of Monetary Economics*, No. 54, pp. 1837-47.

Burriel, P., F. de Castro, D. Garrote, E. Gordo, J. Paredes and J. Pérez (2010), "Fiscal Policy Shocks in the Euro Area and the US: An Empirical Assessment", ECB, Working Paper, No. 1133.

Busetti, F., A. Felettigh and R. Tedeschi (2011), "Esportazioni italiane e domanda potenziale all'interno e all'esterno dell'area dell'euro: Stime econometriche disaggregate e uno scenario Controfattuale", Bank of Italy, mimeo.

Caprioli, F. and S. Momigliano (2011), "The Effects of Fiscal Shocks with Debt-stabilizing Budgetary Policies in Italy", Temi di Discussione, No. 839, Banca d'Italia.

Chahrour, C., S. Schmitt-Grohe and M. Uribe (2010), "A Model-based Evaluation of the Debate on the Size of Tax Multiplier", NBER, Working Paper, No. 16169.

Christiano, L., M. Eichenbaum and S. Rebelo (2009), "When is the Government Spending Multiplier Large?", Northwestern University, manuscript.

Chung, H. and E.M. Leeper (2007), "What Has Financed Government Debt", University of Indiana, manuscript.

Coenen, G., C. Erceg, C. Freedman, D. Furceri, M. Kumhof, R. Lalonde, D. Laxton, J. Lindé, A. Mourougane, D. Muir, S. Murusula, C. de Resende, J. Roberts, W. Roeger, S. Snudden, M. Trabandt and J. in't Veld, (2010), "Effects of Fiscal Stimulus in Structural Models", IMF, Working Paper No. 10/73.

Cogan, J., T. Cwik, V. Wieland and J. Taylor (2009), "New Keynesian Versus Old Keynesian Government Spending Multipliers", ECB, Working Paper, No. 1090.

Corsetti, G., A. Meier and G. Mueller (2009), "Fiscal Stimulus with Spending Reversal", IMF, Working Paper, No. 09106.

Corsetti, G., K. Meier and G. Mueller (2010), "Debt Consolidation and Fiscal Stabilization of Deep Recessions", *American Economic Review*, No. 100, pp. 41-5.

Creel, J., P. Monperrus-Veroni and F. Saraceno (2005), "Discretionary Policy Interactions and the Fiscal Theory of the Price Level: A SVAR Analysis on French Data", OFCE, Document de travail, No. 12.

Edelberg, W., M. Eichenbaum and J. Fisher (1999), "Understanding the Effects of a Shock to Government Purchases", *Review of Economic Dynamics*, Issue I, pp. 166-206, January.

Engen, E. and R.G. Hubbard (2004), "Federal Government Debts and Interest Rates", NBER, Working Paper, No. 1068, pp. 1419-31.

Fatás, A. and I. Mihov (2001), "The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence", CEPR, Discussion Paper, No. 2760.

Fazzari, S.M., J. Morley and I. Panovska (2012), "State Dependent Effects of Fiscal Policy", mimeo.

Favero, C. and F. Giavazzi (2007), "Debt and the Effect of Fiscal Policy", NBER, Working Paper, No. 12822.

Gale, W.G. and P.Orzag (2002), "The Economic Effects of Long-term Fiscal Discipline", Urban Institute-Brookings Institution Tax Policy Center, Discussion Paper, December.

Giordano, R., S. Momigliano, S. Neri and R. Perotti (2008), "The Effects of Fiscal Policy in Italy: Evidence from a VAR Model", *European Journal of Political Economy*, No. 23, pp. 707-33.

Golinelli, R. and S. Momigliano (2008), "The Cyclical Response of Fiscal Policies in the Euro Area. Why Do Results of Empirical Research Differ So Strongly?", Temi di Discussione, No. 654, Bank of Italy.

Hamilton, J. and M. Flavin (1986), "On the Limitations of Government Borrowing: A Framework for Empirical Testing", *American Economic Review*, No. 76, pp. 808-19.

Henry, J., P. Hernández de Cos and S. Momigliano (2004), "The Short-term Impact of Government Budgets on Prices: Evidence from Macroeconometric Models", ECB, Working Paper, No. 396.

Ilzetzki, E., E. Mendoza and C. Vegh (2009), "How Big are Fiscal Multipliers?", CEPR, Policy Insight, No. 39.

IMF (2012), "World Economic Outlook".

ISTAT (2010), "Rapporto annuale: La situazione del paese nel 2010".

King, R. and C. Plosser (1985), "Money, deficit and inflation", Carnagie Rochester Conference Series on Public Policy, No. 22, pp. 147-96.

Kirchner, M., J. Cimadomo and S. Hauptmeier (2010), "Transmission of Government Spending Shocks in the Euro Area: Time Variation and Driving Forces", mimeo.

Marcellino, M. (2006), "Some Stylized Facts on Non-systematic Fiscal Policy in the Euro Area", *Journal of Macroeconomics*, No. 28, pp. 461-79.

Miller, S.M. and F.S. Russek (1996), "Do Federal Deficits Affect Interest Rates? Evidence from Three Econometric Methods", *Journal of Macroeconomics*, No. 18, pp. 403-28.

Mountford, A. and H. Uhlig (2002), "What Are the Effects of Fiscal Policy Shocks?", CEPR, Discussion Paper, No. 3338.

Neri, S. (2001), "Assessing the Effects of Monetary and Fiscal Policy", Temi di discussione, No. 425, Bank of Italy.

Parker, J. (2011), "On Measuring the Effects of Fiscal Policy in Recessions", Northwestern University, manuscript.

Perotti, R. (2004), "Estimating the Effects of Fiscal Policy in OECD Countries", IGIER, mimeo.

Ramey, V. (2008), "Identifying Government Spending Shocks: It's All in the Timing", NBER, Working Paper, No. 15464.

Ramey, V. and M. Shapiro (1997), "Costly Capital Reallocation and the Effects of Government spenDing", NBER, Working Paper, No. 6283.

Romer, D. and C. Romer (2010), "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks", *American Economic Review*, No. 100, pp. 763-801.

Sims, C., J. Stock and M. Watson (1990), "Inference in Linear Time Series Models with Unit Roots", Econometrica, No. 58.

Spilimbergo, A., A. Symansky and M. Schindler (2009), "Fiscal Multipliers", IMF, Staff Position Note, May 09/11.

Trehan, B. and C. Walsh (1991), "Testing Intertemporal Budget Constraints: Theory and Applications to U.S. Federal Budget and Current Account Deficits", *Journal of Money, Credit and Banking*, No. 23, pp. 210-23.

FISCAL CONSOLIDATION IN THE MIDST OF THE CRISIS

Francesco Di Comite,[*] *Gabriele Giudice,*[*] *Julia Lendvai*[*] *and Ingrid Toming*[*]

1 Introduction

We analyse the key aspects of the dramatic fiscal consolidation in Latvia in 2008-11 and the linkages between fiscal policy and short-term economic growth in a small open economy. Amidst financial turmoil and the unwinding of extreme imbalances, the Latvian economy contracted by almost 25 per cent from peak to trough. As the government undertook a massive consolidation (of over 15 per cent of GDP, kicking in mainly in July 2009), the economy bounced back more rapidly than anyone expected. After mid-2010, contraction yielded to sustained growth, while the ambitious fiscal targets under the EU/IMF Balance-of-Payments programme were consistently over-achieved. This experience of large-scale consolidation during a major economic correction provides valuable insights into the mechanism of fiscal adjustment.

Before identifying preconditions and contributing factors to such an outcome, we first need to correctly measure the changes in public finances which took place over the period. However, as a significant part of the adjustment is missed by standard measures of fiscal effort, a bottom-up approach is also needed. A review of quarterly GDP and budgetary results helps understanding the effective sequencing of fiscal adjustment and economic activity. We then review the composition of the consolidation and compare its implementation with the literature on the optimal mix of measures. Subsequently, using the European Commission's QUEST model, we review the short-term multipliers of the main measures undertaken in Latvia and discuss their potential longer-term effect on the economy. The results are compared with the effective economic outturn. To explain differences, we discuss the effects of the external environment, the use of EU funds (as a partial substitute for domestic financing) and confidence effects which could have altered multipliers in the midst of the crisis.

The chapter is organised as follows. In the next section we analyse how fiscal consolidation can be duly measured in periods of high volatility. Section 3 reviews the consolidation performance of Latvia, considering both its composition and timing, and compares it to its Baltic neighbours. Section 4 provides estimates of long- and short-term effects of the fiscal consolidation on the Latvian economy. It also looks at whether non-Keynesian effects may have occurred, offsetting the standard multipliers, and provides a tentative measurement of their relevance. Section 5 presents the main lessons we can draw from the Latvian experience. Section 6 concludes.

2 Measuring fiscal consolidation

In the recent case of Latvia, nobody can really argue about the fact that the budgetary adjustment over the period 2008-11 has been huge. However, nailing down a number to identify the amount of such consolidation is not easy. The problems of measuring fiscal effort have been extensively discussed in academic literature. There are two main approaches to determining the size of fiscal consolidation: one based on changes in the cyclically adjusted primary balance (CAPB),[1] also called "conventional" or "top-down" approach, and one based on policy action, also

---

* European Commission.

1 Whenever data is available, using the change in structural balance (cyclically-adjusted balance corrected for one-off and temporary measures) is of course a preferable measure. The change in structural balance, measured following the methodology described in Giorno *et al.* (1995) and Girouard *et al.* (2005) is also used in the context of the EU fiscal policy surveillance.

Table 1

The Extent of Latvian Fiscal Consolidation, as Captured by Different Measures
(percent of GDP)

	2009	**2010**	**2011**
Change in cyclically-adjusted primary balance (CAPB), European Commission 2012 Spring Forecast	0.5	1.3	3.3
Change in structural primary balance, European Commission 2012 Spring Forecast	0.8	1.6	1.9
Consolidation measures as reported by the government (Convergence Programme April 2012)	9.5	4.0	2.3

Source: Commission Services.

called "historical" or "bottom-up" approach. These two methods have sometimes led to quite different results in terms of measuring the fiscal effort. For example, the following reasons for deviations are explored in IMF (2010) and Guajardo *et al.* (2011):

- the change in CAPB does not capture unrecorded one-off and temporary measures and their subsequent reversals (though this bias is removed when structural balances are considered);
- the cyclical adjustment methodology does not sufficiently capture changes in tax bases during periods of sharp contractions of economic activity, notably changes in stock and house prices, fall in consumption or wage bill as a share of GDP etc.

Moreover, the difficulty of determining the cyclical position in real time implies an additional uncertainty when calculating cyclically-adjusted fiscal figures. This is amplified in periods of significant adjustment in the economy, such as the past few years in Latvia.

However, even if measurement problems could be completely eliminated, there are still situations where these two approaches would produce diverging results. Firstly, the policy action approach measures the impact of discretionary fiscal policy against the unchanged policy scenario, while the cyclically-adjusted balance aims at capturing a non-cyclical increase or decrease in the ratio of revenue or expenditure to GDP. The results could in particular diverge for large expenditure items (like social benefits and public sector wage bill) when their recurrent indexation (not captured by policy action) leads to changing their ratio to GDP. Secondly, the cyclical adjustment is based on potential GDP and when the potential output itself (or its measure) changes, this could automatically lead to a change in the cyclically-adjusted balance due to the rigidity of expenditure (or revenue elasticity being different from unity). Thus, in the case of a falling potential output (or its statistical revision) a policy action might be needed just to keep the cyclically adjusted primary balance ratio stable.

Latvia is one of countries where these two measures give particularly diverse results for the period of fiscal consolidation, especially in 2009 (see Table 1),[2] therefore the potential sources of difference for that year are discussed in more detail below.

[2] The discrepancy of a similar magnitude, amounting to 7½-9 per cent of GDP depending on the way of measuring the CAPB, was also recorded in Ireland in the same year, see Guajardo et al and European Commission (2011b).

A closer look at the developments in 2009 shows that the CAPB suffers from the following set of issues, related mainly to composition effects and changes therein not captured by the established methodology:

- it underestimates the effects on indirect taxes of an absorption boom and bust: The recent cycle was characterised in Latvia (as well as in other Baltic economies, see European Commission, 2010) by particularly large swings in domestic demand: in 2006-08 the domestic demand exceeded production by around 20 per cent and was reflected in a sizeable current account deficit. This trend reversed sharply in 2008-09 along with falling confidence and contracting credit supply. Unfortunately, the effect of this extreme domestic demand cycle is improperly captured by the cyclical adjustment of the general government's budgetary position used in the EU budgetary surveillance, which adjusts budgetary components based on fluctuations in *output* rather than *absorption*. The alternative measure, which would allow capturing also the effect of the cycle on the tax base for indirect taxes, is an "*absorption gap*" (see Lendvai *et al.*, 2011), which aims at capturing the effect of the current account being above or below the current account norm determined by fundamentals, similarly to the way output gap measures fluctuations of output around its potential level. For Latvia, such a correction of the cyclically-adjusted balance would imply higher underlying deficits in 2005-08, but lower underlying deficit in 2009. Overall, this adjustment could reduce the discrepancy between "top-down" and "bottom-up" measurements of the fiscal effort by **2.6 percentage points** of the difference. An alternative explanation is offered in Sancak et al (2010), where the authors analysed responsiveness of tax revenue and in particular VAT to the business cycle. They found that on average a one percentage point increase (decrease) in the output gap corresponds to 1¾ percentage point increase (decrease) in VAT revenue; the key channels through which the output gap affects the revenue are found to be shifts in consumption patterns towards goods and services with higher (lower) VAT rates and lower (higher) tax evasion during economic expansions (contractions). This approach could explain **1.4 percentage points of GDP** fall in indirect taxes in Latvia in 2009, which occurred against sizeable indirect tax increases in that year, but would still leave about 1 per cent of GDP unexplained, suggesting that the effect of these shifts in behaviour might have been even stronger in Latvia than for the panel of countries covered in the study.
- it underestimates the effects of a reversal in labour taxes: as already discussed above, the standard cyclical adjustment methodology, including the one used by the European Commission, takes into account variations in GDP but not variations between individual tax bases, therefore large fluctuations among the latter will be omitted by the cyclical adjustment methodology. In Latvia in 2009 a particularly large change was observed in the ratio of compensation of employees to GDP, which dropped from 50.8 per cent in 2008 to 46.7 per cent in 2009. It should be noted that in comparison to average historical levels of the ratio (43.1 per cent in 2000-10), both years reflected very high wage growth that took place at the peak of the cycle, but it nevertheless helps to explain why the decline in labour taxes is not fully captured by the adjustment. If the ratio of compensation of employees to GDP would have remained the same in 2009 as it was in 2008, this could have resulted in additional labour taxes in the magnitude of **1.2 percentage points** of GDP;
- it underestimates the cyclical impact of increase in social outlays: The change in unemployment benefits in response to changing cyclical conditions (captured by the cyclical component of expenditure in calculations of the cyclically-adjusted balance) suggests an increase in unemployment outlays in Latvia in 2009 by LVL 43 m or 0.3 per cent of GDP. However, actual data indicates that unemployment benefits increased by LVL 83 m (0.6 per cent of GDP) in that year. Moreover, the cyclical adjustment of expenditure only captures an increase in unemployment benefits, while expenditure on sickness and disability benefits similarly increased in 2009 by some LVL 50 m (0.3 per cent of GDP) above its level of 2008, which can

be only to a minor extent explained by discretionary policy changes. The possible explanation why the actual increase in crisis-related social spending surpassed the one predicted by the cyclical adjustment by some **0.5 percentage points** of GDP could lie in a behavioural change, as previously inactive part population started looking for job (and/or benefit) opportunities amidst plummeting confidence;

- it does not take into account one-off and temporary factors: As discussed above, the cyclically-adjusted budgetary indicators do not take into account one-off and temporary factors and exceptional costs, for which reason it is preferable to use – whenever available – the structural balance when measuring the fiscal effort. Indeed, there have been large exceptional costs related to the stabilisation of the financial sector in Latvia in 2009-11 related to Parex Bank, with overall impact of 0.9 per cent of GDP in 2009, 1.7 per cent of GDP in 2010 and 0.2 per cent of GDP in 2011. These costs were, however, partly offset by a government's decision to retain in the publicly managed pension system part of the social tax previously transferred to privately managed pension funds. The overall impact of temporary and exceptional measures resulted in a **0.3 percentage points** worsening of the general government balance in 2009.

On the other hand, it is also true that the **consolidation amount expressed by the government does not include all measures which should have been recorded as discretionary policy.** As discussed above, around half of the difference between the change in cyclically-adjusted primary balance and the policy action approach can be related to factors not fully captured by the conventional cyclical adjustment. However, there are also factors not covered by the government's estimate of discretionary policy that affect the cyclically-adjusted balance. Notably, non-cyclical social benefits (particularly pensions) increased considerably in 2009, by 2.3 percentage points of GDP (Republic of Latvia, 2009a). These increases covered sizeable indexation of pensions due a lagged effect of years of high wage growth,[3] an increase in pension supplement for pre-1996 years of service (which was initially intended only for lowest pensions but eventually extended to all pensioners),[4] as well as some other increases. The largest part of this increase – approximately two thirds – related to pension indexation and did not constitute a discretionary policy change, but the remaining third is simply omitted from the government's policy action estimate. At the same time, the cyclically-adjusted balance likely captures all of this increase in social spending that took place in 2009, offering another sizeable explanatory factor for the difference between two approaches. Moreover, given the limited and unsophisticated nature of the social safety net in Latvia before the crisis, it became clear as the crisis evolved that the system cannot fully cope with the cyclical impact. For this reason, ensuring adequate social safety net has been from the onset an important part of the stabilisation programme, with additional social safety net measures amounting up to 1 per cent of GDP in 2009 (Republic of Latvia, 2009b), although in practice their impact was somewhat lower in that year. The combined effect of these social benefit increases could thus account for approximately another **3 percentage points** of GDP of the difference between the two approaches. Overall, the possible sources of discrepancy between the change in cyclically-adjusted primary balance and policy action approach in 2009 are summarised in Table 2.

Overall, the evidence provided in this section points to the need to be very cautious in using CAPB as the only measure for the assessment of fiscal consolidation, and to duly consider also the

---

3 Until 2009, pensions were indexed in Latvia twice a year, in April and October, on the basis of CPI and social security benefits' trends. Particularly sizeable indexation of pensions took place in the course of 2008. As a result, an average old age pension in December 2008 was higher by 32 per cent than in January 2008 (according to data published on the website of the State Social Insurance Agency); following the introduction of supplementary pensions from January 2009 an average old age pension increased further by about 7 per cent. The pension indexation has been suspended from 2009 (until end-2013, according to current plans), although average pensions continue increasing somewhat as the share of new retirees, who tend to have higher pensions, gradually rises.

4 The government tried to reverse part of this increase through the 2009 supplementary budget, but this was rejected by a Constitutional Court ruling, leaving social benefits at higher level.

Table 2

Contribution of Different Items to the Discrepancy Between Alternative Measures of Fiscal Consolidation

	Impact
Cyclical adjustment underestimating fall in indirect taxes	1.4 – 2.6 pps
Cyclical adjustment underestimating fall in labour taxes	1.2 pps
Cyclical adjustment not capturing behavioural changes in social benefits outlays	ca 0.5 pps
Exceptional financial sector costs net of other temporary measures	0.3 pps
Policy action approach "missing" expansionary elements	up to 3.0 pps
TOTAL	**up to 7½ pps**
+ uncertainties related to real time estimates of output gap, differences in measurement methodologies, etc.	…

Source: Commission Services.

"bottom-up" approach for the analysis and policy conclusions.

3 Fiscal consolidation in Latvia and comparison with the other Baltic countries

As established in the previous section, the Latvian authorities have implemented – in particular in 2009 and 2010 – a very substantial fiscal consolidation, although measuring its magnitude is a complicated issue due to very abrupt changes that took place in the Latvian economy over the period of economic adjustment. Some insights into the mechanism of consolidation could, however, be obtained by going into a more detailed analysis of the adjustments, and by comparing evolution of fiscal indicators in Latvia to those of the other Baltic economies, given that economic developments have been similar and all three countries have implemented a broadly comparable fiscal adjustment over the period of 2009-11.

As a starting point, one could observe that total-revenue-to-GDP ratio in Latvia stayed unchanged between 2007 and 2011, while the tax-to-GDP ratio actually declined despite numerous and sizeable tax measures. Broadly similar developments took place in Lithuania, while in Estonia both revenue-to-GDP and tax-to-GDP ratios increased over the same period. The expenditure-to-GDP ratio increased sharply in all three countries between 2007 and 2010, but only Latvia succeeded in bringing the level of government consumption back to the 2007 level already by 2010 (see Table 3). These developments are analysed here in further detail.

3.1 The revenue side of the fiscal consolidation

On the revenue side, it is important to note that the share of tax revenue to GDP declined in Latvia in 2009, compared to 2008, despite very sizeable tax measures that entered into force from the beginning of that year and amounted in total to 3.3 per cent of GDP (*ex ante* estimate), of which 2.6 per cent on the side of consumption taxes. While partly explained by falling revenue elasticities discussed above, this contrasts developments in Estonia, where the share of taxes, including consumption taxes, to GDP actually increased in 2009, even though main tax measures only entered into force from the second half of the year (see Figure 1).

Table 3

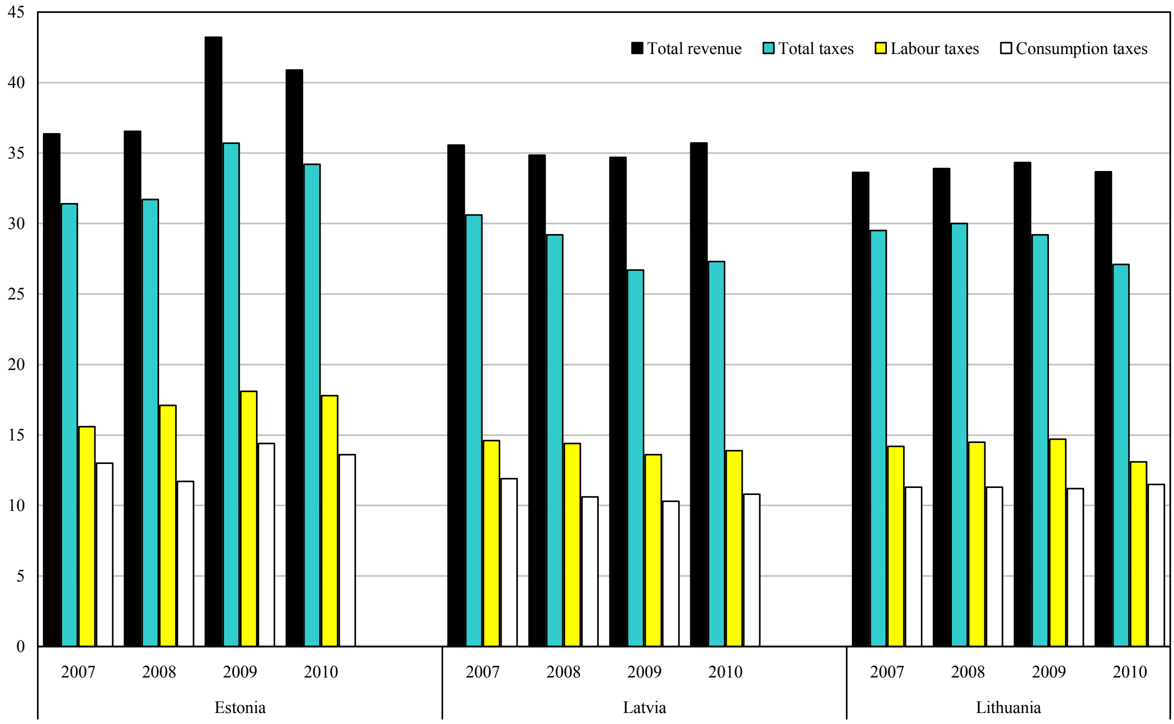
Evolution of Revenues and Expenditures in the Baltics
(percent of GDP)

	Latvia					Lithuania					Estonia				
	2007	**2008**	**2009**	**2010**	**2011**	**2007**	**2008**	**2009**	**2010**	**2011**	**2007**	**2008**	**2009**	**2010**	**2011**
Total revenue	35.6	34.9	34.7	35.7	35.6	33.6	33.9	34.3	33.7	32.0	36.4	36.5	43.2	40.9	39.2
o.w. tax burden	30.8	29.7	27.0	27.5	27.9	29.9	30.4	29.8	27.5	26.5	31.5	31.8	36.1	34.1	33.2
Total expenditure	36.0	39.1	44.4	43.9	39.1	34.6	37.2	43.8	40.9	37.5	34.0	39.5	45.2	40.6	38.2
o.w. gov. consumption	17.8	20.0	19.6	17.5	15.6	17.8	19.2	22.0	20.5	18.9	16.4	19.2	22.0	20.9	19.5
o.w. social transfers	7.1	8.1	12.6	12.5	10.8	9.1	10.9	15.2	13.0	11.2	8.5	10.5	14.0	13.1	11.7
General government balance (EDP)	**–0.4**	**–4.2**	**–9.8**	**–8.2**	**–3.5**	**–1.0**	**–3.3**	**–9.4**	**–7.2**	**–5.5**	**2.4**	**–2.9**	**–2.0**	**0.2**	**1.0**

Source: Commission Services.

Figure 1

Ratio of Total Revenue, Tax Revenue and Revenues from Main Tax Categories to GDP in the Baltics



Source: Eurostat, *Taxation Trends 2012*.

As already discussed above, the cyclical impact of falling tax elasticities could explain approximately half of "missing" consumption taxes in Latvia in 2009 (while in Estonia this effect seems to have taken place earlier, in 2008). Indeed, VAT compliance dropped very substantially in Latvia between 2007 and 2009 – considerably more than in other two Baltic countries – and still remains the lowest in three countries, even though before the crisis the indicator was above that of Lithuania; similar trends can be observed with regard to VAT C-efficiency.[5]

Labour tax developments in Latvia in 2009 are less contradictory and their fall in relation to GDP can be largely explained by wage bill dynamics discussed above; such a sharp decline in the ratio of wage bill to GDP was only observed in Latvia. Nevertheless, the fall in implicit tax rate on labour in crisis years (to some extent offset by retaining some of revenue that previously was redirected to the mandatory pension pillar from 2009) suggests that compliance rates have fallen with respect to labour taxes as well.

All these factors point to a considerable deterioration of tax compliance in Latvia in 2008-09, which occurred alongside sizeable measures to increase tax revenue. The resulting level of statutory tax rates is now somewhat higher in Latvia than in other two Baltic countries, but the tax efficiency appears the lowest.[6] To bring the level of statutory rates closer to those of other Baltic economies, the Latvian Parliament passed on 24th May 2012 changes to tax laws that foresee lowering VAT rate from 22 to 21 per cent from July 2012 and gradually lowering the personal income tax from 25 to 20 per cent over the three year period starting from 2013.

EU funds have also played a very prominent role as a revenue source particularly in last years, due to a combination of factors: firstly, major projects related to the 2007-13 financial perspective became operational with a few years lag and, secondly, intensifying the absorption of EU funds was a deliberate policy to support the domestic demand in crisis years in line with the conditionality of the EU Balance-of-Payments assistance programme. However, this also implies that while it should be possible to sustain comparable level of capital revenue in the short term, in the medium term the amounts will decline as respective financing is exhausted; this might already be the case for current transfers in the short term.

3.2 The expenditure side of the fiscal consolidation

Scaling back expenditure played the crucial part in the Latvian fiscal consolidation strategy – according to the authorities' estimates, savings on the expenditure side amounted to around 10 per cent of GDP over the period of 2008-11, of which 6.7 per cent of GDP in 2009 alone. The expenditure side consolidation was centred on cuts in government consumption, which mostly cover public sector wages and good and services procured by the government. At the same time, social benefits remained broadly intact throughout the crisis, with an increase in some categories.

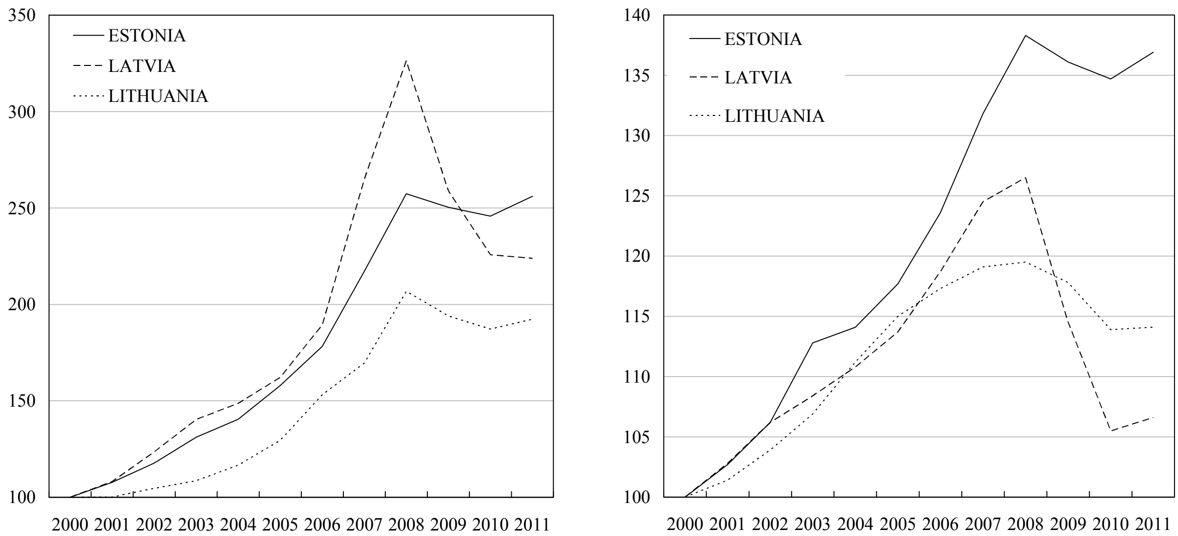
Statistical indicators confirm that the magnitude of consolidation on the side of government consumption was unprecedented and constitutes the most remarkable feature of Latvia's fiscal adjustment, with government consumption contracting by a fifth in real terms between 2008 and 2010 and, even more shocking, by almost a third in nominal terms over the same period. Latvia was the only country in the Baltics to bring the ratio of government consumption to GDP back to the level of 2007 already in 2010, despite a substantial fall in economic activity, underlying that government spending was cut most substantially in Latvia among the Baltic economies (see

5 VAT compliance ratio measures proportion of VAT actually collected in relation to theoretically possible collection, based on the value of private consumption and scope of application of standard and reduced VAT rates (using HICP weights). VAT C-efficiency uses only standard VAT rate and overall consumption, thus measuring both compliance and policy gap.

6 Based on the analysis of consumption and labour taxes; taxation of capital cannot be easily compared across the Baltics, notably due to a different system in use in Estonia, where only distributed profits are taxed.

Figure 2

Evolution of Government Consumption in the Baltics
(left panel: values, right panel: volumes)



Source: Eurostat.

Figure 2). However, consumption also increased most during the boom years in Latvia, in particular on the side of prices. This unsustainable trend prior to the crisis – as well as the fact that very decisive measures were taken in 2009-10 to bring the government consumption back to sustainable levels – has been acknowledged by Åslund and Dombrovskis (2011). For example, according to the book half of 75 state agencies (in a country with a population of about 2 million) were to be closed down according to the 2009 stabilisation programme.

Among government functions, health related expenditure, defence expenditure and education expenditure stick out as areas most affected by the cuts in Latvia: between 2008 and 2010, health related expenditure declined by approximately 27 per cent (while "only" by 10 and 7 per cent in Estonia and Lithuania respectively) and education related expenditure declined by about 26 per cent (compared with a decline of about 10 per cent in both Estonia and Lithuania). However, in particular in education expenditure also increased most rapidly prior to the crisis in Latvia. The defence budgets were substantially decreased in all three countries, but again most notably – by almost a half – in Latvia, by over a quarter in Lithuania and by around tenth in Estonia. The provision of general government services also declined most notably in Latvia. At the same time, expenditure on economic affairs (which among other things reflect EU funds absorption) actually increased in Latvia over the period of 2008-10, while declining most notably in Estonia and to a lesser extent in Lithuania.

Both Åslund and Dombrovskis (2011) and World Bank (2010) shed some light on these exceptional developments with regard to healthcare and education sectors: both sectors were in a need of radical reforms to align the provision of services to demographic trends and to improve efficiency. These reform plans were available, but the implementation was delayed due to the lack of political support. The crisis – which revealed the need to bring public finances on a sustainable path – acted as a catalyst for reforms, which were implemented over a very short period of time. The World Bank (2010) later noted that "Latvia has achieved years' worth of difficult structural reforms in the short space of just a few months".

Another area where existing reform plans might have helped to implement the expenditure-side consolidation, was the administrative territorial reform enacted from 1 July 2009. As a result of this reorganisation, one administrative level was completely abolished and the number of territorial units declined from 548 to 119 (110 municipalities and 9 republican cities). The reform had no direct link to the consolidation strategy and had been prepared for years, but is nevertheless likely to have had a positive impact on local governments' finances. Similarly to developments at the level of general government discussed above, expenditure of municipalities increased fastest among the three Baltic countries prior to the crisis in Latvia, but also declined most abruptly in 2009-10.

3.3 The timing of the consolidation

Although fiscal consolidation officially started at the end of 2008, when the Latvian authorities turned to the EU, the IMF and regional neighbours for the financial assistance that resulted into the Balance-of-Payment assistance programme, it was not until the second half of 2009 that the bulk of consolidation actually took place. On February 2009, in fact, the government fell over concerns about its handling the economic crisis and its inability to impose the austerity measures agreed with the international lenders, leading to the formation of a new government in March 2009 whose explicit mandate was to implement the agreed fiscal austerity.

Given the deterioration of the economy during the first months of 2009 and the inability of his predecessor to actually implement the consolidation measures, the newly appointed government needed to act quickly and decisively to restore confidence and redress the situation. For this reason Prime Minister Valdis Dombrovskis and the international lenders agreed on the need to clearly front-load the consolidation and to adopt the necessary measures as soon as possible, adopting in the supplementary budget of June 2009 all the necessary measures to keep the government deficit below the 10 per cent of GDP, and to implement a progressive consolidation bringing the deficit below the threshold of 8.5 per cent in 2010 and 6 per cent of GDP in 2011, ultimately correcting the excessive deficit (i.e. bringing the deficit below 3 per cent of GDP) by 2012.

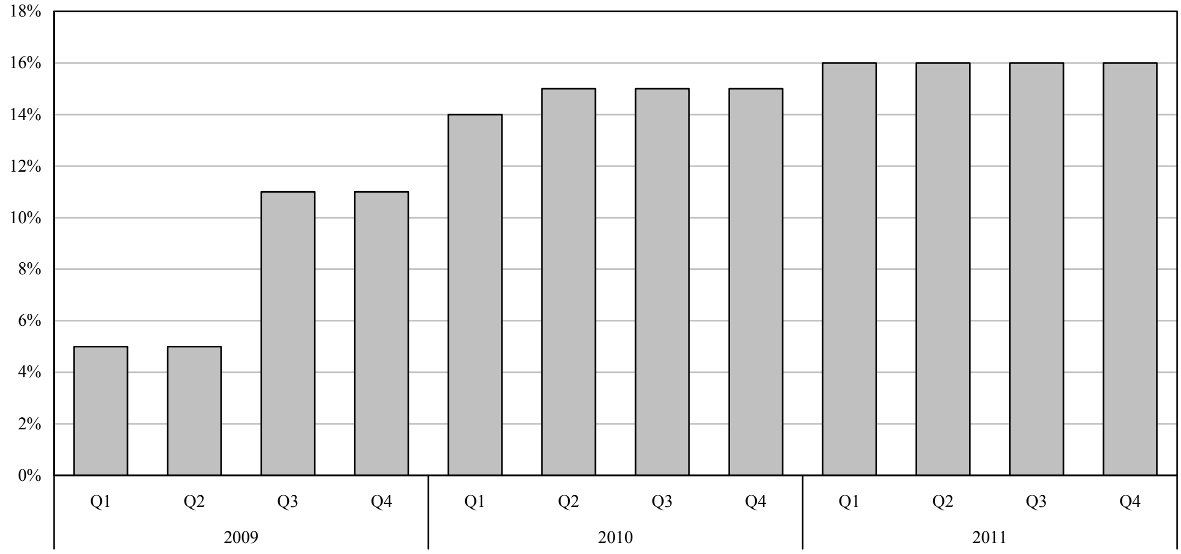
In June 2009 a massive set of measures of over 4 per cent of GDP were adopted with the supplementary budget 2009, and in July measures concerning the 2010 budget were already proposed by the government and negotiated with international lenders with a view to reassure about the subsequent steps. Finally, in November 2009 an additional package of fiscal adjustment was adopted, entering immediately into force and defining the key elements of fiscal consolidation in 2010. It can thus be said that the bulk of the consolidation (about 10 per cent of GDP) was actually designed and adopted in less than six months, in the course of the second half of 2009. This represented a strongly front-loaded and credible adjustment, which affected market's perception of the Latvian situation already from the beginning of 2010. In Figure 3 we report a tentative quarterly accounting of the effective entry into force of the measures, where the series has been built on the basis of government's *ex ante* commitments and expenditures have been checked against *ex post* reporting.

4 Fiscal consolidation and economic activity

In this section we analyse the interplay between fiscal consolidation and growth in Latvia. Fiscal multipliers of the above-mentioned measures are presented and compared to the actual GDP data. What is remarkable about the Latvian experience is that significant cuts in government expenditures and tax hikes coincided with a robust economic recovery, pointing to the existence of relevant non-Keynesian effects offsetting the contractionary Keynesian effects of fiscal consolidation.

Figure 3

Quarterly Accounting of Fiscal Consolidation Measures in Latvia
(percent of quarterly GDP)



Source: Commission Services.
Note: the series has been built on the basis of government's *ex ante* commitments and projections for revenues, with expenditures checked through *ex post* reporting.

4.1 *Fiscal multipliers in the long and short term*

Using the latest version of the Commision-developed dynamic stochastic general equilibrium (DSGE) model, QUEST (Ratto *et al.*, 2009), in this subsection we compare the fiscal multipliers of various budgetary measures and get some insights on the theoretical impact of composition of the Latvian consolidation measures. QUEST is a large-scale open economy new-Keynesian model used for policy analysis.[7] The model economy is described by optimal decisions of households and firms. There are three production sectors: a construction sector and two manufacturing sectors producing traded and non-traded final consumption goods.

The model features three types of households:

- A share of households are "Ricardian": they own capital and have unlimited access to financial markets; their consumption decisions are based on the life-time income hypothesis;
- Another share of households are "collateral-constrained": they have limited access to credit markets and can only get indebted against the value of their collateral (housing stock) up to an exogenously given level;
- The third type of households is so-called hand-to-mouth consumers: they do not have access to financial markets and consume their after-tax labour income and transfer earnings in every given period.

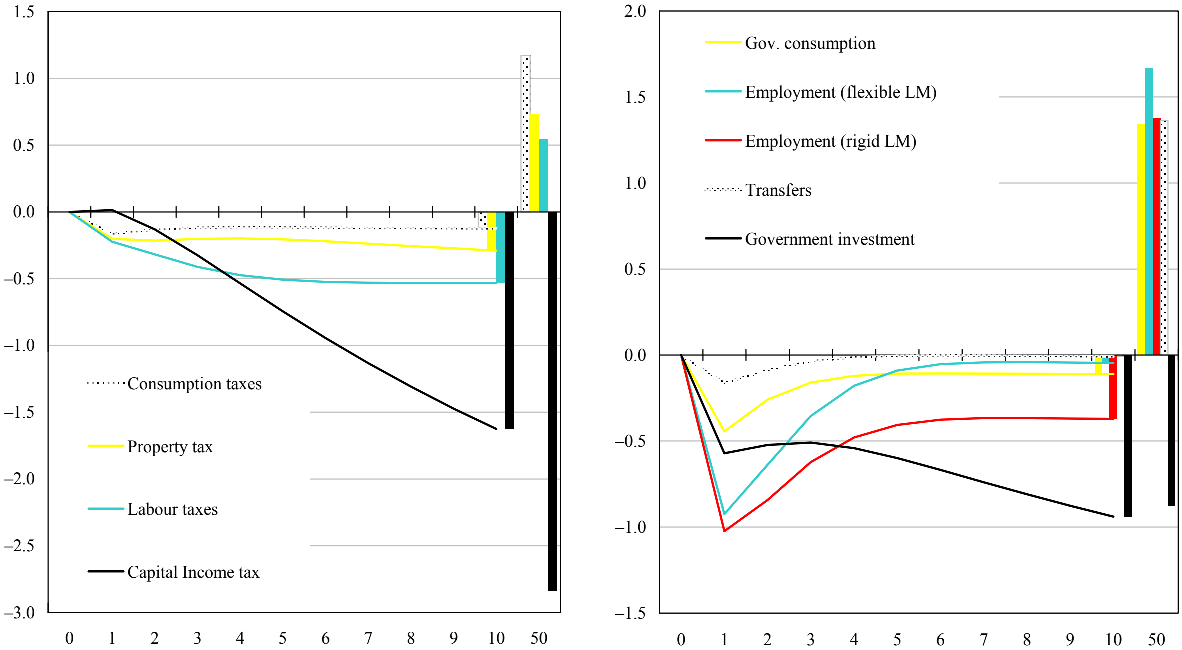
Fiscal policy is described by a rich set of fiscal instruments. The government can raise revenues by a tax on consumption (VAT), on personal income (PIT), on corporate income (CIT) or on immovable property, via social security contributions and finally via a lump-sum tax. The fiscal

---

7 For a comprehensive review of alternative structural models used for policy analysis, see Cogan *et al.* (2010) or Coenen *et al.* (2012), where a comparison of IMF, ECB and QUEST models can be found.

Figure 4

Annual Impact Analysis of Different Tax Hikes and Expenditure Cuts Leading to a Permanent Budgetary Consolidation of 1 Percent of GDP, Simulated with QUEST II



Source: Commission Services.

authority spends on government consumption, government investment, unemployment insurance benefit payments, and transfers. Government consumption is further broken down into intermediate government consumption (unproductive expenditures) and compensation of employees (which equals government output following standard national account practices). The government budget does not need to be in balance every period. Fiscal deficits are financed by changes in the public debt. The model is closed down by a debt rule according to which one of the above instruments reacts endogenously to stabilise debt in the long-run at its target. The presence of non-Ricardian households allows for Keynesian transmission channels of fiscal policy.

The model incorporates various real, nominal as well as financial frictions to match the dynamic response of the economy to standard shocks. It was calibrated to the Latvian economy for size, openness, trade shares and relative size of each component of GDP. In addition, the monetary policy is characterised by a fixed exchange rate regime.[8]

To evaluate the economic impact of fiscal consolidation, this section looks at the multipliers of fiscal consolidation of a given size achieved by different instruments. In particular, Figure 4 displays the impact of different tax hikes and expenditure cuts leading to a permanent budgetary consolidation of (*ex ante*) 1 per cent of GDP using one instrument at the time. Given the model's assumptions about long-run real and nominal growth rates, a 1 per cent of GDP reduction in the fiscal deficit corresponds to a 27 per cent of GDP reduction in the long-run debt target. In the simulations in this section it is assumed that fiscal space gained by the long-run debt reduction is used to decrease labour income taxes over time.

[8] For a detailed description of the model see, e.g., Lendvai and Roeger (2010).

The simulations suggest that fiscal consolidations have a negative impact on economic activity in the short run. Over time, however, if the fiscal space is used to reduce distortionary taxes (labour income taxes in the simulations), the effect of the consolidation turns out to be positive for most of the instruments. Further, as can be seen in the figure, the results indicate that expenditure cuts may have a larger impact on GDP than tax hikes, although this effect tends to turn around very quickly. The model also confirms that VAT and property taxes are less distortionary than labour income tax, whereas tax on capital income leads to a reduction of the economy's capital stock over time and thereby leads to a significant reduction in production as well.

On the expenditure side, a reduction in transfers and unproductive government investment leads to the smallest short-run negative impact on GDP. Reduction in the compensation of employees (either via public wage cuts or by lay-offs in the public sector) may have significant negative effects on total GDP in the short-run. Over time, the reaction of the economy depends on the flexibility of the labour market (wages and movements of employees between sectors): the more flexible the labour market, the more private GDP will pick up in response to the reduction of public employment or public wages – and hence the less negative/the more positive the long-run effect will be. Finally, a reduction in productive government investment spending reduces productivity in the private sector and therefore turns out to be rather detrimental for overall economic activity over a longer horizon.

The simulations provide a benchmark that can be used to assess the impact of the composition of fiscal consolidation on growth in the short and long term (when fiscal space gained through the consolidation can be used to reduce distortive taxes). As the above discussion suggests, an optimal mix of measures would have implied higher consumption taxes on the revenue side and cuts in government consumption and employment on the expenditure side, especially as far as the long-term benefits are concerned. It is important to notice that the actual effects of the cuts in public employment depend on how flexible is the labour market and, more precisely, on how smoothly workers can move from the public to the private sector. In the case of Latvia one could safely argue that labour market institutions are rather supportive of high labour turnover and thus we can reasonably expect the flexible labour market multiplier to provide better guidance than the rigid labour market multiplier in forecasting the effects of consolidation on Latvian GDP.

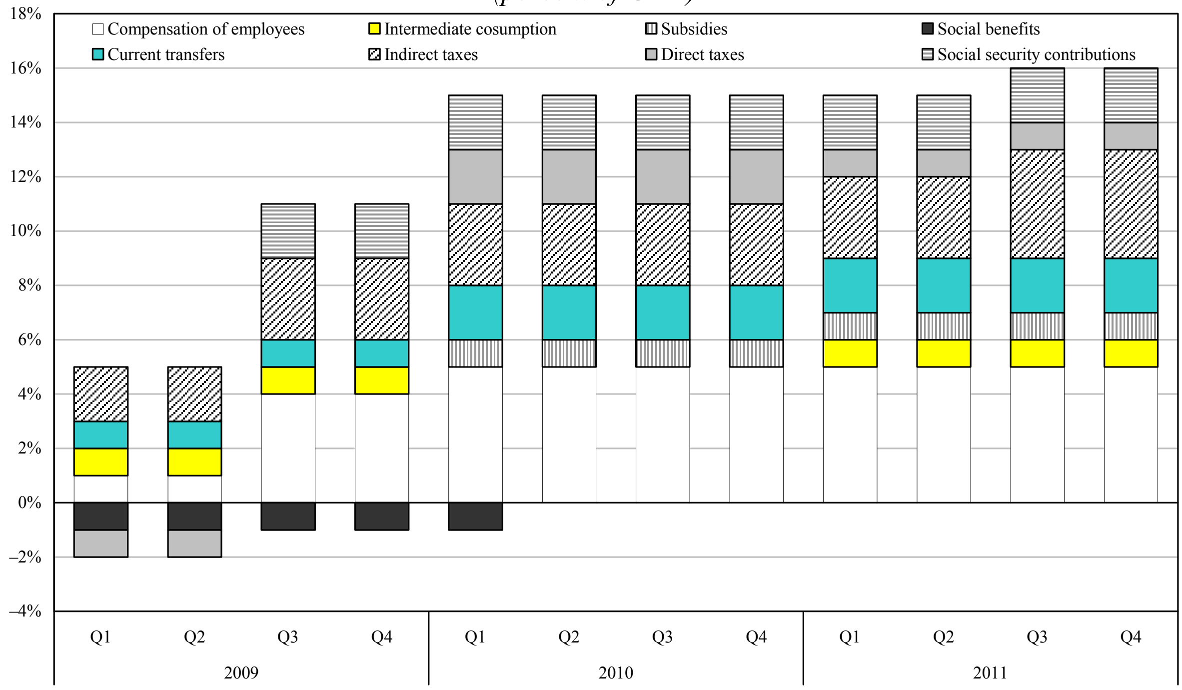
As we can see from Figure 5, those measures were indeed prominent in the actual composition of the fiscal consolidation undertaken by the Latvian government under the supervision of the international lenders. In particular, public employment (in the form of both wage cuts and reductions in the number of employees) stands out as the most important single item of consolidation over time, followed by indirect taxes (composed mostly of consumption taxes).

Latvia's fiscal consolidation was therefore clearly designed to maximise long-term gains, but what about the short term effects? A quarterly accounting of fiscal consolidation can allow us to identify how the fiscal multipliers associated with the timing of consolidation may have affected GDP growth in each quarter. It should be kept in mind, however, that it is virtually impossible to have a precise quarterly accounting of the fiscal measures, as it entails a certain degree of arbitrariness in the imputation of policies formally implemented during the year and for which is it not possible to monitor the effective implementation. This implies that also the multipliers' estimation may be affected and should be interpreted as indicators of the order of magnitude of the effects rather than as precise numbers.

Figure 6 illustrates the economic effect of the Latvian fiscal consolidation undertaken since 2009 based on simulations with the QUEST model. The simulation assumes that the consolidation takes place against a high deficit baseline which is assumed to be long-lasting before the consolidation is announced in 2009q1. Further, it is assumed that the entire set of consolidation measures is announced in 2009q1 and that it is believed to be permanent and perfectly credible.

Figure 5

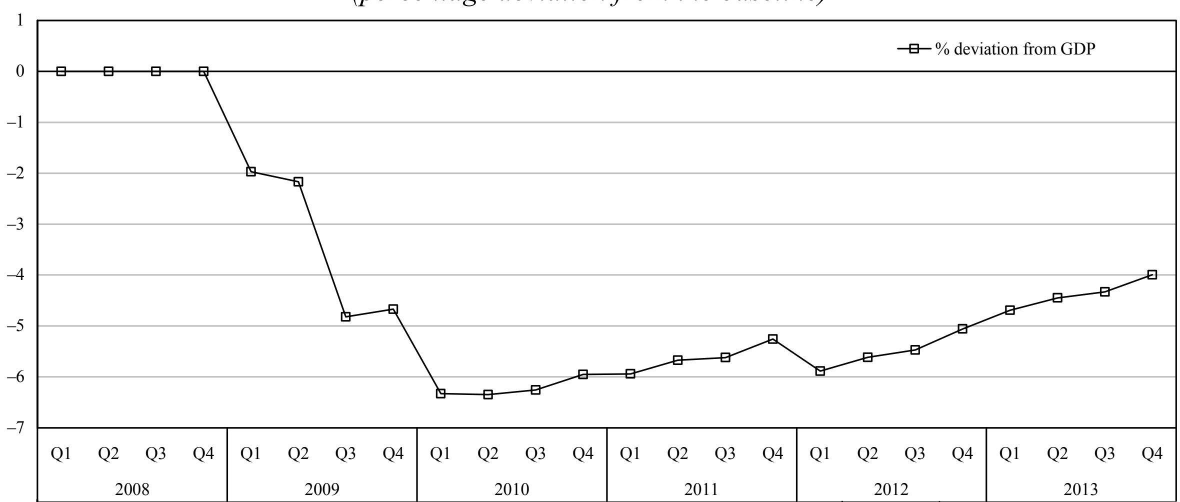
Quarterly Disaggregation of Fiscal Consolidation in Latvia, by Individual Measures
(percent of GDP)



Source: Commission Services.
Note: the series has been built on the basis of government's *ex ante* commitments and projections for revenues, with expenditures checked through *ex post* reporting.

Figure 6

Quarterly Impact on GDP of the Actual Mix of Latvian Fiscal Consolidation Measures Simulated with QUEST II
(percentage deviation from the baseline)

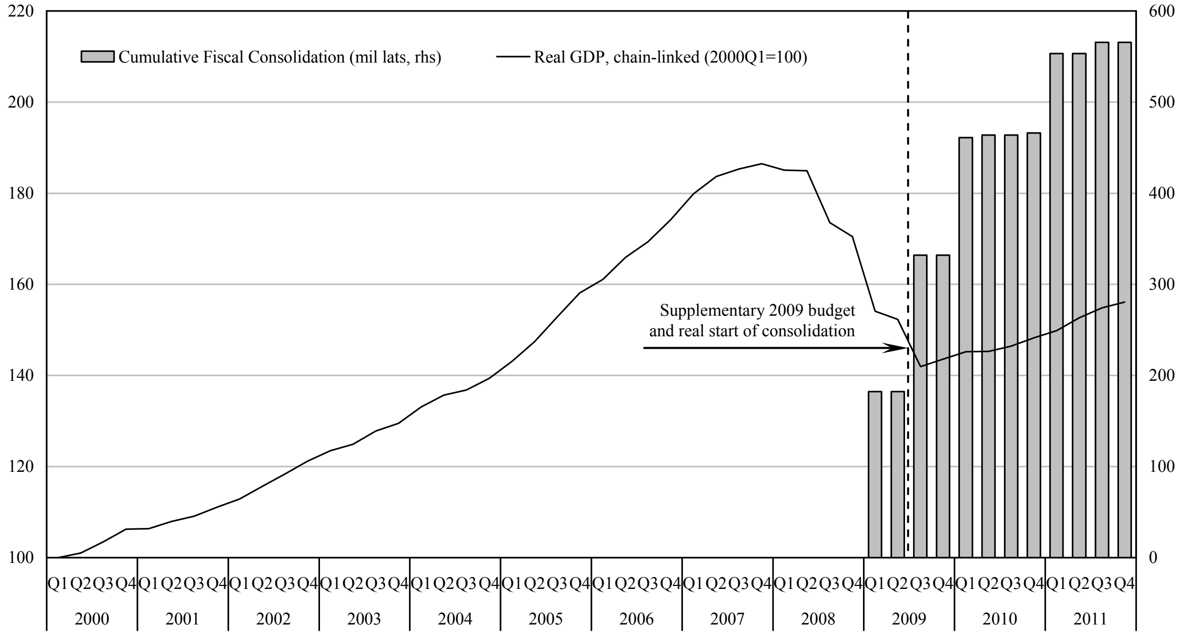


0	0	0	0	–2	–2.2	–4.8	–4.7	–6.3	–6.3	–6.3	–6	–5.9	–5.7	–5.6	–5.3	–5.9	–5.6	–5.5	–5.1	–4.7	–4.4	–4.3	–4

Source: Commission Services.

Figure 7

Quarterly Account of GDP Growth and Fiscal Consolidation Measures



Source: Commission Services.

The fiscal space resulting from the consolidation is used to reduce lump-sum taxes – the least distortionary tax in the model. While it may be argued that lump-sum tax is an artificial instrument which is not available in practice to policy makers, this assumption allows us to attribute as little as possible positive confidence effects to the short-run impact of the fiscal consolidation in our simulations.

The simulation results suggests that the negative effects of consolidation were expected to kick in progressively as the consolidation plan unfolded, reaching more than 6 per cent of GDP in the first quarter of 2010 and then fading away slowly, as the effects of additional measures in the following quarters played against the recovery from the effects of the first negative shocks. In a way, this series can be interpreted as showing the theoretical short-term pain the Latvian economy could have endured in the absence of non-Keynesian effects.

However, a quarterly look at the time pattern of total consolidation undertaken and GDP growth reveals that GDP growth reversed to positive almost immediately after serious consolidation started in the second half of 2009 following the supplementary budget measures envisaged in July (see Figure 7) and by the end of 2011 real GDP was already 10 per cent higher than 2 years earlier and, remarkably, 56 per cent higher than it was at the beginning of the decade. In order to understand what may have caused such a quick recovery in the presence of significant fiscal consolidation, in the next section we investigate what role non-Keynesian effects may have played in the post-crisis Latvia.

4.2 Non-Keynesian effects

Since the seminal contribution of Giavazzi and Pagano (1990), we know that under certain conditions fiscal consolidation can trigger non-Keynesian effects as strong, or even stronger, than standard contractionary Keynesian effects on demand (Giavazzi *et al.*, 2000). When this happens, consolidation may turn out to be expansionary and result in a quick rebound of the economy, of the kind observed in Latvia in late 2009. In other words, the sign and magnitude of fiscal multipliers depend on particular conditions under which fiscal policy is implemented. As noted by Alesina and Ardagna (1998), the main channel through which non-Keynesian effects are activated is aggregate demand: a serious fiscal tightening may indeed increase both consumption and investment, as wealth rises when future tax burdens decline and interest rates decline when credibility is restored and inflation or default risks abate. Indeed, the improvement in the fiscal position may immediately affect consumer confidence, business confidence, and in particular it may lead to a reduction in risk premia which influence the economy's borrowing costs and thereby also the cost of capital. For this effect to produce an expansion, though, the tightening must be sizeable, credible, and occur after a period of stress when the budget is quickly deteriorating and public debt is building up (Afonso, 2010; Giudice *et al.*, 2007). The new EU member states, in particular, seem to be prone to such growth-enhancing consolidation (Rzonca and Cizkowicz, 2005).

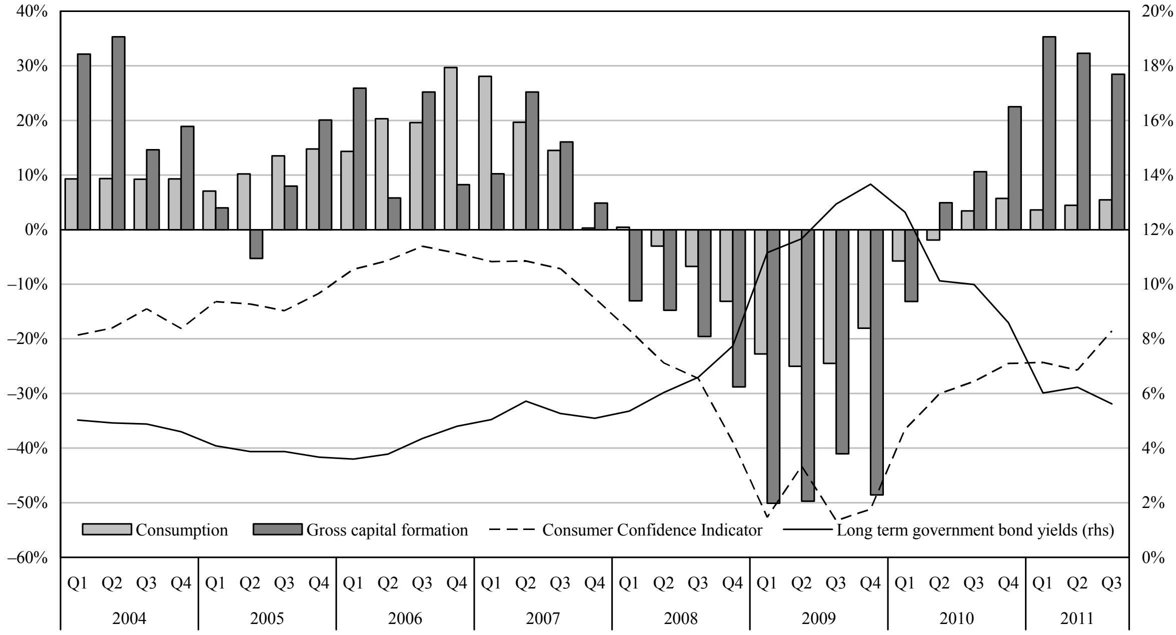
An increase in consumer confidence may raise current consumption through expectation of higher future income and the willingness to consume today part of the expected future gains. Consumers could both expect taxes to be lower in the future, as a consequence of current consolidation or their gross income to be higher due to an improvement in the fundamentals of the economy. A similar argument can be made for entrepreneurs, who may anticipate higher consumer expenditure and start investing in the economy to have enough capacity to match demand as soon as it picks up. Arguably, these effects are consistently accounted for in the QUEST model, leading to the scenario portrayed in Figures 4 and 6, where it is assumed that the fiscal space gained by the consolidation over time is used to reduce non-distortionary taxes. However, two key determinants of economic performance for small open economies such as shocks in external demand and in country risk premia due to developments in the international financial markets cannot be introduced endogenously in the simulation, even if their directly affect investments and capital formation. For this reason we analyse them separately and then link them to the results of the simulation to determine their likely impacts.

In addition, if undertaken through spending cuts rather than tax increases, fiscal consolidation is likely to produce growth-enhancing gains in external competitiveness. Cuts in government consumption, and in particular in public wages and public employment can spill over to the private sector and abate the costs of domestic manufacturing, leading to gains in international market shares. The process may be more or less quick depending on the particular labour market institutions of the country undertaking the cuts, but eventually the increased availability of labour and lower wages in the public sector are bound to map into a more efficient production process. However, it is worth noting that while volumes exported increase the effect on value of exports is partly offset by the decrease in export prices, so that in some simulations the overall effect in terms of value added is not necessarily very strong.

In the case of Latvia, there is some evidence on the activation of all these channels of economic expansion triggered by fiscal consolidation, each following a slightly different timing. This could contribute to explain the pace of recovery from the crisis. The connection between the renewed confidence in the Latvian Government and risk premia, investments and consumption can be seen from Figure 8. After a constant deterioration of confidence in 2008 and most of 2009, reflecting the impact of the financial crisis first and Government financial sustainability then, it can be seen how the Latvian Government's decision to undertake bold actions to consolidate its fiscal position (mid-2009) was resulted in an improvement in consumer confidence, investments and consumption, whereas risk premia first stabilised and then decreased, at a time in which standard Keynesian wisdom would have predicted further recession due to the contraction in public consumption.

Figure 8

Quarterly Series of Consumption, Capital Formation, Consumer Confidence Indicator and Long-term Government Bond Yields



Source: Commission Services.

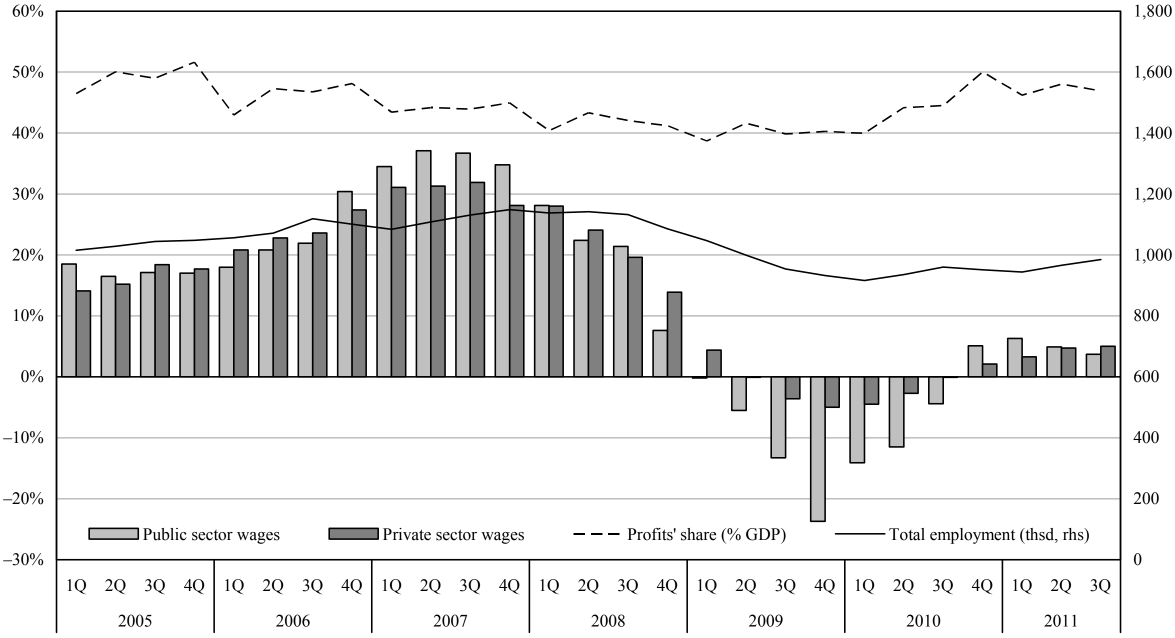
There could be different reasons for the consumption and investments to increase so rapidly between the last quarter of 2009 and the first half of 2010. For example, an increase in consumption and investment could have been driven by higher wages in the private sector or gains in total employment, or also it could have come from a sudden increase value added in export-oriented sectors, due to an increase in Latvian competitiveness or an increase in external demand. We investigate these channels and find no evidence to support them. As a matter of fact, wages and total employment actually decreased as a result of the Government-led internal devaluation strategy, as shown in Figure 9. In addition the profits' shares in the economy remained constant while the economy contracted, meaning that lower wage bill didn't lead immediately to higher profits to reinvest in the economy.

Indeed, the positive impact of lower wages on the growth of value added in manufacturing took some quarters before materialising, as Figure 10 shows. Real wages began decreasing in 2009, but value added in the manufacturing sector started to pick up substantially only during the second half of 2010 and in 2011, which means it cannot be used to explain the recovery in real terms of growth of gross value added in the private sector observed since the second quarter of 2009.

If not from higher total wage bill or profits, the recovery in consumption observed since the second half of 2009 may then have been triggered by an increase in exports, as firms may have consumed more intermediate or capital goods to serve foreign markets. This has been typically an important channel in previous cases of growth in the short run after a substantial fiscal consolidation, but again it again does not seem to apply to the Latvian case. In Figure 11 we show the contribution to nominal GDP growth of different components of GDP and, at first sight, it may appear that the evolution of net exports contributed positively to growth in 2009, reducing GDP contraction by more than 10 per cent of GDP every quarter.

Figure 9

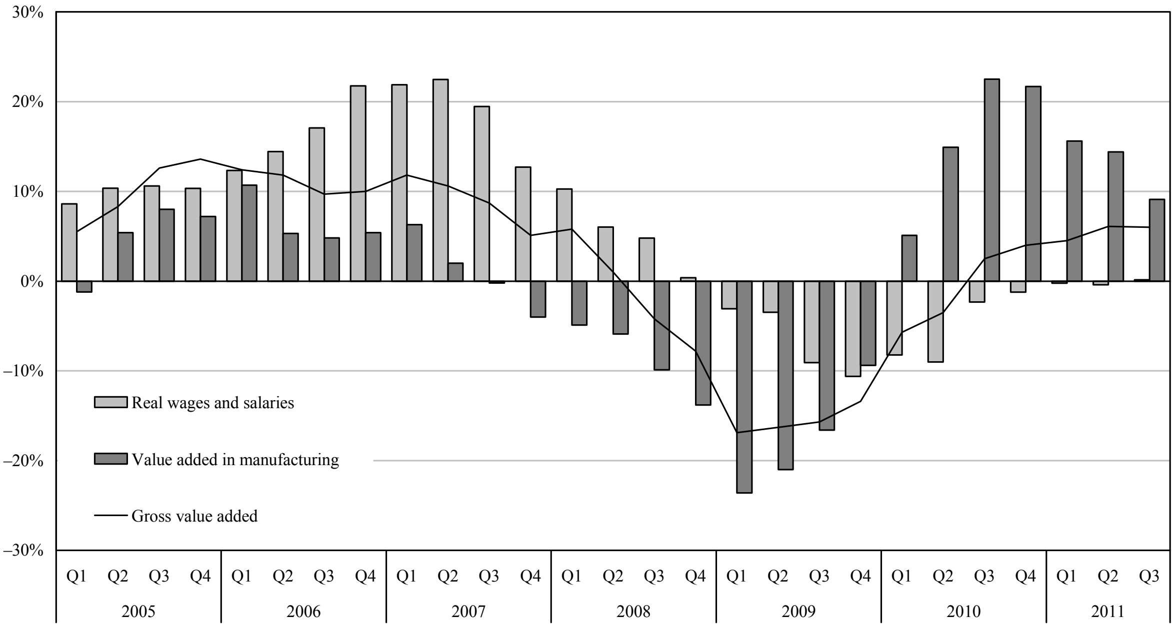
Quarterly Series of Private and Public Sector Wages *(annual changes)*, Profit's Share of GDP and Employment



Source: Commission Services.

Figure 10

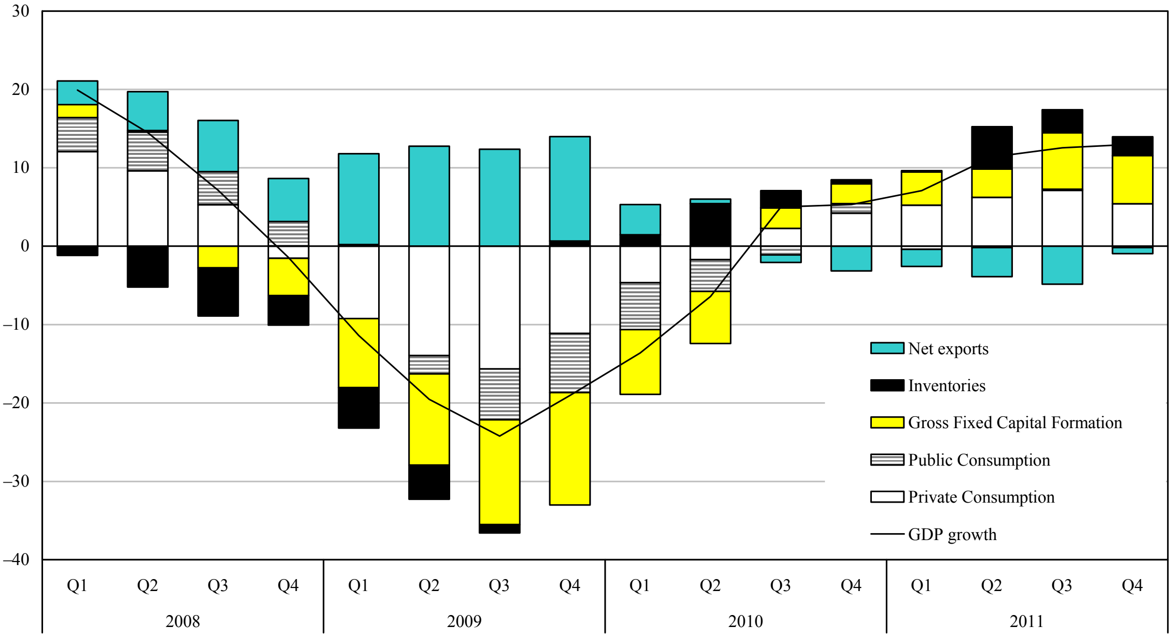
Quarterly Series of Real Wages, Value Added in Manufacturing and Gross Value Added *(annual changes)*



Source: Commission Services.

Figure 11

Quarterly Series of Contribution to Annual Nominal GDP Growth of All GDP Components



Source: Commission Services.

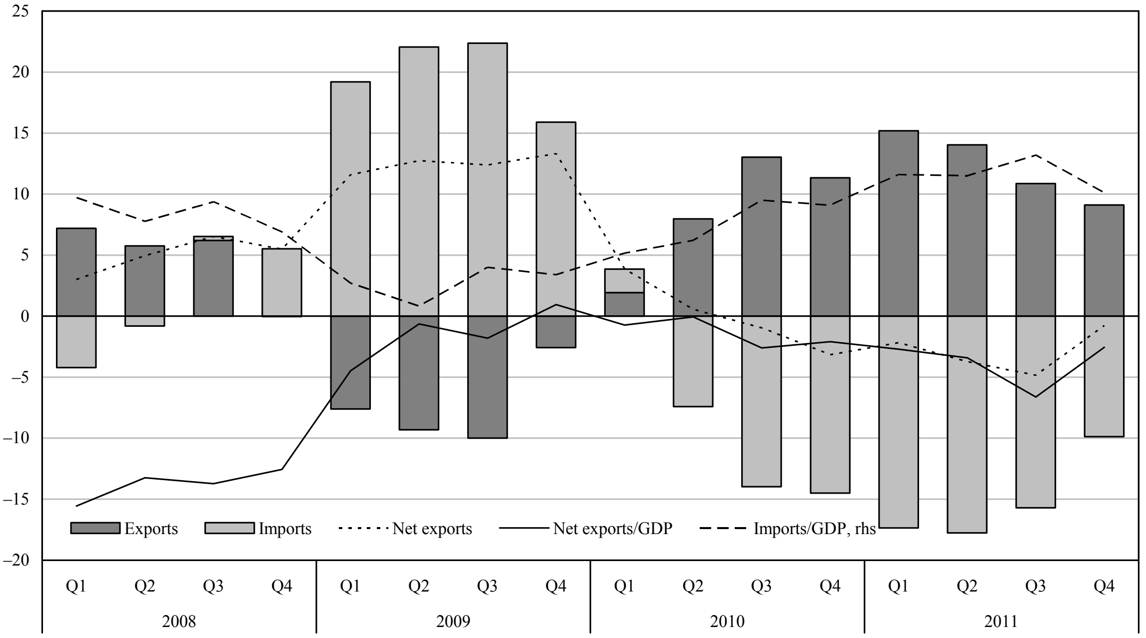
Unfortunately, a deeper observation of the dynamics behind the positive contribution of the external sector in the 2009 figures shows that exports contracted significantly and it was just an even greater contraction of imports that tilted the net trade balance on the positive side. This can be seen clearly from Figure 12, where the contribution to GDP growth of net exports is disaggregated into imports' and exports' contribution. In the second and third quarters of 2009 the contribution of imports' contraction to GDP growth was above 20 per cent of GDP, which accounts for a big share of the contemporaneous contraction in private consumption and gross fixed capital formation shown in Figure 11 (between 25 and 30 per cent of GDP). Indeed the ratio of import over total GDP (measured on the right axis of Figure 12) shrank from 60 to 40 per cent between 2008Q1 and 2009Q2. It is true that in 2009Q4 Latvian trade balance was positive for the first time in more than a decade, but it was only because between 2008Q1 and 2009Q4 total imports dropped by more than 1/3 and total exports by 1/5, so it would be fair to say that Latvian external adjustment happened despite and not thanks to external demand dynamics.

Summing up, all the available evidence point in the direction of suggesting that during the second half of 2009 competitiveness gains and external demand did not play a significant role in kick-starting Latvian economy. However, it should be noted that this outcome was probably driven by the extremely weak external demand due to the global spread of the financial crisis and could thus not be directly compared to previous episodes of export-driven expansionary fiscal consolidation happening during more favourable external conditions.

Still, even in the absence of external support, Latvian economy did start to recover as soon as consolidation kicked in, leaving as the only possible explanation a recovery of confidence. A clear sign of this can be seen in the financial sector, as the financial openness of the country allowed capitals to flow easily in and out of the country in response to policy action and confidence in the stability of the economy. As we can see in Figure 13, net flows of foreign direct investments and

Figure 12

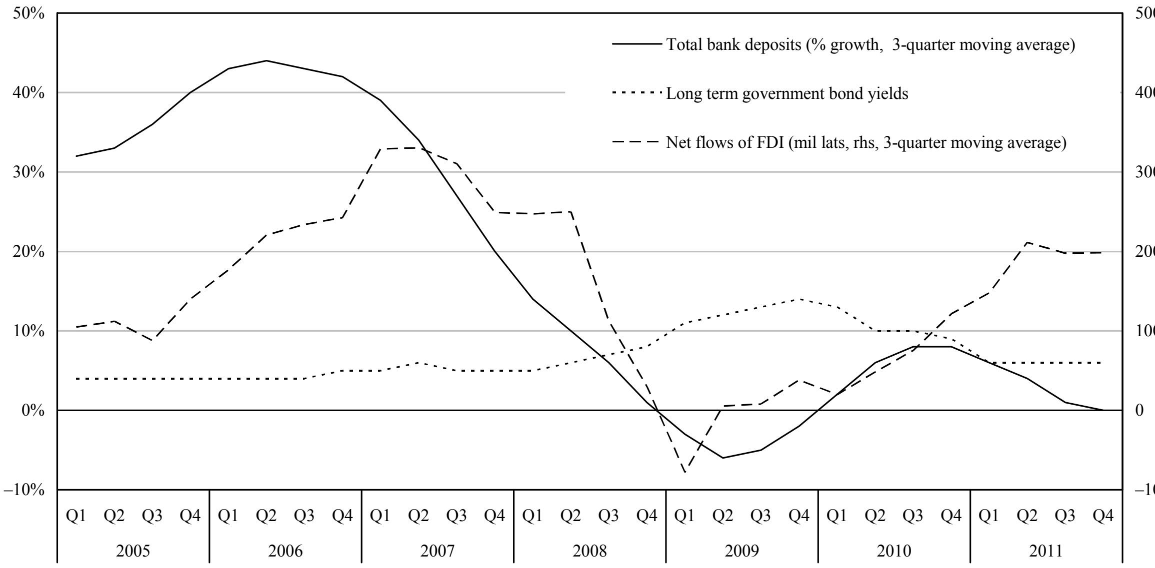
Quarterly Series of Annual Net Export's Contribution to GDP Split into Imports and Exports, and Imports/GDP Ratio in Percentage Points



Source: Commission Services.

Figure 13

Quarterly Evolution of Long-term Government Bond Yields, Total Bank Deposits and Net Flows of FDI



Source: Bank of Latvia.
Note: Total bank deposits are expressed in terms of millions of lats, total bank deposit annual growth is shown percentage points. Both series show 3-quarter moving averages.

residents' total bank deposits fled the country during the crisis quarters, but came back as soon as fiscal consolidation started. The series clearly mirror the investment and consumption series shown in Figure 8. As the consolidation measures kicked in, from the second half of 2009, it can be noticed that Latvian residents stopped withdrawing their savings from the banking system and foreigners started investing again in the country. This clearly shows how related are capital flows and foreign investments to the level of confidence in the country, which is in turn closely linked to the government action.

The general lesson we can draw from the impact of fiscal consolidation on the Latvian economy is that in a flexible and open economy a bold intervention by the government and the international community to restore confidence can trigger important non-Keynesian effects which may even completely offset standard Keynesian multipliers.

In addition, the immediate response of the confidence-related channels of non-Keynesian reaction and the lagged response of external competitiveness can have the additional advantage of resulting in a prolonged stimulus as a result of the two effects kicking in at different times. This feature may provide the government a comfortable period of economic growth after a crisis which can be used to enact the due structural reforms.

4.3 A tentative measure of non-Keynesian effects

As we observed in the previous sections, however, the short term negative effects of fiscal consolidation never fully materialised in the Latvian experience as the economy started recovering just as the bulk of the fiscal consolidation kicked in, from the second half of 2009. In order to give an idea of the unexpected linkages between economic growth and fiscal consolidation, we plot in Figure 14 the previously estimated Keynesian effects of fiscal consolidation on GDP against the evolution of real GDP in the quarters following the consolidation. Normalising GDP using the first quarter of consolidation, 2009Q1, and taking it as a baseline, we consider the percentage difference of each quarter from the baseline. Comparing the gap between the deviation of actual GDP from the baseline and the theoretical deviations that should have arisen from the fiscal multipliers of the measures, we can have a rough estimate of the magnitude of the non-Keynesian effects. It can be noted that real GDP contracted up to the third quarter of 2009, but then bounced back between the end of 2009 and 2010, at a time in which the Keynesian effects associated with the additional consolidation should have dragged it down. Even if we know that many additional factors not included in our simulation may have contributed to determine this gap, the difference between these two series point to the presence of significant non-Keynesian effects rising from 1 per cent of GDP in 2010Q1 to 7 per cent in 2011Q4.

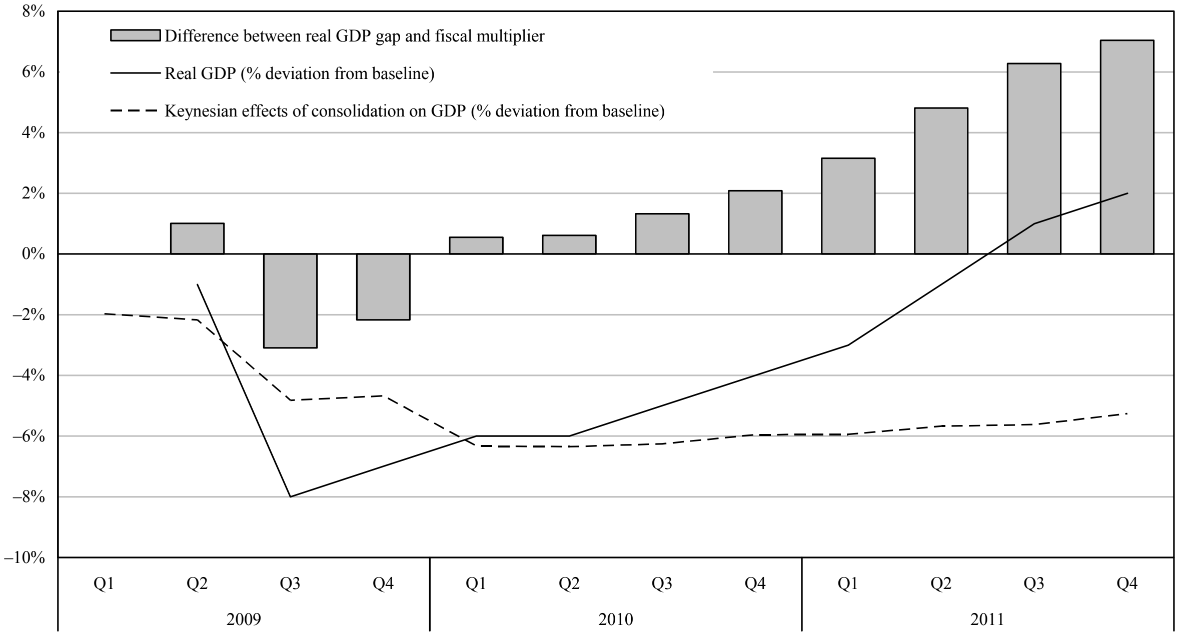
We should keep in mind that this is a rather conservative estimate, since in our QUEST simulation external demand was assumed to be stable, whereas in Figure 12 we have shown that exports dropped by 9-10 per cent of GDP in 2009Q2 and 2009Q3, even if the overall contribution of trade to GDP growth was positive due to a more than proportional contraction in imports.

Interestingly enough, the evolution of the consumer confidence indicator introduced in Section 3 follows closely our indicator of non-Keynesian impact of fiscal consolidation on the economy, as can be seen from Figure 15, this pointing to the relevance of the recovery of consumer confidence as a possible source of non-Keynesian effect.

In addition, we may notice that a similar improvement, starting from the second half of 2009 and consolidating in 2010, can be seen in the evolution of indicators of financial confidence such as the credit default swap (CDS) spreads and the interbank market rates, shown in Figure 16.

Figure 14

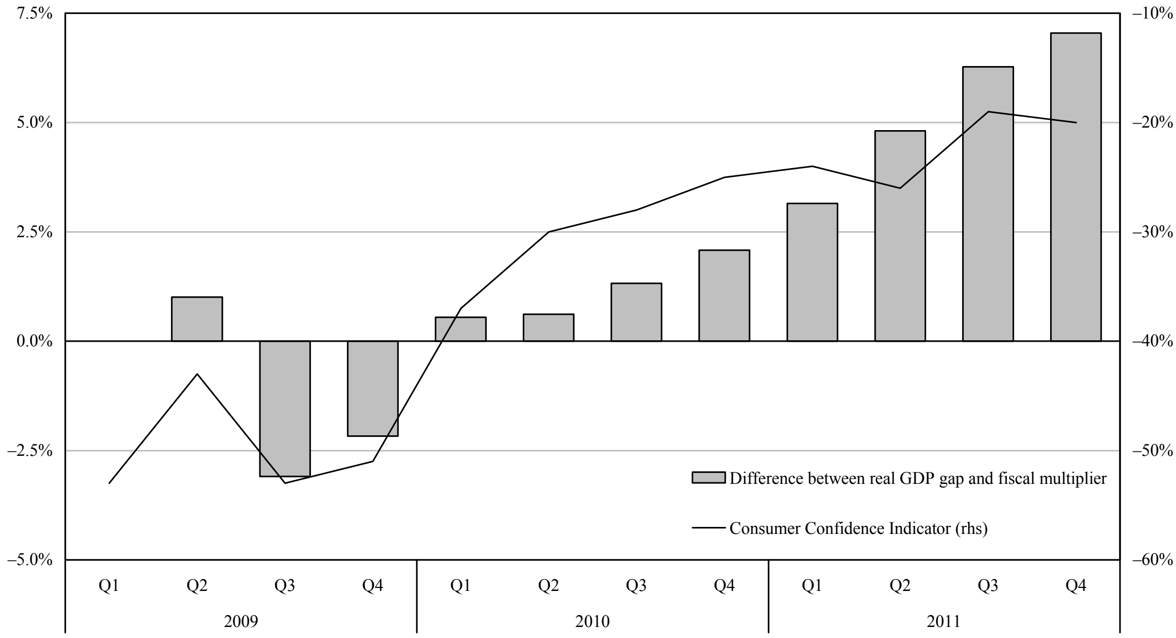
Difference Between Real GDP Gap and GDP Gap Simulated with QUEST II Considering the Actual Composition of Fiscal Consolidation, by Quarters



Source: Commission Services.

Figure 15

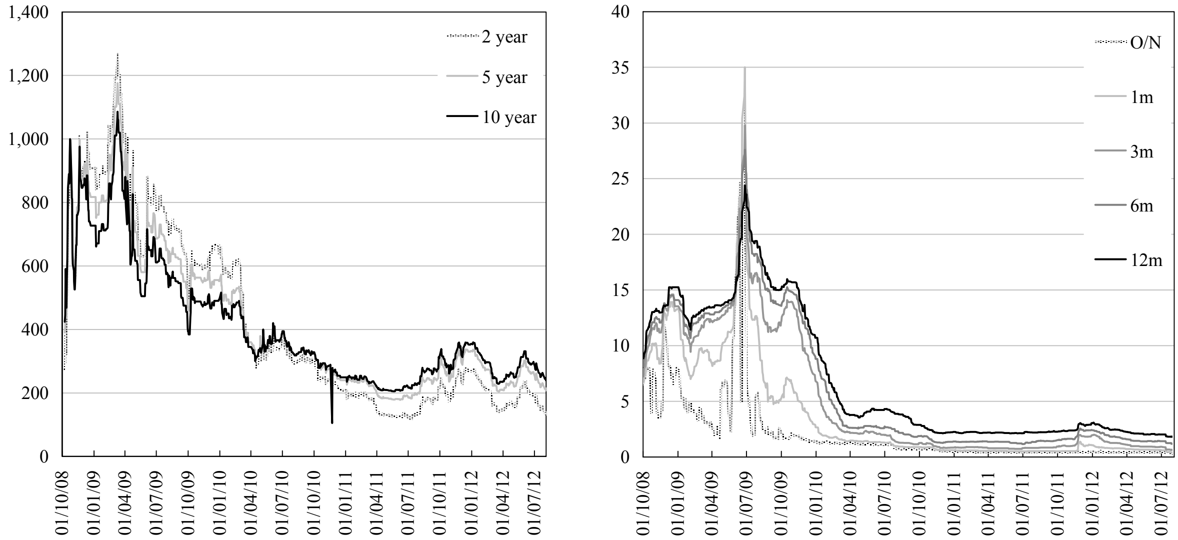
Quarterly Evolution of the Difference Between Real GDP Gap and GDP Gap Simulation, Plotted Against the Evolution of the Consumer Confidence Indicator



Source: Commission Services.

Figure 16

Evolution of Two Financial Confidence Indicators:
the CDS Spreads *(left panel)* **and the Interbank Market Rates (***right panel)*



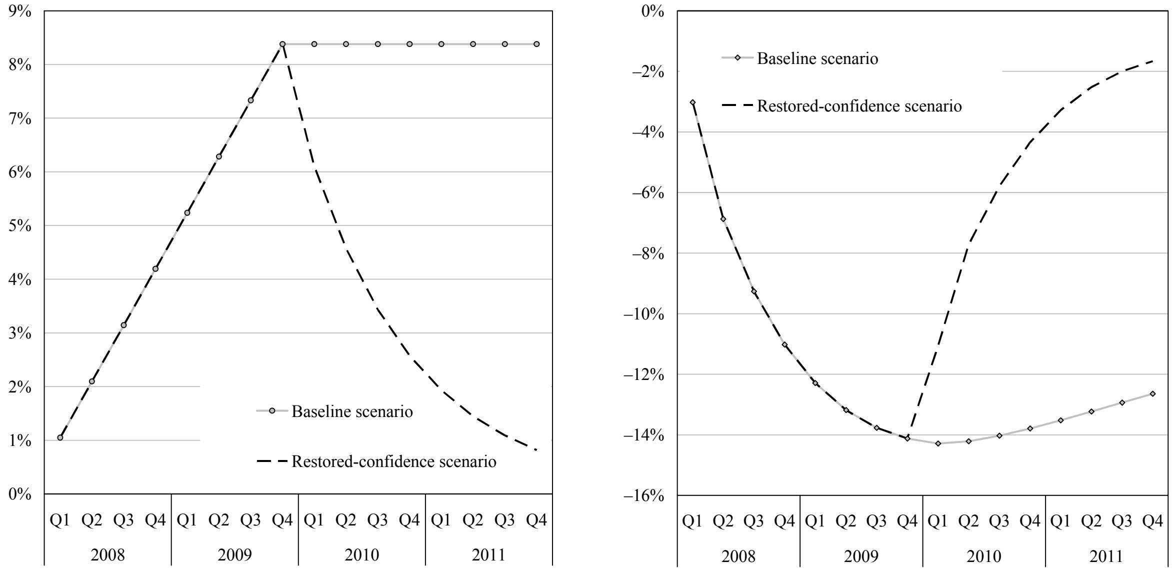
Source: Reuters EcoWin.
Note: CDS spread are expressed in basis points. Interbank market rates are RIGIBOR, fixing, in percentages.

Since also government bond yields and financial risk premia have been identified in the literature as sources of non-Keynesian effects, we further investigate their potential role in the Latvian case, turning again to a QUEST simulation and showing the results in Figure 17. The risk premium in the model drives a wedge between the domestic and the world interest rates and concerns domestic borrowing costs for each the households, the corporate and the public sector. As far as the small open economy is indebted to the rest of the world, this risk premium will also constitute a wealth transfer to external economies. The baseline scenario shows the large negative effect of a persistent annualised 800 basis point increase in the spread starting from 2008Q1. This roughly matches the pattern of government bond yields and CDS spreads for Latvia in 2008 and 2009 with the assumption that, absent the measures taken by the government in 2009, risk premia would have remained persistently high over the following years. The reversal scenario shows the effect of the drop in spreads back from 800 basis points to close to around 100 basis points by 2012. The sudden reversal has a positive effect on economic activity which converges back to its pre-2008 level relatively quickly following the reversal.

The reversal in the Latvian yields may arguably be linked to the firm fiscal consolidation measures undertaken by the government. As such, the above scenario underlines the likely pre-eminence of the financial channel in triggering the observed non-Keynesian effects. In other words, the consolidation measures helped bring back confidence in the financial markets and allowed Latvia to dispel the negative effects associated with the very high risk premia it was experiencing before the government took action. It is worth noting that the simulated size of the shock is rather significant, reaching 14 per cent of GDP at its peak. The link between the reduction in bond yields and recovery can be seen in Figure 18, where we plot the evolution of Latvia's real GDP (black dashed line) against the GDP trend simulated by the QUEST model (red solid line) and the bond yields shock. It seems reasonable to attribute part of the merits of the quick recovery to the normalisation of the risk premia, which allowed firms and consumers to gain a better access to the financial markets.

Figure 17

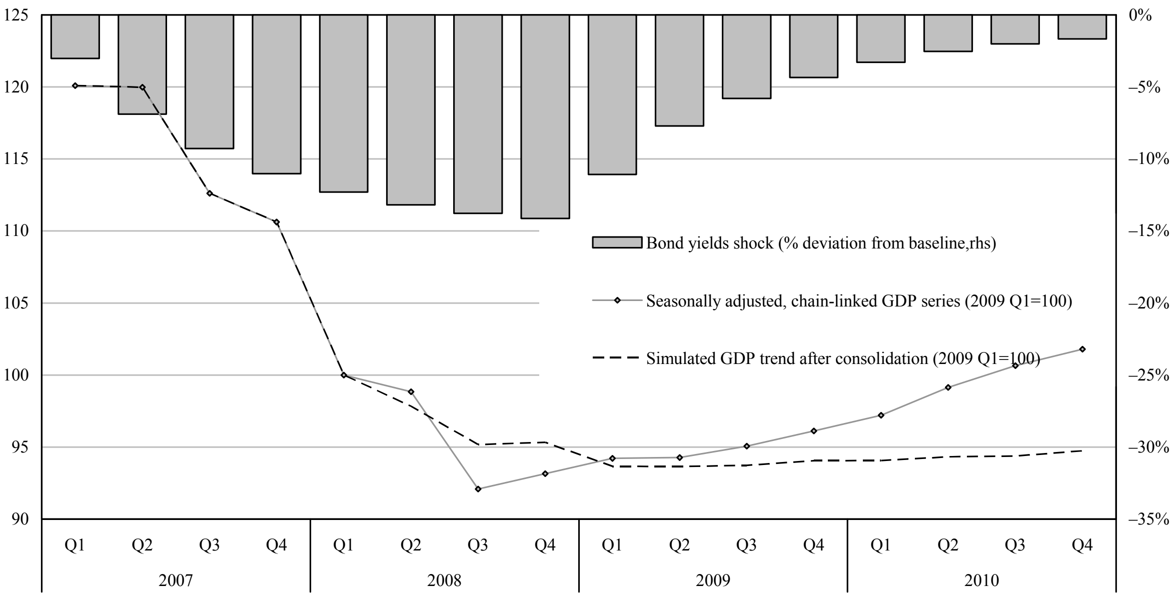
Impact on GDP ***(right panel)*** **of a Financial Confidence Shock of the Magnitude Experienced by Latvia During the Balance-of-payment Crisis, as Captured by the Long-term Government Bond Yields Spread** ***(left panel)***



Source: Commission Services.
Note: Real data until 2009Q4, long-term government bond yields spread being the deviation from the Latvian average in the previous 10 years, then QUEST simulation.

Figure 18

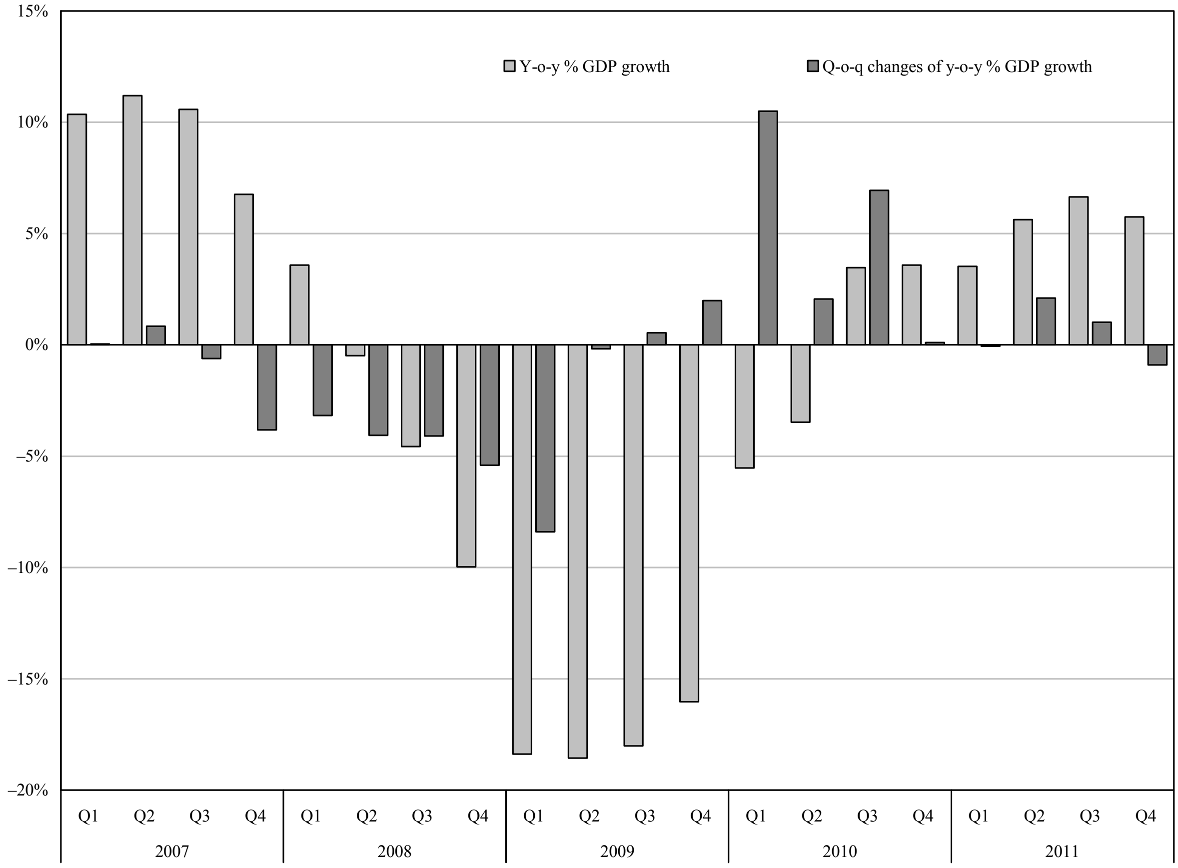
Simulated Impact on GDP of the Financial Confidence Shock and of the Impact of Fiscal Consolidation, Plotted Against the Actual Series of Real GDP



Source: Commission services. Note: QUEST simulation for the simulation for the GDP trend after consolidation starts after 2009Q1 and takes into account the impact of fiscal consolidation but not the impact of the financial confidence shock, thus plotted separately.

Figure 19

Annual GDP Growth, by Quarter, and Quarterly Changes in Annual GDP Growth



Source: Commission Services.

Since consumer and business confidence has often been identified as the main driver of non-Keynesian effects in the literature, this observation seems to reinforce our intuition that the particular characteristics of Latvian fiscal consolidation managed to offset the short-term negative impact of fiscal consolidation.

As a final remark, it is worth seeing how Latvian economic growth was affected by the consolidation measures not compared to our simulations, but in its own sake. As a complement to the GDP figures in level provided in Figure 7, Figure 19 shows the year-on-year GDP changes and their quarterly changes. It can be noticed that while the situation keeps on deteriorating for the entire 2007 and 2008, increasingly bad growth performances, the economy reacts to the austerity measures by first stabilising, in middle of 2009 and then rebounding strongly by the beginning of 2010, even if positive year-on-year changes could be observed only by the second half of 2010.

Summing up, our analysis suggest strongly that credible, bold, front-loaded and well-designed measures managed to convince Latvians and foreign investors, between 2009 and 2010, that the worst was over and the country was back again on a sustainable path. This renewed confidence in the country immediately alleviated the economic pain caused by prohibitive risk premia for government bonds and has triggered the equivalent of a cost-free economic stimulus to the economy when it was most needed.

5 What lessons from Latvia?

Latvia's experience represents a remarkable example of how fiscal consolidation should be undertaken to maximise long-term benefits and at the same time provide relief to the economy in the short term. The foundations of its success lied on the following essential elements:

- **Timing:** a rapid response is crucial when the economy and the budget are getting out of hands, but time is needed for surgical and meaningful action. It is therefore essential to have a large enough financial package and a long enough horizon to avoid across-the-board cuts;
- **Size:** when trends are wrong, everybody, including markets, must be impressed by the size of action. Going big can change mind sets and attitudes. Much of what has been done has been large from the beginning: wage adjustment, employment, reforms in key areas such as education, health and the organisation of the public administration;
- **Trust:** at the end, what drives the economy is the behaviour of agents. This is strongly affected by credibility of policies, but even more by the trust in the counterparts;
- **Country-specific analysis:** the adjustment of Latvia defied much of conventional economics. There must be courage to challenging some of its assertions, when new ground is being broken. Every economy is different at any given time. While there are similarities, one should not overlook key differences;
- **Prudence:** in devising an adjustment, one should not bank on uncertain benefits. Markets and observers have asymmetric reactions. Better results lead at best to a progressive increasing credibility. But any credibility can be quickly lost because of a small negative underperformance. A certain distance must thus always be kept from the edge;
- **Effective Communication:** effective communication is needed to spell out misinterpretation and to persuade actors that the policy objectives are achievable. Telling the "hard-truth", explaining what needs and can be done, reminding about the final objective, have been key elements of the Balance-of-Payments assistance programme that supported Latvia's fiscal consolidation.

6 Conclusions

The unprecedented fiscal consolidation efforts undertaken by Latvia in 2009 represent an ideal case study to have a fresh look at the short-term relation between fiscal policy and GDP growth. Especially on the expenditure side, the Latvian consolidation strategy was characterised by a careful design of measures, based on strategic plans rather than across-the-board cuts in several important areas. The bold, decisive, targeted and front-loaded nature of Latvian consolidation appear to have contributed to trigger non-Keynesian effects so relevant as to offset the standard negative Keynesian reaction to spending cut and tax hike (which were in themselves minimised by the growth-friendly composition of the consolidation). Government intervention and international lenders' guidance certainly halted a downward spiral and was accompanied by a sudden recovery in confidence which is likely to have prompted a quick rebound of consumption and investments in the private sector. With negative effects limited and positive ones kicking in in a sequential manner, this consolidation rapidly drove the Latvian economy on a sustained growth path.

There could be several conditions that allowed the consolidation to work so well. First of all, the fiscal sector in Latvia over-expanded so rapidly in boom years preceding the crisis that it could not pose much resistance to its downsizing. Second, even though it grew rapidly before the crisis, the size of the public sector in Latvia and in the Baltics in general has historically been smaller than in the rest of the European Union. This implies that the impact of fiscal multipliers is more limited than in other European countries, as more scope is left for the private sector's behaviour to determine the ultimate effects on growth. Finally, the economic contraction and loss of confidence were so serious at the onset of the crisis years that they could have amplified the effects of the following rebound.

It is also worth mentioning that the availability of EU funds may have offset part of the cuts in public expenditure. A study commissioned by the Latvian Ministry of Finance (SSE, 2011) estimated that the impact EU funds on the Latvian economy amounted to 4 per cent of real GDP in 2009 and 5.2 per cent in 2010. Even if that is not a significant increase with respect to 2008, when the estimated impact on the economy was 3.9 per cent of GDP, it may be argued that the crowding out effects of EU funds should be lower in a phase of economic contraction. Credit should however be given to the Latvian authorities and to the Commission for having secured the co-funding of such expenditure during the consolidation, which was achieved by higher cuts to other current expenditures.

All in all, important lessons that can be drawn from the Latvian experience. Good judgements on country-specific issues, right timing and sufficient size of intervention were key elements for Latvian success, but for their potential benefits to be fully tapped, they had to be accompanied by mutual trust across decision makers, prudence and effective communication. It was this particular combination of features that allowed the consolidation measures to restore confidence and significantly offset the possible negative impacts of consolidation on the economy. Latvia showed that the trade-off between short-term pain and long-term gain can be avoided if intervention is sufficiently well designed.

REFERENCES

Afonso, A. (2010), "Expansionary Fiscal Consolidations in Europe: New Evidence", *Applied Economics Letters*, Taylor and Francis Journals, Vol. 17, No. 2.

Alesina, A. and S. Ardagna (1998), "Tales of Fiscal Adjustment", *Economic Policy*, CEPR, CES, MSH, Vol. 13, No. 27.

Åslund, A. and V. Dombrovskis (2011), "How Latvia Came Through the Financial Crisis", Peterson Institute for International Economics, May 2011.

Coenen, G., C.J. Erceg, C. Freedman, D. Furceri, M. Kumhof, R. Lalonde, D. Laxton, J. Linde, A. Mourougane, D. Muir, S. Mursula, C. de Resende, J. Roberts, W. Roeger, S. Snudden, M. Trabandt and J. in't Veld (2012), "Effects of Fiscal Stimulus in Structural Models", *American Economic Journal: Macroeconomics*, American Economic Association, Vol. 4, No. 1.

Cogan, J.F., T. Cwik, J.B. Taylor and V. Wieland (2010), "New Keynesian Versus Old Keynesian Government Spending Multipliers", *Journal of Economic Dynamics and Control*, Elsevier, Vol. 34, No. 3.

European Commission (2010), "Cross Country Study: Economic Policy Challenges in the Baltics", *European Economy*, Occasional Paper, No. 58, Directorate-General for Economic and Financial Affairs.

European Commission (2011a), "European Union Balance of Payments Assistance to Latvia", *EC Staff Report from the 5th Review Mission to Riga*, Directorate-General for Economic and Financial Affairs.

European Commission (2011b), "The Economic Adjustment Programme for Ireland", *European Economy*, Occasional Paper, No. 76, Directorate-General for Economic and Financial Affairs.

European Commission (2012), *Taxation Trends in the European Union*, Publications Office of the European Union.

Giavazzi, F., T. Jappelli and M. Pagano (2000), "Searching for Non-linear Effects of Fiscal Policy: Evidence from Industrial and Developing Countries", *European Economic Review*, Elsevier, Vol. 44, No. 7.

Giavazzi, F. and M. Pagano (1990), "Can Severe Fiscal Contractions Be Expansionary? Tales of Two Small European Countries", in NBER (ed.), *NBER Macroeconomics Annual 1990*, Vol. 5, pp. 75-122.

Giorno, C., P. Richardson, D. Roseveare and P. van den Noord (1995), "Estimating Potential Output, Output Gaps and Structural Budget Balances", OECD, Economics Department, Working Paper, No. 152.

Girouard, N. and C. André (2005), "Measuring Cyclically-adjusted Budget Balances for OECD Countries", OECD, Economics Department, Working Paper, No. 434.

Giudice, G., A. Turrini and J. in't Veld (2007), "Non-Keynesian Fiscal Adjustments? A Close Look at Expansionary Fiscal Consolidations in the EU", *Open Economies Review*, Springer, Vol. 18, No. 5.

Guajardo, J., D. Leigh and A. Pescatori (2011), "Expansionary Austerity: New International Evidence", IMF, Working Paper, No. WP/11/158.

IMF (2010), *World Economic Outlook*, October 2010, Chapter 3.

Lendvai, J., L. Moulin and A. Turrini (2011), "From CAB to CAAB? Correcting Indicators of Structural Fiscal Positions for Current Account Imbalances", *European Economy*, Economic Paper, No. 442, April.

Lendvai, J. and W. Roeger (2010), "External Imbalances in the Baltics – Catching Up or Imbalances", *European Economy*, Economic Paper, No. 398, January.

Ratto, M., W. Roeger and J. in't Veld (2009), "QUEST III: An Estimated Open-economy DSGE Model of the Euro Area with Fiscal and Monetary Policy", *Economic Modelling*, Elsevier, Vol. 26, No. 1.

Republic of Latvia (2009a), *Convergence Programme of the Republic of Latvia 2008-2011*, January.

—————— (2009b), *Latvia: Letter of Intent*, 27 June.

—————— (2012), *Convergence Programme of the Republic of Latvia 2012-2015*, April.

Rzonca, A. and P. Cizkowicz (2005), "Non-Keynesian Effects of Fiscal Contraction in New Member States", European Central Bank, Working Paper, No. 519.

Sancak, C., R. Velloso and J. Xing (2010), "Tax Revenue Response to the Business Cycle", IMF, Working Paper, No. WP/10/71.

Stockholm School of Economics – Riga (2011), "Evaluation of the Impact of EU Funds on the Economy of Latvia", available at: http://www.esfondi.lv/upload/Petijumi_un_izvertejumi/2ndStageReportEnglishUltimateFinal15Dec.docx

World Bank Human Development Sector Unit (2010), *Latvia – From Exuberance to Prudence, A Public Expenditure Review of Government Administration and the Social Sectors*, September.

COMMENTS ON SESSION 1
THE SHORT-TERM IMPACT OF FISCAL POLICY

Jan Babecký[*]

Comments on "Fiscal Multipliers and Fiscal Consolidations" by Ray Barrell, Dawn Holland and Ian Hurst and "Fiscal Multipliers: How Much Bang for the Buck?" by Glenn Follette and Byron Lutz

Let me start by thanking the organisers for inviting me and giving an opportunity to discuss these two papers. The first paper, "Fiscal multipliers and Fiscal Consolidations" by Ray Barrell, Dawn Holland and Ian Hurst presents an empirical evidence on fiscal multipliers based on simulations using the National Institute Global Econometric Model (NiGEM). An assessment of fiscal consolidation is performed for 18 OECD countries, focusing on actual fiscal programs for 2010-12 and on the expected fiscal actions for 201-20. A series of NiGEM simulations is used covering alternative horizons ranging from 2006-11 to 2015-27. An important feature of the authors' approach is accounting for forward-looking behaviour of financial matters, via an implicit incorporation of the role of expectations. The key result is that a tighter monetary policy reduces output growth in the short run but – due to a lower debt stock – contributes to (sustainable) output growth in the long run.

The second paper, "Fiscal multipliers: How Much Bang for the Buck?" by Glenn Follette and Byron Lutz, presents a narrative evidence on fiscal multipliers for the U.S., based on the survey of the empirical literature and the FRB staff's macroeconomic model. Assessment of the effectiveness of the U.S. fiscal policy in stimulating aggregate demand is conducted for a series of policy measures implemented between 1953 and 2010, including the 2008/2010 stimulus package. The main result is that the increases in the deficit helped boosting demand. Nevertheless direct multipliers were less than one, largely due to a reliance on tax cuts.

Let me comment on issues common to both papers. First, one can observe a large variation in the reported multipliers. The first paper attempts to relate the differences in multipliers to country-specific features (e.g., country size, degree of openness, and the degree of dependence on consumption and current income) and such variables as labour market flexibility and path-through of policies (e.g., a rise in VAT) into prices. The second paper shows that while direct multipliers in the U.S. are relatively low (by international standards), still there is an important variation of multipliers over time.

What can be learnt from such a variation in multipliers? Let me recall one relevant statement by Leeper (2010) regarding the variety of empirical estimates of fiscal multipliers: "One clear message emerges from (this) vast literature: estimates of multipliers are all over the map, providing empirical support for virtually any policy conclusion. The diversity of findings, often based on the same U.S. time series data, highlights the difficulties in obtaining reliable estimates of fiscal effects and points to the need for systematic analyses that confront fiscal policy's complexities." (p. 19).

It would be worth examining the underlying reasons of such diversity in multipliers, for example to investigate the role of methods (e.g., sample size, econometric technique, time period covered), the role of measurement of multipliers (expenditure/spending, short-/long-term, measures of dynamics) and the role of econometric specification (*i.e.*, the control variables) and the quality of studies used. Given the topic of this year's workshop – "Fiscal Policy and Growth" – let me

---

[*] Czech National Bank.

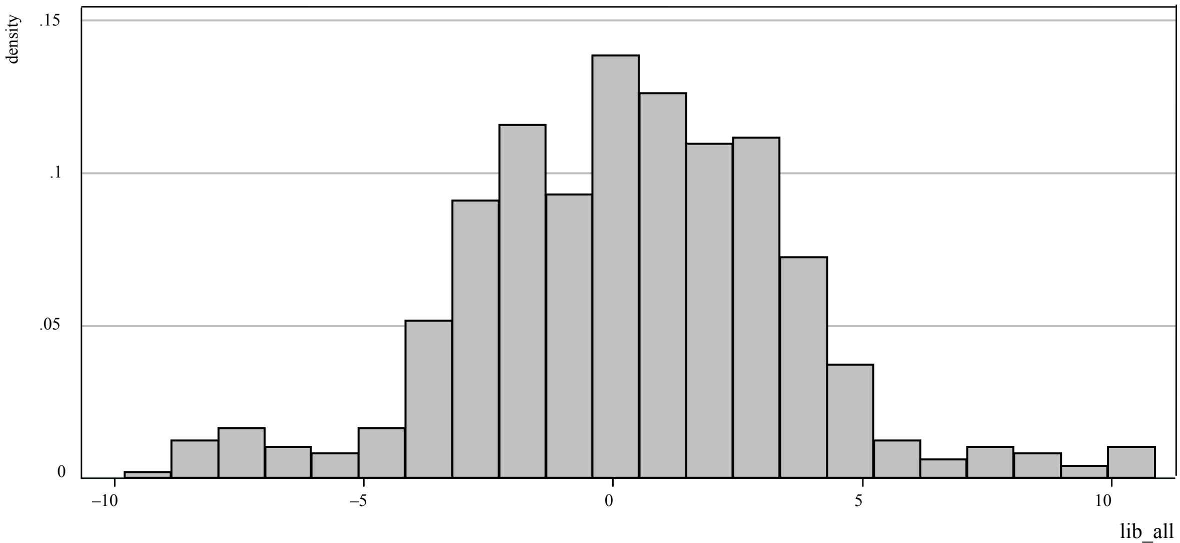
illustrate whether some lessons could be taken from the existing literature on the issue of "Structural Reform and Growth".

Based on the review of about fifty studies for transition economies, Babecký and Campos (2011) ask similar types of questions to those arising in the fiscal policy context: what is the impact of reform on growth? What are the short-run costs versus long-run benefits? A summary of the reform-growth nexus could be illustrated on the following three figures. Overall, considering the pool of available estimates of the effect of structural reform on growth (more precisely, *t*-statistics of the estimates, in order to allow for a comparability across studies which use different units of measurement, different specifications, etc.), Figure 1 shows the variety of estimates ranging from negative to positive ones, with the average effect of reform on growth being close to zero.

If one separates short-term and long-term effects, the histograms change. For example, in the short-run, the link between reforms and growth becomes negative (Figure 2), suggesting that the reforms are characterized by non-negligible real costs. These costs are offset over time, when benefits from implementing structural reform become materializing (Figure 3). Nevertheless, both Figure 2 and Figure 3 still demonstrate a large variety of the estimates.

Figure 1

Link Between Structural Reform and Growth: Overall Effect

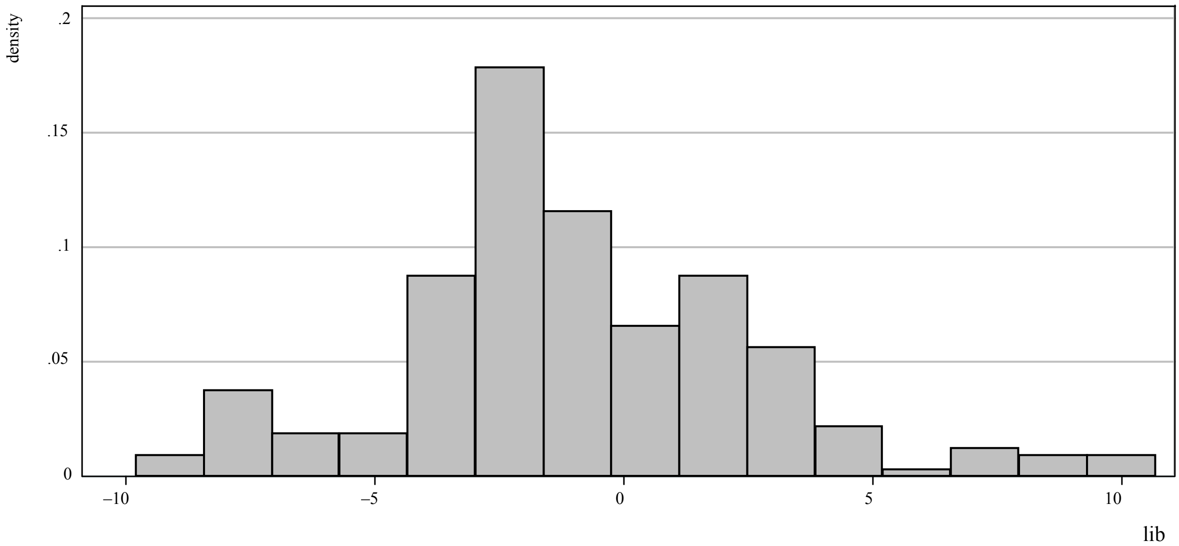


Note: Histogram of the *t*-statistics of coefficients of structural reforms on economic growth: 515 coefficients from the 46 papers (Figure 1 in Babecký and Campos, 2011).

This variation could be further explored employing the methods of quantitative review of literature – Meta-Regression Analysis (MRA), along with the above mentioned lines method-measurement-specification. Thus, it might be worth applying the MRA for a similar type of questions addressed in the fiscal policy – growth literature. Rusnak (2011) and Gechert and Will (2012) are two perspective applications of MRA to government spending multipliers. Apart from understanding the reasons which are behind differences in the estimate of multipliers, MRA can also help identifying the "best-practice" specification.

Figure 2

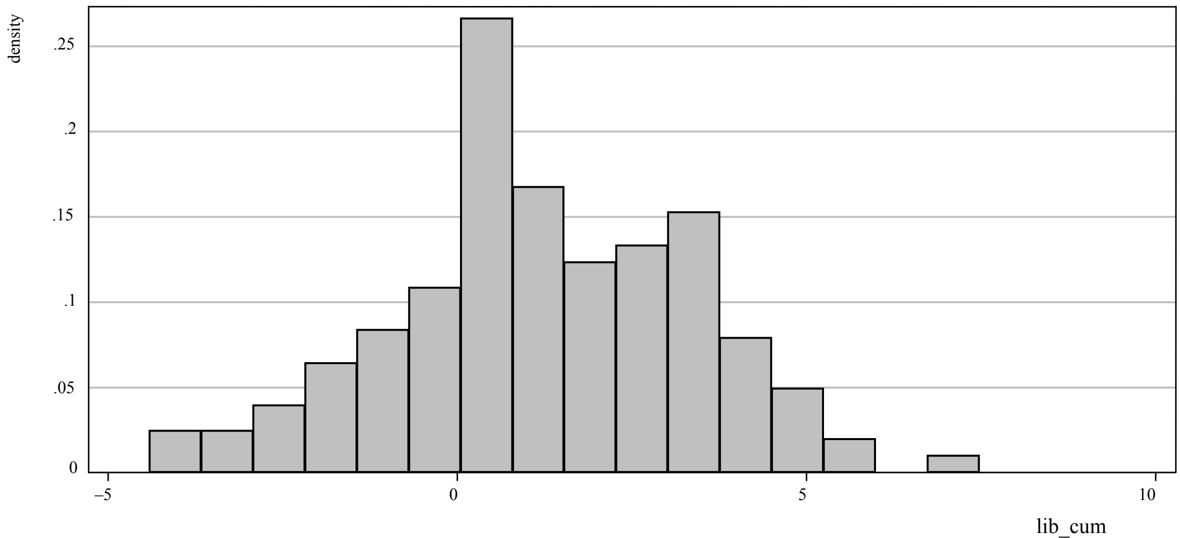
Link Between Structural Reform and Growth: Short-run Effect



Note: Histogram of the t-statistics of coefficients of contemporaneous structural reforms on economic growth: 234 coefficients (Figure 2 in Babecký and Campos, 2011).

Figure 3

Link Between Structural Reform and Growth: Long-run Effect



Note: Histogram of the t-statistics of coefficients of cumulative effect of structural reforms on economic growth: 276 coefficients (Figure 3 in Babecký and Campos, 2011).

Further directions for future research could include such issues as (i) the role of debt sustainability expectations (current analysis is largely done under assumption of constant risk premia); (ii) the impact of consolidation on risk premia; and (iii) fiscal stress testing (how changes in output growth would affect public finances).

REFERENCES

Babecký, J. and N.F. Campos (2011), "Does Reform Work? An Econometric Survey of the Reform-Growth Puzzle", *Journal of Comparative Economics*, Vol. 39, No. 2, pp. 140-58.

Gechert, S. and H. Will (2012), "Fiscal Multipliers: A Meta Regression Analysis", Macroeconomic Policy Institute (IMK), Working Paper, No. 97/2012. Available at: http://www.boeckler.de/pdf/p_imk_wp_97_2012.pdf

Leeper, E.M. (2010), "Monetary Science, Fiscal Alchemy", NBER, Working Paper, No. 16510, Available at: http://www.nber.org/papers/w16510.pdf

Rusnak, M. (2011), "Why Do Government Spending Multipliers Differ? A Meta-Analysis", paper presented at the 2011 *Meta-Analysis of Economics Research Network Colloquium*, Cambridge. Available at: http://www.hendrix.edu/maer-network/default.aspx?id=40672

COMMENTS ON SESSION 1
THE SHORT-TERM IMPACT OF FISCAL POLICY

FISCAL POLICY AND ITS MACROECONOMIC EFFECTS – A DISCUSSION

Adi Brender[*]

1 Introduction and framework

Caprioli and Momigliano's (C&M) paper: "The Macroeconomic Effects of Fiscal Policy Shocks during Good and Bad Times" examines the macroeconomic effects of fiscal policy changes using an extended VAR model. The paper provides a unique analysis and useful insights but there are avenues through which it could be improved.

The key background policy question of the paper is whether fiscal policy can and should be used to smooth the business cycle. To answer this question it is important to identify whether fiscal policy is effective and under what conditions, what are the potential "side effects" of using fiscal policy, and are there alternative measures (e.g., monetary policy) that can reach the same results more effectively or with less side effects. A related issue, particularly relevant in the current European economic setting, is to identify the optimal consolidation pace in various states of the economy.

Table 1

Main Characteristics of Caprioli and Momigliano's Model

Model Characteristic	Caprioli & Momigliano
Framework	VAR
Fiscal policy effect on growth	non-linear – binary
Fiscal policy effect on unemployment	not discussed
Monetary policy	Short-term interest rates included
Present/future tradeoff	No
Public debt	Debt excluded
Sample	Italian time series
Data	Quarterly
Constraints on policy	None
Fiscal measure	Central government consumption

Table 1 depicts the model's main characteristics. A contribution that stands out is the non-linearity of fiscal policy effects introduced by C&M and its interaction with monetary policy. On the down side a key caveat of the model for policy interpretation purposes is that it does not impose an intertemporal constraint on fiscal policy.

* Bank of Israel.

Table 2

Analytical Division of Data

	Status of the Economy	
Fiscal policy		
	Recession	**Close to full Employment**
Consolidation	+	
Expansion	+	

A useful way to characterize the policy issues and economic environment on which the paper focuses is presented in Table 2. C&M use data on fiscal measures in all the cells of the table, but focus their analysis on policy during recessions; that is the left side of the table. It could be useful if they, as well as other papers that analyze the macroeconomic effects of fiscal policy, provided a breakdown of the data points in the sample according to the classification in this table because theory suggests potential differences between the cells. Otherwise, it is not clear that there are sufficient observations in each cell to allow drawing policy conclusions about that policy environment.

2 Specific comments

The main point of C&M's paper is clear, analytically robust, and supported by the empirical analysis: the fiscal multiplier depends on the state of the economy. The analysis is conducted on Italian quarterly data over 30 years, and finds that fiscal shocks have a positive effect on private GDP for several quarters in recessions and no effect in expansions. They also show that part of this effect is due to interest rates that rise in response to fiscal expansions in growth periods but not in recessions.

An important condition for the sustainability of policies, however, is that the average impact of fiscal shocks on the economy would be close to zero. If the average is positive, policy-makers may be tempted to use fiscal policy to constantly inject stimulus into the economy. As mentioned repeatedly in the literature, without a non-linear effect of the debt to GDP ratio on the growth rate or welfare, there is no consistent optimal policy. Hence the debt to GDP ratio explodes. In the case of C&M, the technical result is that the effect of a fiscal shock on private GDP is positive and fades after a few quarters in recessions, and is insignificant in expansions. Accordingly, when one adds the impact of the fiscal stimulus itself on GDP, the effect is always positive and there is no "cost" to expanding the deficit. For a policy-maker this implies that whenever an "old" shock fades it is time to boost the economy with another stimulus, leading to an unsustainable policy. It would be worthwhile to introduce the appropriate constraints in the model, e.g., in the form of an effect on long-term growth, yields and risk premiums, to wrap the analytical framework of the model. Using a continuous non-linear model, rather than a binary "two states of the world" approach, may generate the required non-linearity. I should stress that this comment does not relate to the empirical results of the paper – which do not show such persistence of expansions – but to the analytical interpretation of the results.

A useful feature of this paper is the explicit introduction of short-term interest rates into the model. This allows to examine potential interactions between monetary and fiscal policies, and to experiment with alternative policy options that tradeoff between the two. One question in this context, however, is whether the interaction between fiscal policy and interest rates is still relevant for a single country in the Euro zone. The authors could provide some insights on this by allowing for different effects of this relationship in the pre-Euro and Euro periods. Another question that merits more attention is whether the response of short-term interest rates reflects only monetary policy, or also the sentiment of investors, consumers and financial markets – due to the different signaling value of fiscal shocks in the various environments represented by the cells of Table 2.

A more technical point, but conceptually important, is the choice of the indicator for the state of the economy. C&M use the average growth over several quarters as the indicator from which they derive the classification of the state of the economy to recession or expansion. Theory, however, is more focused on "stock" variables such as the output gap or capacity utilization. This feature seems to be important when one examines the estimation results which derive the state of the economy variable from capacity utilization; in that case there is no significant effect of fiscal shocks on private GDP in either state of the economy. The choice of which variable is used to characterize the state of the economy is particularly important in periods like the current one where a big drop in GDP almost 4 years ago had been followed by growth, but not one that was sufficient to fully close the output gap.

Finally, a useful extension of the analysis would be to examine whether the magnitude of the effects of positive and negative fiscal shocks is similar during recession periods. In the current period, where strong incentives exist for both fiscal consolidation and stimulus, such an analysis may provide important insights to policy-makers.

3 Conclusion and potential extensions

The C&M paper offers useful insights on the macroeconomic effects of fiscal policy. Specifically, incorporating the role of monetary policy response to fiscal shocks offers a potential avenue for evaluating trade-offs between alternative policy instruments. Combining these with some form of a welfare function that guides policy makers and introducing an explicit cost to fiscal expansions would facilitate a broader picture of these tradeoffs and of the considerations in designing fiscal policy.

Of particular importance in setting such future frameworks would be to internalize some insights from our accumulated experience with fiscal policy. Specifically, it seems quite clear that in the absence of crises or external incentives political leaders almost never find a "good time to cut". Hence, monetary policy may be a preferable instrument for counter-cyclical purposes to the extent possible, or until the "liquidity trap" is approached. Such an analysis may also highlight some of the costs of large monetary unions, such as the Euro zone, where country-specific monetary policy is not available. Adding the required features for such an analysis to a fiscal framework would make models more relevant for genuine policy analysis. The paper discussed here is a useful contribution in this direction and such extensions may make it and even greater one.

COMMENTS ON SESSION 1
THE SHORT-TERM IMPACT OF FISCAL POLICY

Walpurga Köhler-Töglhofer[*]

First of all, I would like to thank our host for the excellent organisation of this event and the lavish supply of food for thought. As in past years, the variety of insights presented and the breadth of items discussed has provided us all with an intellectually enriching atmosphere. I have been invited to discuss two papers from the first session: "Fiscal Consolidation in the Midst of the Crisis: Lessons from Latvia", by Francesco Di Comite, Gabriele Giudice, Julia Lendvai and Ingrid Toming, and "Fiscal Adjustment, Job Creation, Job Destruction" by Alfonso Arpaia and Alessandro Turrini. However, I will focus primarily on just one of them, namely Di Comite *et al.*'s paper about the short-term impact on growth of the huge fiscal consolidation package that was implemented in Latvia in mid-2009.

The first session of this workshop revisited the long-standing discussions about the short-term impact of fiscal policy on growth. In our most recent discussion in 2010, we dealt with the issue of whether EU-coordinated fiscal stimulus packages (alongside an expansive monetary policy) would have short- and medium-term growth-enhancing effects. This time, however, the discussion centred on the growth impact of indispensable fiscal adjustment measures in response to – in some cases, exceedingly – worrisome fiscal developments. Budget deficits and government debts have soared in nearly all advanced countries on account of the economic crisis and unprecedented fiscal measures taken to limit demand shortfalls and support the financial sector. The rapid worsening of public finances in some countries, notably Latvia, Hungary, Greece, Ireland and Portugal, has resulted in rapidly rising financing costs, as investors have lost confidence in the countries' ability to service their debt. In many cases, international and/or supranational financial assistance measures have become necessary. As investors' trust has waned, fiscal policymakers have been faced with the enormous challenge of finding ways to mitigate the problems.

What can be done in the face of unsustainable deficit and debt developments and rapidly rising financing costs, coupled with a weak macroeconomic environment at risk of falling back into recession, is "a matter of bitter controversy" (see Perotti, 2011) – as has also been observed in recent debates about finding effective solutions to the European debt crisis.

According to traditional Keynesian views, fiscal adjustment measures typically entail short-term costs in terms of economic growth and rising unemployment – a view that recent literature on the topic tends to confirm (see also IMF, 2010). According to this paradigm, the aftermath of a recession is the worst time to start fiscal consolidation. So what then is the proper fiscal policy reaction if an undue delay of indispensable fiscal adjustments would ultimately give rise to even greater adjustment costs, as the government debt accumulated in the interim would necessitate an even greater fiscal correction later on? Above all, what are the short-term costs of fiscal consolidations pursued in a credible and consistent manner in cases where the starting position of a country is particularly precarious or when financing conditions on the markets have already become prohibitive? Could not properly designed fiscal adjustment measures in such precarious circumstances pave the way back to growth?

[*] Oesterreichische Nationalbank (OeNB), Deputy Head of the Economic Analysis Division.
E-mail: walpurga.koehler-toeglhofer@oenb.at
The views expressed here are those of the author and do not necessarily reflect the position of the OeNB.

Some economists – based on arguments derived from the literature on non-Keynesian fiscal effects and on "successful fiscal consolidation" – have proposed that governments should rapidly curb deficits, even if the country concerned has not yet fully recovered. If done properly, namely by reducing spending rather than by increasing taxes, and perhaps also with the support of necessary growth-stimulating structural reforms, fiscal adjustments would not be harmful; on the contrary, they might even result in a boost to GDP – although it should be stressed that there is "scarce" empirical evidence to support this view.

Di Comite *et al.* found that non-Keynesian fiscal effects have emerged in Latvia. They investigated the economic impact of the huge fiscal consolidation package that Latvia had had to implement in 2009 and called for lessons to be learned from that experience. The global economic and financial crisis had hit Latvia comparatively hard. Like Estonia and Lithuania, Latvia had become increasingly vulnerable already before the crisis due to a gradually overheating economy caused by huge booms in credit, construction and housing, and consumption, accompanied by high or steadily rising current account deficits and external debts. In addition, the high private sector net external debt was held in foreign currencies.[1] Even before the onset of the crisis, it was already clear that these imbalances would have to be corrected; however, the start of the crisis simply amplified the magnitude of correction. When the crisis unfolded, Latvia experienced the sharpest GDP contradiction of all countries worldwide. The country even had to ask for international help to stave off possible insolvency. However, Latvia managed to resist giving up its fixed exchange rate peg to the euro.

In the middle of 2009, the newly elected Latvian government implemented a second, truly sizeable consolidation package that amounted to about 15 per cent of GDP, with the bulk of the measures coming into effect immediately. The package comprised drastic expenditure cuts, including painful nominal wage cuts in both the public and private sectors. But after a few quarters, the Latvian economy had started to rebound slowly.

Growth implications of fiscal consolidations are dependent on the size and, in particular, the composition of the consolidation measures; the speed with which measures are implemented is probably also a key factor. Thus looking first at the composition of the fiscal adjustment, I would agree with Di Comite *et al.* that "the (growth-) favourable composition of the adjustment in Latvia has contained the negative effects of Keynesian fiscal multipliers".

According to the study, the lessons to draw from this case are:

1) with respect to timing, that "a rapid response is crucial",
2) with respect to size, that "going big can set mindsets and attitudes",
3) with respect to composition, that it is necessary to "do it in a growth-friendly way with a priority on spending cuts",
4) that "credible policies" are decisive,
5) and that "prudence and effective communication" are also of the essence.

In addition, I would also add "ownership, commitment and fairness" to this list of essential prerequisites.

Di Comite *et al.* concluded that the Latvian experience has shown that the trade-off between short-term pain and long-term gain can be avoided if intervention is sufficiently well-designed. The

---

[1] The credit, housing and consumption boom was fuelled by capital inflows in the form of FDIs, portfolio investments and loans. The highly credible fixed exchange rate peg to the euro created an incentive to borrow in foreign currency; due to a rising domestic inflation, negative real interest rates pushed up demand for loans and amplified the boom. The observable rapid economic growth fuelled expectations that high growth would continue; this also encouraged people to borrow. Moreover, fiscal policy also behaved in Latvia in a highly pro-cyclical manner. Consequently, both supply and demand factors contributed to the emergence of substantial credit booms, housing booms and consumption booms and an overheated economy (see also Darvas, 2009).

hypothesis that Latvia's huge frontloaded fiscal adjustment triggered non-Keynesian effects or growth-stimulating effects is debatable, however. An antagonistic, if not sarcastic, take on this issue was expressed by Krugman: "a few more successes like this and Latvia will be back in the Stone Age";[2] in other words, the economic and social cost of Latvia's fiscal adjustment was actually very high, despite its "growth-friendly" composition.

Non-Keynesian effects would arise if GDP growth shows an immediate positive reaction to a negative fiscal shock via positively affecting consumer confidence, business confidence and/or external competitiveness. However, according to Latvijas Banka's Macroeconomic Developments Report, Latvia's GDP decreased by nearly 20 per cent (year on year) and by almost 7 per cent (quarter on quarter) in the third quarter of 2009. Gross fixed capital formation slumped by nearly 40 per cent (quarter on quarter, in real terms) in the first quarter and showed negative growth rates until the second quarter of 2010; the private consumption level shrank by more than 30 per cent between the first quarter of 2008 and the second quarter of 2009 and continued to shrink by about 3 per cent (quarter on quarter) also in the third quarter of 2009. From 2007 to 2010, Latvia's GDP fell by about 25 per cent, while unemployment rose to more than 20 per cent and poverty increased dramatically.

Household sentiment started to improve in the fourth quarter of 2009, but only at an exceptionally low level. Given this extraordinary reduction in GDP, and thus in income, the more optimistic perception of households *vis-à-vis* future developments in their budgets and the economy probably arose from a conviction that the lowest point of the economic downturn had been reached. The following development would therefore be consistent with a mean-reverting process that had started in early 2010. The "recovery" of the Latvian economy at the beginning of 2010 was actually very weak: in the third quarter of 2010, private consumption reached the same level as in the second quarter of 2009; with gross fixed capital formation, it was not until the third quarter of 2011 that it reached a level comparable with that of the second quarter of 2009. Moreover, other countries – notably some of Latvia's important trading partners – picked up in the third quarter of 2009; according to the IMF's Fifth Review Report, exports had already started to increase in the second quarter of 2009. Moreover, at the beginning of 2010, household sentiment may have been affected positively by a ruling of the Constitutional Court of the Republic of Latvia in favour of pensioners, stating that the part of the pensions that had been withheld would have to be reimbursed and that no further withholding of pensions would be allowed.

The second presentation, submitted by Arpaia and Turrini, focused on the impact of fiscal adjustment measures on job creation and job destruction and the share of long-term unemployment in high-versus-low employment protection environments. The findings were based on a dynamic panel regression for EU countries, mostly from the mid-1990s until 2009-10. The regressions were run for four different dependent variables – cyclical unemployment, job separation rates, job finding rates and long-term unemployment shares – with two different kinds of fiscal explanatory variables, namely "action-based" as opposed to "structural-based", and were run separately for countries with low employment protection and those with a high level of protection. The presentation encompassed only the bunch of regressions. The main result from the regression exercise was that the impact of fiscal adjustment on worker flows and share of long-term unemployment is different with respect to the employment protection regime. As this impact seems to be worse in countries with high levels of employment protection, the authors conclude that liberalising the labour market (*i.e.*, easing employment protection legislation), while at the same time consolidating public finance, does not need to be associated with higher unemployment; indeed it may actually mitigate the impact of consolidation on long-term unemployment.

[2] See http://krugman.blogs.nytimes.com/2011/07/18/lats-of-luck/

Growth implications of employment protection regimes and labour market regulations feature strongly in the debate on EU-2020. Their interaction with fiscal adjustment measures is of primary interest to policymakers. Thus, the absence of any theoretical foundation or discussion is problematic. Given this shortcoming, the authors would be well-advised to abstain from too strong – and probably premature – policy conclusions. Furthermore, it is even more problematic to try to draw policy conclusions with respect to the intensity of employment protection legislation without taking into account their interaction with other labour market institutions, such as unemployment benefits.

REFERENCES

Darvas, Z. (2009), "The Baltic Challenge and Euro-area Entry", Bruegel Policy Contribution, Issue 2009/13.

IMF (2010), "Will It Hurt? Macroeconomic Effects of Fiscal Consolidation", *World Economic Outlook*, October, Chapter 3, pp. 93-124.

Perotti, R. (2011), "The 'Austerity Myth': Gain Without Pain", BIS, Working Paper, No. 362.

Session 2

GOVERNMENT BUDGETS AND POTENTIAL GROWTH

DEBT AND GROWTH: NEW EVIDENCE FOR THE EURO AREA

Anja Baum,[*] Cristina Checherita-Westphal[**] and Philipp Rother[***]

Against the background of the euro area sovereign debt crisis, our paper investigates the relationship between public debt and economic growth and adds to the existing literature in the following ways. First, we use a dynamic threshold panel methodology in order to analyse the non-linear impact of public debt on GDP growth. Second, we focus on 12 euro area countries for the period 1990-2010, therefore adding to the current discussion on debt sustainability in the euro area. Our empirical results suggest that the short-run impact of debt on GDP growth is positive and highly statistically significant, but decreases to around zero and loses significance beyond public debt-to-GDP ratios of around 67 per cent. This result is robust throughout most of our specifications, in the dynamic and non-dynamic threshold models alike. For high debt-to-GDP ratios (above 95 per cent), additional debt has a negative impact on economic activity. Furthermore, we can show that the long-term interest rate is subject to increased pressure when the public debt-to-GDP ratio is above 70 per cent, broadly supporting the above findings.

Non-technical summary

The fiscal situation remains challenging in much of the developed world, particularly in the euro area. Market concerns with respect to fiscal sustainability in vulnerable euro area countries have grown and spread to other countries. Against this background, empirical research has started to focus on estimates of the impact of public debt on economic activity.

Looking at the debt-growth nexus literature, two characteristics become apparent. First, only few studies focus on euro area countries. This is insofar surprising as the euro area/EMU offers economic dynamics that are rarely found anywhere else in the world. Moreover, this group of countries is in need of special attention given the current sovereign debt crisis. Second, most of the empirical studies still rely on linear estimation frameworks. Only more recently has the focus been shifting to non-linear threshold analyses, inter alia by employing the threshold panel methodology developed by Hansen (1999). However, all of these studies focus exclusively on non-dynamic panel models, which might lead to inconsistent results due to the persistence of GDP growth rates. To our best knowledge our paper is the first to account for this problem through application of a dynamic threshold framework. Comparing the results from dynamic and non-dynamic threshold estimations provides an idea not only about the robustness of the impact of debt on growth, but also about the robustness of the estimated optimal debt ratios.

Our paper adds to the existing literature in the following ways. First, we use a dynamic threshold panel methodology in order to analyse the non-linear impact of public debt on GDP growth. Second, in comparison to the majority of empirical studies we analyse the short-run relationship between public debt and economic growth using yearly data. Third, our focus on EMU

[*] University of Cambridge, Faculty of Economics, Austin Robinson Building, Sidgwick Avenue Cambridge, CB3 9DD, UK. E-mail: ab862@cam.ac.uk

[**] European Central Bank, Fiscal Policies Division, Kaiserstrasse 29, D-60311 Frankfurt am Main, Germany. E mail: cristina.checherita@ecb.europa.eu

[***] European Central Bank, Fiscal Policies Division, Kaiserstrasse 29, D-60311 Frankfurt am Main, Germany. E-mail: philipp.rother@ecb.europa.eu

We are grateful to participants of an ECB seminar, and in particular to B. Born, G. Callegari, J. Cimadomo, F. Holm-Hadulla, A. Klemm, G. Koester, D. Nautz and D. Tambakis for helpful comments on previous versions of the paper. The views expressed in this paper are those of the authors and do not necessarily reflect those of the European Central Bank.

data provides the opportunity to make specific policy inference, adding to the current discussion on the sustainability of debt dynamics in the euro area.

Our empirical results suggest the following. The short-run impact of debt on GDP growth is positive, but decreases to close to zero beyond public debt-to-GDP ratios of around 67 per cent (*i.e.*, up to this threshold, additional debt has a stimulating impact on growth). This result is robust throughout most of our specifications, in the dynamic and non-dynamic threshold model alike. For really high debt ratios (above 95 per cent), additional debt has a negative impact on economic activity. Confidence intervals for the thresholds are tight, that is (63; 69) for the lower threshold and broader at about (80; 100) for the upper one. Furthermore, we can show that the long-term interest rate is subject to increased pressure when the public debt-to-GDP ratio is above 70 per cent, broadly supporting the above findings.

1 Introduction

The current sovereign debt crisis with its epicenter in the euro area has forcefully revived the academic and policy debate on the economic impact of public debt. Market concerns with respect to fiscal sustainability in vulnerable euro area countries have grown and spread to other countries. Against this background, empirical research has started to focus on estimates of the impact of public debt on economic activity, inter alia by attempting to unveil possible non-linearities.

Nonetheless, the empirical literature on this topic remains scarce (see, for example, Schclarek, 2004; and Reinhart and Rogoff, 2010) and only few studies employ a non-linear impact analysis and are of particular interest for our paper. One of these is a contribution by Kumar and Woo (2010), who use dummy variables for pre-determined ranges of debt to show non-linear effects in a sample of emerging and advanced economies. They find that only very high (above 90 per cent of GDP) levels of debt have a significant and negative impact on growth. Another recent contribution is provided by Checherita and Rother (2010). Expressing growth as a quadratic functional form of debt in a sample of twelve euro area countries over a period starting in 1970, they find significant evidence for a concave (inverted U-shape) relationship. The debt turning point, beyond which debt starts having a negative impact on growth, is found at about 90-100 per cent of GDP.[1]

Papers that relate more closely to the non-linear panel threshold methodology we use in this analysis include the work by Chang and Chiang (2009) and Cecchetti, Mohanty and Zampolli (2011). Both of these papers employ the threshold methodology for non-dynamic panels. Chang and Chiang (2009) analyse a sample of 15 OECD countries and use yearly observations for the period 1990-2004. In a generalisation of the Hansen (1999) multiple regime panel threshold model, they run a regression of GDP per capita growth on the debt-to-GDP ratio and find two debt-to-GDP threshold values, 32.3 per cent and 66.25 per cent. Interestingly, the impact of the debt ratio is positive and significant in all three regimes, higher in the middle regime and lower in the two outer regimes. They thus cannot support the crowding-out view if the debt-to-GDP ratio is more than the threshold value.[2] Cecchetti, Mohanty and Zampolli (2011) use a sample of 18 OECD countries for the period 1980-2010 and obtain a threshold for government debt at 85 per cent of GDP. In contrast to Chang and Chiang (2009), they find a negative impact on growth in the high debt regime.

---

1 Confidence intervals for the debt turning points provided in Checherita and Rother (2010) suggest that the negative growth effect of high debt may start already from levels of around 70-80 per cent of GDP.

2 Chang and Chiang (2009) apply a panel smooth transition regression (PSTR), with a continuous transition function depending on an observable transition variable. In their additive version of this model, the transition function becomes an indicator function, with $I[A] = 1$ when event A occurs, and 0 otherwise. As a consequence, the additive PSTR model is equivalent to the multiple regime threshold model developed by Hansen (1999).

Going through the current empirical debt-growth nexus literature, three characteristics become apparent. First, none of the above mentioned papers uses a dynamic panel threshold approach. Because of the likely persistence in the economic growth rate, the neglect of such a dynamic specification might lead to inconsistent results. Including such dynamics, on the other hand, allows us to capture the effect of debt on growth after controlling for growth persistence, and in this way it is well suited for estimating short-run relationships. To our best knowledge, the current paper is the first to estimate a dynamic threshold model for the debt-growth nexus and then to compare the results of dynamic and static panel estimations. It thus also provides an idea about the robustness of results across different methodologies.

Second, most of the above papers study the long-term impact of debt on growth (Schclarek, 2004; Reinhart and Rogoff, 2010; Kumar and Woo, 2010; Checherita and Rother, 2010).[3] So far, the only exception has been Chang and Chiang (2009), who use exclusively yearly data and thus capture a short-term impact comparable to our focus. On the same note, most of the literature on short-term growth effects analyses fiscal multipliers of shocks to government expenditure or taxes (see Hemming *et al.*, 2002; and van Riet, 2010, for relevant surveys), and if the role of debt is accounted for, its influence is indirect. IMF (2008), for instance, finds that the impact of discretionary fiscal impulses on real GDP growth is contingent on the level of debt, *i.e.*, it is positive and larger at low government debt levels (relative to the sample average). Differently from these studies, the objective of the present paper is to investigate the direct (short-term) impact of debt on growth.

Third, Checherita and Rother (2010) has been so far the only paper focussing exclusively on euro area countries. This is surprising as the EMU offers economic dynamics that are rarely found elsewhere in the world. Moreover, with the current sovereign debt crisis, the euro area would be in need of particular attention, while averaging across OECD countries makes policy inferences difficult.

To summarise, our paper adds to the existing literature in the following ways. First, we use a dynamic threshold panel methodology, inter alia by adapting the methodology proposed in Hansen and Caner (2004), and use it to analyse the non-linear impact of public debt on GDP growth. To our best knowledge, a comparable approach has been applied only once before, in a contribution by Kremer, Bick and Nautz (2009), who analyse the non-linear impact of inflation on growth. Second, we study the short-run relationship between public debt and economic growth using yearly data. Third, our focus on EMU data provides the opportunity to make specific policy inference, adding to the current discussion on the sustainability of debt dynamics in the euro area.

The paper is organised as follows. Section 2 describes the employed methodology and Section 3 presents the data. The estimation results are shown in Section 4. Section 5 employs several robustness exercises, including a broad extension of the explanatory variable set and an analysis of the impact of debt on long-term interest rates. Section 6 concludes.

2 Methodology

In order to account for the persistence of the growth rate, we need a threshold model that allows for endogeneity. Caner and Hansen (2004) develop a threshold methodology for dynamic models, which has to be extended to a panel framework. With several differences as explained

[3] Checherita and Rother (2010) use both yearly data for the dependent variable (and one year-lagged debt data), as well as 5-year overlapping and non-overlapping averages (with debt measured at the beginning of the 5-year period and estimates corrected in all cases for time autocorrelation), but do not find radically different results across the various specifications. Cecchetti, Mohanty and Zampolli (2011) use the (less conventional) long-term approach by employing only the 5-year overlapping average growth rates.

below, the extension we apply here has been first suggested by Kremer *et al.* (2009), who analyse the non-linear impact of inflation on growth within an Arellano and Bover (1995) estimation.[4]

The starting point for the threshold analysis is the specification of a linear model, which in the present case is a balanced panel of the form:

$$y_{it} = \mu_i + \chi y_{i,t-1} + \alpha X_{it} + u_{it} \tag{1}$$

$y_{i,t}$ is the dependent variable of country i at time t, $y_{i,t-1}$ is the endogenous regressor, in our case the lagged dependent variable, μ_i are the country specific fixed effects and X is a set of explanatory regressors. The error term μ_{it} is independent and identically distributed with mean zero and finite variance. The linear model can be estimated following the Arellano and Bond (1991) dynamic panel approach.[5]

We estimate the dynamic threshold model following the approach by Caner and Hansen (2004), who develop an estimator and an inference theory for models with endogenous variables and an exogenous threshold variable. Since Caner and Hansen (2004) do not apply their procedure to panel data we first have to make their framework suitable to deal with the country-specific fixed effects. While in a non-dynamic panel model the individual effects μ_i can be removed by mean differencing, in the dynamic panel mean differencing leads to inconsistent estimates due to the fact that the lagged dependent variable will always be correlated with the mean of the individual errors and thus with all of the transformed individual errors (see Arellano, 2003, p. 17). As an alternative we apply a strategy as first suggested in Kremer *et al.* (2009) and use forward orthogonal deviations[6] (1995). The method subtracts the average of all future available observations of a variable and makes it possible to maintain the uncorrelatedness of the error terms.[7]

The dynamic panel threshold model can be represented with:

$$y_{it} = \mu_i + \chi y_{i,t-1} + \alpha' x_{it} + \beta_1 d_{it} I(z_{it} \leq z^*) + \beta_2 d_{it} I(z_{it} > z^*) + u_{it} \tag{2}$$

where x is a set of regime independent control variables, d is the set of variables allowed to switch between regimes, and l is an indicator function taking on the value 1 if the value of the threshold series z is below a specific threshold value z^*.

In the estimation of the dynamic panel model, we first run a reduced form regression of the endogenous variable on a set of instruments. For the lagged GDP growth rate we use higher lags of GDP growth as instruments and we can then replace $y_{i,t-1}$ in equation (2) with its predicted values $\hat{y}_{i,t-1}$.

After the reduced form regression the threshold model can be estimated, with the specific threshold value being determined following the strategy by Hansen (1999). The procedure includes three essential steps:

1) first, we conduct a series of least squares (LS) minimisations. That is, we estimate model (2) with 2SLS for each value of the threshold series z. The corresponding LS estimates of the

---

4 An alternative approach for a dynamic threshold model can be found in Cimadomo (2007). He extends the Hansen (1999) approach by a two stage procedure. In the first step, the autoregressive coefficient is estimated from a linear regression. In the second stage this coefficient is treated asknown and fixed in the non-linear panel regression model.

5 In contrast to our paper, Kremer *et al.* (2009) employ the Arellano and Bover (1995) estimator, as they focus on the central role of initial income for growth convergence. Due to the endogeneity of the lagged level of GDP, the application of Arellano and Bover (1995) is necessary. Since we focus on growth persistence and a short-run impact analysis, the Arellano and Bond (1991) estimation is more appropriate.

6 Programming codes for forward orthogonal deviations can be obtained from http://www.cemfi.es/ arellano.

7 An empirical Monte Carlo proof for the advantage of orthogonal deviations over mean deviations is found in Hayakawa (2009).

parameters and the sum of squared residuals are kept;[8]

2) in a second step the threshold value z^* is selected as the one which minimises the sum of squared residuals;
3) in a third step we test for the significance of the chosen z^*. Since the threshold value is not identified under the null of linearity, the distribution of a standard *F*-statistic is not chi-square. Hansen (2000) therefore proposes a bootstrap procedure with which the asymptotic null distribution of the heteroscedasticity adjusted test statistic can be approximated.[9]

Hence, we test for the threshold significance using the test statistic:

$$F_T = \sup_{z \varepsilon S} F_T(z) \tag{3}$$

where:

$$F_T(z) = T\left(\frac{\tilde{\sigma}_T^2 - \hat{\sigma}_T^2(z)}{\hat{\sigma}_T^2(z)}\right) \tag{4}$$

where $\hat{\sigma}_T^2 = \frac{1}{t}\sum_{t=1}^{T}\hat{u}_t^{\,2}$ is the estimated residual variance of the threshold and $\tilde{\sigma}_T^2$ is the residual variance of the corresponding linear model. Details of the testing procedure are described in the Appendix.

If we find a significant threshold value z^*, the slope coefficients of equation (2) are estimated with GMM.[10] For a more efficient weighting matrix in the coefficient estimation, we prefer the general GMM to the 2SLS estimator, and repeatedly predict the residuals to construct new covariance matrices of the moments after the initial 2SLS estimate.

We also allow for the possibility of more than one threshold and therefore more than two regimes (see Hansen 1999), but since a second threshold value turns out to be insignificant in most of the specifications it will be ignored in the following analysis, unless specified otherwise.

3 Data

3.1 Structural considerations

The model is estimated for 12 euro area countries Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal and Spain for yearly data starting with 1990. Using this relatively short time span offers a couple of advantages. First, the shorter period covers more accurately the process of EMU preparation and implementation and is thus less prone to structural changes and more comparable with today's economic conditions. More importantly, the debt-to-GDP ratio is found to be non-stationary upon inclusion of the previous decade (the 1980s). Using the longer time span we would not be able to fully rely on the results of

---

8 This step is repeated for each value of the threshold series on a specified subset of the series, which should be trimmed in order to assure a minimum number of observations in the resulting subsamples. In the non-dynamic model, the 2SLS estimator reduces to the simple LS estimator.

9 We test the null hypothesis of linearity against threshold non-linearity also allowing heteroscedasticity in the error terms. Caner and Hansen (2004) provide evidence that the distribution theory in Hansen (2000) is applicable to the case of 2SLS estimation. However, a full distribution theory for dynamic panels has not yet been provided (we thank Bruce Hansen for his comments). The specific coefficients on the explanatory variables of the dynamic model should thus be considered carefully. Since on the other hand the non-dynamic panel estimation might give inconsistent results due to omitted lagged variables, the direct comparison of both approaches will give an idea about the range in which the coefficients lie.

10 The slope coefficients of non-dynamic model are estimated by OLS.

the threshold testing procedure and, consequently, on the obtained threshold values.[11] Given the above, we base our main estimation models on the period 1990-2007/10 (we do, however, include a discussion on results from the year 1980 onwards in the robustness section).

We are analysing the impact of one-year lagged debt-to-GDP ratios on annual real GDP growth rates. We thus obtain a near contemporaneous effect, which gives us an idea of the short-term debt impact. Hence, a positive impact of debt on growth could be interpreted as a stimulating effect of additional debt. However, the possibility that long-term effects of high debt might be negative cannot be ruled out based on the yearly analysis.

3.2 *Endogenous, regime-dependent variable and other control variables*

The data used originates primarily from the European Commission AMECO database. The endogenous variable is the real GDP growth rate. As the single regime-dependent and threshold variable we use the debt-to-GDP ratio. Since this can be correlated with a range of other factors impacting on growth, we also control for a broad set of other explanatory variables. In the benchmark specification, we include the gross fixed capital formation as a share of GDP, trade openness (defined as imports plus exports as a share of GDP), and a dummy signalling the period of effective EMU membership. Moreover, under the robustness tests we control for other potentially relevant variables as identified in the theoretical and empirical growth literature, such as the initial level of GDP per capita, population growth, secondary education, a measure for the old dependency ratio, the unemployment rate, the budget balance and long- and short-run interest rates.

4 Estimation

4.1 *Benchmark model*

The benchmark model for the 12 EMU countries over the period 1990-2007 (first, excluding the current crisis years) is estimated in the following specification:

$$y_{it} = \mu_i + \chi y_{i,t-1} + \alpha_1 OPEN_{i,t-1} + \alpha_2 GCF_{i,t-1} + \alpha_3 EMU_{it}$$
$$+ \beta_1 d_{i,t-1} I(d_{i,t-1} \leq d^*) + \beta_2 d_{i,t-1} I(d_{i,t-1} > d^*) + u_{it} \tag{5}$$

where y is the GDP growth rate, $OPEN$ is the trade openness measure, GCF is the ratio of gross capital formation to GDP, EMU is the dummy variable which signals the EMU membership, and d is the debt-to-GDP series, with d^* being the debt-to-GDP threshold value. For the dynamic model, $y_{(t-1)}$ is replaced by the predicted values $\hat{y}_{(t-1)}$ obtained from the structural first stage regression of $y_{(t-1)}$ on the lags of $y_{(t-2)}$ to $y_{(t-8)}$. Of course, GDP growth in the structural equation could be dependent on more than one lag. However, we find a second and higher lags to be insignificant in all of our specifications, and therefore they will be ignored in the following analysis.

Table 1 shows the benchmark results for the non-dynamic and the dynamic panel threshold estimation. We can see some differences between the two models, but for both the direction and the significance of the coefficients are comparable. As such, trade openness has a significantly positive effect on GDP growth, the coefficient on investment is positive but insignificant and the EMU dummy is significantly negative. In the dynamic model the strongest impact on current growth comes from the past growth rate itself.

11 Details on the distribution theory can be found in Hansen (2000).

Table 1

Benchmark Results, 1990-2007

Variable	Non-dynamic Panel		Dynamic Panel	
$y_{(t-1)}$			0.4583^{***}	*(0.1055)*
Openness	0.0148^{**}	*(0.0064)*	0.0172^{**}	*(0.0078)*
GCF	0.0539	*(0.0401)*	0.0184	*(0.0396)*
EMU	-0.0070^{**}	*(0.0034)*	-0.0099^{***}	*(0.0031)*
$d_{(t-1)}$ if $d \le d^*$	0.0697^{***}	*(0.0209)*	0.0668^{***}	*(0.0148)*
$d_{(t-1)}$ if $d > d^*$	0.0082	*(0.0095)*	0.0124	*(0.0104)*
Threshold Estimate	$d^* = 0.6640$		$d^* = 0.6644$	
Bootstrap p-value	0.0630		0.0780	
Confidence Intervals	$0.6287 < d^* < 0.6831$		$0.6287 < d^* < 0.6908$	

Standard errors in brackets.
$^{*}/^{**}/^{***}$ indicate significance levels at the 10/5/1 per cent level.

Independent of the specifications, both models find a debt threshold value of around 0.664, which is significant at the 10 per cent level with p-values of 0.063 and 0.078 for the non-dynamic and the dynamic model, respectively. This threshold value splits the observations of the non-dynamic (dynamic) panel into 128 (125) observations in the lower, and 88 (91) observations in the upper regime. When the debt ratio is below 66.4 per cent of GDP, the impact of additional debt is significantly positive in both specifications, with coefficients corresponding to around 0.07 percentage point increase in the annual growth rate after a 1 percentage point increase in the debt-to-GDP ratio. If the debt ratio is above the threshold value, the impact reduces to values around zero, which are therefore insignificant.

This is a very strong result. Additional debt might have a positive impact on GDP growth due to stimulus effects of fiscal policy. However, once a debt threshold is reached this positive effect disappears or becomes insignificant.

4.2 Including the years 2008-10

We re-estimate the model including the crisis years 2008 to 2010. The results for the two threshold models are presented in Table 2. The threshold value of the non-dynamic model increases slightly to 71.7 per cent. At the same time, the regime-independent coefficients change notably compared to the benchmark results, with the GCF being the only positive and significant variable. The impact of debt on GDP growth also changes substantially. For the extended period, it is significantly positive in the lower regime, and significantly negative in the upper regime, while now diverging more in absolute size between the two specifications.

Including the years 2008 to 2010 in the dynamic specification gives the high threshold value of 95.6 per cent, which is significant at the 10 per cent level with a p-value of 0.098, resulting in

Table 2

Benchmark Results, 1990-2010

Variable	Non-dynamic Panel		Dynamic Panel	
$y_{(t-1)}$			0.3218^{***}	*(0.1245)*
Openness	–0.0082	*(0.0072*	0.0014	*(0.0058)*
GCF	0.1126^{**}	*(0.0529)*	0.0147	*(0.0568)*
EMU	–0.0071	*(0.0045)*	-0.0091^{**}	*(0.0036)*
$d_{(t-1)}$ if $d \leq d^*$	0.0470^{***}	*(0.0182)*	0.0351^{***}	*(0.0107)*
$d_{(t-1)}$ if $d > d^*$	-0.0411^{***}	*(0.0144)*	-0.0588^{***}	*(0.0200)*
Threshold Estimate	$d^* = 0.717$		$d^* = 0.956$	
Bootstrap p-value	0.0960		0.0980	
Confidence Intervals	$0.6287 < d^* < 0.7809$		$0.8140 < d^* < 1.0344$	

Standard errors in brackets.
$^*/^{**}/^{***}$ indicate significance levels at the 10/5/1 per cent level.
Threshold of 0.717 splits the sample into 168 observations in the lower and 85 in the upper regime.
Threshold of 0.956 splits the sample into 198 observations in the lower and 55 in the upper regime.

198 observations in the lower, and 55 observations in the upper regime.[12] Except for trade openness the regime-independent coefficients are more robust to changes in the time span than in the non-dynamic model (hence, the lagged GDP is significantly positive, GCF insignificant and the EMU dummy significantly negative). However, the changes for the regime dependent debt variable are comparable to the non-dynamic panel. In the lower regime, the impact of debt is positive at 0.035 per cent, while in the upper regime we obtain a larger negative impact of –0.059 per cent (both values being significant).

With a coefficient of 0.035 the impact in the lower regime decreases strongly compared to the value of 0.067 in the specification without the years 2008-10. However, since the introduction of the higher debt threshold leads to an average estimate over almost the entire original sample (plus a few new observations), we re-estimate the dynamic model with a second threshold, combining the multiple threshold estimation strategy by Hansen (1999) with our framework. We fix the first threshold at 95.6 per cent, and test for a second threshold in the lower sample. We indeed find a second threshold corresponding to the smallest sum of squares again to be 0.664, but it is insignificant with a p-value of 0.147. For illustration purposes the estimation results including the second threshold are shown in Table 3. Compared to the results of the dynamic model presented in Table 2, the debt impact in the lowest sample is now higher (0.0496), while the value of the second regime is insignificant and close to zero up to the threshold of 95.6 per cent of GDP. Afterwards, the debt impact remains negative, highly statistically significant and similar in size.

Hence, our results suggest that debt can have a stimulus effect on growth in the EMU up to a value of between 60 and 70 per cent of GDP. Above that, the growth impact becomes first insignificant, before turning negative for very high debt-to-GDP ratios.

---

[12] The reason for a higher threshold when the years 2008-10 are included is that the point of highest significance of the one break we are looking at shifts upwards. Using a data set up to 2007, we had only few observations with debt higher than 95 per cent of GDP.

Table 3

Second Threshold Value – Dynamic Panel

Variable	Dynamic	
$y_{(t-1)}$	0.3221***	(0.1245)
Openness	–0.0001	(0.0058)
GCF	0.0200	(0.0567)
EMU	–0.0092**	(0.0037)
$d_{(t-1)}$ if $d \leq 0.664$	0.0496***	(0.0137)
$d_{(t-1)}$ if $0.664 \leq d \leq 0.956$	0.0146	(0.0114)
$d_{(t-1)}$ if $d > 0.956$	–0.0591***	(0.0200)

Standard errors in brackets.
*/**/*** indicate significance levels at the 10/5/1 per cent level.
The two thresholds split the sample into 154 observations in the lower regime, 44 in the middle regime, and 55 in the upper regime.

5 Robustness

To make sure that our results are robust throughout a broader range of specifications, we conduct a variety of additional tests. Those include further explanatory variables, an extension of the time frame, further endogeneity tests, an analysis of influential euro area countries, and an analysis employing the real sovereign long term interest rate as the dependent variable. For most of the robustness tests, the results of the benchmark specification can be supported and remain consistent.

5.1 Including further explanatory variables

Next to lagged GDP growth, trade openness, gross capital formation and the dummy for EMU membership, we consecutively include further explanatory variables to test for robustness of the results. These are population growth, the old dependency ratio, the unemployment rate, secondary education, GDP per capita, the general government budget balance and primary budget balance (in ratios to GDP), private gross capital formation (replacing the aggregate variable) and the long and short term interest rates. All variables included are lagged one year compared to the dependent variable in order to avoid further endogeneity. Table 4 shows the results for the threshold dynamic model. Altogether, there are comparatively few changes in the coefficients and their significance, no matter which other variable is included.[13] Furthermore, for all the specifications the estimated threshold associated with the smallest sum of squares is 66.4 per cent, and the threshold value remains significant at the 10 per cent level. The debt coefficients of the two regimes are mostly comparable to the benchmark specification. Only for the last two columns the debt impact is smaller, but it is still significant and positive in the lower, and very close to zero in the upper regime.

[13] This is also true if the explanatory variables are used without or with two lags instead.

Table 4

Robustness, Dynamic Model – Non-linear, 1990-2007

Variable	(1)	(2)	(3)	(4)	(5)[a]	(6)	(7)	(8)	(9)	(10)[b]
$y_{(t-1)}$	0.4679*** (0.1080)	0.4448*** (0.1046)	0.4473*** (0.1414)	0.4592*** (0.1073)		0.4278*** (0.1224)	0.4551*** (0.126)	0.4183*** (0.1115)	0.3496*** (0.1008)	0.4135*** (0.1081)
Openness	0.0176** (0.0079)	0.0176** (0.0077)	0.0188** (0.0088)	0.0176** (0.0077)	0.0326*** (0.0139)	0.0153* (0.009)	0.0141 (0.0090)	0.0158* (0.0088)	0.0214*** (0.0074)	0.0285*** (0.0098)
GCF	–0.0024 (0.0599)	0.0004 (0.0476)	–0.0218 (0.0492)	0.0193 (0.0402)	–0.0182 (0.0472)	–0.0408 (0.0468)	–0.0411 (0.0475)		–0.0144 (0.0382)	–0.0318 (0.0362)
EMU	–0.0096*** (0.0031)	–0.0095*** (0.0032)	–0.0112*** (0.0041)	–0.0101** (0.0031)	–0.00001 (0.0042)	–0.0052 (0.0041)	–0.0046 (0.0037)	–0.0053 (0.0035)	–0.0117*** (0.0030)	–0.0120 (0.0114)
Population growth	0.2643 (0.5179)									
Old ratio		–0.1693 (0.2067)								
Unemployment			–0.1024 (0.1288)							
Secondary education				0.0000 (0.0001)						
GDP per Capita					–1.8948 (1.278)					
Budget balance						–0.0153 (0.0668)				
Primary budget balance							–0.0409 (0.0623)			
GCF Private								–0.0543 (0.0544)		
Long run interest rates									–0.4230*** (0.1086)	
Short run interest rates										–0.3390*** (0.0781)
$d_{(t-1)}$ if $d \leq 0.664$	0.0697*** (0.0159)	0.0670*** (0.0144)	0.075**** (0.0191)	0.0669*** (0.0152)	0.0626*** (0.0142)	0.0718*** (0.0151)	0.0707*** (0.0145)	0.0730*** (0.0145)	0.0491*** (0.0137)	0.0396*** (0.0114)
$d_{(t-1)}$ if $d \leq 0.664$	0.0157 (0.0122)	0.0120 (0.0105)	0.0211 (0.0139)	0.0137 (0.0115)	0.0054 (0.015)	0.0125 (0.0117)	0.0112 (0.0119)	0.0133 (0.0115)	0.0048 (0.0127)	0.0055 (0.0106)
Bootstrap *p*-value	0.085	0.069	0.10	0.075	0.10	0.084	0.092	0.10	0.070	0.080
Confidence Region	$0.6287<d^*<0.6908$	$0.6287<d^*<0.6908$	$0.6287<d^*<0.6831$	$0.6287<d^*<0.6908$	$0.6287<d^*<0.6698$	$0.6287<d^*<0.6908$	$0.6287<d^*<0.6908$	$0.6287<d^*<0.6908$	$0.6127<d^*<0.6831$	$0.6287<d^*<0.7210$

Standard errors in brackets.
a) Non-dynamic estimation since lagged GDP per capita and lagged GDP growth rate are highly correlated.
b) Estimation excludes Luxembourg due to data limitations.
$^{*}/^{**}/^{***}$ indicate significance levels at the 10/5/1 per cent level.

Table 5

Alternative Endogenous Variables, Threshold Panel, 1990-2007

Dependent Variable	**Potential GDP Growth**		**GDP Growth**		**GDP Growth**	
Estimation Method	**Dynamic Panel (a)**		**Dynamic Panel (b)**		**IV 2SLS (c)**	
$y_{(t-1)}$	0.8562***	(0.0344)	0.2209**	(0.1008)	0.4234***	(0.1102)
Openness	0.0038**	(0.0018)	0.0310**	(0.0089)	0.0132**	(0.0058)–
GCF	–0.0356***	(0.0096)	–0.0569	(0.0572)	0.0246	(0.0413)
EMU	–0.0020***	(0.0008)	–0.0117***	(0.0043)	–0.0077**	(0.0037)
$d_{(t-1)}$ if $d \leq d^*$	0.0163***	(0.0028)	0.0867***	(0.0177)	0.0583***	(0.0119)
$d_{(t-1)}$ if $d > d^*$	0.0041	(0.0030)	0.0185	(0.0149)	–0.0016	(0.0161)
Threshold Estimate	$d^* = 0.6644$		$d^* = 0.6640$		$d^* = 0.6640$	
Bootstrap p-value	0.026		0.085		0.058	
Confidence Intervals	$0.6287 < d^* < 0.7170$		$0.6287 < d^* < 0.6908$		$0.6287 < d^* < 0.6831$	

(a) $y_{(t-1)}$: potential GDP growth; (b) $y_{(t-1)}$: output gap;
(c) $y_{(t-1)}$: GDP growth, debt/GDP as second endogenous variable.
Standard errors in brackets.
*/**/*** indicate significance levels at the 10/5/1 per cent level.

5.2 *Including the period 1980-1989*

As discussed above, the non-stationarity of the debt-to-GDP variable if the years 1980-89 are included causes the resulting threshold estimates to be potentially unreliable. We do, however, re-estimate the model including the foregoing decade to examine whether our implications are generally stable. The estimation suggests that while the obtained linear (regime-independent) coefficients do not change significantly, including the previous decade leads to insignificant threshold estimates.[14] Although insignificant, the two debt-to-GDP ratios associated with the lowest sum of squares lie on average around 0.20 and 0.67, depending on the specification. The lower values can be explained by the lower average debt ratios prevailing in the 80s.

---

[14] This result does not change if dummy variables for the 90s or the years 2008-2010 are included. The results in this subsection are available upon request.

5.3 An alternative endogenous variable / Dealing with endogeneity

In addition to using the GMM estimation[15] to further control for the possibility of endogeneity problems we estimate the dynamic panel with the growth rate of potential GDP instead, where the first lag of the dependent variable, $y_{(t-1)}$, is instrumented with longer lags of the GDP growth rate. The results are shown in the first column of Table 5. The employed endogenous GDP variable has little impact on the significance and size of the threshold value and the debt coefficients, as well as on the direction of the regime-independent variables (the only change is observed in the significance of GCF, which is now significantly negative). The threshold estimate is again 66.4 per cent, being statistically significant at the 5 per cent level. The impact of debt below the threshold decreases, but is still positive and significant, while the impact above 66.4 per cent remains insignificant.

As an alternative, we replace the lagged GDP growth rate in the benchmark specification with the lagged output gap, which is again instrumented with further lags of GDP growth. The results are shown in the second column of Table 5. The coefficient on the output gap series, $y_{(t-1)}$ is positive and significant, while the threshold value and all of the remaining coefficients are comparable to the benchmark specification.

Another endogeneity issue might arise from the debt variable itself. That is, we can expect reverse causation between GDP growth rates and debt levels (low growth rates are likely to result in higher debt-to-GDP ratios). Even though the positive values of the debt coefficients in the benchmark estimation rule out the possibility of reverse causation almost entirely, we still control for endogeneity to check if the results are altered significantly. If this was the case, we could suspect further endogeneity problems. We would like to continue estimating the dynamic panel when debt endogeneity is taken into account. Unfortunately it is impossible to split the instrumented debt-to-GDP series within the construction of the GMM estimator. Therefore we have to limit our estimation to a less efficient (albeit still consistent) 2SLS estimation of the following form:

1) in a first step, the lagged GDP growth rate and the lagged debt-to-GDP series are regressed on higher lags of both variables plus all the exogenous regressors. We then predict the values for both lagged GDP growth and lagged debt-to-GDP;
2) the threshold testing procedure is similar to the benchmark estimation, only with the regime dependent series being the predicted values for debt/GDP;
3) based on the threshold value, the coefficients are estimated using OLS. The resulting coefficients are the 2SLS estimators.

The third column of Table 5 shows the results from the described regression approach. As can be seen, the coefficients differ only negligibly from the benchmark results.

5.4 Influential countries

Based on the benchmark specification, we first exclude two sets of countries, those with the highest and those with the lowest debt-to-GDP ratios over time. Excluding Luxembourg – the country with the lowest debt-to-GDP ratios – has no significant impact on the results. The same is true if we exclude Belgium or Italy, the two countries with the highest average debt ratios. Even if the two countries are excluded together (resulting in a sample with only 10 countries) the

[15] See Caselli *et al.* (1996) who proposed to use GMM as a way to deal with endogeneity problems in the context of panel growth regressions and Durlauf *et al.* (2005) for a related discussion.

coefficients change only marginally and the significant debt-to-GDP threshold value is again 66.4 per cent.

Next to the outliers of high and low debt ratios, we conduct the exclusion exercise for all the remaining countries (excluding one country at a time). Only two countries seem to have an impact on the debt threshold: Greece and Ireland. Excluding Greece or/and Ireland results in a debt threshold of 45 per cent.[16] The coefficients of debt on GDP growth in the two regimes are comparable to the benchmark results, positive and significant for debt ratios below, insignificant and close to zero above the threshold value.

However, we would like to mention that the exclusion of countries is conducted only as an econometric exercise and is of limited value to our analysis. Not only could we lose significant spillover effects, but we are also specifically interested in the most significant values for the (old) euro area as a whole over the period of our analysis and not only for a subset of countries.

5.5 *Influence on the interest rate*

Finding a significant debt threshold gives rise to the question why its impact on growth becomes smaller once a certain threshold value is reached. Among other channels, higher public debt is likely to be associated by investors with higher sovereign risk premia, which could be translated into higher long-term interest rates. In turn, this may lead to an increase in private interest rates and a decrease in private spending growth, both by households and firms (see Elmendorf and Mankiw 1999), which is likely to dampen output growth. While the empirical findings on the relationship between public debt and long-term interest rates are diverse, a significant number of recent studies suggest that high debt may contribute to rising sovereign yield spreads (see Codogno *et al.* 2003; Schuknecht *et al.* 2010 and Attinasi *et al.* 2009, among others) and ultimately sovereign long-term interest rates (Ardagna *et al.* 2007, Laubach 2009).

In order to examine this hypothesis, we run a non-dynamic threshold estimation of the form:

$$\begin{aligned} INT_{it} &= \mu_i + \alpha_1 INT^s_{i,t-1} + \alpha_2 GDP_{i,t-1} + \alpha_3 OPEN_{i,t-1} + \alpha_4 EMU_{it} \\ &\quad + \beta_1 d_{i,t-1} I(d_{i,t-1} \le d^*) + \beta_2 d_{i,t-1} I(d_{i,t-1} > d^*) + u_{it} \,, \end{aligned} \tag{6}$$

INT is the sovereign long-term real interest rate, INT^S is the short-term real interest rate, which is included to capture monetary policy effects, GDP is the growth rate of GDP, and as before $OPEN$ is the trade openness measure, EMU is the dummy variable which signals the EMU membership, and d is the debt-to-GDP series, with d^* being the threshold value. The explanatory variables are broadly in line with Ardagna *et al.* (2007).[17]

Both interest rate series are de-trended, applying linear trend filtering from 1990. The resulting coefficients for the two periods 1990-2007 and 1990-2010 are presented in Table 6. For both time periods we find a threshold value of 73.8 per cent, significant at 10 per cent, and respectively, at 1 per cent level. Below this threshold, the impact of additional debt decreases the long-run interest rates.[18] Once the threshold is reached, we observe an increasing pressure on the

[16] The results of estimations with Greece and Ireland excluded one at a time are comparable with those resulting from a combined exclusion.

[17] Ardagna *et al.* (2007) estimate the response of long-term interest rates in a panel of 16 OECD countries, over the years 1975-2002. Comparable to our specification, they use the nominal interest rate on 10-year government bonds as the dependent variable, and GDP growth, interest rates on 3-month Treasury bills, inflation and deficit as explanatory variables, a baseline specification which is close to the one employed in our paper.

[18] For a detailed discussion on reasons for the negative impact of debt on interest rates below a threshold value, we refer to Section 3, specifically 3.2 in Ardagna *et al.* (2007).

Table 6

Interest Rates, Non-dynamic Threshold Model

Years	**1990-2007**		**1990-2010**	
INT^S	0.2860^{***}	(0.0551)	0.3881^{***}	(0.0442)
GDP	–0.0801	(0.0509)	–0.0491	(0.0452)
Openness	-0.0172^{**}	(0.0073)	–0.0087	(0.0059)
EMU	0.0077^{**}	(0.0030)	0.0062^{**}	(0.0028)
$d_{(t-1)}$ if $d \leq d^*$	-0.0406^{***}	(0.0089)	-0.0288^{***}	(0.0077)
$d_{(t-1)}$ if $d > d^*$	0.0079	(0.0122)	0.0283^{***}	(0.0086)
Threshold Estimate	$d^* = 0.7380$		$d^* = 0.7380$	
Bootstrap *p*-value	0.078		0.009	
Confidence Intervals	$0.6287 < d^* < 0.7709$		$07220 < d^* < 0.8180$	

Dependent variable: long-term real sovereign interest rates.
Standard errors in brackets.
*/**/*** indicate significance levels at the 10/5/1 per cent level.

interest rate. This is true especially for the longer period, for which the coefficient on the upper regime debt ratio is highly statistically significant and positive. These results are broadly in line with Ardagna *et al.* (2007): using debt in a quadratic functional form, they find a non-linear effect of public debt on long-term interest rates, with a negative impact when the debt-to-GDP ratio is below 65 per cent and a positive impact when the ratio is above this threshold.[19] The resulting crowding-out of economic activity helps explaining why the impact of additional debt on the economy decreases with the size of debt, and might even become negative above certain threshold values.

6 Conclusion

Our paper analyses the short-run impact of debt-to-GDP ratios on GDP growth, using one year lagged debt ratios in a non-linear threshold panel model. The empirical results suggest the following. The short-run impact of debt on GDP growth is positive, but decreases to close to zero and loses significance beyond public debt-to-GDP ratios of around 67 per cent. This result is robust throughout most of our specifications, in the dynamic and non-dynamic threshold models alike. For high debt ratios (above 95 per cent) the impact of additional debt has a negative impact on economic activity. The confidence intervals for the thresholds are generally tight, at about (63; 69) for the lower threshold and broader at about (80; 100) for the upper threshold.

[19] Ardagna *et al.* (2007) further include a panel VAR estimation, which does not account for any form of non-linearity. Clearly, applying the threshold methodology to a VAR specification would be an interesting extension. It is, however, beyond the scope of this paper.

Various robustness tests show that the lower threshold value reacts only marginally to changes in the number of control variables and countries included. The only departure from 67 per cent as the most significant debt threshold value occurs when we include the years before 1990 and the crisis years 2008-10. However, in both cases tests for further thresholds reveal that 67 per cent is associated with the value resulting in the (second) smallest SSR. We further show that the long-term interest rate is subject to increased pressure when the public debt-to-GDP ratio is above 70 per cent, broadly supporting the above findings.

Our results suggest that the positive short term economic stimulus from additional debt decreases drastically when the initial debt level is high, and might even become negative. The reverse would imply that when the debt ratio is very high, reducing it would have beneficial effects for annual growth. On the other hand, in case of low debt levels, reducing the debt further would tend to reduce growth in the short run, in line with conventional Keynesian multipliers (while the long-term effect may differ). Hence, in light of the attempt to defend increasing debt with economic stimulus reasons, our results are supportive only if the initial debt level is below a certain threshold.

APPENDIX
THRESHOLD TESTING

The pointwise *F*-statistic is:

$$F_T = \sup_{z \varepsilon S} F_T(z) \tag{7}$$

where:

$$F_T(z) = T(\frac{\tilde{\sigma}_T^2 - \hat{\sigma}_T^2(z)}{\hat{\sigma}_T^2(z)}) \tag{8}$$

with $\tilde{\sigma}_T^2$ being the estimated residual variance of the corresponding linear model. The threshold value is not identified under the null of linearity and consequently the distribution of the standard *F*-statistic is not chi-square (Hansen 2000). We can approximate the asymptotic distribution with the following bootstrap procedure:

Compute y_t^* iid $N(0,1)$ random draws and regress y_t^* on X_t and on $X_t(z)$ to obtain the residual variances $\tilde{\sigma}_T^{*2}$ and $\tilde{\sigma}_T^{*2}(z)$, respectively. Repeated bootstrap draws from the test statistic:

$$F_T^* = \sup_{z \varepsilon S} F_T^*(z) \tag{9}$$

with:

$$F_T^*(z) = T(\frac{\tilde{\sigma}_T^{*2} - \hat{\sigma}_T^{*2}(z)}{\hat{\sigma}_T^{*2}(z)}) \tag{10}$$

can then be used to approximate the asymptotic null distribution of F_T. The distribution of F_T^* converges weakly in probability to the null distribution of F_T under the alternatives for Γ_2 and the asymptotic bootstrap *p*-value is obtained by counting the percentage of bootstrap samples for which the bootstrap statistic F_T^* exceeds the statistic F_T.

Accounting for possible heteroscedasticity in the error terms, the standard *F*-statistic is replaced by a heteroscedasticity-consistent Wald or Lagrange Multiplier test:

$$L_T = \sup_{z \varepsilon S} L_T(z) \tag{11}$$

with:

$$L_T(z) = (R\hat{\delta}(z))' [R(M_T(z)^{-1} V_T(z) M_T(z)^{-1})R']^{-1} \, R\,\hat{\delta}(z) \tag{12}$$

where $R = (1\ I)$ is the selector matrix, $M_T(z) = \sum X_t(z) X_t(z)'$ and $V_T(z) = \sum X_t(z) X_t(z)' \hat{u}_t^2$.

REFERENCES

Arellano, M. and S. Bond (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *The Review of Economic Studies*, No. 58, pp. 277-97.

Arellano, M. and O. Bover (1995), "Another Look at the Instrumental Variable Estimation of Error-components Models", *Journal of Econometrics*, Vol. 68, No. 1, pp. 29-51.

Arellano, M. (2003), *Panel Data Economics*, Oxford University Press.

Ardagna, S., F. Caselli and T. Lane (2007), "Fiscal Discipline and the Cost of Public Debt Service: Some Estimates for OECD Countries", *The B.E. Journal of Macroeconomics*, Vol. 7, No. 1, Article 28.

Attinasi, M.G., C. Checherita and C. Nickel (2009), "What Explains the Surge in Euro Area Sovereign Spreads During the Financial Crisis of 2007-09?", ECB, Working Paper, No. 1131, also published in *Public Finance and Management*, Vol. 10, No. 4, pp. 595-645, 2010.

Caner, M. and B.E. Hansen (2004), "Instrumental Variable Estimation of a Threshold Model", *Econometric Theory*, No. 20, pp. 813-43.

Caselli, F., G. Esquivel and F. Lefort (1996), "Reopening the Convergence Debate: A New Look at Cross Country Growth Empirics", *Journal of Economic Growth*, Vol. 1, No. 3, pp. 363-89.

Cecchetti, S.G., M.S. Mohanty and F. Zampolli (2011), "The Real Effects of Debt", BIS, Working Paper, No. 352.

Chang, T. and G. Chiang (2009), "The Behavior of OECD Public Debt: A Panel Smooth Transition Regression Approach", *The Empirical Economics Letters*, Vol. 8, No. 1.

Checherita, C. and P. Rother (2010), "The Impact of High and Growing Government Debt on Economic Growth – An Empirical Investigation for the Euro Area", ECB, Working Paper, No. 1237, forthcoming in *European Economic Review*.

Cimadomo, J. (2007), "Fiscal Policy in Real Time", Centre D'Etudes Prospectives et D'Informations International (CEPII), Working Paper, No. 2007-10.

Codogno, L., C. Favero and A. Missale (2003), "Yield Spreads on EMU Government Bonds", *Economic Policy*, No. 18, pp. 505-32.

Durlauf, S., P. Johnson and J. Temple (2005), "Growth Econometrics", in P. Aghion and S.N. Durlauf (eds.), *Handbook of Economic Growth*, 1A, Amsterdam, North-Holland.

Elmendorf, D. and N. Mankiw (1999), "Government Debt", in J. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, 1C, pp. 1615-69, Amsterdam, North-Holland.

Hansen, B.E. (1999), "Threshold Effects in Non-dynamic Panels: Estimation, Testing and Inference", *Journal of Econometrics*, Vol. 93, No. 2, pp. 345-68.

—————(2000), "Sample Splitting and Threshold Estimation", *Econometrica*, Vol. 68, No. 3, pp. 575-603.

Hayakawa, K. (2009), "First Difference or Forward Orthogonal Deviation – Which Transformation Should be Used in Dynamic Panel Data Models? A Simulation Study", *Economics Bulletin*, Vol. 29, No. 3, pp. 2008-17.

Hemming, R., M. Kell and S. Mahfouz (2002), "The Effectiveness of Fiscal Policy in Stimulating the Economic Activity – A Review of the Literature", IMF, Working Paper, No. WP/02/2008.

IMF (2008), "Fiscal Policy as a Countercyclical Tool", IMF, World Economic Outlook, Chapter 5, October.

Kremer, S., A. Bick and D. Nautz (2009), "Inflation and Growth: New Evidence From a Dynamic Panel Threshold Analysis", SFB 649, Discussion Paper, No. 2009-036.

Kumar, M.S. and J. Woo (2010), "Public Debt and Growth", IMF, Working Paper, No. WP/10/174.

Laubach, T. (2009), "New Evidence on the Interest Rate Effects of Budget Deficits and Debt", *Journal of the European Economic Association*, Vol. 7, No. 4, pp. 858-85.

Reinhart, C.M. and K.S. Rogoff (2010), "Growth in a Time of Debt", NBER, Working Paper, No. 15639.

Schclarek, A. (2004), "Debt and Economic Growth in Developing Industrial Countries", mimeo.

Schuknecht, L., J. von Hagen and G. Wolswijk (2010), "Government Bond Risk Premiums in the EU Revisited: The Impact of the Financial Crisis", ECB, Working Paper, No. 1152.

van Riet, A. (ed.) (2010), "Euro Area Fiscal Policies and the Crisis", ECB, Occasional Paper, No. 109.

PUBLIC DEBT AND GROWTH

Manmohan S. Kumar[*] and Jaejoon Woo[*]

This paper examines the impact of high public debt on long-run economic growth in a panel of advanced and emerging economies over four decades, while taking into account various estimation issues including reverse causality and endogeneity. Threshold effects, non-linearities, and differences between advanced and emerging market economies are also explored. High initial public debt is found to be significantly and consistently associated with slower subsequent growth, controlling for other determinants of growth. The adverse effect largely reflects a slowdown in labor productivity growth mainly due to reduced investment and slower growth of capital stock. Extensive robustness checks confirm the results.

1 Introduction

The recent global economic and financial crisis has led to an unprecedented increase in public debt across the world. By the end of 2012, public debt is expected to reach about 107 per cent of GDP in advanced economies – its highest level in 50 years. This has raised serious concerns about fiscal sustainability and their economic impact for many advanced economies amid the current European sovereign debt crisis. What are the effects on longer-term growth of high public debt? This is an important policy question. Surprisingly, however, there has been little systematic empirical analysis in the literature, despite the existence of a very large empirical growth literature (see, for example, Aghion and Durlauf, 2005).[1]

Public debt has important influence over the economy both in the short- and the long run. The conventional view is that debt can stimulate aggregate demand and output in the short run, but crowds out capital and reduces output in the long run (see Elmendorf and Mankiw, 1999 for a literature survey). This paper concerns the long-run effects of public debt. Standard growth theory predicts that an increase in government debt leads to slower growth: a temporary decline in growth along the transition path to a new steady state in the neoclassical model, such as the Solow model, and a permanent decline in growth in the endogenous growth model (Saint-Paul, 1992). Building on Barro's (1990) endogenous growth model with public good services, Aizenman *et al.* (2007) also show that with effective upper bound on tax revenue due to distortions and imperfect tax

---

[*] International Monetary Fund, Washington, D.C. 20431.

The authors would like to thank Emanuele Baldacci, John Berdell, Carlos Caceres, Giovanni Calligari, Yongsung Chang, Cristina Checherita, Carlo Cottarelli, Rafael Domenech, Balazs Egert, Julio Escolano, Phil Gerson, Atish Ghosh, Alfred Greiner, Fuad Hasanov, William Hauk, Gerhard Illing, John Janssen, Julian Di Giovanni, George Karras, Jun Il Kim, Jun-Kyung Kim, Daniel Leigh, Paolo Mauro, Sandro Modigliano, Carmen Reinhart, Helmut Reisen, Andre Sapir, Jeffrey Wooldridge, Zheng Zhang, and seminar participants at the IMF, European Commission Conference on Public Debt and Economic Growth, Banca d'Italia Public Finance Workshop, World Bank-KDI School International Conference on Fiscal Policy and Management, Beijing Forum, Yonsei University, Korea Development Institute (KDI), and Korea Institute of Public Finance (KIPF) for helpful comments and discussions. Julia Guerreiro provided excellent research assistance. The opinions expressed in the paper are those of the authors and should not be held to represent those of the IMF or its Member countries. Correspondence: jwoo@imf.org or mkumar@imf.org

[1] A notable partial exception is Reinhart and Rogoff (2010) who examine economic growth and inflation at different levels of government debt in advanced and emerging economies based on long historical data series. However, their study only considers correlations between debt and growth, and does not take into account other determinants of growth via econometric analysis as well as issues such as reverse causality (*i.e.*, low growth can lead to large public debt). After the publication of the working paper version of our paper (Kumar and Woo, 2010), subsequent studies by others examined much smaller samples of countries and obtained the results that are quantitatively similar to ours: Checherita and Rother (2010) in 12 Euro economies for 1970-2008 and Cecchetti *et al.* (2011) in 18 OECD countries for 1980-2006. However, they mostly focus on identifying the threshold level of debt above which debt becomes harmful to growth. They do not explore the channels through which debt can affect growth nor consider the interaction between growth, debt, and a country's economic and financial position *vis-à-vis* the rest of the world (or currency composition of debt), not to mention the lack of rigorous discussion on related econometric issues.

enforcement, an increase in (initial) debt lowers the productive government spending, which reduces the return to capital and growth subsequently.

High debt may adversely affect medium- and long-run growth via several channels: high public debt can adversely affect capital accumulation and growth via higher long-term interest rates (Gale and Orzag, 2003; Baldacci and Kumar, 2010), higher future distortionary taxation (Barro, 1979; Dotsey, 1994) and lower future public infrastructure spending (Aizenmann *et al.*, 2007), higher inflation (Sargent and Wallace 1981; Barro 1995; Cochrane 2011), and greater uncertainty about prospects and policies. In more extreme cases of a debt crisis, by triggering a banking or currency crisis, these effects can be magnified (Burnside *et al.*, 2001; Hemming *et al.*, 2003). Also, high debt is likely to constrain the scope for countercyclical fiscal policies, which may result in higher volatility and further lower growth (Aghion and Kharroubi, 2007; Woo, 2009).

The purpose of this paper is to examine empirically the effects of high public debt on economic growth. To our knowledge, this paper presents the first econometric evidence on the impact of *initial* high public debt on *subsequent* growth of real GDP per capita in a panel of advanced and emerging economies for the period of 1970-2008 by carefully applying various econometric techniques. Here it is worth emphasizing that the paper uses *initial* level of government debt to examine the impact on *subsequent* growth over the next five to twenty years (or longer) so that it avoids reverse causality. Evidence strongly suggests an inverse relationship between initial debt and subsequent growth, controlling for other determinants of growth: on average, a 10 percentage point increase in the initial debt-to-GDP ratio is associated with a slowdown in real per capita GDP growth of around 0.2 percentage points per year, with the impact being somewhat smaller in advanced economies. This order of magnitude is robust to various specifications, estimation methods, samples and periods. There is some evidence of non-linearity with higher levels of initial debt (above around 90 per cent of GDP) having more significantly negative effects on subsequent growth.

Moreover, we find that the impact on growth of initial debt is conditional on a country's economic and financial position *vis-à-vis* the rest of the world and that the currency composition of public debt matters. The adverse impact of debt on growth is larger when the net foreign asset (NFA) position is low or the portion of foreign-currency denominated debt as a share of total public debt is high. Growth accounting exercises imply that the adverse effect largely reflects a slowdown in labor productivity growth mainly due to reduced investment and slower growth of capital stock, rather than through slower growth of TFP or human capital. Additional evidence on the impact of initial debt on subsequent investment renders strong support to this conclusion. We conduct extensive robustness checks. The results are robust to a number of alternative specifications, which control for the variables usually identified as the main determinants of economic growth (Sala-i-Martín *et al.*, 2004), as well as to different samples and periods. In particular, we carefully address a variety of econometric issues including reverse causality, endogeneity, and outliers.

Our paper is related to a few studies that have looked at the impact of *external* (public and private) debt on economic growth *exclusively* in the context of *low income economies*. Most of these studies were motivated by the "debt overhang" hypothesis – a situation where a country's debt service burden is so heavy that a large portion of output accrues to foreign lenders and consequently creates disincentives to invest (Krugman, 1988; Sachs, 1989). Imbs and Rancière (2009) and Pattillo *et al.* (2002, 2004) find a non-linear effect of external debt on growth: that is, a negative and significant impact on growth at high debt levels (typically, over 60 per cent of GDP), but an insignificant impact at low debt levels. Besides the differences in estimation strategies, however, we examine the growth impact of *public debt* in the context of *advanced (and emerging)* *economies* that is largely domestic and denominated in domestic currency,[2] which may have different implications for the magnitude of growth impact and the operating channel(s), compared to those of external debt in the context of low income countries.

The rest of the paper is organized as follows: Section 2 briefly describes data and some stylized facts relating to public debt and growth; Section 3 discusses a number of methodological issues and estimation strategy, and then presents the main panel regression results on the relationship between debt and growth, followed by Section 4 Growth Accounting. Section 5 concludes. Appendixes 1-3 provide additional discussion regarding country sample, data sources and growth accounting.

2 Data and stylized facts

Data for the key variables such as GDP, population, investment, and government size are obtained primarily from the latest version 7.0 of Penn World Table (Heston *et al.*, 2011). Fiscal data including government debt are primarily from the IMF's World Economic Outlook database, and other variables are from World Bank's World Development Indicators, Barro and Lee (2011). The availability of data on public debt and other variables included in the regression dictated the sample size: the main analysis is based on a panel of 38 advanced and emerging economies with a population of over 5 million for the period 1970-2008, while we also present the results using the full sample of 79 countries (including advanced, emerging, and developing countries) without imposing a population size restriction (see Appendices 1-2 for the country list and data sources).

Some stylized facts: First, data on government debt and growth clearly show that there is a negative correlation between *initial* government debt and *subsequent* growth of real per capita GDP. Figure 1 shows a scatter plot of initial debt against subsequent growth of real per capita GDP over five-year periods in the sample of countries with population of over 5 million. According to the OLS fitted line, the coefficient of initial debt is –0.024. Taken at face value (*i.e.*, ignoring the potential endogeneity problem, and not controlling for other growth determinants), it suggests that a 10 percentage point increase in initial debt-to-GDP ratio is associated with a subsequent slowdown in per capita GDP growth of 0.24 percentage points. At shown below, this magnitude turns out to be surprisingly consistent with that obtained using robust econometric analysis. Similarly, *initial* debt is negatively associated with both *subsequent* growth of capital per worker (Figure 2) and domestic investment over 5-year periods (Figure 3).

Second, the *subsequent* growth rate of per capita GDP over five-year periods during high *initial* debt episodes (above 90 per cent of GDP) is on average lower than that during low *initial* debt episodes (below 30 per cent of GDP) across various groups of countries (Figure 4). In advanced economies, the difference in the average growth rates between low initial debt and high initial debt episodes is 0.9 percentage points; in emerging economies, it is more than twice that (1.7 percentage points). This pattern is consistent with econometric results discussed later. Similarly, the average growth differential in G7 countries between low and high initial debt periods is 1.7 percentage points. In the full sample (including developing countries), the growth differential is 2.8 percentage points. (See Appendix Table 10 for summary statistics on average growth rates of real GDP per capita, output per worker, TFP, capital stock per worker, and average levels of domestic investment at different levels of initial government debt for various country groupings).[3]

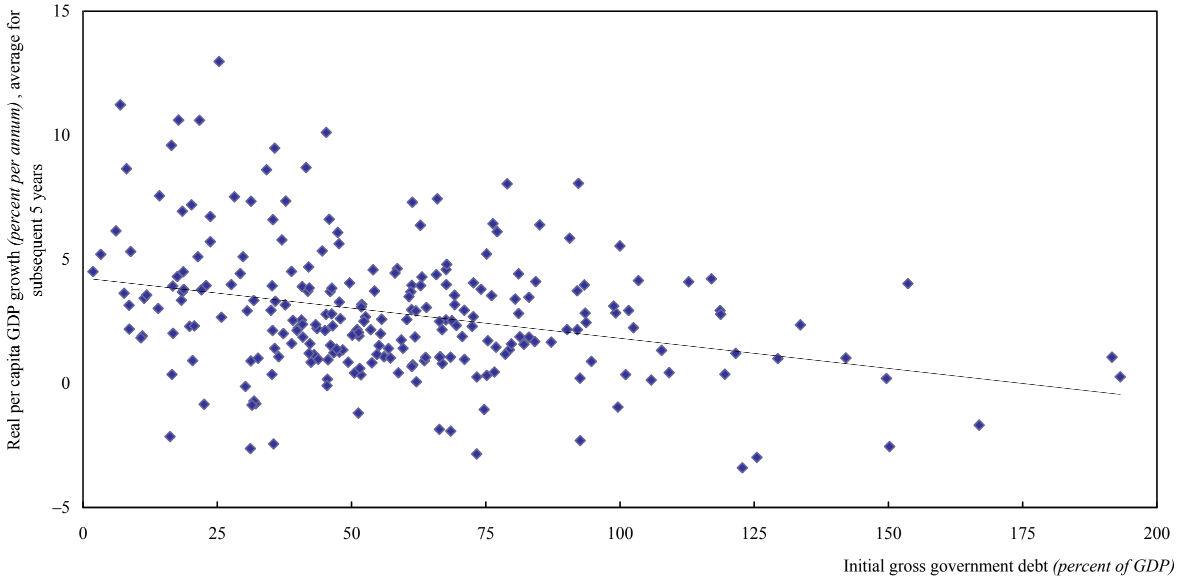
---

[2] This is not only true of advanced economies throughout the sample period, but also of emerging economies in the recent decades during which the portion of domestic-currency denominated debt has been increasing sharply.

[3] Also, high *initial* government debt levels at the start of recession are associated with a slower *subsequent* recovery and longer duration of recovery. See Woo *et al.* (2012) for details.

Figure 1

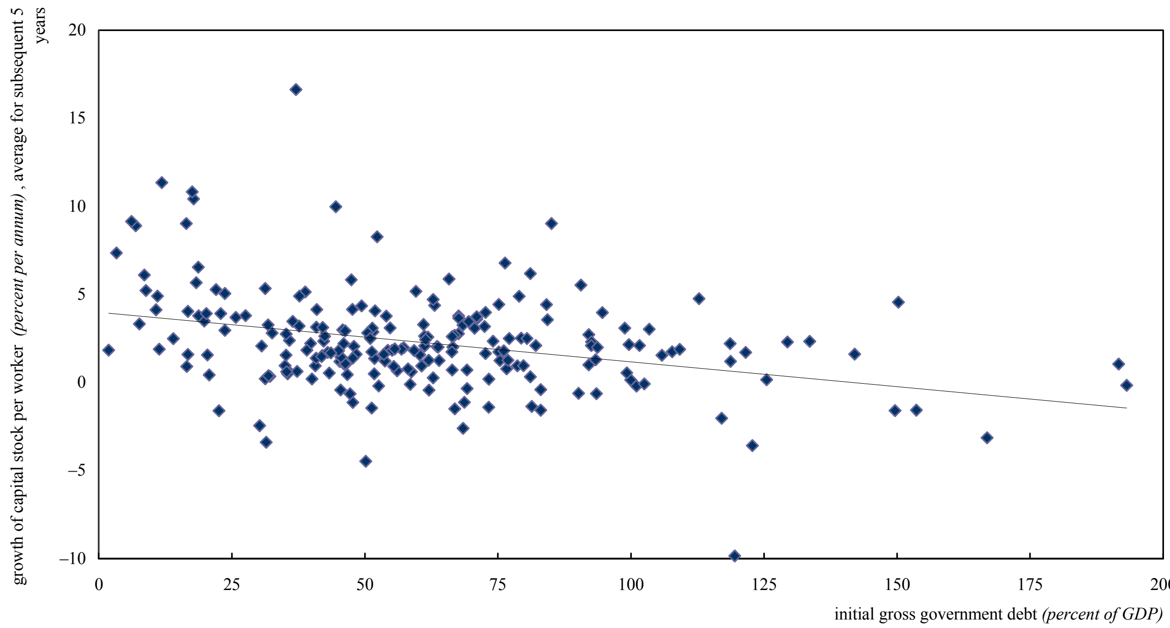
Initial Government Debt and Subsequent Growth of per Capita Real GDP Over Five-year Periods



Fitted line: Growth =4.24–0.024*Initial debt, where the initial debt coefficient is significant at 1 per cent.
Source: Authors' calculation.

Figure 2

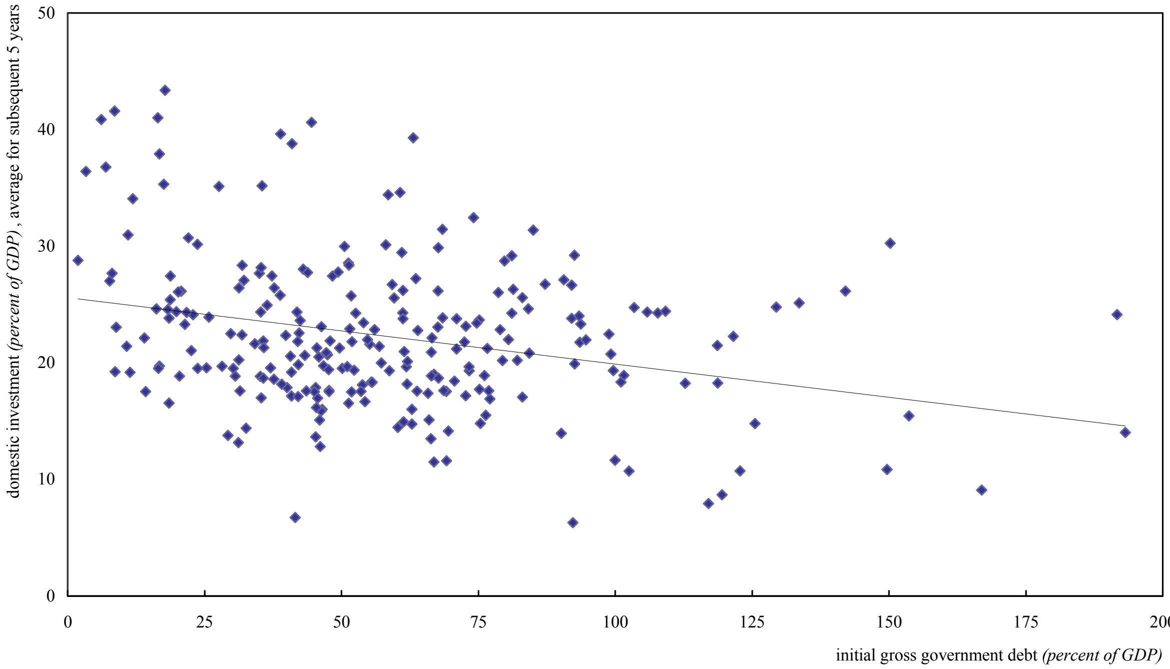
Initial Government Debt and Subsequent Growth of Capital Stock per Worker Over Five-Year Periods



Fitted line: Growth of capital per worker=3.99–0.028*Initial debt, where the debt coefficient is significant at 1 per cent.
Source: Authors' calculation.

Figure 3

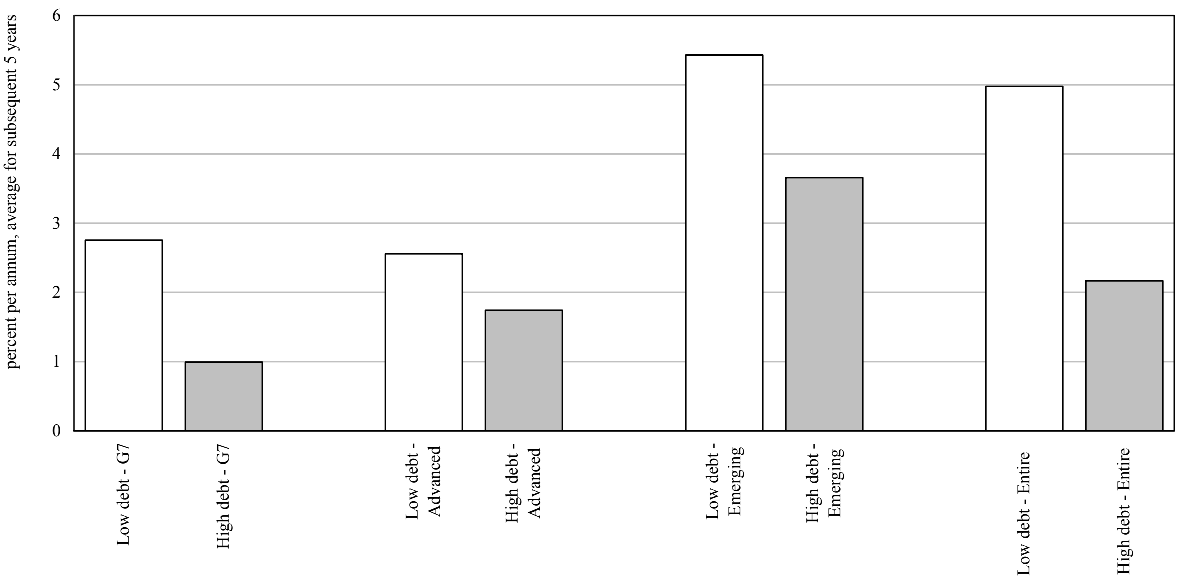
Initial Government Debt and Subsequent Domestic Investment over Five-Year Periods



Fitted line: Investment=25.6–0.057*Initial debt, where the debt coefficient is significant at 1 per cent.
Source: Authors' calculation.

Figure 4

Subsequent Growth of Real GDP per capita
Between High and Low Initial Government Debt Episodes
(low debt <30% of GDP and high debt>90% of GDP)



Source: Authors' calculation.

3 Econometric analysis

3.1 Model specification

The formal analysis focuses on the medium/longer-run relationship between initial government debt and subsequent economic growth, while exploiting both cross-sectional and time-series dimensions of the data. Our panel spans 39 years from 1970 to 2008, and comprises eight non-overlapping five-year periods (1970-74, 1975-79, … , 2000-04, 2005-08), except for the last period spanning four years. In addition, cross-country OLS regressions are estimated for longer time periods – for example, two or three decades (see Appendix Tables 11-12 for the results).

The baseline panel regression specification is as follows:

$$y_{i,t} - y_{i,t-\tau} = \alpha y_{i,t-\tau} + \boldsymbol{X}_{i,t-\tau}\beta + \gamma Z_{i,t-\tau} + \eta t + \nu_i + \varepsilon_{i,t} \tag{1}$$

where a period is a five-year time interval (*i.e.*, τ=4); t denotes the end of a period and $t-\tau$ denotes the beginning of that period; i denotes country; y is the logarithm of real per capita GDP; ν_i is the country-specific fixed effect; ηt is the time-fixed effect; $\varepsilon_{i,t}$ is an unobservable error term; $\boldsymbol{X}_{i,t-\tau}$ is a vector of economic and financial variables; $Z_{i,t-\tau}$ is the initial government debt (in percent of GDP).[4]

A core set of explanatory variables that have been shown to be consistently associated with growth in the literature is fully taken into account.[5] The variables X in the baseline specification are as follows: (i) initial level of real GDP per capita, to capture the catching-up process; (ii) human capital, to reflect the notion that countries with an abundance of it are more likely to have a greater ability to attract investors, absorb ideas from the rest of the world, and engage in innovation activities (Grossman and Helpman, 1991). As a proxy for human capital, we use the log of average years of secondary schooling in the population over age 15 in the initial year, taken from Barro and Lee (2011); (iii) initial government size (as measured by government consumption share of GDP) is also included, in the light of the robust results obtained by Sala-i-Martín *et al.* (2004);[6] (iv) initial trade openness (sum of export and import as a percent of GDP); (v) initial financial market depth (liquid liabilities as a percent of GDP); (vi) initial inflation as measured by CPI inflation (to be precise, logarithm of (1+inflation rate)); (vii) terms of trade growth rates (averaged over each time period); (viii) a measure of banking crisis incidence is also included (based on Reinhart and Reinhart, 2008), reflecting Reinhart and Rogoff's (2009) finding that banking crises are typically accompanied by large increases in government debt. At the same time, banking crises typically result in slow growth; (ix) fiscal deficit is included to take into account the finding that fiscal deficits are negatively associated with longer-run growth (see Fischer, 1993; Baldacci *et al.*, 2004).

To check the robustness of results, parsimonious specifications are tried and additional variables also considered, such as population (a proxy of country size), aged-dependency ratio (a

---

4 To be precise, the average growth rate of real per capita GDP per year over the period $t-\tau$ and t is $(y_{i,t} - y_{i,t-\tau})/\tau$, which is actually used in the empirical application of equation (1). All the explanatory variables in $\mathbf{X}_{i,t-\tau}$ are measured at the beginning of period, except for the terms of trade growth, incidences of banking crisis, and fiscal deficit that are measured over the period $t-\tau$ and t.

5 In particular, the findings of Sala-i-Martín *et al.* (2004) and Sala-i-Martín (1997) are closely followed in selecting the core set of growth determinants.

6 Also, it can be motivated by a consideration of fiscal sustainability. Huang and Xie (2008) derive a fiscal sustainability frontier in an endogenous growth framework, and show that higher levels of government spending reduce the sustainable level of government debt. This implies that estimating a threshold effect on growth based on a widely used single-dimensional perspective of fiscal sustainability such as debt in excess of a particular level may be difficult. What matters is the ability to finance any given level of debt, which in part depends on the availability of savings and the preferences of the savers. Related, Woo (2003) finds that financial market depth is one of the robust determinants of public deficits for various estimation techniques and extensive robustness checks including an extreme-bounds analysis. Thus, a measure of financial depth is included in the baseline regression.

proxy for population aging), investment,[7] fiscal spending volatility, urbanization, private saving, and checks and balances or constraints on executive decision-making (as a proxy of durable institutionalized constraints; see Glaeser *et al.*, 2004).

In addition to taking into account the "core set" of growth determinants which are mostly embodied in the initial conditions, it is worth emphasizing that our estimation uses *initial* level of debt to examine the impact on *subsequent* growth over the next five to two decades (or longer) and thereby avoid the reverse causality problem. Reverse causality may not be a trivial issue as slower economic growth can lead to high debt buildup, rather than high debt lowering growth.[8] However, most of other studies (for example, Checherita and Rother, 2010; Patillo *et al.*, 2002, 2004) have run regressions of growth on the contemporaneous debt ratios, compounding the potential reverse causality problem.

3.2 Sources of bias and estimation strategies

There are a number of sources of biases that can cause inconsistent estimates of the coefficients in panel growth regressions.[9] Yet, each of the estimators involves some trade-off: estimators that may seem attractive to address a specific econometric problem can lead to a different type of bias. For example, when an omitted variables bias coexists with measurement errors that are likely in the cross-country data, dealing with the first problem may exacerbate the second. With this in mind, we employ a variety of estimation techniques, such as pooled OLS, robust regression, between estimator (BE), fixed effects (FE) panel regression, and system GMM (SGMM) dynamic panel regression (Blundell and Bond, 1998). Speaking of the important sources of biases, the first is the omitted-variables bias (so-called heterogeneity bias) resulting from possible correlation between country-specific fixed effects (ν_i) and the regressors, affecting the consistency of pooled OLS and BE (between estimator) estimates. The second is the endogeneity problem due to potential correlation between the regressors and the error term, which would affect the consistency of pooled OLS, BE and FE. Specific to dynamic panels, there is a dynamic panel bias which will make FE estimates inconsistent.[10] The third is classical measurement errors (errors in variables) in the independent variables, which affects the consistency of pooled OLS, BE, and FE estimator, although the bias tends to be exacerbated in FE and moderated in BE.

Specifically, the BE estimator (which applies the OLS to a single cross-section of variables averaged across time periods) tends to reduce the extent of measurement error via time averaging of the regressors, but does not deal with the omitted-variables bias; pooled OLS and BE suffer from both heterogeneity bias and measurement errors but will reduce the heterogeneity bias because other things equal, measurement errors tend to reduce the correlation between the regressors and the country fixed effects; FE addresses the problem of the omitted-variables bias via controlling for

---

7 The *proximate* causes of growth, such as investment or capital per worker, are not included in the core set of growth determinants, but are examined in the growth accounting exercises instead. Nonetheless, we check whether including investment in the regression changes the estimated coefficients of initial government debt.

8 Easterly (2001) argues that slow growth contributed to debt explosion in the developing countries in 1980s. However, Imbs and Rancière's (2009) findings contradict Easterly's argument in an event study of external debt: investment actually builds up *prior to* the onset of debt overhang, which argues against the possibility that an investment slump predates the overhang and explains the debt build-up. Related, Reinhart *et al.* (2012) find that public debt overhang episodes are lasting long (typically for more than a decade), and thus refute the view that the negative association between public debt and growth is caused mainly by debt buildups during recessions.

9 See Durlauf *et al.* (2005) for more details on econometric issues in the empirical growth literature.

10 To see this more clearly, one can rewrite the equation (1) as $y_{i,t} = (1+\alpha)y_{i,t-\tau} + \mathbf{X}_{i,t-\tau}\beta + \gamma Z_{i,t-\tau} + \eta_t + \nu_i + \varepsilon_{i,t}$. The endogeneity bias (often called dynamic panel bias) arises due to inevitable correlation between $y_{i,t-\tau}$ and ν_i in the presence of lagged dependent variable because $y_{i,t-\tau}$ is endogenous to the fixed effects (ν_i) in the error term. In the FE, the fixed effects (ν_i) are eliminated via within-transformation, but there is now a correlation between the transformed lagged dependent variable and the transformed error term, causing the FE to be inconsistent and biased downward.

fixed-effects, but tends to exacerbate the measurement error problem, relative to BE and OLS. This measurement error bias under FE tends to get even worse when the explanatory variables are more time-persistent than the errors in the measurement (Hauk and Wacziarg, 2009).[11] Furthermore, in the dynamic panel setting, the within-transformation in the estimation process of FE introduces a correlation between transformed lagged dependent variable and transformed error, which also makes FE inconsistent. Theoretically, the dynamic panel GMM estimator addresses a variety of biases such as the omitted-variables bias, endogeneity, and measurement errors (as long as instruments are uncorrelated with the errors in measurement, for example, if they are white noise as in the classical case), but it may be subject to a weak instruments problem (Roodman, 2009; Bazzi and Clemens, 2009). While the SGMM that is used in this paper is generally more robust to weak instruments than the difference GMM, it can still suffer from weak instrument biases.[12] In sum, it is difficult to see which estimator yields the smaller *total bias* in the presence of various sources of bias a priori.

However, an important conclusion from the Monte Carlo study of growth regressions by Hauk and Wacziarg (2009) is that the BE performs the best among the four estimators (pooled OLS, BE, FE, and difference GMM) in terms of the extent of *total bias* on each of the estimated coefficients in the presence of both potential heterogeneity bias and a variety of measurement errors.[13] Therefore, the BE and SGMM estimators are the preferred estimation techniques in this paper, while we utilize the other techniques also.

As further robustness checks, we also run a single cross-country regression of the type that is most commonly used in the empirical growth literature for longer time periods. This helps address the issue that the five-year time interval in the panel may not be long enough to smooth out short-term business cycle fluctuations. The cross-country regression results (including the order of magnitude of the coefficients) however turn out to be broadly similar to those from panel regressions. On the other hand, the least squares estimates tend to be sensitive to outliers, either observations with unusually large errors or influential observations with unusual values of explanatory variables (often called leverage points). In an extensive evaluation of growth regressions in relation to macroeconomic policy variables, Easterly (2005) argues that some of the large effects on growth of a policy variable in the earlier empirical studies are often caused by outliers that represent "extremely bad" policies. Thus, to ensure that our results are not unduly driven by outliers, robust regression is also implemented.[14]

---

11 Intuitively, the within-transformation (*i.e.*, demeaning) under FE may exacerbate the measurement error bias by decreasing the signal-to-noise ratio (Grilliches and Hausman, 1986).

12 A standard test of weak instruments in dynamic panel GMM regressions does not currently exist (Bazzi and Clemens, 2009). See Stock *et al.* (2002) on why the weak instrument diagnostics for linear IV regression do not carry over to the more general setting of GMM.

13 The BE estimator applies the OLS to perform estimating of the following equation:

$$\overline{y_{i,} - y_{i,-1}} = \alpha \overline{y_{i,-1}} + \overline{X_{i,-1}}\beta + \gamma \overline{Z_{i,-1}} + v_i + \bar{\varepsilon}_i$$

where the upper bar indicates the average of each variable across time periods (up to eight periods), for example, $\overline{X_{i,-1}} = \sum_t X_{i,t-\tau} / T_i$. Thus, time-fixed effects are not appropriate and suppressed by the BE. As one can see, the BE estimator does not correspond to the cross-sectional estimator most commonly used in the literature in which in which the dependent and explanatory variables are averaged, say, over 1970-2008, except for the initial income level in 1970.

14 It is essentially an iterated re-weighted least squares regression in which the outliers are dropped (if Cook's distance is greater than 1) and the observations with large absolute residuals are down-weighted.

3.3 Basic results

The main results for advanced and emerging economies are presented in Table 1. Columns 1-4 show that the coefficients of initial debt are negative and are significant at the 1-5 per cent levels, with their values ranging from –0.015 to –0.030 across the various estimation techniques.[15] The BE regression in column 1 suggests that a 10 percentage points of GDP increase in initial debt is associated with a slowdown in subsequent growth in real GDP per capita of around 0.25 percentage points per year. The pooled OLS and FE in columns 2 and 3 yield results similar to that of the BE regression, although their estimates of initial debt coefficient become somewhat smaller (around –0.02). The SGMM estimate of initial debt coefficient is also in a similar range (–0.03) and significant at the 1 per cent level.

The coefficients on other explanatory variables (initial income per capita, average years of schooling, financial market development, inflation, banking crisis, and fiscal deficit) are of the expected sign and mostly significant at conventional levels across various estimation techniques. The OLS and FE estimators are likely to be biased in the opposite direction in the context of lagged dependent variables in short panels, with OLS biased upwards, and FE downwards. The *consistent* GMM estimator should lie between the two (Bond 2002). In the growth regressions, this means that the OLS understates the convergence rate (reflected by the coefficient of initial income per capita), while the FE estimator overstates it. Consistent with this reasoning, the OLS coefficient of initial real per capita GDP is –1.88, whereas the FE coefficient is –3.92. The SGMM coefficient of the initial income per capita (–2.34) is between those two estimates, indicating that the reported SGMM estimate in column 4 is likely to be a *consistent* parameter estimate of the convergence rate.

Consistency of the SGMM estimator depends on the validity of the instruments. We consider two specification tests, suggested by Arellano and Bover (1995) and Blunedell and Bond (1998). The first is a Hansen J-test of over-identifying restrictions, which tests the overall validity of the instruments by analyzing the sample analog of the moment conditions used in the estimation process. This indicates that we cannot reject the null hypothesis that the full set of orthogonality conditions are valid (p-value=0.65).[16] The second test examines the hypothesis that the error term $\varepsilon_{i,t}$ is not serially correlated. We use an Arellano-Bond test for autocorrelation, and find that we cannot reject the null hypothesis of no second-order serial correlation in the first-differenced error terms (p-value=0.24).[17]

The regressions in columns 2-4 do not include the time-fixed effects. It is possible that global factors can simultaneously affect both domestic growth and public debt which may bias the results toward finding a stronger relationship between debt and growth. At the same time, however, as global factors can be correlated with domestic fiscal or economic variables, one can expect that the inclusion of time-fixed effects may understate the estimated effects of these variables. Columns 5-7 include time-fixed effects in the regression to allow for global factors. The pooled OLS and SGMM coefficients of initial debt remain significant at 5-10 per cent, and the size of

[15] In the OLS and robust regressions, dummies for OECD, Asia, Latin America, and sub-Saharan Africa are included. Results for robust regressions are similar to those of pooled OLS, so they are not reported to save space.

[16] Importantly, the difference-in-Hansen tests of exogeneity of instrument subsets do not reject the null hypothesis that the instrument subsets for the level equations are orthogonal to the error (p-value=0.34), that is, the assumption that lagged differences of endogenous explanatory variables that are being used as instruments in levels is uncorrelated with the errors. This is the additional restriction that needs to be satisfied for the SGMM estimator.

[17] The dynamic panel GMM can generate too many instruments, which may overfit endogenous variables and run a risk of a weak-instruments bias (Roodman, 2009; Bazzi and Clemens, 2009). Given that, one recommendation when faced with a weak-instrument problem is to be parsimonious in the choice of instruments. Roodman (2009) suggests restricting the number of lagged levels used in the instrument matrix or collapsing the instrument matrix or combining the two. Some studies including Beck and Levine (2004) use the technique of collapsing instrument matrix. The reported SGMM results in our paper are obtained by combining the "collapsed" instrument matrix with lag limits.

Table 1

Baseline Panel Regression – Growth and Initial Government Debt, 1970-2008 (Five-year Period Panel)
Sample: Advanced and Emerging Economies (with Population of Over 5 Million)
(dependent variable: real per capita GDP growth)

Explanatory Variables	(1) BE	(2) Pooled OLS	(3) FE	(4) SGMM	(5) Pooled OLS	(6) FE	(7) SGMM
Initial real GDP per capita	–2.123***	–1.877**	–3.924***	–2.336***	–1.707**	–4.744**	–2.229***
	(–5.02)	*(–2.54)*	*(–2.74)*	*(–3.47)*	*(–2.14)*	*(–2.36)*	*(–2.95)*
Initial years of schooling	4.813***	3.143**	3.388	4.508*	3.136**	2.394	3.161
	(3.94)	*(2.57)*	*(1.64)*	*(1.93)*	*(2.55)*	*(1.07)*	*(1.55)*
Initial inflation rate	2.151	–2.100***	–2.630***	–2.666**	–2.457***	–2.454***	–2.678**
	(0.82)	*(–3.32)*	*(–5.38)*	*(–2.49)*	*(–3.21)*	*(–5.81)*	*(–2.05)*
Initial government size	0.109**	0.109**	0.147	0.162	0.111**	0.055	0.138
	(2.06)	*(2.43)*	*(1.68)*	*(1.36)*	*(2.38)*	*(0.70)*	*(1.23)*
Initial trade openness	–0.002	–0.004	0.023*	–0.013**	–0.005	0.023	–0.004
	(–0.43)	*(–0.78)*	*(1.73)*	*(–2.03)*	*(–1.11)*	*(1.57)*	*(–0.57)*
Initial financial depth	0.022**	0.020**	0.001	0.035***	0.023**	0.006	0.027**
	(2.15)	*(2.13)*	*(0.07)*	*(3.18)*	*(2.50)*	*(0.64)*	*(2.31)*
Terms of trade growth	0.204**	–0.013	0.009	–0.032	–0.017	–0.003	–0.044*
	(2.33)	*(–0.52)*	*(0.33)*	*(–1.14)*	*(–0.70)*	*(–0.13)*	*(–1.97)*
Banking crisis	–1.077	–0.617	–0.638***	–1.033	–0.612*	–0.513*	–1.838
	(–0.61)	*(–1.58)*	*(–2.96)*	*(–1.55)*	*(–1.75)*	*(–1.98)*	*(–1.24)*
Fiscal deficit	0.028	–0.044***	–0.047***	–0.046***	–0.045***	–0.035***	–0.062***
	(0.80)	*(–4.27)*	*(–4.07)*	*(–2.96)*	*(–4.72)*	*(–3.50)*	*(–3.10)*
Initial government debt	–0.025**	–0.022***	–0.015**	–0.030***	–0.018**	–0.004	–0.019*
	(–2.28)	*(–3.29)*	*(–2.17)*	*(–4.14)*	*(–2.34)*	*(–0.67)*	*(–1.89)*
Arellano-Bond AR(2) test *p*-value[1]				0.65			0.45
Hansen *J*-statistics (*p*-value)[2]				0.24			0.29
Number of observations	166	166	166	166	166	166	166
R^2	0.68	0.51	0.39		0.58	0.51	
Time-fixed effects	N/A	No	No	No	Yes	Yes	Yes

Note: Heteroskedasticity and country-specific autocorrelation consistent *t*-statistics are in parentheses. Time dummies are not reported. Levels of significance: *** 1%, ** 5%, * 10%. In the OLS regressions, dummies for OECD, Asia, Latin America, and Sub-Saharan Africa are also included in each regression (not reported to save space). FE refers to the fixed-effects panel regressions and BE is the between estimator. For the dynamic panel estimation, a two-step system GMM (SGMM) with the Windmeijer's finite-sample correction for the two-step covariance matrix.
1) The null hypothesis is that the first-differenced errors exhibit no second-order serial correlation.
2) The null hypothesis is that the instruments used are not correlated with the residuals.

those coefficients is reduced as expected. The estimated effects suggest that a 10 percentage point increase in the initial debt-to-GDP ratio is associated with a slowdown in growth of per capita GDP around 0.2 per cent per year.

In contrast, the FE results on initial debt turn out to be particularly sensitive to whether time-fixed effects are included or not in the regression (compare column 6 with column 3). The FE coefficient of initial debt is now insignificant and reduced to –0.004. It is well known in the literature that the FE can bias toward zero the slope estimates on the determinants of the steady-state level of income – the accumulation and depreciation variables in the Solow model (Islam, 1995). Given that the FE estimator tends to identify parameters on the basis of within-country variation, compared to cross-sectional alternatives such as pooled OLS and BE, it is not surprising that the within-country variation in each of regressors (especially time-persistent variables) is further reduced once time-fixed effects are accounted for.[18] Moreover, the measurement error bias can also be exacerbated under FE. With these caveats, time-fixed effects are included in the remaining regressions.

3.4 Robustness of results

A variety of robustness checks were conducted: First, to account for the possibility that there may have been structural changes over the sample period, including changes in global trend growth or global risk factors, time-fixed effects were included. In addition, we restricted the sample to the second half of the period to check whether there are significant changes in the estimated coefficients. Thus columns 1-4 in Table 2 repeat the same sets of regressions (BE, pooled OLS, FE, and SGMM) for the period of 1990-2008. The results are quite similar to those for the entire period. Except for the FE estimate, the impact of initial debt is significant, ranging from –0.020 to –0.024, indicating that a 10 percentage point increase in initial debt-to-GDP ratio is associated with decline in per capita GDP growth of around 0.2-0.24 per cent per year.

Second, columns 5-8 and 9-12 of Table 2 replicate the regression exercises for 46 advanced and emerging economies and the full sample of 79 countries (46 advanced and emerging economies and 33 developing countries) regardless of the population size for the entire period, respectively. Again, the results are broadly the same as those from the 38 advanced and emerging economies with a population of over 5 million, although the size of the debt coefficients becomes slightly smaller.

Third, Table 3 presents the results based on a parsimonious specification that excludes the fiscal deficit term.[19] The coefficients of initial debt are negative and significant at 1-5 per cent, ranging from –0.014 to –0.026, except for the FE result in which the coefficient of initial debt loses statistical significance (columns 1-4). It is noteworthy that the BE estimates of initial debt coefficient are stable around 0.21 to 0.26 across different samples, periods, and specifications. Using average debt instead of initial debt also yields a similar range of –0.019 to –0.030 for the debt coefficients under BE, OLS and SGMM, which are all significant at 1-10 per cent (columns 5, 6 and 8), except for the FE in column 7.

Fourth, additional variables are considered, such as population size (a proxy of country size), aged-dependency ratio (a proxy of population aging), investment, fiscal volatility, urbanization, and checks and balances or constraints on executive decision-making (as a proxy of durable

[18] With the time-fixed effects included, the coefficients of years of schooling and initial debt are often insignificant under FE in contrast to those under SGMM, as one can see throughout this paper.

[19] Qualitatively similar results are obtained in various parsimonious specifications, such as also dropping a measure of banking crisis and/or financial market depth.

Table 2

Robustness Checks—Time Period and Sample

(dependent variable: real per capita GDP growth)

Explanatory Variables	(1) BE	(2) Pooled	(3) FE	(4) SGMM	(5) BE	(6) Pooled	(7) FE	(8) SGMM	(9) BE	(10) Pooled	(11) FE	(12) SGMM
	Period: 1990-2008 Sample: OECD and Emerging Economies				**Period: 1970-2008 Sample: OECD and Emerging Economies Without Population Size Restriction**				**Period: 1970-2008 Sample: Full Sample (Including Developing Countries) Without Population Size Restriction**			
Initial real GDP per capita	–1.794***	–1.711**	–3.325*	–2.376**	–1.796***	–1.074*	–5.843***	–2.072*	–0.962***	–1.021**	–4.495**	–1.566**
	(–4.67)	*(–2.22)*	*(–1.99)*	*(–2.21)*	*(–4.37)*	*(–1.80)*	*(–3.09)*	*(–1.96)*	*(–2.79)*	*(–2.09)*	*(–2.13)*	*(–2.12)*
Initial years of schooling	3.815***	3.491***	–0.784	3.903	3.768***	1.809*	4.629**	2.956	1.550*	0.887	2.624	2.346*
	(3.35)	*(2.78)*	*(–0.17)*	*(0.92)*	*(3.10)*	*(1.68)*	*(2.56)*	*(0.87)*	*(1.79)*	*(0.98)*	*(1.11)*	*(1.79)*
Initial inflation rate	1.258	–2.918***	–2.308***	–1.717	2.227	–1.201**	–2.262***	–1.112	2.727	0.324	–0.899	–0.251
	(0.51)	*(–3.19)*	*(–4.33)*	*(–1.14)*	*(0.92)*	*(–2.14)*	*(–5.37)*	*(–0.93)*	*(1.14)*	*(0.46)*	*(–1.12)*	*(–0.33)*
Initial government size	0.120**	0.119**	0.074	0.205*	0.030	–0.018	–0.039	–0.180*	–0.020	–0.026	–0.023	–0.092
	(2.41)	*(2.45)*	*(0.68)*	*(1.73)*	*(0.77)*	*(–0.44)*	*(–0.56)*	*(–1.75)*	*(–0.63)*	*(–1.00)*	*(–0.41)*	*(–1.23)*
Initial trade openness	0.001	–0.007	0.030*	–0.006	0.009**	0.003	0.015	0.003	0.003	0.004	0.002	0.000
	(0.19)	*(–1.55)*	*(1.76)*	*(–0.72)*	*(2.38)*	*(0.78)*	*(1.63)*	*(0.24)*	*(0.83)*	*(1.29)*	*(0.15)*	*(0.03)*
Initial financial depth	0.016*	0.027**	0.002	0.032	0.002	0.001	0.007	–0.001	–0.000	–0.004	–0.006	0.006
	(1.71)	*(2.68)*	*(0.13)*	*(1.66)*	*(0.27)*	*(0.07)*	*(0.76)*	*(–0.06)*	*(–0.05)*	*(–0.60)*	*(–0.54)*	*(0.39)*
Terms of trade growth	0.223***	–0.016	–0.018	–0.049	0.187**	–0.001	0.008	–0.046	–0.033	0.028	0.062**	0.024
	(2.79)	*(–0.29)*	*(–0.36)*	*(–0.94)*	*(2.14)*	*(–0.04)*	*(0.31)*	*(–1.03)*	*(–0.64)*	*(0.92)*	*(2.05)*	*(0.74)*
Banking crisis	0.632	–0.358	–0.576	–1.233	–1.445	–0.867**	–0.837***	–1.003	–3.566**	–1.357***	–1.026***	–1.861***
	(0.38)	*(–0.68)*	*(–1.15)*	*(–0.90)*	*(–0.80)*	*(–2.23)*	*(–2.80)*	*(–1.16)*	*(–2.32)*	*(–3.85)*	*(–3.53)*	*(–3.21)*
Fiscal deficit	0.009	–0.055***	–0.046***	–0.057*	0.050*	–0.037***	–0.045***	–0.045**	–0.028**	–0.034***	–0.041***	–0.035**
	(0.27)	*(–4.18)*	*(–2.92)*	*(–1.71)*	*(1.72)*	*(–3.40)*	*(–4.25)*	*(–2.46)*	*(–2.17)*	*(–3.80)*	*(–5.50)*	*(–2.13)*
Initial government debt	–0.024***	–0.020**	–0.008	–0.023*	–0.019*	–0.020**	–0.011*	–0.021*	–0.021***	–0.017***	–0.011*	–0.016*
	(–2.85)	*(–2.26)*	*(–0.65)*	*(–2.02)*	*(–1.94)*	*(–2.62)*	*(–1.78)*	*(–1.74)*	*(–3.22)*	*(–3.31)*	*(–1.66)*	*(–1.83)*
Arellano-Bond AR(2) test *p*-value[1]				0.42				0.59				0.59
Hansen J-statistics (*p*-value)[2]				0.13				0.98				0.36
Number of observations	124	124	124	124	208	208	208	208	297	297	297	297
R^2	0.72	0.61	0.44		0.56	0.44	0.51		0.37	0.36	0.43	
Time-fixed effects	N/A	Yes	Yes	Yes	N/A	Yes	Yes	Yes	N/A	Yes	Yes	Yes

Note: Heteroskedasticity and country-specific autocorrelation consistent *t*-statistics are in parentheses. Time dummies are not reported. Levels of significance: *** 1%, ** 5%, * 10%. In the OLS regressions, dummies for OECD, Asia, Latin America, and Sub-Saharan Africa are also included in each regression (not reported to save space). FE refers to the fixed-effects panel regressions and BE is the between estimator. For the dynamic panel estimation, a two-step system GMM (SGMM) with the Windmeijer's finite-sample correction for the two-step covariance matrix.

1) The null hypothesis is that the first-differenced errors exhibit no second-order serial correlation.

2) The null hypothesis is that the instruments used are not correlated with the residuals.

Table 3

Robustness Checks – Parsimonious Specification: Advanced and Emerging Economies

(dependent variable: real per capita GDP growth)

Explanatory Variables	**(1) BE**	**(2) Pooled OLS**	**(3) FE**	**(4) SGMM**	**(5) BE**	**(6) Pooled OLS**	**(7) FE**	**(8) SGMM**
Initial real GDP per capita	–2.007***	–2.068**	–5.835**	–2.545***	–1.722***	–1.786**	–6.157***	–2.014**
	(–5.08)	*(–2.41)*	*(–2.59)*	*(–3.37)*	*(–4.45)*	*(–2.17)*	*(–3.25)*	*(–2.36)*
Initial years of schooling	4.576***	3.486**	1.404	6.493**	3.393***	2.749**	1.057	3.654*
	(3.89)	*(2.68)*	*(0.51)*	*(2.42)*	*(2.93)*	*(2.25)*	*(0.38)*	*(1.91)*
Initial inflation rate	1.469	–1.276*	–1.692***	–0.683	2.467	–1.376	–2.318*	–4.405
	(0.60)	*(–1.73)*	*(–5.52)*	*(–0.97)*	*(1.20)*	*(–1.30)*	*(–1.79)*	*(–1.49)*
Initial government size	0.117**	0.093**	0.001	0.011	0.094*	0.084*	0.009	0.264
	(2.26)	*(2.03)*	*(0.01)*	*(0.08)*	*(1.88)*	*(2.01)*	*(0.12)*	*(1.14)*
Initial trade openness	–0.004	–0.001	0.038***	0.000	–0.005	–0.002	0.030**	–0.005
	(–0.79)	*(–0.15)*	*(2.83)*	*(0.04)*	*(–1.16)*	*(–0.58)*	*(2.59)*	*(–0.45)*
Initial financial depth	0.024**	0.017*	0.002	0.005	0.024**	0.020**	0.002	0.026
	(2.47)	*(1.98)*	*(0.32)*	*(0.51)*	*(2.61)*	*(2.21)*	*(0.30)*	*(1.38)*
Terms of trade growth	0.169**	0.005	0.003	–0.014	0.006	–0.007	0.021	–0.031
	(2.24)	*(0.15)*	*(0.11)*	*(–0.46)*	*(0.07)*	*(–0.25)*	*(0.67)*	*(–1.06)*
Banking crisis	–0.880	–0.483	–0.402	–1.311	–2.004	–1.199***	–1.208***	–0.614
	(–0.50)	*(–1.21)*	*(–1.48)*	*(–0.85)*	*(–1.35)*	*(–2.74)*	*(–2.97)*	*(–0.42)*
Initial government debt	–0.026**	–0.014**	0.010	–0.014*				
	(–2.39)	*(–2.12)*	*(1.36)*	*(–1.95)*				
Government debt, average					–0.030***	–0.019**	–0.004	–0.023*
					(–2.87)	*(–2.36)*	*(–0.56)*	*(–1.86)*
Arellano-Bond AR(2) test p-value[1]				0.08				0.14
Hansen J-statistics (p-value)[2]				0.33				0.27
Number of observations	166	166	166	166	181	181	181	181
R^2	0.67	0.52	0.45		0.59	0.49	0.47	
Time-fixed effects	N/A	Yes	Yes	Yes	N/A	Yes	Yes	Yes

Note: Heteroskedasticity and country-specific autocorrelation consistent t-statistics are in parentheses. Time dummies are not reported. Levels of significance: *** 1%, ** 5%, * 10%. In the OLS regressions, dummies for OECD, Asia, Latin America, and Sub-Saharan Africa are also included in each regression (not reported to save space). FE refers to the fixed-effects panel regressions and BE is the between estimator. For the dynamic panel estimation, a two-step system GMM (SGMM) with the Windmeijer's finite-sample correction for the two-step covariance matrix.

1) The null hypothesis is that the first-differenced errors exhibit no second-order serial correlation.

2) The null hypothesis is that the instruments used are not correlated with the residuals.

institutional quality; see Glaeser *et al.*, 2004). The results do not change appreciably (Table 4). Columns 1-4 add the log of initial population to the baseline specification: the coefficients of initial debt are negative and significant at 5 per cent level except for the FE in column 3 in which it is insignificant. According to the BE, OLS, and SGMM, the estimated effects of initial debt suggest that a 10 percentage point increase of initial debt-to-GDP ratio is associated with slowdown in growth of per capita GDP of around 0.18 to 0.25 per cent per year. In contrast, the coefficients of population size are insignificant except for FE in which it becomes significant.

The results when initial domestic investment (as a percent of GDP) is added to the baseline specification are shown in columns 5-8 of Table 4. Under OLS and SGMM, the coefficients of initial debt ratio are significant at 5 per cent level, whereas the coefficients of investment are of the expected positive sign and significant at 5 per cent under BE and OLS. Under SGMM, the investment coefficient becomes insignificant, and its coefficient size is slightly smaller than that under BE. However, the FE estimates of the coefficients of initial debt and initial investment are not only insignificant, but the coefficient of initial investment even changes its sign to negative.

In columns 9-12 of Table 4, we include a measure of fiscal spending volatility (as measured by a logarithm of standard deviations of annual growth in real general government expenditures) in the regressions. Recently, Fatás and Mihov (2003) have argued that excessive discretionary fiscal policies that are not related to dealing with business cycle fluctuations can lead to higher output volatility and lower growth.[20] At the same time, this excessive fiscal activism may lead to a large debt buildup. According to this view, excessive fiscal discretion may be an underlying force behind the negative relation between government debt and growth. If this is so, one may expect the coefficient of initial debt in the growth regression to become weaken or at least to get smaller in its absolute value, once the fiscal volatility term is included in the regression. However, our analysis does not find evidence in support of this view.[21] The coefficients of fiscal volatility are insignificant, and even change sign across different estimations. By contrast, the coefficients of initial debt remain largely significant, and the size of estimated coefficients is quite similar to that in the baseline regressions.

Finally, we run a single cross-country regression of the type that is most commonly used in the empirical growth literature for longer time periods. The cross-country regression results are presented in Appendix Tables 11 and 12. They are remarkably similar to the above panel regression results. In particular, the size of estimated initial debt coefficients which is around –0.02~–0.03 is remarkably similar to that found in the baseline panel regression.

3.5 Non-linearities and differences between advanced and emerging economies

To explore potential non-linearities, Table 5 (columns 1-4) shows regressions that include the interaction terms between initial debt and dummy variables for three ranges of initial debt: Dum_30 for low debt (below 30 per cent of GDP); Dum_30-90 for medium debt (30-90 per cent of GDP); and Dum_90 for high debt (over 90 per cent of GDP). The coefficients of low initial debt (*i.e.*, initial debt*Dum_30) are all insignificant and of the positive sign, which seems to suggest that

---

20 Ideally, the measure of fiscal policy volatility (that is, excessive discretionary policy changes undertaken for reasons other than smoothing out business cycle fluctuations) can be constructed in a more sophisticated manner. For example, it can be obtained as a standard deviation of the residuals from time-series regression of government spending growth on macroeconomic variables such as output growth and inflation. Given such a short time duration of each period, it is impossible to run a meaningful time-series regression for each five-year period. However, the qualitative behavior of such a measure of fiscal volatility is very similar to that of a crude measure of fiscal volatility as used in this paper (Woo, 2009).

21 While there is significant evidence that fiscal volatility is positively correlated with output volatility and that output volatility is negatively associated with growth (Fatás and Mihov, 2003; Ramey and Ramey, 1995), there is little analysis in the literature regarding the relationship between government debt and fiscal behavior such as fiscal volatility or fiscal cyclicality.

Table 4

Robustness Checks – Additional Variables: Advanced and Emerging Economies

(dependent variable: real per capita GDP growth)

Explanatory Variables	(1) BE	(2) Pooled OLS	(3) FE	(4) SGMM	(5) BE	(6) Pooled OLS	(7) FE	(8) SGMM	(9) BE	(10) Pooled OLS	(11) FE	(12) SGMM
Initial real GDP per capita	–1.798***	–1.581**	–4.361***	–2.478**	–2.412***	–2.506***	–3.832	–2.909***	–2.110***	–1.737**	–4.762**	–1.830**
	(–3.39)	*(–2.14)*	*(–2.76)*	*(–2.43)*	*(–6.07)*	*(–2.82)*	*(–1.64)*	*(–2.74)*	*(–4.32)*	*(–2.17)*	*(–2.36)*	*(–2.53)*
Initial years of schooling	4.611***	2.994**	–1.364	6.483*	4.385***	3.729***	2.057	5.403	4.818***	3.037**	2.358	3.173
	(3.73)	*(2.52)*	*(–0.48)*	*(1.68)*	*(3.93)*	*(3.08)*	*(0.94)*	*(1.60)*	*(3.86)*	*(2.49)*	*(1.08)*	*(1.14)*
Initial inflation rate	2.481	–2.313***	–2.642***	–5.741	2.099	–2.659***	–2.484***	–5.742	2.140	–2.351***	–2.444***	–3.296*
	(0.94)	*(–3.15)*	*(–5.48)*	*(–0.90)*	*(0.89)*	*(–3.53)*	*(–5.54)*	*(–0.94)*	*(0.80)*	*(–2.94)*	*(–5.15)*	*(–1.68)*
Initial government size	0.094*	0.109**	0.079	0.251	0.128**	0.119**	–0.010	0.174	0.110*	0.108**	0.055	0.245**
	(1.72)	*(2.44)*	*(0.91)*	*(0.95)*	*(2.64)*	*(2.71)*	*(–0.14)*	*(1.08)*	*(1.98)*	*(2.31)*	*(0.70)*	*(2.17)*
Initial trade openness	0.002	–0.001	0.042***	–0.009	0.001	–0.003	0.020	–0.012	–0.003	–0.004	0.023	0.002
	(0.34)	*(–0.21)*	*(3.08)*	*(–0.75)*	*(0.15)*	*(–0.95)*	*(1.18)*	*(–1.06)*	*(–0.39)*	*(–0.87)*	*(1.49)*	*(0.27)*
Initial financial depth	0.015	0.021**	0.007	0.019	0.021**	0.024***	0.005	0.025	0.022*	0.022**	0.005	0.019
	(1.20)	*(2.38)*	*(0.88)*	*(0.89)*	*(2.32)*	*(3.12)*	*(0.53)*	*(1.48)*	*(1.97)*	*(2.38)*	*(0.64)*	*(1.46)*
Terms of trade growth	0.219**	–0.014	–0.012	–0.028	0.300***	–0.011	–0.005	–0.026	0.205**	–0.017	–0.003	–0.048**
	(2.47)	*(–0.56)*	*(–0.50)*	*(–0.63)*	*(3.45)*	*(–0.47)*	*(–0.20)*	*(–0.45)*	*(2.24)*	*(–0.65)*	*(–0.13)*	*(–2.35)*
Fiscal deficit	0.039	–0.043***	–0.032***	–0.041*	0.015	–0.047***	–0.038***	–0.064	0.028	–0.044***	–0.035***	–0.043***
	(1.07)	*(–4.68)*	*(–3.83)*	*(–1.71)*	*(0.45)*	*(–5.21)*	*(–3.68)*	*(–1.58)*	*(0.76)*	*(–4.54)*	*(–3.45)*	*(–2.84)*
Banking crisis	–1.506	–0.687*	–0.298	–0.747	–0.434	–0.543	–0.391	–2.481	–1.059	–0.597*	–0.510*	–1.523
	(–0.83)	*(–2.02)*	*(–1.07)*	*(–0.55)*	*(–0.27)*	*(–1.38)*	*(–1.34)*	*(–0.99)*	*(–0.58)*	*(–1.73)*	*(–1.98)*	*(–1.12)*
Initial government debt	–0.025**	–0.018**	0.003	–0.018**	–0.015	–0.014**	–0.008	–0.025*	–0.025**	–0.017**	–0.004	–0.014
	(–2.24)	*(–2.49)*	*(0.50)*	*(–2.29)*	*(–1.40)*	*(–2.59)*	*(–1.17)*	*(–1.74)*	*(–2.24)*	*(–2.36)*	*(–0.66)*	*(–1.28)*
Initial population size (log)	0.275	0.200	9.096***	0.094								
	(1.01)	*(1.02)*	*(2.81)*	*(0.09)*								
Initial investment					0.106**	0.076**	–0.079	0.080				
					(2.65)	*(2.50)*	*(–1.42)*	*(0.78)*				
Fiscal volatility									0.031	–0.194	–0.016	0.133
									(0.06)	*(–0.76)*	*(–0.07)*	*(0.28)*
Arellano-Bond AR(2) test *p*-value[1]				0.12				0.25				0.27
Hansen *J*-statistics (*p*-value)[2]				0.88				0.24				0.99
Number of observations	166	166	166	166	166	166	166	166	166	166	166	166
R^2	0.69	0.59	0.56		0.75	0.61	0.53		0.68	0.59	0.51	
Time-fixed effects	N/A	Yes	Yes	Yes	N/A	Yes	Yes	Yes	N/A	Yes	Yes	Yes

Note: Heteroskedasticity and country-specific autocorrelation consistent *t*-statistics are in parentheses. Time dummies are not reported. Levels of significance: *** 1%, ** 5%, * 10%. In the OLS regressions, dummies for OECD, Asia, Latin America, and Sub-Saharan Africa are also included in each regression (not reported to save space). FE refers to the fixed-effects panel regressions and BE is the between estimator. For the dynamic panel estimation, a two-step system GMM (SGMM) with the Windmeijer's finite-sample correction for the two-step covariance matrix.

1) The null hypothesis is that the first-differenced errors exhibit no second-order serial correlation.

2) The null hypothesis is that the instruments used are not correlated with the residuals.

Table 5

Panel Regression – Different Levels of Initial Debt and Advanced vs. Emerging Economies

(dependent variable: real per capita GDP growth)

Explanatory Variables	(1) BE	(2) Pooled OLS	(3) FE	(4) SGMM	(5) BE	(6) Pooled OLS	(7) FE	(8) SGMM
Initial real GDP per capita	–2.014***	–1.875***	–4.912**	–2.227***	–2.796***	–2.539***	–4.705**	–2.897***
	(–5.13)	*(–2.79)*	*(–2.65)*	*(–3.14)*	*(–4.51)*	*(–2.96)*	*(–2.35)*	*(–4.07)*
Initial years of schooling	4.377***	3.185***	2.260	3.988	4.691***	3.127***	2.232	2.074
	(3.77)	*(3.10)*	*(1.00)*	*(1.42)*	*(3.91)*	*(2.79)*	*(1.03)*	*(1.06)*
Initial inflation rate	1.551	–2.773***	–2.329***	–2.352**	0.503	–3.213***	–2.390***	–9.852**
	(0.59)	*(–3.67)*	*(–5.06)*	*(–2.65)*	*(0.18)*	*(–3.17)*	*(–5.17)*	*(–2.31)*
Initial government size	0.135**	0.127***	0.033	0.199**	0.096*	0.086*	0.056	0.293**
	(2.65)	*(3.06)*	*(0.40)*	*(2.03)*	*(1.82)*	*(2.02)*	*(0.70)*	*(2.65)*
Initial trade openness	–0.003	–0.005	0.026*	–0.007	–0.002	–0.005	0.023	–0.005
	(–0.65)	*(–1.37)*	*(1.77)*	*(–1.02)*	*(–0.30)*	*(–1.18)*	*(1.56)*	*(–0.76)*
Initial financial depth	0.023**	0.023***	0.006	0.026***	0.022**	0.024***	0.005	0.032***
	(2.18)	*(3.02)*	*(0.68)*	*(2.84)*	*(2.24)*	*(2.87)*	*(0.57)*	*(3.06)*
Terms of trade growth	0.183*	–0.018	–0.003	–0.038	0.235**	–0.008	–0.002	–0.050**
	(1.93)	*(–0.65)*	*(–0.18)*	*(–1.23)*	*(2.66)*	*(–0.32)*	*(–0.10)*	*(–2.26)*
Fiscal deficit	0.011	–0.046***	–0.033***	–0.045**	0.019	–0.050***	–0.034***	–0.059***
	(0.32)	*(–4.75)*	*(–3.14)*	*(–2.23)*	*(0.53)*	*(–4.94)*	*(–3.24)*	*(–3.69)*
Banking crisis	–1.270	–0.563	–0.468	–0.612	–0.992	–0.588*	–0.506*	–1.163
	(–0.72)	*(–1.60)*	*(–1.61)*	*(–0.83)*	*(–0.57)*	*(–1.75)*	*(–1.94)*	*(–1.13)*
Initial debt*Dum below30	0.016	0.0002	0.017	0.030				
	(0.17)	*(0.01)*	*(0.65)*	*(1.25)*				
Initial debt*Dum 30 90	–0.037	–0.028**	0.007	–0.015				
	(–1.43)	*(–2.66)*	*(0.79)*	*(–1.26)*				
Initial debt*Dum above90	–0.010	–0.015***	–0.001	–0.015***				
	(–0.79)	*(–2.79)*	*(–0.08)*	*(–2.91)*				
Initial debt*Dum advanced					–0.017	–0.012**	–0.005	–0.014*
					(–1.35)	*(–2.19)*	*(–0.75)*	*(–1.95)*
Initial debt*Dum emerging					–0.044**	–0.042***	0.001	–0.038*
					(–2.62)	*(–2.97)*	*(0.08)*	*(–1.95)*
Arellano-Bond AR(2) test *p*-value[1]				0.34				0.14
Hansen *J*-statistics (*p*-value)[2]				0.86				0.85
Number of observations	166	166	166	166	166	166	166	166
R^2	0.75	0.62	0.52		0.7	0.61	0.51	
Time-fixed effects	N/A	Yes	Yes	Yes	N/A	Yes	Yes	Yes

Note: Heteroskedasticity and country-specific autocorrelation consistent *t*-statistics are in parentheses. Time dummies are not reported. Levels of significance: *** 1%, ** 5%, * 10%. In the OLS regressions, dummies for OECD, Asia, Latin America, and Sub-Saharan Africa are also included in each regression (not reported to save space). FE refers to the fixed-effects panel regressions and BE is the between estimator. For the dynamic panel estimation, a two-step system GMM (SGMM) with the Windmeijer's finite-sample correction for the two-step covariance matrix.

1) The null hypothesis is that the first-differenced errors exhibit no second-order serial correlation.

2) The null hypothesis is that the instruments used are not correlated with the residuals.

relatively low levels of public debt is not significantly harmful to growth. In the OLS, the coefficient of medium level of debt (initial debt*Dum_30-90) is significant at 5 per cent, and its estimated coefficient is –0.028. But they are all insignificant in other estimations (BE, FE and SGMM). By contrast, the coefficients of high debt (initial debt*Dum_90) are negative and significant at 1 per cent under OLS, and SGMM.

Interestingly, the negative effect of initial debt on growth in advanced economies tends to be smaller than that in emerging economies. Columns 5-8 in Table 5 use the interaction terms between initial debt and dummy variables for advanced and emerging economies.[22] The coefficients of both interaction terms are negative and significant at various levels, except for the FE results and the coefficient of the initial debt*Dum_advanced term in BE. Under BE, OLS, and SGMM, the coefficients of initial debt in advanced economies range from –0.012 to –0.017, whose absolute size is smaller than that of emerging economies (–0.038 to –0.044): a 10 percentage point increase in initial debt-to-GDP ratio is associated with growth slowdown around 0.12-0.17 per cent in advanced economies, compared to 0.38-0.4 per cent in emerging economies.[23] This may reflect limited borrowing capacity of emerging economies due to less-developed domestic financial markets or fragile access to international capital markets.

3.6 Net foreign asset position, foreign liabilities, and domestic vs. foreign currency-denominated portion of public debt

An important question that arises is whether and the extent to which the impact on growth of initial debt is conditional on a country's economic and financial position *vis-à-vis* the rest of the world. For example, does the NFA (net foreign asset) position of a country or aggregate foreign liabilities matter for the magnitude of the relationship between public debt and growth?[24] Is it the case that the adverse impact of high debt on growth would be smaller if at the same time the aggregate foreign liabilities of a country are relatively low? This could be related to the fact that high public debt is being financed by private domestic savings rather than from abroad. Conversely, excessive foreign liabilities may compound the fiscal vulnerability arising from public debt *per se*, to the extent that foreign creditors may be more sensitive to changes in global risk appetite, or they may have shorter time horizons. Another channel could be in terms of signaling: high public debt when foreign liabilities are also high may indicate that the imbalances facing a country are broader than just the public sector. Similar arguments could be used with regard to the NFA, rather than only foreign liabilities *per se*.

In order to investigate this issue, we considered the NFA and foreign liabilities (as percent of GDP) as an additional variable, as well as an interactive term. It is the case that the bilateral correlation between government debt and the NFA or foreign liabilities is low (correlation coefficients are –0.10 and 0.11, respectively), and neither the NFA nor foreign liabilities are not significant in growth regressions, as shown in columns 1-4 of Table 6 (the results on foreign liabilities are not reported). However, the logic of the above argument would suggest that the interaction of initial public debt with NFA or liabilities might be more important. This was assessed by examining the interaction of debt with a dummy that took a value of 1 if the NFA exceeded the

---

22 See Appendix 1 for the list of advanced and emerging economies.

23 The same pattern is also found in the regressions on components of output per worker growth that the negative effects on growth of high debt are greater in emerging economies than in advanced economies.

24 The current sovereign debt crisis in Europe suggests that there is a strong correlation between the NFA positions and sovereign yields, indicating the market perceptions of fiscal risks associated with high debt (such as debt default and fiscal unsustainability) may depend on the NFA position. Conversely, some commentators observe that the currently very low yields on Japanese government bonds despite the very high level of debt (about 230 per cent of GDP) are possibly due to its high level of NFA in addition to Japan's haven status.

Table 6

Panel Regression – Different Levels of Initial NFA and Foreign Liabilities

(dependent variable: real per capita GDP growth)

Explanatory Variables	(1) BE	(2) Pooled	(3) FE	(4) SGMM	(5) BE	(6) Pooled	(7) FE	(8) SGMM	(9) BE	(10) Pooled	(11) FE	(12) SGMM
Initial real GDP per capita	–2.127*** (–4.95)	–1.698** (–2.23)	–4.772** (–2.29)	–1.852** (–2.51)	–2.273*** (–5.43)	–1.863** (–2.66)	–4.754** (–2.38)	–2.182*** (–3.86)	–1.909*** (–4.40)	–1.816** (–2.35)	–4.949** (–2.46)	–1.881*** (–2.78)
Initial years of schooling	4.760*** (3.81)	3.044** (2.51)	2.345 (1.04)	2.580 (1.11)	4.458*** (3.72)	3.076*** (2.92)	2.396 (1.08)	3.749*** (2.77)	5.066*** (4.22)	3.308*** (2.76)	2.250 (1.04)	1.592 (0.67)
Initial inflation rate	2.019 (0.75)	–2.397*** (–3.20)	–2.483*** (–5.82)	–1.402 (–1.21)	2.874 (1.12)	–2.098*** (–3.06)	–2.418*** (–5.53)	–1.905 (–1.58)	–0.277 (–0.09)	–2.621*** (–3.57)	–2.527*** (–5.84)	–2.514** (–2.15)
Initial government size	0.108* (2.00)	0.115** (2.44)	0.057 (0.73)	0.142 (1.50)	0.096* (1.86)	0.115** (2.70)	0.059 (0.74)	0.114 (0.68)	0.117** (2.26)	0.117** (2.61)	0.053 (0.67)	0.111 (1.32)
Initial trade openness	–0.003 (–0.50)	–0.006 (–1.28)	0.023 (1.51)	0.008 (1.07)	0.0003 (0.06)	–0.004 (–1.17)	0.024 (1.58)	–0.007 (–1.40)	–0.001 (–0.13)	–0.002 (–0.39)	0.026* (1.90)	0.003 (0.28)
Initial financial depth	0.019 (1.47)	0.021** (2.18)	0.006 (0.66)	0.018 (1.24)	0.014 (1.29)	0.021** (2.62)	0.006 (0.66)	0.027** (2.33)	0.020* (1.98)	0.022** (2.26)	0.006 (0.64)	0.021* (1.66)
Terms of trade growth	0.199** (2.22)	–0.016 (–0.62)	–0.003 (–0.13)	–0.034 (–0.77)	0.167* (1.92)	–0.021 (–0.90)	–0.004 (–0.17)	–0.034 (–0.99)	0.161* (1.81)	–0.022 (–0.95)	–0.007 (–0.28)	–0.051*** (–2.75)
Fiscal deficit	0.028 (0.79)	–0.044*** (–4.80)	–0.035*** (–3.59)	–0.034 (–1.44)	0.021 (0.62)	–0.045*** (–5.40)	–0.035*** (–3.52)	–0.044 (–1.59)	–0.0002 (–0.00)	–0.050*** (–5.03)	–0.039*** (–3.55)	–0.067*** (–2.78)
Banking crisis	–0.943 (–0.52)	–0.570 (–1.66)	–0.525* (–1.88)	–2.219* (–1.96)	–1.468 (–0.85)	–0.510 (–1.46)	–0.489* (–1.83)	–1.077 (–1.19)	–0.672 (–0.38)	–0.550 (–1.56)	–0.485* (–1.81)	–0.427 (–0.54)
Initial government debt	–0.024** (–2.14)	–0.017** (–2.40)	–0.004 (–0.72)	–0.015* (–1.81)								
Initial NFA (net foreign assets)	0.003 (0.39)	0.005 (0.84)	–0.002 (–0.21)	–0.013 (–1.26)								
*Initial debt*Dum NFA above median*[3]					–0.020* (–1.80)	–0.015** (–2.64)	–0.004 (–0.60)	–0.023* (–1.84)				
Initial debt*Dum NFA below median					–0.042*** (–2.88)	–0.029*** (–3.17)	–0.006 (–0.70)	–0.029* (–1.95)				
Initial debt*Dum Foreign Liabilities below_75percentile[4]									–0.013 (–0.99)	–0.015* (–1.98)	–0.003 (–0.38)	–0.017* (–1.85)
Initial debt*Dum Foreign Liabilities above_75percentile									–0.036*** (–2.81)	–0.025*** (–2.74)	–0.010 (–1.19)	–0.025* (–1.71)
Arellano-Bond AR(2) test p-value[1]				0.16				0.28				0.36
Hansen J-statistics (p-value)[2]				0.47				0.16				0.90
Number of observations	166	166	166	166	166	166	166	166	166	166	166	166
R^2	0.68	0.59	0.51		0.71	0.61	0.51		0.7	0.59	0.52	
Time-fixed effects	N/A	Yes	Yes	Yes	N/A	Yes	Yes	Yes	N/A	Yes	Yes	Yes

Note: Heteroskedasticity and country-specific autocorrelation consistent t-statistics are in parentheses. Time dummies are not reported. Levels of significance: *** 1%, ** 5%, * 10%. In the OLS regressions, dummies for OECD, Asia, Latin America, and Sub-Saharan Africa are also included in each regression (not reported to save space). FE refers to the fixed-effects panel regressions and BE is the between estimator. For the dynamic panel estimation, a two-step system GMM (SGMM) with the Windmeijer's finite-sample correction for the two-step covariance matrix.

1) The null hypothesis is that the first-differenced errors exhibit no second-order serial correlation.

2) The null hypothesis is that the instruments used are not correlated with the residuals.

3) The median value of NFA in the sample of 36 advanced and emerging economies is –17 per cent of GDP.

4) The 75 percentile level of foreign liabilities in the sample of 36 advanced and emerging economies is 89 per cent of GDP.

sample median value (–17 per cent of GDP), or if foreign liabilities were greater than the 75th percentile (89 per cent of GDP), and 0 otherwise. The results are shown in columns 5-8 and 8-12 of Table 6, respectively. The results bear out the basic hypothesis: when foreign liabilities are high or NFA low, the adverse impact of public debt on growth is about *one and a half to two* times as large as is the case otherwise. These results are striking from an economic perspective, and statistically significant. Perhaps what they are really implying is the notion that if the economy as a whole is operating essentially outside its means, the impact of high public debt on growth is substantially worse than when it is operating within it.

Next, we turn to the question of whether the currency composition of public debt also matters. The larger the portion of foreign-currency denominated debt as a share of total public debt, the larger the extent of exposure to foreign currency risk. This is related to the "Original Sin" problem highlighted by Eichengreen and Hausmann (1999), which could have adverse macroeconomic consequences. If a country affected by original sin has net foreign debt, then this country is likely to have a currency mismatch in its national balance sheet and large swings in the real exchange rate will have an effect on aggregate wealth and affect a country's ability to service its debt. As a consequence, original sin tends to make debt riskier, increase volatility, and affect a country's ability to conduct an independent monetary policy. Table 7 shows the results when we included the interaction of debt with a dummy that took a value of 1 if the domestic-currency portion exceeded the sample median value (89 per cent of total debt), or if it is greater than the 25th percentile (59 per cent of total debt), and 0 otherwise. The regression coefficients of the interaction terms are mostly significant and of the expected sign. They suggest that when the foreign-currency debt portion is large, the negative impact of public debt on growth can be more than twice as large as is the case otherwise.

4 Growth accounting

A detailed growth accounting exercise was also undertaken to explore channels (factor accumulation versus total factor productivity) through which government debt influences growth.[25] Taking a standard neoclassical framework, we consider a Cobb-Douglas production function $Y=AK^{\alpha}(HL)^{1-\alpha}$, where α is capital income share; K is physical capital; L is labor input; H is human capital; and A is TFP (total factor productivity). In terms of per worker, the production function can be written as $y=Ak^{\alpha}H^{1-\alpha}$, where $y=Y/L$ (output per worker) and $k=K/L$ (capital per worker). Then, growth of output per worker ($\dot{y}/y$) can be decomposed to TFP growth ($\dot{A}/A$) and contributions from growth of capital per worker ($\dot{k}/k$) and growth of human capital ($\dot{H}/H$).

$$\dot{y}/y = \dot{A}/A + \alpha(\dot{k}/k) + (1-\alpha)(\dot{H}/H) \tag{2}$$

Table 8 presents results from panel regression on output per worker growth and its components (TFP growth ($\dot{A}/A$) and growth of capital per worker ($\dot{k}/k$)), using the same baseline specification (Equation 1).[26] First, the coefficients of initial debt in the regressions of output per worker growth are significant at 5-10 per cent under BE, OLS, and SGMM, ranging from –0.012 to –0.022, whereas it becomes insignificant under FE (columns 1-4). The estimated

[25] See Appendix 3 for details about the growth accounting. The relation between labor force participation and initial debt is also examined, but the results are not significant (not reported).

[26] In terms of regression specification, y now denotes the logarithm of output per worker (Y/L) in the regressions on growth of output per worker (columns 1-4 of Table 8); y is the logarithm of level of TFP in the TFP growth regressions (columns 5-8); y is the logarithm of capital stock per worker (K/L) in the regressions on growth of capital stock per worker (columns 9-12). In the investment regressions of Table 9, the dependent variable is the average level of domestic investment (percent of GDP) over the period t and t-τ.

Table 7

Panel Regression – Domestic vs. Foreign Currency-Denominated Portion of Public Debt

(dependent variable: real per capita GDP growth)

Explanatory Variables	(1) BE	(2) Pooled OLS	(3) FE	(4) SGMM	(5) BE	(6) Pooled OLS	(7) FE	(8) SGMM
Initial real GDP per capita	–2.531***	–2.092***	–4.927**	–2.337**	–2.178***	–1.856**	–4.818**	–2.688**
	(–4.79)	*(–2.96)*	*(–2.32)*	*(–2.29)*	*(–4.40)*	*(–2.44)*	*(–2.35)*	*(–2.37)*
Initial years of schooling	5.311***	3.293***	3.195	4.209	5.054***	3.110**	3.030	2.578
	(4.01)	*(3.10)*	*(1.32)*	*(1.54)*	*(3.63)*	*(2.52)*	*(1.22)*	*(0.74)*
Initial inflation rate	0.946	–2.471***	–2.393***	–3.002**	2.136	–2.652***	–2.401***	–2.521*
	(0.30)	*(–3.53)*	*(–5.90)*	*(–2.28)*	*(0.69)*	*(–2.98)*	*(–4.73)*	*(–1.67)*
Initial government size	0.081	0.091*	0.086	0.182	0.111*	0.112**	0.095	0.118
	(1.30)	*(2.01)*	*(1.19)*	*(1.64)*	*(1.80)*	*(2.32)*	*(1.24)*	*(1.05)*
Initial trade openness	–0.002	–0.005	0.025	–0.012*	–0.001	–0.004	0.026	0.001
	(–0.32)	*(–0.93)*	*(1.51)*	*(–1.72)*	*(–0.18)*	*(–0.90)*	*(1.48)*	*(0.12)*
Initial financial depth	0.018	0.017**	0.005	0.026*	0.022	0.023**	0.004	0.024*
	(1.40)	*(2.08)*	*(0.50)*	*(1.84)*	*(1.54)*	*(2.36)*	*(0.41)*	*(1.97)*
Terms of trade growth	0.211**	0.004	0.003	–0.032	0.212**	–0.018	–0.000	–0.040*
	(2.27)	*(0.14)*	*(0.10)*	*(–0.99)*	*(2.18)*	*(–0.72)*	*(–0.00)*	*(–1.70)*
Banking crisis	–1.613	–0.832*	–0.588*	–0.501	–0.547	–0.612	–0.577*	–2.577
	(–0.67)	*(–2.03)*	*(–2.00)*	*(–0.34)*	*(–0.23)*	*(–1.33)*	*(–1.98)*	*(–1.48)*
Fiscal deficit	0.008	–0.051***	–0.036***	–0.074***	0.028	–0.047***	–0.035***	–0.063***
	(0.19)	*(–4.36)*	*(–3.24)*	*(–4.01)*	*(0.66)*	*(–4.61)*	*(–3.11)*	*(–4.43)*
Initial debt*Dum_domdebt_below25pctile[3]	–0.047**	–0.054***	–0.039***	–0.060*				
	(–2.35)	*(–2.86)*	*(–2.79)*	*(–1.94)*				
Initial debt*Dum_domdebt_above25pctile	–0.021*	–0.017**	–0.004	–0.023*				
	(–1.72)	*(–2.50)*	*(–0.77)*	*(–1.74)*				
Initial debt*Dum_domdebt_belowMedian[4]					–0.025	–0.028**	–0.011	–0.033**
					(–1.63)	*(–2.71)*	*(–1.04)*	*(–2.24)*
Initial debt*Dum_domdebt_aboveMedian					–0.025*	–0.018**	–0.006	–0.019**
					(–1.90)	*(–2.40)*	*(–0.87)*	*(–2.20)*
Arellano-Bond AR(2) test *p*-value[1]				0.68				0.89
Hansen J-statistics (*p*-value)[2]				0.41				0.55
Number of observations	151	151	151	151	151	151	151	151
R^2	0.7	0.63	0.51		0.67	0.6	0.51	
Time-fixed effects	N/A	Yes	Yes	Yes	N/A	Yes	Yes	Yes

Note: Heteroskedasticity and country-specific autocorrelation consistent *t*-statistics are in parentheses. Time dummies are not reported. Levels of significance: *** 1%, ** 5%, * 10%. In the OLS regressions, dummies for OECD, Asia, Latin America, and Sub-Saharan Africa are also included in each regression (not reported to save space). FE refers to the fixed-effects panel regressions and BE is the between estimator. For the dynamic panel estimation, a two-step system GMM (SGMM) with the Windmeijer's finite-sample correction for the two-step covariance matrix.

1) The null hypothesis is that the first-differenced errors exhibit no second-order serial correlation.

2) The null hypothesis is that the instruments used are not correlated with the residuals.

3) The 25 percentile level of domestic currency-denominated public debt portion in the sample 36 advanced and emerging economies is 59 per cent of total public debt.

4) The median level of domestic currency-denominated public debt portion in the sample 36 advanced and emerging economies is 89 per cent of total public debt.

Table 8

Growth Accounts and Panel Regression: Advanced and Emerging Economies

Explanatory Variables	(1) BE	(2) Pooled OLS	(3) FE	(4) SGMM	(5) BE	(6) Pooled OLS	(7) FE	(8) SGMM	(9) BE	(10) Pooled OLS	(11) FE	(12) SGMM
	dependent variable: growth of output per worker				*dependent variable: growth of TFP*				*dependent variable: growth of capital stock per worker*			
Lagged dependent variable[1]	–1.728**	–2.034**	–6.198**	–2.338**	–2.851***	–3.783***	–9.309***	–2.768**	–0.425	–0.515	–3.698*	–2.547
	(–2.25)	*(–2.40)*	*(–2.47)*	*(–2.71)*	*(–3.33)*	*(–4.62)*	*(–3.95)*	*(–2.61)*	*(–0.50)*	*(–0.79)*	*(–1.88)*	*(–1.57)*
Initial years of schooling	3.669***	2.649*	–1.829	3.894	2.507***	1.858**	–3.418	2.016	2.089	1.240	–1.809	10.654**
	(3.09)	*(2.01)*	*(–0.63)*	*(1.40)*	*(3.29)*	*(2.62)*	*(–1.33)*	*(1.12)*	*(1.26)*	*(0.91)*	*(–0.40)*	*(2.58)*
Initial inflation rate	1.443	–1.830**	–2.928***	–4.783	1.565	–1.241**	–2.260***	–4.515	0.190	–2.450***	–2.824***	–8.658
	(0.44)	*(–2.34)*	*(–5.72)*	*(–1.35)*	*(0.72)*	*(–2.04)*	*(–5.06)*	*(–1.64)*	*(0.04)*	*(–3.18)*	*(–4.73)*	*(–1.03)*
Initial government size	0.134**	0.104**	–0.076	0.102	0.070*	0.052	–0.031	0.143*	0.182**	0.114*	–0.330***	0.388
	(2.32)	*(2.28)*	*(–0.69)*	*(1.08)*	*(1.80)*	*(1.60)*	*(–0.27)*	*(1.73)*	*(2.31)*	*(1.96)*	*(–3.20)*	*(1.62)*
Initial trade openness	–0.009	–0.005	0.006	–0.009	–0.003	–0.001	0.016	–0.004	–0.016	–0.011	–0.015	–0.026
	(–1.14)	*(–1.06)*	*(0.47)*	*(–1.14)*	*(–0.56)*	*(–0.37)*	*(1.23)*	*(–0.28)*	*(–1.48)*	*(–1.52)*	*(–1.30)*	*(–0.93)*
Initial financial depth	0.030**	0.023**	0.012	0.026**	0.021**	0.017***	0.010	0.023*	0.025	0.015	0.003	0.027
	(2.33)	*(2.27)*	*(1.39)*	*(2.13)*	*(2.39)*	*(2.81)*	*(1.22)*	*(2.03)*	*(1.43)*	*(1.29)*	*(0.49)*	*(1.12)*
Terms of trade growth	0.342**	–0.038	–0.023	–0.059	0.237**	–0.021	–0.011	–0.048**	0.305*	–0.019	–0.007	–0.022
	(2.69)	*(–1.24)*	*(–0.82)*	*(–1.52)*	*(2.73)*	*(–0.89)*	*(–0.45)*	*(–2.19)*	*(1.79)*	*(–0.32)*	*(–0.15)*	*(–0.16)*
Banking crisis	–0.484	–0.033**	–0.027**	–0.010	–0.165	–0.032***	–0.022**	–0.033	–0.271	–0.010	–0.014	0.068
	(–0.26)	*(–2.55)*	*(–2.30)*	*(–0.28)*	*(–0.13)*	*(–3.53)*	*(–2.61)*	*(–1.00)*	*(–0.11)*	*(–0.77)*	*(–1.42)*	*(0.66)*
Fiscal deficit	0.061	–0.430	–0.539	–0.273	0.020	–0.327	–0.466	0.108	0.091	–0.128	–0.118	0.612
	(1.48)	*(–0.88)*	*(–1.29)*	*(–0.43)*	*(0.74)*	*(–0.77)*	*(–1.37)*	*(0.11)*	*(1.59)*	*(–0.28)*	*(–0.31)*	*(0.68)*
Initial government debt	–0.022*	–0.012*	0.005	–0.020**	–0.009	–0.004	0.009	–0.008	–0.034*	–0.023**	–0.014*	–0.045**
	(–1.78)	*(–1.75)*	*(0.69)*	*(–2.45)*	*(–1.11)*	*(–1.16)*	*(1.33)*	*(–0.60)*	*(–2.06)*	*(–2.13)*	*(–1.79)*	*(–2.04)*
Arellano-Bond AR(2) test *p*-value[2]				0.16				0.9				0.14
Hansen J-statistics (*p*-value)[3]				0.25				0.42				0.28
Number of observations	159	159	159	159	159	159	159	159	159	159	159	159
R^2	0.75	0.5	0.45		0.79	0.51	0.44		0.58	0.41	0.55	
Time-fixed effects	N/A	Yes	Yes	Yes	N/A	Yes	Yes	Yes	N/A	Yes	Yes	Yes

Note: Heteroskedasticity and country-specific autocorrelation consistent *t*-statistics are in parentheses. Time dummies are not reported. Levels of significance: *** 1%, ** 5%, * 10%. In the OLS regressions, dummies for OECD, Asia, Latin America, and Sub-Saharan Africa are also included in each regression (not reported to save space). FE refers to the fixed-effects panel regressions and BE is the between estimator. For the dynamic panel estimation, a two-step system GMM (SGMM) with the Windmeijer's finite-sample correction for the two-step covariance matrix.

1) The log of initial level of output per worker for columns 1-4; the log of initial level of TFP for Columns 5-8; and the log of initial level of capital stock per worker for columns 9-12, respectively.

2) The null hypothesis is that the first-differenced errors exhibit no second-order serial correlation.

3) The null hypothesis is that the instruments used are not correlated with the residuals.

coefficients of initial debt from the preferred estimators (BE and SGMM) indicate that a 10 percentage point increase in initial debt-to-GDP ratio is associated with a slowdown in growth of labor productivity (output per worker) of around 0.2 per cent per year.

Columns 5-8 show the regression results for TFP growth. There seems to be significant (conditional) convergence in the level of TFP, as indicated by the significant and negative coefficients of the log of initial level of TFP (in the first row). However, the coefficients of initial debt are insignificant across all four regressions, while they have a negative sign (except for FE). The estimated coefficients of initial debt under BE and SGMM are around –0.01.

The regression results for growth of capital per worker are stronger (columns 9-12). The initial debt coefficients are all significant at the conventional levels across estimation techniques, ranging from –0.014 to –0.045. Since the capital income share (α) is assumed to be 0.35 in the growth accounting exercise, the estimated coefficients of initial debt under BE and SGMM suggest that a 10 percentage point increase in initial debt-to-GDP ratio induces slowdown in growth of output per worker around 0.1-0.2 per cent per year via the channel of reduced growth in capital per worker. Taken together, the individual effects of initial debt on TFP growth and capital per worker growth roughly add up to 0.2-0.3 per cent per year, which is approximately in line with the regression outcomes for growth of output per worker shown in columns 1-4. However, there are no significant effects on human capital growth from debt and are not reported.

Table 9 presents panel regressions for domestic investment (percent of GDP, averaged over each five-year time period). Columns 1-3 show the regression results using the baseline specification except for the dependent variable which is the average domestic investment. The coefficients of initial debt are all significant at 1-10 per cent, ranging from –0.06 to –0.1. Columns 4 and 5 present the dynamic panel SGMM regressions in which the lagged term of the average investment is included instead of initial income per capita. The coefficient of initial debt in column 4 is significant at 5 per cent, and its estimate suggests that a 10 percentage point increase in initial debt-to-GDP ratio is associated with decline in domestic investment by about 0.4 percentage points of GDP. Column 5 includes interaction terms between initial debt and dummy variables for advanced and emerging economies. The coefficients of both interaction terms are significant at 5-10 per cent, and the estimated effects suggest that the adverse impact on domestic investment from debt in emerging economies is almost twice as large as that in advanced economies.

In addition, we considered the potential relationship between high debt and macroeconomic volatility. Intuitively, high debt may not only increase uncertainty about economic prospects and policies but also raise vulnerability to crises, leading to greater macroeconomic volatility. A simple scatter plot of macroeconomic volatility against initial government debt suggests a mild positive correlation. We ran regressions on macroeconomic volatility as measured by the log of standard deviation of annual real GDP growth rates using the baseline specification. The coefficient of initial debt in the regressions for volatility is only significant and of expected positive sign under FE when time-fixed effects are not included. However, they are all insignificant in all other estimations (with or without time dummies). Similarly, the coefficient of high debt (as captured by the interaction term, initial debt*Dum_90) is only significant under FE with no time-fixed effects included, as is the coefficient of initial debt for advanced economies (*i.e.*, initial debt*Dum_advance) in a separate FE regression (not reported to save space).

From the growth accounting perspective, therefore, the adverse effects on growth of initial debt largely reflect a slowdown in labor productivity growth mainly due to reduced investment and slower growth of capital per worker.

Table 9

Panel Regression on Investment: Advanced and Emerging Economies

Explanatory Variables	(1) BE	(2) Pooled OLS	(3) FE	(4) SGMM	(5) SGMM
Lagged dependent variable				0.763***	0.773***
				(8.35)	*(5.62)*
Initial real GDP per capita	–3.028*	2.645	8.700***		
	(–1.90)	*(0.89)*	*(3.76)*		
Initial years of schooling	3.361	–3.261	–2.197	5.029	–0.682
	(0.73)	*(–0.74)*	*(–0.34)*	*(1.56)*	*(–0.27)*
Initial inflation rate	–10.390	–1.632	–2.371***	–3.305*	–4.949***
	(–1.05)	*(–0.81)*	*(–3.15)*	*(–1.71)*	*(–3.27)*
Initial government size	–0.027	–0.056	–0.429**	0.367*	0.147
	(–0.14)	*(–0.32)*	*(–2.31)*	*(1.75)*	*(0.73)*
Initial trade openness	–0.011	0.000	–0.051*	–0.043***	–0.027***
	(–0.54)	*(0.02)*	*(–1.88)*	*(–2.94)*	*(–3.08)*
Initial financial depth	0.046	0.010	–0.009	0.031	0.022
	(1.19)	*(0.27)*	*(–1.00)*	*(1.56)*	*(1.43)*
Terms of trade growth	–0.157	0.062	0.069	0.200**	0.144*
	(–0.48)	*(0.70)*	*(0.91)*	*(2.42)*	*(1.81)*
Fiscal deficit	0.161	–0.002	–0.058***	–0.017	–0.069
	(1.21)	*(–0.07)*	*(–4.67)*	*(–0.31)*	*(–1.36)*
Banking crisis	1.178	–0.488	0.663	–1.519	–1.240
	(0.18)	*(–0.32)*	*(0.71)*	*(–1.06)*	*(–0.38)*
Initial government debt	–0.110**	–0.057*	–0.055***	–0.041**	
	(–2.64)	*(–1.67)*	*(–5.12)*	*(–2.48)*	
Initial debt*Dum_advanced					–0.032***
					(–2.94)
Initial debt*Dum_emerging					–0.077**
					(–2.61)
Arellano-Bond AR(2) test *p*-value[1]				0.54	0.79
Hansen *J*-statistics (*p*-value)[2]				0.59	0.40
Number of observations	166	166	166	159	159
R^2	0.45	0.48	0.53		
Time-fixed effects	N/A	Yes	Yes	Yes	Yes

Note: Heteroskedasticity and country-specific autocorrelation consistent *t*-statistics are in parentheses. Time dummies are not reported. Levels of significance: *** 1%, ** 5%, * 10%. In the OLS regressions, dummies for OECD, Asia, Latin America, and Sub-Saharan Africa are also included in each regression (not reported to save space). FE refers to the fixed-effects panel regressions and BE is the between estimator. For the dynamic panel estimation, a two-step system GMM (SGMM) with the Windmeijer's finite-sample correction for the two-step covariance matrix.

1) The null hypothesis is that the first-differenced errors exhibit no second-order serial correlation.

2) The null hypothesis is that the instruments used are not correlated with the residuals.

5 Concluding Remarks

Given the sharp increase in advanced country sovereign debt as a result of the global economic and financial crisis, there have begun to be serious concerns about its broader economic and financial market impact including an acute sovereign debt crisis in Europe. In particular, a number of observers have alluded to the risk that large debts may discourage capital accumulation and reduce economic growth. This could occur through higher long-term interest rates, higher future distortionary taxation, higher inflation, greater vulnerability to a debt crisis, and reduced scope for future counter-cyclical fiscal policy. If growth is indeed reduced, fiscal sustainability issues are likely to be exacerbated, with further adverse consequences.

Empirical evidence, based on a range of econometric techniques, strongly suggests an inverse relationship between initial debt and subsequent growth, controlling for other determinants of growth: on average, a 10 percentage point increase in the initial debt-to-GDP ratio is associated with a slowdown in real per capita GDP growth of around 0.2 percentage points per year, with the impact being smaller (around 0.15) in advanced economies and/or smaller when (net) foreign liabilities are relatively high. Also, the currency composition of public debt matters. There is some evidence of non-linearity, with only high (above 90 per cent of GDP) levels of debt having a significant negative effect on growth. This adverse effect largely reflects a slowdown in labor productivity growth, mainly due to reduced investment and slower growth of the capital stock per worker. On average, a 10 percentage point increase in initial debt ratio is associated with a decline of investment by about 0.4 percentage points of GDP, with a larger impact in emerging economies. Various robustness checks yield largely similar results. They underline the need to take measures to not just stabilize public debt but to place them on a downward trajectory in the medium and long term.

APPENDIX 1
COUNTRY LIST

The sample of countries is dictated by the availability of data. The following 38 advanced and emerging economies with a population of over 5 million are included in the baseline panel regressions.

Country	Country
Australia	Japan
Austria	Korea
Belgium	Malaysia
Brazil	Mexico
Canada	Netherlands
Chile	Pakistan
China	Peru
Colombia	Philippines
Czech Republic*	Poland
Denmark	Portugal
Egypt	Russian Federation*
France	Slovak Republic*
Germany	South Africa
Greece	Spain
Hong Kong	Sweden
Hungary	Switzerland
India	Turkey
Indonesia	United Kingdom
Italy	United States

Note:
1. Three countries with the asterisk mark (*) in the above list are not included in the growth accounting exercise because necessary data in computing TFP are not available.
2. Eight additional countries are also available in the panel regressions for all available 46 advanced and emerging economies without the over-5-million-population size restriction: Finland, Iceland, Ireland, Israel, Jordan, Norway, New Zealand, and Singapore.

3 Thirty three developing countries that are included in the full sample of 79 countries are: Barbados, Bolivia, Bulgaria, Costa Rica, Croatia, Cyprus, Ecuador, Gambia, Guinea-Bissau, Guyana, Honduras, Iran, Jamaica, Kuwait, Lesotho, Mauritania, Mauritius, Mozambique, Nicaragua, Panama, Romania, Rwanda, Senegal, Slovenia, Sri Lanka, Sudan, Swaziland, Syria, Togo, Trinidad & Tobago, Tunisia, Uganda, and Uruguay.

4 The list of advanced economies includes Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Netherlands, New Zealand, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States, which were the OECD member nations as of 1990, except for Turkey which is classified as an emerging market economy.

APPENDIX 2
DESCRIPTION OF DATA

A. Dependent variables

The following dependent variables are measured over the five-year period in the panel (or the relevant time period in the cross-country regression).

1) Growth of real per capita GDP, PWT7.0 (2011)
2) Growth of output per worker, PWT7.0 (2011)
3) TFP growth, constructed using PWT7.0 (2011) and Barro and Lee (2011)
4) Growth of capital per worker PWT7.0 (2011)
5) Domestic investment (percent of GDP), PWT7.0 (2011)
6) Volatility of output (log of standard deviation of annual real GDP growth rates over the five-year period), PWT7.0 (2011)

B. Explanatory variables

Initial values of explanatory variables – for example, initial real GDP per capita or initial government size – are measured at the measured at the beginning of each five-year period in the panel (or the relevant time period in the cross-country regression). Otherwise, the variables, such as terms of trade growth or average government debt, are averaged over the five-year period.

1) Initial real GDP per capita (in log), PWT7.0 (2011)
2) Initial average years of schooling of population of age over 15 (in log), Barro and Lee (2011)
3) Initial government size (percent of GDP), PWT7.0 (2011)
4) Initial trade openness (percent of GDP), PWT7.0 (2011)
5) Initial inflation rate (log of $(1+\pi)$), WDI (2011)
6) Initial financial market depth (liquid liabilities, percent of GDP), WDI (2011)
7) Terms of trade growth (in percent), IMF, WEO (2011)
8) Banking crisis (total number of incidences over five-year period), Reinhart and Reinhart (2008)
9) Initial population size (in log), PWT7.0 (2011)
10) Fiscal deficit (percent of GDP), IMF, WEO (2011)
11) Population growth (in percent), PWT7.0 (2011)
12) Initial domestic investment (percent of GDP), PWT7.0 (2011)
13) Fiscal volatility (log of standard deviation of annual growth rates of real general government expenditures over the five-year period), WDI (2011)
14) Aged-dependency ratio (ratio of population of age over 65 to working population), WDI (2011)
15) Urbanization, WDI (2011)
16) Checks and balances, Database of Political Institutions (2009)
17) Constraints on executive decision-making, Polity IV (2009)
18) Initial gross government debt (percent of GDP), IMF, WEO (2011)
19) Average gross government debt (percent of GDP), IMF, WEO (2011)

APPENDIX 3
GROWTH ACCOUNTING

Taking a standard neoclassical approach, let us consider a Cobb-Douglas production function $Y=AK^{\alpha}(HL)^{1-\alpha}$, where α=capital income share; K=physical capital; L=labor input; H=human capital; and A= TFP (total factor productivity). In terms of per worker, the production function can be written as $y=Ak^{\alpha}H^{1-\alpha}$, where $y=Y/L$ (output per worker) and $k=K/L$ (capital per worker). Then, growth of output per worker ($\dot{y}/y$) can be decomposed to TFP growth ($\dot{A}/A$) and contributions from growth of capital per worker ($\dot{k}/k$) and growth of human capital ($\dot{H}/H$):

$$\dot{y}/y = \dot{A}/A + \alpha(\dot{k}/k) + (1-\alpha)(\dot{H}/H).$$

The growth accounting is consistent with a wide range of alternative production functional forms linking the factor inputs and output. It is only necessary to assume a degree of competition sufficient so that the earnings of the factors are proportionate to their factor productivity. Then we can measure TFP growth rates, using the shares of income paid to the factors to measure their importance in the production process as described above (see Caselli, 2005 for details about TFP). Since consistent measures of factor income shares are often difficult to obtain for individual countries, most studies assume that income shares are identical across time and space. Yet, Gollin (2002) provides strong evidence in support of such an assumption of constant income shares across time and space, which is consistent with the Cobb-Douglas function approach. Also, Bernanke and Gürkaynak (2001) find no systematic tendency for labor shares to vary with real GDP per capita or the capital-labor ratio nor systematic tendency to rise or fall over time, and most estimated labor income shares lie between 0.6 and 0.8, the average being 0.65. In this paper, we tried both a fixed labor share of 0.65 and actual income shares from Gollin (2002) and Bernanke and Gürkaynak (2001). The results using alternative income share measures are very similar, suggesting that using a fixed labor income share is not a serious problem.

We construct a new data set on TFP for a large number of developed and developing countries in the period 1970-2008. National income and product account data and labor force data are obtained from the latest version 7.0 of the Penn World Table (Heston *et al.*, 2011). To construct the labor quality index for human capital (H), we take average years of schooling in the population over 15 years old from the international data on educational attainment by Barro and Lee (2011). We follow Hall and Jones (1999) to give larger weight to more educated workers as follows: $H = e^{\phi(E)}$, where E is average years of schooling; the function $\phi(E)$ is piece linear with slope of 0.134 for $E \le 4$, 0.101 for $4 < E \le 8$; and 0.068 for $8 < E$. The rationale behind this functional form for human capital is as follows. The wage of a worker with E years of education is proportional to her human capital. Since the wage-schooling relationship is widely believed to be log-linear, this would imply that human capital (H) and education (E) would have a log-linear relation as well, such as $H=\exp(const \times E)$. However, international data on education-wage profiles (Psacharopulos, 1994) suggests that in sub-Saharan Africa (which has the lowest levels of education), the return to one extra year of education is about 13.4 per cent, the world average is 10.1 per cent, and the OECD average is 6.8 per cent.

We estimate the capital stock, K, using the perpetual inventory method: $K_t = I_t + (1-\delta)K_{t-1}$, where I_t is the investment and δ is the depreciation rate. Data on I_t are from PWT 7.0 as real aggregate investment in PPP. For many countries in our sample, investment data go back to as early as 1950-55. We estimate the initial value of the capital stock, say, in year 1950 as I1950/(g+δ) where g is the average compound growth rate between 1950 and 1960, and δ is the depreciation rate (δ =0.06 is assumed). We further adjust these capital stocks for the portion of residential capital stock that is not directly related to production activity.[27] Batteries of consistency checks suggest that our estimates of TFP growth are reasonable.

[27] PWT 5.6 provides data on residential capital per worker as a fraction of nonresidential capital per worker for 63 countries. For these countries, we use the average ratio of nonresidential capital to total capital to impute the nonresidential capital stock in our data set. For the remaining countries, we assume that nonresidential capital is two-thirds of the total capital, which is about the average value of 0.69 for the countries for which the data are available.

APPENDIX

Table 10

Level of Initial Government Debt, Growth, and Investment, 1970-2008: Countries with a Population of Over 5 Million

Group of Countries	Initial Debt below 30 per cent of GDP	Initial Debt between 30 and 60 per cent of GDP	Initial Debt between 60 and 90 per cent of GDP	Initial Debt above 90 per cent of GDP
		Average: Real per capita GDP Growth Rate *(annualized over the subsequent 5 years)*		
Entire	5.0	2.7	2.6	2.2
Advanced[1]	2.6	1.8	2.1	1.7
Emerging	5.4	3.1	2.9	3.7
Developing	6.6	4.4	3.1	2.2
		Average: Output per worker Growth Rate *(annualized over the subsequent 5 years)*		
Entire	4.4	1.9	2.0	1.7
Advanced	2.3	1.2	1.6	1.5
Emerging	4.7	2.3	2.3	3.4
Developing	5.9	3.3	2.4	1.6
		Average: TFP Growth Rate *(annualized over the subsequent 5 years)*		
Entire	1.3	0.3	0.7	1.1
Advanced	0.3	0.1	0.5	0.4
Emerging	2.0	0.8	0.7	2.4
Developing	2.1	-0.3	1.1	1.4
		Average: Capital stock per worker Growth Rate *(annualized over the subsequent 5 years)*		
Entire	4.6	2.4	2.2	1.5
Advanced	4.2	1.8	2.2	2.1
Emerging	5.8	1.8	1.9	0.9
Developing	2.5	5.7	2.3	1.2
		Average: Domestic Investment *(percent of GDP over the subsequent 5 years)*		
Entire	25.8	21.7	21.6	18.5
Advanced	25.2	20.7	21.9	23.9
Emerging	30.5	22.1	21.8	16.4
Developing	21.0	23.7	21.0	15.8

Note: Initial debts are the government gross debt to GDP (percent) in the first year of each five-year sub-period (*i.e.*, 1970, 1975, 1980, 1985, 1990, 1995, 2000, 2005). Average growth rates (percent per annum) are over each five-year sub-period (*i.e.*, 1970-74, 1975-79, 1980-84, 1985-89, 1990-94, 1995-99, 2000-04, 2005-08).

1) Advanced economies are defined as the OECD Members as of 1990, excluding Turkey, which is classified as an emerging economy.

Table 11

Cross-country Regression – Government Debt and Real per Capita GDP Growth: Advanced and Emerging Economies (Without Restriction on Population Size)

(dependent variable: real per capita GDP growth)

Explanatory Variables	(1) OLS 1975-2008	(2) OLS 1985-2008	(3) OLS 1990-2008	(4) OLS 1995-2008	(5) OLS 2000-2008	(6) OLS 1990-2008	(7) OLS 1995-2008	(8) OLS 2000-2008
Initial real GDP per capita	1.862	–2.928*	–2.464***	–1.726**	–0.480	–1.353	–1.121*	–0.494
	(1.91)	*(–2.00)*	*(–4.44)*	*(–2.37)*	*(–0.58)*	*(–1.63)*	*(–1.84)*	*(–0.61)*
Initial years of schooling	0.393	0.576	2.462**	2.944**	1.021	1.419	2.204**	1.286
	(0.38)	*(0.50)*	*(2.66)*	*(2.08)*	*(0.63)*	*(1.15)*	*(2.09)*	*(0.82)*
Initial inflation rate	8.395**	–1.578	0.400	8.932**	1.628	–0.059	2.831**	1.300
	(4.37)	*(–0.77)*	*(0.99)*	*(2.12)*	*(0.43)*	*(–0.38)*	*(2.19)*	*(0.38)*
Initial government size	–0.127*	–0.024	–0.027	0.021	0.114**	–0.020	0.020	0.101*
	(–2.86)	*(–0.40)*	*(–0.85)*	*(0.58)*	*(2.25)*	*(–0.72)*	*(0.57)*	*(1.96)*
Initial trade openness	0.012*	0.016	0.010**	0.014***	0.001	0.008	0.004	–0.0002
	(3.93)	*(1.39)*	*(2.18)*	*(3.04)*	*(0.21)*	*(1.43)*	*(0.81)*	*(–0.04)*
Terms of trade growth	0.039	–0.036	–0.192	–0.189*	0.071	–0.195	–0.124	0.049
	(0.54)	*(–0.20)*	*(–1.13)*	*(–1.97)*	*(0.78)*	*(–1.31)*	*(–1.60)*	*(0.61)*
Banking crisis			–0.428	–0.728	0.061	0.082	–0.825	–0.044
			(–1.26)	*(–1.33)*	*(0.11)*	*(0.22)*	*(–1.60)*	*(–0.08)*
Initial government debt	–0.020**	–0.009	–0.018***	–0.029***	–0.020			
	(–4.49)	*(–1.07)*	*(–3.29)*	*(–3.73)*	*(–1.65)*			
Government debt, average						–0.021**	–0.022**	–0.018*
						(–2.21)	*(–2.68)*	*(–1.83)*
Number of observations	10	20	30	37	44	42	46	46
R^2	0.99	0.60	0.85	0.67	0.63	0.53	0.51	0.62

Note: Heteroskedasticity-consistent t-statistics are in parentheses. Levels of significance: *** 1%, ** 5%, * 10%. An intercept term and dummies for OECD, Asia, Latin America, and Sub-Saharan Africa are included in each regression, except for column (1) in which the number of observations is small relative to the number of covariates (not reported to save space).

Table 12

Growth Accounting and Cross-Country Growth Regression: Advanced and Emerging Economies

(without restriction on population size)

Explanatory Variables	(1) OLS 1990-2008	(2) OLS 1995-2008	(3) OLS 1990-2008	(4) OLS 1995-2008	(5) OLS 1990-2008	(6) OLS 1995-2008	(7) OLS 1990-2008	(8) OLS 1995-2008	(9) OLS 1990-2008	(10) OLS 1995-2008	(11) OLS 1990-2008	(12) OLS 1995-2008
	dependent variable: growth of real output per worker				*dependent variable: growth of TFP*				*dependent variable: growth of capital stock per worker*			
Initial real GDP per capita	–2.278***	–1.490**	–1.219	–1.033	–1.810***	–1.070**	–1.276**	–1.001***	–1.438*	–1.080	–0.041	–0.119
	(–4.35)	*(–2.19)*	*(–1.44)*	*(–1.68)*	*(–5.14)*	*(–2.57)*	*(–2.52)*	*(–2.87)*	*(–1.89)*	*(–1.21)*	*(–0.04)*	*(–0.16)*
Initial years of schooling	2.653***	3.076**	1.692	2.620**	2.972***	2.810***	2.352***	2.790***	1.350	2.387	0.004	1.300
	(2.90)	*(2.10)*	*(1.40)*	*(2.17)*	*(4.37)*	*(3.04)*	*(3.12)*	*(3.79)*	*(0.86)*	*(1.24)*	*(0.00)*	*(0.85)*
Initial inflation rate	0.739*	11.195**	0.079	3.680*	0.762**	7.907***	0.239	2.529**	0.029	8.876	–0.440	2.710
	(1.89)	*(2.54)*	*(0.33)*	*(1.91)*	*(2.84)*	*(3.08)*	*(1.41)*	*(2.04)*	*(0.05)*	*(1.23)*	*(–1.39)*	*(0.98)*
Initial government size	–0.030	0.038	–0.033	0.015	–0.026	0.038*	–0.026	0.019	–0.037	0.006	–0.038	–0.012
	(–0.87)	*(1.10)*	*(–1.01)*	*(0.40)*	*(–1.51)*	*(1.86)*	*(–1.24)*	*(0.84)*	*(–0.68)*	*(0.13)*	*(–0.99)*	*(–0.25)*
Initial trade openness	0.010**	0.013**	0.007	0.002	0.011***	0.011***	0.009**	0.005	–0.002	0.004	–0.006	–0.008
	(2.35)	*(2.64)*	*(1.14)*	*(0.35)*	*(3.47)*	*(4.05)*	*(2.30)*	*(1.54)*	*(–0.32)*	*(0.60)*	*(–0.94)*	*(–1.27)*
Terms of trade growth	–0.063	–0.187*	–0.089	–0.171**	–0.054	–0.165**	–0.031	–0.138**	–0.082	–0.071	–0.176	–0.098
	(–0.43)	*(–1.80)*	*(–0.64)*	*(–2.29)*	*(–0.59)*	*(–2.64)*	*(–0.38)*	*(–2.66)*	*(–0.33)*	*(–0.44)*	*(–1.07)*	*(–1.00)*
Banking crisis	–0.014	–0.628	0.432	–0.837	0.030	–0.467	0.372	–0.299	–0.345	–0.204	0.092	–1.295*
	(–0.04)	*(–1.01)*	*(1.15)*	*(–1.55)*	*(0.14)*	*(–1.28)*	*(1.59)*	*(–0.93)*	*(–0.62)*	*(–0.23)*	*(0.18)*	*(–1.75)*
Initial government debt	–0.021***	–0.029***			–0.012***	–0.018***			–0.020*	–0.027*		
	(–3.33)	*(–2.86)*			*(–3.93)*	*(–3.21)*			*(–1.77)*	*(–1.80)*		
Government debt, average			–0.020**	–0.017**			–0.010	–0.008			–0.026**	–0.026**
			(–2.08)	*(–2.20)*			*(–1.68)*	*(–1.68)*			*(–2.33)*	(–2.69)
Number of observations	30	36	44	45	30	36	44	45	30	36	44	45
R^2	0.85	0.64	0.48	0.46	0.87	0.69	0.56	0.51	0.65	0.42	0.45	0.38

Note: Heteroskedasticity-consistent t-statistics are in parentheses. Levels of significance: *** 1%, ** 5%, * 10%. An intercept term and dummies for OECD, Asia, Latin America, and Sub-Saharan Africa are included in each regression (not reported to save space).

REFERENCES

Aghion, P., and S. Durlauf (2005), *Handbook of Economic Growth*, Vol. 1A and 1B, Amsterdam, North-Holland.

Aghion, P. and E. Kharroubi (2007), "Cyclical Macro Policy and Industry Growth: The Effect of Countercyclical Fiscal Policy", Harvard University, Working Paper.

Aizenman, J., K. Kletzer and B. Pinto (2007), "Economic Growth with Constraints on Tax Revenues and Public Debt: Implications for Fiscal Policy and Cross-country Differences", NBER, Working Paper, No. 12750, Cambridge (MA).

Arellano, M. and S. Bond (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, Vol. 58, No. 2, pp. 277-97, April.

Arellano, M. and O. Bover (1995), "Another Look at the Instrumental Variables Estimation of Error-Components Models", *Journal of Econometrics*, No. 68, pp. 29-51.

Baldacci, E., A. Hillman and N. Kojo (2004), "Growth, Governance and Fiscal Policy Transmission Channels in Low-income Countries", *European Journal of Political Economy*, Vol. 20, pp. 517-49.

Baldacci, E. and M. Kumar (2010), "Fiscal Deficits, Public Debt and Sovereign Bond Yields", IMF, Working Paper, No. 10/184, Washington (D.C.).

Barro, R. (1979), "On the Determinants of the Public Debt", *Journal of Political Economy*, Vol. 85, No. 5, pp. 940-71.

—————— (1990), "Government Spending in a Simple Model of Endogenous Growth", *Journal of Political Economy*, Vol. 98, No. 5, pp. S103-125.

—————— (1995), "Inflation and Economic Growth", NBER, Working Paper, No. 5326, Cambridge (MA).

Barro, R. and J. Lee (2011), "A New Data Set of Educational Attainment in the World, 1950-2010", NBER, Working Paper, No. 15902.

Bazzi, S. and M. Clemens (2009), "Blunt Instruments: On Establishing the Causes of Economic Growth", Center for Global Development, Working Paper, No. 171, Washington (D.C.).

Beck, T. and R. Levine (2004), "Stock Markets, Banks, and Growth: Panel Evidence", *Journal of Banking and Finance*, Vol. 28, No. 3, pp. 423-42.

Bernanke, B. and R. Gürkaynak (2001), "Is Growth Exogenous? Taking Mankiw, Romer, and Weil Seriously?", in B. Bernanke and K. Rogoff (eds.), *NBER Macroeconomics Annual 2001*, pp. 11-57, Cambridge (MA), NBER.

Blundell, R. and S. Bond (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models", *Journal of Econometrics*, Vol. 87, No. 1, pp. 115-43.

Bond, S. (2002), "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice", Institute for Fiscal Studies, Working Paper, No. 09/02, London.

Burnside, C., M. Eichenbaum and S. Rebelo (2001), "Prospective Deficits and the Asian Currency Crisis", *Journal of Political Economy*, Vol. 109, No. 6, pp. 1155-97.

Caselli, F. (2005), "Accounting for Cross-country Income Differences", in P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, Amsterdam, North-Holland.

Cecchetti, S., M. Mohanty and F. Zampolli (2011), "The Real Effects of Debt", Bank for International Settlements, Working Paper, No. 352, August.

Checherita, C. and P. Rother (2010), "The Impact of High and Growing Government Debt on Economic Growth: An Empirical Investigation for the Euro Area", European Central Bank, Working Paper, No. 1237, August.

Cochrane, J. (2011), "Understanding Policy in the Great Recession: Some Unpleasant Fiscal Arithmetic", *European Economic Review*, No. 55, pp. 2-30.

Dotsey, M. (1994), "Some Unpleasant Supply Side Arithmetic", *Journal of Monetary Economics*, pp. 507-24.

Durlauf, S., P. Johnson and J. Temple (2005), "Growth Econometrics", in P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, Amsterdam, North-Holland.

Easterly, W. (2005), "National Policies and Economic Growth", in P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, Amsterdam, North-Holland.

Eichengreen, B. and R. Hausmann (1999), "Exchange Rates and Financial Fragility", paper presented at the symposium *New Challenges for Monetary Policy*, 26-28 August, Jackson Hole (WY).

Elmendorf, D. and N.G. Mankiw (1999), "Government Debt", in J.B. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, Vol. 1C, Amsterdam, North-Holland.

Fatás, A. and I. Mihov (2003), "The Case for Restricting Fiscal Policy Discretion", *Quarterly Journal of Economics*, Vol. 118, No. 4, pp.1419-47.

Fischer, S. (1993), "The Role of Macroeconomic Factors in Growth", *Journal of Monetary Economics*, Vol. 32, No. 3, pp. 485-512.

Gale, W. and P. Orszag (2003), "The Economic Effects of Long-term Fiscal Discipline", Urban-Brookings Tax Policy Center, Discussion Paper, No. 8, Washington (D.C.), Brookings Institution.

Glaeser, E., R. La Porta, F. López-de-Silanes and A. Shleifer (2004), "Do Institutions Cause Growth?", *Journal of Economic Growth*, Vol. 9, pp. 271-303.

Gollin, D. (2002), "Getting Income Shares Right", *Journal of Political Economy*, Vol. 110, No. 2, pp. 458-74.

Grilliches, Z. and J. Hausman (1987), "Errors in Variables in Panel Data", *Journal of Econometrics*, Vol. 31, No. 1, pp. 93-118.

Grossman, G. and E. Helpman (1991), *Innovation and Growth in the Global Economy*, Cambridge (MA), MIT Press.

Hall, R. and C.I. Jones (1999), "Why Do Some Countries Produce So Much More Output per Worker than Others?", *Quarterly Journal of Economics*, Vol. 114, pp. 83-116.

Hauk, W. and R. Wacziarg (2009), "A Monte Carlo Study of Growth Regressions", *Journal of Economic Growth*, Vol. 14, No. 2, pp. 103-47.

Hemming, R., M. Kell and A. Schimmelpfennig (2003), "Fiscal Vulnerability and Financial Crises in Emerging Market Economies", IMF, Occasional Paper, No. 218, Washington (D.C.).

Heston, A., R. Summers and B. Aten (2011), *Penn World Table Version 7.0*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, June.

Huang, H. and D. Xie (2008), "Fiscal Sustainability and Fiscal Soundness", *Annals of Economics and Finance*, Vol. 9, No. 2, pp. 239-51.

Imbs, J. and R. Rancière (2005), "The Overhang Hangover", World Bank, Policy Research, Working Paper, No. 3673, Washington (D.C.).

Islam, N. (1995), "Growth Empirics: A Panel Data Approach", *Quarterly Journal of Economics*, Vol. 110, No. 4, pp.1127-70.

Krugman, P. (1988), "Financing Versus Forgiving a Debt Overhang", *Journal of Development Economics*, Vol. 29, pp. 253-68.

Kumar, M. and J. Woo (2010), "Public Debt and Growth", IMF, Working Paper, No. 10/174, Washington (D.C.).

Pattilo, C., H. Poirson and L. Ricci (2002), "External Debt and Growth", IMF, Working Paper, No. 02/69, Washington (D.C.).

—————(2004), "What Are the Channels through which External Debt Affects Growth?", IMF, Working Paper, No. 04/15, Washington (D.C.).

Ramey, G. and V. Ramey (1995), "Cross-country Evidence on the Link between Volatility and Growth", *American Economic Review*, Vol. 85, pp. 1138-51.

Reinhart, C. and V. Reinhart (2008), "Capital Flow Bonanzas: An Encompassing View of the Past and Present", in J. Frankel and F. Giavazzi (eds.) *NBER International Seminar in Macroeconomics 2008*, Chicago, Chicago University Press.

Reinhart, C., V. Reinhart and K. Rogoff (2012), "Debt Overhang: Past and Present", NBER, Working Paper, No. 18015.

Reinhart, C. and K. Rogoff (2009), "The Aftermath of Financial Crises", *American Economic Review*, Vol. 99, No. 2, pp. 466-72, May.

—————(2010), "Growth in a Time of Debt", *American Economic Association Papers and Proceedings*, Vol. 100, No. 2, pp. 1-9, May.

Roodman, D. (2009), "A Note on the Theme of Too Many Instruments", *Oxford Bulletin of Economics and Statistics*, Vol. 71, No. 1, pp. 135-58.

Sachs, J. (1989), "The Debt Overhang of Developing Countries", in G.A. Calvo, R. Findlay, P. Kouri and J.B. de Macedo (eds.), *Debt Stabilization and Development*, pp. 80-102, Oxford and Cambridge (MA), Basil Blackwell.

Sala-i-Martín, X. (1997), "I Just Ran Two Million Regressions", *American Economic Review*, Papers and Proceedings, pp. 178-83, May.

Sala-i-Martín, X., G. Doppelhofer and R. Miller (2004), "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach", *American Economic Review*, Vol. 94, No. 4, pp. 813-35.

Saint-Paul, G. (1992), "Fiscal Policy in an Endogenous Growth Model", *Quarterly Journal of Economics*, Vol. 107, No. 4, pp. 1243-59.

Sargent, T. and N. Wallace (1981), "Some Unpleasant Monetarist Arithmetic", *Quarterly Review*, Federal Reserve Bank of Minneapolis, Fall.

Stock, J.H., J.H. Wright and M. Yogo (2002), "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments" *Journal of Business and Economic Statistics*, Vol. 20, No. 4, pp. 518-29.

Woo, J. (2003), “Economic, Political and Institutional Determinants of Public Deficits”, *Journal of Public Economics*, Vol. 87, No. 3, pp. 387-426, March.

——————(2009), “Why Do More Polarized Countries Run More Procyclical Fiscal Policy?”, *Review of Economics and Statistics*, Vol. 91, No. 4, pp. 850-70, November.

Woo, J., T. Kinda and M. Poplawski-Ribeiro (2012), “Economic Recovery from Recessions, Fiscal Policy, and Structural Reforms”, IMF, work in progress.

World Bank (2011), *World Development Indicators CD-ROM*, Washington (D.C.).

DYNAMIC LABOR SUPPLY WITH TAXES: THE CASE OF ITALIAN COUPLES

Maria Rosaria Marino, Marzia Romanelli* and Martino Tasso**

Labor force participation rate among married women in Italy is particularly low. In order to better understand the role played by the tax and benefit system on this phenomenon, we build and estimate a structural dynamic life-cycle model of household labor supply, saving, and consumption behavior. The model features several sources of heterogeneity in the characteristics of the members of the couple and it incorporates most of the fiscal rules which have an effect on the net incomes of the agents. The parameters of the model are estimated using cross-sectional and longitudinal data for the 2004-10 period. We use the estimated model to simulate a few counterfactual policies and study their effect on labor supply and poverty. In this version of our work we present some preliminary estimates and simulations.

1 Introduction

Government decisions about how to raise revenue have obviously a large impact on households' choices. The design of these policies can foster economic growth through the labor supply channel. Interventions in this area face a trade-off between the desire to increase the welfare in the poorest strata of the population and the need to avoid negative effects on the labor supply. In many developed countries these interventions take the form of special provision of the tax scheme or work-related cash benefits. Because the fixed cost of working is likely to be related to the number of children in the family these instruments vary accordingly. Moreover, a long series of studies have found that the margin which is more likely to be affected by these policies is the participation one for single and married women.

The role of taxes and family benefits on household labor supply and consumption decisions has been a topic of deep research interest for a long time. The works of Eckstein and Wolpin (1989), Sheran (2007), and Eckstein and Lifshitz (2011) are examples of contributions to the modelling of female labor supply in a dynamic framework. On the other hand, relatively few studies which estimate such complex models allow for a full specification of taxes and welfare benefits: the works of Haan and Prowse (2010) on joint retirement decisions of German workers, and Keane and Wolpin (2010) on labor supply effects of the Earned Income Tax Credit in the United States are exemplary of this strand of the literature. Other scholars decided to calibrate, rather than estimate, their models (see, for example, the recent contribution of Blundell *et al.*, 2011).

The introduction or the extension of cash benefits in several countries over the last twenty years created the opportunity for the study of the various effects of these policy tools. The works by Eissa and Liebman (1995) and Meyer (2002) deal with the effects of different extensions of the

* Bank of Italy.

E-mails: mariarosaria.marino@bancaditalia.it, marzia.romanelli@bancaditalia.it, martino.tasso@esterni.bancaditalia.it

Martino Tasso acknowledges the funding of the Bank's research fellowship.

We are indebted to Sandro Momigliano and Maura Francese. We also thank Roberto Bonsignore, Fabrizio Colonna, Domenico Depalo, Valentina Michelangeli, Cecilia Moretti, Gilles Mourre, Pietro Rizza, Stefania Zotteri, the participants to the Bank of Italy 14th Public Finance Workshop, and all our colleagues at the Public Finance Division of the Research Department for a series of useful comments. We acknowledge the use of the computing facilities at the Bank of Italy and we are indebted to Giarcarlo Marra for technical assistance. The views expressed in this study are our own and do not necessarily reflect those of the Bank of Italy. Of course we are responsible for any error. Comments on this very preliminary and incomplete draft are very welcome.

Earned Income Tax Credits in the United States; Blundell *et al.* (2000) studies the English Working Families Tax Credits instead.

The Italian labor market is characterized by a particularly low participation rate among women. According to data collected by Eurostat, this rate among women between the age of 15 and 64 was just 51.1 per cent in 2010 (up from 46.3 per cent ten years earlier). The same figure was between 64.4 and 70.8 per cent in the EU, United Kingdom, Germany, France, and Spain. The average degree of labor market attachment by married women is even lower. A few studies have dealt with the effects of the Italian tax system on this outcome. A series of simulations of alternative tax systems are presented in Colombino and Del Boca (1990), Aaberge *et al.* (1999), and Aaberge *et al.* (2004). More recently, Marcassa and Colonna (2011) present some extremely interesting evidence of the high implicit tax rates imposed by the Italian tax system on the second earners. All these studies, while accounting for the main features of the tax scheme and simulating the likely effects of hypothetical reforms, model the labor supply decisions of the households in a static framework.

We contribute to this strand of literature by building and estimating a dynamic life-cycle model of household labor supply and saving decisions. Our model incorporates fiscal rules in place in the period 2005-11, as well as the main features of the family allowances. The agents in the model are heterogeneous in terms of human capital (education and on-the-job experience), and the families differ also by the number of children. We use a two-step approach to estimate the parameters of our model; like in French (2005), we recover the estimates of the parameters in the wage equations separately from the preferences. We use the method of simulated moments (or indirect inference) to estimate the values of the parameters in the agents' utility function. In this, our approach is similar to that of the study by Van der Klaauw and Wolpin (2008) on the effect of social security reforms on retirement and savings decisions by elderly in the United States.

Dynamics enters our model in several ways. First of all, agents accumulate human capital while working (like in Imai and Keane (2004)): when comparing the costs and the benefits of participation, married women take into account the fact that each additional year in the market has long-lived effects. Secondly, households are allowed to accumulate and decumulate assets, thus providing a mechanism through which they can ensure against adverse shocks on the labor market. Finally, like in all life-cycle models, agents are forward looking, and they react not only to the implementation of policies, but also to their announcement. That is, they are allowed to intertemporally adjust both consumption and labor supply.

The goal of our research is to build a model which can be used to assess the effect of changes in the tax-benefit system on female participation to the labor market. In this version of our model we present the results of a set of highly preliminary experiments. In particular, we simulate the effects of policies which could be used to increase the female participation rate directly via an increase in the household net labor income or, indirectly, giving support to the low income households which are the ones where the female participation rate is particularly low (Marcassa and Colonna, 2011). Our results are consistent with the prediction of the economic theory. In general, an increase in households' non-labor income decreases the overall poverty (in terms of head-count ratio) but lowers the incentives of married women to participate in the labor market. On the contrary, policies aimed at increasing the return of the hours worked have positive effects on both dimensions.

The rest of the paper is organized as follows. Section 2 deals with the main features of the Italian labor market, while section 3 introduces the model, explaining our solution method as well. In section 4 we illustrate the main features of the Italian fiscal system, as well as those of the family allowances. Sections 5 and 6 provide respectively an illustration of the econometric technique and

Table 1

Activity and Employment Rates (15 to 64 Years)

Country	Activity Rate					Employment Rate				
	1997	2007	2008	2009	2010	1997	2007	2008	2009	2010
European Union (EU)	67.9	70.4	70.8	70.9	71.0	60.7	65.3	65.8	64.5	64.1
Euro area (EA)	66.2	70.9	71.3	71.3	71.4	58.6	65.6	65.9	64.5	64.2
Germany (DE)	70.6	75.6	75.9	76.3	76.6	63.7	69	70.1	70.3	71.1
Spain (ES)	62.4	71.6	72.6	73.0	73.4	49.5	65.6	64.3	59.8	58.6
France (FR)	68.1	69.9	70.0	70.5	70.5	59.6	64.3	64.8	64.0	63.8
Italy (IT)	*58.2*	*62.5*	*63*	*62.4*	*62.2*	*51.3*	*58.7*	*58.7*	*57.5*	*56.9*
United Kingdom (UK)	75.4	75.5	75.8	75.7	75.5	69.9	71.5	71.5	69.9	69.5

Source: Eurostat.

the data sources we use. Some preliminary results are presented in Sections 7 and 8, while Section 9 concludes, providing a guideline for our ongoing and future work.

2 The Italian labor market

The Italian labour market is characterized by participation and employment rates considerably lower than those of the other major European economies (Table 1) and well below the objective set by the Europe 2020 strategy. Although the decade preceding the 2008 financial crisis has seen a substantial improvement in both dimensions, the gap is still far from closing. The economic crisis has further deteriorated the picture. In particular in the years 2008-10, differently from the other largest EU countries, Italy has shown a decline not only in the employment rate but also in the participation to the market.

The positive dynamics in employment observed up to the pre-crisis period was determined mainly by the expansion in part-time and temporary contracts, whose shares increased by 6.8 and 5.3 percentage points respectively in the period 1997-2007 (more than 2 and 4 times the EU average). Moreover, unemployment in Italy was and still is more likely to be of long term duration with respect to the other EU countries: in 2007 the unemployment spell was at least 12 months for more than 47.4 per cent of the Italian unemployed workers while the EU average was 42.7 per cent; in 2010 the incidence of long term unemployment increased in Italy up to 48.4 per cent, while an opposite trend was observed on average in the other EU countries (39.9 per cent).

The aggregate data hide the large disparities that affect different groups of workers and that have led to an increasing dualism of the labour market. In particular, the poor performance of the labour market partly reflects its segmentation which tends to segregate the young and the women. Indeed, these are the dimensions along which Italy records some of the largest gaps. Differences by gender and age are well reflected in activity and employment rates (Table 2).

With respect to the other European countries, the young and the female workers are particularly distressed. The participation rate registered on average in Italy in 2010 in the age group 15-24 is lower than the corresponding value for the EU economies by almost 15 p.p. (23 percentage points with respect to Germany and more than 30 percentage points compared to UK). For what concerns employment the picture is analogous, with rates largely below the other major EU countries.

Table 2

Activity and Employment Rates by Sex and Age Groups, 2010
(percent)

Age Group	Activity Rate							Employment Rate						
	EU	**EA**	**DE**	**ES**	**FR**	**Italy**	**UK**	**EU**	**EA**	**DE**	**ES**	**FR**	**Italy**	**UK**
Males														
15-24	46.1	45.5	53.7	45.1	42.9	**33.2**	61.8	36.2	35.9	47.9	25.6	33.4	**24.3**	48.5
25-49	92.4	92.9	93.6	93.2	94.8	**89.5**	92.1	84.3	84.3	86.9	75.7	87.3	**83.3**	85.9
50-54	88.0	89.8	90.9	88.4	91.5	**88.9**	87.7	81.7	83.5	84.8	75.6	86.3	**85.1**	82.4
55-64	58.9	58.2	70.8	63.9	45.2	**49.6**	69.1	54.6	53.8	65.0	54.7	42.1	**47.6**	65.0
15-64	77.6	78.2	82.3	80.7	74.9	**73.3**	81.7	70.1	70.4	76	64.7	68.1	**67.7**	74.5
Females														
15-24	39.7	39.5	48.9	40.1	35.6	**23.4**	56.4	31.8	31.6	44.6	24.2	27.2	**16.5**	46.6
25-49	79.0	78.9	81.4	80.3	84.2	**65.7**	78.7	71.7	71.0	76.4	64.4	76.7	**59.3**	74.1
50-54	73.9	73	80.9	66.7	81.2	**57.8**	78.3	68.9	67.8	76.1	56.6	75.8	**55.1**	75.5
55-64	41.2	40.9	54.5	38.5	40.0	**27.0**	51.1	38.6	38.0	50.5	33.2	37.4	**26.2**	49.5
15-64	64.4	64.5	70.8	65.9	66.1	**51.1**	69.4	58.2	57.9	66.1	52.3	59.7	**46.1**	64.6

Source: Eurostat.

Particularly affected are the women, whose participation and employment rates in 2010 were the lowest within the EU (with the exception of Malta). The gap between men and women is also impressive: it is almost double than what can be observed on average in the EU, both in terms of participation and employment rates (respectively 22.2 and 21.6 percentage points in Italy vs. 13.2 and 11.9 on average in the EU in 2010). Moreover, the gender gap enlarges sensibly in case of married workers with children and in correspondence of lower levels of education attainment (Table 3).

3 Setup of the model

We model the household's problem in a standard dynamic framework. We also assume that the decision maker is the household. The agent chooses how much to consume and how many hours to work to maximize her lifetime utility. A series of state variables affect the decision process: the agent takes into account the level of accumulated assets, and the realized labor incomes of all the components of the household, as well as the cost related to raising children under different labor market participation scenarios. Clearly, expectations about the future play a role too. Moreover, the agent knows the structure of the tax-and-transfer system and its effect of the family net income under different circumstances.

For the sake of simplicity, for the moment being, we assume that the husband is always employed in a full time-job (except when he is retired). This assumption greatly simplifies the treatment of the problem, is broadly in line with empirical data, and is not unusual in this kind of literature (see for example Eckstein and Wolpin (1989)). On the other hand, the wife can be in one of the following three states: out of the labor force, employed in a part-time job, or employed in a full-time occupation. Both husband and wife receive a new job offer at the beginning of each period. The log hourly wages follow a Mincer-type structure:

$$\log(e^h_{jt}) = \alpha^h_0 + \alpha^h_1 age^h_{jt} + \alpha^h_2 agesq^h_{jt} + \alpha^h_3 edu^h_{jt} + \varepsilon^h_{jt} \tag{1}$$

$$\log(e^w_{jt}) = \alpha^w_0 + \alpha^w_1 edu^w_{jt} + \alpha^w_2 \exp^w_{jt} + \alpha^w_3 \exp sq^w_{jt} + \alpha^w_4 pt_{jt} + \varepsilon^w_{jt} \tag{2}$$

$$\varepsilon^i_{jt} \sim N(0,\sigma^{2,i}), \quad \forall i \in \{h,w\} \tag{3}$$

The fact that women's wage equation depends on the accumulated experience allows us to incorporate in the model a new channel through which labor supply decisions (and therefore tax policy ones) may have long-lasting effects. The coefficient α_4^w captures the penalty in the hourly wage that a woman incurs when she works in a part-time occupation.

Once a member of the family reaches the age of 65, he or she retires and gets a pension which is a deterministic function of her income in the last year of employment. Every individual dies with certainty at age 85. Since wives and husbands are not necessarily the same age, the model accounts for possible periods of widowhood too.

The recursive problem can be written as follows:

$$V_t\,(X^h_t, X^w_t, A_t) = \max_{\{l^w, A_{t+1}\}}\left\{U_t + \beta E[V_{t+1}(X^h_{t+1}, X^w_{t+1}, A_{t+1})]\right\}$$

subject to:

$$\frac{A_{t+1}}{(1+r)} = A_t + \tau_t[e^h l^h + e^w l^w] - c_t - K_t$$

Table 3

Gender Employment Rate Gap by Highest Level of Education Attained and Household Composition, 2010
(percent)

Country	Single Adult with Children	Single Adult without Children	Adult Living in a Couple with Children	Adult Living in a Couple without Children
Total				
EA	–13.1	–5.0	–22.6	–11.8
DE	–11.6	0.7	–23.3	–10.3
ES	–9.7	–8.7	–22.6	–14.2
FR	–15.6	–5.8	–15.6	–5.8
Italy	*–11.2*	*–11.7*	*–34.2*	*–21.4*
UK	–17.7	–0.9	–18.3	–12.6
Pre-primary, primary and lower secondary education				
EA	–18.7	–12	–35.1	–18.6
DE	na	–1.9	–35.3	–19.6
ES	–18.6	–15.3	–32.0	–22.9
FR	–23.4	–7.0	–24.8	–6.0
Italy	*–19.6*	*–20.7*	*–49.0*	*–28.9*
UK	–18.3	–3.2	–26.6	–22.2
Upper secondary and post-secondary non-tertiary education				
EA	–11.3	–4.8	–22.1	–7.6
DE	–7.3	–0.1	–20.9	–7.3
ES	3.5	–10.8	–22.4	–10.4
FR	–16.2	–8.6	–17.3	–3.1
Italy	*–10.2*	*–8.3*	*–30.9*	*–14.2*
UK	–14	0.8	–17.9	–9.9
First and second stage of tertiary education				
EA	–7.8	–1.1	–13.8	–5.6
DE	–14.4	3.1	–17.8	–5.5
ES	–9.3	–4.2	–15.7	–2.7
FR	–2.7	–2.5	–10.7	–6.3
Italy	*–6.8*	*–6.2*	*–17.6*	*–10.2*
UK	–11.1	0.1	–14.4	–5.1

Souce: Eurostat.

where A_t is the household's net wealth at the beginning of period t, l^h and l^w are the number of hours supplied on the labor market by husband and wife respectively, and τ_t a function which replicates the main features of the tax-and-benefit system in year t. c_t is household consumption, while K_t is the cost of childcare in period t: it depends on whether there are children in the household in that period, and on the mother's labor market participation.

For the moment being, a quite simple specification is chosen for the utility function:

$$U_t = \frac{\left(\dfrac{c_t}{n_t}\right)^{1+\eta}}{1+\eta} - \phi \cdot \frac{l_t^{1+\gamma}}{1+\gamma} \tag{4}$$

The household cares about both the level of consumption and the number of hours worked. In particular, η is the coefficient of relative risk aversion, while φ and γ measure the extent of the disutility of working. This specification of the preferences has been used often by the literature on dynamic labor supply (see Imai and Keane, 2004 and Keane, 2011).

One of the main drawbacks of the standard life-cycle model is its inability to replicate well the shape of consumption pattern over time. Adjusting for the demographic characteristics of the household can help to solve this problem: consumption is hump-shaped, it tracks income, and peaks when the head of the household is in her late thirties (Fernandez-Villaverde and Krueger, 2002). To accommodate for demographics, we rescale consumption in the utility function by dividing it by the equivalent number of household members, n_t, like in Laibson *et al.* (2007) and in Attanasio and Wakefield (2010).[1]

3.1 Solution of the model

As explained above, the dynamic programming involves several continuous and discrete state variables, making a full solution infeasible in this case. Therefore, we follow an approximation method which has become customary in this kind of large estimable dynamic models (Keane and Wolpin, 1994). In a nutshell, this approach is based on choosing a random subset of the points in the state space at each point in time and solve for the optimal value function there, while approximating the same function elsewhere on the basis of a flexible function of the state variables. The solution of the model is then obtained through value function iteration, starting from the last period and working backwards. The shocks are approximated numerically through Monte Carlo integration.

The solution of the dynamic programming allows us to obtain the optimal choices of the agents in each possible situation. Because of that, we can simulate the life of our households from the first period in which we observe them in the data onwards. For each household we simulate 20 realizations of the wage shocks for both members of the couple in each period. Our simulations involve about 20,000 wage offers in each period. For each of them, and for each possible labor supply choice, we compute the income of the members of the family, net of taxes and social security contributions and the implied level of family allowances. These simulations are at the basis of our econometric strategy to recover the preference parameters.[2]

---

1 We divide total household consumption by the square root of the number of household members.

2 In order to deal with the computational burden implied by the very high number of computations, we choose Fortran 90 as programming language and we parallelize both the value function iteration and the simulation with the OpenMP libraries. Our program runs in parallel on as many as 32 processors.

Table 4

Income Brackets and Tax Rates

2005-06		2007-11	
Income Brackets (euros)	**Tax Rates**	**Income Brackets (euros)**	**Tax Rates**
0-26,000	23%	0-15,000	23%
26,000-33,500	33%	15,000-28,000	27%
33,500-100,000	39%	28,000-55,000	38%
Above 100,000	43%	55,000-75,000	41%
		Above 75,000	43%

4 The Italian tax and benefit system

The model incorporates the main features of the Italian tax-benefit system: the personal income tax (so-called Irpef) and family allowances.

Irpef is a "personal" and progressive tax. Its amount depends on specific characteristics of the taxpayer (occupation, household composition, specific expenses of a personal nature, and so on) and it is calculated applying increasing tax rates to specified income brackets (see Table 4). Horizontal and vertical equity are granted through deductions from taxable income (as for the period 2005-06) or tax credits (as for the years 2007-11) for work-related expenses and dependent people (Tables 5, 6, and 7). The amount of both instruments is inversely related and linearly dependent from income, ensuring different degrees of progressivity for different sources of income and family structures.

On the basis of these characteristics Irpef has become, since its introduction, the main tool for income redistribution policies in Italy, *i.e.*, policies aimed at alleviating the tax burden on households with low income and a large number of components. This is especially true since the Italian tax system lacks more appropriate redistribution tools, such as subsidies or a negative tax programs able to support people with tax liabilities smaller than tax credits (so-called "incapienti").

Family allowances are tax exempt public cash transfers to families with incomes below certain levels. To be eligible for these cash transfers, the sum of taxable salaries and pension incomes of the components of the household has to be at least 70 per cent of the gross family income. The amount of family allowances increases with the size of the household but it is inversely related to gross household income. Family income brackets are established by law every July and revalued each year by the percentage change in average annual index of consumer prices for the families of workers and employees, while the amount of the allowances remain unchanged. Family income limits are higher for lone parents and those with disabled persons.

The model contains the main characteristics of the Italian tax-benefit system in force in the period 2005-11 and allows the simulation of alternative schemes related to different features of Irpef and family allowances.

Table 5

Tax Deductions, 2005-06

Income Source	**Maximum Amount (DEDB) (euros)**	**Dependent People**	**Maximum Amount (DEDF) (euros)**
Dependent worker	7,500	Spouse	3,200
Pensioner	7,000	Child	2,900
Self-employed	4,500	Child younger than 3 years	3,450
Other	3,000	Child with handicap	3,700

Using:

$$\boxed{x_i = \frac{26{,}000 + DEDB - y}{26{,}000}}$$

Using:

$$\boxed{x_i = \frac{78{,}000 + DEDF - y}{78{,}000}}$$

Amount: $\begin{cases} 0, & if\ x_i \le 0 \\ x_i * DED, & if\ 0 < x_i < 1 \\ DED, & if\ x_i \ge 1 \end{cases}$

5 Econometric strategy

The goal of our econometric exercise is to estimate the parameters in the utility function of the agents. In this preliminary version of our work we focus only on the coefficient of relative risk aversion and the parameters of the disutility of working. Possible extensions, including heterogeneity in the preferences are left for the future version of this work. We identify these parameters by searching for the vector of values which minimizes a weighted distance between our observed data and the behavior of the agents simulated by our model. The strategy is that of the so-called Method of Simulated Moments (or Indirect Inference), as in McFadden (1989). More formally, the econometric problem can be explained as follows:

$$\hat{\theta} = \arg\min\{g(\theta)'Wg(\theta)\}$$

and:

$$g(\theta)' = \left[m_1^D - m_1^S(\theta),\ldots,m_J^D - m_J^S(\theta)\right]$$

where m_j^D be the j^{th} moment in the data and m_j^S the j^{th} simulated moment. The latter is found as an average across all the simulated individual observations, that is as $m_j^S = \dfrac{1}{NS} \sum m_j^s(\theta)$ where θ is the vector of parameters to be estimated.

The weighting matrix W is a diagonal matrix whose entries on the main diagonal are the inverse of the variances on the sample moments.

For the moment being, the moments used include the proportion of families in which wives participate to the labor force, work full-time, as well as the mean value of net worth. The pattern in the accumulation of the assets by the households is used to identify the coefficient of relative risk aversion, as in previous studies, such as those by Cagetti (2003) and Gourinchas and Parker (2002). The parameters governing the scale and the shape of the disutility from working are identified by the share of observations in each labor market status.

Table 6

Tax Credits for Work-related Expenses, 2007-11

Income Source	**Income Brackets** *(euro)*	**Tax Credit** *(euro)*
Dependent worker	0-8,000	1,840
	8,000-15,000	1,338+502*[(15,000–y)/7,000]
	15,000-55,000	1,338*[(55,000–y)/40,000]
	Above 55,000	0
	Plus:	
	23,000-24,000	10
	24,000-25,000	20
	25,000-26,000	30
	26,000-27,700	40
	27,700-28,000	25
Pensioner aged less than 76	0-7,750	1,725
	7,750-15,000	1,255+470*[(15,000–y)/7,500]
	15,000-55,000	1,255*[(55,000–y)/40,000]
	Above 55,000	0
Pensioner aged 76 and more	0-7,750	1,783
	7,750-15,000	1,297+486*[(15,000–y)/7,250]
	15,000-55,000	1,297*[55,000–y)/40,000]
	Above 55,000	0
Self-employed	0-4,800	1,104
	4,800-55,000	1,104*[(55,000–y)/50,200]
	Above 55,000	0

In order to obtain the optimal value of the parameters, our algorithm has to iterate between the solution of the model (and the simulation of the optimal behavior of our agents) and the minimization of the objective function. Because the objective function is likely to be discontinuous, we adopt a minimization algorithm which is based on the function values only, namely the Nelder and Mead (1965) method.

In order to alleviate the computational burden of the estimation, we choose to proceed in two steps, estimating the wage equations separately from the preference parameters. This approach is similar to that of French (2005), among others. This strategy is dictated mostly by the fact that a single dataset cannot provide all the needed information: in particular we use a different data source to estimate the wage offers, gross of any tax and social security contribution.

Table 7

Tax Credits for Dependent People, 2007-11

Dependent People	**Income Brackets** *(euro)*	**Tax Credit** *(euro)*
Spouse	0-15,000	800–110*[y/15,000]
	15,000-40,000	690
	40,000-80,000	690*[(80,000–y)/40,000]
	Above 80,000	0
	Plus:	
	29,000-29,200	10
	29,200-34,700	20
	34,700-35,000	30
	35,000-35,100	20
	35,100-35,200	10
Child	Aged 3 or more	$\boxed{(800 \cdot nchild) \cdot \frac{((95{,}000 + 15{,}000 \cdot (nchild - 1)) - y)}{(95{,}000 + 15{,}000 \cdot (nchild - 1))}}$
	Younger than 3	$\boxed{(900 \cdot nchild) \cdot \frac{((95{,}000 + 15{,}000 \cdot (nchild - 1)) - y)}{(95{,}000 + 15{,}000 \cdot (nchild - 1))}}$
	With handicap	(1)
	More than 3 children	(2)
Other dependent people		$\boxed{(750 \cdot nother) \cdot \frac{(80{,}000 + 15{,}000 \cdot (nother - 1) - y)}{(80{,}000 + 15{,}000 \cdot (nother - 1))}}$

(1) Previous formulas but 800 and 900 euros are increased by 200 euros.
(2) Maximum amount augmented by 200 euros for each child after the first one.

6 Data

We use two main sources of data. Data about family composition and asset accumulation come from the Bank of Italy Survey on Household Income and Wealth (SHIW). Data about gross labor incomes come from several waves of the EU Community Statistics on Income and Living Conditions (EU-SILC) survey. Observations are matched on the basis of comparable background information about both members of the couple. All monetary values are expressed in 2010 euros using the official price indexes computed by the Italian National Statistical Office (ISTAT).

Bank of Italy has been collecting a nationally representative household survey since the 1960s. The SHIW collects information about sources of income and wealth allocation for about 8,000 households. Since 1989, it features a longitudinal component. About half of the families are

Table 8

Descriptive Statistics

	Average	S.D.	Observations
Family-level data:			
Net worth	159,854	139,014	559
Number of kids	1.62	0.93	559
Individual-level data:			
Wife partecipation	0.51	0.5	559
Wife full-time work	0.39	0.49	559
Wife years of education	9.45	2.22	559
Husband years of education	9.33	2.15	559
Wife age	40.36	6.21	559
Husband age	43.58	6.21	559

Source: our calculations on the SHIW 2004 sample. Data in 2010 euros.

interviewed in up to five waves. Given its detailed information on assets, this dataset has been used widely in previous studies[3] and it is well suited for our research goal.

We use four continuous waves of the SHIW dataset: from 2004 to 2010, the most recent one. We focus only on married individuals, who are out of the labor force or dependent workers in each wave. Our selection decision is dictated by the fact that the rules for the determination of taxable income and some features of the tax structure are different for self-employed with respect to employees. We plan on extending our analysis to single individuals in future versions of this study. We drop very few observed households who accumulated an extremely high or extremely low level of assets. Since the SHIW is a rotating panel, our resulting sample is unbalanced. We observe 559 households in 2004: almost 70 percent of them are followed until 2010, more than 80 per cent until 2008. Overall, our resulting sample is composed of 2,792 individuals-years observations.

Table 8 reports some simple unweighted descriptive statistics about our sample in 2004. The average net worth is slightly lower than 160,000 euros. Only one every two married women is employed, while only about two fifths of them works full-time. The number of children per family is about two and it is about constant in our sampled families across the six observed years.

The EU-SILC survey is released annually within the European Statistical System. The survey aims at collecting cross-country comparable micro-data on income, poverty and social exclusion at European level. Starting in 2003 in six member states, it currently covers all EU countries. The database has both a cross-sectional and a longitudinal dimension. Concerning Italy, the survey started in 2004. The reference population is made of private households residing in the country and their current members. The sample design is a rotational one articulated in four groups

[3] See, for example, Jappelli and Pistaferri (2000).

drawn according to a stratified two-stage selection (where in the first stage municipalities are selected and in the second one households). The design attaches to each household (and to each member in the same household) a sample weight adjusted for non-response and external sources (such as the population distribution by age and sex). Over the period 2004-09 the average number of households interviewed each year is about 21,700, corresponding to 54,800 individuals (46,700 aged 15 or above). The Italian section of the EU-SILC survey includes some methodological peculiarities regarding in particular some sources of personal income, including earnings. The recorded data are indeed controlled and integrated with administrative data, via an exact match at individual level based on taxpayer identification numbers (ISTAT, 2008). This process allows for minimizing the under-reporting of the income data, making them more reliable.

In the estimation of the employee income generating process, we pool the 2004-09 waves together and select individuals aged between 25 and 55. We further restrict our sample by considering only employees and non-working women, ending up with 41,761 observations. Income is defined as the gross monthly earnings for employees, which includes only monetary earnings in the main job, gross of tax and social contributions.[4] We build hourly wages dividing these amounts by the reported number of hours worked.

Some parameters are kept constant during the estimation; this is the case of the discount rate β, which is set to 0.98, and of the annual return rate on financial investments r, which is set to 1.5 per cent, in line with other studies. Data from the 2009 survey on consumption conducted by ISTAT is used to parametrize the childcare costs, which vary according to the labor market status of the mother.

7 Preliminary results

As explained above, we estimate the parameters of the models in two separate steps. First, we estimate the wage functions separately for men and women, then we use these results to parametrize the model and estimate the preference parameters.

The log wage equations are estimated using standard techniques: ordinary least squares for men, maximum likelihood, with sample selection correction, for women. The results are shown in Table 9. As expected, the wage profile is hump-shaped. The return of an additional year of education is about 3.3 per cent for men and 4.4 per cent for women. Experience has a positive and significative effect on offered wages for women (one additional year on the job increases offered hourly wage by about 3 per cent). Part-time jobs come with a significative penalty: *ceteris paribus*, hourly wages are about 6 per cent lower than in full-time occupations.

As regards the preliminary estimates of the preference parameters (see Table 10), we find a coefficient of relative risk aversion of –2.76, which is within the range of the existing estimates. Moreover, working is associated with a sizable disutility, which varies with the number of hours worked. The standard errors around our estimates of the preference parameters are quite low.

The fit of the model to the observed data is quite good. The main features are reported in Table 11. Even though the model slightly underpredicts the average level of net worth in each wave, the asset distribution mirrors quite closely that observed in the data (Table 12). The model predicts very closely the average proportion of wives who are participating to the labor market, and the average proportion of full-time employees. In terms of net wages, the unconditional net income in 2006 is around 20,000 euros for men, while it is around 8,000 euros for women.

---

[4] We use the variable PY200G.

Table 9

First Stage Estimates

	Men Coeff	(se)	Women Coeff	(se)
Age	0.0374	(0.0028)	-	
*Age*2	–0.0003	(0.0000)	-	
Experience	-		0.0343	(0.0014)
*Experience*2	-		–0.0005	(0.0000)
Part-time	-		–0.0637	(0.0066)
Education	0.0334	(0.0006)	0.0441	(0.0007)
Married	0.0751	(0.0050)	0.0693	(0.0050)
Constant	1.087	(0.0545)	1.472	(0.0179)
Observations:	42,343		41,761	
Method:	OLS		Heckit	

Table 10

Preference Parameters

η	φ	γ
–2.757	3.046	–0.078
(0.009)	(0.026)	(0.007)

8 Policy experiments (preliminary)

The model is used to simulate the effects of four main changes to the tax-benefit system on the female participation rate and on the overall poverty.[5] The policy exercises can be divided in two main groups: changes aimed at increasing the non-labor income of the households in the lowest part of the income distribution and changes which directly influence labor income. In particular, the policy experiments belonging to the first group include: i) a 20 per cent increase in family allowances; ii) a possible refund of at most 400 euros to households whose net tax liabilities are negative (so-called *incapienti*); iii) a 35 per cent rise in child-related tax credits. The fourth simulation which consists of a 30 per cent increase in work-related tax credits affects directly labor income.

5 We define as poor a household whose net income is below the relative poverty line reported by the National Statistical Office (Istat). It should be noticed that such poverty line is calculated in terms of consumption expenditure. However in general in the lowest part of the income distribution consumption and net income tend to be of the same magnitude. As measure of poverty we consider the head-count ratio.

Table 11

Fit of the Model

	Year	Data	Model
Female participation:			
	2006	51.6	52.5
	2008	54.4	53.5
	2010	52.5	54.2
Female full-time employment:			
	2006	37.6	37.6
	2008	39.9	40.4
	2010	40.1	42
Family net wealth:			
	2006	185,113	153,996
	2008	194,900	141,849
	2010	202,386	133,026

Table 12

Distribution of the Assets in 2006
(thousands of 2010 euros)

Percentile	Data	Model
5%	3	5
10%	8	10
25%	59	40
50%	165	128
75%	278	227
90%	394	348
95%	479	435

All the experiments are announced in 2004 and implemented in 2007 (except the one concerning family allowances which is applied since 2005). This is because in 2005 and 2006 tax credits were replaced by tax deductions. The time lag allows us to also test to which extent these policies would create some inter-temporal shift in labor supply.

With respect to the baseline scenario (which simulate the actual tax-benefit system) all policy alternatives produce a reduction in net revenue amounting to around 4 per cent (defined as the algebraic sum of tax revenue, net of tax credits, of social security contributions and tax expenses for family allowances).

The model is used to simulate the optimal choices of about 10,000 families over their life-cycle, starting from the end of 2004. These optimal choices are obtained solving the dynamic programming using the optimal parameters estimated in section 7.

The main results are summarized in Table 13, which illustrate the effects of the simulated policies on the female participation rates, full-time jobs and poverty head-count ratio.

It is important to bear in mind that the treatment of unemployment in the current version of the model may play a crucial role. In particular, our model assumes that there are no frictions in the labor market. Being aware of the relevance of such assumption, it will be relaxed in the next version of the model.

As far as results as concerned, the policy experiments reduce, as expected, the overall head-count ratio. They however differ for the magnitude of the effect. In particular, it goes from a minimum of –0.4 percentage points, in the case of partially refundable tax credits, to –1.7 percentage points when an increase in child-related tax credits is implemented. Generally, the two alternatives involving tax credits produce effects which are almost twice that of the other designed policies.

Concerning the impact on the female participation rate, the policy experiments aimed at increasing the households' non-labor income are not effective, and sometimes even detrimental. In particular, an increase in the family allowances, which are not dependent from the active position of the second earner but only from the household overall income, would negatively affect both labor supply and full-time employment. This is due to the inverse relation between the amount of family allowances and household income. The same effect is obtained increasing proportionally child-related tax credits or making all tax credits (including those for the spouse) partially refundable. On the other hand, when only the work-related tax credits are increased wives' labor supply in general rises (both in terms of part-time and full-time employment). The initial decrease we observe in 2006 is exclusively due to inter-temporal shifts in labor supply related to the time lag between the announcement of the policy and its implementation. Therefore, overall, this policy experiment is the only one successful in reaching both higher female participation rates and lower headcount ratios.

9 Conclusions and agenda for ongoing work

In this work, we build and estimate a large dynamic life-cycle model of labor supply, consumption, and asset accumulation for a sample of Italian families, which were observed between 2004 and 2010. The model allows for heterogeneity across agents, and incorporates the main features of the tax-and-benefit schemes in place at that time. The goal of our research is to build a tool that could be used in the future to run a series of policy experiments in the area of taxation and labor supply. The Italian labor market is characterized by a low participation rate of married women. As highlighted by a series of previous works, the tax code may play an important role. In a set of highly preliminary results, we show the possible effect on labor supply of a short

Table 13

Policy Simulations
(preliminary)

	Year	Female Participation	Female Full-time Employment
Baseline:			
	2006	52.46	37.65
	2008	53.46	40.35
	2010	54.24	42.01
Head-count ratio in 2010: 7.24 per cent			
Increasing family allowances by 20 per cent:			
	2006	48.45	35.64
	2008	49.55	38.10
	2010	50.54	40.17
Change in net revenue in 2010: –4.10 per cent			
Change in head-count ratio in 2010: –0.84 per cent			
Making all tax credits refundable up to 400 euros:			
	2006	51.39	37.23
	2008	49.72	36.98
	2010	50.33	39.27
Change in net revenue in 2010: –4.50 per cent			
Change in head-count ratio in 2010: –0.38 per cent			
Increasing child-related tax credits by 35 per cent:			
	2006	51.40	36.73
	2008	52.76	39.14
	2010	53.32	41.19
Change in net revenue in 2010: –4.27 per cent			
Change in head-count ratio in 2010: –1.65 per cent			
Increasing work-related tax credits by 30 per cent:			
	2006	50.97	36.35
	2008	54.06	41.69
	2010	54.63	43.18
Change in net revenue in 2010: –4.35 per cent			
Change in head-count ratio in 2010: –1.34 per cent			

We compute net revenue as the algebraic sum of tax revenue, net of tax credits, of social security contributions and tax expenses for family allowances.

list of partial reforms to the system. This work can be extended in different directions. First of all, we plan to enrich the specification of the utility function, so that some forms of both observed and unobserved heterogeneity could be accounted for. This would give us the opportunity to study the differential effects of hypothetical reforms on different sectors of the population. Moreover, allowing for different *types* in the population would allow for a better treatment of the initial conditions.

The estimation of the risk aversion coefficient requires that our model captures the main aspects of the risks to which Italian families are exposed. This is unlikely to be the case in the present form of our study: in particular, we are working to incorporate a better treatment of unemployment into the setup of the model.

Both the introduction of unobserved permanent heterogeneity, and the introduction of labor market rationing through unemployment shocks are likely to increase the degree of persistence in the observed behavior of the simulated agents. We expect these features to lower the magnitude of our simulated responses to reforms to the tax and benefit system.

Finally, extending the study to a sample of single adults could allow us to investigate the role of preferences in the distribution of resources inside the household and the potential effects of taxation schemes, including those family based, on different sectors of the population.

REFERENCES

Aaberge, R., U. Colombino and S. Strøm (1999), "Labour Supply in Italy: An Empirical Analysis of Joint Household Decisions, with Taxes and Quantity Constraints", *Journal of Applied Econometrics*, Vol. 14, No. 4, pp. 403-22.

—————(2004), "Do More Equal Slices Shrink the Cake? An Empirical Investigation of Tax-transfer Reform Proposals in Italy", *Journal of Population Economics*, Vol. 17, No. 4, pp. 767-85.

Attanasio, O. and M. Wakefield (2010), "The Effects on Consumption and Saving of Taxing Asset Returns", in J. Mirrlees (ed.), *Dimensions of Tax Design: The Mirrlees Review*, pp. 675-736, Oxford University Press, Oxford and New York.

Blundell, R., M. Dias, C. Meghir and J. Shaw (2011), "The Long-term Effects of In-work Benefits in a Life-cycle Model for Policy Evaluation", CeMMAP, Working Paper.

Blundell, R., A. Duncan, J. McCrae and C. Meghir (2000), "The Labour Market Impact of the Working Families' Tax Credit", *Fiscal Studies*, Vol. 21, No. 1, pp. 75-104.

Cagetti, M. (2003), "Wealth Accumulation Over the Life Cycle and Precautionary Savings", *Journal of Business and Economic Statistics*, Vol. 21, No. 3, pp. 339-53.

Colombino, U. and D. Del Boca (1990), "The Effect of Taxes on Labor Supply in Italy", *Journal of Human Resources*, pp. 390-414.

Eckstein, Z. and O. Lifshitz (2011), "Dynamic Female Labor Supply", *Econometrica*, Vol. 79, No. 6, pp. 1675-726.

Eckstein, Z. and K.I. Wolpin (1989), "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience", *Review of Economic Studies*, pp. 375-90.

Eissa, N. and J. Liebman (1995), "Labor Supply Response to the Earned Income Tax Credit", National Bureau of Economic Research, Working Paper.

Fernandez-Villaverde, J. and D. Krueger (2002), "Consumption Over the Life Cycle: Facts from Consumer Expenditure Survey Data", National Bureau of Economic Research, Working Paper.

French, E. (2005), "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour", *Review of Economic Studies*, Vol. 72, No. 2, pp. 395-427.

Gourinchas, P. and J. Parker (2002), "Consumption Over the Life Cycle", *Econometrica*, Vol. 70, No. 1, pp. 47-89.

Haan, P. and V. Prowse (2010), "A Structural Approach to Estimating the Effect of Taxation on the Labour Market Dynamics of Older Workers", *The Econometrics Journal*, Vol. 13, No. 3, pp. S99–S125.

Imai, S. and M. Keane (2004), "Intertemporal Labor Supply and Human Capital Accumulation", *International Economic Review*, Vol. 45, No. 2, pp. 601-41.

ISTAT (2008), "L'indagine europea sui redditi e le condizioni di vita delle famiglie (Eu-Silc)", Collana Metodi e Norme, No. 37.

Jappelli, T. and L. Pistaferri (2000), "The Dynamics of Household Wealth Accumulation in Italy", *Fiscal Studies*, Vol. 21, No. 2, pp. 269-95.

Keane, M. (2011), "Labor Supply and Taxes: A Survey", *Journal of Economic Literature*, Vol. 49, No. 4.

Keane, M.P. and K.I. Wolpin (1994), "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence", *Review of Economics and Statistics*, Vol. 76, No. 4, pp. 648-72.

——————(2010), "The Role of Labor and Marriage Markets, Preference Heterogeneity, and the Welfare System in the Life Cycle Decisions of Black, Hispanic, and White Women", *International Economic Review*, Vol. 51, No. 3, pp. 851-92.

Laibson, D., A. Repetto and J. Tobacman (2007), "Estimating Discount Functions with Consumption Choices Over the Lifecycle", National Bureau of Economic Research, Working Paper.

Marcassa, S. and F. Colonna (2011), "Taxation and Labor Force Participation: The Case of Italy", THEMA (THéorie Economique, Modélisation et Applications), Université de Cergy-Pontoise, Working Paper.

McFadden, D. (1989), "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration", *Econometrica*, Vol. 57, No. 5, pp. 995-1026.

Meyer, B. (2002), "Labor Supply at the Extensive and Intensive Margins: The EITC, Welfare, and Hours Worked", *The American Economic Review*, Vol. 92, No. 2, pp. 373-79.

Nelder, J. and R. Mead (1965), "A Simplex Method for Function Minimization", *The Computer Journal*, Vol. 7, No. 4, p. 308.

Sheran, M. (2007), "The Career and Family Choices of Women: A Dynamic Analysis of Labor Force Participation, Schooling, Marriage, and Fertility Decisions", *Review of Economic Dynamics*, Vol. 10, No. 3, pp. 367-99.

Van der Klaauw, W. and K.I. Wolpin (2008), "Social Security and the Retirement and Savings Behavior of Low-income Households", *Journal of Econometrics*, Vol. 145, No. 1, pp. 21-42.

DO PUBLIC POLICIES OF A NET-REVENUE-MAXIMIZING GOVERNMENT ALSO PROMOTE INFORMALITY?

Nivedita Mukherji and Fuad Hasanov***

This paper examines the effects of fiscal and regulatory policies on the size of a country's informal economy and its government's net revenue. Introducing two types of formal goods with only one having a substitute in the informal economy, this paper finds that changes in public policies influence not only the size of the informal economy, they influence the composition of production within the formal sectors as well. Public policies that impact informality often have differential impact on the two types of formal production. This redistribution of production within the formal sector influences the impact of policies on the government's net revenue. The paper also allows some formal producers to evade taxes and informal producers to pay bribes. Tax evasion and the necessity of informal producers to pay bribes to hide their informal status further influence how public policies impact informality and distribute production within the formal sectors. Prior research on informality largely ignores multiple formal goods and fails to account for the differential impact of policies on the different formal sectors. These effects are further amplified when tax evasion and bribes are taken into consideration.

1 Introduction

In recent years the issue of production in informal sectors has drawn considerable attention. De Soto (1989) provides valuable information regarding factors which promote the development of informal markets. Although it has been recognized for long that the presence of these markets may adversely affect an economy, it is only recently that serious theoretical and empirical studies of the issue are being conducted.[1]

A large portion of the current literature has studied the effects of regulations and taxation on the size of the informal economy.[2] See Schneider and Enste (2000) for a review of many such studies. While this literature focuses on how government tax and regulatory policies promote the growth of informal economies, there is insufficient attention given to the reasons behind such policies. Marcouiller and Young (1995), Azuma and Grossman (2008) and Mukherji (2004) are some theoretical papers that study the possible rationale behind such government policies. These papers view the governments of proprietory or predatory states as agents that maximize tax revenue net of public services (termed net revenue by Azuma and Grossman and graft by Marcoullier-Young and Mukherji). Azuma and Grossman (2008) find that the distribution of productive endowments and access to private substitutes of public services impact public policies that induce some producers to operate in the informal sector. Hibbs and Pichulescu (2009) also incorporate public services and the quality of public institutions in a model of informality. They find that the incentive to operate in the informal sector is influenced by the quality of institutions

* Department of Economics, Oakland University, Rochester, MI 48309, USA. E-mail: mukherji@oakland.edu

** International Monetary Fund, Washington, DC 20431, E-mail: fhasanov@imf.org

The views expressed herein are those of the author and should not be attributed to the IMF, its Executive Board, or its management.

[1] Papers such as Viramani (1989), Goswami *et al.* (1991), Besley and McLaren (1993), Shleifer and Vishny (1993), Jain (1998), Tanzi (1994, 1998), Bardhan (1997), Johnson, Kaufmann, Zoido-Lobaton (1998a, b) view informality to be a result of corruption of officials, such as tax collectors, and show that the government is better off if such corruptions can be eliminated. Loayza (1996), Sarte (2000), Loayza, Oviedo, Serven (2005) study the adverse impact of the informal economy on the economy's growth path.

[2] Feige (1989), Cebula (1997), Johnson, Kaufmann, Zoido-Lobaton (1998a and b), Friedman, Johnson, Kaufman and Zoido-Lobaton (1999), Ihrig and Moe (2001), Fugazza and Jacques (2004) and Chong and Gradstein (2007) are some recent papers in this literature.

and governance available to private sector producers. Marcoullier and Young (1995) show that in some cases a “black hole” of graft exists when public policies aimed at maximizing graft almost drive the formal sector out of existence. Mukherji (2004) extends Marcoullier and Young’s model by endogenizing the labor supply decision of households and challenges the “black hole” result.

This paper extends the theoretical models in Marcoullier and Young (1995) and Mukherji (2004) to further examine how public policies affect informality and net revenue in a richer model. The paper’s extensions involve i) introducing government regulations ii) increasing the number and types of goods produced by the economy, iii) allowing some formal producers to evade taxes, and iv) allowing informal producers to pay bribes to stay informal. Since the empirical literature finds a strong relationship between regulations and informality,[3] the extension related to regulation is natural. The paper extends the number of goods to simply recognize that most informal goods are produced in both formal and informal sectors and that some goods like automobiles are produced in formal sectors alone. Finally, it is well documented that many formal producers evade taxes and informal producers pay many bribes to remain informal. Hence these extensions are also natural.

Schneider and Enste (2000) cautions that the conventional result that higher taxes increase informality may not be robust and must be studied in a general equilibrium context that takes into the account the impact of taxes on individual labor-leisure decisions and demand and supply of formal and informal goods. The results of this paper demonstrate that indeed in a richer model, the conventional results may not hold. Dessy and Pallage (2001) also find ambiguous effects of tax policy on informality and caution against “simple-minded” policy recommendation based on taxation.

The extensions noted above are found to have significant impact on results. The inclusion of a formal sector that has no informal counterpart introduces some interesting sectoral redistributions of production in response to policy changes. These are further amplified when tax evasion is possible and informal producers must pay bribes to maintain their status. For example, when neither tax evasion nor bribes are allowed, informality increases as the tax rate increases. This is consistent with other papers in the literature. However, when tax evasion is allowed, a higher tax rate increases the price of the good that has no informal counterpart and causes sectoral redistribution of production within the two formal sectors of the economy. This effect is further affected when informal producers must pay bribes. The interaction of the tax evasion and bribes effects reduces the the positive impact of higher tax rates on informal production. It is possible for higher tax rates to actually reduce informality if the price effect noted above is strong enough. The rearrangement of production within the two formal sectors also impacts how higher taxes affect overall tax revenue. Existing theoretical literature on informality concentrates only on the movement of labor and production between the formal and informal sectors. This paper demonstrates that public policies impact the distribution of production also within formal sectors. If this effect is ignored, the results capture only a portion of the full impact of public policies on informality and net revenue.

Robinson and Slemrod (2011) suggest that when multiple types of taxes and methods of enforcement exist, the impact of taxation on informality is influenced by the complexity of the system. Consistent with Dessy and Pallage (2001) these studies show that the effect of taxation and other public policies on informality is more complex than what some prior research suggests.

Since some production such as large scale manufacturing always remains formal, some taxes are evaded, and informal producers routinely pay bribes, it is important to incorporate them in the study of informality. To our knowledge, there is no other paper in the literature that examines this

[3] Johnson, Kaufmann and Shleifer (1997) and Friedman, Johnson, Kaufman and Zoido-Lobaton (1999) show that higher regulations of all types increase the size of the informal economy.

interaction in the context of informality. The results related to net revenue demonstrate that public policies influence the two formal sectors in opposite directions in most cases. Hence even if a change in policy increases informality, it may decrease production and revenue of one formal industry but increase the same for another. The net impact on net revenue depends on the strengths of these two opposing effects on tax revenue. Existing literature that mainly considers the presence of one formal sector fails to account for this inter-sectoral redistribution of production in the formal economy as a result of changes in public policy.

These results then also raise concerns about the choice of net revenue as the maximand for a government otherwise interested in policies that promote informal production. While theoretically it appears sensible to assume that a proprietory state would be interested in maximizing tax revenue net of some minimal productive services it must provide, the paper finds that the factors that contribute to informality do not necessarily increase net revenue. This suggests that if one needs to understand the motivations behind policies that promote informality, an alternative objective function is perhaps called for. Some metric measuring government extraction from publicly funded projects might be a better alternative.

Major implications of the relationship between public policies and both informality and net revenue are investigated empirically using data from about 50 countries. To our knowledge this paper provides the first attempt in the literature to empirically measure net revenue to study the impact of public policies on it in the context of informality. The empirical results related to informality and regulations are mostly consistent with existing literature. If indicators of democracy/bureaucracy and corruption are included in the estimation, regulations fail to have a significant impact on informality. This result is consistent with the results found in Chong and Gradstein (2007). The results on taxation and public services differ from other studies. The paper finds that higher taxes reduce informality and not increase it. This supports the theoretical result of the paper but is generally at odds with many other empirical studies cited above. Additionally, the existing literature argues that higher public services entice producers to operate in the formal sector and reduce informality. It also increases tax revenue (see Johnson and Kauffman, 1998b). While this paper finds that higher public services increase net revenue in most cases, it also increases informality. Unlike regulation, if indicators of democracy/bureaucracy and corruption are included in the estimation, public services and taxes continue to have a statistically significant impact on informality.

The empirical results related to net revenue show that higher taxes, lower regulations, and higher public services increase net revenue. Furthermore, countries with higher income, good democratic/bureaucratic and corruption indicators have higher net revenue. These are the factors that also reduce informality. These empirical results then raise concerns about the choice of net revenue as the maximand for a government otherwise interested in policies that promote informal production. While theoretically it appears sensible to assume that a proprietory state would be interested in maximizing tax revenue net of some minimal productive services it must provide, empirically the paper finds generally a negative correlation between factors that contribute to informality and the factors that increase net revenue.

Due to the lack of reliable data for countries run by dictatorships it is difficult to compare their graft or net revenue with the net revenue of other countries. However, the strength and robustness of the relationships found here for a very diverse group of countries question the ability of a government to extract increasing amounts of net revenue for itself by pursuing economically detrimental public policies. Thus policies that promote informality do not increase net revenue empirically, with the exception of public services. If public services are used to improve a country’s institutions, law and order, bureaucracy, infrastructure and such, in the long run these improvements will reduce informality.

The rest of the paper is organized as follows. Section 2 describes the theoretical model, Section 3 addresses the key theoretical results, Section 4 includes an empirical investigation, and Section 5 provides concluding remarks.

2 Description of the economy

The model-economy analyzed here is similar to the one used in Mukherji (2004) and Marcouiller and Young (1995). Individuals in this economy produce two distinct goods, H and J. Unlike Mukherji's and Marcouiller-Young's papers, one of these two goods, denoted by H, can be produced in either the formal sector or an informal sector since its production can be concealed. If it is produced in the formal sector it is called F. Otherwise it is called I. Production of the other good, J, however cannot be concealed and hence must occur in the formal sector alone. All production requires some public services, g. If production of a good occurs in the informal economy, producers have only partial access to these public services. Hence, informal producers must bear the cost of acquiring private substitutes of necessary excludable public services to remain productive.

All formal production is taxed at the rate τ. Since good H is concealable, producers of F can evade taxes. Tax evasion of good J is not possible since output is costlessly verifiable by the government.

2.1 Description of production functions

2.1.1 Good F (Good H produced in the formal sector)

Recall that output of good H can be concealed. To reduce the incidence of tax evasion that concealment makes possible, the government requires all formal producers of good H, that is producers of F, to comply with some regulations. These regulations, represented by R, determine the government's success in catching such evasions. That remains the sole purpose of regulations in this economy. In the simplest case, R is also the probability that a firm will be caught in its efforts to evade taxes. If caught, a firm pays a penalty at a rate ν. The effective tax rate in that case becomes $\tau(1+\nu) \equiv T$.

A formal producer has the choice to truthfully report all production or to conceal it. Truthful reporting necessitates paying taxes at the rate τ while efforts to conceal leads to an expected tax rate of $R\tau(1+\nu) = R*T$. If $\tau < R*T$, all formal producers will truthfully report their production. If $\tau \geq R*T$, however, producers will misreport their earnings. After-tax return to the producers of F then depends on the above tax-regulatory situation.

Case 1: $\tau < R*T$

After-tax output when all firms truthfully report their production is given by:

$$Y_F = (1-\tau)\psi*((1-R)l_F)^{1-\phi} g^{\phi} \tag{1}$$

This production function demonstrates that output depends on the amount of labor, l, and access to productive public services, g. Production in this economy is organized in units where the owner is the sole provider of labor. Hence l_F in equation (1) denotes the amount of labor supplied by a producer of good F. The term $(1-R)$ multiplying labor supply captures the reduction in productive labor services caused by regulations. ψ is a technology parameter and ϕ is a positive fraction capturing the elasticity of output to public services.

Case 2: $\tau \geq R*T$

In this scenario all firms choose to conceal their production. Hence after-tax production is given by:

$$Y_F = (1-R+R(1-T))\psi*((1-R)l_F)^{1-\phi} g^{\phi}) \tag{2}$$

Recall that a firm successfully evades taxes with probability 1–*R* and is caught with probability *R*. In case it evades, it keeps the entire output. Otherwise it retains only the fraction 1–*T*. Hence the term 1–*R*+*R*(1–*T*) in the above equation. The remaining variables and parameters are as described above.

2.1.2 Good I (Good H produced in the informal sector)

The informal sector producing good H works much like the formal sector, except that output here is not taxed and producers do not have to comply with any regulations. Producers here, however, do not have access to all public services. While some infrastructure related public services such as roads are available to all producers, certain other services are only partially available at best. Informal producers may expend some resources in the form of bribes to gain increased access to these services and in some cases provide private substitutes of these services. Thus, they have to divert some of their labor services for gaining more complete access to partially available public services and/or for the production of substitutes of the public services enjoyed by producers in the formal sector.

An informal producer is assumed to have full access to only a fraction γ of the public services g available to producers in the formal sector. By expending some effort they can increase that fraction to $\gamma + s$, where $0 < s < 1$ also represents the fraction of labor diverted for this purpose. The production function of the informal good I is then given by:

$$Y_I = \psi[(1-s)l_I]^{1-\phi}[(\gamma+s)g]^{\phi}) \tag{3}$$

A positive solution for the fraction s requires the assumption $\phi > \gamma 1 + \gamma$.

Informal producers get caught by the authorities with probability π. This probability is assumed to be proportional to the ratio of informal to total population. That is:

$$\pi = \theta\, n_I(N) \tag{4}$$

where n_I equals the number of people who produce in the informal sector, N equals total population, and θ is a positive parameter reflecting the government's success in capturing informal producers. The positive relationship between the probability π and the ratio of informal to total population is based on the observation that it is much easier to escape the authorities if a very small fraction of producers produce informally than if a much larger fraction did. The government's incentive to go after these producers will also tend to increase as the proportion rises. Once caught, however, these producers have to give up their entire output. Hence expected output of an informal producer is $(1-\pi)Y_I$.

2.1.3 Good J

This good is produced in the formal sector alone and cannot be concealed from the government. Hence production here is not subject to regulations. The production function is similar to that of good H and is given by:

$$Y_J = \delta\psi l_J^{1-\phi} g^{\phi} \tag{5}$$

where δ is a positive constant indicating that the technology used by this sector is different from the technology used in the production of good H. The elasticities of output to labor and government services are assumed to be the same as those for good H to keep the problem tractable.

2.2 Preferences and optimal consumption-labor supply decisions

The producers of goods H (F, I) and J are individuals who choose the amount of labor they supply by balancing the disutility of labor and the consumption it makes possible. The utility function of a representative producer-consumer is as follows:

$$U(H_i, J_i, l_i) = [H_i^{\sigma-1\sigma} + J_i^{\sigma-1\sigma}]^{\sigma\sigma-1} - \alpha l_i \tag{6}$$

$i = F, I, J$. This utility function shows that individuals derive utility from the consumption of goods H and J and leisure. σ is the elasticity of substitution between the two goods and α is a parameter denoting the weight of leisure in the utility function. Assuming that the output of good H produced formally and informally are indistinguishable, utility is a function of H.

2.2.1 Consumption and labor supply decisions of producers of good F

Case 1: $\tau < R*T$

When the tax and regulatory structure is such that producers report their production truthfully to the government, the budget constraint producers of F face is as follows:

$$H_F + pJ_F = (1-\tau)\psi * ((1-R)l_F)^{1-\phi} g^{\phi} \tag{7}$$

The formal good H is treated as the numeraire in this economy and p is the price of good J in terms of good H. Producers of F choose their consumption and labor supplies by maximizing the utility given by equation (6) subject to the above budget constraint. Routine calculations yield:

$$H_F = p^{\sigma} J_F \tag{8}$$

$$l_F = ((1-\phi)\psi\alpha)^{1\phi}(1-\tau)^{1\phi}(1-R)^{(1-\phi)\phi}(1+p^{1-\sigma})^{1\phi(\sigma-1)} g \tag{9}$$

Substituting from equations (8) and (9) in the budget constraint, consumption of the formal good is given by:

$$H_F = (1-\tau)^{1\phi}\psi^{1\phi}(1-R)^{(1-\phi)\phi}(1-\phi\alpha)^{1-\phi\phi}(1+p^{1-\sigma})^{1-\phi\sigma\phi(\sigma-1)} g \tag{10}$$

Indirect utility of producers of the formal good, V_F, then equals:

$$V_F = \phi(1-\phi\alpha)^{1-\phi\phi}\psi^{1\phi}(1-\tau)^{1\phi}(1-R)^{(1-\phi)\phi}(1+p^{1-\sigma})^{1\phi(\sigma-1)} g \tag{11}$$

Case 2: $\tau \geq R*T$

Case 2 parallels Case 1. The only difference here is the after-tax term in the solutions. The budget constraint in this case changes to:

$$H_F + pJ_F = (1-R*T)\psi*((1-R)l_F)^{1-\phi} g^{\phi} \tag{12}$$

The solutions are changed as follows:

$$l_F = \left((1-\phi)\psi\alpha\right)^{1\phi}(1-R*T)^{1\phi}(1-R)^{(1-\phi)\phi}(1+p^{1-\sigma})^{1\phi(\sigma-1)}g \tag{13}$$

$$H_F = (1-R*T)^{1\phi}\psi^{1\phi}(1-R)^{(1-\phi)\phi}\left(1-\phi\alpha\right)^{1-\phi\phi}(1+p^{1-\sigma})^{1-\phi\sigma\phi(\sigma-1)}g) \tag{14}$$

$$V_F = \phi\left(1-\phi\alpha\right)^{1-\phi\phi}\psi^{1\phi}(1-R*T)^{1\phi}(1-R)^{(1-\phi)\phi}(1+p^{1-\sigma})^{1\phi(\sigma-1)}g \tag{15}$$

2.2.2 Consumption and labor supply decisions of producers of good I

The budget constraint facing these producers is given by:

$$H_I + pJ_I + B = (1-\pi)\psi[(1-s)l_I]^{1-\phi}[(\gamma+s)g]^{\phi} \tag{16}$$

In this equation B represents the amount of bribes or additional expenses expended by these producers to remain informal.[4] s, as described above, is the fraction of labor services diverted by these producers to increase their access to public and/or private substitutes of public services.

Maximizing equation (6) subject to equation (16) results in the following optimal solutions:

$$s = \phi - \gamma(1-\phi) \tag{17}$$

$$H_I = p^{\sigma}J_I \tag{18}$$

$$l_I = \left((1-\phi)\psi\alpha\right)^{1\phi}(1-\pi)^{1\phi}(1+p^{1-\sigma})^{1\phi(\sigma-1)}(1-s)^{1-\phi\phi}(\gamma+s)^{1\phi}g \tag{19}$$

$$H_I = (1-\pi)^{1\phi}\psi^{1\phi}\left((1-\phi)\alpha\right)^{1-\phi\phi}(1+p^{1-\sigma})^{1-\phi\sigma\phi(\sigma-1)}(1-s)^{1-\phi\phi}(\gamma+s)^{1\phi}g - B(1+p^{1-\sigma})^{-1} \tag{20}$$

$$V_I = \phi\left((1-\phi)\alpha\right)^{1-\phi\phi}\psi^{1\phi}(1-\pi)^{1\phi}(1+p^{1-\sigma})^{1\phi(\sigma-1)}(1-s)^{1-\phi\phi}(\gamma+s)^{1\phi}g - B(1+p^{1-\sigma})^{1\sigma-1} \tag{21}$$

where V_I is the indirect utility of the informal producers.

2.2.3 Consumption and labor supply decisions of producers of good J

The problem faced by these producers parallels the one faced by the producers of good *F*. The optimal choices of consumption and leisure are also similar and are as follows:

$$l_J = \left((1-\phi)\delta\psi\alpha\right)^{1\phi}(1-\tau)^{1\phi}p^{1\phi}(1+p^{1-\sigma})^{1\phi(\sigma-1)}g \tag{22}$$

$$H_J = (1-\tau)^{1\phi}(\delta\psi)^{1\phi}p^{(1-\phi)\phi}\left(1-\phi\alpha\right)^{1-\phi\phi}(1+p^{1-\sigma})^{1-\phi\sigma\phi(\sigma-1)}g \tag{23}$$

Indirect utility of the producers equals:

[4] If producers in the formal sector have to pay bribes instead of informal producers as discussed in this paper, a negative value is assigned to *B*.

$$V_J = \phi(1-\phi\alpha)^{1-\phi\phi}(\delta\psi)^{1\phi}(1-\tau)^{1\phi}p^{1\phi}(1+p^{1-\sigma})^{1\phi(\sigma-1)}g \tag{24}$$

2.3 Equilibrium allocation of labor

In this economy, producers can freely move from one production to another. With such free mobility, for these three sectors to co-exist, utilities in all three sectors must be identical, that is $V_F = V_I = V_J$. The price that sets $V_F = V_J$, is given by:

$$p = (1-R)^{1-\phi}\delta \tag{25}$$

if $\tau < R*T$ and:

$$p = (1-R*T)(1-R)^{1-\phi}\delta(1-\tau) \tag{26}$$

if $\tau \geq R*T$.

Result 1

The price of good J is higher when taxes are evaded. This price decreases as regulations increase. The relationship between the price and the tax rate depends on the tax and regulatory condition of the economy. If they are such that producers of F truthfully report their earnings, changes in taxes do not affect the price. If the tax-regulatory structure causes producers of F to evade taxes ($1 > R(1+\nu)$), the price increases as the tax rate increases.

This result follows directly from equations (25) and (26). As regulations increase, the indirect utility of producers of good F decreases. This increases the utility of producers of good *J*. To restore equality of utilities the price of good J must decrease. A reduction in the price increases the utility of the producers of good F (the buyers of the good whose price is falling) and decreases the utility of the suppliers of good *J*. Hence a rise in regulations reduces the price of the good exempt from regulations.

When the tax rate increases it affects the producers of goods F and J equivalently if producers of good F do not evade taxes. In that event the price p does not change. If the producers of good F evade their taxes, however, taxes impact the price *p*. Differentiation of the price p in equation (26) with respect to the tax rate τ shows that the derivative is positive if $1-R(1+\nu) > 0$. (Recall ν is the penalty for tax evasion). Since $\tau \geq R*T = R*\tau(1+\nu)$ is the same as $1 \geq R*(1+\nu)$, the price of good J increases as the tax rate increases. This shows that as long as the probability of getting caught, R, and the penalty for getting caught, ν, are reasonably small compared to the tax rate, an increase in the tax rate increases the price of good *J*. This is because the marginal impact of a one unit increase in the tax rate on the producers of good *F*, $R(1+\nu)$ is less than its impact on producers of J which results in a more adverse effect on the utility of the producers of good *J*. This is compensated by an increase in the price of *J*. The condition $1 \geq R*(1+\nu)$ also indicates that the price is higher when taxes are evaded. Hence the result.

For the informal production of good H to occur in equilibrium in this economy, the utility of these producers must equal the utility of producers in other sectors. Setting $V_I = V_F$ yields:

$$\pi = 1-1(1-s)^{1-\phi\phi}(\gamma+s)^{1\phi}\left[B(1+p^{1-\sigma})^{1-\phi\phi(1-\sigma)}\phi\psi^{1\phi}(1-\phi\alpha)^{1-\phi\phi}g + (1-\tau)^{1\phi}(1-R)^{1-\phi\phi}\right]^{\phi} \tag{27}$$

if $\tau < R*T$, but:

$$\pi = 1-1(1-s)^{1-\phi\phi}(\gamma+s)^{1\phi}\left[B(1+p^{1-\sigma})^{1-\phi\phi(1-\sigma)}\phi\psi^{1\phi}\left(1-\phi\alpha\right)^{1-\phi\phi}g+(1-R*T)^{1\phi}(1-R)^{1-\phi\phi}\right]^{\phi} \quad (28)$$

if $\tau \geq R*T$. Recall that the probability of getting caught in the informal sector is proportional to the fraction of the population working there. Thus, having determined π in equations (27) and (28), the number of producers in the informal sector directly follows from equation (4).[5] Thus:

$$n_I = N\theta\left|1-1(1-s)^{1-\phi\phi}(\gamma+s)^{1\phi}\left\{B(1+p^{1-\sigma})^{1-\phi\phi(1-\sigma)}\phi\psi^{1\phi}\left(1-\phi\alpha\right)^{1-\phi\phi}g+(1-\tau)^{1\phi}(1-R)^{1-\phi\phi}\right\}^{\phi}\right| \quad (29)$$

if $\tau < R*T$, but:

$$n_I = N\theta\left|1-1(1-s)^{1-\phi\phi}(\gamma+s)^{1\phi}\left\{B(1+p^{1-\sigma})^{1-\phi\phi(1-\sigma)}\phi\psi^{1\phi}\left(1-\phi\alpha\right)^{1-\phi\phi}g+(1-R*T)^{1\phi}(1-R)^{1-\phi\phi}\right\}^{\phi}\right| \quad (30)$$

if $\tau \geq R*T$.

Given the solution for n_I, the number of producers who produce either good F or produce good J equals $N - n_I \equiv n$. Market clearing conditions in the goods market determine the distribution of producers in the two formal product markets.

Demand for good H comes mainly from the producers of good J since the formal and informal producers of good H use portions of their own production for consumption. The supply of good H equals the portion that remains after personal consumption of the formal and informal producers of H. Demand for good J equals the demand by the formal and informal producers of good H. The supply of good J equals the demand for good H by the producers of good J divided by the price of good J. This market clearing condition is given by the following equation:[6]

$$n_J H_J = n_F H_F p^{\sigma-1} + n_I H_I p^{\sigma-1} \quad (31)$$

It follows from the condition $n_F + n_J = n \equiv N - n_I$ and equation (31) that:

$$n_F = nH_J - 1p^{\sigma-1}n_I H_I H_F p^{\sigma-1} + H_J \quad (32)$$

$$n_J = nH_F p^{\sigma-1} + 1p^{\sigma-1}n_I H_I H_F p^{\sigma-1} + H_J \quad (33)$$

It follows from the equality of indirect utilities of producers producing F and J that:

$$H_F = pH_J \quad (34)$$

Equating indirect utilities of producers of F and I yields:

$$H_I = H_F + (1-\phi)\phi B(1+p^{1-\sigma})^{-1} \quad (35)$$

---

5 Note that if the relationship between π and n_I, as given in equation (4), was assumed to be non-linear, there would be no qualitative impact on the solution for n_I and hence results.

6 Note from equation (8) that $J_F = H_F P^{\sigma}$. With n_F producers of good F, total demand for good J by them equals $n_F H_F P^{\sigma}$. The value of that in terms of good H is obtained by multiplying this amount by the price p. Similar calculations explain the second term on the right hand side of equation (31).

Result 2

When informal producers must pay bribes, the loss in utility caused by the bribe is compensated in the form of higher output and consumption made possible by the lack of taxes, regulations, and free access to some public services.

This result follows from equation (35). Informal producers have a direct cost in the form of bribes that formal producers do not bear. For indirect utilities to be equalized across sectors, as is evident from a comparison of V_F and V_I, the indirect utility informal producers derive from consumption and leisure to offset bribery costs must exceed the indirect utility formal producers derive from the same factors. This is made possible by the higher output informal producers succeed in appropriating for themselves because of their ability to evade taxes, avoid regulations, and gain partial access to free public services. Comparison of V_I and V_F shows that the reduction in utility caused by the bribe, $B(1+p^{1-\sigma})^{1\sigma-1}$ is compensated in the form of higher consumption of goods H and J due to the increased output made possible by evading taxes and regulations. This extra amount equals:

$$\phi((1-\phi)\alpha)^{1-\phi\phi}\psi^{1\phi}(1+p^{1-\sigma})^{1\phi(\sigma-1)}g\left[(1-\pi)^{1\phi}(1-s)^{1-\phi\phi}(\gamma+s)^{1\phi}-(1-R*T)^{1\phi}(1-R)^{(1-\phi)\phi}\right]$$

This expression shows that this advantage increases with higher regulations and public services and thereby increases π and n_I. It also increases with higher taxes if the direct effect on it dominates the impact of taxes on the price p.

Using equations (34) and (35), the number of producers of goods F and J simplify to:

$$n_F = Np^{2-\sigma}+1-n_I-(1-\phi)\phi Bn_I H_F(1+p^{\sigma-2})(1+p^{1-\sigma}) \quad (32')$$

$$n_J = N1+p^{\sigma-2}+(1-\phi)\phi Bn_I H_F(1+p^{\sigma-2})(1+p^{1-\sigma}) \quad (33')$$

These equations complete the determination of all endogenous variables.

The above solutions for n_F, n_J, and n_I show that if informal producers do not pay any bribes, that is $B=0$:

$$n_F = N1+p^{2-\sigma}-n_I \quad (36)$$

since $H_I = H_F$. Also:

$$n_J = Np^{2-\sigma}1+p^{2-\sigma} \quad (37)$$

$$n_I = N\theta\left[1-1(1-s)^{1-\phi\phi}(\gamma+s)^{1\phi}\left\{(1-\tau)^{1\phi}(1-R)^{1-\phi\phi}\right\}^{\phi}\right] \quad (38)$$

if $\tau < R*T$. This expression is appropriately adjusted if $\tau > T*R$. The following result follows from a comparison of the solutions for number of producers when $B>0$ and when $B=0$.

Result 3

When informal producers pay bribes, the size of the informal economy is lower than when $B=0$. *The increase in the size of the formal economy caused by the reduction in informality is entirely absorbed by sector* F. *The bribe, however, causes an additional direct effect on the formal sector by moving some producers away from sector* F *to sector* J. *The number of producers of good* J *increases but the number of producers of good* F *may or may not increase when* $B>0$.

This result follows directly from equations (32) and (33). The ambiguity in the change for good F occurs because it experiences an increase due to the decrease in informal producers but experiences a loss of producers to industry J. The net change depends on which of these changes is stronger.

The following section examines the impact of government services, regulations and taxes on the distribution of producers and net revenue.

3 Impact of public services, regulations, and taxation on informality and net revenue

The last section showed that the government's tax and regulatory policies can shift production to the informal sector and also motivate some formal producers to evade taxes. A question that remains is what motivates governments to adopt policies that motivate such behaviors.

Marcouiller-Young (1995), Mukherji (2004) and Azuma-Grossman (2008) consider the government's objective to be the maximization of graft or tax revenue net of productive public services particularly in the context of predatory states. The objective of this section is to determine the relationship between this net revenue or graft and public policy instruments such as public services, tax rates, and regulations. The objective is not to determine the tax rate, regulation, and public services that maximize net revenue. Rather, the objective here is to examine how net revenue responds to each of these policy instruments for given values of the other two. This helps to answer questions such as: given the current level of public services and regulatory environment, can a government increase net revenue by taxing more?

As defined in Marcouiller and Young (1995) and Mukherji (2004), net revenue (or graft) equals tax revenue net of public services. In this paper tax revenue is obtained from the formal production of goods H and J. Thus net revenue, denoted by G, equals:

$$G = n_F R^* T\psi^* ((1-R)l_F)^{1-\phi} g^{\phi} + n_J \tau\delta\psi l_J^{1-\phi} g^{\phi} - g \tag{39}$$

Public policies impact this net revenue by changing production and by changing the sectoral distribution of producers. Analysis of this revenue is based on the assumption that the degree of substitutability between the two goods in consumption is not large ($\sigma < 1$). It follows from the solutions of labor supplies that higher taxes and regulations reduce labor supplies while higher public services increase them and these changes will have the expected changes on net revenue. That is, the decrease in labor supply as a result of higher taxes will interact with the direct impact of the higher tax rate and produce a Laffer curve type relationship. In this economy, these changes interact with the movement of labor within different sectors of the formal economy and from the formal to the informal economy. Interestingly, a sector may be impacted by regulation not because production there is subject to regulation but because regulations drive producers of other goods there. These movements are influenced by the possibility to evade taxes and the necessity to pay bribes in the informal economy, among other factors.

Result 4

When $B > 0$ *tax revenue generated by industry* J *increases. Tax revenue generated by industry* F *may increase or decrease.*

This result is a direct consequence of Result 3 which shows that the size of the informal sector is reduced. This increases the number of producers of F. However, an additional movement of producers from F to J occurs as a result of the bribe. If the decrease in F due to this effect exceeds the rise in F due to the reduction in informality, tax revenue from F will decline; otherwise it will increase. The unambiguous increase in the number of producers of J will increase tax revenue generated by that industry.

An examination of how public policies impact the distribution of producers and tax revenues follows.

3.1 Change in the tax rate

Changes in policy variables impact sectoral distribution of labor and net revenue in three ways: 1) through their direct impact, 2) by changing the price/the price channel and 3) by changing the impact of bribes on utilities of producers. The net effect is the combined effects of these three changes. The analyses below separate these effects to gain a better understanding of the changes.

Case 1: No tax evasion and no bribes

To gain an understanding of how public policies impact informality and G, it is instructive to start from the simplest case: there is no tax evasion and informal producers do not pay any bribes, that is $\tau < R*T$ and $B = 0$.

Equations (25) and (34)-(36) show that in such a situation, higher taxes do not impact the price p and the number of producers who produce good J. Higher taxes, however, increase the size of the informal economy and reduce the number of producers of F.

The solutions for l_F and l_J show that both decrease as the tax rate rises. Hence as the tax rate increases there is a decrease in the number of producers of good F and the amount of labor supplied by these producers, The negative effects of these on tax revenue is mitigated by the increase in revenue generated by the higher rate. This is also true for good J with the exception that there is no decline in number of producers here. The combined effects of the higher rate directly on revenue and indirectly through its impact on labor supply and number of producers of good F generate a Laffer curve type relationship between the tax rate and revenue.

Case 2: Tax evasion and no bribes

If the possibility of tax evasion is allowed, the main difference with Case 1 is that now the price p becomes a function of the tax rate. This creates an additional channel through which taxes impact both the sectoral distribution of producers and net revenue. Result 1 based on equation (26) shows that the price increases as the tax rate increases. Equations (36)-(38) show that as p increases, n_F decreases but n_J increases. This effect reinforces the decrease in n_F due to the direct effect of the tax change discussed in Case 1. The price change does not impact the size of the informal sector but increases the number of producers of J. Thus tax collection from production of good J increases but tax collection from production of F is reduced as the higher price drives producers away from good F to good J. The impact of this redistribution on net revenue will depend on the tax generating capacity of the two formal sectors.

Case 3: Tax evasion and bribes

If $B > 0$, equation (30) shows that n_I becomes smaller due to the additional term

$B(1+p^{1-\sigma})^{1-\phi\phi(1-\sigma)}\phi\psi^{1\phi}$. This term rises as the tax rate rises (see Result 1). Hence this will reduce the positive effect of higher taxes on informality due to the direct and price effects noted above. Informality then will not increase as much, or in the more extreme situation, decrease when $B>0$.

The relative reduction in informality will directly cause an increase in n_F (see Result 3). The change in the tax rate also impacts the term multiplying Bn_I in equations (32') and (33'). Substituting for H_F it follows that the term increases as the tax rate increases. If n_I is increased by the higher tax rate, this additional factor causes a decline in the number of producers of F. All of these producers move to sector J.

Result 5

When taxes are not evaded and informal producers do not pay bribes, higher taxes increase the number of informal producers. All of these producers are diverted from the formal sector F; there is no impact on number of producers of J. When taxes are evaded, the price of J increases and some producers move to industry J from F as taxes are increased. There is no additional impact on informality. However, if informal producers have to pay bribes, an increase in the tax rate may or may not increase informality. If informality increases, the producers will be drawn from good F. There will be a further loss of producers from good F to good J. The overall impact on net revenue depends on this redistribution and the revenue generating capacities of the two industries F and J.

3.2 Change in regulation

An increase in regulation decreases the price when both taxes are evaded and when they are not. The reduction in price becomes larger when taxes are evaded as equation (26) shows. So the impact of a change in regulation on sectoral distribution of producers and their labor supply will be in the same direction for these two cases. Hence these two cases are not treated separately for changes in R.

Case 1: Tax evasion and no bribes

When $B=0$, an increase in R increases n_I. This follows from equation (30). Also equation (36) can be rearranged as:

$$n_F+n_I=N1+p^{2-\sigma}$$

Since p decreases as R increases, the right side of the above equation increases implying that n_J decreases. While n_I increases and n_J decreases, the impact on n_F is less clear. Higher regulation drives more producers to become informal but the lowering of the price of good J stimulates some producers to good F. The price effect should be dominated by the direct impact of regulations on formal production. Hence higher regulations are expected to decrease n_F.

The reduction in n_J decreases the tax revenue from this sector as R increases. The higher R is also expected to reduce n_F and labor supply. This is offset by the increase in revenue brought about by the increased ability to catch tax evaders due to the increase in regulations. Hence the net impact of higher regulation on net revenue depends on the relative strengths of the positive and negative effects on tax collection, number of producers, and labor supplies.

Case 2: Tax evasion and bribes

When $B > 0$, equation (30) shows that the number of informal producers is smaller. However, the increase in n_I as R increases is larger. This outflow of producers to the informal sector occurs from the sector producing F. There is also a redistribution of some producers between goods F and J from equations (32), (33), and (14)). This redistribution is proportional to n_I and follows $(1+p^{1-\sigma})^{(1-\phi)\phi(1-\sigma)}1+1p^{2-\sigma}$. While n_I increases, the other term decreases with a rise in regulations. If the net change is an increase, the number of producers of F is further reduced. Otherwise the decline in F is less sharp. Net revenue depends on this redistribution.

Result 6

When $B = 0$, an increase in R increases n_I but n_J and n_F decrease. Higher regulations reduce tax revenue from industry J. Higher regulations increase tax revenue from industry F only if the direct effect of higher tax collection as a result of the increased regulation is strong enough to offset the reduction in n_F and l_F. Otherwise, net revenue will decrease with higher regulation. If $B > 0$, higher regulations will divert some producers away from F to I, further reducing revenue from F. Higher regulations additionally will cause some redistribution of producers between goods F and J. If there is an increase in the number of producers of J as a result of this redistribution, it offsets the negative impact on production of J due to the price effect. The overall impact on net revenue will depend on the net flow of producers between the sectors and the revenue generating capacities of the two formal sectors.

3.3 *Change in public services*

The price p does not depend on government services g. Like the regulation case there is no benefit in separating out the possibility of no tax evasion since there is no additional impact through the price channel brought about by tax evasion. The presence of bribes, however, matters.

Case 1: Tax evasion and no bribes

When $B = 0$, g has no impact on n_F, n_J, or n_I. However, g increases labor supplies l_F and l_J. Substitution of these labor supplies in the net revenue equation shows that the revenues are linear functions of g. Hence an increase in g increases net revenue if net revenue is positive and decreases it if net revenue is negative. If net revenue is negative, a decrease in g to 0 will eliminate the deficit by eliminating production. This is similar to Marcoullier-Young's "black hole" result with the exception that informal production will also stop.

Case 2: Tax evasion and bribes

When $B > 0$, equation (30) shows that n_I is smaller but increases as g increases. This increase occurs because the higher public service increases the value of the additional consumption informal producers enjoy as compensation for the bribes they pay. This motivates more producers to become informal.

Overall tax revenue will be higher than when $B = 0$ but declining as g increases and induces an increase in informality. If $n_I g$ increases as g increases there is an additional movement of producers out of good F to good J.

Result 7

When informal producers do not pay any bribes, there is no sectoral redistribution of producers as a result of change in government services. Net revenue increases if it is positive and decreases if it is negative. When informal producers pay bribes, the number of producers of good F is reduced and higher public services may further reduce this number. Some of these producers move to the informal sector while some may move to good J. Hence the overall number of producers in the formal sector declines and mitigates the positive effects of higher public services on production and labor supplies.

The results highlight the importance of the sectoral redistribution of production in determining the impact of public policies on net revenue. The results also show that public policies can have different impacts on different types of formal production. That is, the impacts they have on goods that have close substitutes in the informal sector (good F) are often the exact opposite of the effects they have on goods that are produced formally only (good J). This is summarized in the following result.

Result 8

When goods with substitutes in the informal sector coexist with goods which can only be produced in the formal sector and informal producers pay bribes, government tax and regulatory policies that increase informality may also increase production of the good which has no informal substitute. The loss to the economy due to higher informality may be offset by the increased production of this formal good.

These results highlight the significant sectoral redistribution of production caused by tax and regulatory policies. Policies that promote and increase informality may positively benefit an industry that has no direct connection to informal production. Thus policymakers need to be aware of redistribution of production within the formal sector since it has significant impacts on production and net revenue.

4 Empirical investigation

The previous sections developed a model that examined the combined roles of multiple goods, tax evasion, and bribery on the relationship between public policies and informality and public policies and net revenue. This section investigates empirically these relationships when such differences in economic environments for conducting business are taken into account. Lack and unreliability of cross-country data on tax evasion, bribery and relative price of goods which have informal substitutes and goods which do not, limit the scope of conducting a full-scale empirical test of the theoretical model. Nonetheless, data on governance and corruption indicators allow for the possibility of capturing the general business environment that foster activities such as tax evasion and the burden of conducting business in the formal economy. There is no formal test of the price effect of public policies and the sectoral redistribution of production within the formal sector as a result of changes in public policies. So the scope of the empirical investigation of this section is limited to the main objective of the paper - do public policies that promote informality also increase net revenue? While several papers have studied the public policy such as taxation and regulation and informality relationship empirically as the introductory section shows, there is no study that empirically considers how these policies also impact net revenue.

4.1 Data and descriptive statistics

The informal economy data come from Schneider (2004). Schneider estimates the size of the informal economy using a dynamic multiple-indicators multiple-causes framework. The informal economy is specified as a latent (unobservable) variable and various causes and indicators of the informal economy are used as observable variables. This method captures more than one "indicator" of the shadow economy as well as considers more than "one cause" in estimating the size of the informal sector. Three major types of causes identified in the literature include the burden of taxation, the burden of regulation, and citizens' attitude toward the state ("tax morality"). Three major types of indicators for the size of the shadow economy are monetary indicators (monetary transactions), developments in the labor market (movement of labor), and the developments in the production market (movement of inputs). Schneider compiles the size of the shadow economy for 145 countries for 1999-2000, 2000-01 and 2002-03. In this paper, the 1999-2000 data are used to conduct a cross-sectional analysis.

The tax rate used is the top marginal individual income tax rate obtained from the World Tax Database published by the University of Michigan. The series provide comprehensive data coverage across time and countries. The regulation variable is taken from the Heritage Foundation's component of the Index of Economic Freedom (with higher values indicating more regulation). As discussed below, to control for the quality of institutions, a democracy/bureaucracy measure that is the sum of democratic accountability and bureaucratic quality provided by the PRS Group's International Country Risk Guide (ICRG) is used. Additionally, a measure of corruption provided by ICRG (with higher values indicating better institutions) is also used. These two measures capture the general economic environment that foster activities such as tax evasion and bribery. It is worth noting that the bribery considered in the theoretical part of the paper deals with bribery in the informal sector only. Log real per capita GDP is used as another control variable and is taken from the World Bank's World Development Indicators.

Data on tax revenues and productive expenditures, necessary to compute net revenue, are obtained from the Government Finance Statistics yearbook's consolidated accounts (budgetary, extra budgetary, and social security) of the central government, published annually by the International Monetary Fund (IMF). The data expressed as percentages of GDP are available at the NYU's Development Research Institute (DRI) website. The series, however, exclude state and local government expenditures. While the tax revenue data are available, measuring government productive services is not straightforward. From a theoretical standpoint, these services include productive services that are part of formal sector firms' production functions. These services also impact firms in the informal sector although to a lesser extent. Thus to measure productive government services, government expenditures are defined as the sum of the expenditures on public order and safety, fuel and energy, and transportation and communications. Of course, this is not a perfect measure but given data limitations, it should provide a useful benchmark.[7] In addition,

[7] The issue of measurement error in the expenditure variable needs to be taken seriously since the variable is also a regressor thus potentially resulting in the errors-in-variables problem. In our estimations we use an instrumental variable approach that should mitigate this problem.

since education and health could probably be considered as productive government services affecting firms' output, an alternative analysis including these expenditures is also conducted.

To mitigate measurement problems and business cycle effects in the data, 5-year averages taken over 1995-1999 are used, except for the GDP variable that uses only 1995 data. The use of the beginning-of-the-period data reduces possible endogeneity problems and thus GDP is not instrumented in estimations below. In total, data are available for 75 countries for net revenue and productive government expenditures. However, in estimations that follow, only about 50 observations are used since there are missing data for other variables.[8]

Table 1 presents descriptive statistics of the variables used. It also gives a list of the countries that are included in the study. The choice of countries is exclusively driven by data availability considerations. The average size of the informal economy in the data is about 30 per cent of the official GDP with a range from 8.6 to 67.1 per cent. Interestingly, the informal economy has a negative correlation –0.25) with individual income tax rate, but perhaps not surprisingly, a positive correlation with regulation and institutional measures (higher values indicate stronger institutions, so the correlation coefficients are negative). The relationship with productive government expenditures excluding education/health is positive but relatively small (0.09). The average net revenue relative to GDP is about 18 per cent and with education/health expenditures, it is about 12.5 per cent. The correlation of the net revenue measure with productive government expenditures is mainly negative. Yet interestingly, expenditures with education/health and the other measure of net revenue (revenue less expenditures excluding education/health) is positive at about 0.23, which is perhaps due to education/health expenditures being incorporated in the net revenue. Lastly, higher values of net revenue are associated with higher taxes but with less regulation and better institutions.

The countries sorted by net revenue excluding education/health are shown in Table 2. Since net revenue as defined in this paper does not mean government corruption, the pattern in the data is not as straightforward. Generally, more developed countries have higher net revenues suggesting that these countries generate larger tax revenues in excess of productive government expenditures. In addition, given the definition of net revenue, it may seem that instead of measuring government's "profit", a proxy is calculated for budget surplus or deficit. However, the relationship between these measures is very weak with a correlation of less than 0.1.[9]

4.2 Estimation and results

To analyze the effects of productive government expenditures, taxes, and regulation on the informal economy and net revenue, the following equations are specified:

$$Informal_j = \alpha_0 + \alpha_1 Expend_j + \alpha_2 Tax_j + \alpha_3 Regul_j + X\delta + \varepsilon_j \qquad (40)$$

for $j = 1,2,\ldots J$.

---

[8] Future work can probably incorporate more data into the analysis and also use panel data to check for robustness of the results.

[9] The net revenue estimations discussed in the next section have also been estimated using surplus/deficit as a dependent variable. The OLS and GMM results produce mostly insignificant coefficients except for the coefficient on regulation in some instances. The coefficient on expenditures, in contrast to the net revenue estimations, is negative but insignificant in all but a few estimations at the 10 per cent level (using GMM). The GMM-CUE approach (discussed in the next section) also produces insignificant coefficients in most estimations. However, with the expenditures variable excluding education/health, in estimations using log GDP per capita and democracy/bureaucracy variables, a negative coefficient on expenditures with significance at 5 per cent (but not 1 per cent) and 10 per cent, respectively are obtained. In summary, given that other variables are insignificant and the expenditures variable is insignificant or marginally significant in a few estimations (yet with a different sign), there does not seem to be a statistical relationship between the regressors and the surplus/deficit variable.

Table 1

Descriptive Statistics and Correlation Matrix

	Informal Economy	Net Revenue	Net Revenue (educ/health)	Deficit	Expenditures	Expenditures (educ/health)	Indidual Income Tax Rate	Regulation	Log Real GDP per Capita	Democracy/ Bureaucracy	Corruption
Mean	29.98	18.34	12.52	–2.33	2.98	8.8	35.03	2.95	8.19	6.99	3.76
Standard Deviation	13.04	9.66	8.84	2.9	1.88	4.56	14.95	0.85	1.49	2.26	1.21
Minimum	8.6	–6.69	–15.98	–8.7	0.01	0.02	0	1	5.09	1.52	1.37
Maximum	67.1	39.67	30.88	9.97	12.33	21.66	61.1	5	10.69	10	6
Observations	64	75	75	75	75	75	59	71	74	61	61
Correlation Matrix											
Informal Economy	1										
Net Revenue	–0.255	1									
Net Revenue (educ/health)	–0.254	0.938	1								
Deficit	–0.241	0.085	0.096	1							
Expenditures	0.089	–0.166	–0.363	–0.117	1						
Expenditures (educ/health)	–0.006	0.231	–0.101	–0.055	0.763	1					
Indiv. Income Tax Rate	–0.249	0.411	0.43	0.006	–0.273	–0.062	1				
Regulation	0.414	–0.267	–0.19	–0.273	–0.183	–0.284	0.171	1			
Log Real GDP per Capita	–0.538	0.485	0.461	0.264	–0.119	0.076	0.012	–0.598	1		
Democracy+Bureaucracy	–0.514	0.523	0.513	0.124	–0.194	0.071	0.236	–0.343	0.71	1	
Corruption	–0.552	0.555	0.54	0.112	0.078	0.188	0.155	–0.353	0.636	0.762	1

$$NetRevenue_j = \beta_0 + \beta_1 Expend_j + \beta_2 Tax_j + \beta_3 Regul_j + X\gamma + \varepsilon_j \tag{41}$$

for $j = 1,2,\ldots J$.

Expend, *Tax*, and *Regul* variables are productive government expenditures with and without education/health, individual income tax rates, and regulation, respectively. The variable X includes log real GDP per capita and institutional measures (democracy/bureaucracy or corruption) that capture the general economic environment.

Several estimators are used for the above equations. The first estimator used is OLS. However, since the regressors could be endogenous in the above specifications resulting in inconsistent estimates, the generalized method of moments, GMM (Hansen, 1982), and the continuously updated GMM (CUE) of Hansen, Heaton, and Yaron (1996), estimators are used. The CUE has been shown to have better properties in small samples (Hansen, Heaton, and Yaron, 1996) and in the presence of weak instruments (Stock and Wright, 2000 and Stock, Wright, and Yogo, 2002). Four different instrument sets are also used: (i) constant, log real GDP per capita in 1995, latitude, and lagged values of expenditure, individual income tax rate, regulation, and corruption (averaged over 1990-94); (ii) the first set plus two interaction terms of lagged expenditure and lagged tax rate with a developing country dummy; (iii) the first set plus dummies for South Asia and British legal origin (other region and legal origin dummies are insignificant in the first stage regressions); and (iv) the first set and all regional and legal origin dummies (10 dummies).

In the above instrument sets, when the democracy/bureaucracy variable is used as a regressor, its lag rather than lagged corruption variable is used. In using lagged values of the regressors as instruments, it is assumed that the regressors are predetermined; namely, the innovation/error term is uncorrelated with the past values of regressors (a similar assumption is made in panel data models).[10] This allows for the use of GMM or CUE to obtain consistent estimates. The validity of the instruments are tested by using Hansen's (1982) *J*-test of overidentifying restrictions. Additionally, to obtain right inferences, relevant instruments are necessary. The Cragg-Donald (CD) (1993) statistic for weak instruments is used to assess the strength of the instruments in the first stage regressions. Using lagged variables rather than just regional and legal origin dummies helps alleviate the weak instrument problem as indicated by the CD statistic. Lastly, with cross-sectional regressions, country-specific effect can correlate with the regressors or instruments. Since panel data are not used, country effect cannot be differenced out. The check on this issue is the *J*-test of overidentifying restrictions, and if the test does not reject the validity of the instrument set, the equations are less likely to be misspecified.

4.2.1 The informal economy

Table 3 shows the estimation results for the informal economy as a dependent variable.[11] Estimations using government productive expenditures with and without education/health as well as using OLS and CUE are presented.[12] The instrument set used is (i) discussed above and is based on the high Cragg-Donald statistic indicating the relevance of the instruments (Stock and Yogo,

---

[10] Thus log real GDP per capita in 1995 is a valid instrument.

[11] The outlier observation for expenditures, Kuwait, is omitted in informal economy estimations, and Bahrain and Kuwait are omitted in net revenue estimations. The data for these countries have large expenditures (Bahrain: 6.8 per cent and Kuwait: 8.4 per cent with a mean of 3 per cent and standard deviation of 1.9 per cent for 75 observations of the data) and small tax revenues relative to total revenues (Bahrain: 0.31 and Kuwait: 0.04). Adding these observations to the estimations produces imprecise coefficients on expenditures and in the case of the informal economy, on tax rates as well.

[12] The GMM estimates are close to those using CUE but have a higher precision. Thus the GMM estimates result in stronger inference. Yet for the sake of brevity and since CUE is a better estimator, the CUE results are reported.

Table 2

Net Revenue, Expenditures, and Deficit
(average, 1995-99, sorted)

Country Code	Country Name	Net Revenue	Expenditures	Net Revenue	Expenditures	Deficit
NLD	Netherlands	39.667	3.084	27.449	15.302	–2.967
LUX	Luxembourg	35.998	4.98	30.882	10.096	2.276
SVN	Slovenia	34.042	3.858	24.134	13.766	–0.557
ISR	Israel	32.565	2.551	20.124	14.992	–2.402
SWE	Sweden	32.559	2.679	29.646	5.593	–3.333
GBR	United Kingdom	32.057	1.83	24.892	8.995	–2.096
DNK	Denmark	31.729	1.925	27.27	6.384	0.139
LSO	Lesotho	31.362	5.013	14.711	21.665	–0.957
SVK	Slovak Republic	31.361	4.056	18.804	16.612	–3.399
POL	Poland	30.781	2.706	24.097	9.39	–1.549
NOR	Norway	30.762	2.739	26.547	6.954	0.344
HUN	Hungary	30.124	3.346	24.104	9.366	–3.031
SYC	Seychelles	30.011	4.458	19.501	14.968	–6.265
IRL	Ireland	29.033	1.769	18.918	11.884	–0.522
CZE	Czech Republic	28.417	3.893	18.318	13.991	–0.777
BLR	Belarus	26.421	2.85	23.723	5.548	–1.829
ESP	Spain	26.203	2.052	22.819	5.435	–4.85
DEU	Germany	25.686	1.33	19.219	7.796	–1.912
EST	Estonia	25.164	4.915	16.247	13.831	0.183
FIN	Finland	25.14	2.421	20.166	7.394	–4.614
URY	Uruguay	24.992	1.189	21.079	5.102	–1.667
MLT	Malta	24.506	3.796	15.763	12.539	–6.737
LVA	Latvia	23.495	3.858	17.739	9.614	–1.626
BGR	Bulgaria	22.325	3.84	18.997	7.168	–2.858
TUN	Tunisia	22.129	3.248	14.09	11.287	–2.527
AUS	Australia	21.88	0.872	16.262	6.49	–0.042
ROM	Romania	21.723	3.243	16.372	8.594	–3.634
ISL	Iceland	21.698	3.975	10.904	14.769	–0.687
CYP	Cyprus	21.532	3.792	15.192	10.132	–3.83
ZWE	Zimbabwe	21.223	4.182	10.768	14.637	–8.464
TTO	Trinidad and Tobago	21.024	2.985	14.611	9.398	0.203
CHE	Switzerland	20.098	1.918	13.928	8.088	–1.037
MAR	Morocco	19.717	4.145	13.147	10.715	–4.397
BRA	Brazil	18.719	1.072	15.799	3.993	–7.311
USA	United States	18.062	0.836	13.546	5.352	–0.344
CHL	Chile	17.998	1.033	11.757	7.274	1.162

Country Code	Country Name	Net Revenue	Expenditures	Net Revenue	Expenditures	Deficit
GRC	Greece	17.589	2.62	12.223	7.986	–7.128
RUS	Russian Federation	17.55	1.277	16.523	2.304	–4.775
MYS	Malaysia	17.42	3.259	11.16	9.519	2.335
CAN	Canada	17.392	1.168	16.407	2.153	–1.666
BHS	Bahamas, The	16.652	2.727	9.706	9.673	–0.42
PAN	Panama	15.749	1.64	5.582	11.807	0.418
CRI	Costa Rica	15.501	2.696	6.462	11.735	–1.996
KOR	Korea, Rep.	15.501	2.79	11.645	6.647	–0.32
MUS	Mauritius	15.428	2.906	9.357	8.977	–2.045
EGY	Egypt, Arab Rep.	15.413	2.968	9.901	8.479	–1.02
VUT	Vanuatu	15.113	3.676	7.874	10.915	–2.606
TUR	Turkey	14.929	2.238	10.788	6.379	–7.359
TJK	Tajikistan	14.076	2.535	12.938	3.673	–4.656
BDI	Burundi	13.726	1.34	9.18	5.887	–5.021
SGP	Singapore	13.715	2.107	9.268	6.555	9.97
LKA	Sri Lanka	13.522	2.505	9.397	6.63	–7.089
IDN	Indonesia	13.447	1.399	11.84	3.006	–0.02
SYR	Syrian Arab Republic	13.362	3.455	10.197	6.62	–0.736
MNG	Mongolia	13.357	2.804	11.441	4.721	–8.698
MDV	Maldives	13.072	7.538	–0.569	21.178	–4.908
THA	Thailand	12.287	3.741	6.785	9.244	–3.046
DOM	Dominican Rep.	11.846	3.078	7.854	7.07	0.365
MEX	Mexico	11.596	1.277	7.333	5.541	–0.609
ARG	Argentina	11.278	1.235	10.018	2.494	–1.58
KAZ	Kazakhstan	10.791	2.057	8.309	4.539	–4.041
BOL	Bolivia	10.019	3.898	4.798	9.119	–2.327
CMR	Cameroon	9.417	0.81	7.212	3.015	0.841
IND	India	8.889	0.233	8.317	0.805	–4.918
SLV	El Salvador	8.681	3.705	4.972	7.414	–1.807
YEM	Yemen, Rep.	8.38	3.328	1.707	10.001	–3.15
COL	Colombia	8.299	2.032	3.284	7.047	–4.364
MDG	Madagascar	7.598	0.749	4.748	3.599	–1.452
IRN	Iran, Islamic Rep.	7.538	4.362	1.591	10.31	–1.022
NPL	Nepal	4.482	4.23	1.215	7.497	–4.088
MMR	Myanmar	2.503	1.406	1.244	2.666	–2.16
BHR	Bahrain	0.981	6.799	-4.599	12.379	–4.755
HRV	Croatia	0.037	0.006	0.027	0.016	–0.001
BTN	Bhutan	–5.177	12.33	–12.523	19.675	–0.004
KWT	Kuwait	–6.689	8.386	–15.976	17.674	–7.059

Table 3

Estimation of the Informal Economy

	Excluding Education/Health										Including Education/Health									
	OLS					CUE					OLS					CUE				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Expenditures	3.24	1.62	2.15	3.73	2.97	3.79	2.46	2.75	4.47	4.72	0.90	0.59	0.58	0.93	0.71	1.58	0.50	0.80	1.59	1.43
Standard error	1.32	1.58	1.28	1.30	1.66	1.63	1.53	1.06	1.11	1.29	0.51	0.51	0.56	0.55	0.57	0.60	0.60	0.44	0.46	0.53
p -value	0.02	0.31	0.10	0.01	0.08	0.02	0.11	0.01	0.00	0.00	0.08	0.25	0.31	0.10	0.22	0.01	0.41	0.07	0.00	0.01
Indiv. Income Tax Rate	–0.36	–0.35	–0.32	–0.31	–0.32	–0.73	–0.19	–0.28	–0.26	–0.28	–0.40	–0.38	–0.35	–0.35	–0.36	–0.84	–0.30	–0.35	–0.34	–0.35
Standard error	0.17	0.13	0.15	0.14	0.13	0.18	0.18	0.14	0.15	0.15	0.18	0.14	0.15	0.15	0.13	0.22	0.17	0.15	0.16	0.15
p-value	0.04	0.01	0.03	0.03	0.01	0.00	0.27	0.05	0.08	0.05	0.03	0.01	0.03	0.02	0.01	0.00	0.07	0.02	0.03	0.02
Regulation	7.48	1.65	2.93	3.27	1.86	6.79	–0.47	–0.26	1.22	2.04	8.65	2.19	3.60	4.76	2.10	9.13	0.51	1.51	4.93	3.03
Standard error	1.58	1.85	2.73	2.01	2.00	2.30	2.16	2.25	1.99	3.02	1.81	1.95	2.64	2.09	2.23	2.13	2.17	2.31	2.04	3.21
p-value	0.00	0.38	0.29	0.11	0.36	0.00	0.83	0.91	0.54	0.50	0.00	0.27	0.18	0.03	0.35	0.00	0.81	0.51	0.02	0.35
Log of GDP/Capita		–5.95			–2.27		–6.17			1.03		–6.18			–3.72		–6.54			–1.69
Standard error		1.66			2.20		1.28			3.06		1.49			2.11		1.34			2.64
p-value		0.00			0.31		0.00			0.74		0.00			0.09		0.00			0.52
Democracy+Bureaucracy			–2.84					–4.49					–2.95					–4.93		
St. error			1.44					1.07					1.48					1.04		
p-value			0.06					0.00					0.05					0.00		
Corruption				–5.72	–4.33				–6.89	–7.84				–5.41	–3.24				–7.46	–5.97
Standard error				1.41	2.11				1.52	3.36				1.42	1.97				1.58	2.78
p-value				0.00	0.05				0.00	0.02				0.00	0.11				0.00	0.03
Constant	12.99	84.18	48.92	44.39	64.80	29.81	83.56	67.64	52.02	44.75	11.77	84.86	49.18	41.96	75.22	24.18	90.15	69.20	44.96	60.65
Standard error	9.71	20.61	17.63	10.73	19.34	11.83	16.66	14.01	10.83	23.96	8.84	17.17	14.82	8.54	18.77	11.91	16.39	11.66	8.49	24.21
p-value	0.19	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.06	0.19	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.01
R^2	0.36	0.53	0.46	0.55	0.56						0.34	0.53	0.44	0.51	0.55					
Adjusted R^2	0.32	0.48	0.40	0.51	0.51						0.30	0.49	0.39	0.46	0.49					
J-test						5.09	3.23	1.31	0.81	0.69						5.42	1.78	1.52	0.47	0.07
p-value						0.17	0.20	0.52	0.67	0.40						0.14	0.41	0.47	0.79	0.80
CD stat						9.19	10.22	9.39	9.02	4.27						18.39	14.56	10.66	10.85	4.10
Number of observations	49	49	47	47	47	46	46	44	44	44	49	49	47	47	47	46	46	44	44	44

Notes: Heteroskedasticity-consistent standard errors.
Instrument set: Log GDP/capita in 1995, latitude, and lagged expenditure, individual tax rate, regulation, and corruption (1990-94).
When Democracy+Bureaucracy variable is used in estimations, lagged democracy/bureaucracy rather than lagged corruption is used in the instrument set. *J*-test: Test of overidentifying restrictions.
CD stat: Cragg-Donald statistic for weak instruments.

2004).[13] In addition, in all estimations, the *J*-test of overidentifying restrictions does not reject the null hypothesis of the validity of the instruments.

The impact of government expenditures without education/health using the OLS estimator is positive and large with a coefficient between 1.5 and 4. It is, however, imprecise in two of four specifications. Since there could be endogeneity problems with the OLS estimator, the CUE is examined. The coefficients are more precise and larger, about 2.5 to 4.5. These numbers imply that everything else constant, a one percentage point increase in productive government expenditures relative to GDP, increases the informal sector by 2.5-4.5 percentage points of official GDP. This is a large impact and the theory above confirms this finding.

Including education/health into the expenditures measure produces low and imprecise coefficients using OLS. However, the CUE results in more precise estimates. The parameters are smaller than in the estimation without education/health – at approximately 1.5. Perhaps the effect is smaller and is not as precise as before since the inclusion of education/health expenditures does not impact firms' incentives immediately. It may take years before a more educated and healthy workforce may impact the firms' decision in terms of the benefits and costs of operating in the informal economy.

The results also show that size of the informal economy increases with more regulation, worse institutions or lower level of development, and lower individual income tax rates. The coefficients are statistically significant (at 5 or 10 per cent) but including regulation and institutions or log GDP variables together results mostly in an imprecise coefficient on regulation perhaps suggesting some collinearity issues. The inclusion of both corruption and GDP variables confirms the significance of corruption and results in similar parameter estimates. It is not surprising that higher regulation and worse institutions imply a higher informal sector. This also suggests that the cost of these factors on the formal economy is stronger than on the informal and drives production to the informal sector. A higher cost of the formal sector suggests that the value of B in the theoretical section of the model should perhaps be negative. Another interesting result is that higher income tax rates imply a lower informal sector of the economy. However, this is consistent with the theoretical findings (see Section 3.1 for details). Finally, another important result is that higher public services that increase informality as well as net revenue (see the next section), may not promote informality in the longer run. Although one percentage increase in public services increases informality by about 4.5 per cent, improving corruption environment from that of Bulgaria to that of Australia results in a decrease of the informal sector by about 7 per cent. With higher public services on law and order, infrastructure, communications, country's institutions would improve and thus reduce informality, which is confirmed by the empirical results. Thus, if the goal of states is to use such a policy to maximize net revenue, it may not be an informality-increasing policy in the longer run.

4.2.2 Net revenue

Table 4 presents estimations for net revenue. Excluding education/health and using the OLS estimator, the impact of productive government expenditures is positive and statistically significant at 5 or 10 per cent level. It seems that the expenditure variable creates a problem of simultaneity since expenditures are subtracted from tax revenues to arrive at net revenue while the same expenditures variable is also used as a regressor. However, it is precisely because expenditures are

[13] Using more instruments that include regional and legal origin dummies [instrument set (iv)], reduces the CD statistic to about 4, which is indicative of weak instruments. These estimations result in a higher precision of our estimates; however, given a weak instrument set, we cannot rely much on the inference. The results using instrument sets (ii) and (iii) are in general similar to those using (i). However, the CD statistic is smaller in size compared to that of instrument set (i).

Table 4

Estimation of Net Revenue

	Excluding Education/health										Including Education/health									
	OLS					CUE					OLS					CUE				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Expenditures	1.87	2.61	2.19	1.40	2.28	3.10	2.61	3.18	1.65	2.15	0.19	0.34	0.29	0.10	0.20	0.32	0.19	0.32	–0.09	–0.08
Standard error	0.77	0.76	0.81	0.82	0.85	1.00	0.63	0.80	0.79	0.86	0.21	0.20	0.24	0.22	0.19	0.24	0.22	0.30	0.31	0.31
p-value	0.02	0.00	0.01	0.09	0.01	0.00	0.00	0.00	0.04	0.01	0.37	0.09	0.24	0.64	0.30	0.19	0.39	0.28	0.77	0.80
Indiv. Income Tax Rate	0.19	0.15	0.11	0.12	0.13	0.59	0.07	0.09	0.06	0.07	0.14	0.10	0.06	0.07	0.07	0.27	0.05	–0.03	–0.01	–0.01
Standard error	0.10	0.08	0.09	0.09	0.08	0.19	0.08	0.10	0.10	0.09	0.09	0.07	0.08	0.08	0.07	0.11	0.08	0.09	0.09	0.09
p-value	0.06	0.07	0.22	0.18	0.11	0.00	0.36	0.40	0.54	0.42	0.11	0.16	0.47	0.38	0.33	0.01	0.48	0.76	0.90	0.90
Regulation	–4.31	–0.18	–2.23	–2.86	–1.01	–3.56	2.79	0.30	–0.76	0.44	–3.02	0.94	–0.45	–1.22	0.10	–2.49	4.82	2.65	0.98	1.13
Standard error	1.05	1.24	1.46	1.08	1.18	1.39	1.61	1.32	1.19	1.78	1.03	1.19	1.16	1.00	1.09	0.98	1.61	1.45	1.21	1.99
p-value	0.00	0.89	0.13	0.01	0.39	0.01	0.08	0.82	0.52	0.81	0.00	0.43	0.70	0.23	0.93	0.01	0.00	0.07	0.42	0.57
Log of GDP/Capita		4.01			2.94		5.29			1.68		3.71			1.84		5.65			0.19
Standard error		0.66			1.17		0.88			2.01		0.78			1.02		1.01			1.96
p-value		0.00			0.02		0.00			0.40		0.00			0.08		0.00			0.92
Democracy+Bureaucracy			1.73					3.41					1.78					3.42		
Standard error			0.77					0.59					0.61					0.68		
p-value			0.03					0.00					0.01					0.00		
Corruption				2.94	1.12				5.35	3.69				3.11	2.01				5.23	5.06
Standard error				0.79	1.17				0.95	2.10				0.73	1.01				0.96	2.01
p-value				0.00	0.35				0.00	0.08				0.00	0.05				0.00	0.01
Constant	20.70	–25.87	3.99	9.09	–16.82	–1.89	–43.14	–18.02	–4.94	–18.22	16.21	–26.59	–1.96	2.52	–13.49	7.49	–51.96	–20.92	–7.64	–9.17
Standard error	6.04	9.11	10.02	7.14	10.81	9.31	10.63	7.86	5.55	16.21	5.39	8.92	7.13	6.28	8.56	5.92	11.70	8.19	6.08	16.89
p-value	0.00	0.01	0.69	0.21	0.13	0.84	0.00	0.02	0.37	0.26	0.00	0.00	0.78	0.69	0.12	0.21	0.00	0.01	0.21	0.59
R^2	0.35	0.52	0.48	0.50	0.52						0.21	0.42	0.40	0.44	0.45					
Adjusted R^2	0.31	0.48	0.44	0.46	0.46						0.17	0.38	0.35	0.40	0.38					
J-test						9.72	3.22	0.49	0.72	0.19						10.70	3.82	0.81	0.26	0.25
p-value						0.02	0.20	0.78	0.70	0.66						0.01	0.15	0.67	0.88	0.62
CD stat						12.56	13.27	10.69	10.06	2.86						22.65	15.07	12.80	11.58	2.86
Number of observations	56	55	52	52	51	52	52	48	48	48	56	55	52	52	51	52	52	48	48	48

Notes: Heteroskedasticity-consistent standard errors.
Instrument set: Log GDP/capita in 1995, latitude, and lagged expenditure, individual tax rate, regulation, and corruption (1990-1994).
When Democracy+Bureaucracy variable is used in estimations, lagged democracy/bureaucracy rather than lagged corruption is used in the instrument set. *J*-test: Test of overidentifying restrictions.
CD stat: Cragg-Donald statistic for weak instruments.

subtracted from tax revenues, they are no longer part of the net revenue measure, which should avoid the simultaneity problem. Nonetheless, the expenditure variable could be endogenous along with other regressors; that is, they could be correlated with the innovation/error term, so the CUE was used. The coefficient becomes larger in magnitude and more precise. The estimations imply that if productive government expenditures increase by one percentage point relative to GDP, net revenue rises by about 2-3 percentage points relative to GDP. However, introducing education/health into the expenditures variable results in a very small and insignificant coefficient. This suggests that health/education expenditures may not have an immediate impact on net revenue, and it may take time before the benefits of better health and education are reaped through higher productivity and higher tax revenues.

Similar to the informal sector estimations, the impact of the level of development and institutions variables is highly statistically significant and large indicating that worse institutions and lower level of development decrease net revenue. Tax rates positively affect net revenue while regulation has a negative impact. However, the impact of taxes is small (0.1) and statistically insignificant. *J*-test of overidentifying restrictions rejects the null at 5 per cent level in specifications using only a regulation variable. Introducing GDP or institutions variables, the regulation variable becomes statistically insignificant. Interestingly, the coefficient becomes positive, which implies higher regulation increases net revenue,[14] and significant at 10 per cent in a couple of estimations using the CUE and mostly in estimations including education/health. However, the evidence of positive impact is not conclusive, and the coefficient is statistically significant in only a couple of estimations.

5 Conclusion

The paper finds that the inclusions of tax evasion by formal producers, bribes paid by informal producers, and multiple types of goods significantly affect how public policies affect informality and net revenue.

Changes in public policies cause changes in the price of the good that has no informal sector. This price change causes changes in the number of producers of this formal good. Often these producers are drawn from the formal good that has an informal sector. Hence public policies shift producers within the two formal sectors. The literature on informality largely fails to account for this production redistribution.

Furthermore, when informal producers pay bribes to maintain their status, informality is reduced. The producers that remain informal, however, derive more utility from direct consumption than their formal counterparts to compensate for the loss of income and utility caused by the bribe. This additional utility is made possible by avoiding regulations and taxes and equivalently captures the value of the bribe to an informal producer in terms of lost utility. These utility effects depend on the values taken by public policy variables. This factor further impacts the distribution of producers between the various sectors.

As public policies redistribute production, it often impacts the two formal sectors in opposite directions. Whether tax revenue rises in response to a policy change depends on the relative responsiveness of the two sectors to policy instruments. Hence the paper demonstrates the importance of taking into consideration multiple formal sectors and bribes in studies of informality.

Empirically, the paper finds:

[14] The positive coefficient on regulation is also consistent with the theory presented.

- Productive public expenditures increase net revenue. Once education and health expenditures are added, the result becomes statistically insignificant. As mentioned above, expenditures related to health and education have more longer term than immediate effect on current production. Hence, the results without education and health may be more appropriate for the current study.
- Taxes have a positive but small impact on net revenue. Once institutional variables are considered, taxes fail to have any statistically significant effect on net revenue.
- The impact of regulation on net revenue is mixed. Estimations which yield a significant impact of regulations show that if GDP is included in the estimation, higher values of regulations increase net revenue. However, if GDP is not included, in most other instances where regulation has a significant effect, higher regulations are associated with lower net revenue.
- GDP and institutional variables have a large and statistically strong impact on net revenue. They also show that countries with better institutions and higher level of development have higher net revenue.

These results show that it is possible to increase net revenue by having higher taxes, more regulations, and higher public services. With the exception of taxes (which has a small, if any, effect on net revenue) these factors also increase informality. The results also show that to achieve higher net revenue, institutional reforms in the form of better bureaucratic quality and democratic accountability and less corruption are desirable. Once these institutional factors are introduced, while public services continue to remain significant, the effects of regulations and taxes on net revenue weaken. Furthermore, good institutions are usually not present in countries with predatory governments. Hence to understand why countries engage in policies that increase informality, researchers may want to consider an alternative objective.

REFERENCES

Azuma, Y. and H. Grossman (2008), “A Theory of the Informal Sector”, *Economics and Politics*, Vol. 20, No. 1, pp. 62-79.

Bardhan, P. (1997), “Corruption and Development: A Review of Issues”, *Journal of Economic Literature,* Vol. 35, No. 3, pp. 1320-46.

Besley, T.J. and J. Mclaren (1993), “Taxes and Bribery: The Role of Wage Incentives”, *The Economic Journal*, pp. 119-41.

Cebula, R.J. (1997), “An Empirical Analysis of the Impact of Government Tax and Auditing Policies on the Size of the Underground Economy: The Case of the United States 1993-94”, *American Journal of Economic Society*, Vol. 56, No. 2, pp. 173-85.

Chong, A. and M. Gradstein (2007), “Inequality and Informality”, *Journal of Public Economics*, Vol. 91, pp. 159-79.

de Soto, Hernando (1989), *The Other Path: The Invisible Revolution in the Third World*, June Abbot, translation, New York (NY), Harper and Row.

Feige, E.L. (ed.) (1989), *The Underground Economies. Tax Evasion and Information Distortion*, Cambridge (MA), Cambridge University Press.

Friedman, E., S. Johnson, D. Kaufmann and P. Zoido-Lobaton (1999), “Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries”, World Bank, Discussion Paper.

Goswami, O., I.N. Garg and A. Sanyal (1991), “Bureaucratic Corruption, Income Tax Enforcement and Collection”, mimeo.

Jain, A. (ed.) (1998), *Economics of Corruption*, Boston (MA), Kluwer Academic Publishers.

Johnson, S., D. Kaufmann and A. Schleifer (1997), “The Unofficial Economy in Transition”, Brookings Papers on Economic Activity, Vol. 1997, No. 2, pp. 159-239.

Johnson, S., D. Kaufmann and P. Zoido-Lobaton (1998a), “Regulatory Discretion and the Unofficial Economy”, *American Economic Review,* Vol. 88, No. 2, pp. 387-92.

—————(1998b), “Corruption, Public Finances and the Unofficial Economy”, The World Bank Discussion Paper.

Loayza, N.V. (1996), “The Economics of the Informal Sector: A Simple Model and Empirical Evidence from Latin America”, Carnegie-Rochester Conference Series on Public Policy, No. 45, pp. 129-62.

Loayza, N.V., A.M. Oviedo and L. Serven (2005), “The Impact of Regulation on Growth and Informality – Cross-country Evidence”, The World Bank, Policy Research, Working Paper.

Marcouiller, D. and L. Young (1995), “The Black Hole of Graft: The Predatory State and the Informal Economy”, *American Economic Review*, Vol. 85, No. 3, pp. 630-46.

Mukherji, N. (2004), “Government Policies and Graft in an Economy with Endogenous Labor Supply”, *Journal of Development Economics*, Vol. 73, pp. 423-34.

Sarte, P.D. (2000), “Informality and Rent-seeking Bureaucracies in a Model of Long-run Growth”, *Journal of Monetary Economics*, No. 46, pp. 173-97.

Schneider, F. and D. Enste (2000), “Shadow Economies: Size, Causes and Consequences”, *Journal of Economic Literature*, Vol. 38, No. 1, pp. 77-114, March.

Schneider, F. (2002), “Size and Measurement of the Informal Economy in 110 Countries Around the World”, The World Bank, Working Paper, July.

Shleifer, A. and R.W. Vishny (1993), “Corruption”, *Quarterly Journal of Economics,* Vol. 108, No. 4, pp. 559-617.

Tanzi, V. (1994), “Corruption, Governmental Activities, and Markets”, IMF, Working Paper, No. 99, pp. 1-20.

———— (1998), “Corruption Around the World: Causes, Consequences, Scope, and Cures”, IMF, Working Paper, No. 63, pp. 1-39.

Virmani, A. (1989), “Indirect Tax Evasion and Productive Efficiency”, *Journal of Public Economics*, No. 39, pp. 223-37.

ECONOMIC PERFORMANCE, GOVERNMENT SIZE, AND INSTITUTIONAL QUALITY

António Afonso[*, **] *and João Tovar Jalles*[**, ***]

We outline a growth model with an explicit government role, where more government resources reduce the optimal level of private consumption and per worker output. For an unbalanced country panel we use different proxies for government size and institutional quality. Our results, consistent with the model, show a negative effect of the size of government on growth. Similarly, institutional quality has a positive impact on real growth, and government consumption is consistently detrimental to growth. Moreover, the negative effect of government size on growth is stronger the lower institutional quality, and the positive effect of institutional quality on growth increases with smaller government size. The negative effect on growth of the government size variables is more mitigated for Scandinavian legal origins, and stronger at lower levels of civil liberties and political rights.

1 Introduction

Governments tend to absorb a sizeable share of society's resources and, therefore, they affect economic development and growth in many countries.[1] Throughout history high levels of economic development have been attained with government intervention. Where government did not exist, little wealth was accumulated. However, despite necessary, government intervention is not a sufficient condition for prosperity, if it leads to the monopolization of the allocation of resources and other important economic decisions, and societies do not succeeded in attaining higher levels of income.[2]

In addition, economic progress is limited when government is zero per cent of the economy (absence of rule of law, property rights, etc.), but also when it is closer to 100 per cent (the law of diminishing returns operates in addition to, e.g., increased taxation required to finance the government's growing burden – which has adverse effects on human economic behaviour, namely on consumption decisions). This idea is related to the so-called "Armey Curve", after Richard Armey, who borrowed a graphical technique popularized by Arthur Laffer, whose crucial underpinnings were already present in Dupuit (1844). Friedman (1997) suggested that the threshold where government's role in economic growth is between 15-50 per cent of the national income.

The existing literature also presents mixed results as to the relationship between government size and economic development (for a recent survey see Bergh and Henrekson, 2011). Important differences in existing research concern the measurement of government size, the type of countries studied (rich vs. poor) and the time span considered. On the one hand, the former may impact economic growth negatively due to government inefficiencies, crowding-out effects, excess burden

* ISEG/UTL - Technical University of Lisbon, Department of Economics; UECE – Research Unit on Complexity and Economics. UECE is supported by FCT (Fundação para a Ciência e a Tecnologia, Portugal), E-mail: aafonso@iseg.utl.pt

** European Central Bank, Directorate General Economics, Kaiserstraße 29, D-60311 Frankfurt am Main, Germany. E-mail: antonio.afonso@ecb.europa.eu and joao.jalles@ecb.europa.eu

*** University of Aberdeen, Business School, Edward Wright Building, Dunbar Street, AB24 3QY, Aberdeen, UK. E-mail: j.jalles@abdn.ac.uk

1 According to the Wagner's Law the scope of the government usually increases with the level of income because government has to maintain its administrative and protective functions, its attempts to ensure the proper operation of market forces and provision of social and cultural (public) goods.

2 Public choice explanations of government growth are discussed in Holcombe (2005).

of taxation, distortion of the incentives systems and interventions to free markets (Barro, 1991; Bajo-Rubio, 2000). Indeed, several studies report that the efficiency of government spending can increase, either by delivering the same amount of services with fewer resources or by using more efficiently existing spending levels (see Afonso *et al.*, 2005, 2011; Angelopoulos *et al.*, 2008). Moreover, Slemrod (1995) and Tanzi and Zee (1997) find a negative impact if the size of government exceeds a certain threshold. The rationale behind this argument is that in countries with big governments the share of public expenditures designed to promote private sector productivity is typically smaller than in countries with small governments (Folster and Henrekson, 2001). On the other hand, government activities may also have positive effects due to beneficial externalities, the development of a legal, administrative and economic infrastructure and interventions to offset market failures (Ghali, 1998; Dalagamas, 2000). On the debate between the positive vs. negative effects of government growth, Grossman (1988) suggested that a non-linear model was preferred in explaining its impact on total economic output.

Our motivation also comes from Guseh (1997) who presents a model that differentiates the effects of government size on economic growth across political systems in developing countries. Growth in government size has negative effects on economic growth, but the negative effects are three times as great in non-democratic systems as in democratic systems.

Our paper includes several contributions: i) we first outline a growth model allowing for an explicit government role, we characterize the conditions underlying the optimal path of the economy and determine the steady-state solutions for the main aggregates; ii) we analyse a wide set of 108 countries composed of both developed and emerging and developing countries, using a long time span running from 1970-2008, and employing different proxies for government size and institutional quality to increase robustness; iii) we build new measures of extreme-type political regimes which are then interacted with appropriate government size proxies in non-linear econometric specifications; iv) we make use of recent panel data techniques that allow for the possibility of heterogeneous dynamic adjustment around the long-run equilibrium relationship as well as heterogeneous unobserved parameters and cross-sectional dependence (e.g. Pooled Mean Group, Mean Group, Common Correlated Pooled estimators, inter alia); and vi) we also deal with potentially relevant endogeneity issues.

Our results show a significant negative effect of the size of government on growth. Similarly, institutional quality has a significant positive impact on the level of real GDP per capita. Interestingly, government consumption is consistently detrimental to output growth irrespective of the country sample considered (OECD, emerging and developing countries). Moreover, i) the negative effect of government size on GDP per capita is stronger at lower levels of institutional quality, and ii) the positive effect of institutional quality on GDP per capita is stronger at smaller levels of government size.

On the other hand, the negative effect on growth of the government size variables is more attenuated for the case of Scandinavian legal origins, while the negative effect of government size on GDP per capita growth is stronger at lower levels of civil liberties and political rights

The remainder of the paper is organised as follows. Section two presents the theoretical model, which underlies and motivates the empirical specifications. Section three addresses data-related issues. Section four elaborates on the econometric methodology and presents and discusses our main results. Section five concludes the paper.

2 Model and econometric specification

In this section we present a growth model that relates output and government size and it will provide the theoretical motivation for our empirical (panel) analysis in Section 3. Our model fits within a broader literature that expands a Barro (1991)-type model where government plays an active role.[3] We consider a typical economy with a constant elasticity of substitution utility function of the representative agent given by:

$$U = \int_0^\infty e^{-\gamma t} \frac{c_t^{1-\theta} - 1}{1-\theta} dt \tag{1}$$

where c is per capita consumption, θ is the intertemporal substitution and γ is the (subjective) time discount rate or rate of time preference (a higher γ implies a smaller desirability of future consumption in terms of utility compared to utility obtained by current consumption. Population (which we assume identical to labour force, L) grows at the constant rate n, that is, $L_{it} = L_{i0}e^{n_i t}$. Output in each country i at time t is determined by the following Cobb-Douglas production function:

$$Y_{it} = K_{it}{}^{\alpha} G_{it}{}^{\beta} (A_{it} L_{it})^{1-\alpha-\beta}, 0 < \alpha < 1,\ 0 < \beta < 1,\ 0 < \alpha + \beta < 1 \tag{2}$$

Y is the final good, used for private consumption, G is public consumption expenditure, which proxies for government size, and K is investment in physical capital. We consider the case of no depreciation of physical capital. The output used to produce G equals qG (which one can think of as being equivalent to a crowding-out effect in private sector's resources). A is the level of technology and grows at the exogenous constant rate μ, that is, we have

$$A_{it} = A_{i0} e^{\mu_i t + I_{it} \rho_i} \tag{3}$$

with I_{it} being a vector of institutional quality, political regime, legal origin and other related factors that may affect the level of technology and efficiency in country i at time t, and ρ_i is a vector of (unknown) coefficients related to these variables. In this framework, the state of labour-augmenting technology (A) depends not only on exogenous technological improvements determined by μ, but also on the level of institutional quality (such as the rule of law), the degree of democratic political foundations, etc. Institutions may be critical in facilitating technological breakthroughs, which may not occur without appropriate sound institutional environments. The presence of efficient and effective institutions ensures that labour can be used for productive purposes, instead of being wasted with red tape or rent seeking activities (North, 1990; Nelson and Sampat, 2001).

We begin by writing down the resource constraint for this economy in per worker terms, given by:

$$\dot{K}_t = Y_t - C_t - qG_t \Leftrightarrow \dot{k}_t = y_t - c_t - qg_t - nk_i \tag{4}$$

where $\dot{K}_t$ is the time derivative of physical capital and small letters represent per worker terms (after scaling down by L).

---

[3] Peden and Bradley (1989) employ a theoretical model of output growth to derive an equation that controls for cyclical influencces and distinguishes the effects of government growth on the economic base from the effects on the economic growth rate. Lee (1992) and Devarajan *et al.* (1996) expand Barro's model, allowing different kinds of government expenditures to have different impacts on growth. At a more disaggregated level, distinguishing between productive and non-productive spending, Glomm and Ravikumar (1997) and Kneller *et al.* (1999) are able to determine the optimal composition of different kinds of expenditure, based on their relative elasticities. Similarly, Chen (2006) investigates the optimal composition of public spending and its relationship to economic growth.

We now write the conditions that characterize the optimal path for the economy and determine the steady-state solution for private and public consumption and income per worker. The optimal path is the solution of:

$$\begin{aligned} &\max_{c_t, g_t} \int_0^\infty e^{-\gamma t} \frac{c_t^{1-\theta} - 1}{1-\theta} dt \\ &s.t.: \dot{k}_t = k_t^{\alpha} g_t^{\beta} A_t^{1-\alpha-\beta} - c_t - q g_t - n k_i \end{aligned} \tag{5}$$

Solving the Hamiltonian's corresponding first order conditions and after some manipulations yields:[4]

$$\begin{aligned} k^* &= A \left(\frac{\alpha}{\theta\mu + \gamma + n} \right)^{\frac{1-\beta}{1-\alpha-\beta}} \left(\frac{\beta}{q} \right)^{\frac{1-\beta}{1-\alpha-\beta}} \\ g^* &= A^{\frac{1-\alpha-\beta}{1-\beta}} \left(\frac{\beta}{q} \right)^{\frac{1-\beta}{1-\alpha-\beta}} k^{*\frac{\alpha}{1-\beta}} \\ y^* &= k^{*\alpha} g^{*\beta} A^{1-\alpha-\beta} \\ c^* &= y^* - (n+\mu)k^* - qg^* \end{aligned} \tag{6}$$

A special case occurs when $\alpha + \beta = 1$ and $n = \mu = 0$ in which there is no transition dynamics and the economy is always in the balanced growth path.

We refrain from making full considerations on the model's solution, but one, in particular, is worth making:[5] an increase in *q* (which implicitly proxies the overall size of the public sector translating the fact that more resources are needed/required to finance *G*) reduces both the optimal level of private consumption per worker (and physical capital per worker) and, more importantly, the optimal level of output per worker in this model economy.

Turning to econometric specification, in the steady state, output per effective worker ($\hat{y}_{it} = Y_{it} / A_{it} L_{it}$) is constant while output per worker ($y_{it} = Y_{it} / L_{it}$) grows at the exogenous rate μ. In general, output in effective worker terms evolves as $\hat{y}_{it} = (k_{it})^{\alpha} (g_{it})^{\beta}$ and in (raw) worker terms, output evolves according to $y_{it} = A_{it} (k_{it})^{\alpha} (g_{it})^{\beta}$. Taking logs on both sides we get $\ln y_{it} = \ln A_{it} + \alpha \ln k_{it} + \beta \ln g_{it}$, and using (3) and the fact that in (2) we have $(A_{it} L_{it})^{1-\alpha-\beta}$ entering the utility function, we obtain,

$$\ln y_{it} = A_0 + (1-\alpha-\beta)\mu_i t + (1-\alpha-\beta)\rho_i I_{it} + \alpha \ln k_{it} + \beta \ln g_{it} . \tag{7}$$

Equation (7) describes the evolution of output per worker (or labour productivity), as a function of a vector of institutional and political related variables, which may change over time, the size of the public sector or government, the level of physical capital and the exogenous growth rate of output. Given the production function relationship, (7) is valid both within and outside the steady-state and this is important, particularly, if one makes use of static panel data techniques for estimation purposes. Moreover, it is not dependent on assumptions on the behaviour of savings, hence offering a reasonable basis for estimation. Based on (7), we will use both a linear and non-linear specification (in which interaction or multiplicative terms are included), as follows:

---

4 The derivation is available upon request.

5 In an alternative setting in which the government introduces a tax over total income (or production) to finance public consumption, the overall conclusion (with respect to the effect of government size) does not change.

$$\ln y_{it} = b_0 + b_1 t + b_3 I_{it} + b_4 \ln k_{it} + b_5 \ln g_{it} + \varepsilon_{it} \quad (8)$$

$$\ln y_{it} = b_{0i} + b_1 t + b_3 I_{it} + b_4 \ln k_{it} + b_5 \ln g_{it} + b_6 (I_{it} g_{it}) + \eta_{it} \quad (9)$$

where the b's are (unknown) parameters to be estimated, I_{it} and g_{it} denote the proxies for institutional quality and government size, respectively, and ε_{it} and η_{it} are model specific error terms satisfying the usual assumptions of zero mean and constant variance. Equations (8) and (9) provide the basis for the empirical models to be estimated in Section 3.

Finally, the variation of causality between government size and growth detected in cross-section and time-series papers suggests that there are important differences in the way in which governments influence economic performance across countries. We argue that it may reflect, lato sensu, institutional differences across countries and, while this is a plausible conjecture, there is as yet little direct evidence to confirm that institutions and political regimes make a difference to the way in which governments affect economic outcomes.

3 Data

The dataset consists of an unbalanced panel of observations for 108 heterogeneous countries for the period 1970-2008 in 5-year averages (to overcome short-run business cycle fluctuations as is common practice in the growth literature).[6] Countries are grouped into developed (OECD) and emerging and developing based on the World Bank classification. Annual data on real GDP per capita (y) and gross fixed capital formation (inv) are retrieved from the World Bank' World Development Indicators. We estimate the capital stock (Ky) using the perpetual inventory method, that is, $Ky_t = Inv_t + (1-\delta)Ky_{t-1}$, where Inv_t is the investment and δ is the depreciation rate. Data on Inv_t comes from Summers and Heston's PWT 6.3 as real aggregate investment in PPP. We estimate the initial value of the capital stock (Ky_0), in year 1950 as $Inv_{1950}/(g+\delta)$ where g is the average compound growth rate between 1950 and 1960, and δ is the depreciation rate (set to 7 per cent for all countries and years).

Our proxies of government size (g) will be the respective Gwartney and Lawson's (2008) composite variable (*govsize*). This variable includes government consumption expenditures (as a percentage of total consumption), transfers and subsidies (as a percentage of GDP), the underlying tax system (proxied by top marginal tax rates) and the number of government enterprises. We also make use of total government expenditures (*totgovexp_gdp*), government consumption (*govcons_gdp*) – as in our theoretical model - and, finally, total government debt (*govdebt_gdp*). The first two variables come from a merger between WDI, the IMF's International Financial Statistics (IFS) and Easterly's (2001) datasets.[7] The latter was retrieved from the recent IMF's historical debt series due to Abas *et al.* (2010).

For institutional-related variables (our I) we rely on:[8] i) the Polity 2 (*polity*) measure and regime durability in years (*durable)* (from Marshall and Jaegger's Polity's 4 database), ii) Freedom House's Political Rights (*pr*), Civil Liberties (*cl*) and composite index (*fh*),[9] iii) the corruption

6 Summary statistics and correlation matrices are omitted for economy of space but they are available upon request.

7 The classification of the data is described in IMF (2001).

8 The interested reader should refer to the original sources for the full definition of the variables used.

9 Constructed by simply averaging Political Rights and Civil Liberties.

perception index (*cpi*) (from the Transparency International database).[10] iv) an index of democratization (*demo*) due to Vanhanen (2005), v) a governance index (*governance*)[11] from Kaufman *et al.* (2009) (World Bank project), vi) the political system (*ps*), a dummy variable that takes a value zero for presidential regime, the value one for the assembly-elected presidential regime and two for parliamentary regime (from the Database of Political Institutions), and vii) countries' legal origins, English (*bri*), French (*fre*), German (*ger*) or Scandinavian (*sca*)[12] (from La Porta *et al.*, 1999).[13]

For robustness purposes we will also make use of factor analysis and combine different sets of institutional-related variables (in particular, *pr*, *cl*, *polity*, *demo* and *cpi*) and then look at the first common factor. However, the sampling technique is unfortunately restricted to the fact that cross-country data are limited in the country coverage and vary widely across different data sources. This limitation creates an incomplete data issue and poses a problem for the Principal Component Analysis (PCA) that we wish to employ. Indeed, PCA is based on an initial reduction of the data to the sample mean vector and sample covariance matrix of the variables, and this cannot be estimated from datasets with a large proportion of missing values (Little and Rubin, 1987).[14] Hence, imputation is required prior to extracting the first principal component.[15] The Expectation-Maximization Algorithm (EMA) as suggested by Dempster *et al.* (1977) is used to fill in missing data. This algorithm is based on iterating the process of regression imputation and maximum likelihood and it consists of two steps: the first step, the "E (expectation)-step" computes expected values (conditional on the observed data) and the current estimates of the parameters. Using the estimated "complete data", in the second step or "M-step", the EMA re-estimates the means, variances and covariances using a formula that compensates for the lack of residual variation in the imputed values.[16]

The first principal component is normalized in such a way that high values indicate higher institutional quality. Our standardized index, *EMA_PCA*, can be written as:[17]

$$EMA\_CA = 0.78cl + 0.89pr + 0.92polity + 0.69demo + 0.34cpi$$

In addition, the first principal component explains 73.6 per cent of the total variance in the standardized data.[18] This aggregate index will be used in the empirical analysis below.

---

10 See Goel and Nelson (1998) for a disaggregated analysis on the effect of government size on corruption.

11 This is the result of averaging six variables: voice and accoutability, political stability, government effectiveness, regulatory quality, rule of law and control of corruption.

12 There is no risk of multicollinearity since "socialist" legal origin is not included explicitly on the right-hand-side as an explanatory variable.

13 Data sources and definitions are provided in the Appendix.

14 Moreover, the lack of data also increases the degree of uncertainty and influences the ability to draw accurate conclusions.

15 The varimax rotation method is chosen.

16 The EMA assumes that the data are missing at random (MAR) and in order to check that the MAR assumption can be applied to the measures of institutional quality, a test analysis called "separate variance *t*-test", in which rows are all variables which have 1 per cent missing or more, and columns are all variables, is carried out. The *p*-values are more than 5 per cent meaning that missing cases in the row variable are not significantly correlated with the column variable and this, can be considered as MAR.

17 A likelihood ratio test was used to examine the "sphericity" case, allowing for sampling variability in the correlations. This test comfortably rejects sphericity at the 1 per cent level with a Kaiser-Meyer-Olkin measure of sampling adequacy equal to 0.831.

18 Given that the PCA is based on the classical covariance matrix, which is sensitive to outliers, we take one further step by basing it on a robust estimation of the covariance (correlation) matrix. A well suited method is the Minimum Covariance Determinant (MCD) – we implement Rousseeuw and Van Driessen's (1999) algorithm. After re-computing the same measure with the MCD version we obtain similar results, meaning that outliers are not driving our factor analysis (the correlation coefficient between the two equals 98,04 per cent, statistically significant at 1 per cent level).

4 Methodology and results

4.1 Baseline results

Equations (8) and (9) can be estimated directly using panel data techniques, which allow for both cross-section and time-series variation in all variables and present a number of advantages *vis-à-vis* standard Barro-type pooled cross-section estimation approaches (see Greene, 2003).

Table 1.a and 1.b present our first set of results for the pooled OLS and fixed-effects specifications, respectively (the former is presented for completeness). Both tables are divided into two panels (A and B) covering different proxies for institutional quality (eight in total). At this point, we use Gwartney and Lawson's government size measure only and discuss its individual inclusion in our regression of interest as well as its interaction with a variable I_{it}.

A few remarks are worth mentioning. There is a positive effect of the capital stock on the level of real GDP per capita throughout the different specifications regardless of the institutional variable employed. One also finds a consistent and statistically significant negative coefficient on the government size (less so when fixed-effects are used, see Table 1.b). Its coefficient varies between 0.03 and 0.11 across the two tables, meaning that an increase in government size by 10 percentage points is associated with a 0.3 to 1.1 per cent lower annual growth. This order of magnitude is consistent with previous studies. Similarly, institutional quality has a consistent and statistically significant positive impact on the level of real GDP per capita (more mitigated with fixed-effects). Finally, when statistically significant the interaction term is negative, meaning that i) the negative effect of government size on GDP per capita is stronger at lower levels of institutional quality, and ii) the positive effect of institutional quality on GDP per capita is stronger at smaller levels of government size. The interaction term means that the marginal effect of government size will differ at different levels of institutional quality. However, this result depends on the proxy used for I_{it}. Nevertheless, we obtain in most regressions considerably high R-squares. Moreover, when regional dummies are included, coefficients keep their statistical significance and sign.

If we redo the exercise with the EMA_PCA variable instead, for both pooled OLS and fixed-effects estimators, Table 2 shows meaningful results for the size of the government and for the institutional quality index, when OLS is considered.

4.2 Endogeneity and dynamic panel estimation

In the analysis of empirical production functions, the issue of variable endogeneity is generally of concern. Moreover, instead of estimating static equations, we now allow for dynamics to play a role. A negative correlation between government size and economic growth does not imply causality. In fact, the most obvious reason (among many) to suspect reverse causality a problem is that welfare states social insurance schemes act as automatic stabilizers. Hence, we reformulate our regression equation(s) and take real GDP growth per capita as our dependent variable being a function of lagged real GDP per capita, investment (gross fixed capital formation as percentage of GDP), a government-size proxy and an interaction term (with an institutional quality proxy) – as common practice in the empirical growth literature. We estimate this new specification by means of the Arellano-Bover system-GMM estimator[19] which jointly estimates the

[19] The GMM approach estimates parameters directly from moment conditions imposed by the model. To enable identification the number of moment conditions should be at least as large as the number of unknown parameters. Moreover, the mechanics of the GMM approach relates to a standard instrumental variable estimator and also to issues such as instrumental validity and informativeness.

Table 1.a

Results of OLS Estimation, with Interaction Terms

Sample	**Full**											
Estimator	**Pooled OLS**											
Spec.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Institutional Proxy	cl			pr			polity			demo		
ln k	0.942^{***}	0.908^{***}	0.941^{***}	1.032^{***}	0.999^{***}	1.031^{***}	1.086^{***}	1.025^{***}	1.080^{***}	0.954^{***}	0.905^{***}	0.958^{***}
	(0.043)	*(0.042)*	*(0.044)*	*(0.044)*	*(0.043)*	*(0.045)*	*(0.038)*	*(0.039)*	*(0.040)*	*(0.041)*	*(0.039)*	*(0.041)*
g	-0.064^{***}	-0.039^{**}	–0.037	-0.076^{***}	-0.040^{**}	–0.070	-0.061^{***}	–0.027	–0.036	-0.028^{**}	–0.004	-0.067^{**}
	(0.013)	*(0.016)*	*(0.050)*	*(0.016)*	*(0.017)*	*(0.058)*	*(0.017)*	*(0.017)*	*(0.026)*	*(0.014)*	*(0.015)*	*(0.031)*
I	0.220^{***}	0.201^{***}	0.255^{***}	0.112^{***}	0.107^{***}	0.120^{*}	0.021^{***}	0.024^{***}	0.043^{**}	0.025^{***}	0.024^{***}	0.016^{**}
	(0.026)	*(0.023)*	*(0.064)*	*(0.021)*	*(0.018)*	*(0.072)*	*(0.005)*	*(0.005)*	*(0.020)*	*(0.003)*	*(0.002)*	*(0.007)*
$I^{*}g$			–0.006			–0.001			–0.004			-0.002^{*}
			(0.010)			*(0.011)*			*(0.003)*			*(0.001)*
Latin America	-0.240^{***}			-0.297^{***}			-0.337^{***}			-0.275^{***}		
	(0.070)			*(0.072)*			*(0.071)*			*(0.064)*		
Asia	-0.773^{***}			-0.783^{***}			-0.842^{***}			-0.848^{***}		
	(0.092)			*(0.100)*			*(0.098)*			*(0.085)*		
Africa	–0.015			0.099			0.032			–0.011		
	(0.110)			*(0.119)*			*(0.112)*			*(0.099)*		
N	437	437	437	437	437	437	448	448	448	476	476	476
R^2	0.923	0.934	0.923	0.909	0.924	0.909	0.897	0.915	0.897	0.917	0.931	0.918

Sample	**Full**											
Estimator	**Pooled OLS**											
Spec.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Institutional Proxy		cpi			governance			ps			pc	
ln k	0.813*** *(0.048)*	0.828*** *(0.042)*	0.805*** *(0.047)*	0.763*** *(0.058)*	0.771*** *(0.055)*	0.758*** *(0.056)*	1.182*** *(0.045)*	1.150*** *(0.049)*	1.183*** *(0.045)*	1.249*** *(0.039)*	1.205*** *(0.047)*	1.252*** *(0.039)*
g	–0.007 *(0.015)*	–0.003 *(0.015)*	–0.109** *(0.053)*	–0.039** *(0.018)*	–0.037* *(0.020)*	–0.080*** *(0.027)*	–0.041* *(0.023)*	–0.009 *(0.023)*	–0.034* *(0.021)*	–0.039 *(0.025)*	–0.017 *(0.026)*	0.034 *(0.064)*
I	0.200*** *(0.017)*	0.201*** *(0.016)*	0.103** *(0.042)*	0.563*** *(0.061)*	0.574*** *(0.051)*	0.240* *(0.126)*	0.001 *(0.036)*	0.053* *(0.032)*	0.085 *(0.178)*	0.182* *(0.109)*	0.047 *(0.104)*	0.674 *(0.425)*
I*g			–0.017** *(0.007)*			–0.054*** *(0.021)*			–0.014 *(0.031)*			–0.084 *(0.072)*
Latin America		0.088 *(0.067)*			0.120 *(0.092)*			–0.317*** *(0.097)*			–0.254*** *(0.096)*	
Asia		–0.579*** *(0.077)*			–0.528*** *(0.111)*			–0.755*** *(0.148)*			–0.547*** *(0.150)*	
Africa		0.289*** *(0.105)*			0.219 *(0.151)*			0.126 *(0.167)*			0.062 *(0.152)*	
N	240	240	240	176	176	176	258	258	258	225	225	225
R^2	0.954	0.964	0.955	0.950	0.958	0.951	0.919	0.932	0.919	0.935	0.942	0.936

Note: The models are estimated by Pooled OLS. The dependent variable is the logarithm of real GDP per capita. A time trend has been included but is not reported for reasons of parsimony. Robust heteroskedastic-consistent standard errors are reported in parenthesis below each coefficient estimate. A constant term has been estimated but it is not reported for reasons of parsimony. *, **, *** denote significance at 10, 5 and 1 per cent levels.

Table 1b

Results of FE Estimation, with Interaction Terms

Estimator	FE							
Spec.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Institutional Proxy	cl		pr		polity		demo	
ln k	0.691***	0.692***	0.687***	0.688***	0.575***	0.574***	0.609***	0.605***
	(0.078)	*(0.079)*	*(0.077)*	*(0.078)*	*(0.079)*	*(0.080)*	*(0.079)*	*(0.080)*
g	–0.006	–0.005	–0.005	–0.010	–0.029**	–0.038***	–0.018	–0.042**
	(0.016)	*(0.024)*	*(0.016)*	*(0.019)*	*(0.011)*	*(0.012)*	*(0.014)*	*(0.017)*
I	0.009	0.011	0.013	0.022	0.009***	0.004	0.002	0.005*
	(0.013)	*(0.036)*	*(0.010)*	*(0.028)*	*(0.003)*	*(0.007)*	*(0.002)*	*(0.003)*
I*g		0.003		0.006		–0.002*		–0.001**
		(0.006)		*(0.005)*		*(0.001)*		*(0.001)*
N	437	437	437	437	448	448	476	476
R2	0.823	0.824	0.825	0.826	0.836	0.839	0.821	0.826

Estimator	FE							
Spec.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Institutional Proxy	cpi		governance		ps		pc	
ln k	0.611***	0.611***	0.215	0.245*	0.586***	0.582***	0.588***	0.590***
	(0.152)	*(0.151)*	*(0.152)*	*(0.130)*	*(0.141)*	*(0.141)*	*(0.157)*	*(0.154)*
g	–0.002	–0.006	–0.015*	–0.021**	0.033	–0.058***	0.034	0.026
	(0.007)	*(0.019)*	*(0.008)*	*(0.009)*	*(0.024)*	*(0.020)*	*(0.029)*	*(0.059)*
I	0.004	0.012	0.128**	0.247**	–0.032	0.256*	–0.041	–0.094
	(0.013)	*(0.019)*	*(0.061)*	*(0.112)*	*(0.041)*	*(0.136)*	*(0.040)*	*(0.293)*
I*g		0.001		0.018		–0.043**		0.009
		(0.003)		*(0.013)*		*(0.020)*		*(0.054)*
N	240	240	176	176	258	258	225	225
R^2	0.722	0.723	0.468	0.488	0.767	0.785	0.748	0.748

Note: The models are estimated by Fixed-Effects. The dependent variable is the logarithm of real GDP per capita. A time trend has been included but is not reported for reasons of parsimony. Robust heteroskedastic-consistent standard errors are reported in parenthesis below each coefficient estimate. A constant term has been estimated but it is not reported for reasons of parsimony. *, **, *** denote significance at 10, 5 and 1 per cent levels.

Table 2

Results of OLS and FE Estimation, with Interaction Terms. PCA-based Institutional Measure

Estimator	OLS		FE	
Spec.	**1**	**2**	**3**	**4**
ln k	0.976***	0.970***	0.675***	0.676***
	(0.048)	*(0.050)*	*(0.079)*	*(0.079)*
g	–0.066***	–0.046*	–0.018	–0.019
	(0.015)	*(0.024)*	*(0.014)*	*(0.016)*
I	0.423***	0.307***	–0.016	–0.029
	(0.064)	*(0.113)*	*(0.035)*	*(0.057)*
I*g		0.029		0.003
		(0.026)		*(0.012)*
N	411	411	411	411
R^2	0.913	0.913	0.821	0.821

Note: The models are estimated by Fixed-Effects. The dependent variable is the logarithm of real GDP per capita. A time trend has been included but is not reported for reasons of parsimony. Robust heteroskedastic-consistent standard errors are reported in parenthesis below each coefficient estimate. A constant term has been estimated but it is not reported for reasons of parsimony. *, **, *** denote significance at 10, 5 and 1 per cent levels.

equations in first differences, using as instruments lagged levels of the dependent and independent variables, and in levels, using as instruments the first differences of the regressors.[20] Intuitively, the system-GMM estimator does not rely exclusively on the first-differenced equations, but exploits also information contained in the original equations in levels.

Another contribution of our study is the construction of new (and more meaningful) democracy measures based on the variable *polity* (described in the Appendix). The role of political systems and democracy in particular, on the government size-growth relationship is assessed by regressing three structural aspects of democracy (to be defined below) on 5-year averages of real GDP per capita growth rates.[21] Indeed, *polity* does not capture two important dimensions of political regimes – either their newness (following, for example, democratization or a return to authoritarian rule) or their more established (consolidated) nature.

Therefore, Rodrik and Wacziarg (2005) define a major political regime change to have occurred when there is a shift of at least three points in a country's score on *polity* over three years or less. Using this criterion we define new democracies (ND=1) in the initial year (and subsequent four years) in which a country's *polity* score is positive and increases by at least three points and is sustained, ND=0 otherwise. Established democracies (ED=1) are those new democratic regimes that have been sustained following the 5 years of a new democracy (ND). In any subsequent year, if established democracies (ED) fail to sustain the status of ND, ED=0. Using these criteria, they define sustained democratic transitions (SDT) as the sum of ND and ED. They use the same procedure, mutatis mutandis, to define new autocracies (NA), established autocracies (ES) and sustained autocratic transition (SAT).

This yields six distinct binary-type measures of the character of political regimes – ND, ED, NA, EA, SDT, and SAT – for most years during 1970-2008. Finally, Rodrik and Wacziarg (2005) define small regime changes (SM) as changes in *polity* from one year to the next that are less than three points.[22] A recent empirical application of these measures to explain the impact of extreme-type political regimes on economic performance can be found in Jalles (2010). There are several advantages from creating these new measures, which allow us to distinguish the impact of new and established electoral democracies and autocracies on economic development, and also to assess the impact of sustained democratic and autocratic transitions on economic growth.

Endogeneity[23] between right-hand side measures of democracy and autocracy and a standard set of control variables is corrected for by taking a system-GMM (SYS-GMM) approach – as detailed above. As suggested in Mauro (1995), La Porta *et al.* (1997), Hall and Jones (1999), Acemoglu *et al.* (2001) and Dollar and Kraay (2003), the democracy measures are instrumented by:

1 the durability (age in years) of the political regime type (*durable*) retrieved from Marshall and Jaeggers' database;[24]

---

20 As far as information on the choice of lagged levels (differences) used as instruments in the differences (levels) equation, as work by Bowsher (2002) and, more recently Roddman (2009) has indicated, when it comes to moment conditions (as thus to instruments) more is not always better. The GMM estimators are likely to suffer from "overfitting bias" once the number of instruments approaches (or exceeds) the number of groups/countries (as a simple rule of thumb). In the present case, the choice of lags was directed by checking the validity of different sets of instruments.

21 An equation with real GDP per capita growth as the dependent variable is motivated by (standard) augmentation of Solow-Swan type models with a government size proxy (similarly to our production function in Section 2) and following Barro and Sala-i-Martín's (1992) and Mankiw *et al.*'s (1992) approaches.

22 Thus SM = 1 for a small regime change and SM = 0 otherwise.

23 And also the existence of possible measurement errors when accounting for democracy.

24 The average age of the party system is also used in Przeworski *et al.* (2000) and Beck *et al.* (2001). This potential instrument is also in line with Bockstette, Chanda and Putterman (2002) who document the use of the state antiquity index as an appropriate instrument for institutional quality.

2 *latitude* (from La Porta *et al.*, 1998, 1999): Hall and Jones (1999) launched the general idea that societies are more likely to pursue growth-promoting policies, the more strongly they have been exposed to Western European influence, for historical or geographical reasons. In this context, other two possible instruments could be common and civil law, translating the type of legal origin of each country;

3 ethnic fragmentation (*ethnic*) (from Alesina *et al.*, 2003): on a broad level, the role of ethnic fragmentation in explaining the (possible) growth effect of democracy can be derived from the literature on the economic consequences of ethnic conflict. It has been shown that the level of trust is low in an ethnically divided society (Alesina and La Ferrara, 2000). Moreover, the lack of co-operative behaviour between diverse ethnic groups, leads to the tragedy of the commons as each group fights to divert common resources to non-productive activities (e.g., Mauro, 1995).[25]

Table 3 reports the results with the four proxies for government size defined in Section 3 and splitting the sample into OECD, emerging and developing countries groups.[26] Focusing on the full sample first we observe that the Gwartney and Lawson's government size measure appears with a statistically significant negative coefficient. When interacted with SAT it has a negative and statistically significant coefficient, meaning that in autocratic countries increased government size has greater negative effect on output growth. The reverse is true for democratic countries, whose negative impact of government size is mitigated but remains mostly negative. The remaining proxies keep the statistically negative coefficient, but interaction terms lose economic and statistical relevance. For the OECD sub-group the individual effects of the different proxies of government size are similar but interaction terms are never statistically significant. Developing countries report a statistically negative coefficient on government consumption expenditure and debt-to-GDP ratio, with the latter having a lesser detrimental effect in democratic countries. All in all, government consumption is the proxy that is more consistently and clearly detrimental to output growth.

More stringent empirical tests on the role of democracy on the government size-growth relation were carried out, for robustness purposes (similarly to Rock, 2009). We defined "extreme" democratic transitions as those where the *polity* variable is greater than 5. In these instances, a new sustainable democratic transitions variable, SDT1=1 when *polity*>5, otherwise SDT1=0. Similarly, a new sustainable autocratic transitions variable was created, SAT1=1 when *polity*<–5, otherwise SAT1=0. The logic behind this construction is to test for the impact of democracy and autocracy on growth in cases where countries' governments are closer to either pure democracies or pure autocracies.[27] Results (not shown) using the new SAT1 and SDT1 variables do not qualitatively change the results presented in Table 3 and discussed above.

We also assessed the importance of political-institutional measures, specifically legal origins. From Table 4 a first general conclusion is that interaction terms with a Scandinavian legal origin dummy yields the higher (in absolute value) estimated coefficients (when significant), compared with other legal origins. More particularly, in specification 4 and 5, for the full sample and OECD respectively, the government debt-to-GDP ratio and government size appear with a

---

25 Other similarly possible instruments are the historical settler mortality or population density in 1500, as in Acemoglu and Robinson (2005), the constitutional initiative which allows citizens to amend or demand a revision of the current constitution (as in Poterba, 1996), the share of population that speaks any major European language – *Eurfrac* –, inter alia. For the three instruments chosen the exclusion restriction is that durability, latitude and ethnic fragmentation do not have any impact on present economic growth other than their impact on democracy.

26 In the great majority of our system-GMM regressions the Hansen-*J*-statistic is associated with p-values larger than 10 per cent. This statistic tests the null hypothesis of correct model specification and valid overindentifying restrictions, *i.e.*, validity of instruments.

27 The cut-off point for defining these measures of democracy/autocracy was taken directly from Marshall and Jaeggers (http://www.systemicpeace.org/polity/polity4.htm).

Table 3

Results of Estimations Controlling for Endogeneity (with Interaction Terms of New Political Systems' Measures)

Sample	**All**				**OECD**				**Emerging**				**Developing**			
Estimation	**SYS-GMM**															
Spec.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
gfcf_gdp	–0.25 (0.192)	0.11* (0.661)	0.13** (0.058)	0.14** (0.036)	0.67* (0.363)	–0.07 (0.188)	–0.06 (0.155)	0.07 (0.222)	0.66** (0.262)	0.02 (0.306)	0.29** (0.137)	0.28* (0.155)	–0.15 (0.203)	0.24*** (0.088)	0.13* (0.075)	0.12* (0.073)
Government size proxy	govsize	Totgovexpp	Govcons	Govdeb	govsize	Totgovexpp	Govcons	Govdebt	govsize	Totgovexpp	Govcons	Govdebt	govsize	Totgovexpp	Govcons	Govdebt
g	–2.37** *(1.088)*	–0.20*** *(0.049)*	–0.37*** *(0.122)*	–0.02*** *(0.005)*	–1.88** *(0.871)*	–0.20 *(0.158)*	–0.79*** *(0.273)*	0.02 *(0.062)*	–1.51 *(1.525)*	–0.14 *(0.139)*	0.16 *(0.340)*	–0.02 *(0.034)*	–1.64 *(1.937)*	–0.14 *(0.087)*	–0.33** *(0.154)*	–0.02*** *(0.004)*
g*SAT	–0.70* *(0.393)*	0.03 *(0.027)*	–0.05 *(0.056)*	–0.01 *(0.005)*	0.18 *(0.206)*	0.08 *(0.138)*	0.23 *(0.380)*	0.04 *(0.056)*	–17.61* *(10.570)*	0.03 *(0.182)*	–0.49** *(0.211)*	0.03 *(0.025)*	–0.14 *(1.677)*	–0.11* *(0.060)*	0.06 *(0.101)*	0.01 *(0.010)*
g*SDT	0.78** *(0.354)*	0.04 *(0.045)*	–0.01 *(0.057)*	0.02*** *(0.003)*	–0.05 *(0.141)*	–0.04 *(0.124)*	0.02 *(0.273)*	0.01 *(0.054)*		–0.12 *(0.166)*	–0.03 *(0.148)*	–0.01 *(0.028)*	–0.29 *(2.086)*	0.16** *(0.069)*	0.05 *(0.115)*	0.01*** *(0.004)*
Observations	383	1757	3653	3200	116	716	938	849	117	454	868	779	170	642	1,964	1,677
Hansen (*p*-value)	0.04	1.00	1.00	1.00	0.89	1.00	1.00	1.00	0.95	1.00	1.00	1.00	0.38	1.00	1.00	1.00
AB AR(1) (*p*-value)	0.02	0.00	0.00	0.00	0.15	0.01	0.00	0.01	0.05	0.01	0.00	0.00	0.08	0.00	0.00	0.00
AB AR(2) (*p*-value)	0.29	0.00	0.01	0.04	0.36	0.00	0.01	0.06	0.14	0.04	0.19	0.32	0.39	0.11	0.03	0.13

Note: The models are estimated by system GMM (SYS-GMM). The dependent variable is real GDP per capita growth. "*SDT*" and "*SAT*" stand for sustained democratic transition and sustained autocratic transition – for more details refer to the main text. Robust heteroskedastic-consistent standard errors are reported in parenthesis below each coefficient estimate. The Hansen test evaluates the validity of the instrument set, *i.e.*, tests for over-identifying restrictions. AR(1) and AR(2) are the Arellano-Bond autocorrelation tests of first and second order (the null is no autocorrelation), respectively. Also a constant term, lagged dependent variable and a time trend have been included but are not reported for reasons of parsimony. *, **, *** denote significance at 10, 5 and 1 per cent levels.

Table 4

Results of Estimations Controlling for Endogeneity (with Interaction Terms of Legal Origins' Type)

Sample	**All**				**OECD**				**Emerging**				**Developing**			
Estimation	**SYS-GMM**															
Spec.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
gfcf_gdp	–0.19	0.12*	0.16***	0.14***	1.13***	–0.09	–0.12	0.30	0.67***	–0.06	0.14	0.22**	–0.13	0.28***	0.09	0.11
	(0.287)	*(0.065)*	*(0.052)*	*(0.054)*	*(0.345)*	*(0.146)*	*(0.140)*	*(0.187)*	*(0.255)*	*(0.400)*	*(0.145)*	*(0.110)*	*(0.291)*	*(0.083)*	*(0.068)*	*(0.066)*
Government size proxy	govsize	Totgovexpp	Govcons	Govdebt	govsize	Totgovexpp	Govcons	Govdebt	govsize	Totgovexpp	Govcons	Govdebt	govsize	Totgovexpp	Govcons	Govdebt
g	–0.11	–0.14	–1.02***	–0.12*	–7.06*	–0.27	–0.80	–0.19	–0.05	–0.31	0.58	–0.02	15.74	–1.30**	–1.11**	–0.51*
	(0.287)	*(0.299)*	*(0.327)*	*(0.061)*	*(3.946)*	*(0.775)*	*(0.926)*	*(0.154)*	*(2.929)*	*(0.396)*	*(0.395)*	*(0.020)*	*(14.481)*	*(0.602)*	*(0.465)*	*(0.282)*
g*british	–4.77	–0.04	0.61*	0.10*	5.58	–0.22	–0.54	0.33	–3.28	0.42	–1.48***	0.11	–19.14	1.28**	0.80	0.48*
	(4.481)	*(0.319)*	*(0.371)*	*(0.062)*	*(4.154)*	*(0.992)*	*(0.936)*	*(0.410)*	*(4.053)*	*(0.792)*	*(0.560)*	*(0.157)*	*(14.805)*	*(0.648)*	*(0.543)*	*(0.279)*
g*french	–1.71	0.01	0.72**	0.11*	5.50	0.24	0.21	0.20	2.70	0.15	–0.72*	–0.04	–20.12	1.25**	0.66	0.51*
	(3.190)	*(0.326)*	*(0.362)*	*(0.061)*	*(4.069)*	*(0.910)*	*(1.688)*	*(0.142)*	*(4.094)*	*(0.540)*	*(0.410)*	*(0.039)*	*(16.637)*	*(0.573)*	*(0.505)*	*(0.281)*
g*german	1.17	0.36	0.99	0.17*	3.88	–0.35	–0.83	0.33	–	–	–	–	–	–	–	–
	(2.167)	*(0.426)*	*(0.836)*	*(0.101)*	*(4.741)*	*(0.746)*	*(1.701)*	*(0.217)*								
g*scandinavian	–0.87	–0.13	0.785	0.21**	7.01	0.24	0.29	0.39*	–	–	–	–	–	–	–	–
	(2.782)	*(0.537)*	*(0.682)*	*(0.087)*	*(5.294)*	*(1.219)*	*(1.220)*	*(0.216)*								
Observations	393	1886	4010	3483	116	794	1,006	910	111	462	894	798	178	677	2,201	1,858
Hansen (*p*-value)	0.34	1.00	1.00	1.00	0.90	1.00	1.00	1.00	0.93	1.00	1.00	1.00	0.37	1.00	1.00	1.00
AB AR(1) (*p*-value)	0.02	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.11	0.00	0.00	0.00
AB AR(2) (*p*-value)	0.15	0.00	0.00	0.01	0.76	0.00	0.02	0.04	0.31	0.02	0.29	0.30	0.15	0.03	0.00	0.05

Note: See note in Table 3 for details. "British", "French", "German" and "Scandinavian" denote British, French, German and Scandinavian legal origins, respectively.

(statistically) negative coefficient; however, this effect on growth is mitigated particularly if a country has a Scandinavian legal origin.[28] For developing countries, both French and British legal origins appear with statistically significant positive interaction term coefficients when the government size proxy is total government expenditures.

As suggested by Ram (1986) another possible specification is the use of the growth rate of the government size proxy. We also test this specification to determine its impact on growth across political systems or levels of institutional quality. All variables are retained except G_{it} that is now replaced by dG_{it} / G_{it} together with the corresponding interaction terms.[29] Comparing with our previous results the coefficients of the linear term of government size proxies are positive and statistically significant in two out of five specifications. According to Conte and Darrat (1988) Ram's specification is suitable for testing short-term growth effects, while the specification used in this paper assesses the effects of government size on the underlying growth rate. Growth and development are long-run concepts whereas management of aggregate demand, a Keynesian prescription, is basically a short-term concept. Hence, while short-term measures of government may have a positive impact on an economy, the impact of government on the underlying growth rate generally differs between political regimes and legal origins as found in this paper (a comparable robustness analysis is available upon request).

Further in our inspection, similar regressions, where the I_{it} variable is now replaced with the composite Freedom House index, were estimated.[30] Two main results are worth mentioning: i) government size keeps its statistically significant negative sign, but its interaction with the Freedom House index yields a statistically negative coefficient (for the full sample), suggesting that the negative effect of government size on GDP per capita growth is stronger at lower levels of civil liberties and political rights; and ii) for the OECD sub-group debt has a statistically significant negative coefficient estimate and its interaction with the Freedom House index results in a negative estimate significant at 5 per cent level.

4.3 Robustness checks

One concern when working with time-series data is the possibility of spurious correlation between the variables of interest (Granger and Newbold, 1974). This situation arises when series are not stationary, that is, they contain stochastic trends as it is largely the case with GDP and investment series. The advantage of panel data integration is twofold: firstly, the tests are more powerful than the conventional ones: secondly, cross-section information reduces the probability of a spurious regression (Barnerjee, 1999). Results of first (Im, Pesaran and Shin, 1997; Maddala and Wu, 1999) and second generation (Pesaran, 2007) panel integration tests (not shown) suggest that we can accept most conservatively that non-stationarity cannot be ruled out in our dataset.

In face of this finding, it seems that the time-series properties of the data play an important role: we suggest that the bias in our models is the result of non-stationary errors, which are introduced into the fixed-effects and GMM equations by the imposition of parameter homogeneity. Hence, careful modelling of short-run dynamics requires a slightly different econometric approach. We assume that (8), or (9), represents the equilibrium which holds in the long-run, but that the

---

[28] Bergh and Henrekson (2011) propose two explanations for why countries (such as Scandinavian ones) with high taxes (hence, larger government size) are able to enjoy above average growth (which supports the absence of conclusive or statistically significant coefficients). One is that these countries have higher social trust; another is that their larger governments compensate for high taxes and spending by implementing market-friendly policies in other areas.

[29] The full table is available upon request.

[30] Ibidem.

dependent variable may deviate from its path in the short-run (due, e.g., to shocks that may be persistent). There are often good reasons to expect the long-run equilibrium relationships between variables to be similar across groups of countries, due e.g. to budget constraints or common technologies (unobserved TFP) influencing them in a similar way. In fact, in line with discussions in the empirical growth literature for modelling the "measure of our ignorance" we shall assume that the long-run relationship is composed of a country-specific level and a set of common factors with country-specific factor loadings.

The parameters of (8) and (9) can be obtained via recent panel data methods. Indeed, at the other extreme of panel procedures, based on the mean of the estimates (but not taking into account that certain parameters may be the same across groups), we have the Mean Group (MG)[31] estimator (Pesaran and Smith, 1995) and as an intermediate approach the Pooled Mean Group (PMG)[32] estimator, which involves both pooling and averaging (Pesaran *et al.*, 1999). These estimators are appropriate for the analysis of dynamic panels with both large time and cross-section dimensions, and they have the advantage of accommodating both the long-run equilibrium and the possibly heterogeneous dynamic adjustment process.

Therefore, a second step in our empirical approach is to make use of the Common Correlated Effects Pooled (CCEP) estimator that accounts for the presence of unobserved common factors by including cross-section averages of the dependent and independent variables in the regression equation and where averages are interacted with country-dummies to allow for country-specific parameters. In the heterogeneous version, the Common Correlated Effects Mean Group (CCEMG), the presence of unobserved common factors is achieved by construction and the estimates are obtained as averages of the individual estimates (Pesaran, 2006). A related and recently developed approach due to Eberhardt and Teal (2010) was termed Augmented Mean Group (AMG) estimator and it accounts for cross-sectional dependence by inclusion of a "common dynamic process".[33]

We base our panel analysis on the unrestricted error correction ARDL(*p,q*) representation:

$$\Delta y_{it} = \phi_i y_{it-1} + \beta'_i x_{it-1} + \sum_{j=1}^{p-1} \lambda_{ij} \Delta y_{it-j} + \sum_{q=1}^{q-1} \gamma'_{ij} \Delta x_{it-j} + \mu_i + u_{it}, i = 1,2,...,N; t = 1,2,...,T \tag{10}$$

where y_{it} is a scalar dependent variable, x_{it} is the $k \times 1$ vector of regressors for group *i*, μ_i represents the fixed effects, ϕ_i is a scalar coefficient on the lagged dependent variable. β'_i's is the $k \times 1$ vector of coefficients on explanatory variables, λ_{ij}'s are scalar coefficients on lagged first-differences of dependent variables, and γ_{ij}'s are $k \times 1$ coefficient vectors on first-differences of explanatory variables and their lagged values. We assume that the disturbances u_{it}'s in the ARDL model are independently distributed across *i* and *t*, with zero means and constant variances. Assuming that $\phi_i < 0$ for all *i*, there exists a long-run relationship between y_{it} and x_{it} defined as:

$$y_{it} = \theta'_i y_{it-1} + \eta_{it}, i = 1,2,...,N; t = 1,2,...,T \tag{11}$$

where $\theta'_i = -\beta_i' / \phi_i$ is the $k \times 1$ vector of the long-run coefficients, and η_{it}'s are stationary with possible non-zero means (including fixed effects). Equation (10) can be rewritten as:

[31] The MG approach consists of estimating separate regressions for each country and computing averages of the country-specific coefficients (Evans, 1997; Lee *et al.*, 1997). This allows for heterogeneity of all the parameters.

[32] This estimator allows the intercepts, short-run coefficients and error variances to differ freely across groups, but the long-run coefficients are constrained to be the same. The group-specific short-run coefficients and the common long-run coefficients are computed by the pooled maximum likelihood estimation.

[33] We thank Markus Eberhardt for making his code available.

$$\Delta y_{it} = \phi_i \eta_{it-1} + \sum_{j=1}^{p-1} \lambda_{ij} \Delta y_{it-j} + \sum_{q=1}^{q-1} \gamma'_{ij} \Delta x_{it-j} + \mu_i + u_{it}, i = 1,2,...,N; t = 1,2,...,T \quad (12)$$

where η_{it-1} is the error correction term given by (11), hence ϕ_i is the error correction coefficient measuring the speed of adjustment towards the long-run equilibrium.

Table 5 presents our first set of robustness results, and it includes for each sub-sample both the PMG and MG estimates using different proxies for institutional quality entering in linear form together with the Gwartney and Lawson government size variable. For the OECD sub-group we get a positive and statistically significant coefficient on democracy in specification 4 and three statistically negative coefficients of government size when using the MG estimator. One should expect rich countries to get a negative correlation between government size and growth if thought in terms of the Olson's (1982) mechanism: organized interest groups tend to evolve, and struggle to get advantages for themselves in the form of transfers or legislation, which have a side effect, delaying the regular functioning and growth of economy. The scope for interest group action is likely to be greater in countries with larger governments, where there is increased potential for profits from rent-seeking activities, leading to a greater diversion of resources to unproductive ends (Buchanan, 1980). In a recent paper, Bergh and Karlsson (2010) also uncovered a detrimental growth effect of larger governments in a panel of rich countries using the Bayesian Average over Classical Estimates approach. For both emerging and developing countries (Panels B and C) statistical significance of government size is hard to find,[34] but the institutional proxy is statistically significant for emerging countries (*pr*, political rights, and democracy), and for developing countries (*cl*, civil liberties).

The MG estimator provides consistent estimates of the mean of the long-run coefficients, though these will be inefficient if slope homogeneity holds. Under long-run slope homogeneity, the pooled estimators are consistent and efficient. The hypothesis of homogeneity is tested empirically in all specifications using a Hausman-type test applied to the difference between MG and PMG. Under the null hypothesis the difference in the estimated coefficients between the MG and the PMG estimators is not significant and the PMG is more efficient. The p-value of such a test is also present in Table 6.a, and only for the OECD the null is rejected, being the MG estimator more efficient, and the long-run slope homogeneity rejected.

An equivalent set of results (not shown) with the interaction term between government size and an institutional proxy of interest reveals shows that in the case of the OECD the interaction term is negative and statistically significant for the polity indicator instance. However, the government size is not significant. In the case of developing countries, with the polity variable, government size negatively affects the level of per capita GDP, institutional quality appears with positive and statistically significant estimate and, we get a negative interaction coefficient.

We redo the exercise but similarly to Tables 3 and 4 allow for other proxies of government size to play a role (see Table 6). Only estimated coefficients of the government size proxy, the institutional quality PCA-based measure and the interaction term are reported for reasons of parsimony (full results are available upon request). We present different econometric specifications mainly for robustness and completeness. All in all, we get negative and statistically significant coefficients on total government expenditure, government consumption and public debt-to-GDP ratio irrespectively of the sample under scrutiny. Our results are in line with Romero-Avila and Strauch (2008) who found a negative a significant effect from government consumption (and

---

34 In poor countries public sectors are typically small, and the relationship between government size and growth can even be positive (because a state typically succeeds in collecting taxes when successful at providing the stability necessary for economic activity – sound institutions – to start growth) – see Besley and Persson (2009).

Table 5

Results of Estimations Allowing for Heterogeneous Technology Parameters but Homogeneous Factor Loadings

Panel A

Sample	**OECD**							
Estimator	**PMG**				**MG**			
Spec.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Institutional variable	cl	pr	polity	demo	cl	pr	polity	demo
ln k	0.73*** *(0.090)*	0.55*** *(0.082)*	0.71*** *(0.085)*	0.54*** *(0.104)*	0.68*** *(0.101)*	0.68*** *(0.097)*	0.39*** *(0.068)*	0.47*** *(0.105)*
G	–0.01 *(0.010)*	–0.01 *(0.009)*	–0.01 *(0.011)*	–0.00 *(0.012)*	–0.02* *(0.012)*	–0.02** *(0.010)*	–0.01* *(0.009)*	–0.02 *(0.012)*
I	0.01 *(0.005)*	0.00 *(0.006)*	0.00 *(0.002)*	0.001** *(0.001)*	0.01 *(0.013)*	0.00 *(0.007)*	0.00 *(0.002)*	0.00 *(0.002)*
Error Correction	–0.75*** *(0.192)*	–0.46*** *(0.156)*	–0.79*** *(0.000)*	–0.65*** *(0.000)*	–0.57 *(0.852)*	–0.62 *(0.904)*	–0.88 *(0.909)*	–0.79 *(0.837)*
Hausman test for homogeneity (p-value)	0.05	0.03	0.01	0.03				

Panel B

Sample	**Emerging**							
Estimator	**PMG**				**MG**			
Institutional variable	cl	pr	polity	demo	cl	pr	polity	demo
ln k	0.88*** *(0.173)*	0.94*** *(0.163)*	0.76*** *(0.200)*	1.33*** *(0.340)*	–0.12 *(0.642)*	0.28* *(0.155)*	–0.09 *(0.391)*	–0.69 *(0.544)*
G	–0.01 *(0.020)*	–0.00 *(0.014)*	–0.01 *(0.011)*	–0.01 *(0.020)*	–0.02 *(0.028)*	–0.02 *(0.024)*	0.01 *(0.031)*	0.01 *(0.029)*
I	0.01 *(0.007)*	0.02* *(0.120)*	–0.01 *(0.007)*	0.01* *(0.004)*	0.02 *(0.040)*	–0.02 *(0.021)*	0.01 *(0.019)*	0.00 *(0.008)*
Error Correction	–0.69*** *(0.000)*	–0.72*** *(0.001)*	–0.75*** *(0.000)*	0.83*** *(0.002)*	–0.90*** *(0.172)*	–0.51 *(1.43)*	–0.71*** *(0.181)*	–0.92*** *(0.177)*
Hausman test for homogeneity (p-value)	0.31	0.02	0.31	0.26				

Panel C

Sample	**Developing**							
Estimator	**PMG**				**MG**			
Institutional variable	cl	pr	polity	demo	cl	pr	polity	demo
ln k	0.33*** *(0.091)*	0.11 *(0.110)*	0.63*** *(0.109)*	0.45*** *(0.113)*	0.81*** *(0.255)*	0.79*** *(0.234)*	0.52*** *(0.193)*	0.68*** *(0.230)*
g	0.01 *(0.007)*	0.01 *(0.004)*	0.003 *(0.009)*	0.001 *(0.009)*	–0.02 *(0.021)*	–0.02 *(0.018)*	–0.01 *(0.011)*	–0.02* *(0.012)*
I	–0.01 *(0.008)*	–0.01 *(0.012)*	0.01 *(0.012)*	–0.001 *(0.002)*	0.03** *(0.016)*	–0.02 *(0.016)*	0.00 *(0.020)*	0.003 *(0.003)*
Error Correction	–0.54*** *(0.001)*	–0.18*** *(0.001)*	–0.72*** *(0.000)*	–0.60*** *(0.000)*	–0.76*** *(0.085)*	–0.71*** *(0.088)*	–0.25 *(0.249)*	–0.93*** *(0.128)*
Hausman test for homogeneity (p-value)	0.11	0.85	0.15	0.18				

Note: The models are estimated by either PMG or MG estimators. The dependent variable is the logarithm of real GDP per capita. A time trend has been included but is not reported for reasons of parsimony. Hausman test for homogeneity: under the null hypothesis the difference in the estimated coefficients between the MG and PMG estimators, it is not significant and PMG is more efficient. *, **, *** denote significance at 10, 5 and 1 per cent levels.

Table 6

Results of Estimations Allowing for Homogeneous and/or Heterogeneous Technology Parameters and Factor Loadings, With and Without Interaction Terms. PCA-based Institutional Measure. Different Government Size Proxies

Sample	OECD				Emerging				Developing			
Estimator	**OLS**	**MG**	**CCEP**	**AMG**	**OLS**	**MG**	**CCEP**	**AMG**	**OLS**	**MG**	**CCEP**	**AMG**
Spec.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
totgovexp_gdp	0.00	–0.002***	–0.01***	–0.00*	–0.03***	0.00	–0.001***	0.00	–0.00	–0.00	–0.001***	–0.00
	(0.001)	*(0.001)*	*(0.001)*	*(0.001)*	*(0.005)*	*(0.001)*	*(0.001)*	*(0.002)*	*(0.003)*	*(0.001)*	*(0.001)*	*(0.001)*
I	1.02***	0.02	0.014	–0.49	0.43***	–2.60	0.01	–4.29	0.65***	–3.91	0.01	–0.00
	(0.059)	*(2.491)*	*(0.032)*	*(2.903)*	*(0.068)*	*(2.598)*	*(0.010)*	*(4.293)*	*(0.039)*	*(3.894)*	*(0.017)*	*(0.019)*
govcons_gdp	–0.02***	0.00	–0.02***	0.00	–0.06***	–0.00	–0.001**	–0.00	–0.02***	0.00	–0.003**	–0.00
	(0.005)	*(0.002)*	*(0.002)*	*(0.002)*	*(0.006)*	*(0.002)*	*(0.002)*	*(0.002)*	*(0.003)*	*(0.002)*	*(0.001)*	*(0.002)*
I	0.93***	1.56	0.04***	3.89**	0.46***	–0.01	0.00	–0.00	0.63***	–0.04	–0.00	–0.02
	(0.058)	*(1.056)*	*(0.012)*	*(1.768)*	*(0.058)*	*(0.017)*	*(0.010)*	*(0.016)*	*(0.028)*	*(0.027)*	*(0.011)*	*(0.022)*
govdebt_gdp	0.00	–0.00	–0.001***	–0.00	–0.001***	–0.00	0.00	–0.001**	–0.002**	–0.00	–0.001***	–0.002**
	(0.001)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.001)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.002)*	*(0.000)*	*(0.001)*
I	1.09***	1.17	0.04***	1.99	0.45***	0.00	–0.01	–0.01	0.62***	–2.86	0.00	–2.86
	(0.053)	*(1.988)*	*(0.013)*	*(2.410)*	*(0.062)*	*(0.020)*	*(0.011)*	*(0.019)*	*(0.031)*	*(2.414)*	*(0.011)*	*(2.628)*
totgovexp_gdp	–0.001*	4.42	0.01***	–0.26	–0.03***	6.94	–0.001***	–0.00	0.00	–0.02	–0.01***	–0.01*
	(0.003)	*(5.179)*	*(0.001)*	*(0.747)*	*(0.005)*	*(6.946)*	*(0.001)*	*(0.002)*	*(0.003)*	*(0.020)*	*(0.001)*	*(0.006)*
I	1.16***	152.49	0.01	–10.31	0.76***	243.48	0.03	0.07	0.28**	–0.40	0.12***	0.12
	(0.091)	*(180.465)*	*(0.033)*	*(16.802)*	*(0.229)*	*(243.301)*	*(0.028)*	*(0.083)*	*(0.118)*	*(0.837)*	*(0.039)*	*(0.251)*
I*g	–0.00*	–4.53	0.00	0.22	–0.01*	–6.96	–0.00	–0.00	–0.02***	0.01	–0.004***	–0.00
	(0.003)	*(5.162)*	*(0.001)*	*(0.624)*	*(0.007)*	*(6.959)*	*(0.001)*	*(0.003)*	*(0.005)*	*(0.027)*	*(0.001)*	*(0.009)*
govcons_gdp	–0.09***	–2.04	0.00	–2.66	–0.06***	0.68	–0.01***	–0.63	–0.02***	–0.17	–0.003***	–0.16
	(0.014)	*(2.120)*	*(0.004)*	*(2.215)*	*(0.006)*	*(0.980)*	*(0.002)*	*(0.743)*	*(0.003)*	*(0.173)*	*(0.001)*	*(0.175)*
I	0.26*	–46.66	0.11***	0.78*	0.73***	12.56	0.16***	–12.10	0.78***	–10.40	0.09***	–10.57
	(0.155)	*(32.780)*	*(0.039)*	*(0.394)*	*(0.179)*	*(19.236)*	*(0.028)*	*(14.459)*	*(0.077)*	*(10.266)*	*(0.024)*	*(10.325)*
I*g	–0.10***	1.74	–0.01***	2.37	–0.02*	–0.68	–0.01***	0.64	–0.01**	0.30	–0.01***	0.31
	(0.012)	*(1.775)*	*(0.003)*	*(1.907)*	*(0.010)*	*(0.981)*	*(0.002)*	*(0.743)*	*(0.005)*	*(0.290)*	*(0.001)*	*(0.292)*
govdebt_gdp	–0.00	–0.26	–0.001***	–0.32	–0.002***	0.89	0.00	0.41	–0.00	0.24	–0.002***	0.20
	(0.002)	*(0.288)*	*(0.000)*	*(0.271)*	*(0.001)*	*(1.096)*	*(0.000)*	*(0.476)*	*(0.000)*	*(0.188)*	*(0.000)*	*(0.204)*
I	0.91***	–9.52	0.05**	–9.93	0.60***	15.50	–0.02	7.53	0.72***	1.64	0.00	5.23
	(0.104)	*(9.635)*	*(0.019)*	*(9.260)*	*(0.119)*	*(21.701)*	*(0.017)*	*(9.332)*	*(0.049)*	*(4.870)*	*(0.014)*	*(5.012)*
I*g	–0.002*	0.24	–0.00	0.29	–0.001*	–0.90	0.00	–0.42	–0.002**	–0.24	–0.00	–0.34
	(0.002)	*(0.256)*	*(0.000)*	*(0.241)*	*(0.002)*	*(1.096)*	*(0.000)*	*(0.476)*	*(0.001)*	*(0.307)*	*(0.000)*	*(0.342)*

Note: The models are estimated by Pooled OLS, MG, CCEP or AMG estimators. The dependent variable is the logarithm of real GDP per capita. *, **, *** denote significance at 10, 5 and 1 per cent levels.

transfers) on economic growth. We refrain from making a detailed analysis. Still, for instance, specifications 7 and 11 for the emerging and developing countries groups and with the government consumption as a proxy for government size show a negative effect of government consumption, and a positive effect of the PCA-based institutional measure. Finally, there is a negative interaction term: i) the negative effect of government consumption on GDP per capita is stronger at lower levels of institutional quality, and ii) the positive effect of institutional quality on GDP per capita increases at smaller levels of government consumption.

5 Conclusion

We outlined a growth model with an explicit government role showing that more resources required to finance government spending reduce both the optimal level of private consumption and of output per worker. Following up on that theoretical motivation we perform an empirical panel analysis with 108 countries from 1970-2008, employing different proxies for government size and institutional quality.

Therefore, we provide additional evidence on the issue of whether “too much” government is good or bad for economic progress and macroeconomic performance, particularly when associated with differentiated levels of (underlying) institutional quality and alternative political regimes.

Moreover, we make use of recent panel data techniques that allow for the possibility of heterogeneous dynamic adjustment around the long-run equilibrium relationship as well as heterogeneous unobserved parameters and cross-sectional dependence (e.g., Pooled Mean Group, Mean Group, Common Correlated Pooled estimators, inter alia); we also deal with potentially relevant endogeneity issues.

Our results allow for several conclusions regarding the effects on economic growth of the size of the government: i) there is a significant negative effect of the size of government on growth; ii) institutional quality has a significant positive impact on the level of real GDP per capita; iii) government consumption is consistently detrimental to output growth irrespective of the country sample considered (OECD, emerging and developing countries); iv) moreover, the negative effect of government size on GDP per capita is stronger at lower levels of institutional quality, and the positive effect of institutional quality on GDP per capita is stronger at smaller levels of government size. Therefore, our empirical results are consistent with the growth model presented in the paper.

In addition, the negative effect on growth stemming from the government size variables is more attenuated for the case of Scandinavian legal origins, while the negative effect of government size on GDP per capita growth is stronger at lower levels of civil liberties and political rights.

APPENDIX
VARIABLES AND SOURCES

Variable	Definition/Description	Acronym	Source
REAL GDP per capita		*Gdppc*	World Bank's Word Development Indicators (WDI)
gross fixed capital formation (% GDP)		*Gfcf_gdp*	WDI
Public investment (% GDP)		*Pubinv_gdp*	WDI and AMECO for advanced countries
real aggregate investment in PPP		*Inv*	Summers and Heston's PWT 6.3
Government size	Composite variable (*govsize*). This variable includes government consumption expenditures (as percentage of total consumption), transfers and subsidies (as percentage of GDP), the underlying tax system (proxied by top marginal tax rates) and the number of government enterprises.	*govsize*	Gwartney and Lawson (2008)
Central Government Debt (% GDP)		*Govdebt_gdp*	IMF (Abas *et al.*, 2010)
Total Government Expenditure (% GDP)		*Totgovexp_gdp*	WDI, IMF IFS, Easterly (2001)
Public Final Consumption Expenditure (% GDP)		*Govcons_gdp*	WDI, IMF IFS, Easterly (2001)
Polity 2	The polity score is computed by subtracting the autoc score (autocracy index) from the democ score (democracy index); the resulting unified polity scale ranges from +10 (strongly democratic) to –10 (strongly autocratic). Refer to the database's supporting documentation for more details.	*polity*	Marshall and Jaegger's Polity's 4 database
Political Rights	Political rights enable people to participate freely in the political process, including the right to vote freely for distinct alternatives in legitimate elections, compete for public office, join political parties and organizations, and elect representatives who have a decisive impact on public policies and are accountable to the electorate.	*pr*	Freedom House
Civil Liberties	Civil liberties include freedom of speech, expression and the press; freedom of religion; freedom of assembly and association; and the right to due judicial process.	*cl*	Freedom House
corruption perception index	The CPI focuses on corruption in the public sector and defines corruption as the abuse of public office for private gain. The CPI Score relates to perceptions of the degree of corruption as seen by business people, risk analysts and the general public.	*cpi*	Transparency International database
index of democratization	This index combines two basic dimensions of democracy – competition and participation – measured as the percentage of votes not cast for the largest party (Competition) times the percentage of the population who actually voted in the election (Participation).	*demo*	Vanhanen (2005)
governance index	This is the result of averaging 6 variables: voice and accoutability, political stability, government effectiveness, regulatory quality, rule of law and control of corruption.	*governance*	Kaufman *et al.* (2009)
legal origins	English, French, German or Scandinavian	*bri*, *fre*, *ger* and *sca*	La Porta *et al.*, 1999
Regime durability	The number of years since the most recent regime change (defined by a three point change in the p_polity score over a period of three years or less) or the end of transition period defined by the lack of stable political institutions (denoted by a standardized authority score).	*Durable*	Marshall and Jaegger's Polity's 4 database
latitude		*latitude*	La Porta *et al.*, 1999
ethnic fragmentation	Reflects probability that two randomly selected people from a given country will not belong to the same ethnolinguistic group. The higher the number, the more fractionalized society.	*ethnic*	Alesina *et al.*, 2003

Countries in the dataset

Afghanistan, Albania, Algeria, Angola, Argentina, Armenia, Australia, Austria, Azerbaijan, Bahrain, Bangladesh, Belarus, Belgium, Benin, Bhutan, Bolivia, Botswana, Brazil, Bulgaria, Burkina Faso, Burundi, Cambodia, Cameroon, Canada, Chad, Chile, China, Colombia, Comoros, Congo, Dem. Rep., Congo, Rep., Costa Rica, Cote d'Ivoire, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Ecuador, Egypt, Arab Rep., El Salvador, Equatorial Guinea, Estonia, Finland, France, Gabon, Gambia, The, Georgia, Germany, Ghana, Greece, Guatemala, Guinea, Guyana, Hungary, Iceland, India, Indonesia, Iran, Islamic Rep., Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kazakhstan, Kenya, Kiribati, Korea, Rep., Kuwait, Kyrgyz Republic, Lao PDR, Latvia, Lebanon, Lesotho, Liberia, Libya, Lithuania, Luxembourg, Macedonia, FYR, Madagascar, Malawi, Malaysia, Mauritania, Mauritius, Mexico, Mongolia, Montenegro, Morocco, Namibia, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Qatar, Romania, Russian Federation, Rwanda, Samoa, San Marino, Saudi Arabia, Senegal, Serbia, Sierra Leone, Singapore, Slovak Republic, Slovenia, Somalia, South Africa, Spain, Sri Lanka, Sudan, Suriname, Swaziland, Sweden, Switzerland, Syrian Arab Republic, Thailand, Trinidad and Tobago, Tunisia, Turkey, Uganda, Ukraine, United Arab Emirates, United Kingdom, United States, Uruguay, Uzbekistan, Venezuela, RB, Vietnam, Yemen, Rep., Zambia, Zimbabwe.

REFERENCES

Abas, A., N. Belhocine, A. El Ganainy and M. Horton (2010), "A Historical Public Debt Database", IMF, Working Paper, No. 10/245.

Acemoglu, D. and J. Robinson (2005), *Economic Origins of Dictatorship and Democracy*, Cambridge University Press.

Acemoglu, D., S. Johnson and J.A. Robinson (2001),"The Colonial Origins of Comparative Development: An Empirical Investigation", *American Economic Review*, No. 91, pp. 1369-401.

Afonso, A., L. Schuknecht and V. Tanzi (2005). "Public Sector Efficiency: An International Comparison", *Public Choice*, Vol. 123, No. 3-4, pp. 321-47.

—————(2011), "Income Distribution Determinants and Public Spending Efficiency", *Journal of Economic Inequality*, Vol. 8, No. 3, pp. 367-89.

Alesina, A. and E. La Ferrara (2000), "The Determinants of Trust", NBER, Working Paper, No. 7621.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat and R. Wacziarg (2003), "Fractionalization", *Journal of Economic Growth*, No. 8, pp. 155-94.

Angelopoulos, K., A. Philippopoulos and E. Tsionas (2008), "Does Public Sector Efficiency Matter? Revisiting the Relation Between Fiscal Size and Economic Growth in a World Sample", *Public Choice*, Vol. 137, No. 1-2, pp. 245-78.

Bajo-Rubio, O. (1991), "A Further Generalization of the Solow Growth Model: The Role of the Public Sector", *Economic Letters*, Vol. 68, pp. 79-84.

Banerjee, A. (1999), "Panel Data, Unit Roots and Cointegration: An Overview", *Oxford Bulletin of Economics and Statistics*, special issue on "Testing Unit Roots and Cointegration Using Panel Data", *Theory and Applications*, No. 61, pp. 607-29, November.

Barro, R. (1991), "Economic Growth in a Cross Section of Countries", *Quarterly Journal of Economics*, No. 106, pp. 407-43.

Barro, R. and X. Sala–i–Martín, (1992), "Convergence", *Journal of Political Economy*, No. 100, p. 223-51.

Beck, T., G. Clarke, A. Groff, P. Keefer and P. Walsh (2001), "New Tools and New Tests in Comparative Political Economy: The Database of Political Institutions", *World Bank Economic Review*, No. 15, pp. 165-76.

Bergh, A. and M. Karlsson (2010), "Government Size and Growth: Accounting for Economic Freedom and Globalization", *Public Choice*, No. 142, pp. 195-213.

Bergh, A. and M. Henrekson (2011), "Government Size and Growth: A Survey and Interpretation of the Evidence", *Journal of Economic Surveys*, Vol. 25, No. 5, pp. 872-97.

Besley, T. and T. Persson (2009), "The Origins of State Capacity: Property Rights, Taxation, and Policy", *American Economic Review*, Vol. 99, No. 4, pp. 1218-44.

Bockstette, V, A. Chanda and L. Putterman (2002), "States and Markets: The Advantage of an Early Start", *Journal of Economic Growth*, Vol. 7, No. 4, pp. 347-69.

Browsher, C.G. (2002), "On Testing Overidentifying Restrictions in Dynamic Panel Data Models", CEPR, Discussion Paper, No. 3048, London.

Buchanan, J. (1980), “Rent-seeking and Profit-seeking”, in J.M. Buchanan and G. Tullock (eds.), *Towards a Theory of the Rent–seeking Society*, College Station (TX).

Chen, B. (2006), “Economic Growth with Optimal Public Spending Composition”, *Oxford Economic Papers*, No. 58, pp. 123-36.

Coakley, J., A.M. Fuertes and R. Smith (2006), “Unobserved Heterogeneity in Panel Time Series Models”, *Computational Statistics and Data Analysis*, Vol. 50, No. 9, pp. 2361-80.

Conte, M.A. and A.F. Darrat (1988), “Economic Growth and the Expanding Public Sector”, *Review of Economics and Statistics*, Vol. 70, No. 2, pp. 322-30.

Dalamagas, B. (2000), “Public Sector and Economic Growth: The Greek Experience”, *Applied Economics*, No. 32, pp. 277-88.

Dempster, A.P., N.M. Laird and D.B. Rubin (1977), “Maximum Likelihood from Incomplete Data Via the EM Algorithm”, *Journal Royal Statistical Society B*, No. 39, pp. 1–22.

Devarajan, S., V. Swaroop and H. Zou (1996), “The Composition of Public Expenditure and Economic Growth”, *Journal of Monetary Economics*, No. 37, pp. 313-44.

Dollar, D. and A. Kraay (2003), “Institutions, Trade and Growth”, *Journal of Monetary Economics*, No. 50, pp. 133-62.

Dupuit, J. (1844), “De la Mesure de l’Utilité des Travaux Publiques”, in *Annales des Ponts et Chaussées, Vol. 8*, translated and reprinted in: K. Arrow and T. Scitovski (eds.) (1969), *AEA Readings in Welfare Economics*, AEA, pp. 255-83.

Easterly, W.R. (2001), “The Lost Decades: Developing Countries Stagnation in Spite of Policy Reform 1980-1998”, *Journal of Economic Growth*, Vol. 6, No. 2, pp. 135-57.

Eberhardt, M. and F. Teal (2010), “Productivity Analysis in Global Manufacturing Production”, Oxford University, Discussion Paper, No. 515, November.

Evans, P. (1997), “How Fast Do Economies Converge”, *Review of Economics and Statistics*, No. 46, pp. 1251-71.

Fölster, S. and M. Henrekson (2001), “Growth Effects of Government Expenditure and Taxation in Rich Countries”, *European Economic Review*, No. 45, pp. 1501-20.

Friedman, M. (1997), “If Only the US Were as Free as Hong Kong”, *Wall Street Journal*, July 8, A14.

Ghali, K.H. (1998), “Government Size and Economic Growth: Evidence from a Multivariate Cointegration Analysis”, *Applied Economics*, Vol. 31, pp. 975-87.

Glomm, G. and B. Ravikumar (1997), “Productive Government Expenditures and Long-run Growth”, *Journal of Economic Dynamics and Control*, Vol. 21, pp. 183-204.

Goel, R. and M. Nelson (1998), “Corruption and Government Size: A Disaggregated Analysis”, *Public Choice*, Vol. 97, No. 1-2, pp. 107-20.

Granger, C.W.J. and P. Newbold (1974), “Spurious Regressions in Econometrics”, *Journal of Econometrics*, No. 2, pp. 111-20.

Greene, W. (2003), *Econometric Analysis*, Pearson Education Inc. (NJ).

Grossman, P. (1988), “Government and Economic Growth: A Non–linear Relationship”, *Public Choice*, Vol. 56, No. 2, pp. 193-200.

Guseh, J.S. (1997), “Government Size and Economic Growth in Developing Countries: A Political–economy Framework”, *Journal of Macroeconomics*, No. 19, pp. 175-92.

Gwartney, J. and R.A. Lawson (2008), "Economic Freedom of the World: 2008 Annual Report", available at: http://freetheworld.com/

Hall, R.E. and C. Jones (1999), "Why Do Some Countries Produce So Much More Output per Worker than Others?", *Quarterly Journal of Economics*, No. 114, pp. 83-116.

Holcombe, R. (2005), "Government Growth in the Twenty–first Century", *Public Choice*, Vol. 124, No. 1-2, pp. 95-114.

Im, K.S., M.H. Pesaran and Y. Shin (2003), "Testing for Unit Roots in Heterogeneous Panels", *Journal of Econometrics*, No. 115, pp. 53-74.

IMF (2001), "A Manual on Government Finance Statistics 2001", (GFSM 2001), IMF, Washington (D.C.), available at: http://www.imf.org/external/pubs/ft/gfs/ manual/gfs.htm

Jalles, J.T. (2010), "Does Democracy Foster or Hinder Growth? Extreme–type Political Regimes in a Large Panel", *Economics Bulletin*, Vol. 30, No. 2, pp. 1359-72.

Kapetanios, G., M.H. Pesaran and T. Yamagata (2011), "Panels with Non-stationary Multifactor Error Structures", *Journal of Econometrics*, Vol. 160, No. 2, pp. 326-48.

Kaufmann, D., A. Kraay and M. Mastruzzi (2009), "Governance Matters VIII: Aggregate and Individual Governance Indicators for 1996-2008", World Bank, Policy Research Paper, No. 4978, available at: http://ssrn.com/abstract=1424591

Kneller, R., M. Bleaney and N. Gemmell (1999), "Fiscal Policy and Growth: Evidence from OECD Countries", *Journal of Public Economics*, No. 74, pp. 171-90.

La Porta, R., F. López-de-Silanes, A. Shleifer and R. Vishny (1997), "Legal Determinants of External Finance", *Journal of Finance*, No. 52, pp. 1131-50.

—————— (1998), "Law and Finance", *Journal of Political Economy*, No. 106, pp. 1113-55.

—————— (1999), "The Quality of Government", *Journal of Law, Economics and Organization*, Vol. 15, No. 1, pp. 222-79.

Lee, J. (1992), "Optimal Size and Composition of Government Spending", *Journal of the Japanese and International Economies*, No. 6, pp. 423-39.

Lee, K., M.H. Pesaran and R.P. Smith (1997), "Growth and Convergence in a Multi–country Empirical Stochastic Solow Model", *Journal of Applied Econometrics*, No. 12, pp. 357-92.

Little, J.A. and D.B. Rubin (1987), *Statistical Analysis with Missing Data*, John Wiley and Sons, New York (NY).

Maddala, G.S. and S. Wu (1999), "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test", *Oxford Bulletin of Economics and Statistics*, Vol. 61, pp. 631-52, November.

Mankiw, N.G., D. Romer and D.N. Weil (1992), "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics*, No. 107, pp. 407-37.

Mauro, P. (1995), "Corruption and Growth", *Quarterly Journal of Economics*, No. 110, pp. 681-712.

Nelson, R.R. and B.N. Sampat (2001), "Making Sense of Institutions as a Factor Shaping Economic Performance", *Journal of Economic Behavior and Organization*, No. 44, pp. 31-54.

North, D. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, Cambridge (MA).

Olson, M. (1982), *The Rise and Decline of Nations: Economic Growth, Stagflation, and Social Rigidities*, Yale University Press, New Haven (CN).

Peden, E. and M. Bradley (1989), "Government Size, Productivity and Economic Growth: The Post–war Experience", *Public Choice*, Vol. 61, No. 3, pp. 229-45.

Pesaran, M.H. (2006), "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure", *Econometrica*, Vol. 74, No. 4, pp. 967-1012.

—————(2007), "A Simple Panel Unit Root Test in the Presence of Cross Section Dependence", *Journal of Applied Econometrics*, Vol. 22, No. 2, pp. 265-312.

Pesaran, M.H. and R.P. Smith (1995), "Estimating Long–run Relationship from Dynamic Heterogeneous Panels", *Journal of Econometrics*, No. 68, pp. 79-113.

Pesaran, M.H., Y. Shin and R.P. Smith (1999), "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels", *Journal of American Statistical Association*, No. 94, pp. 31-54.

Poterba, J.M. (1996), "Budget Institutions and Fiscal Policy in the U.S. States", *American Economic Review*, Vol. 86, No. 2, pp. 395-400.

Przeworski, A., M. Alvarez, J. Cheibub and F. Limongi (2000), "Democracy and Development: Political Regimes and Economic Well–being in the World, 1950-1990", Cambridge University Press, New York (NY).

Ram, R. (1986), "Government Size and Economic Growth: A New Framework and Some Evidence from Cross–section and Time Series Data", *American Economic Review*, No. 76, pp. 191-203.

Rock, M. (2009), "Has Democracy Slowed Growth in Asia?", *World Development*, Vol. 37, No. 5, pp. 941-52.

Rodrik, D. and R. Wacziarg (2005), "Do Democratic Transitions Produce Bad Economic Outcomes?", *American Economic Review*, Vol. 95, No. 3, pp. 50-55.

Romero–Ávila, D. and R. Struch (2008), "Public Finances and Long–term Growth in Europe: Evidence from a Panel Data Analysis", *European Journal of Political Economy*, Vol. 24, No. 1, pp. 172–91.

Roodman, D.M. (2009), "A Note on the Theme of Too Many Instruments", *Oxford Bulletin of Economics and Statistics*, Vol. 71, No. 1, pp. 135-58.

Rousseeuw, P.J. and K. Van Driessen (1999), "A Fast Algorithm for the Minimum Covariance Determinant Estimator", *Technometrics*, No. 41, pp. 212-23.

Slemrod, J. (1995), "What Do Cross–country Studies Teach About Government Involvement, Prosperity, and Economic Growth?", Brookings Papers on Economic Activity, No. 2, pp. 373-431.

Tanzi V. and H. Zee (1997), "Fiscal Policy and Long–run Growth", IMF, Staff Papers, No. 44, pp. 179-209.

Vanhanen, T. (2005), "Measures of Democracy 1810-2004" [computer file]. FSD1289, version 2.0 (2005-08-17), Tampere: Finnish Social Science Data Archive [distributor].

FISCAL POLICY AND ECONOMIC GROWTH: THE CASE OF ALBANIA

Gerti Shijaku[*] and Arlind Gjokuta[**]

This discussion material analysis the effects of fiscal policy on the economic growth in the case of a small open developing country, Albania, by employing an endogenous growth model on a GMM approach. The results obtained show that government revenue growth has a higher effect on economic growth than government expenditure. The impact of revenue and expenditure on growth were analysed by categorising tax revenue into distortionary and non-distortionary, whilst government expenditure were divided into productive and non-productive. Under such composition we found that revenue sub-categories reduce growth, while distortionary taxation has much larger and statistically significant effect. Besides, the parameter values show that growth is effected positively by productive expenditure and negatively by non-productive. This material also analysis the impact of public debt on growth and finds that the size of public debt is negatively related to growth rate.

1 Introduction

The role of fiscal policy (FP) on economic growth has driven several studies both on the theoretical and on the empirical fronts. Modern macroeconomic literature emphasises both the short run and the long run objectives of FP (Romer, 2006). In the short run it can be used to counter output cyclicality and/or stabilise volatility in macro variables, which is descriptively same as the effects of the short run monetary policy. Further for the long-run, FP and the debt financing methods can also affect both demand and supply side of the economy. The subject on the effects of FP on economic growth is quite relevant, since the development of appropriate fiscal instruments could lead to a persistent and sustainable boost on economic growth. Thus, the aim of this paper is to examine the fiscal policy-growth relationship in the case of a small open developing country, Albania, as it is crucial to know how public activities through taxation and expenditure policies have served as an incentive to growth.

By the end of the '90s and during the last decade, Albanian economic policies aimed at maintaining macroeconomic stability, enable poverty-reducing and non-inflationary economic growth policies and achieving fiscal consolidation through budget deficit and public debt reduction. As such public finance saw major reformation aiming at government expenditure cuts and boosting revenues, expanding the tax base, simplifying and implementing new tax system, promoting tax intensive through reducing tax burden on business, and reducing informality and tax evasion. Tax revenues witnessed major reductions in custom duties rate due to Free Trade Agreements under the Stabilization and Association agreement with the European Union, the CEFTA and World Trade Organization membership. This was followed by considerable raise in national, local and excise tax level, cuts in social contributions and small business tax and the changes in the threshold for VAT registration. In addition, tax legislation changes were finalised with the elimination of all exclusions and facilitations under the old tax system, the approval of a 10 per cent flat income tax in 2007 and the reduction of the profit tax to only 10 per cent in 2008.

Further, the Albanian economy took advantages of macroeconomic stimulus in the form of fiscal expansion during 2007-09, mainly as a result of previous work to consolidate the fiscal

[*] Bank of Albania, Research Department, E-mail: gshijaku@bankofalbania.org

[**] Bank of Albania, Monetary and Operational Department, E-mail: agjokuta@bankofalbania.org

The views expressed herein are of the authors and do not necessarily reflect the views of the Bank of Albania.

position and the anchoring of macroeconomic policies and public expectations. Albanian economy, hence, was faced with the effect of global crisis enjoying a counter-cyclical FP during 2009, reinforcing the trend that began during the period 2007-08. On the other hand, apparently these economic incentives mitigated the adverse effects that had on the Albanian economy the global financial crisis.

In this case the questions coming up relate to the analysis of what are the concrete effects of fiscal policies on economic growth, in the case of Albania? Have they stimulated economic growth? This discussion paper focused on how the government activities, namely composition of expenditures and revenues, affect the long run growth rate? The answer to these questions is quite difficult because the transmission operation mechanisms of the effects of FP are quite complex and above all the effects take time to be displayed fully.

To our best knowledge, fiscal-growth relationship has only recently been empirically studied in the case of Albania, so far. In a recent discussion material, Mançellari (2011) studied the effects of FP in Albania based on a model with four macroeconomics variables, namely FP, Gross Domestic Product (GDP), interest rates and the prices level, through a SVAR and impulse responses approach. The analysis was based on the methodology developed by Blanchard and Perotti (2002). The main findings of this paper, concluded that FP does affect economic activity, cuts in tax burden has the highest cumulative GDP multiplier and the GDP multiplier of capital expenditure is greater that current expenditure multiplier.

In this paper, unlike Mançellari (2011), we contribute to the fiscal-growth subject in the case of Albania in various ways. First, FP is considered to be endogenous, but we based our empirical analysis of fiscal-growth relationship on a different endogenous economic growth model. This approach incorporates the public sector, namely FP, into the growth model of Solow. Second, by doing so, we can include a richer menu of FP effects by identifying and incorporating the specific FP variables as to enhance economic growth in Albania, namely the distortionary and non-distortionary public revenues and productive and non-productive public expenditures. Additionally, we consider the effect of public debt to GDP ratio to examine whether financing capital expenditures through borrowing (indebtedness) has served as growth-promoting or reducing. Finally, we tried to empirically identify the effect of FP throughout different time-samples, mainly 1998-2006 and 1998-2010.

In Section 2 we summarise some key developments in Albanian FP during 1998-2010. The relevant empirical model and the data are outlined in following section. Then, Section 4 presents the empirical results. The material concludes with main findings in Section 5.

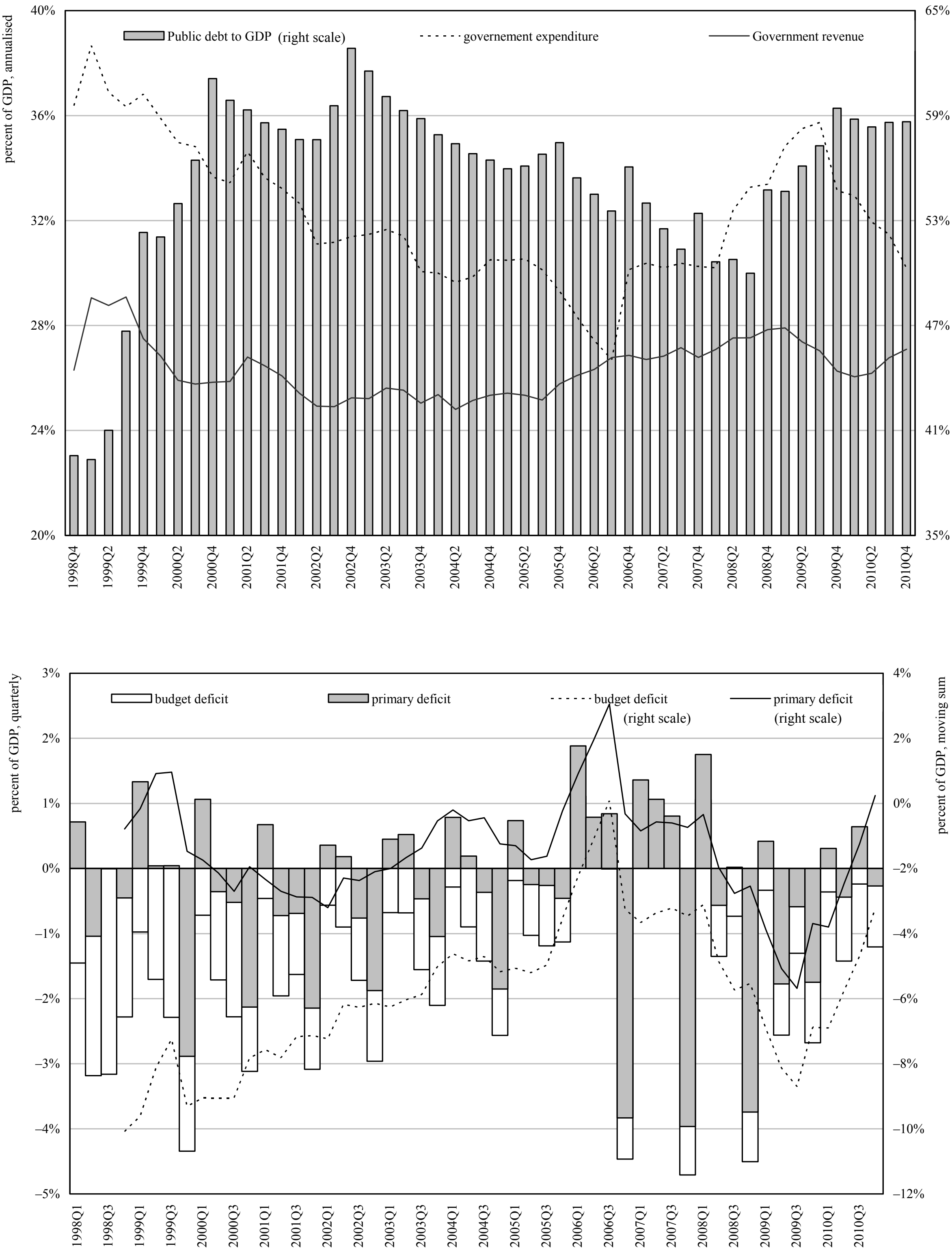
2 Albanian fiscal policy during 1998-2010

Under the IMF program support, the Albanian government focused on maintaining macroeconomic stability, reducing poverty and achieving sustainable non-inflationary economic growth,[1] after gradual orientation towards a market economy in early 1990 and fast improvement of an important part of economic indicators by the end of the ’90s. The government also aimed at achieving fiscal consolidation through budget deficit and public debt reduction through continuous fiscal consolidation. For this reason, public finance has been under continuous scrutiny of major reformation on expenditure and tax collection system. The philosophy of these fiscal reforms was based on the idea of reducing current expenditures (mainly personnel expenditure, subsidies and

[1] See: Enhanced Structural Adjustment Facility (ESAF – 1998-2001), Poverty Reduction and Economic Growth (PREG – 2002-05) and it was extended to Extended Fund Facility (EFF – 2006-09). In January 2009, Albania graduated from the Fund-supported program.

Figure 1

Selected Fiscal Indicators, 1998-2010

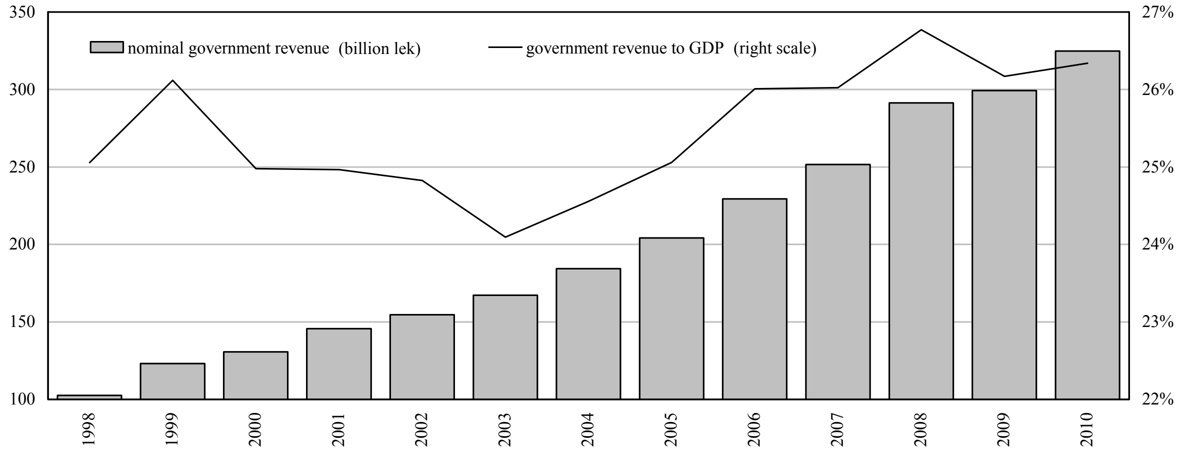


Source: Ministry of Finance.

Figure 2

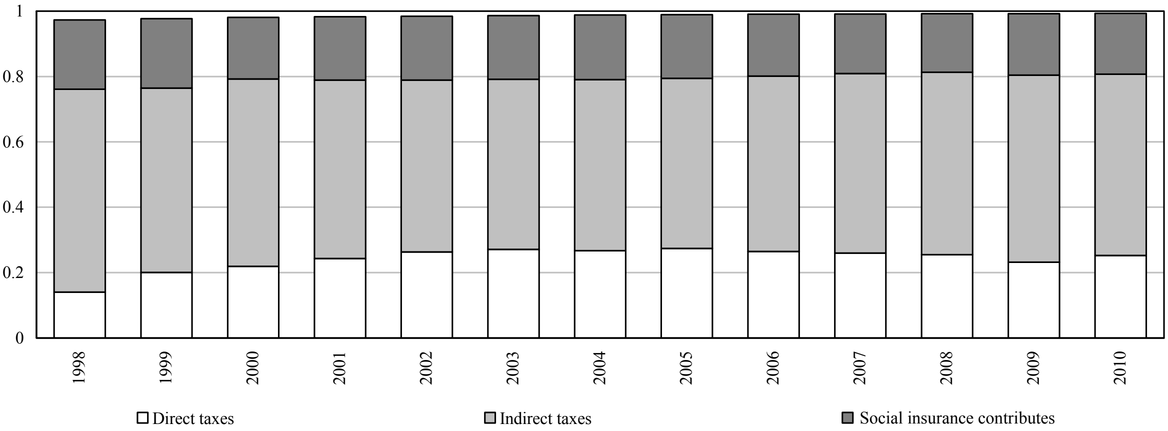
Government Revenue Indicators, 1998-2010

Government Revenue Dynamics

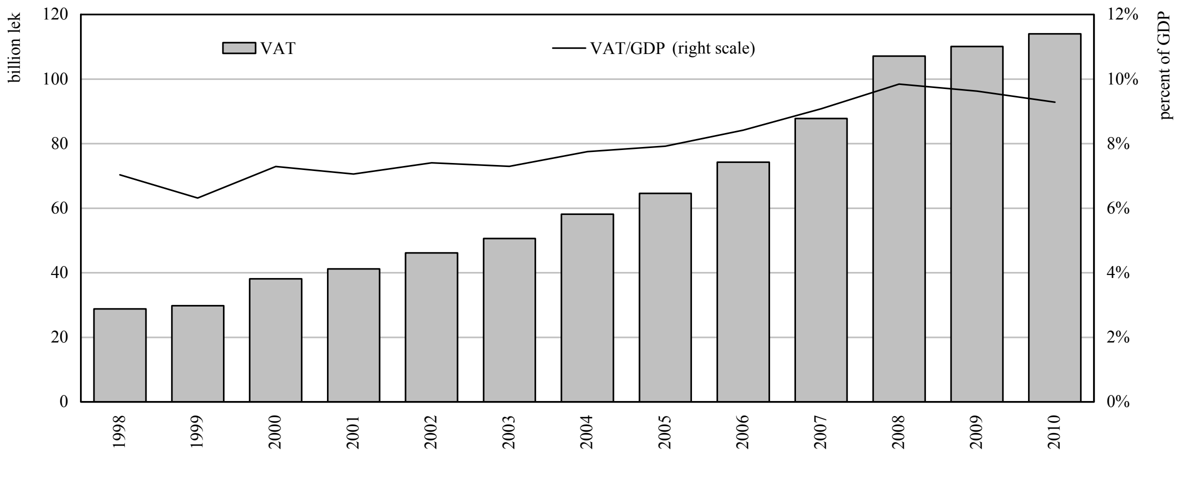


Main Revenue Sources

(percent of GDP)



Value Added Tax Burden

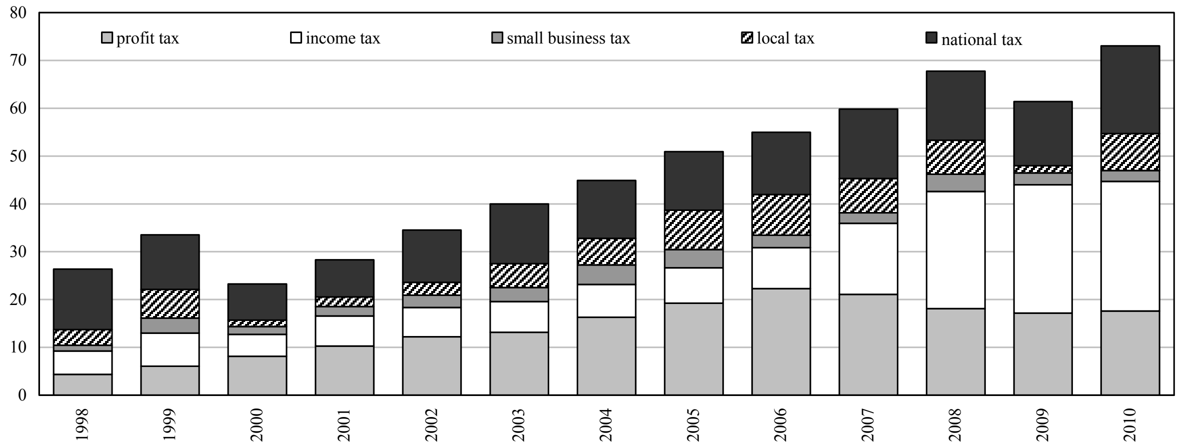


Source: Ministry of Finance.

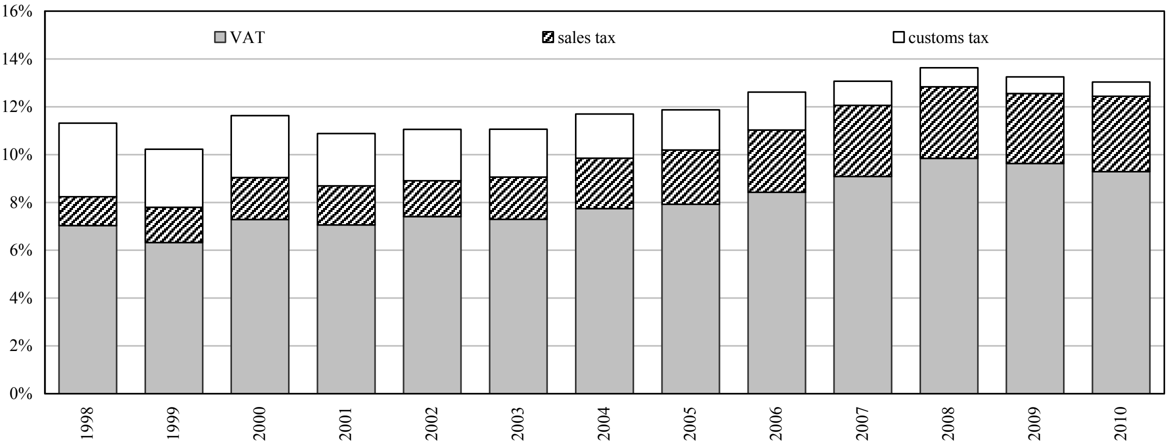
Figure 2 (continued)

Government Revenue Indicators, 1998-2010

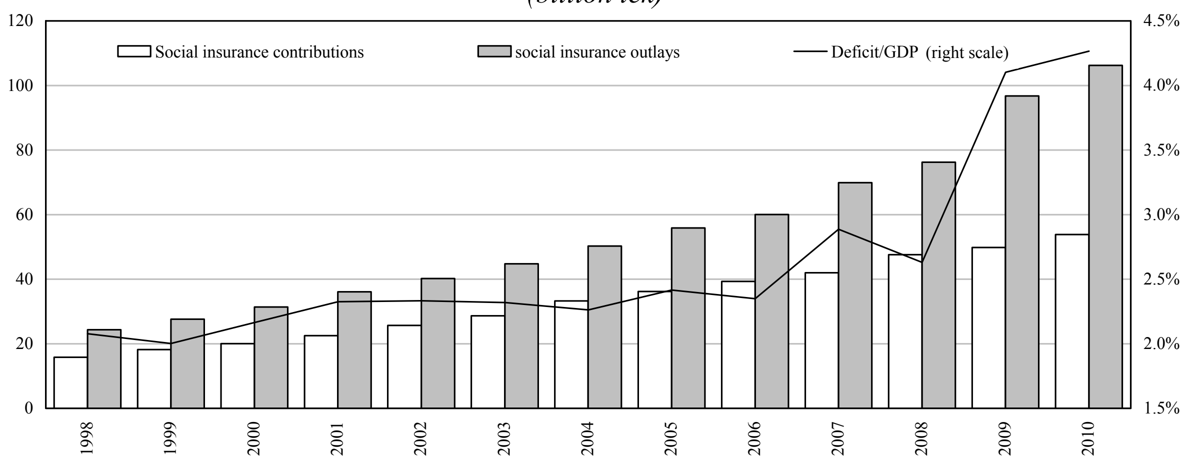
Performance of Direct Tax Burden
(billion lek)



Performance of Indirect Tax Burden



Social Insurance Deficit
(billion lek)



Source: Ministry of Finance.

privatising public-owned companies), expanding the tax base, simplifying and implementing new tax system, promoting tax intensive through reducing tax burden on business, and reducing informality and tax evasion.[2] As a result, budget deficit in 2010 was gradually reduced to 3.2 per cent of GDP from 9.6 per cent in 1998, mainly through cuts in government subsidies, personnel expenditure and interest payments on debt servicing. However, raising budget deficit and public debt during 2007-09 reflected both the action of automatic stabilizers in the form of reduced income and the countercyclical FP through wages and capital expenditure increases.

During the last decade, Albanian tax system also saw major reformations.[3] A series of additional initiatives took place as part of tax legislation changes and were finalised with the approval of a new fiscal package in the second half of 2007. Some of these changes intended to stimulate business incentives and at the same time regenerate more tax revenues. Such reforms consisted of the change from a progressive to a 10 percentage flat income (2007) and profit (2008) tax system and the elimination of all exclusions and facilitations under the old tax system. Besides, there were major reductions in customs duties due to the CEFTA and World Trade Organization membership, the Stabilization and Association agreement with the European Union, etc. Other changes spotted were a considerable rise in national, local and excise tax levels, cuts in social contributions from 42.5 per cent in 2006 to only 17 per cent in 2009, the diminishing of the small business tax to 1.5 per cent in 2006 from 4 per cent in 2005 and a change in the threshold for VAT registration to 5 million ALL turnover per calendar year (2010). All these reforms and structural changes have resulted in a moderated balance growth of government tax revenues, even though increasingly in nominal terms. Indirect taxes such as customs duties, VAT and excise tax are among main indicators of economic activity movements of the country and give the main contribute of tax revenues, reaching round 50 per cent of total level. Profit tax and personal income tax are the main contributors in the group of direct taxes, counting about 13.8 per cent of total revenue in 2010 from only 8 per cent in 1998, even though they are applicable to several categories of income and have been affected by fiscal evasion and non-declaration.

In addition, the public expenditure policies have been focused on promoting sustainable growth and reducing poverty and wealth inequalities. Thus, based on the medium-term fiscal framework (MTFF),[4] a reducing-oriented government expenditure policy aimed at cutting current expenditure to create more funds for strategic capital expenditure identified in the MTFF. As a result, total public expenditure to GDP ratio has shown a declining tendency from 35 per cent in 1998 to approximately 29 per cent in 2010. Current expenditures to GDP ratio have been diminishing, decreasing in 2010 to 24.4 per cent from 28.7 per cent in 1998, even though they capture more than 80 per cent of total expenditure. During this period, personnel (26 per cent), interest payment (18 per cent) and social contribution (27 per cent) represent the highest percentage share of the total current expenditure. Although, FP is oriented to raise wages in the public sector, cuts in personal expenditure are mainly due to reducing the number of employees in the public sector through increasing efficiency and privatisation process and lowering of social contribution expenditure. Interest payments have been diminishing mainly through improvements in government timescale borrowing and cuts in public debt and in interest rates and extending the debt maturity period, followed by considerable raise in social insurance outlays. Further, capital expenditures have on average remained at 6.3 per cent of GDP in the period 1998-2010, even though they have been subject of raise and/or cut based on the Albanian macroeconomic conditions and priorities identified in the MTFF. As such, due to the priorities in infrastructure investment, capital investments reached 8.6 per cent and 8.4 in 2008 and 2009. The distribution of capital

2 See also Shijaku (2009).

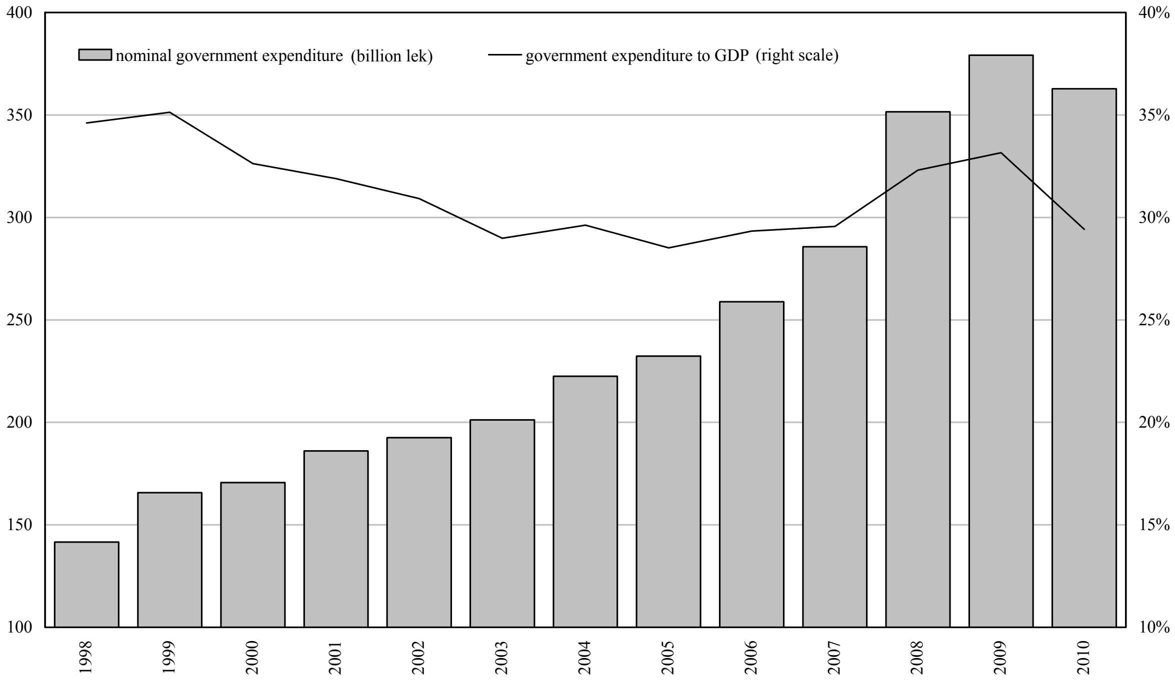
3 Following the introduction of profit (1994) and VAT (1996) tax, the Albanian tax system introduced an income and small and medium business enterprise tax (1998) and customs duties tax (1999).

4 Known also as Medium-Term Economic Program

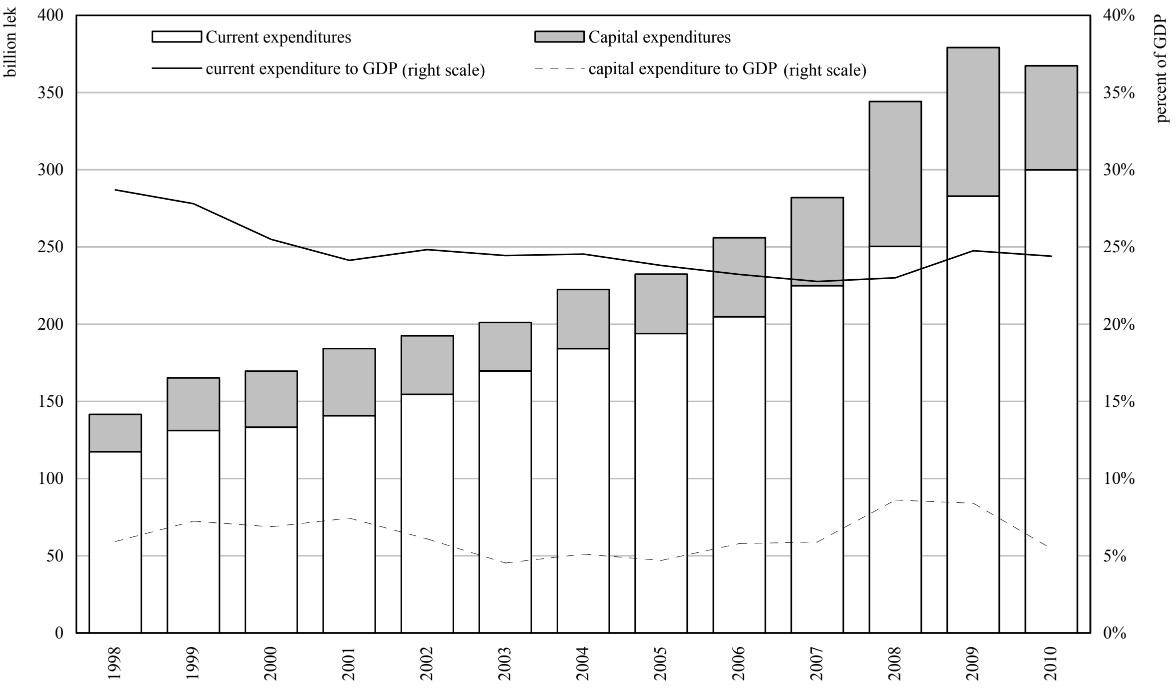
Figure 3

Government Expenditure Indicators, 1998-2010

Government Expenditure Dynamics



Current and Capital Expenditures



Source: Ministry of Finance.

expenditure, in general, was orientated to maintain a relatively high level of spending for areas such as health, education and infrastructure. Mainly these expenditures are financed mostly through domestic borrowing contributing on average by more than 60 per cent.

3 The methodology and data

3.1 Methodology

Neoclassical growth models, based upon the rational expectations assumption, imply that FP can affect only output level but not the long-run growth rate. The steady-state growth rate is driven by the exogenous factors e.g. population growth and technological progress, whilst FP can affect only transition path to this steady state (Judd, 1985). By contrast, under the growth model of Solow, Barro (1990) and Baxter and King (1993) considered a Cobb-Douglas production function and incorporated channels through which FP can determine both the level of the output path and the steady-state growth rate.[5] Instead of only including physical and human capital, the growth rate now depends on the government activity as well, by putting public sector into the production function. To put it formally, we follow Kneller *et al.* (1999) basing the growth model on the following equation:

$$Ø_t = f\left(X_{i_t}, Z_{i_t}\right) \tag{1}$$

or:

$$Ø_t = \alpha + \sum_{i=1}^{k} \beta_i X_{it} + \sum_{j=1}^{m} \gamma_j Z_{jt} + \varepsilon_t \tag{2}$$

where $Ø_t$ is the growth rate of country i at time t, which is a function of conditioning (non-fiscal) variables (X_{it}) based on Solow growth model and fiscal variables (Z_{it}) based on budgetary indicators.[6] Further, α, β_i and γ_j represent the constant term and the slope coefficient of the growth impact of non-fiscal and fiscal variables and $\varepsilon_t \sim iid\ (0, \sigma^2)$ represents the stochastic error term.

Turning to the specification of our model, we build and estimated three variants of endogenous growth model based on identity (2), as follows:

---

5 According to Barro and Sala-i-Martín (1992), output (y) is provided by both private and public sector according to the production function:

$$y = Ak^{1-x} g_x \tag{a}$$

where k represents private capital and g is a publicly provided input. Considering the inter-temporal budget constraint, the government balances its budget in each period by raising a proportional tax on output at rate τ and lump-sum taxes of L, expressed as follows:

$$g + C = L + \tau y \tag{b}$$

where C represents government consumption goods. Taxes on output, in contrast to the lump-sum taxes, will affect private sector incentives to invest in the input goods, such that under such utility function the growth rate will take the form:

$$Ø = \lambda(1-\tau)(1-\alpha)A^{\frac{1}{1-x}}\left(\frac{g}{y}\right)^{\frac{x}{1-x}} - \varepsilon \tag{c}$$

where λ and ε are constant and reflect parameters in the utility function, while the growth rate is decreasing by the rate of (τ) and increasing by the rate of (g). In practice, however government budget is not balanced in every period, so the constraint becomes:

$$g + C + b = L + \tau y \tag{d}$$

where b is budget surplus.

6 Kneller *et al.* (1999) specified a model including investment to GDP ratio, labour force growth rate, net lending, budget surplus, while classified fiscal variables into one of six types. Government revenues are divided into distortionary, non-distortionary and other revenues and government expenditures are classified into productive, non-productive and other expenditures.

$$Ø_t = \alpha + \beta_1^{*}\eta_t + \beta_2^{*}\varphi_t + \beta_3^{*}\mu_t + \beta_4^{*}\tau_t + \beta_5^{*}g_t + \beta_6^{*}\text{debt}_t + \varepsilon_t \tag{3.1}$$

and:

$$Ø_t = \alpha + \beta_1 {*} \eta_t + \beta_2 {*} \varphi_t + \beta_3 {*} \mu_t + \beta_4 {*} \theta_t + \beta_5 {*} \sigma_t + \beta_6 {*} \rho_t + \beta_7 {*} \pi_t + \beta_8 {*} \text{debt}_t + \varepsilon_t \tag{3.2}$$

where $Ø_t$ is Albanian annual real economic growth rate (recongr_yoy); η_t is the fixed gross capital formation[7] (fgcf_ratio); φ_t is the employment annual growth rate (empgr_yoy); μ_t is a proxy for trade openness index measured as the sum of total import + exports to nominal GDP ratio (opentb_ratio); τ_t and g_t represent fiscal indicators and stands for government revenues excluding grants (rev_ratio) and expenditure (exp_ratio); θ_t and σ_t represents revenue counterpart sub-categories, standing for the distorsionary (*disrev_ratio*) and non-distortionary (*nddrev_ratio*) revenues; ρ_t and π_t stand for the expenditure counterpart sub-categories, representing productive (pexp_ratio) and non-productive (npexp_ratio) expenditure; $debt_t$ represents the ratio of public debt to nominal GDP (*debt_ratio*).

From a theoretical point of view, physical and human capitals are the main factors of production in the growth model of Solow. Thus, fixed gross capital formation to GDP ratio (fgcf_ratio) and employment annual growth rate (*EMPGR)* entered the model as explanatory variables. Besides, *EMPGR* controls for business cycle effects on growth (Benos, 2009). Regarding other non-fiscal variables, we used the sum of imports and exports as a proportion of GDP (opentb_ratio), to account for external effects on the economic growth. Regarding fiscal variables, accordingly, we considered some notable exceptions when modelling endogenous fiscal-growth relationship. First, a model suffers from substantial bias coefficients estimation if both sides of budget are not taken into account, given that FP affects output through taxation and expenditures policies (Kneller *et al.*, 1999). Thus, in our model the fiscal variables encounter to capture full effects of FP by entering into the model both government revenues and expenditures indicators. Second, Kneller *et al.* (1999) and Benos (2009) finds out that some types of government expenditures and taxation can be either growth-enhancing or reducing. Hence, following Barro and Sala-i-Martín (2004), the public revenues were categorised into distortionary (disrev_ratio), non-distortionary (nddrev_ratio) and other public revenues (orev_ratio), whislt public expenditure were categorised into productive (pexp_ratio), non-productive (npexp_ratio) and other public expenditures, (oexp_ratio).

Additionally, according to Kneller *et al.* (1999), if budget constraint is fully specified, so that:

$$\sum_{j=1}^{m} X_{jt} = 0$$

One element of Z must be omitted in the estimation of equation (2) in order to avoid perfect colinearity. In other words, this exclusion also offers a proper way to interpret any changes in fiscal variables included in the model. As such, we omitted the variables of other revenue and expenditure from our model, given their relatively size and impact on economic growth and the critical value of the *F*-test based on an omitted variables test and correlation test (Table 3). Finally, empirical models of FP may suffer from bias estimation if they do not impose debt indicators (Favero and Giavazzi, 2007). But, the debt financing methods can affect both the supply and demand side of the economy (Klalid *et al.*, 2007). Besides, as it increases, indebtedness can turn from initially growth-enhancing (or neutral) to eventually growth reducing (Cecchetti *et al.*, 2011). Thus, we have also included in our model public debt to GDP ratio to examine potential effects of the level of indebtedness on growth and to distinguish whether debt is growth-enhancing or reducing.

---

7 Refer also as total capital investments. We also specify the growth model using as proxy the private investment to GDP ratio and found the same results.

The endogenous fiscal-growth model does not place restrictions on the sign of the coefficients. But a negative sign (–) represents a negative impact on growth and vice versa. Kneller *et al.* (1999) suggested that increasing burden of taxation weakens the incentives to invest, hence reducing growth. Government expenditures influence the marginal product of private capital through increase consumption goods and services, henceforth boost growth. Amanja and Morrissey (2006) imply that taxation and expenditure policies can harm or promote growth. A tax system that causes distortions to private agents' investment incentives can retard investment and growth. Analogously, if the system is such that it leads to internationalization of externalities by private agents, it may induce efficiency in resource allocation and thus foster investment and growth. The same applies with the nature of government expenditure, where excessive current expenditure at the expense of investment is likely to discourage growth and vice versa.

In addition, some types of government expenditures and taxation can be either growth-enhancing or reducing. We expect that distortionary taxation weakens the incentives to invest in physical/human capital, hence reducing growth. Benos (2009) reveals that non-distortionary taxation does not affect the above incentives, therefore growth, due to the nature of the utility function assumed for the private agents. However, we would expect that raising non-distortionary taxation would affect production through increasing marginal costs whether tax is levied on producers or consumers. Therefore, if tax is levied on producers it reduces the marginal return to private capital and if it is levied on consumers it effects the incentives to consume more, hence harming growth. Further, an augmenting productive spending financed by non-distortionary taxes will boost growth. But, this effect is ambiguous if distortionary taxation is used. In the latter case, there is a growth-maximizing level of productive expenditure, which may or may not be Pareto efficient (Irmen-Kuehnel, 2008). Rising also non-productive spending financed by non-distortionary taxes will be neutral for growth. But, if distortionary taxes are used the impact on growth will be negative. Besides, if non-productive expenditure serves as means to create consumption based expenditure, then an increase will boost growth.

Finally, as Cecchetti *et al.* (2011) puts forward, the impact of debt burden to growth is ambiguous, given that raising indebtedness can turn from initially growth-enhancing (or neutral) to eventually growth reducing. Public debt burden can smooth consumption not only through lifetime, but also across generations, by providing more human capital and productive technology as long as they are not constrained by macroeconomic instability, distorted policies and institutional weaknesses. It can also provide liquidity services and increase financial intermediation, which can contribute to easing the credit conditions faced by firms and households, thus crowding in private investment and helping growth. Above a certain threshold, however, debt is found to reduce growth as rising indebtedness, including its domestic component, above a country's repayment ability would discourage private investment due to the expectation of higher future taxes (Blavy, 2006). Several types of risk factors related to rising debt would account on raising domestic interest rates, crowding out public investment within the budget and private investment in general, a rowing portion of savings would go towards purchases of government debt, rather than capital investments and higher marginal tax rates may be used to pay rising interest cost, leading to reducing of saving rates and discouraged work. This may harm the economic growth.

In the specified models, we also assumed that there exist some strong potential for endogeneity of the fiscal and debt variables, especially reverse causation (low or negative growth rates are likely to induce higher expenditure–revenues and debt burdens).[8] The models, hence, are estimated by Generalised Moments of Movements (GMM). GMM approach allows the usage of instrumental variables regression to deal with a situation where some of the right-hand side (RHS)

[8] While the economic growth rate is likely to have a linear negative impact on the public debt-to-GDP ratio, high levels of public debt are also likely to be deleterious for growth.

variables are correlated with disturbances due to endogeneity problems.[9] The idea behind instrumental variables is to find a set of variables, termed instruments, which are both correlated with the explanatory variables in the equation and eliminate the correlation between RHS variables and the disturbances. For the GMM estimator to be identified, there must be at least as many instrumental variables as there are parameters to estimate. As such, RHS with four lags are used as the relevant instrumental variables in our GMM models, given also that empirical evidence[10] suggests that there are lagged effects of fiscal and non-fiscal policy on growth. In models for which there are more moment conditions than model parameters, GMM estimation provides a straightforward way to test the specification of the proposed model through the *J*-statistic hypothesis test. A simple application of the *J*-statistic is to test the validity of overidentifying restrictions, under the null hypothesis that the overidentifying restrictions are satisfied.

3.2 Data

The paper considers quarterly data from 1998Q01 to 2010Q04, but we also tried to evaluate the effect of FP prior to the effects of fiscal expansion and reforms after 2007 and also prior to the effects of the recent financial and economic crisis that affected the economic activity in Albania. Thus, we tried to empirically identify the effect of FP throughout different time-samples, mainly 1998-2006 and 1998-2010. The economic growth model is based on capital, labour, trade openness and fiscal variables. The data on fixed gross capital formation, real economic growth and employment rate are taken from the Albanian Institute of Statistics (INSTAT). Quarterly FGCF is interpolated from annual data by linear match last approach using E-views. The series on FGCF and private investment are extended to 2010Q04 by an Autoregressive Integrated Moving Average (ARIMA) forecast process.[11] The data on exports and imports of goods and services are taken from Bank of Albania.

Government expenditure represents the total level and government revenues do not include grants since the later are donations and do not account for the state of the Albanian economic activity.[12] As noted above, within the class of endogenous growth models relevant to this study, results are driven by classification of fiscal variables into different types and a key issue is the allocation of taxes and expenditures, respectively, to distortionary vs. non-distortionary revenues and productive vs. non-productive expenditures. Distortionary government revenue is the sum of profit tax + personal income tax + national taxes and others + revenues from local government + social insurance contributions. Non-distortionary government revenue is the sum of Custom Duties + VAT + Excise Tax. Disaggregation of expenditure relates to the classification of the public expenditures based on budgetary indicators as an alternative solution to the unavailability of the appropriate time series for the public expenditures as in Barro and Sala-i-Martín (2004). Therefore, productive government expenditure is the sum of public capital expenditures. Non-productive government expenditure is the sum of personnel expenditure + subsidies + social insurance outlays + operational & maintenance + other expenditures + electricity compensation + compensation for expropriation + interest cost of bank restructuring + loans to KESH + payment for participation in

---

9 Ordinary Least Square (OLS) and weighted LS (WLS) are biased and inconsistent if right-hand side variables are correlated with the disturbance term.

10 See Amanja and Morrissey (2005) and Burger (2011).

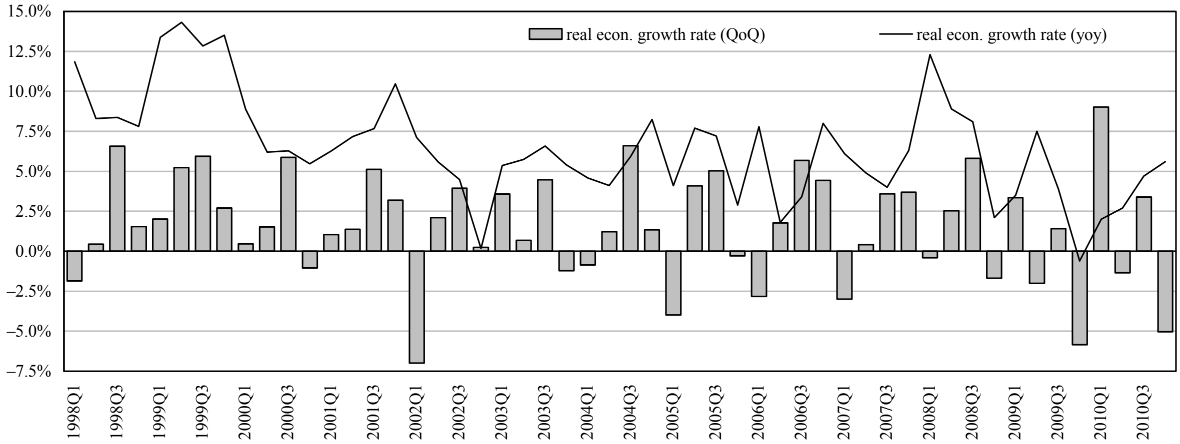
11 The Albanian Institute of Statistics (INSTAT), which produces the official country statistics, has only annual data from 1996 to 2008, which can limit the purpose of this study. Kota (2007) has used the real economic growth rate as a benchmark to generate the data on FGCF for the period 2008-10.

12 We also specify the growth model using as proxy the total government revenue to GDP ratio and found the same results.

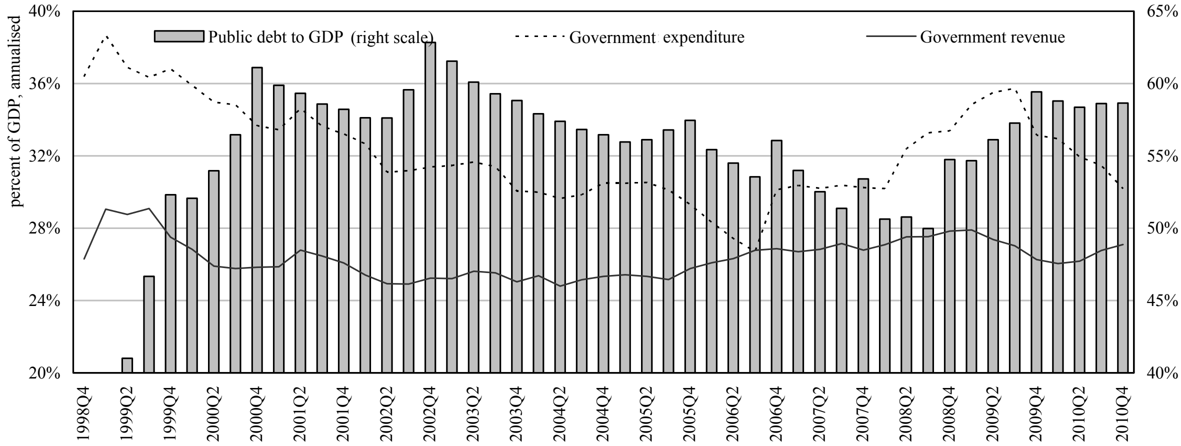
Figure 4

Economic Growth and Explanatory Variables

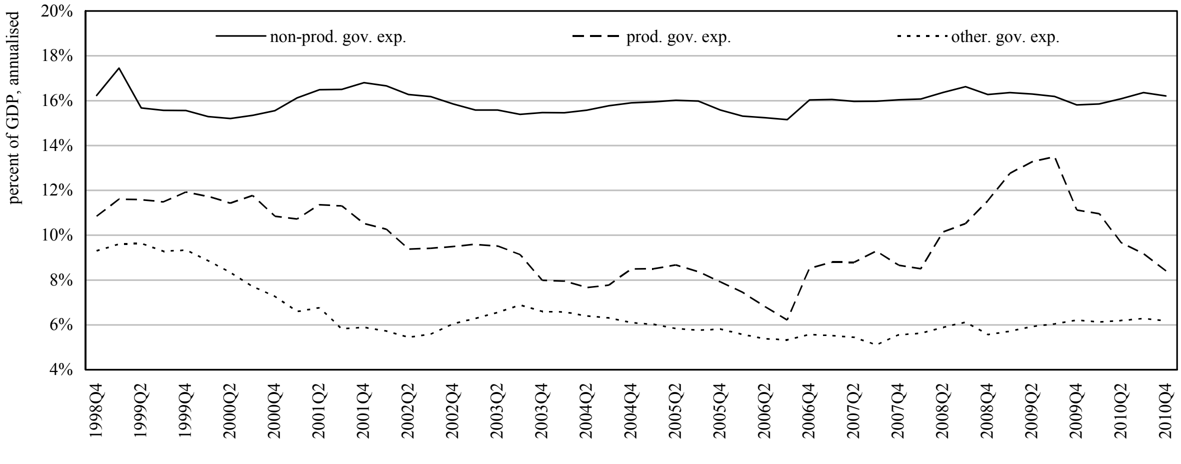
Real Economy Growth Rate



Fiscal Indicators



Government Expenditures

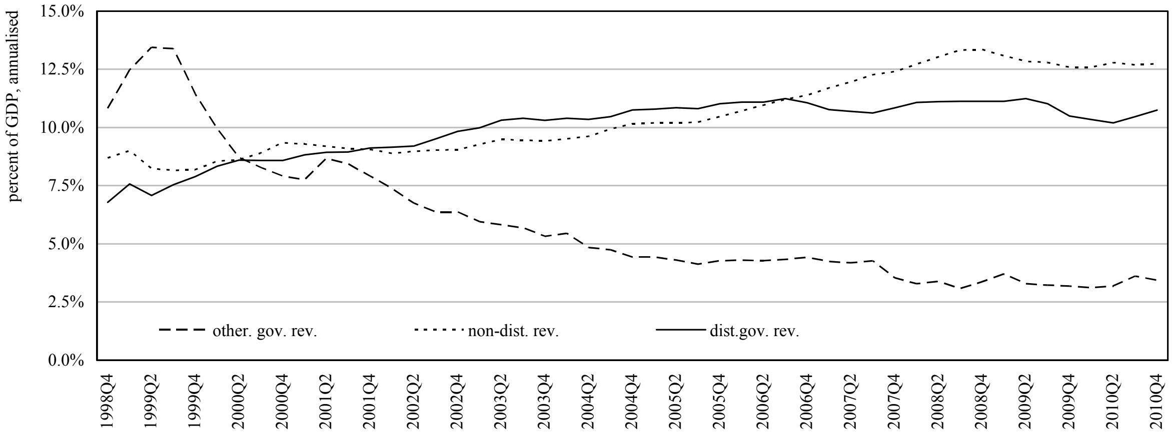


Source: Bank of Albania, Ministry of Finance and INSTAT.

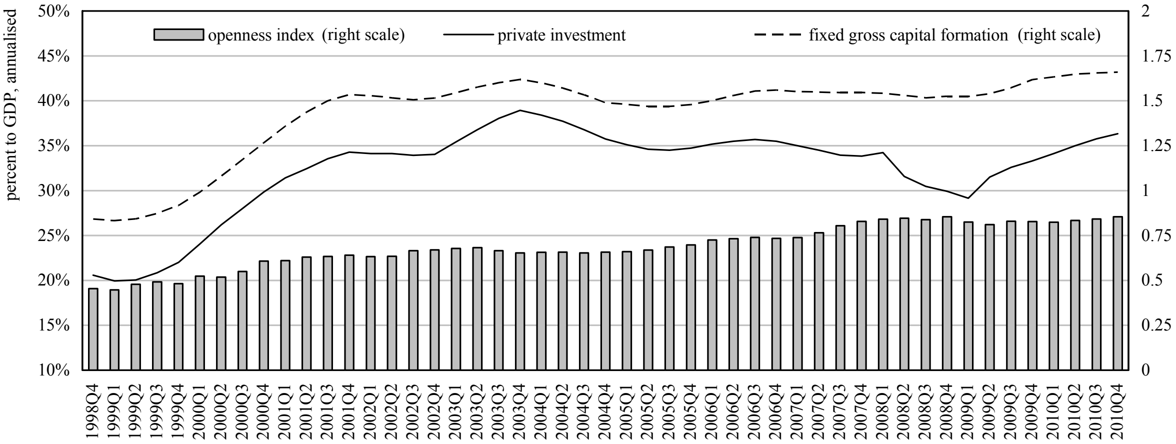
Figure 4 (continued)

Economic Growth and Explanatory Variables

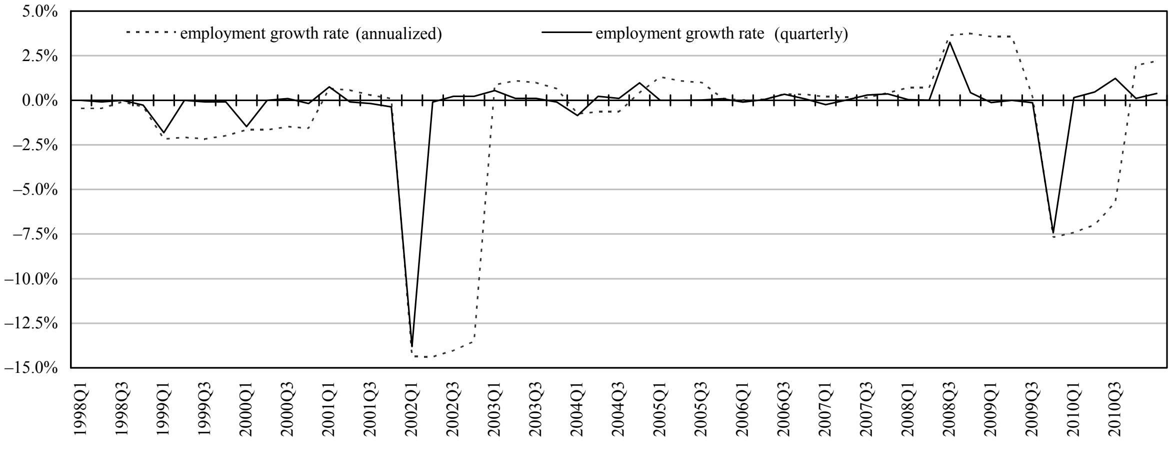
Revenues



Openness Index, Fixed Gross Capital Formation and Private Investment



Employment Growth Rate



Source: Bank of Albania, Ministry of Finance and INSTAT.

BISH capital + energy support. Fiscal data and the public debt are taken from the Ministry of Finance. Data, besides economic and employment growth rate, are generated as a ratio of GDP.[13]

4 Empirical results

Table 1 summarises the results according to the GMM techniques. Coefficients on models (A) of the table is based on the sample time: 1998:01-2010:04 estimation, models (C) add the effects of dummy variables on fiscal reforms, expansionary and effects of financial and economic crisis and model (B) estimate the relationship before these phenomena, respectively the sample time: 1998:01-2006:04. After conducting Augmented Dickey Fuller and Philips Perron unit root tests (Table 2) we find conclusive evidence only on the non-stationary of non-distortionary of government revenue. Hence, ndrev_ratio entered the model in first difference. The results on GMM specification are also based on model diagnostic tests (Table 1). The statistical value of the regression determination coefficient (R^2) and a set of diagnostic tests conducted on the model specification reveal no problems with respect to serial correlation (*Q*-statistic and Squared Residuals) and Hausman test on over-identification of the instrumental variables (*J*-statistic and Coefficient of over-ID and Prob.).

Empirical results in Table 1 demonstrate that the value of the coefficients is statistically significant at conventional levels, regarding the non-fiscal. Employment growth and fixed gross capital formation[14] are estimated to have a positive effect on production growth, confirming the prediction of endogenous growth theory. These are expected since labour and capital are factors of production in most growth models and they support the endogenous growth models. Furthermore, as far as economic openness is concerned, it effects growth negatively.

Regarding the fiscal variables, results (Table 1) show that their effect on growth rate is statistically significant at conventional levels. Growth is affected negatively by government revenues and positively by expenditure policies. Government revenues effected growth more than the government expenditure, given the coefficient size for the estimated sample. This follows the same conclusions drawn by Mançellari (2011). Hence, raising rev_ratio by 1 percent will reduce growth by round .476 per cent and boosting exp_ratio by 1 percent stimulates growth by round .146 per cent. This would re-enforce a theory already expressed by Barro and Sala-i-Martín (1992) that revenues effect growth negatively and expenditure enhance growth. Under Barro and Sala-i-Martín (2004) fiscal decomposition, we found that revenue sub-categories reduce growth, but distortionary taxation has much larger and statistically significant effect. Growth rate will diminish by .6374 per cent in response of 1 percent increase in disrev_ratio (distorsionary revenues) and by round .128 percentage points in response of 1 percent raise in ndisrev_ratio (non-distorsionary revenues). On the other hand, the empirical results show that growth is positively affected by productive expenditure and negatively by non-productive. Productive expenditure has a much higher effect on growth than non-productive expenditure. Raise in pexp_ratio by 1 percent will boost growth positively by round .460 and a 1 percentage point decrease in npexp_ratio will improve growth by more than .272 percentage points. Based on the value of the coefficients, productive expenditures have a larger impact on growth than non-productive expenditure. Based on equation (3.1A), raising any type of revenues or decreasing expenditure by government bring along negative effects in economic growth, but it clearly matters what type of revenue to rise and what type of expenditure to decrease in order to improve the budget balance and at the same time achieve the best results on GDP growth. As such, based on the

---

13 See Afonso and Jales (2011).

14 We also specify the growth model using as a proxy the private investment to GDP ratio (invtprv_ratio) and found relatively the same results.

Table 1

Results of Macroeconomics and Fiscal Indicators on Real Economic Growth Rate (recongr_yoy) Based on GMM Specification Techniques(1)

	Equation (3.1)		Equation (3.2)	
	(A)	(B)	(A)	(B)
C	.415289*	.286233*	.528287*	.482663*
FGCF_RATIO	.183665	.727787	.699920*	.854009*
EMPGR_YOY	.130057*	.038906	.320074*	.027685*
OPENTB_RATIO	–.468705*	–.975659*	–.625912*	–.697641*
	FISCAL VARIABLES			
REV_RATIO	–.476155*	–.432053*		
EXP_RATIO	.146148*	.278665*		
DISREV_RATIO			–.637372*	–.610666*
NDREV_RATIO			–.127742***	–.108589*
PEXP_RATIO			.460154*	.344109*
NPEXP_RATIO			–.271892*	–.104717*
DEBT_RATIO	–.373513*	–.178836*	–.586688*	–.566797*
	DIAGNOSTIC TESTS			
J-static	.198145	.158028	.227737	.222982
Coef. OverID	9.3128	5.2149	10.7036	6.9124
(Prob.)	.9520	.7343	.9986	.9969

Sample Time: (A) – (1998q01 – 2010q04); (B) – (1998q01 – 2006q04).
(1) – variables on the RHS are used as instrumental variables.
Based on: * (1 per cent), ** (5 per cent), *** (10 per cent) level of significance.

value of the coefficients in Table 1 (equations (3.2A) and (3.2B)), if government wishes to boost budget revenue it should choose indirect taxes instead of direct taxes as raising this category has slightly less negative effects on growth. On the other hand, coefficients value suggests that if government wishes to reduce fiscal deficit through expenditure cuts policies it should consider non-productive rather productive expenditure cuts, as the former has a negative effect on economic growth.

Results imply that revenues have a higher negative effect on growth, compared to the estimated coefficient value for the period 1998-2006. The impact of revenue on growth has increased from .432 prior to 2007:01 to round .476 for the whole sample. Results demonstrate that amplifying negative impact is mostly due to extending effects through distortionary taxation policies. Their negative effect on growth has increased by round .0267 points compared to only round .192 points raise in non-distortionary negative impact. However, the impact of expenditure on growth is weaker compared to the estimated coefficient value for the period 1998-2006. The positive impact of expenditure on growth has shrunk to only .146 points compared to .279 it was prior to 2007, given the size of the coefficient. Considering the sub-categories of government expenditure, results imply that productive and non-productive expenditure have a higher respectively effect on growth after 2006. These reflect the attitude of the counter-cyclical FP through capital and wages increase. This proves that rising capital expenditure has provided bigger positive impact on growth and has also mitigated the negative affects that global financial and economic crisis had on the Albanian economy. This confirms findings by Bachmann and Sims (2011) that raising government investments, especially during downturns, boost business confidence. The positive effect of pexp_ratio on growth has increased by round .116 percentage points and the negative effect on npexp_ratio has gone up by round .167 per cent. First, these implying effects reflect mainly the attitude of the counter-cyclical FP through capital and wages raise in the period 2007-09. This, as Afonso (2006) puts forward, reveals the Albanian public sector efficiency on resource allocation and output scores maximisation. Second, the diminishing impact of expenditure on growth is mostly due to raising negative impact of non-productive expenditures.

Further, findings show that the coefficient on debt ratio is statistically significant at conventional levels and negatively related to growth rate. This effect is even greater compared to the estimated coefficient value for the period 1998-2006. This, according to Cecchetti *et al.* (2011), suggests that debt burden is above a threshold of growth-enhancing. Hence, raising debt burden reduce growth. According to results by Shijaku (2011) in the verge of raising cost of borrowing a further increase above Albanian repayment ability or sustainability level would discourage public investment within the budget structure and may crowd-out private investments. In addition, given the magnitude of the coefficients, raising debt ratio to finance capital public investment would crowds out the effects of productive expenditure. Instead, if government wishes to stimulate economic activity through boosting productive expenditure, it should do it through lowering the non-productive expenditure rather than borrowing instruments.

5 Conclusion

Albania FP has been under continuous scrutiny of major reformation on expenditure and tax collection system. The philosophy of these fiscal reforms was based on the idea of reducing current expenditures and boosting government revenues. The Albanian economy took advantages of macroeconomic stimulus in the form of fiscal expansion ahead of monetary adjustments, during the financial and global crisis. Raising budget deficit and public debt reflected both the action of automatic stabilizers in the form of reduced income and the countercyclical FP through wages and capital expenditure increases and also the cost of fiscal burden as a result of government decision to stimulate the economy, while fiscal incentives were narrowing.

This discussion material analysis the Albanian FP effects upon economic growth based on an endogenous fiscal-growth model. The aim of this paper is not to resolve the raging debate on the ability of FP to affect economic growth, but to examine the case of a small open developing country, Albania. Regarding fiscal variables, the results obtained show that overall growth rate is affected negatively by government revenues and positively by expenditure policies. Considering the parameter magnitude government revenue effected growth more than government expenditure. Categorising tax revenues into distortionary and non-distortionary, we found that government revenues and the sub-categories reduce growth, but distortionary taxation has much larger and significant effect. Further, growth is positively affected by productive expenditure and negatively by non-productive, but the former has a greater impact.

Additionally, based on the coefficient value, empirical results suggest that since 2007 expenditure-growth relationship is weaker, while revenues have a higher negative impact on growth. Results demonstrate that rising revenues negative impact is mainly due to distortionary policies. Expenditure policies reflect the attitude of the counter-cyclical FP through capital and wages increase. Further, findings show that the coefficient value of debt burden is negatively related to growth rate. This effect is statistically significant. This impact is even greater since 2007. Financing government capital investment through borrowing mechanism has stimulated growth, but according to Cecchetti *et al.* (2011) debt burden is above a threshold of growth-enhancing.

APPENDIX

Table 2

Unit Root Tests

Variable	ADF		Phillips-Perron		ADF		Phillips-Perron		ADF		Phillips-Perron	
	Null Hypothesis: Unit Root											
	Level	First Difference	Level	First Difference	Level	First Difference	Level	First Difference	Level	First Difference	Level	First Difference
	(Prob.)											
	Intercept				Intercept and Trend				None			
recongr_yoy	(.0049)	(.0000)	(.0049)	(.0000)	(.0040)	(.0000)	(.0048)	(.0000)	(.0617)	(.0000)	(.0985)	(.0000)
fgcf_ratio	(.0462)	(.0174)	(.3230)	(.0112)	(.2111)	(.0394)	(.7667)	(.0299)	(.8853)	(.0028)	(.9308)	(.0017)
empgr_yoy	(.0000)	(.0000)	(.0000)	(.0001)	(.0000)	(.0000)	(.0000)	(.0001)	(.0000)	(.0000)	(.0000)	(.0001)
opentb_ratio	(.6175)	(.0000)	(.0916)	(.0001)	(.0569)	(.0000)	(.0000)	(.0001)	(.9974)	(.0118)	(.9539)	(.0000)
exp_ratio	(.3667)	(.0001)	(.0000)	(.0001)	(.7335)	(.0000)	(.0000)	(.0001)	(.2894)	(.0000)	(.4514)	(.0000)
npexp_ratio	(.0250)	(.0001)	(.0000)	(.0001)	(.0486)	(.0000)	(.0001)	(.0001)	(.3317)	(.0000)	(.6533)	(.0000)
pexp_ratio	(.4958)	(.0000)	(.0000)	(.0000)	(.8138)	(.0000)	(.0000)	(.0000)	(.5027)	(.0000)	(.0008)	(.0000)
rev_ratio	(.6750)	(.0000)	(.0000)	(.0001)	(.0572)	(.0000)	(.0000)	(.0001)	(.8572)	(.0000)	(.9083)	(.0000)
ndrev_ratio	(.9129)	(.0007)	(.9235)	(.0001)	(.3712)	(.0041)	(.5217)	(.0006)	(.9748)	(.0000)	(.9921)	(.0000)
disrev_ratio	(.0863)	(.0000)	(.0002)	(.0000)	(.0000)	(.0000)	(.0000)	(.0000)	(.9688)	(.0000)	(.8706)	(.0000)
debt_ratio	(.0001)	(.0000)	(.0933)	(.0000)	(.0014)	(.0000)	(.4996)	(.0000)	(.7258)	(.0000)	(.9138)	(.0000)

[a] Automatic lag selection based on Schwarz Info Criterion (SIC).

Table 3

Estimated Results on Redundant Variables Test

Redundant Variables	Null Hypothesis: The Variable is Not Significant for the Model			
	F-statistic	Prob. *F*-statistic	Log Likelihood Ratio	Prob. Chi-square
DISREV_RATIO	1.387526	(0.2460)	1.782925	(0.1818)
NDREV_RATIO	4.385346	(0.0428)	5.434565	(0.0197)
OREV_RATIO	1.230876	(0.2740)	1.584729	(0.2081)
PEXP_RATIO	0.613775	(0.4381)	0.796378	(0.3722)
NPEXP_RATIO	0.245582	(0.6230)	0.320139	(0.5715)
OEXP_RATIO	0.639366	(0.4288)	0.829314	(0.3625)
DEBT_RATIO	2.540964	(0.1190)	3.219037	(0.0728)

Synthesis of results generated using E-views 6.

REFERENCES

Afonso, A. and G. Alegre (2008), "Economic Growth and Budgetary Components – A Panel Assessment for the EU", European Central Bank, Working Paper, No. 848, January.

Afonso, A. and J. Jallers (2011), "Economic Growth and Government Size", European Central Bank, Working Paper, No. 1399, November.

Afonso, A., L. Schuknecht and V. Tanzi (2006), "Public Sector Efficiency Evidence for New EU Member States and Emerging Markets", European Central Bank, Working Paper, No. 581, January.

Baldacci, E., B. Clements and S. Gupta (2003), "When Reducing Fiscal Deficits Makes Sense for Low-income Countries", *Finance & Development*, December.

Bachmann, R. and E. Sims (2011), "Confidence and the Transmission of Government Spending Shocks", NBER, Working Paper, No. w17063.

Bank of Albania, (2010), *Bulletin*, 2010H2.

Barro, R. (1979), "On the Determination of Public Debt", *Journal of Political Economy*, Vol. 87, No. 5, pp. 940-71.

——————(1990), "Government Spending in a Simple Model of Endogenous Growth – Part 2: The Problem of Development: A Conference of the Institute for the Study of Free Enterprise Systems", *Journal of Political Economy*, Vol. 98, No. 5, pp. S103-S125, October.

Barro, R. and X. Sala-i-Martín (2004), *Economic Growth*, 2nd edition, MIT Press, Cambridge (MA).

——————(1992), "Public Finance in Models of Economic Growth", *The Review of Economic Studies*, Vol. 59, No. 4, pp. 645-61, October.

Baxter, M. and R. King (1993), "Fiscal Policy in General Equilibrium", *American Economic Review*, Vol. 83, No. 3, pp. 315-34.

Benos, N. (2009), "Fiscal Policy and Economic Growth: Empirical Evidence from EU Countries", MPRA Paper, No. 19174, posted on 11 December.

Blanchard, O. and R. Perotti (2002), "An Empirical Characterisation of the Dynamic Effects of Changes in Government Spending and Taxes on Output", *The Quarterly Journal of Economics*, pp. 1329-68, November.

Blavy, R. (2006), "Public Debt and Productivity: The Difficult Quest for Growth in Jamaica", International Monetary Fund, Working Paper, No. WP/06/235.

Cecchetti, G.S., M.S. Mohanty and F. Zampolli (2011), "The Real Effects of Debt", Bank for International Settlements, Working Paper, No. 352.

Frankel, J. and D. Romer (1999), "Does Trade Cause Growth?", *American Economic Review*, Vol. 89, No. 3, pp. 379-99.

Irmen, A. and J. Kuehnel (2008), "Productive Government Expenditure and Economic Growth", University of Heidelberg, Discussion Paper, No. 464.

Judd, K. (1985), "On the Performance of Patents", *Econometrica*, Vol. 53, No. 3, pp. 567-86.

Kneller, R., M.F. Bleaney and N. Gemmell (1999), "Fiscal Policy and Growth: Evidence from OECD Countries", *Journal of Public Economics*, No. 74, pp. 171-90.

Kota, V. (2007), "Alternative Methods of Estimating Potential Output in Albania", Bank of Albania, Research Department.

Mançellari, A. (2011), "Macroeconomic Effects of Fiscal Policy in Albania: A SVAR Approach", Bank of Albania.

Shijaku, G. (2009), "The Development and Adoption of Fiscal Policy in Albania", Bank of Albania, Research Department, unpublished, August.

—————(2011), "Sustainability of Fiscal Policy: The Case of Albania", Bank of Albania, Research Department, unpublished.

Socol, C. and G.A. Socol (2009), "The Analysis of Fiscal Policy Management in Romania: Lessons for Emerging Countries", *African Journal of Business Management*, Vol. 3, No. 5, pp. 240-47, May.

World Bank (2006), *Albania Restructuring Public Expenditure to Sustain Growth – A Public Expenditure and Institutional Review*, December.

COMMENTS ON SESSION 2
GOVERNMENT BUDGETS AND POTENTIAL GROWTH

*John Janssen**

I would like to thank Daniele and his team for the invitation to the workshop and the opportunity to comment on two interesting papers in this session. Although New Zealand's public debt levels are relatively low (albeit with relatively high levels of private sector debt), projections suggest that under existing policy, debt-to-GDP ratios are likely to rise (Buckle and Cruickshank, 2012). Hence the possible effects of higher public debt on economic growth are of interest.

1 Comments on "Debt and Growth: New Evidence for the Euro area" by Anja Baum, Cristina Checherita-Westphal and Philipp Rother

The focus of this paper is on the short-term, non-linear impact of public debt on GDP growth in the Euro area. Non-linear effects are captured via the use of a threshold regression model, where the threshold distinguishes the two regimes where the behaviour predicted by the model differs.

In terms of methodology, the paper contributes to the literature by extending the non-dynamic threshold panel methodology of Hansen (1999) to a dynamic setting (Caner and Hansen, 2004). The dynamic effects are captured by adding lagged GDP growth rates to the regression. The endogenous variable is the real GDP growth rate, and control variables include: lagged real GDP growth; openness; the investment-to-GDP ratio; and a dummy variable for EMU entry. Estimation uses annual data for 12 Euro area countries over the period 1980 to 2010.

An important part of the estimation involves finding the threshold debt ratio that divides the sample into two different regimes. The dynamic model is estimated with 2Stage Least Squares (2SLS) for each possible value of the threshold variable, and the corresponding sum of squared residuals (SSR) are calculated. The selected threshold value is the one that gives the smallest SSR. Based on this estimate, the slope parameters are estimated using Generalized Method of Moments (GMM). The results are reported both for the non-dynamic and dynamic panels. The possibility of more than one threshold value (*i.e.*, more than two regimes) is found to be insignificant.

In terms of the results, the short-run impact of additional debt is positive and highly significant at debt-to-GDP ratios below 67 per cent for the benchmark case (1980 to 2007). The impact reduces to zero if debt-to-GDP is above the threshold. A longer sample period, up to 2010, changes the dynamic panel results. The short-run impact of additional debt estimated by the dynamic panel is positive and is highly significant at debt-to-GDP ratio levels below 95.6 per cent. Additional debt has a negative impact on economic activity for high debt-to-GDP ratios above 95 per cent and is statistically significant.

The paper argues that the transmission channel behind the results works through long term interest rates and higher sovereign risk premia. Market sensitivity to debt-related news has recently increased in the Euro area. Therefore, an increase in debt levels today may raise concerns about debt sustainability and signal a tighter fiscal policy in the near future. This is likely to dampen the positive stimulus effects of policy that is the dominating factor behind the results. Therefore, it is

* New Zealand Treasury, P.O. Box 3724, Wellington, New Zealand.
E-mail: john.janssen@treasury.govt.nz
The views expressed are those of the author and do not necessarily reflect the views of the New Zealand Treasury.
Thanks to Tugrul Vehbi of the New Zealand Treasury for assistance in preparing these comments.

also important to understand the sources of debt increase. It could make a difference if the additional debt is simply for financing consumption spending versus productive investment.

2 Comments on "Public Debt and Growth" by Manmohan Kumar and Jaejoon Woo

The focus of this paper is on the relationship between high public debt and long-run economic growth. The paper provides further analysis of the findings of Reinhart and Rogoff (2009, 2010) and addresses several of the perceived shortcomings in that work. The contributions along these lines include: the treatment of the endogeneity by using the approach of Arellano and Bover (1995); using the initial level of debt to avoid the reverse causality problem; using an extensive set of regressors to control for the effects of other determinants of growth; and the use of extensive statistical techniques to validate the results.

The estimation starts with a baseline panel of 38 advanced and emerging economies, covering the period 1970 to 2008 and employing a variety of estimation techniques. The rationale for using different methods is based on the fact that different methods involve different tradeoffs (e.g., measurement error versus omitted variable bias). Alternative time period and country coverage are also considered.

The paper also attempts to determine the channels through which debt affects economic growth by considering its effects in a growth accounting framework (*i.e.*, total factor productivity and growth of output and capital stock per worker). The main result is that a 10 per cent increase in the initial debt-to-GDP ratio reduces the subsequent growth rate by 0.2 per cent per year.

The transmission channel is through a slowdown in labour productivity growth due to reduced investment and slower growth of the capital stock. The paper finds evidence of non-linearity, with higher levels of initial debt-to-GDP (>90%) having a proportionately larger negative effect on subsequent growth. Results appear to be robust to different estimation methods with the exception of the fixed effect estimator, where the debt-to-GDP ratio is insignificant.

The fiscal deficit variable is also found to be highly significant in affecting growth rates. Although removing it and other variables in alternative parsimonious specifications still yields an overall negative relationship. This suggests that both deficits and debt matter for growth. It would be interesting to test the results using net debt instead of gross debt. Although data may not be available for the majority of the countries included in the sample, it might be useful to test the validity of the results for a number of countries where net debt data is available. The selection of the thresholds (*i.e.*, low, medium, high) seems somewhat *ad hoc* – what is the rationale for choosing them? Finally, it would also be interesting to assess the sensitivity of the results given the post 2008 experience. Possible extensions to the paper could include the link to external net liabilities and the maturity structure of public debt (elements of these were included in the tabled version of the paper).

3 General comments

Both of the papers focus on relatively narrow aspects of public sector balance sheets, namely public debt. An important lesson from New Zealand's on-going publication of balance sheet information, including the recently published *Investment Statement,* is the insight that can be gained from decomposition. Table 1 lists some of the New Zealand Government's key balance sheet indicators, together with the net positions of portfolio groupings based upon financial, commercial and social objectives.

Table 1

New Zealand Government Balance Sheet Indicators and Portfolios
(billions of NZD dollars, year ended June 2011)

Indicator *(debt reported as +)*		**Portfolio** *(assets – liabilities)*	
Gross debt	72.4	Financial	(47.2)
Net debt	40.1	Commercial	30.9
Net debt including NZSF	24.0	Social	97.2
Net worth	80.9	**Sum = net worth**	**80.9**

Note: Gross debt is gross sovereign-issued debt excluding central bank settlement cash and bills. Net debt is for the core Crown. The NZSF is the New Zealand Superannuation Fund, an entity designed to partially pre-fund future public pensions. Nominal GDP for the year ended June 2011 was around $NZD 200 billion.
Source: Treasury, 2011 Pre-election Economic and Fiscal Update, 25 October.

Unsurprisingly, Table 1 indicates that the social portfolio, comprising assets and liabilities held to provide public services or protect assets for future generations, dominates the balance sheet. Although the (negative) net worth of the financial portfolio is broadly similar to net debt, the former includes a wider set of financial assets and liabilities. In terms of institutional form, these financial assets and liabilities are organized to achieve particular objectives. For example, there is some partial prefunding of public pensions (via the New Zealand Superannuation Fund) and of state-employee pensions (via the Government Superannuation Fund), some matching of accident liabilities (via the Accident Compensation Corporation), and some buffering against natural disasters (via the Earthquake Commission) and macroeconomic shocks (via net debt). Economic developments over recent years, together with significant earthquakes in the Canterbury region have depleted these last two buffers.

Buckle and Cruickshank (2012) assess the wide range of factors influencing the choice of debt targets in New Zealand, many of which interact with the wider elements and objectives of the balance sheet summarized above. A future path of rising gross public debt will, as the two papers commented on suggest, have implications for New Zealand's future economic growth. Nonetheless, the size of those effects and the nature of the transmission channels will likely be influenced by wider developments in the size and composition of the overall public sector balance sheet.

REFERENCES

Arellano, M. and O. Bover (1995), "Another Look at Instrumental Variables Estimation of Error-components Models", *Journal of Econometrics*, Vol. 68, pp. 29-51.

Buckle, B. and A. Cruickshank (2012), "The Requirements for Long-run Fiscal Sustainability", paper presented at the *New Zealand Association of Economists Conference*, 27-29 June, Palmerston North, New Zealand.

Caner, M. and B.E. Hansen (2004), "Instrumental Variable Estimation of a Threshold Model", *Econometric Theory*, Vol. 20, pp. 813-43.

Hansen, B.E. (1999), "Threshold Effects in Non-dynamic Panels: Estimation, Testing and Inference", *Journal of Econometrics*, Vol. 93, pp. 345-68.

Reinhart, C. and K. Rogoff (2009), "The Aftermath of Financial Crises", *American Economic Review*, Vol. 99, pp. 466-72.

——————(2010), "Growth in a Time of Debt*", American Economic Association*, Papers and Proceedings, Vol. 100, pp. 1-9.

COMMENTS ON SESSION 2
GOVERNMENT BUDGETS AND POTENTIAL GROWTH

*Gilles Mourre**

Comments on "Dynamic Labor Supply with Taxes: The Case of Italian Couples" by Maria Rosaria Marino, Marzia Romanelli and Martino Tasso (Banca d'Italia) and "Do Public Policies of A Net Revenue Maximizing Government Also Promote Informality?" by Nivedita Mukherji and Fuad Hasanov

1 *Focus and complementarities of the two papers*

Both papers consider the impact of tax policy on economic behaviour (labour supply, informality and sectoral structure). The paper by Marino *et al.*, referred to as Paper 1 in the remainder of this discussion, focuses on the labour supply of second earners and the role of the tax and benefit system. It builds on a double consensus in the economic literature: financial incentives to work are key for growth, while labour supply issues are particularly relevant for specific labour-market groups, where elasticities to net earnings is stronger. This policy question is particularly relevant for Italy, where the labour force participation rate among married women is particularly low (see Table 1). The paper by Mukherji and Hasanov, referred to as Paper 2 later on, considers the impact of tax rates on informality and tax revenues. It revisits the consensus in the literature by taking into account sectoral heterogeneity, tax evasion and corruptions and enquires about the possibility of a Laffer curve effect in case of high taxation. This policy issue is of particular relevance for developing countries and EU countries with a large tax burden and high tax non-compliance.

The two papers take very different approaches. While the first one uses micro data on Italy, the second one is based on cross-country macroeconomic indicators. However, the complementarity is blatant between the two papers: they both address two relevant structural features of the economy. They both can also be seen as part of a fiscal optimization exercise. As such, they could also help policy makers to improve the design of fiscal policy, with a view to boosting female participation and reducing poverty (Paper 1) and increasing net revenues, via a modulating tax burden, providing an adequate level of public good and reshaping regulations (Paper 2).

2 *Results*

Paper 1 builds on a micro-econometric model to assess the effect of changes in the tax-benefit system on female labour market participation. Consistently with the prediction of the economic theory, an increase in households' non-labour income (e.g., income support to poor household) is estimated to decrease overall poverty (in terms of head-count ratio) but to lower the incentives of married women to participate in the labour market. In contrast, policies aimed at increasing the return of the hours worked have positive effects on both dimensions.

Paper 2 examines the effects of fiscal and regulatory policies on the size of a country's informal economy and its government's net revenue. Changes in public policies are found to influence not only the size of the informal economy, but also the composition of production within the formal sectors. These effects are amplified when tax evasion and bribes are taken into

* European Commission, Directorate-General for Economic and Financial Affairs (ECFIN) and Free University of Brussels (ULB).

Table 1

Tax Burden on Second Earners and Female Employment Rates

Country	Labour Market Performance[(1)]		Disincentives to Work[(2)]	
	Employment Rate Female (2010)	Employment Rate Male (2010)	Inactivity Trap (67% AW) 2009	Low-wage trap (33% to 67% AW, 2009)
BE	74.4	85.5	46.3	58.0
DE	76.3	86.5	51.0	49.0
EE	73.9	75.7	22.6	23.0
IE	65.7	75.0	35.4	32.0
EL	61.1	85.3	31.9	19.0
ES	63.2	75.7	17.5	18.0
FR	76.7	87.1	38.1	23.0
IT	58.7	83.5	42.5	48.0
CY	76.6	88.4	-	-
LU	72.6	92.0	32.8	29.0
MT	47.8	88.7	33.3	23.0
NL	79.3	90.0	46.8	41.0
AT	79.7	88.7	29.2	39.0
PT	74.6	83.9	21.5	28.0
SI	82.1	85.2	55.8	42.0
SK	70.1	81.4	21.1	34.0
FI	79.2	83.9	29.2	32.0
BG	73.6	77.9	20.1	22.0
CZ	73.4	90.5	33.9	28.0
DK	80.6	85.9	78.8	63.0
LV	73.8	72.9	31.9	30.0
LT	76.1	71.4	39.5	26.0
HU	67.1	77.9	32.0	42.0
PL	71.7	82.6	39.2	31.0
RO	67.2	81.5	26.3	31.0
SE	82.0	88.0	23.9	29.0
UK	74.3	85.4	43.7	31.0
EU-27	72.2	84.8	40.2	36.1
EA-17	71.5	84.8	39.7	37.1

Source: European Commission (2001), "Tax Reforms in EU Member States", *European Economy*, No. 5/2011.

consideration. Productive public expenditures increase net revenue. Taxes are found to have a small positive impact, if any, on net revenue and to increase the informal economy. The impact of regulation on net revenue is mixed. The paper concludes that, to raise net revenue, institutional reforms are needed, aiming at better bureaucratic quality and more democratic accountability with a stepped-up fight against corruption.

3 *Methodologies and issues*

On a methodological standpoint, Paper 1 carries a thorough and very interesting analysis – albeit still preliminary – based on a micro-simulation model with a very rich theoretical specification. The model is extremely useful to simulate the impact of concrete parametric/systemic policy measures in Italy, as it consists of a structural dynamic life-cycle model well-suited to analyse household labour supply, saving, and consumption behaviour. The model captures several sources of heterogeneity regarding members of the couple (human capital and number of children) and incorporates most of the fiscal rules relevant for determining the net income of economic agents. Model parameters are estimated using cross-sectional and longitudinal data over 2004-10, which replicates the state of the Italian economy. The estimated model is used to simulate a few counterfactual policies and study their effect on labour supply and poverty.

Three issues could be taken into account as a valuable extension of the current paper 1. First, it may be worth taking varying risk aversion parameters into account, as unemployment risks are uneven across skill groups, regions and sectors. Second, some important factors are not explicitly taken into account: i) non-monetary incentives (not) to work, such as the supply of child care services, which is very relevant for Italy, ii) urban congestion, iii) costs of public transport. Third, it may also be interesting to examine the effect of moving toward a purely individual determination/calculation of tax and benefits, which are still partly computed at the level of the household (especially on the benefit side).

Paper 2 is well drafted and very policy relevant. It is based on a novel model with an attempt to validate it empirically despite strong data limitation. The model includes several types of goods. The empirical estimation uses cross-section data analysis (OLS, GMM), which benefits from a high data variability but faces serious robustness issues. The paper establishes a very relevant distinction between undeclared work and tax evasion in the formal sector.

However, Paper 2 faces some methodological limitations, which could be highlighted further as caveats, and may deserve some further sensitivity analysis. The theoretical model implies perfect labour mobility, which is not always seen in real life. The empirical results remain very fragile, as the number of observations is still very limited (around 50 observations) and the econometric specifications used consume many degrees of freedom. This poses serious problems of inference. Checking the empirical distribution of residuals may give an indication of the extent of the problem. Moreover, some pooled results may be regime dependent, as there is likely to be a great deal of non-linearity between advanced, emerging and developing economies. Therefore, one may wonder whether the results hold true for the euro area. It might also be worth using another variable of tax pressures instead of the Top Marginal Personal Income Rates. The statutory rate for corporate income taxes could be a candidate in this respect. Beside the role on the overall tax burden (highlighted in Paper 2), other relevant aspects should not be neglected and, at least, be mentioned in the paper: simplicity and stability of tax systems, the structure of taxation, the breadth of tax bases and existence of loopholes and the efficiency of individual tax design. As a more minor technical comment, standardising the institutional variables (using the standard deviation) will help interpret the size of the econometric coefficient.

Session 3

TAXATION, REGULATION AND PUBLIC SERVICES

HOW COSTLY ARE THE PUBLIC SECTOR INEFFICIENCIES?
AN INTEGRATED FRAMEWORK FOR ITS ASSESSMENT

*Jorge Onrubia-Fernández** *and A. Jesús Sánchez-Fuentes**

This paper provides a theoretical framework which integrates the conventional methodology for measuring the productive efficiency and the monetary assessment of social welfare changes associated with public sector performance. Two equivalent measures of social welfare changes generated by an improvement (or worsening) in productive efficiency are deduced using duality theory. The first one is obtained from the cost function, while the second one arises directly from the production function. Moreover, the paper induces the application of the theoretical framework proposed to empirical analysis.

1 Introduction

Nowadays, an essential issue to be analyzed in depth is the relationship between the productive efficiency of public sector and the potential budgetary savings associated with its improvement. Especially for advanced economies in which the current crisis effects are affecting the public finances in a more evident way. Quantifying these budgetary savings strongly constitute an alternative fiscal policy tool which goes beyond the traditional view of a fiscal consolidation (cut spending or tax hikes). This measure is not only helpful for short-term consolidation but also it is required to guarantee a sound long-term growth path.

Since the late eighties, the measurement of productive efficiency has received an increasing interest within the public economics area. This trend is even more evident for some specific sectors typically provided by the public sector: health, education, etc.. This growing literature has mainly focused on developing quantitative methodologies (usually grouped into parametric and non-parametric methods) from which we may achieve empirical measures of (technical, allocative or overall) efficiency with which a number of units – assumed to be homogeneous – have produced the public good(s) and service(s). Thus, all these measures usually provide us one scenario to compare their performance.

Without doubt these contributions measuring the productivity of public services are very useful to improve the management of public resources. However, there is lack of literature connecting these results with the potential budgetary gains that may arise from a reduction of public sector inefficiency.

In this vein, the OECD (2011) has recently highlighted the transcendence of implementing reforms addressed to increase the efficiency of public spending, specially for governments that are currently facing outstanding budgetary imbalances. In particular, the OECD refers to the need to improve the productivity of the public spending on education and health. In the first case, it is estimated that the gradual adoption of best practices in primary and secondary education could save resources around 0.5 per cent of GDP (with country range from 0.2 to 1.2 per cent), without

* Universidad Complutense de Madrid.

Both authors acknowledge comments from D. Santín, M. Hortas-Rico, J.J. Díaz and participants at the *European Central Bank Fiscal Policies Division Seminar* and at the *XIX Encuentro de Economía Pública*. Onrubia acknowledges the financial support of the Spanish Ministry of Science and Innovation (Project ECO2009-1003) and of the Santander-UCM program (project GR35/10A).

Corresponding author: Jorge Onrubia-Fernández. Department of Public Finance and Tax Systems. Universidad Complutense de Madrid – Campus de Somosaguas, Edificio 6, Desp. 6 – E-28223 Madrid (Spain). Phone: +34-91-3942442. Fax: +34-91-3942431. E-mail: jorge.onrubia@ccee.ucm.es

compromising the current educational targets. In the case of health, the resources released by improvements in productive efficiency could be even higher, around 2 per cent of GDP (range by country, between 0.4 and 4.8 per cent).

Moreover, the monetary gains are enormous in terms of social welfare. In this respect, it is important to account not only budgetary savings but also the monetary gains in terms of income and wealth derived from consuming a better education and health. Furthermore, from the marginal cost of public funds perspective, we should also consider the reduction in deadweight losses caused by distortionary taxes which provide these resources released.

The aim of this paper is to provide a theoretical framework which allows consistently integrate the conventional methodology for measuring the productive efficiency and the monetary assessment of social welfare changes associated with the public sector performance, defined in the basis of the output of any public activity. In particular, we deduce two measures of social welfare changes generated by an improvement (or worsening) in productive efficiency associated with the procurement of a public good. The first measure is obtained from the cost function, or in other words, from the supply side, while the second one arises directly from the production function. According to duality theory, both measures are equivalent and deducted from the same set of information.

The rest of the paper is organized as follows. In the second section, we introduce our theoretical framework, upon the basis of the conventional measures of efficiency (Farrell's radial approach). In the third section, we present our integrated approach which combines different dimensions typically involved in policy-makers decisions (welfare changes, measures of inefficiencies, etc.). Finally, the fourth section concludes.

2 The model

2.1 Recent concerns on Public Sector Efficiency (PSE)

The monitoring of public sector activity and the potential derivation of measures of the Public Sector Efficiency (PSE) clearly justify the increasing interest observed on analyses related to the Public Sector Performance (PSP, hereinafter). This section briefly discusses the recent evolution of literature focused on the relevant concept, the Public Sector Efficiency (PSE, hereinafter), which refers to the efficient allocation and production of the public good and services. The existing literature comprises alternative approaches to measure -and to evaluate- the PSP and, consequently, the PSE. A non exhaustive description of how this literature has evolved is next. Firstly, a growing number of studies (Afonso *et al.*, 2005; Borge *et al.* 2008; and Clements, 2002, among others) translated the traditional approach used to analyze the productive efficiency of firms to the case of public sector units (countries, municipalities, schools, hospitals, etc.) with the aim of obtaining empirical measures of the PSE for a set of units and rank them. Secondly, some studies (Borge *et al.* 2008, among others) have also explored the identification of determinants of these empirical measures. An alternative perspective is considered by other authors (see Afonso *et al.*, 2010; and Casiraghi *et al.*, 2009, among others) in order to include the distributional concerns traditionally linked to the public sector activity into the efficiency analysis.

All in all, it can be observed that some caveats are still present. First, most of these analyses have focused on the productive efficiency or technical efficiency (ψ). Thus, they have leaven out of the analysis issues related to the allocative efficiency (γ), a relevant component of the overall efficiency (η). This latter measure is our main interest in this paper. Second, the distributional concerns has not been yet fully incorporated to the analysis, although it is a component mostly involved in policy-makers decisions.

Our paper aims to fulfill all these caveats by combining the elements presented; (i) empirical measures of efficiency, (ii) welfare impact and distributional concerns, (iii) a monetary valuation of inefficiencies measured.

2.2 The public sector

This section introduces the notation used in subsequent sections and models the Public Sector Performance according to a framework which could be adapted to very different analysis.

Our model can be briefly described as follows. The public sector produces a vector of goods and services $X = (x^1, \ldots, x^H)$ which we consider excludable unlike pure public goods.[1] Each x^h is produced by a public agency with the corresponding production function for the case of single output, such that:

$$x^h = f(Y) \tag{1}$$

where $Y = (y_1, \ldots, y_n)$ is a vector of n inputs including fixed capital required for the activity and $f \in S = .\ \{(Y, X)\ :\ Y \text{ can produce } X\}$ with S the set of technologies.

The unitary price for each of these n inputs are included in the vector $W = (w_1, \ldots, w_n)$. Consequently, the total cost of producing x^h (c^h) is defined as:

$$c^h(x^h) = \sum y_i w_i \tag{2}$$

Assuming $H = 1$, for the sake of clarity in the presentation, this theoretical framework allows us to introduce the notation used in posterior sections by defining formally all the standard concepts of efficiency – mentioned above – from the inputs-oriented perspective.[2] First, given the minimum quantity of inputs needed for producing the level of output X (Y^*), technical efficiency (ψ) is defined as the ratio between Y and Y^*, such that:

$$\psi = \frac{\|Y^*\|}{\|Y\|} \tag{3}$$

Second, given the combination of inputs producing X at the minimum cost (Y^{**}), the allocative efficiency (γ) is defined as the following ratio:

$$\gamma = \frac{\|Y^{**}\|}{\|Y^*\|} \tag{4}$$

Third, the overall efficiency can be defined as the product of expressions (3) and (4):

$$\eta = \frac{\|Y^{**}\|}{\|Y\|} \tag{5}$$

Finally, we derive the corresponding expression for η in terms of production costs:[3]

---

1 Rivalry and excludability are assumed to consistently reflect changes in the demand observed for each public good.

2 Analogous definitions can be found in the literature according to the output-oriented measures (see Coelli, 2005) for a detailed comparison of both approaches). There are no divergences in the analyses carried out from both perspectives. Therefore, one of them can be excluded.

3 See Coelli (2005) for a detailed description.

$$\eta = \frac{c^{**}}{c} \tag{6}$$

where c and c^{**} are, respectively, the actual level of production costs and the production costs corresponding to Y^{**}, the efficient combination of inputs when producing X, from the technical and the allocative perspective.

3 PSE analysis: an integrated approach

3.1 *The "expenditure-efficiency" function*

The framework described above can be observed from a different perspective, facing the dual version of the same problem. Under these circumstances, the production of public good (x) and its level of output ($\hat{x}$) may be explained by the expenditure function assumed in production ($c(\hat{x})$), and the degree of overall efficiency ($\eta(\hat{x})$). In other words, an "expenditure-efficiency" function (Φ) which is implicit in the conventional production function of productive factors once the vector of input prices (W) is given:

$$x = f(Y)\big|_W \rightarrow x = \phi(c, \eta)\big|_W \tag{7}$$

First of all, from (6), we can express the budgetary cost of producing a quantity of public good from the vector of inputs (Y^{**}) and the degree of overall efficiency reached in the productive process, η:

$$c(\hat{x}) = \eta^{-1} \sum_{i=1}^{n} y_i^{**} w_i \tag{8}$$

Secondly, by applying the inverse function theorem to the optimal technology f_{**} (that determining the overall efficiency condition, Y^{**}), the optimal quantities of each input (y_i^{**}) to produce $\hat{x}$ are obtained. Note that these values only depend on factor prices and technological parameters of the production function:

$$y_i^{**} = f_{**}^{-1}(\hat{x}, W), i \in \{1, 2, \ldots, n\} \tag{9}$$

Next, by combining (8) and (9), and solving for $\hat{x}$ we derive the expenditure-efficiency function, Φ, as proposed:

$$\hat{x} = \phi(c(\hat{x}), \eta)|_W \tag{10}$$

To translate this general notation to our model, $c(.)$ would be the amount of resources allocated for the provision of the public good, and η the degree of efficiency with which the public agency produces this good.

3.2 Changes in the PSE, welfare impact and monetary valuation

This section presents an integrated approach which allows us to integrate the different dimensions involved in the evaluation of the Public Sector Performance; (i) changes in the degree of efficiency, (ii) welfare impacts linked to public policies, and (iii) monetary valuation of effects. The latter may facilitate the understanding of the inefficiency costs. Moreover, an improvement in the degree of efficiency will help to provide the same public good or service but with a lower level of spending.

For the sake of clarification, we detail our assumptions. First, in the following analysis it is assumed that any change in the degree of efficiency is exogenous. However, as Gibbons (2005) discusses, the existence of internal disturbances in the organizations (misscoordination, lack of incentives, etc.) may be the source of inefficiencies. Second, the social welfare generated by consumption of public good (x) is measured in monetary value in the conventional way, that is, by computing the area under the curve of demand for the good and substracting the cost of the inputs used in its production.[4] Additionally, to obtain accurate measurements of changes in consumer welfare we assume the demand functions involved to be compensated.[5] All in all, this theoretical framework contributes to measure welfare impacts linked to changes (improvements/worsening) in the degree of efficiency (η) with which the public good is produced. This analysis translates Myrick-Freeman and Harrington (1990) framework to our model.

Therefore, using our "expenditure-efficiency" function defined in (10), we have the following social welfare function:

$$\Omega = \Omega(Y, W, \eta) = \int_0^x p(u)du - \sum_{i=1}^n y_i w_i \tag{11}$$

where $p(\cdot)$ is the compensated demand function specified in its inverse form.

From equation (11) one can derive the first order conditions with respect to each inputs used (y_i), such that:

$$\frac{\partial \Omega}{\partial y_i} = p(x)\frac{\partial x}{\partial y_i} - w_i = 0, i = 1, ..., n \tag{12}$$

which determine the input demand functions $y_i^{**}(w_i, \eta)$ for all i. It should be noted here that these values are precisely those corresponding to the optimal vector of production factors, Y^{**}. It allows us to compute the optimal output level of public good for a given level of productive efficiency:

$$x^{**}(\eta) = \varphi\left(y_i^{**}(w_i, \eta), \eta\right) \tag{13}$$

Likewise, we could define the social welfare function associated with the production of this public good by considering the overall productive efficiency (η) as a main argument:

$$\Omega(\eta) = \varpi\left(y_i^{**}(w_i, \eta), \eta\right) \tag{14}$$

Applying the envelope theorem to the algebraic analysis described above, we obtain the following proposition:

---

4 Note that, as we did in the previous sections, hereinafter the notation is simplified to a single public good x to highlight the underlying intuitions.

5 See Willig (1976) for a discussion on the accurate measurement of these areas.

Proposition 1: *The net welfare gain is the value of the marginal contribution, in monetary terms, brought about by a reduction (or increase) of overall inefficiency in the production function, so that:*

$$\frac{\partial\Omega(\cdot,\eta)}{\partial\eta}=p(x^{**})\frac{\partial x^{**}(\cdot,\eta)}{\partial\eta}-\sum_{i=1}^{n}w_i\frac{\partial y_i^{**}(\cdot,\eta)}{\partial\eta}=p(x^{**})\varphi_n\left(y_i^{**}(w_i,\eta),\eta\right) \tag{15}$$

Some interesting implications are next. First, this result defines a relationship between the production function and the changes in welfare computed in the light of modification of the degree of efficiency. Second, it can be observed that, under full productivity of all inputs, the value generated by an infinitesimal improvement in productive efficiency is explained by the increase in the output generated. Third, from a different perspective, this gain could be seen as an approximation (φ_n) to the optimal technology (y_i^{**}).

Next, the dual version of this result is achieved. To do this, from (13) one can define the costs functions related to this production as a function of the optimal level of public good, the vector of inputs associated with the optimal technology and the degree of productive efficiency reached, so that:

$$c=c\left(x^{**}(\eta),\eta\right) \tag{16}$$

Accordingly, we can rewrite (11) as:

$$\Omega=\Omega\left(x^{**},\eta\right)=\int_0^{x^{**}}p(u)du-c\left(x^{**},\eta\right) \tag{17}$$

From this perspective, the social welfare, considered as the difference between consumer's surplus and producer's quasi-rents, is maximized for the level of optimal output determined by the equality between price and marginal cost:

$$p(x^{**})=\frac{\partial c\left(x^{**},\eta\right)}{\partial x} \tag{18}$$

Again, combining (17) and (18), the following proposition emerges:

Proposition 2: *The net welfare gain (loss) is the value of the marginal contribution, in monetary terms, brought about by the reduction (increase) of production cost as a consequence of an improvement (worsening) of the degree of overall inefficiency:*

$$\frac{\partial\Omega\left(x^{**},\eta\right)}{\partial\eta}=-\frac{\partial c\left(x^{**},\eta\right)}{\partial\eta} \tag{19}$$

Proof Given (17), we compute the total derivative with respect to the degree of efficiency (η). That is:

$$\frac{d\Omega\left(x^{**},\eta\right)}{d\eta}=\frac{\partial\Omega\left(x^{**},\eta\right)}{\partial x}\frac{\partial x^{**}}{\partial\eta}+\frac{\partial\Omega\left(x^{**},\eta\right)}{\partial\eta} \tag{20}$$

where:

$$\frac{\partial\Omega\left(x^{**},\eta\right)}{\partial x}=p\left(x^{**}\right)-\frac{\partial c\left(x^{**},\eta\right)}{\partial x} \tag{21}$$

and:

$$\frac{\partial\Omega\left(x^{**},\eta\right)}{\partial\eta}=p\left(x^{**}\right)\frac{\partial x^{**}}{\partial\eta}-\left(\frac{\partial c\left(x^{**},\eta\right)}{\partial x}\frac{\partial x^{**}}{\partial\eta}+\frac{\partial c\left(x^{**},\eta\right)}{\partial\eta}\right) \tag{22}$$

Firstly, as a consequence of (18), we could identify $\frac{d\Omega(x^{**},\eta)}{d\eta}$ and $\frac{\partial\Omega(x^{**},\eta)}{\partial\eta}$.

Next, from (22), grouping conveniently and using again (18), we obtain the proposition.

Corollary: *An improvement in the degree of overall inefficiency always involves an increase in social welfare.*

Again, some interesting conclusions can be derived. First, this result defines a relationship between the costs function and the changes in welfare computed when the degree of efficiency is modified. Second, these results can be understood as follows. The infinitesimal improvements in productive efficiency obtained lead to a reduction in the cost of production and, consequently, they are welfare enhancing. Third, combining Propositions 1 and 2 we obtain that the two welfare measures proposed must coincide due to the duality in the relationship between the production function and the cost function, which is underlying in (equality).

To conclude with this subsection, some interesting lessons could be extracted regarding the application of this approach to empirical analyses. First, the final results would lead to monetary valuations of the changes in the overall efficiency, which becomes a very interesting tool from the policy-makers perspective. Second, our approach integrates elements related to efficiency and others related to the equity, which allows to explore this classical trade-off (next subsection will explore this point in depth). Third, this approach requires an estimate of the production function and the cost function as well, which may limit its application when information on the production procedure and/or the production costs is limited.

3.3 Distributional issues

In this subsection, we analyze how the welfare gains from increased efficiency affect consumers of public goods and public sector itself as the producer. In this respect, we first identify the efficiency gains effects on consumer's welfare. Let Ω^C be the measure of consumer surplus used (usually equivalent or compensatory variation), so that:

$$\Omega^C=\int_0^{x^{**}}p(u)du-p(x^{**})x^{**}(\eta) \tag{23}$$

Then, the consumer's marginal gain is:

$$\frac{\partial\Omega^C}{\partial\eta}=-\frac{\partial p(x^{**})}{\partial x}\frac{\partial x^{**}(\eta)}{\partial\eta}x(\eta) \tag{24}$$

Alternatively, if we consider equation (13):

$$\frac{\partial p(x^{**})}{\partial x}=\frac{\partial p}{\partial x^{**}}\frac{\partial x^{**}}{\partial\eta} \tag{25}$$

Now, from the producer's perspective, we repeat a similar strategy. First, we define the producer's surplus in terms of η:

$$\Omega^S = p(x^{**})x^{**}(\eta) - \sum_{i=1}^{n} y_i^{**} w_i \tag{26}$$

where y_i^{**} is determined by the n input demand functions, $y_i^{**}(w_i, \eta)$.

Again, the producer's marginal gain can be obtained by differentiating the previous expression:

$$\frac{\partial \Omega^S}{\partial \eta} = -\frac{\partial c(x^{**}, \eta)}{\partial \eta} + \frac{\partial p(x^{**})}{\partial x} \frac{\partial x^{**}(\eta)}{\partial \eta} x(\eta) \tag{27}$$

In the light of the previous expressions, the following proposition can be demonstrated:

Proposition 3: *An improvement in the degree of overall inefficiency always lead to an increase in consumer's welfare. By contrast, this welfare gain is not guaranteed in the case of producers of public goods.*

Proof: On the one hand, for consumers, this proof can be reduced to check the signs of the expressions mentioned above. As $\frac{\partial p(x^{**})}{\partial x} \leq 0$ and $x(\eta) > 0$, depending on the sign of $\frac{\partial x^{**}(\eta)}{\partial \eta}$ the consumer's net welfare gain will be positive or negative. The optimal vector of inputs (from the technological and the minimization of costs' perspective) is taken as given in (13). As a consequence, a reduction of inefficiency may, in principle, lead to a decreased level of output – in equilibrium. To clarify this latter statement, we differentiate the first order conditions mentioned above, in equation (18), to achieve the following expression:

$$\frac{\partial p(x^{**})}{\partial x} \frac{\partial x^{**}(\eta)}{\partial \eta} = \frac{\partial^2 c(x^{**}, \eta)}{\partial x^2} \frac{\partial x^{**}(\eta)}{\partial \eta} + \frac{\partial^2 c(x^{**}, \eta)}{\partial x \partial \eta} \tag{28}$$

Grouping conveniently:

$$\frac{\partial x^{**}(\eta)}{\partial \eta} = \frac{\frac{\partial^2 c(x^{**}, \eta)}{\partial x \partial \eta}}{\frac{\partial p(x^{**})}{\partial x} - \frac{\partial^2 c(x^{**}, \eta)}{\partial x^2}}$$

On the one hand, looking at the denominator, it is straightforward to establish that $\frac{\partial p(x^{**})}{\partial x} - \frac{\partial^2 c(x^{**}, \eta)}{\partial x^2} < 0$. On the other hand, any improvement in η lead to reductions in costs. Thus, $\frac{\partial^2 c(x^{**}, \eta)}{\partial x \partial \eta} < 0$ and, consequently, $\frac{\partial x^{**}(\eta)}{\partial \eta}$ is always positive.

All in all, we have proved that consumer's welfare increases can be derived from the response in the production costs to an improvement in overall efficiency.

On the other hand, for producers, using the price-elasticity of public good demand, defined as $\varepsilon = \frac{p(x^{**})}{x\frac{\partial p(x^{**})}{\partial x}}$, which is negative by definition, we can prove that $\frac{\partial \Omega^S}{\partial \eta}$ will only be negative if and only if $\frac{\partial x^{**}(\eta)}{\partial \eta} > \varepsilon \frac{\frac{\partial c(x^{**},\eta)}{\partial \eta}}{p}$.

That is, the difference between the social welfare change and the variation in the consumer surplus.

From Proposition 3, the distribution of welfare gains derived from an improvement in the degree of efficiency may be established. Our results indicate that the determinants are the optimal output response to this increase and the price-elasticity of demand. In short, three different possibilities are achieved:

$$\text{(i) } 0 < \frac{\partial x^{**}(\eta)}{\partial \eta} < \varepsilon \frac{\frac{\partial c(x^{**},\eta)}{\partial \eta}}{p} \Leftrightarrow \frac{\partial \Omega^C}{\partial \eta} > 0, \frac{\partial \Omega^S}{\partial \eta} > 0 \tag{29}$$

$$\text{(ii) } \varepsilon \frac{\frac{\partial c(x^{**},\eta)}{\partial \eta}}{p} < \frac{\partial x^{**}(\eta)}{\partial \eta} \Leftrightarrow \frac{\partial \Omega^C}{\partial \eta} > 0, \frac{\partial \Omega^S}{\partial \eta} < 0 \tag{30}$$

In order to show a different perspective of the conclusions described so far, we consider now an example to illustrate (and reinforce) the underlying intuitions. Moreover, some implications for the empirical application of this approach are discussed.

We consider a scenario in which the overall efficiency to produce the public good x improves between two moments in time, from η_0 to η_1. To quantify the value of social welfare generated by the change in the degree of efficiency, we may choose to integrate, alternatively, one of the two welfare change measures presented in Propositions 1 and 2, respectively, and use $[\eta_0, \eta_1]$ as integration interval:

$$\Delta\Omega = \int_{\eta_0}^{\eta_1} p(x^{**})\varphi_n\left(y_i^{**}(w_i,\eta),\eta\right) = -\int_{\eta_0}^{\eta_1} c(x^{**},\eta) \tag{31}$$

From the empirical point of view, the direct quantification of $\Delta\Omega$ from any of the two alternatives shown in (31) requires to determine the changes in the equilibrium output and in the optimal combination of inputs caused by the change in the degree of productive efficiency. This informational requirement should be added to those previously mentioned when estimating the production and/or cost function.

On the contrary, this computation may be simplified when information on production levels of public good before and after to the change analysed is available. To do this, using (11), we simply need to calculate the difference between initial and final social welfare values:

$$\Delta\Omega = \int_0^{x_1} p(u)du - c(x_1,\eta_1) - \int_0^{x_0} p(u)du + c(x_0,\eta_0) \tag{32}$$

By using this quantification, it can be observed how the potential welfare gains resulting from improved efficiency come from the displacement of the supply curve (as there is a reduction in the cost function). In other words, marginal cost of producing public good goes from $\frac{\partial c(x,\eta_0)}{\partial x}$ to $\frac{\partial c(x,\eta_1)}{\partial x}$.

Following to Myrick-Freeman and Harrington (1990), we can obtain an alternative expression for (32) by incorporating the change experienced by the cost function.

To do this, we use the line integral of its gradient along any path between (x_0, η_0) and (x_1, η_1), and integrate along the line connecting them, such that:[6]

$$\Delta\Omega = \int_{x_0}^{x_1} p(u)du - \int_{\eta_0}^{\eta_1} \frac{\partial c(x_0,\eta)}{\partial \eta} d\eta - \int_{x_0}^{x_1} \frac{\partial c(x_1,\eta)}{\partial x} dx \tag{33}$$

Figure 1 shows the net social welfare gain expressed in (33) (the shaded area marked ΔΩ). For the sake of simplicity, we assume linearity for all the curves involved; both compensated public good demand, and marginal cost functions (pre- and post-).

According to the analysis presented above, we could additionally define welfare changes experienced by consumers and the public sector as public good supplier. On the one hand, consumers enhance their welfare by increasing the area under the compensated demand curve, as a consequence of the equilibrium price decrease, from p_0 to p_1.

Figure 1

Net Social Welfare Gain

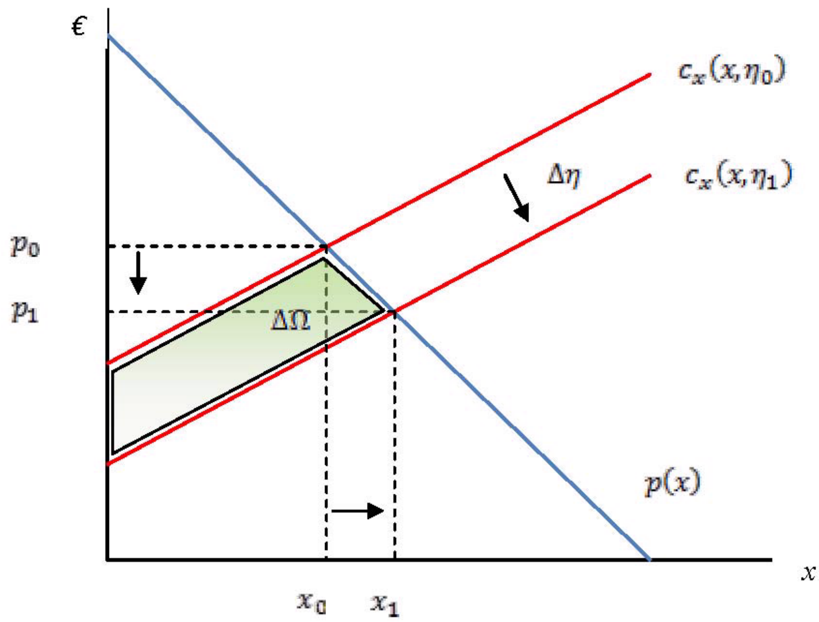


Figure 2 shows the consumers' welfare gain, which is represented by the total upper shaded area. On the other hand, the net change in producer's welfare results from compensating for the decrease in their initial surplus due to the lower resulting price (the patterned upper shaded area) with the new surplus caused by the reduction of costs charted in the new marginal cost function (the lower shaded area marked $\Delta\Omega^S$).

As a consequence, combining this graphical evidence with propositions presented above, we conclude that:

6 See Myrick-Freeman and Harrington (1990) for further details on the underlying method, which is out of the scope of this paper.

i) for any $\eta > 0$, $\Delta\Omega = \Delta\Omega^C + (\Delta\Omega^S - \nabla\Omega^S) > 0$;

ii) we have not any guarantee implying that $(\Delta\Omega^S - \nabla\Omega^S) > 0$.

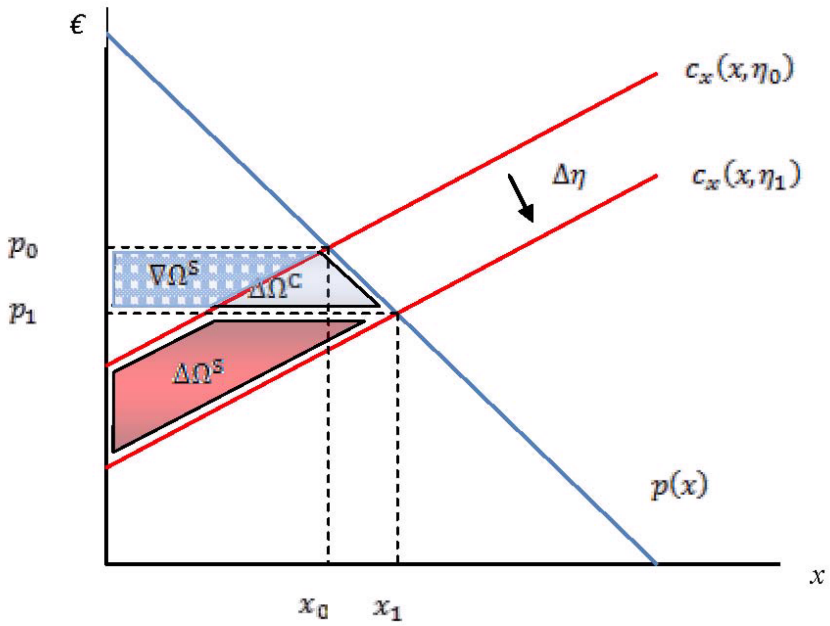
4 Concluding remarks

In the light of the current economic situation, the near future points to intense (supra-/intra-) national social debates on the monitoring of public sector performance (health, education, etc.).

Particularly, advances economies are currently facing issues related to the reorganization of their welfare state. Within this framework, quantifying these budgetary savings strongly constitute an alternative fiscal policy tool which goes beyond the traditional view of a fiscal consolidation (cut spending or tax hikes). This measure is not only helpful for short-term consolidation but also it is required to guarantee a sound long-term growth path.

Figure 2

Consumers' Welfare Gain



In this respect, important policy implications are derived from our results. First, this paper has presented an integrated approach which combines different dimensions involved in the usual policy-makers decisions (efficiency in the production of the public good, welfare impacts and monetary valuation). This proposal satisfies additional features in comparison to the usual methodologies extensively used so far. Mainly, our approach would allow to translate measures of (in)efficiencies into to a monetary value. Second, our proposal may be adapted to be used within a wide variety of empirical applications monitoring and/or evaluating the public sector performance. In this respect, we have identified the information requirements. Finally, we have derived some analytical results which help to understand the underlying intuitions and their linkages.

Finally, this paper links and integrates two different fields growing in parallel so far. On the one hand, empirical analyses monitoring the public sector performance from the production side and, on the other hand, studies analyzing the welfare implications of public policy-makers. For instance, this approach may provide guidance to the design of fiscal consolidation programs, so that they are compatible with a more efficient use of public resources.

REFERENCES

Afonso, A. and C. Scaglioni (2005), “Public Services Efficiency Provision in Italian Regions: A Non-parametric Analysis”, Technical University of Lisbon, Department of Economics at the School of Economics and Management (ISEG), Working Paper, No. 2005/02.

Afonso, A., L. Schuknecht and V. Tanzi (2005), “Public Sector Efficiency: An International Comparison”, *Public Choice*, Vol. 123, No. 3, pp. 321-47.

——————(2010), “Income Distribution Determinants and Public Spending Efficiency”, *Journal of Economic Inequality*, Vol. 8, No. 3, pp. 367-89.

Borge, L.E., F. Torberg and P. Tovmo (2008), “Public Sector Efficiency: The Role of Political and Budgetary Institutions, Fiscal Capacity, and Democratic Participation”, *Public Choice*, Vol. 136, pp. 475-95.

Casiraghi, M., R. Giordano and P. Tommasino (2009), “Behind Public Sector Efficiency: The Role of Culture and Institutions” in S. Barrios, L. Pench and A. Schaechter (eds.), *The Quality of Public Finances and Economic Growth*, European Commission, Occasional Paper, No. 45, March.

Charnes, A., W.W. Cooper and E. Rhodes (1978), “Measuring the Efficiency of Decision Making Units”, *European Journal of Operational Research*, Vol. 2, No. 6, pp. 429-44.

Clements, B. (2002), “How Efficient is Education Spending in Europe?”, *European Review of Economics and Finance*, Vol. 1, No. 1, pp. 3-26.

Coelli, T.J., D.S. Prasada Rao, C.J. O’Donnell and G.E. Battese (2005), *An Introduction to Efficiency and Productivity Analysis*, 2nd edition, Springer.

Debreu, G. (1951), “The Coefficient of Resource Utilization”, *Econometrica*, No. 19, pp. 273-92.

Fakin, B. and A. de Crombrugghe (1997), “Fiscal Adjustment in Transition Economies: Social Transfers and the Efficiency of Public Spending: A Comparison with OECD Countries”, The World Bank, Policy Research, Working Paper, No. 1803, Washington (D.C.).

Farrell, M.J. (1957), “The Measurement of Productive Efficiency”, *Journal of the Royal Statistic Society*, No. 120, pp. 253-81.

Figlio, D.N. and L.W. Kenny (2009), “Public Sector Performance Measurement and Stakeholder Support”, *Journal of Public Economics*, No. 93, pp. 1069-77.

Gibbons, R. (2005), “Four Formal(izable) Theories of the Firm?”, *Journal of Economic Behavior & Organization*, No. 58, pp. 200-45.

Gupta, S. and M. Verhoeve (2001), “The Efficiency of Government Expenditure. Experiences from Africa”, *Journal of Policy Modelling*, No. 23, pp. 433-67.

Myrick Freeman III, A. and W. Harrington (1990), “Measuring Welfare Values of Productivity Changes”, *Southern Economic Journal*, No. 56, pp. 892-904.

OECD (2011), *Economic Policy Reforms 2011. Going for Growth*, OECD, Paris.

Willig, R.D. (1976), “Consumer’s Surplus Without Apology”, *American Economic Review*, No. 66, pp. 589-97.

GROWTH IMPLICATIONS OF STRUCTURE AND SIZE OF PUBLIC SECTORS

Hans Pitlik[*] *and Margit Schratzenstaller*[*]

The relationship between government size and growth has received an enormous attention in the economics literature, and the recent financial crisis has forced this topic back on the agenda. A highly controversial debate in this respect is whether large governments are harmful for growth. Endogenous growth theory provides us with the view that tax structure and the composition of public expenditure may be important for growth, perhaps even more than total tax or expenditure levels. Government size and structure are, however, also reflected in the level and structure of market regulations, which may substitute or complement fiscal intervention.

The study provides an overview of the growth friendliness of fiscal and regulatory structures in a cross-section of EU15- and EU12-members and highly developed OECD countries. Peripheral European (transition) countries are also included, whenever respective data are available. Our analysis is based on several measures capturing the expenditure and the tax side of the budgets, as well as regulatory policies. It is shown that the size and the structure of fiscal and regulatory regimes and, hence, the expected long run-growth impact of government activities, still differ markedly across countries.

1 Introduction

The relationship between government size and growth has received an enormous attention in the economics literature. One of the main questions in this respect is, "are large governments harmful for growth?" While Neoclassical Theory sees only an insignificant role for fiscal policy to impact on the long-run rate of economic growth, Endogenous Growth Theory provides us with the view that fiscal policy can generate permanent effects on the steady state growth rate of output, and not just temporary effects, *i.e.*, on the transitional dynamics towards a higher output level. A number of theoretical models predict that tax structure and the composition of public expenditure may be important for growth, probably even more than total tax or spending levels (e.g., Lucas, 1988; Barro, 1990; Barro and Sala-i-Martín, 1992). Moreover, a non-negligible literature discusses the potential growth effects of international openness or the regulatory regimes on factor and goods markets, which could be seen as a further dimension of public sector size and structure.

Together with the availability of more and better data, both in the cross-section and over time, empirical research on the determinants of economic growth increased remarkably over the last 20 years. Although there is still a substantial model uncertainty leading to a lack of robustness of empirical growth analyses (e.g., Nijkamp and Poot, 2004; Ciccone and Jarocinski, 2010), it is now widely acknowledged that properly designed fiscal and regulatory policies can play an important role in supporting economic growth (e.g., Tanzi and Zee, 1997; Kneller, Bleaney and Gemmell, 1999; Bleaney, Gemmell and Kneller, 2001; Fölster and Henrekson, 2001; Zagler and Durnecker, 2003; Angelopoulos, Economides and Kammas, 2007; Ghosh and Gregoriou, 2008; Romero-Ávila and Strauch, 2008; Gemmell, Kneller and Sanz, 2011). A survey of both older and recent studies, as well as an interpretation of results is available in Bergh and Henrekson (2011).

In this respect it should be emphasized that many empirical analyses focus on developed countries (OECD or EU15), with some notable exceptions (Campos and Coricelli, 2002; Fidrmuc,

---

[*] WIFO, Austria.

This paper is a considerably shortened version of Pitlik and Schratzenstaller (2011).

2003; Bose, Haque and Osborn, 2007; Pushak, Tiongson and Varoudakis, 2007; Baldacci *et al.*, 2008; Bayraktar and Moreno-Dodson, 2010) which concentrate on transition economies and developing countries, respectively. The suitable design of growth-enhancing policies will nevertheless differ substantially across different countries. Accounting for the stage of economic development, the political and institutional environment and (probably) historical legacies of a country, a one-size-fits-all-fiscal and/or regulatory policy in order to promote growth is almost certainly not appropriate. Moreover, the recent Financial Crisis and the Great Recession might lead to a somehow revised view on the role of the state in supporting growth and long-run economic development (Griffith-Jones, Ocampo and Stiglitz, 2010; Blanchard, Dell'Ariccia and Mauro, 2010).

Against this background the purpose of the present paper is to provide a very brief overview of the literature on the growth impact of fiscal (*i.e.*, tax and expenditure) as well as regulatory policies. The main part of the article addresses the question to what extent European and OECD countries (or country groups) suit to concepts of growth-friendly fiscal and regulatory policies.

We proceed as follows. Section 2 is devoted to government expenditure structures. Following a brief discussion of the categorization of public spending categories into "productive" and "unproductive" types, we analyze the development of several spending categories. In a next step we investigate the growth friendliness of expenditure structures. Section 3 presents the tax structures and their evolution over time in a sample of European countries, using adequate macroeconomic and microeconomic indicators. We evaluate the growth friendliness of tax structures and their evolution based on the "tax and growth"-hierarchy derived by the OECD. In Section 4 we turn to the regulation issues. The growth impact of regulatory regimes is less well documented and even more controversially debated than the fiscal size and structure of government. Nevertheless, several empirical investigations support the view that stricter regulation of goods and factor markets is detrimental to economic development. Recent theoretical and empirical research emphasizes the notion of complementarities between institutions and policies in order to enhance growth. Section 5 therefore aims to provide an overall assessment of economic policy regimes and their growth friendliness in a comparative way. Of special interest in this respect is whether there are systematic deficiencies of certain countries (country groups) in providing a combination of growth-friendly economic policies. We will also consider the possibility that some countries provide more (less) regulation (or more/less taxes and expenditure) as a compensation for a lack of (more) reforms in another policy area. Section 6 concludes.

2 Government expenditure

2.1 *Productive vs. unproductive public spending: theoretical background*

The connection between government spending and growth is probably one of the most controversially debated topics in economics. In theory the relationship is ambiguous. On the one hand, government expenditure is deemed an indispensable prerequisite for economic development. The protection and enforcement of private property rights and contracts appear to be the most important factors for economic prosperity and growth. A well-functioning legal system (including expenditure for the courts) and enforcing public order and safety (including the police and the armed forces) are a precondition for economic specialization and the operation of markets (e.g., Hayek, 1960; Buchanan, 1975; North, 1990).

In addition to these essential functions of government, a number of further public goods are considered as potentially growth-enhancing. The operation of a high-quality physical infrastructure as well as basic educational services clearly fall under this category, given that governments will produce or provide these goods more efficiently than markets. At least according to Welfare

Economics, market-failures from public goods, information asymmetries, (network) externalities, and natural monopolies, can be corrected by different categories of public spending (and also by taxation or regulation measures, all subject to cost-benefit-considerations), thus potentially leading to a more efficient allocation of scarce resources through additional government health expenditure, spending on environmental issues, etc.

Beyond such core allocative functions the Musgravian tradition of Public Finance (Musgrave, 1959) advocates a distributional role as well as a stabilization function of government spending. Although not evidently linked with the goal of enhancing economic growth, government spending on these two functions nevertheless has an impact on growth performance, which may be either positive or negative. Higher government spending and a larger public sector may be better able to stabilize the economy if it is hit by macroeconomic shocks (e.g., Fatás and Mihov, 2001), which might also be conducive to longer-run growth (e.g., Ramey and Ramey, 1995; Martin and Rogers, 2000). Higher social transfer spending may not only improve the distribution of income and wealth, and thus satisfy political equity considerations, but may also improve the functioning of labor markets and – under certain circumstances – reduce social conflict in society and thereby enhance growth (e.g., Perotti, 1996).

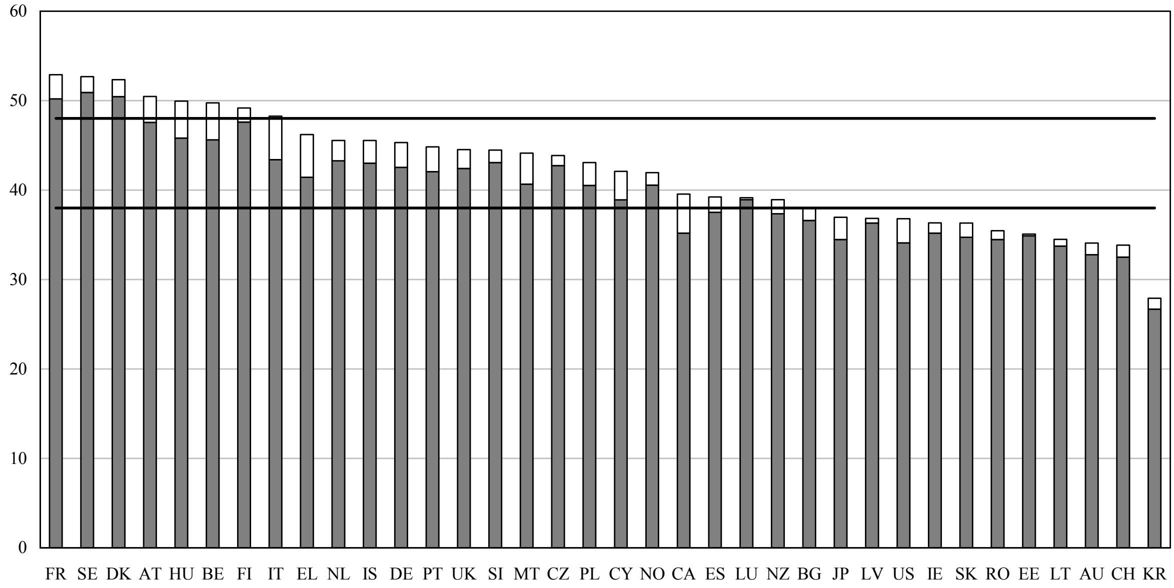
On the other hand, the debate about the appropriate role and size of the state has also shown that in general an ever increasing government sector, as measured by total spending, will slow down or inhibit growth for a number of (partially interconnected) reasons:

- disproportionally increasing distortionary effects of higher levels of taxation to fund increasing expenditures are detrimental for growth, probably also depending on the tax structure. This will be discussed in more detail in Section 3;
- long-run growth effects of most (if not all) public spending categories are subject to diminishing marginal returns, *i.e.*, at higher expenditure levels the marginal productivity of additional public spending is expected to decline. Also, the stage of development of a country will matter. Highly developed countries probably require a different expenditure composition as compared to less developed or transition economies;
- several types of expenditures yet create disincentives for the recipients (households as well as enterprises), leading to a crowding out of productive private spending and a reduction of economic efforts of beneficiaries, which, in turn, impedes growth;
- inside the public bureaucracy resources are often wasted and/or used inefficiently, due to lack of appropriate incentives. Public sector governance will play a crucial role in this respect, as inefficient provision of public services is more likely if institutions are weak. This effect will exacerbate if expenditure levels are high.

Summing up, the theoretical link between government expenditure and economic growth is rather complex. At least, the relationship between public spending and growth appears to be of a non-linear type, depending on factors like type of expenditure under consideration, initial spending level, internal efficiency of public provision, and the level and structure of taxation. In any case there is a theoretical optimum in which a certain level of public expenditure maximizes economic growth, given the disincentive effects of taxation and the level of bureaucratic efficiency. Empirically, these nonlinear effects between spending levels and economic growth are not easy to test because governments do not necessarily prioritize core productive functions of government responsibility over other forms of intervention. Ultimately, as a clear-cut theoretical relation cannot be derived, it is a matter of empirical testing whether and which types of government spending should be classified as "productive" or "unproductive".

Figure 1

Aggregate Government Expenditure Shares
(averages 2004-08; percent of GDP)



Source: EUROSTAT, OECD, and WIFO calculations.

2.2 *Size and structure of government spending*

2.2.1 *Aggregate expenditure*

The most commonly used measure for government size is its expenditure share over GDP. As noted above, there is some evidence that high aggregate spending levels can be an impediment for growth. At least, even if empirical results are sometimes not robust, no recent study finds a positive relationship between long-run growth and high total public expenditure levels.

To get a first impression on the level of government spending, we employ a sample of 36 OECD- and EU27-countries,[1] and display 5-year-averaged values over the years 2004-08 in Figure 1.[2] A 5-year-period is chosen in order to smooth out effects of the business cycle on spending levels. 2009 is not included as during that year most countries' spending-over-GDP ratios are biased upwards, due to a rapid GDP decline plus fiscal stimulus programs as a response to the recent Financial Crisis and the Great Recession.[3] The average 5-year spending level in the sample was 42.1 per cent of GDP, with a minimum of 27.9 per cent (Korea) and a maximum of 52.9 per cent (France). Primary spending levels amounted on average to 39.9 per cent of GDP, with

---

1 The sample includes all 27 EU-members plus all OECD-members that are not members of the EU27, except for Mexico, Israel, Chile and Turkey, both due to a lack of data and structural dissimilarities.

2 If not noted otherwise, we always refer to general government figures. Of course, the degree of decentralization of a country's fiscal responsibilities may also have an effect on the growth effects of government spending. These issues are, however, not dealt with in this paper. See, e.g., Schaltegger and Torgler (2006).

3 Except for Malta and Iceland all countries in the sample increased primary spending over GDP between 2008 and 2009. In Iceland, primary spending already in 2007 exploded from 39.7 to 54.2 per cent of GDP (2008). A simple regression shows that spending increases were somewhat larger in countries with an initially smaller spending level in 2008.

a maximum of 50.9 per cent (Sweden) and a minimum 26.7 per cent in Korea. Interest payments reached on average 2.2 per cent, but Greece and Italy already faced an interest burden of 4.8 per cent of GDP over 2004-08. In any case, interest payments are considered as least productive spending type, as they are exclusively related to past political decisions, and reduce the margin for strategic future-oriented spending of governments currently in office.[4]

Somewhat arbitrarily, we can divide the sample of 36 countries into three sub-samples according to average aggregate spending levels over 2004-08. The group of big spenders consists of countries with a mean expenditure-to-GDP-ratio above 48 per cent.[5] The small government group is made up of countries with average spending levels below 38 per cent of GDP, approximately the mean spending level minus one standard deviation.[6] The medium-spending group consists of countries with a mean expenditure share between 38 and 48 per cent over 2004-08.[7]

2.2.2 Productive vs. non-productive government spending

Preliminaries

The core of endogenous growth models with public spending is that not (only) the total volume of government expenditure is relevant for growth but its composition and, thus, the allocation between expenditure types which are growth enhancing (productive), growth depressing or neutral (non-productive) with respect to economic growth. From the viewpoint of these theories it is in particular the components of government spending that enter directly or as intermediate public inputs the production function of private enterprises which are expected to have a positive impact on a country's growth performance (Barro, 1990; Gemmell, Kneller and Sanz, 2011).

Although the theoretical concept is quite clear it is, however, not so obvious which types of government spending should be counted as productive. Empirical research supports a substantial positive impact of some spending components on growth, but there is still no agreement on which categories. In their survey of the relevant literature Bayraktar and Moreno-Dodson (2010) guess that "[o]ne possible explanation for the mixed results in the literature is sample selection. What we expect is that public spending can improve growth performance of countries only if they are able to use these expenditures productively". This means that the productivity of several public spending types, *i.e.*, their growth-promoting effects, depends critically on the institutional and economic environment of a country.

Another important point of the ongoing debate on productive and non-productive public expenditure is that one should take a more functional perspective. What matters is not the formal economic categorization of several spending types into consumption or investment spending per se, but for which function the money is used. Wages and salaries which are – by definition – a substantial part of government consumption can be employed for highly productive uses (e.g., educational issues) but also for unproductive purposes (e.g., salaries for outdated bureaucracies).

In Table 1 we report a categorization which is based on Gemmell, Kneller and Sanz (2011) with several adaptations and modifications based on European Commission (2002), Barrios and Schaechter (2008) and Bayraktar and Moreno-Dodson (2010). The assignments shown in Table 1

4 The correlation between primary spending and interest spending is only weakly positive (+0.27 in the sample over the years 2001-10).

5 This group is composed of France, Sweden, Denmark, Austria, Hungary, Belgium, Finland and Italy.

6 Korea, Switzerland, Australia, Lithuania, Estonia, Romania, Slovakia, Ireland, the USA, Latvia, Japan and Bulgaria all belong to the small-spender group.

7 Greece, the Netherlands, Iceland, Germany, Portugal, the United Kingdom, Slovenia, Malta, Czech Republic, Poland, Cyprus, Norway, Canada, Spain, Luxembourg and New Zealand (listed from higher to lower shares).

Table 1

Components of Productive and Non-productive Government Spending

Expenditure Type (Theoretical)	Expenditure Type (SNA, COFOG)	Remarks on Productive Impact
Productive		
Core public services	General public administration	Basic services for organization of democracy and public administration
	Public order and safety	Includes spending on police, courts etc.
	Defense	Growth effects disputed, dependent on external threats (?)
Infrastructure spending	Public investment in Economic Affairs	Investment in transport and communication as well as other infrastructure services
	Housing and community services	Predominantly spending for local infrastructures (e.g., water supply)
	Environmental protection	Growth effects disputed
Merit goods/Externalities	Education	Increases productivity of labor, but could also be provided privately in principle
	Health	Increases productivity of labor, but could also be provided privately in principle
Non-productive		
Redistribution	Economic services	Sectoral subsidies, often with sclerotic effects, although some forms of horizontal subsidies (R&D-spending) are productive
	Social protection	Basic social protection may be productive if it improves labor market functions and reduces social tensions
Other	Recreation, culture, religion	Possible indirect positive impact on growth via health channel
Interest payments	Interest payments	Exclusively past-related spending

Source: WIFO compilation, based on Gemmell, Kneller and Sanz (2011). Supplemented by European Commission (2002), Semmler *et al.* (2007); Barrios and Schaechter (2008); Bayraktar and Moreno-Dodson (2010).

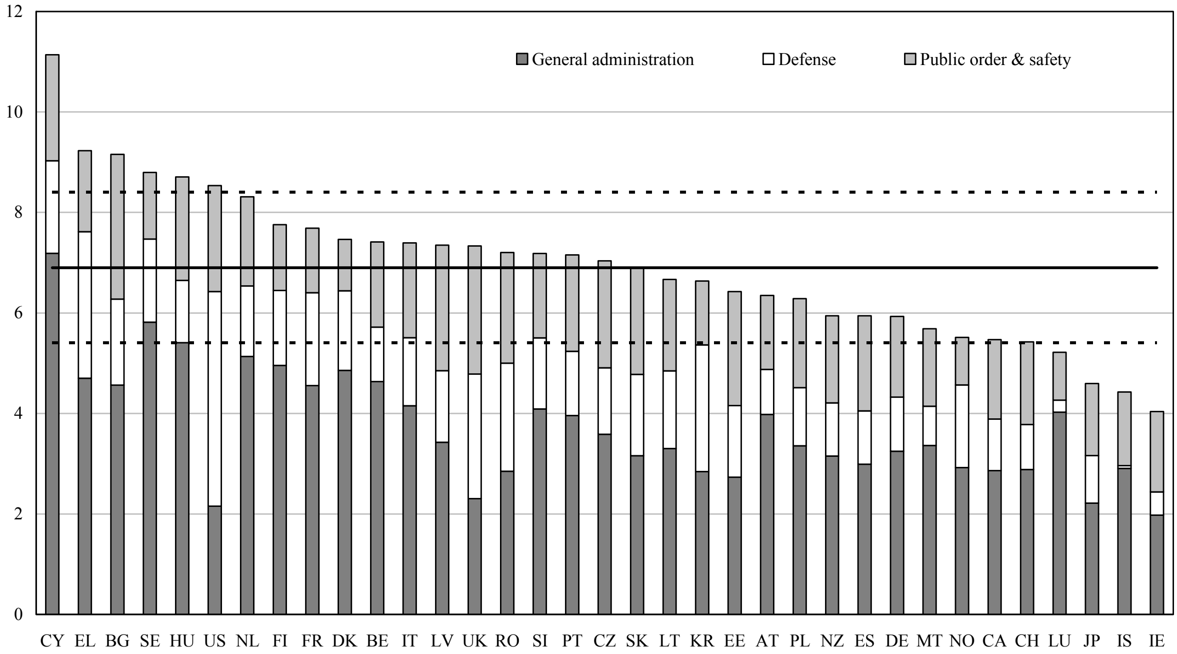
are based on results of macroeconomic research on the impact of fiscal policies. Microeconomic evidence may lead to partly different conclusions.

Core public services

Expenditures for core public services consist of spending for general administration, public order and safety, and defense. Their growth impact stems from the fact that a minimum of public administration services is required in all (democratic) systems, as well as institutions of enforcing law, order and public safety, probably also against external threats.

Figure 2

Government Spending on Core Public Services
(averages 2004-08; percent of GDP)



Source: EUROSTAT, OECD, and WIFO calculations.

Average expenditures on core public services in 35 countries amount to 6.9 per cent of GDP over the years 2004-08.[8] The smallest expenditure ratios (less than 5 per cent of GDP) are found in Ireland, Iceland and Japan; Cyprus, Greece, Belgium, Sweden, Hungary and the USA observe the highest spending on core services in relation to GDP (see Figure 2). In relation to total spending (over the years 2004-08), expenditure on core services on average equal 16.9 per cent, with a range between 9.8 per cent (Iceland) and 26.5 per cent of total spending in Cyprus.

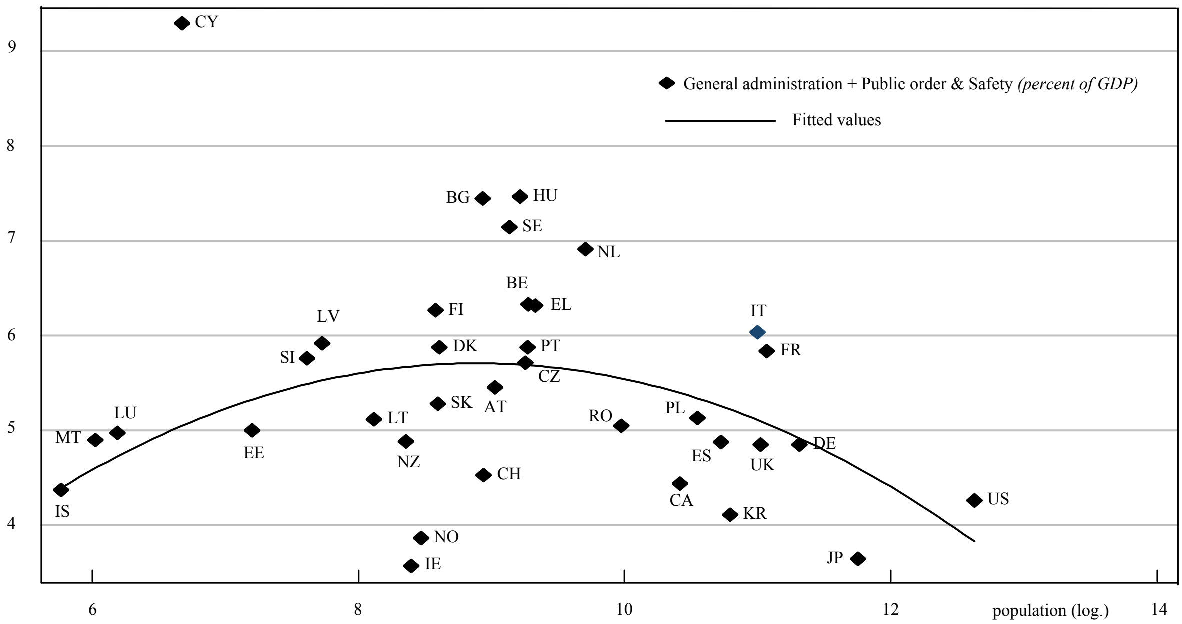
Subtracting defense spending, Figure 3 illustrates no clear evidence that expenditure on general administration and public order and safety are characterized by economies of scale. Neglecting the obvious outlier Cyprus, a hump-shaped relation between population size (in logs) and core public service spending appears to exist, with smaller expenditure ratios in very small and very large countries.

A high quality physical infrastructure is a productivity-enhancing input in private production processes and thus a major driver of a country's growth performance (e.g., Aschauer, 1989; Romp and de Haan, 2007; Crafts, 2009; Egert, Kozluk and Sutherland, 2009). Public infrastructure capital includes utilities and devices for transport and communication, energy and water supply etc. Government spending for infrastructure purposes is frequently approximated by gross fixed investment in the government sector. However, such a statistical recording entails a number of difficult-to-solve problems (e.g., Alegre *et al.,* 2008).

[8] Source: COFOG-databases of EUROSTAT and OECD. Interest spending that is allocated to COFOG-division 1 (General Public Administration) is deducted. For New Zealand, Canada, and Japan, data are only available until 2005/2006/2007. Hence, we calculated an average for shorter time periods. Data for Switzerland include only the years 2007 and 2008, as earlier data are unavailable. Data for Australia are not available.

Figure 3

Government Expenditure on General Administration, Public Order & Safety vs. Population Size
(averages 2004-08; percent of GDP)



Infrastructure spending

Hence, we decided to use a somewhat different classification: According to our definition, infrastructure spending encompasses total government expenditure (current *and* investment spending) in COFOG divisions 5 (Environmental protection) and 6 (Housing and community amenities) plus gross government investment in division 4 (Economic affairs). In our view, this classification captures best of what should be subsumed under the heading of infrastructure spending, which is not necessarily identical to investment expenditure.

Mean infrastructure spending defined along these lines is on average 2.8 per cent of GDP in the sample (averaged over 2004-08).[9] The range is between 1.4 per cent (Denmark) and 5.2 per cent (Czech Republic). The high spending group also includes Korea, Ireland, Japan, and Romania, whereas Austria, Switzerland, the USA, Finland and Belgium all belong to a group with low infrastructure spending (Figure 4). In relation to total government spending, infrastructure expenditure make up on average 7 per cent. Smallest shares of less than 3 per cent of total spending are observed in Denmark and Austria; the highest shares in Korea (16.1 per cent) and Ireland (12.1per cent).

Figure 5 plots infrastructure investment levels over 2004-08 against real GDP per capita (in international US-Dollars (logs) in 2003.[10] A strong negative relation indicates that countries in a catching-up process tend to have higher infrastructure expenditures, whereas countries that already have a high GDP per capita, and presumably a higher quality public capital stock, observe smaller spending in relation to GDP. Smaller government spending on infrastructure may therefore

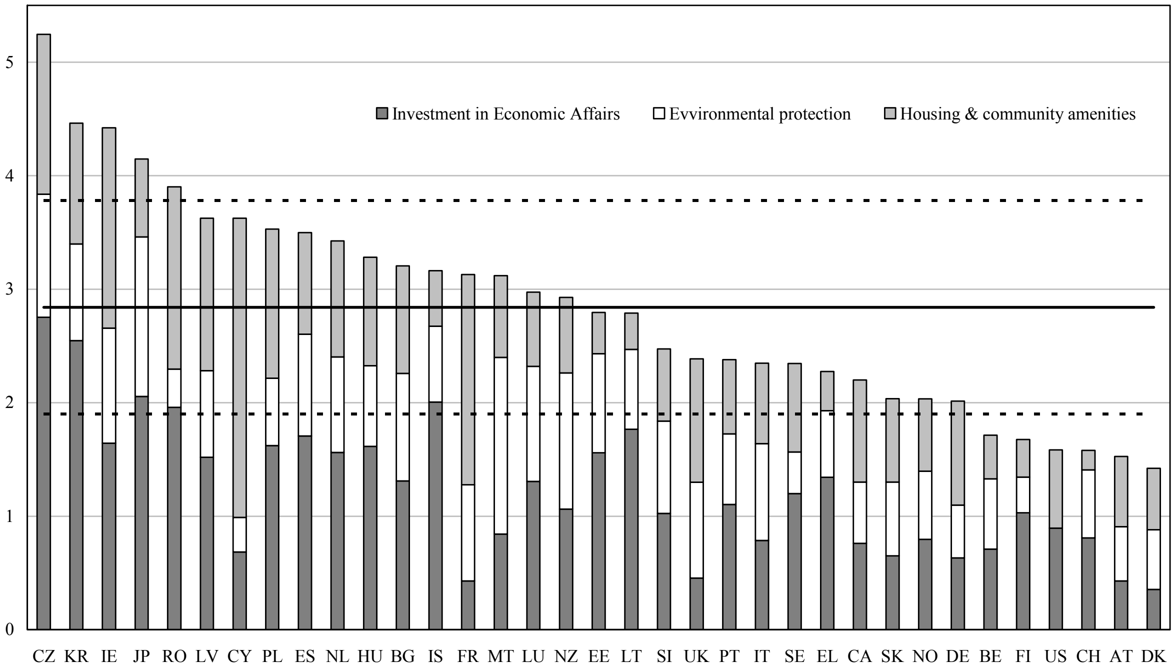
---

9 With respect to data availability and gaps in the data, see footnote 10.

10 Data are from the Penn World Tables 7.0.

Figure 4

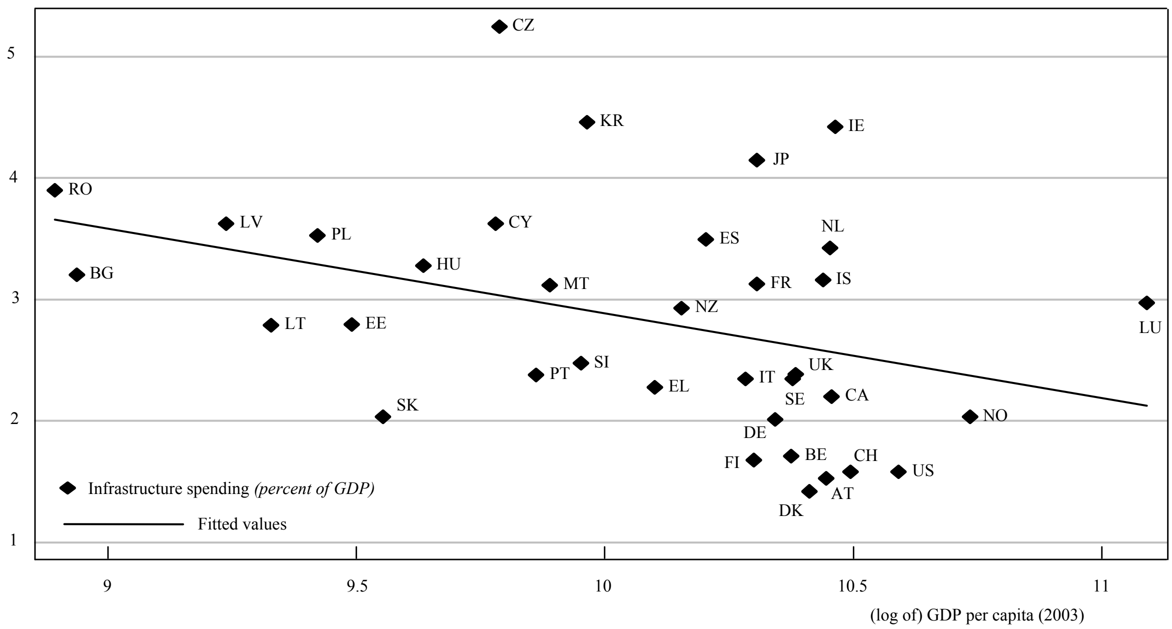
Government Spending on Infrastructure
(averages 2004-08; percent of GDP)



Source: EUROSTAT, OECD, and WIFO calculations.

Figure 5

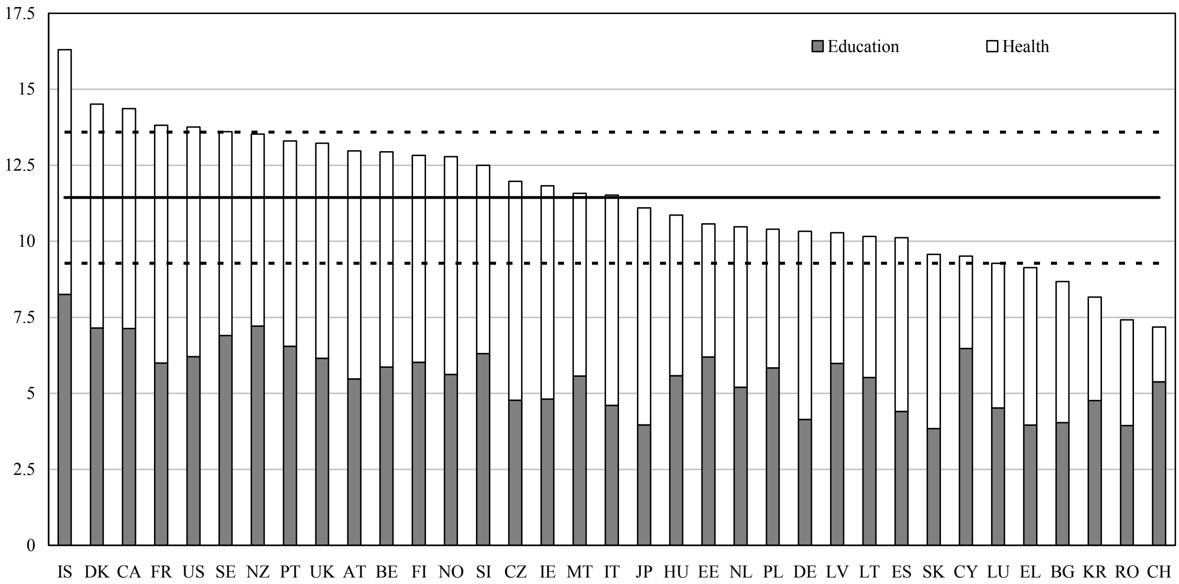
Government Spending on Infrastructure versus GDP per capita 2003



Source: Penn World Tables 7.0, EUROSTAT, OECD, and WIFO calculations.

Figure 6

Government Spending on Education and Health
(averages 2004-08, percent of GDP)



Source: EUROSTAT, OECD, and WIFO calculations.

also be a sign of diminishing returns to public capital (see also Kamps, 2006).[11] Empirical evidence for such a saturation effect is, however, not very strong (Välilä, Kozluk and Mehrotra, 2005), but some country data may be severely biased by off-budget investment that is accounted for as private sector spending.

Spending on merit goods/externalities: education and health

A substantial share of government expenditure of modern Welfare States is devoted to spending on merit goods. The two most prominent examples are education and health spending. With respect to the growth effects of both spending categories the impact of human capital investment is common wisdom now (e.g., Bassanini and Scarpetta, 2002; Baldacci *et al.*, 2008). If public spending on education and health care improve human capital then this should show up in a better growth performance. Especially for economies that operate at the technology frontier human capital investment through education and health care improvements are of crucial importance (e.g., Aghion, 2008).

Redistributive spending

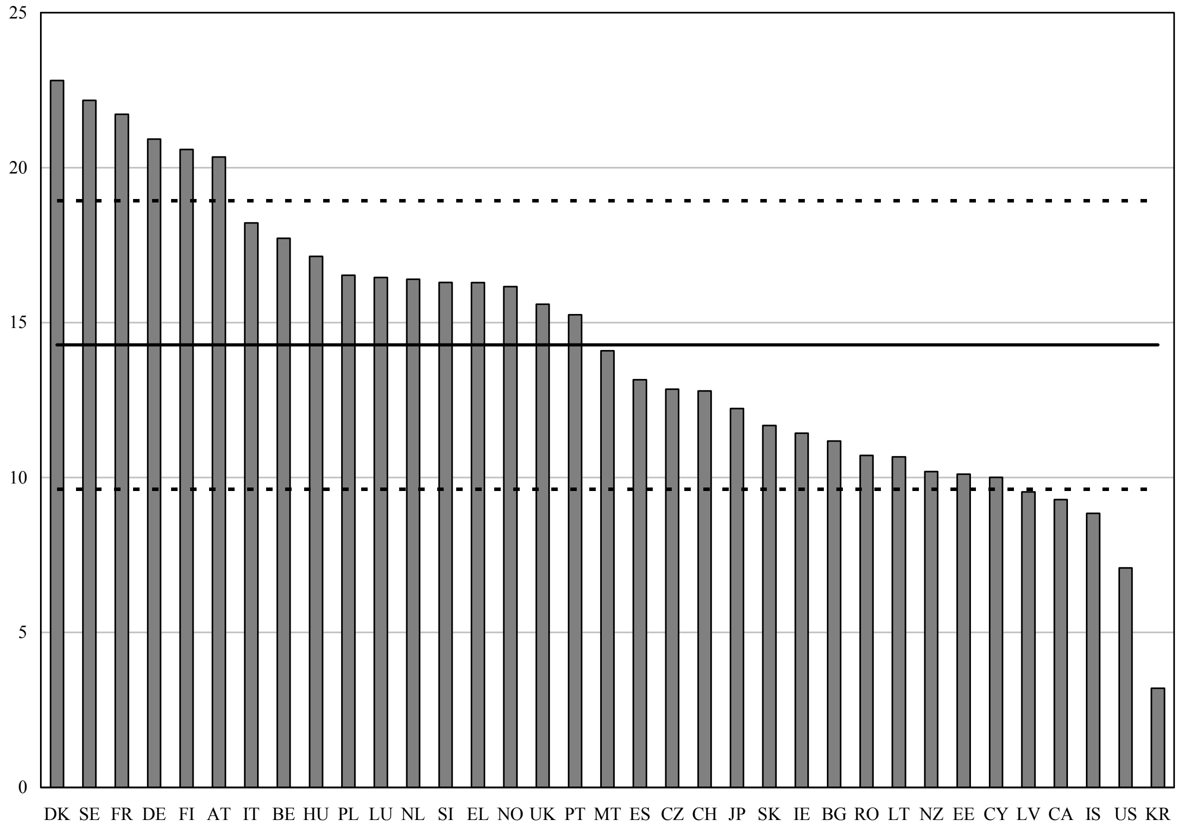
The impact of transfer payments on growth is theoretically ambiguous. On the one hand, redistributive spending may be long-run growth-enhancing if it helps to support and maintain social

[11] In some countries new modes of financing infrastructures by Public-Private-Partnerships or outsourcing may also have contributed to a decline in government investment figures. For an empirical analysis of economic and political factors affecting government investment spending in Europe, see Kappeler and Välilä (2008) or Pitlik (2010).

Figure 7

Government Spending on Social Protection

(averages 2004-08; percent of GDP)



Source: EUROSTAT, OECD, and WIFO calculations.

peace, correct labor market failures or enters as input in private production. Lindert (2004), for example, claims that social welfare spending is almost a "free lunch" without (net) growth deterring effects. Properly designed capital transfers to enterprises may also stimulate growth by promoting private investment. On the other hand, redistributive spending will inhibit growth as it generates disincentives for potential recipients, or stimulate socially unproductive rent seeking (e.g., Murphy, Shleifer and Vishny, 1991). Empirical evidence shows mixed results, although studies that find negative effects of government transfers on economic growth appear to dominate (see e.g., Romero-Ávila and Strauch, 2008, but see also Afonso and Furceri, 2010). Government spending that is predominantly redistributive is generally categorized as non-productive.

Figure 7 displays spending on social protection affairs. It includes cash benefits as well as transfers-in-kind and government services for social protection purposes.[12] Spending on these issues is 20 per cent of GDP or more in Denmark, Sweden, France, Germany, Finland and Austria, whereas Korea, the USA, Iceland, Canada and Latvia spend less than 10 per cent of GDP on social protection. Average government expenditure in the sample is 14.3 per cent of GDP.

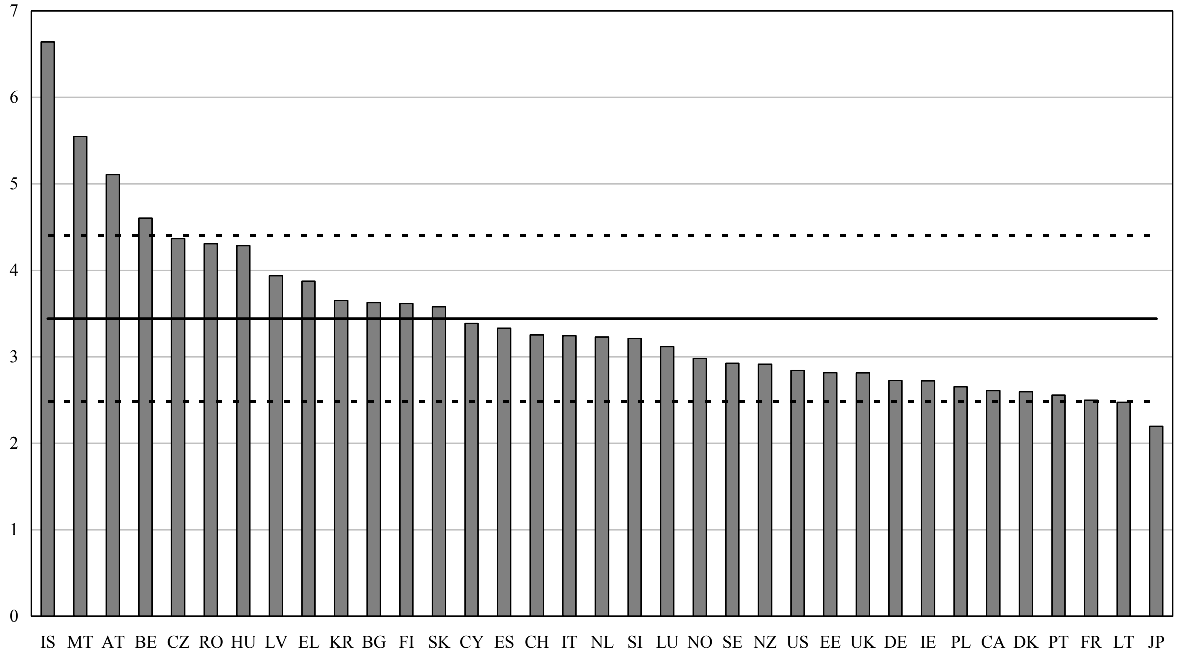
---

[12] Note that this classification does not include health care spending as in the European System of integrated Social Protection Statistics (ESSPROS) categorization of social protection spending.

Figure 8

Government Spending on Economic Affairs

(infrastructure investment deducted; averages 2004-08; percent of GDP)



Source: EUROSTAT, OECD, and WIFO calculations.

A second type of redistributive spending takes the form of sectoral aid for private enterprises. Figure 8 illustrates that average government support over the years 2004-08 was by far highest in Iceland, amounting to almost 7 per cent of GDP. This is, however, due to Iceland's special aid during the banking crisis of 2008, which boosted spending from 3.7 per cent of GDP (2007) to 16.9 per cent.[13] Malta and Austria offer support slightly above 5 per cent of GDP. The average spending level in the sample is 3.4 per cent of GDP. Relatively little support is given by Japan, with slightly more than 2 per cent of GDP.

2.2.3 The overall growth friendliness of government spending

So far, our investigations show that governments in our sample follow very different spending patterns. In particular, we observe clear differences considering the "budget mix" of productive and non-productive expenditure. Table 2 sheds some light on this. In order to investigate the "overall" growth friendliness of a country's spending patterns we simply calculate the share of productive expenditure types (according to our definitions) in total government spending. We use again averages over the years 2004-08 in order to reduce the impact of temporary fluctuations due to singular events. As the general productivity of defense spending is the most controversially debated topic, we differentiate between two definitions of productive expenditures, the first including, and the second excluding military spending. The countries are ranked in order of productive spending without defense.

[13] If the 2008 figure is not used for calculation of the mean, then the Iceland figures drop to 4.1 per cent of GDP.

Table 2

Total Spending and Productive Spending Shares
(averages 2004-08)

Country	Code	Total *(percent of GDP)*	Productive *(percent of total exp.)*	Productive (w/o defense) *(percent of total exp.)*
Korea	KR	27.9	69.1	60.1
New Zealand	NZ	38.9	60.0	57.2
Ireland	IE	36.3	56.0	54.7
Latvia	LV	36.8	57.7	53.8
Cyprus	CY	42.1	57.7	53.3
United States	US	36.8	64.9	53.3
Canada	CA	39.6	55.7	53.1
Iceland	IS	45.5	53.1	53.0
Lithuania	LT	34.5	56.9	52.4
Estonia	EE	35.1	56.4	52.4
Czech Republic	CZ	43.9	55.3	52.3
Japan	JP	36.9	53.8	51.2
Bulgaria	BG	38.0	55.3	50.8
Portugal	PT	44.8	51.0	48.1
Spain	ES	39.2	49.9	47.1
Slovenia	SI	44.5	49.8	46.6
Slovakia	SK	36.3	51.0	46.6
Romania	RO	35.5	52.1	46.0
United Kingdom	UK	44.5	51.6	46.0
Netherlands	NL	45.5	48.8	45.7
Norway	NO	41.9	48.5	44.5
Malta	MT	44.1	46.2	44.4
Poland	PL	43.1	46.9	44.3
Luxembourg	LU	39.1	44.6	44.0
Sweden	SE	52.7	47.0	43.9
Hungary	HU	49.9	45.7	43.3
France	FR	52.9	46.6	43.1
Finland	FI	49.2	45.3	42.3
Belgium	BE	49.8	44.4	42.2
Denmark	DK	52.3	44.7	41.7
Italy	IT	48.3	44.1	41.2
Switzerland	CH	33.8	44.0	41.2
Austria	AT	50.5	41.4	39.6
Greece	EL	46.2	44.7	38.4
Germany	DE	45.3	40.3	38.0

Source: WIFO calculations based on Eurostat and OECD.

Figure 9

Total Spending and Productive Spending Shares (Without Defense Spending) in Total Spending
(averages 2004-08, percent of GDP)

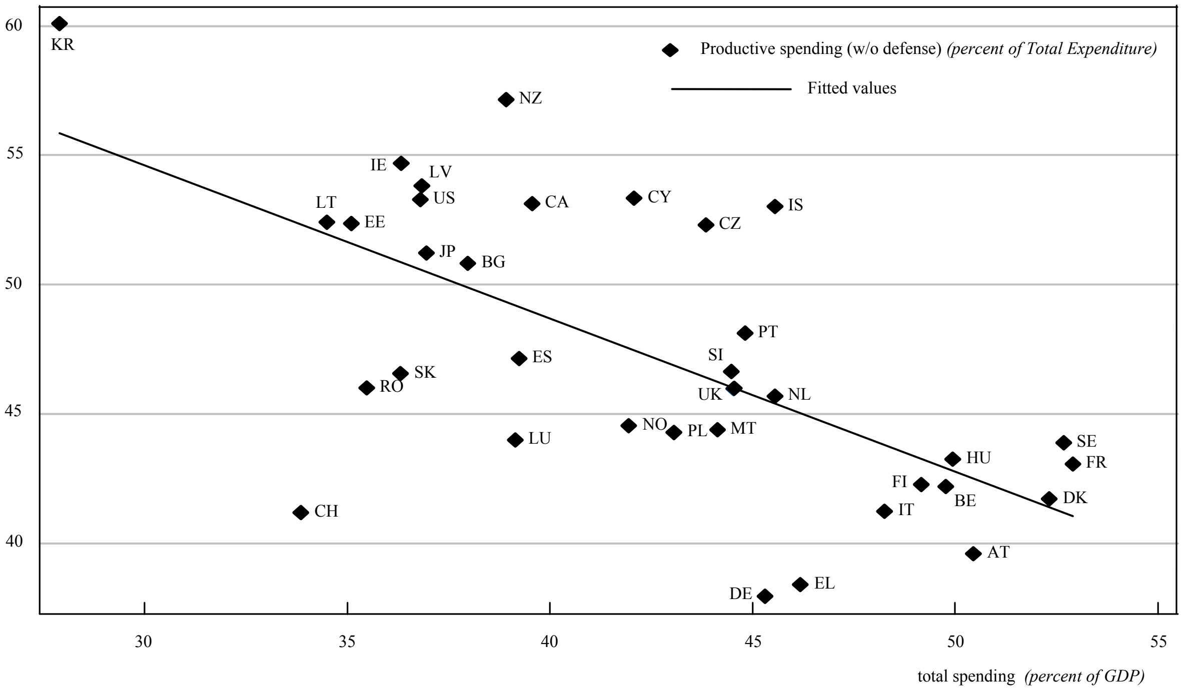


Figure 9 illustrates that there is in general a negative relation between total government spending and productive expenditure shares (without military spending).[14] This is an indication that expansion of government size is mainly due to non-productive spending items.

3 Taxation

The highest budget share of productive spending items is observed for Korea, according to both definitions. Almost 70 per cent of general government expenditure is allocated to productive uses if defense is included, and still more than 60 per cent if defense spending is counted as non-productive. New Zealand and Ireland follow, with a productive spending budget share of 57.2 per cent and 54.7 per cent, respectively. At the lower end of the ranking we find Germany, Greece and Austria with productive budget shares of slightly less than 40 per cent, if military expenditures are excluded. The largest change of productive spending shares when defense spending is included is observed for the USA (+11.6 percentage points), Korea (+9), Greece (+6.3) and Romania (+6.1).

Taxes are the most important revenue source for governments to finance their expenditures. Particularly with the advancement of endogenous growth models implying – in contrast to neoclassical growth theory – that tax policy is able to impact on the long-run growth level itself and not only on the growth rate during the transition of the economy to the steady-state growth rate, the

[14] Results are almost identical if defense spending is included.

relationship between taxes and economic growth has attracted increasing attention. Against the background of the significant increases of public deficits and debt many countries affected by the recent financial and economic crisis are experiencing, the growth friendliness of tax increases to consolidate public budgets currently is of particular interest and an important element of the policy recommendations of the supranational organisations (e.g., European Commission, 2010a, or OECD, 2010a).

3.1 Growth-friendly tax systems: Theoretical background

Physical and human capital, labor supply and technological progress are the crucial determinants of long-run economic growth. To the extent to which taxes influence these growth determinants, they impact on long-run growth. While taxes on capital may dampen savings of private households and firms' investments as well as their innovative activities, taxes on labor may decrease labor supply and demand and adversely affect incentives to invest in human capital. These distortionary effects and disincentives for economic activities of private households and firms may be aggravated by an increasing international integration of goods and factor markets, as a comparatively high tax burden may drive economic activities abroad or may be detrimental for a country's attractiveness for foreign investment or qualified labor (Afonso *et al.*, 2005, Handler *et al.*, 2005).

As, however, the existing theoretical models trying to depict the relationships between taxes and growth or growth-relevant factors, respectively, do not always yield clear-cut results,[15] an increasing number of econometric analyses attempt to tackle this complex question empirically. Therefore in the last three decades an ever-increasing number of empirical studies investigated the influence of taxation on economic growth.[16]

3.2 Growth-friendly tax systems: empirical results

Initially empirical analyses focused on the growth effects of the total level of taxation. However, they only partially support the theoretical expectation of a significant (negative) relationship between the total tax burden and economic growth: Endogeneity problems, the neglect of growth-enhancing expenditures financed by tax revenues, the disregard of taxation structures as well as statistic/conceptual problems in defining the tax ratio limit the explanatory power of the existing empirical studies (Arnold, 2008; Myles, 2009; European Commission, 2010A). The only safe conclusion that may be drawn from the existing empirical evidence is that a high tax ratio does not impact positively on growth (Afonso *et al.,* 2005).

Lately the potential growth impact of the tax structure has attracted more attention than the pure level of the tax burden. The starting point of this more recent empirical work is the assumption – also warranted by theoretical considerations – that different tax categories affect growth with differing intensity and via different channels. In the meantime, a rather large body of empirical analyses has emerged. Most authors focus on growth-relevant effects of specific taxes in a more or less isolated perspective, only few studies examine the growth implications of different tax categories in a comparative perspective.[17]

---

15 For example, it is not clear *ex ante* whether an increase of labor taxes increases or decreases labor supply, as it will have both an income and a substitution effect running in the opposite direction.

16 For recent overviews over relevant empirical work see Schratzenstaller (2007), European Commission (2008) or Myles (2009).

17 Mostly these studies analyse the growth effects of distortionary versus non-distortionary taxes, e.g., Bleaney, Gemmell and Kneller (2001) or Kneller, Bleaney and Gemmell (1999).

Of the latter, a rather recent study by a group of economists associated with the OECD (Johannson *et al.*, 2008) has achieved some prominence and gained considerable attention also among policy-makers. Based on a macroeconomic perspective, a hierarchy of individual taxes with respect to their growth friendliness is derived. Taxes on property have the least growth-dampening effect, followed by taxes on consumption (including environmental taxes in particular). In comparison, personal income taxes (including social security contributions and payroll taxes) are more harmful, and corporate income taxes are most detrimental to growth. This suggests that tax systems relying more on property and consumption taxes display more favourable growth properties than those strongly based on personal and corporate income taxes.

A crucial advantage and the innovative aspect of this approach is that it does not direct an isolated focus on the effects of single tax categories but on the effects of a (revenue-neutral) trade-off between them. However, that the macroeconomic tax structure is of limited use as an indicator for the effective tax burden on individual tax bases, because it does not account for the structure of the overall tax base. Moreover, marginal tax rates shaping incentives for economic decisions of private households and firms are neglected. Thus, an analysis of the tax structure of a given country also include macroeconomic effective tax rates reflecting the distribution of total tax revenues as well as microeconomic (marginal and average) tax rates influencing individual behaviour of private households and firm decisions. Moreover, a complementary look at studies examining growth-relevant effects of individual tax categories certainly is useful to gain deeper insights regarding the concrete channels via which individual tax categories may directly or indirectly impact on economic growth. Two aspects are of particular interests in this respect: namely, the influence of corporate income taxes on firm decisions and of labor taxes on labor supply.

While labor taxes can be assumed to influence various individual decisions shaping the quality and quantity of labor supply (employment in the shadow economy or in non-taxed sectors of the economy, investment in human capital, occupational choices, individual work effort and productivity, etc.), their effect on labor market participation and hours worked has been investigated most intensely and with the most robust results. These can be summarized as follows:[18]

- the influence of labor taxes differs for different demographic groups and educational levels due to differing wage elasticities of labor supply;
- for some groups – e.g., mothers with young children – labor taxes strongly impact on the decision about participation and hours worked;
- the participation decision is rather tax sensitive in the group of lone mothers and men with low qualifications;
- participation as well as hours worked of men in general and highly-qualified men in particular hardly react to labor tax variations.

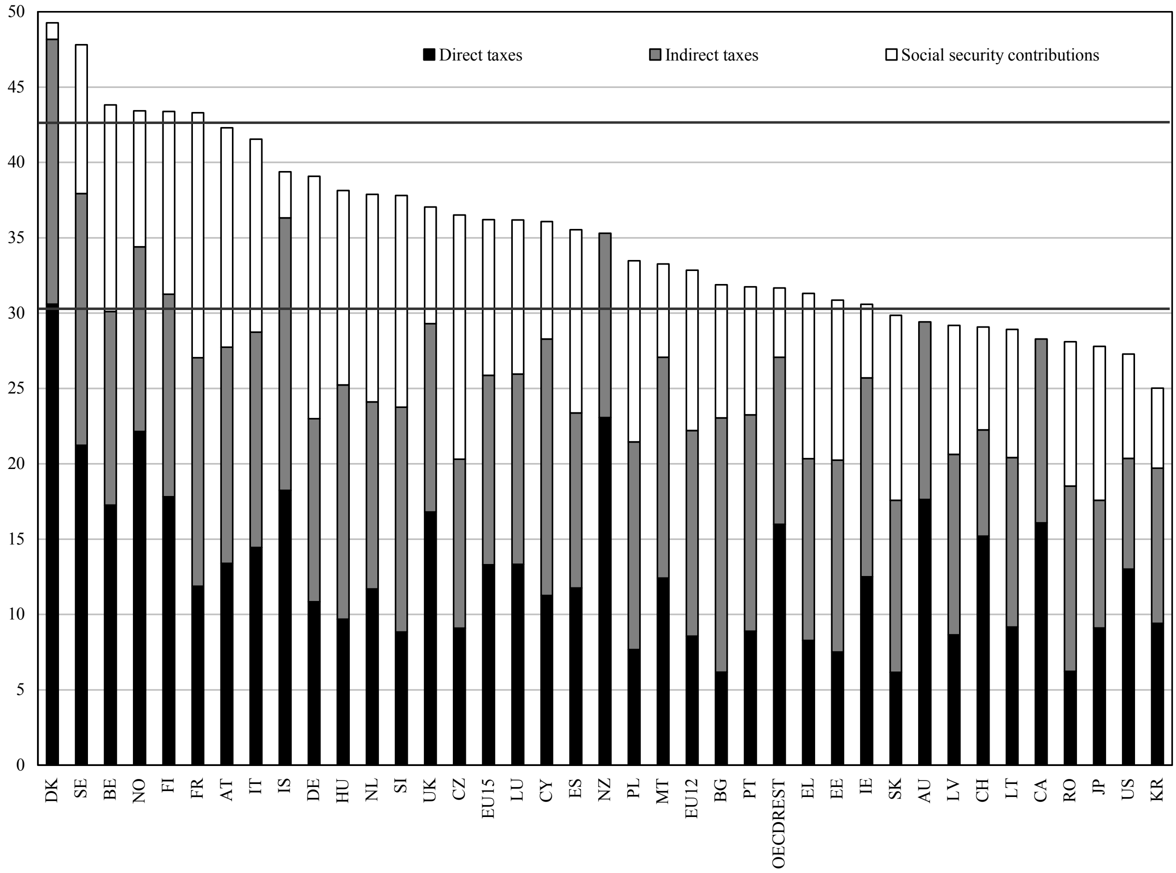
Corporate income taxes influence firm behaviour in various respects. In a rather recent review of the rich empirical evidence, including a meta analysis of studies investigating the influence of taxation on international investment, de Mooij and Ederveen (2008) authors reach the conclusion that the largest tax-base elasticities can be found in empirical studies on profit shifting. Also marginal investment displays a significant elasticity with respect to EMTR, and even more so discrete location decisions.

[18] For the following short summary see the extensive literature reviews by Meghir and Phillips (2008) or Task Force of the Monetary Policy Committee of the European System of Central Banks (2008).

Figure 10

Tax-to-GDP Ratios

(averages 2004-08, percent of GDP)



Sources: European Commission (2011), OECD (2010), *Revenue Statistics 2010*, and WIFO calculations. EU12: new members. OECD: sample countries which are not EU members.

3.3 *Size and structure of taxation*

As already indicated, there are different types of indicators that may be used to measure and evaluate the growth friendliness of tax systems. While the macroeconomic tax structure (*i.e.*, the shares of individual tax categories in total tax revenues or over GDP) can give a first impression concerning (potentially unfavourable) overall tax structures, macroeconomic effective tax rates are required to measure the distribution of the overall tax burden on the respective macroeconomic tax bases. Incentives influencing growth-relevant decisions by firms and individuals are affected by effective microeconomic tax rates.

3.3.1 *Total tax burden and macroeconomic tax structure*

Figure 10 shows the total tax burden (including social security contributions) in percent of GDP (the most common indicator for the overall tax level) for the sample of 36 countries as five-year averages for the period 2004 to 2008. We group – somewhat arbitrarily – the countries regarded in high-tax countries (tax burden above 42 per cent of GDP),[19] in low-tax countries (tax burden below 30 per cent of GDP)[20] and in a group with a medium tax burden (between 30 per cent and 42 per cent of GDP).[21] The country-specific values cover a wide range, from 25 per cent of GDP in South Korea to 49.3 per cent of GDP in Denmark. The average tax level for the rest-OECD countries included in our sample amounts to 31.7 per cent of GDP, for the EU15 countries the average is 36.2 per cent and for the EU12 countries 32.8 per cent.

In a first rough categorization, total tax revenues can be grouped into three main categories: indirect taxes, direct taxes, and social security contributions. Related to GDP, direct taxes dominate on average for the rest-OECD countries in our sample, with 16 per cent; indirect taxes reach 11.1 per cent (see Figure 10). Social security contributions are of considerably smaller significance, with 4.6 per cent of GDP on average for the rest-OECD countries regarded. In the EU12 indirect taxes are clearly dominating on average, with 13.6 per cent of GDP, followed by social security contributions with 10.6 per cent and direct taxes with 8.6 per cent of GDP. In the EU15 the shares of the respective tax categories are comparatively balanced, with direct taxes reaching 13.3 per cent, indirect taxes 12.6 per cent, and social security contributions 10.3 per cent of GDP.

Figure 10 also shows that the shares of these main tax categories in GDP vary considerable between countries. Averaged over the period 2004 to 2008, direct taxes reach 6.2 per cent of GDP in (the flax tax countries) Bulgaria, Romania and the Slovak Republic on the low end, and 30.6 per cent of GDP in Denmark on the high end. Indirect taxes range from 7.1 per cent of GDP in Switzerland to 18.1 per cent in Iceland. While social security contributions make up for 1.1 per cent of GDP in Denmark only, they amount to 16.3 per cent of GDP in France.

3.3.2 Macroeconomic effective tax rates

Macroeconomic or implicit effective tax rates relating total revenues stemming from one tax category to the corresponding tax base and thus reflecting the effective tax burden on individual tax bases are calculated regularly by Eurostat for the EU27 countries plus Iceland and Norway. Eurostat calculates implicit effective tax rates for labor, energy, consumption, and on capital (which are divided further in implicit tax rates on capital and business income and on corporate income). Table 3 contains implicit tax rates for 2000 and 2008 in comparison. On average, implicit tax rates for all macroeconomic tax bases decreased in the EU15. In the EU12, on the other hand, only implicit tax rates on labor and corporate income decreased, while they increased on consumption, energy, and capital.

A closer look at developments in individual countries reveals that they are differently affected by these general trends: Firstly the extent to which tax burdens have changed during the last decade varies considerably across countries. Secondly, about one third of the EU countries regarded are moving against the general trends with regard to implicit tax burdens on labor, capital, and corporate income; in about one fourth of the EU countries analyzed here the implicit tax rate on energy and in half the EU countries the implicit consumption tax rate went down.

---

19 This corresponds approximately to the mean tax ratio plus one standard deviation (41.4 per cent); the resulting group of 8 high-tax countries includes Denmark, Sweden, Belgium, Norway, Finland, France, Austria, and Italy.

20 This corresponds approximately to the mean tax ratio minus one standard deviation (29.1 per cent); the 10 low-tax countries are the Slovak Republic, Australia, Latvia, Switzerland, Lithuania, Canada, Romania, Japan, the United States and South Korea.

21 This is the biggest group with 18 countries, consisting of Iceland, Germany, Hungary, the Netherlands, Slovenia, the United Kingdom, the Czech Republic, Luxembourg, Cyprus, Spain, New Zealand, Poland, Malta, Bulgaria, Portugal, Greece, Estonia, and Ireland.

Table 3

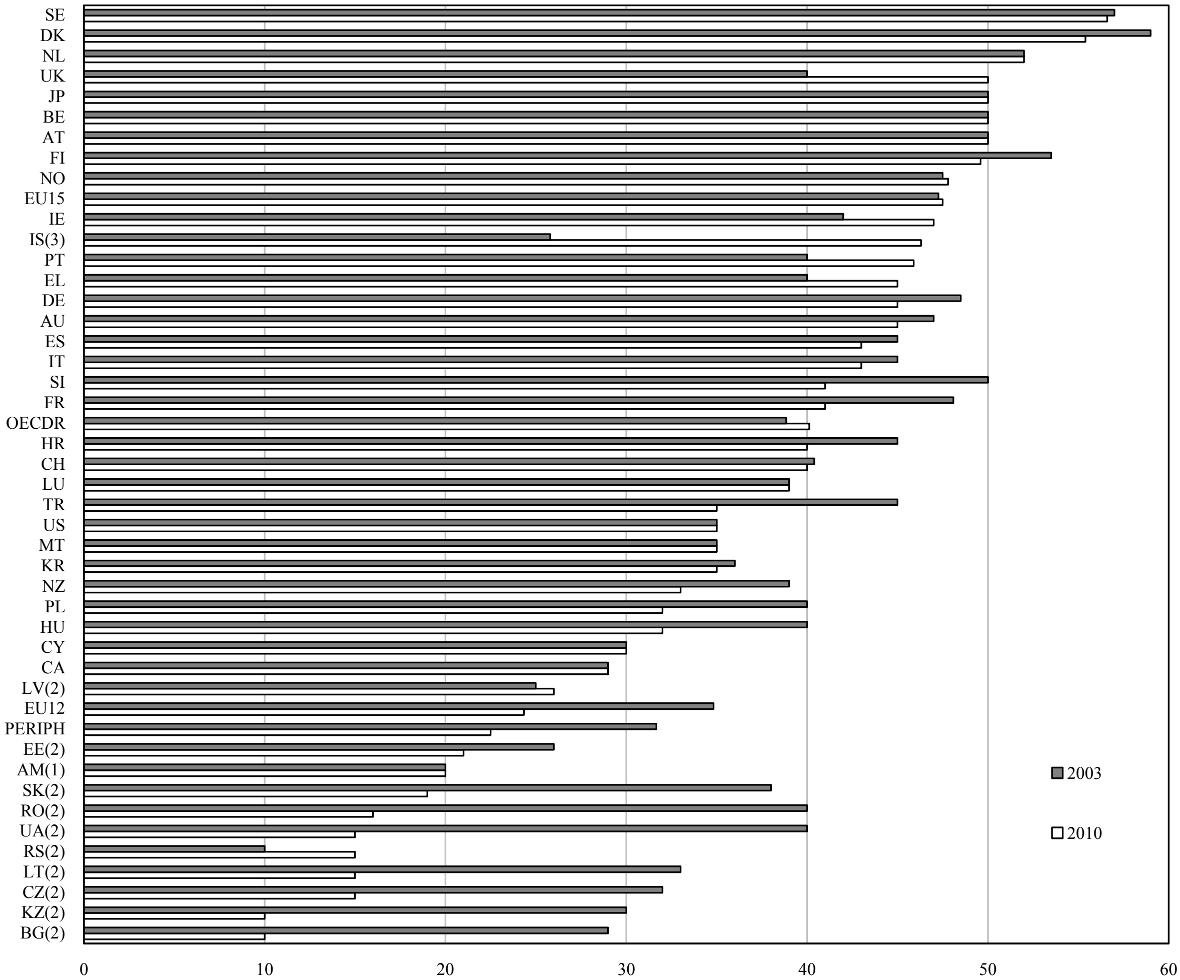
Implicit Tax Rates on Labor, Consumption, Energy, Capital, Corporate Income, EU 27, 2000-08

Country	Labor			Consumption			Energy[1]			Capital			Corporate Income		
	2000	2008	Δ 2000-08	2000	2008	Δ 2000-08	2000	2008[2]	Δ 2000-08	2000[3]	2008[4]	Δ 2000-08	2000[5]	2008[6]	Δ 2000-08
BE	43.6	42.6	–1.0	21.8	21.2	–0.6	92.4	97.1	4.7	29.6	32.7	3.1	24.4	21.4	–3.0
BG	38.7	27.6	–11.1	19.7	26.4	6.8	36.4	71.7	35.3	-	-	-	-	-	-
CZ	40.7	39.5	–1.2	19.4	21.1	1.7	55.2	127.1	71.9	20.9	21.5	0.6	26.2	25.7	–0.5
DK	41.0	36.4	–4.5	33.4	32.4	–1.0	300.8	267.8	–33.1	36.0	43.1	7.1	23.0	24.9	1.9
DE	40.7	39.2	–1.6	18.9	19.8	0.9	192.7	193.8	1.1	28.4	23.1	–5.3	-	-	-
EE	37.8	33.7	–4.1	19.5	20.9	1.5	32.2	71.5	39.3	6.0	10.7	4.8	4.1	8.3	4.3
IE	28.5	24.6	–3.9	25.7	22.9	–2.8	140.5	153.1	12.5	14.9	15.7	0.8	10.0	7.6	–2.4
EL	34.5	37.0	2.5	16.5	15.1	–1.4	117.3	102.0	–15.3	19.9	15.8	–4.1	29.0	18.6	–10.4
ES	28.7	30.5	1.9	15.7	14.1	–1.6	137.8	114.6	–23.2	29.8	32.8	3.0	30.7	34.0	3.3
FR	42.0	41.4	–0.6	20.9	19.1	–1.8	173.2	160.7	–12.5	38.3	38.8	0.4	29.6	29.1	–0.5
IT	42.2	42.8	0.6	17.9	16.4	–1.5	248.7	187.4	–61.3	29.5	35.3	5.8	19.2	31.5	12.3
CY	21.5	24.5	2.9	12.7	20.6	7.8	43.1	110.0	66.9	23.7	36.4	12.6	28.6	37.3	8.7
LV	36.7	28.2	–8.4	18.7	17.5	–1.2	48.3	48.4	0.1	11.2	16.3	5.1	8.6	15.2	6.6
LT	41.2	33.0	–8.2	18.0	17.5	–0.4	58.0	78.5	20.5	7.2	12.4	5.2	3.9	11.1	7.1
LU	29.9	31.5	1.6	23.0	27.1	4.1	164.3	173.3	9.0	-	-	-	-	-	-
HU	41.4	42.4	1.0	27.5	26.9	–0.6	79.7	98.0	18.3	17.1	19.2	2.0	28.7	19.9	–8.8
MT	20.6	20.2	–0.4	15.9	20.0	4.1	142.2	197.0	54.9	-	-	-	-	-	-
NL	34.5	35.4	0.9	23.8	26.7	2.9	154.4	189.8	35.3	20.8	17.2	–3.7	18.5	11.9	–6.6
AT	40.1	41.3	1.2	22.1	22.1	0.0	141.8	150.2	8.4	27.7	27.3	–0.3	27.1	26.1	–1.0
PL	33.6	32.8	–0.8	17.8	21.0	3.2	58.9	108.0	49.0	20.5	22.5	2.0	37.1	20.0	–17.1
PT	27.0	29.6	2.7	18.9	19.1	0.2	111.8	143.4	31.6	33.6	38.6	5.0	25.5	22.6	–2.9
RO	33.5	29.5	–4.0	17.0	17.7	0.7	58.2	26.2	–32.0	-	-	-	-	-	-
SI	37.7	35.7	–2.0	23.5	23.9	0.4	118.3	121.7	3.4	15.7	21.6	5.9	19.6	27.4	7.7
SK	36.3	33.5	–2.8	21.7	18.4	–3.3	42.4	84.6	42.2	22.9	16.7	–6.2	40.2	20.7	–19.4
FI	44.1	41.3	–2.7	28.5	26.0	–2.5	108.7	114.5	5.8	36.1	28.1	–7.9	30.4	19.3	–11.1
SE	46.0	42.1	–3.8	26.3	28.4	2.2	182.0	190.1	8.1	43.2	27.9	–15.3	41.0	23.2	–17.8
UK	25.3	26.1	0.7	18.9	17.6	–1.4	249.5	180.2	–69.3	44.7	45.9	1.2	31.0	22.2	–8.8
EU 15	36.5	36.1	–0.4	22.1	21.9	–0.3	167.7	161.2	–6.5	30.9	30.2	–0.7	26.1	22.5	–3.6
EU 12	35.0	31.7	–3.3	19.3	21.0	1.7	64.4	95.2	30.8	16.1	19.7	3.6	21.9	20.6	–1.3

[1] Energy taxes in Euro per tons of oil equivalent (TOE), base year: 2000; [2] Iceland 2006; Greece, France, Malta 2007; [3] Ireland 2002; [4] Greece 2006, Norway 2007; [5] Ireland 2002; [6] Greece, Portugal 2006.
Source: European Commission (2010b), and WIFO calculations.

Figure 11

Personal Income Tax Rate, 2003-10



Source: KPMG (2010). (1) Introduction of flat tax in 2011; (2) Flat tax; (3) Introduction of flat tax in 2007, abolished in 2010.

3.3.2 Microeconomic tax rates

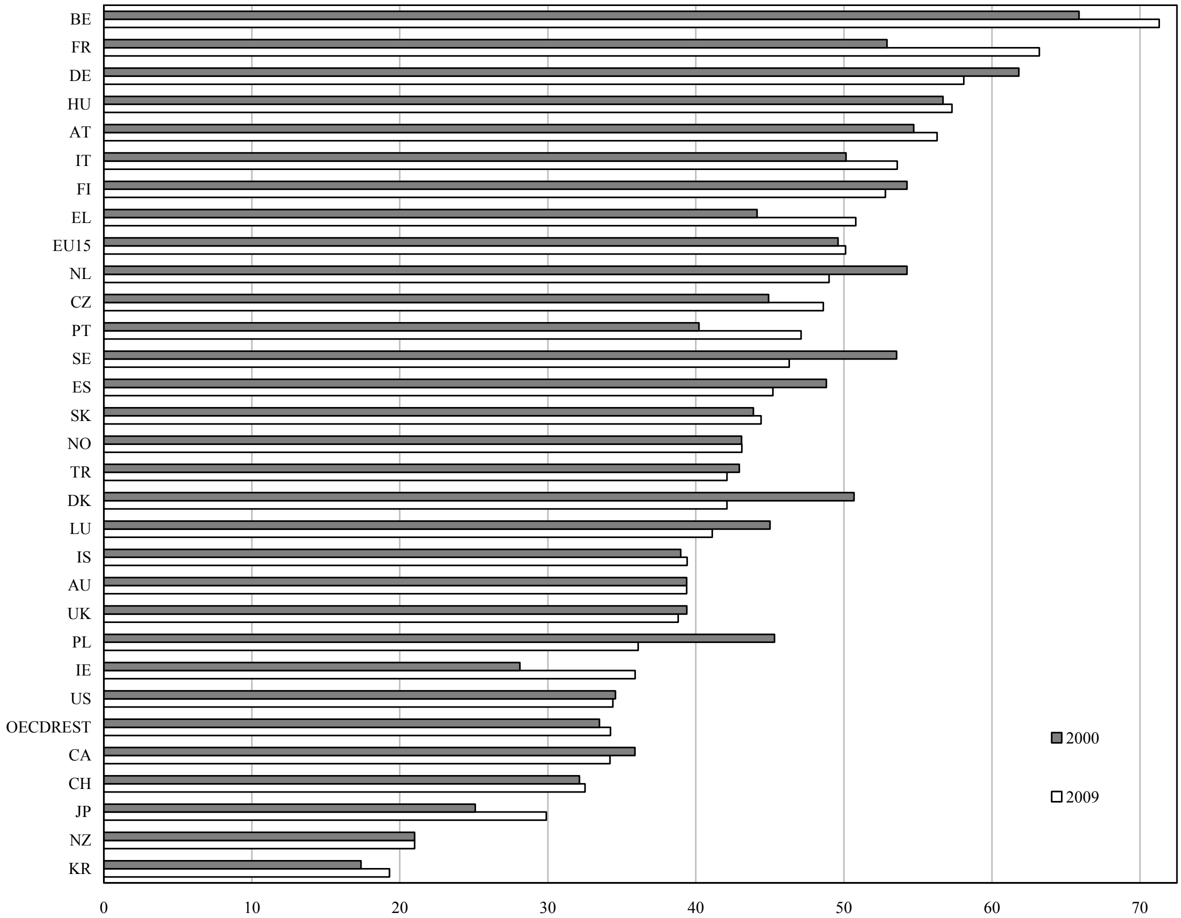
Evaluations of the effects of taxes on labor supply and investment need to be based on microeconomic tax rates. Ideally, these should be forward looking, as the tax burden of the past is of limited relevance for future decisions of economic agents about, for example, investment or labor supply.

3.3.2.1 Microeconomic tax rates on labor

We start with a look at top income tax rates for our sample of 36 countries, which we enrich by 6 peripheral European countries (Croatia, Ukraine, Kazakhstan, Armenia, Republic of Serbia, Turkey). Between 2003 and 2010, a clear downward trend of personal income tax rates can be observed for the EU12 and the peripheral European countries, where the average top income tax rate went down from 34.8 per cent in 2003 to 24.3 per cent in 2010 and from 31.7 to 22.5 per cent,

Figure 12

Marginal Tax Wedge, 67 per cent of Gross Labor Income, 2000-09



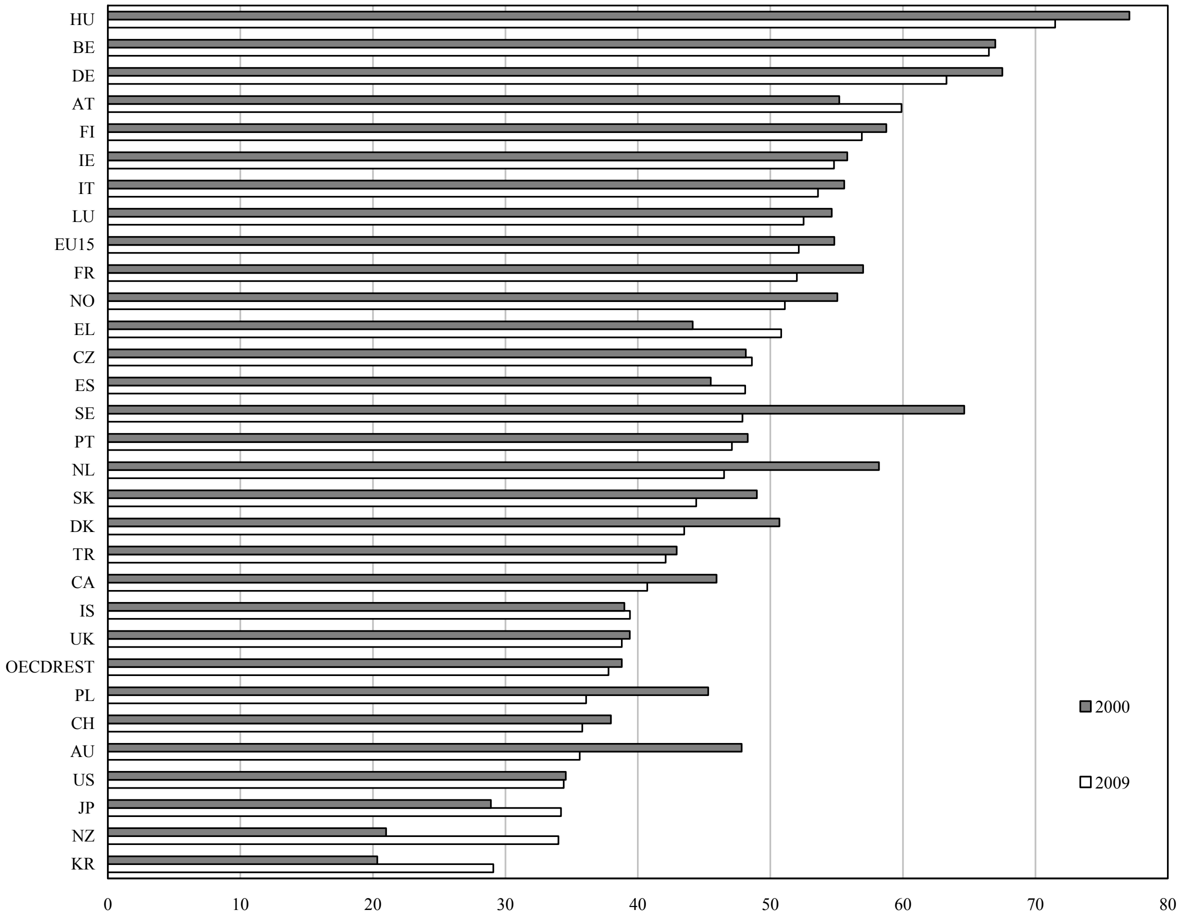
Source: OECD (2011).

respectively. In the EU15 countries, on the other hand, top income tax rates stagnated on average, amounting to 47.5 per cent in 2010. In the rest-OECD countries analyzed here the average top income tax rate increased from 38.9 to 40.1 per cent.

To assess the incentive effects of personal income taxation with regard to labor supply, a focus on top personal income tax rates is far too narrow, however. Firstly, tax sensitivity of labor supply of workers in the top income groups – as the results of the overwhelming majority of empirical studies reported above show – is rather limited; tax elasticity is much higher in lower income groups. Secondly, marginal tax rates are important for decisions about the numbers of hours worked; the participation decision, however, is influenced by average tax rates which also take into account the rules to determine the tax base. Thirdly, to identify the incentive effects of taxation for labor supply all relevant taxes need to be considered: As can be seen in the macroeconomic data above, the majority of countries do not only levy wage taxes, but also social security contributions on labor incomes. Thus, to derive a more complete picture of the possible incentive effects of labor taxation, effective marginal as well as average microeconomic tax rates for different income groups with different tax rate elasticities of labor supply must be determined, which include personal income taxes as well as social security contributions.

Figure 13

Marginal Tax Wedge, 100 per cent of Gross Labor Income, 2000-09



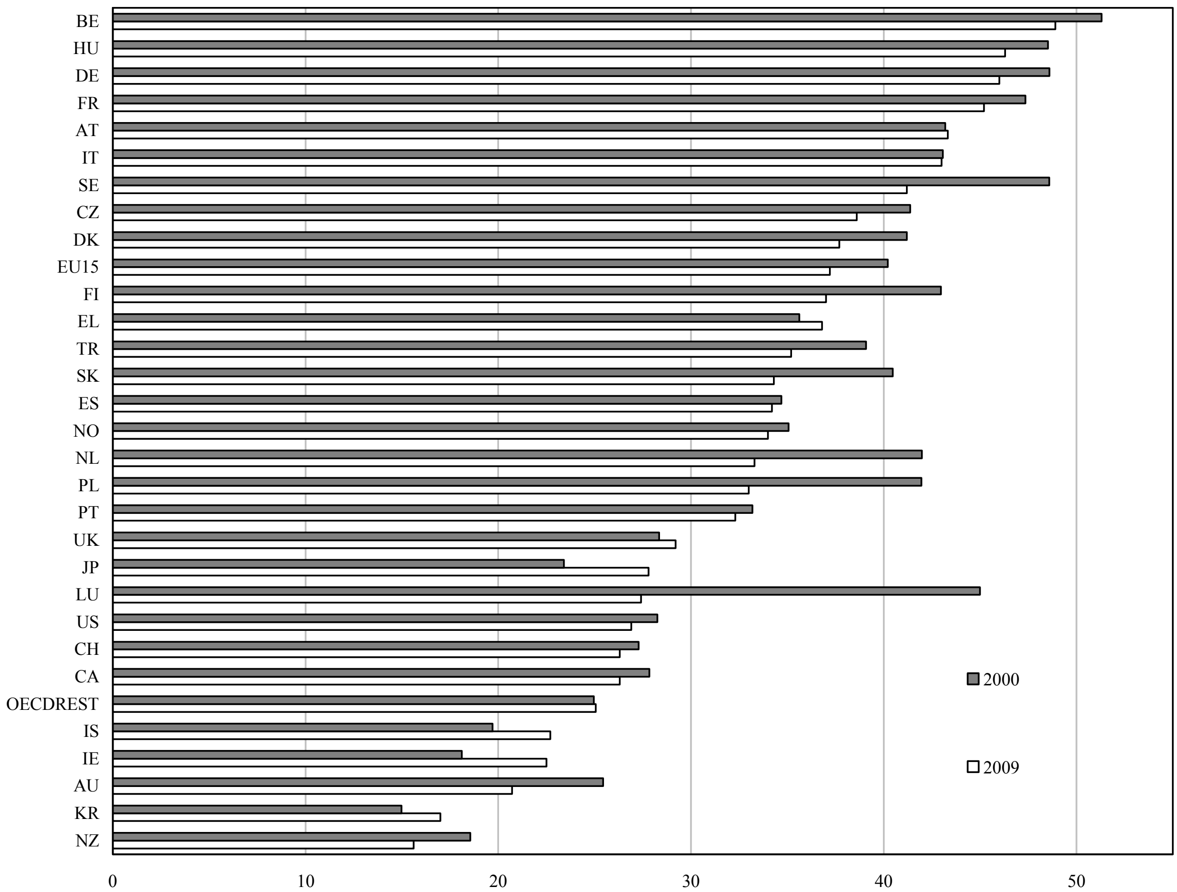
Source: OECD (2011).

Effective marginal and average tax wedges including personal income taxes and social security contributions are calculated regularly by the OECD. For sake of complexity reduction, we choose from the considerable selection of family constellations and income sizes the OECD offers two simple cases: a single earner with 67 per cent of an income (as representative for a rather low income group), and a single earner with an average income. In Figures 12 to 15, marginal and average tax wedges (resulting from wage tax and social security contributions minus cash benefits), respectively, are presented in comparison for the years 2000 and 2009.

For low income earners, in the EU15 the marginal tax wedge slightly rose on average between 2000 and 2009, to a rather high level of 50.1 per cent: Thus it approached the marginal tax rate for an average earner, who faced a marginal tax wedge of 52.1 per cent in 2009 (compared to 54.8 per cent in 2000). The marginal tax wedge for low incomes was lowest in South Korea (19.3 per cent) and highest in Belgium (71.3 per cent). Average incomes were burdened with the lowest marginal tax wedge in South Korea (29.1 per cent) and with the highest marginal tax wedge in Hungary (71.5 per cent). The average tax wedge for the EU15 went down by about 3 percentage points both for low incomes (to 37.2 per cent) and average incomes (to 41.6 per cent). The average tax wedge for low and for average incomes was lowest in New Zealand (15.6 and 18.4 per cent,

Figure 14

Average Tax Wedge, 67 per cent of Gross Labor Income, 2000-09



Source: OECD (2011).

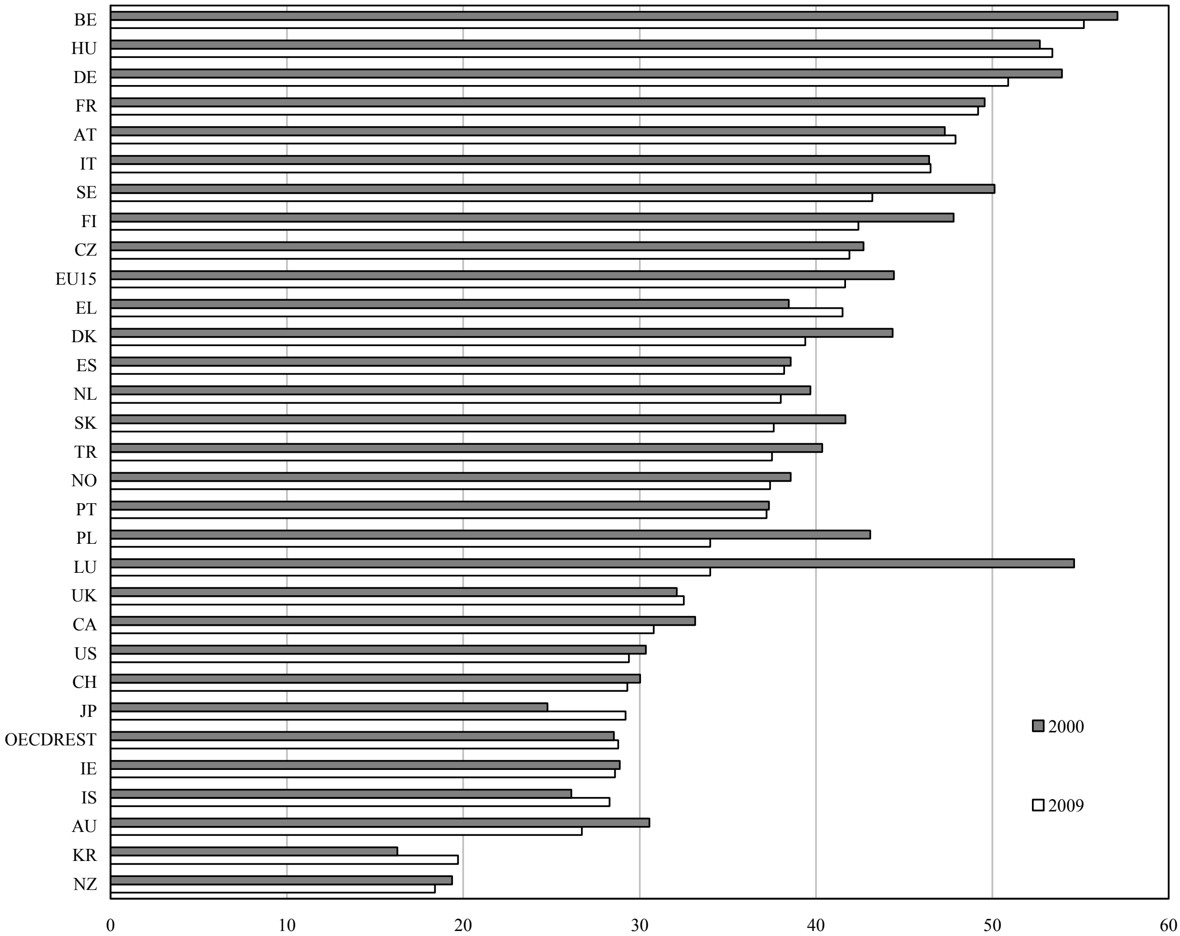
respectively). Low as well as average incomes faced the highest average tax wedge in Belgium (48.9 and 55.2 per cent, respectively). Interestingly, during the past decade the marginal tax wedge for low incomes went down in only about half the countries regarded, while the marginal tax wedge for average incomes as well as the average tax wedges for low and average incomes went down in a clear majority of countries.

3.3.2.2 Microeconomic corporate income tax rates

As mentioned above, a number of recent empirical studies corroborate the theoretical expectation that firm decisions – also in an international context – are influenced by corporate taxation. Hereby statutory corporate income tax rates as well as effective marginal (EMTR) and average (EATR) tax rates are relevant. Figure 16 shows that in our sample of 36 countries plus 10 peripheral European countries statutory corporate income tax rates fell markedly between 1995 and 2010. Only one country (Finland) slightly increased its corporate income tax rate, in 6 other countries (among them the 3 peripheral countries Montenegro, Armenia, and Belarus, but also Malta, Norway, and the United States) it remained constant. Again, the most marked reduction took place in the EU12 countries, where the average corporate income tax rate went down from 31.8 to

Figure 15

Average Tax Wedge, 100 per cent of Gross Labor Income, 2000-09



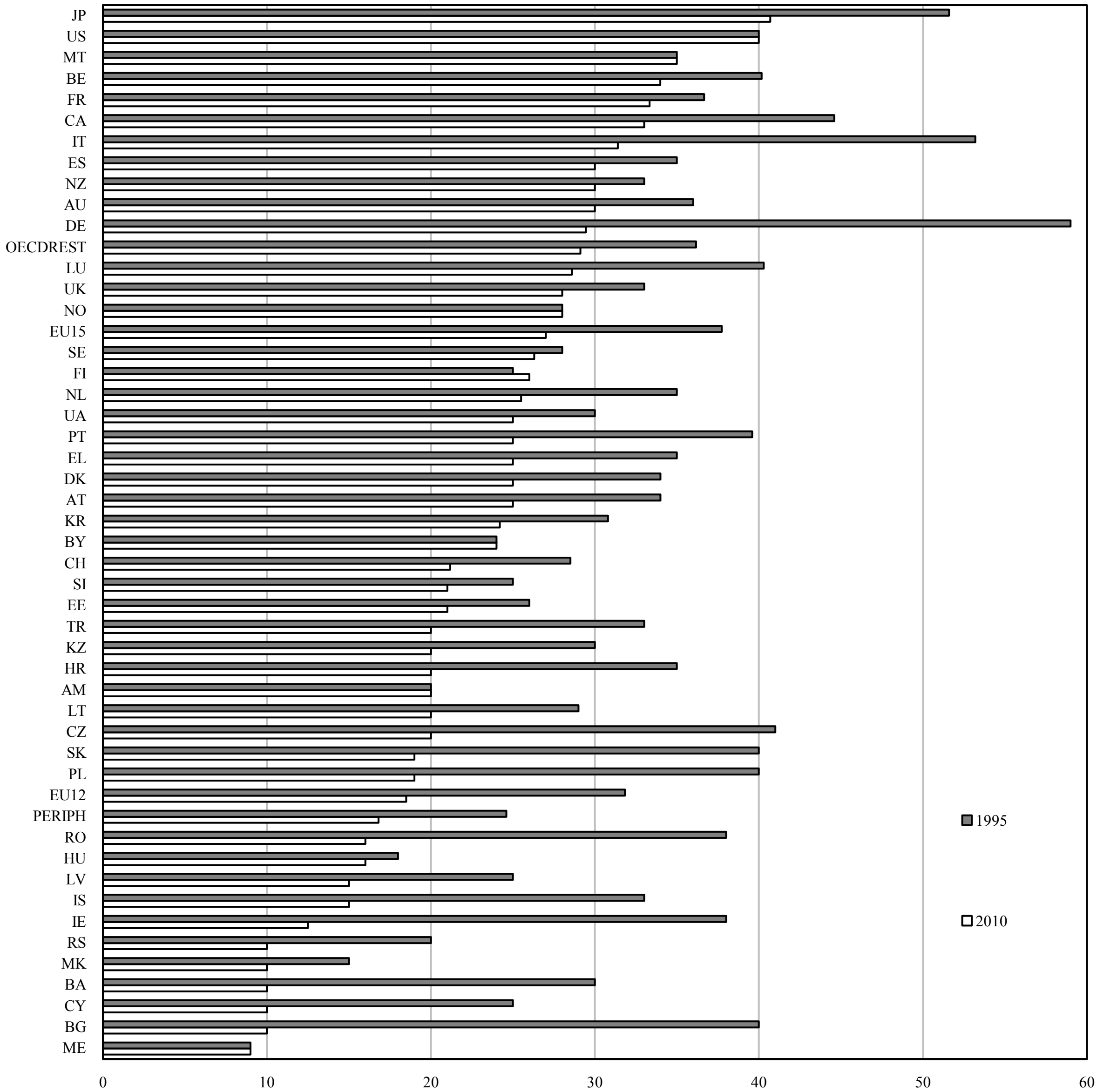
Source: OECD (2011).

18.5 per cent. But also the fall in the EU15 countries (from an average of 37.7 per cent in 1995 to 27 per cent in 2010) as well as in the European peripheral countries (from 24.6 to 16.8 per cent) is considerable. Less pronounced is the upward trend in the group of rest-OECD countries included in our sample; here the average statutory corporate income tax rate fell from 36.2 to 29.1 per cent. The distance between the high-tax and the low-tax countries narrowed down since the mid-Nineties, and while in 1995 3 countries in our sample of 46 countries had a corporate income tax rate of over 50 per cent, 2010 only 2 countries remained in which the corporate income tax rate reached about 40 per cent; it was below this threshold in all other countries.

Table 4 contains EMTR and EATR for all 27 EU countries plus 5 developed OECD countries as well as 3 European periphery countries for 2009 compared to 1998. On average EMTR and EATR were reduced in the rest-OECD countries, from 24.1 to 22 per cent and from 27.4 to 25.9 per cent, respectively. In the EU15, EMTR fell from 23.6 to 19 per cent, in the EU12 from 20.4 to 11.9 per cent. EATR went down from 30.7 to 25.1 per cent in the EU15 and from 27.4 to 17 per cent in the EU12. In this sample of 35 countries, EATR went up in 3 countries only and EMTR increased in 5 countries only; constant EATR and EMTR, respectively, can be observed in 2 identical countries.

Figure 16

Corporate Income Tax Rates, 1995-2010



Sources: KPMG (2010), and WIFO calculations. Earliest data 1995, except for Korea: 1997, Croatia, Kazakhstan, Macedonia: 1999, Serbia: 2002.

3.4 Conclusions

Table 5 gives an overview of the ranks of the countries regarded here (as far as available) with respect to the indicators presented above, whereby higher values of the tax burden indicators imply higher ranks. Of particular interest appears the relationship between the total tax burden on the one hand and the individual tax burden indicators on the other hand. However, a more detailed analysis of the relationships between the individual tax burden indicators goes beyond the scope of the study.

Table 4

Effective Average (EATR) and Marginal Corporate (EMTR) Tax Rates, 1998-2009

Country	EATR			EMTR		
	1998	2009	Δ 1998-2009	1998	2009	Δ 1998-2009
Austria	29.7	22.7	–7.0	20.2	17.4	–2.8
Belgium	34.5	24.7	–9.8	22.7	–5.1	–27.8
Bulgaria	32.0	8.8	–23.2	21.2	5.5	–15.7
Canada[1]	37.1	32.9	–4.2	38.6	32.8	–5.8
Cyprus	27.5	10.6	–16.9	24.4	9.5	–14.9
Czech Republic	26.4	17.5	–8.9	23.0	11.2	–11.8
Denmark	30.0	22.5	–7.5	21.5	16.7	–4.8
Estonia	22.4	16.5	–5.9	13.4	3.6	–9.8
Finland	25.9	23.6	–2.3	21.5	18.1	–3.4
France	39.8	34.6	–5.2	36.8	34.9	–1.9
Germany	41.2	28.0	–13.2	37.9	21.7	–16.2
Greece	30.4	21.8	–8.6	20.5	14.1	–6.4
Hungary	19.0	19.5	0.5	18.7	15.5	–3.2
Ireland	9.4	14.4	5.0	7.8	13.3	5.5
Italy	32.0	27.4	–4.6	9.7	20.8	11.1
Japan[1]	41.7	41.3	–0.4	42.8	41.9	–0.9
Latvia	22.7	13.8	–8.9	17.5	10.8	–6.7
Lithuania	23.0	16.8	–6.2	6.7	8.3	1.6
Luxembourg	32.6	25.0	–7.6	22.4	16.5	–5.9
Malta	32.2	32.2	0.0	26.9	26.9	0.0
Netherlands	32.3	23.7	–8.6	27.2	19.6	–7.6
Norway[1]	26.4	26.5	0.1	23.1	23.3	0.2
Poland	32.4	17.5	–14.9	25.3	13.7	–11.6
Portugal	33.4	23.7	–9.7	25.5	17.1	–8.4
Romania	34.0	14.8	–19.2	26.0	11.9	–14.1
Slovakia	36.7	16.8	–19.9	30.8	11.3	–19.5
Slovenia	20.9	19.1	–1.8	10.5	14.5	4.0
Spain	36.5	32.8	–3.7	35.4	33.4	–2.0
Sweden	23.8	23.2	–0.6	17.9	17.4	–0.5
Switzerland[1]	18.8	18.7	–0.1	12.5	12.4	–0.1
United Kingdom	29.7	28.3	–1.4	27.3	28.9	1.6
United States[1]	38.3	37.4	–0.9	35.9	35.1	–0.8
Croatia[1]	16.5	16.5	0.0	6.9	6.9	0.0
Macedonia[1]	13.3	7.9	–5.4	8.8	1.9	–6.9
Turkey[1]	26.8	17.9	–8.9	19.6	12.6	–7.0
EU 15	30.7	25.1	–5.7	23.6	19.0	–4.6
EU 12	27.4	17.0	–10.4	20.4	11.9	–8.5
OECD rest	27.4	25.9	–1.6	24.1	22.0	–2.0

(1) Earliest data: 2005.
Source: European Commission (2010b), and WIFO calculations.

Table 5

Country-specific Ranks with Respect to Tax Burden Indicators

Country	Total Tax Burden	Share of Growth-dampening Taxes	Top Personal Income Tax Rate [1]	Marginal Tax Wedge 100%	Average Tax Wedge 100%	Corporate Income Tax Rate	EMTR	EATR
Australia	28	13	13	24	26	8		
Austria	7	4	4	4	5	18	13	17
Belgium	3	7	5	2	1	4	32	12
Bulgaria	22	36	36	n.a.	n.a.	35	30	32
Canada	32	11	29	19	20	6	5	4
Cyprus	17	34	28	n.a.	n.a.	36	28	31
Czech Republic	15	9	34	12	9	26	26	23
Denmark	1	16	2	18	11	19	16	18
Estonia	25	25	31	n.a.	n.a.	24	31	27
Finland	5	10	8	5	8	16	12	15
France	6	14	18	9	4	5	3	3
Germany	10	8	14	3	3	11	9	8
Greece	24	22	15	11	10	20	20	19
Hungary	11	26	26	1	2	30	18	20
Iceland	9	33	11	20	25	32		
Ireland	26	31	10	6	24	34	22	29
Italy	8	15	16	7	6	7	10	9
Japan	34	1	6	26	23	1	1	1
Korea	36	35	22	28	27	22	n.a.	n.a.
Latvia	29	24	30	n.a.	n.a.	33	27	30
Lithuania	31	21	35	n.a.	n.a.	27	29	25
Luxembourg	16	19	21	8	17	12	17	11
Malta	21	32	23	n.a.	n.a.	3	7	6
Netherlands	12	17	3	16	13	17	11	13
New Zealand	19	20	25	27	28	9	na	na
Norway	4	5	9	10	15	13	8	10
Poland	20	29	27	22	18	28	21	24
Portugal	23	28	12	15	16	21	15	14
Romania	33	30	33	n.a.	n.a.	31	24	28
Slovakia	27	23	32	17	14	29	25	26
Slovenia	13	18	19	n.a.	n.a.	25	19	21
Spain	18	12	17	13	12	10	4	5
Sweden	2	3	1	14	7	15	14	16
Switzerland	30	6	20	23	22	23	23	22
United Kingdom	14	27	7	21	19	14	6	7
United States	35	2	24	25	21	2	2	2

[1] Out of a sample of 28 countries.
Source: WIFO.

4 Regulation

4.1 The regulatory framework and economic growth

A further dimension of government size is the intensity of regulation. Governments provide the framework for market transactions by setting the rules for voluntary exchange and market entry (and sometimes also: exit). Government regulations impose restrictions on individual market participants' actions and thereby limit the range of opportunities. On the one hand, a minimum set of regulations is a pre-condition for the functioning of markets and competition so that they can unfold their productivity enhancing power. A good regulatory framework reduces transaction costs on goods and factor markets and thus contributes to growth. Moreover, regulations may also improve the allocation of resources by channeling economic behavior of market participants in order to correct market failures from asymmetric information, externalities or natural monopoly markets. On the other hand, overly rigid regulatory systems can be an obstacle to economic growth if the set of implemented rules impedes welfare-enhancing voluntary transactions. Regulatory restraints can be so strict that they prevent an economy to respond quickly to technological change and to allocate scarce resources to their most productive uses.

While too little regulation is bad for growth because the necessary framework for competitive markets is not provided, too much regulation can be bad for growth if it restricts competition (by entry limitations) and voluntary exchange. A lack of competition in markets can thwart incentives for productivity improvements and therefore lead to reduced innovation dynamics through barriers to entrepreneurship (Aghion *et al.*, 2001, Cincera and Galgau, 2005). Severe regulations place an additional burden on economic activities and thus reduce the rate of return from investment in physical or human capital. As such, the burdens from regulation are similar to burdens of taxation. Structural policies and regulations which influence the working properties of markets can therefore contribute to cost differences in goods and factor markets. In case of excessive entry regulations, a liberalization or de-regulation can improve allocative efficiency by reducing monopoly rents and bringing prices in line with marginal costs. Also, enhanced competition will raise the productive efficiency of an economy by changing incentives for businesses. Moreover, a more open economy with reduced entry restrictions is also more attractive to foreign trade and investment (Nicodème and Sauner Leroy, 2007; Djankov, 2009). Finally, regulation also can serve as a means for state enforced re-distribution towards organized special interest groups. Achieving regulatory protection from competition is therefore a goal in socially unproductive rent seeking (Posner, 1975).

Seen from this view, the theoretical problems regarding the choice of an "optimal degree of regulation" are not too different from the questions with respect to the optimal fiscal size of government.[22]

Empirical evidence on the growth effects of the regulatory framework almost always points to the advantages of less heavily regulated markets. A number of empirical papers find that a more market-friendly regulatory environment is conducive to economic growth performance, and that too strict regulatory policies and lack of competition in markets are at the heart of a disappointing growth performance, specifically in some OECD nations (e.g., Dutz and Hayri, 1999; Griffith, Harrison and Simpson, 2006; Nicodème and Sauner Leroy, 2007). Nicoletti and Scarpetta (2003) find that productivity growth is boosted by reforms that promote private corporate governance and competition, and claim that "… entry-limiting regulation may hinder the adoption of technologies, possibly by reducing competitive pressures, technology spillovers, or the entry of new high-tech

[22] Wright (2004) even develops a similar theoretically hump-shaped relation between regulation intensity and growth performance as in Figure 1 of this paper.

firms". Alesina *et al.* (2005) report that a more competitive environment is good for growth as it stimulates private business investment. Fernandes (2008) finds a positive impact of de-regulation on productivity in the services sector in transition economies. Djankov, McLiesh and Ramalho (2006) use data from the World Bank's Doing Business reports as objective measures of business regulations in 135 countries. They find that countries with less regulation grow faster. Dawson (2006) reports a significant negative relationship between a broad measure of economic regulation and growth. Similar results are found when measures of credit market and business regulations are used.

Although it is still an ongoing debate, the vast majority of theoretical models and empirical papers conclude that trade is good for growth (e.g., Grossman and Helpman, 1991; but see also Rodriguez and Rodrik, 2001). The international division of labor is generally supposed to be a major driver for world-wide development. Restrictions on international trade – tariffs, quotas, hidden administrative regulations etc. – are therefore suspected to be growth depressing. What is more controversial among economists is whether freedom of international capital movements is unequivocally good for growth (e.g., Klein, 2005; Edwards, 2007). Even before the recent Financial Crisis a number of economists advocated capital controls as a means to protect local producers and financial markets at a developmental stage (e.g., Stiglitz, 2002).

The most heavily disputed regulations are concerned with labor market issues. On the one hand, market imperfections like asymmetric information and distribution of market power between employers and employees require some protection for workers through labor market legislation (Beetsma and Debrun, 2003). On the other hand, restrictive regulation of labor markets can easily cause sclerotic labor markets that are an obstacle to efficient allocation and growth. Empirical evidence on the growth effects of restrictive labor market regulations is scarce. Most empirical studies are rather concerned with employment effects. Rigid labor market institutions are frequently seen as a fundamental cause for high and persistent unemployment in a number of European countries (e.g., Blanchard and Wolfers, 2000). Though empirical evidence is somewhat scarce, at least some empirical studies indicate that growth in industrial countries – especially in the European economies – could be enhanced by lower de facto labor market regulation (Calderon and Chong, 2005).

4.2 Regulatory policies

In this sub-section we provide an overview of the degree of regulation in OECD and EU27 economies, as well as in a number of countries in the European periphery. Yet, whereas fiscal size can in principle be measured – though only imperfectly and involved with a lot of problems – the quality of regulations governing markets is even more difficult to gauge, as it is not the mere number of laws that is decisive. Instead of introducing a vast number of different indicators and measurement systems for regulatory policies in this sub-section, we employ the most comprehensive composite Economic Freedom of the World-index from the Fraser Institute, which is based on data from various international sources. We take the data from the most recent edition of the Economic Freedom of the World-report (Gwartney and Lawson, 2010) which provides data for the degree of regulation of certain markets and businesses up to 2008. We concentrate on the following dimensions of the *efw*-index:

- the regulation of international trade and capital flows,
- the regulation of domestic credit markets,
- the regulation of business in general, and
- the regulation of labor markets.

Table 6 displays the results for 2008.

Table 6

Intensity of Market Regulations According to Economic Freedom of the World Sub-indices, 2008

Country	Code	International Trade and Capital	Domestic Credit	Domestic Business	Domestic Labor	Summary*
New Zealand	NZ	7.9	10.0	7.8	8.5	8.6
Denmark	DK	7.7	9.5	7.4	7.5	8.0
Canada	CA	7.1	9.5	7.1	8.3	8.0
Ireland	IE	8.2	9.0	6.9	7.6	7.9
Australia	AU	6.7	9.5	6.7	8.5	7.9
United Kingdom	UK	7.6	9.0	6.7	8.0	7.8
United States	US	7.6	7.7	6.7	9.2	7.8
Slovakia	SK	8.1	10.0	5.3	7.7	7.8
Netherlands	NL	8.3	9.5	6.4	6.7	7.7
Estonia	EE	8.0	10.0	7.3	5.6	7.7
Switzerland	CH	6.8	9.0	7.0	7.9	7.7
Belgium	BE	8.0	9.4	6.3	6.9	7.7
Czech Republic	CZ	7.8	9.3	5.6	7.7	7.6
Iceland	IS	5.7	9.3	7.7	7.7	7.6
Bulgaria	BG	7.6	9.5	5.4	7.7	7.6
Hungary	HU	8.1	8.8	6.0	7.1	7.5
Luxembourg	LU	8.1	9.5	7.0	5.3	7.5
Austria	AT	7.6	9.4	6.8	5.9	7.4
Latvia	LV	7.3	9.2	6.1	7.1	7.4
Sweden	SE	7.7	9.5	7.1	5.1	7.4
Japan	JP	6.1	8.9	6.1	8.2	7.3
Finland	FI	7.4	9.8	6.9	5.1	7.3
France	FR	7.3	9.2	6.2	5.6	7.1
Malta	MT	7.1	9.4	4.6	7.0	7.0
Cyprus	CY	7.1	9.5	6.1	5.3	7.0
Lithuania	LT	7.5	9.2	5.7	5.6	7.0
Slovenia	SI	7.3	9.0	6.0	5.4	6.9
Romania	RO	7.4	7.5	5.9	6.7	6.9
Norway	NO	6.5	9.3	6.6	4.9	6.8
Spain	ES	7.0	9.3	5.8	5.1	6.8
Poland	PL	7.1	8.7	4.9	6.5	6.8
Italy	IT	7.1	7.9	5.4	6.3	6.7
Korea	KR	7.1	9.3	6.1	4.0	6.6
Germany	DE	7.7	8.2	6.6	3.9	6.6
Portugal	PT	7.2	7.6	5.9	5.2	6.5
Greece	EL	6.4	7.6	5.7	4.4	6.0
sample mean		7.4	9.1	6.3	6.5	7.3
Georgia	GE	7.7	8.7	7.5	7.3	7.8
Montenegro	ME	7.2	9.6	5.3	7.9	7.5
Kyrgyzstan	KG	7.4	9.2	6.4	6.2	7.3
Croatia	HR	6.5	9.4	5.1	6.3	6.8
Armenia	AM	6.6	9.0	5.3	6.1	6.8
Bosnia and Herzegovina	BA	6.2	8.9	5.2	6.7	6.8
Albania	AL	6.3	8.1	6.1	5.8	6.6
Serbia	RS	6.7	8.7	4.8	5.7	6.5
Turkey	TR	6.4	7.5	6.3	4.4	6.2
Ukraine	UA	6.5	8.1	3.7	6.3	6.2
sample mean		7.2	8.6	6.1	5.7	6.9

* Simple average of the four regulation sub-indices, WIFO calculations.
Source: Gwartney and Lawson (2010).

International trade and capital flows

Also as a consequence of integration of international goods and capital markets through various international treaties, the countries in the sample observe a high level of trade and capital markets liberalization in 2008. On a 0-to-10-point-scale, average regulation index level is 7.4, lying in a range between 8.3 (Netherlands) and 5.7 (Iceland) (see Table 6). Trade and international capital movements are also reasonably liberalized in the 10 countries of the European periphery for which data are available. On average, the liberalization level is 7.2 points, with Georgia (7.7) having a regulatory regime that provides liberties comparable to Sweden or the USA.

Credit market regulations

This sub-index measures the extent to which the banking industry is dominated by private firms and whether foreign banks are permitted to compete in the market. It also indicates the extent to which credit is supplied to the private sector and whether controls on interest rates interfere with the market in credit. The average liberalization level of domestic credit markets in 2008 was 9.1, only a few countries (Portugal, Greece, Romania, Italy, and the USA) observed a liberalization level that is slightly less than 8 points on the scale.

Business regulations

The index of private business regulation identifies the extent to which regulatory policies and bureaucratic procedures restrain entry and reduce competition. In order to score high in this sub-index, governments must allow predominantly markets to determine prices and refrain from regulatory activities that retard entry into business and increase the cost of production. On average, the countries in the OECD/EU27 sample arrive at a liberalization level of 6.3, which is far lower than the international trade regulations level. While New Zealand and Iceland observe the highest level of de-regulation of product markets, especially Malta and Poland appear to have still a high potential to liberalize and, thus, enhance competition on domestic markets. According to the results of most empirical studies, this would boost growth in these countries. OECD (2005b), hence, expected a substantial increase of GDP per capita growth in the EU15 if competition-restraining regulations were abandoned.

Labor market regulations

The least regulated labor markets according to the *efw*-index can be found in the Anglo-Saxon Welfare States (USA, Australia, New Zealand, Canada, UK) as well as in Japan. Continental Europe, especially Germany in 2008, is lagging behind.[23] Greece, Spain, and Portugal also faced more rigid labor market regulations.

Summary index

Taking the simple mean of these four regulation-indices, New Zealand is the least regulated country in the sample, while Greece is the most heavily regulated. The countries in the European Periphery observe somewhat more economic regulation than the ones of the developed countries sample. Yet, the differences in 2008 are not very pronounced.

---

[23] In the meantime Germany put in place a number of labor market reforms which will probably improve its score of the labor market regulation index.

Figure 17 shows a positive relationship between the level of GDP per capita and the state of market liberalization in 2008, taking also into account countries from the European Periphery sample. A simple bi-variate cross-country regression indicates that the interrelation between both variables is statistically significant at a 1 per cent level of confidence.

Figure 18 illustrates development of the summary regulation index over time in four country groups. While markets are already highly liberalized in EU15 and further OECD countries, the EU12 and the European Periphery observed a liberalization of regulatory policies over time. Until 2008 the differences between the country groups have been substantially reduced.

5 Interplay between expenditures, taxation and regulation

5.1 The role of policy complementarities

Having analyzed separately the spending, taxation and regulation patterns of the countries in our sample, the focus of this section will be placed on the interplay of the respective policies. Although often neglected in theoretical as well as empirical investigations, complementarities between policies can play an important in role for the growth friendliness of entire policy packages. As reforms are mutually interdependent, a country's economic policy package needs coherence, or, "economic complementarities", "… in a sense that the effectiveness of one policy depends on the implementation of other policies" (Orszag and Snower*,* 1998). Neglecting such interdependencies between policies can result in a wrong assessment of the economic effects of single policy measures (Aziz and Wescott, 1997).

The role of the interaction between certain economic policies in promoting growth has only recently received significant attention in the empirical growth literature. Aziz and Wescott (1997) consider measures for international openness, macro stability and size of government in a sample of 76 developing countries, and report that – analyzed separately – virtually none of these policies is significant in boosting growth over a 10 year period from 1985-95. Introducing a concept of complementarities between these different policies, they find that countries which have high quality of policies in all three measures (or at least only one "medium quality policy") have a significantly higher probability to observe higher growth.

Chang, Kaltani and Loayza (2009) find that the growth-promoting effect of trade openness depends on complementary reforms which help a country take advantage of international competition. Their estimates show that trade openness can reduce or increase growth, depending on the status of the complementary reforms in the areas educational investment, financial depth, inflation stabilization, public infrastructure quality, governance, labor-market flexibility, ease of firm entry, and ease of firm exit. This clearly indicates that the growth effects of an increase in international trade openness depend positively on the progress made in other policy areas. Bokaky and Freund (2004) also find that increased trade does not stimulate growth in economies with substantial regulatory interventions, it may even reduce growth in countries with excessive government regulation. In a similar vein, Gwartney, Holcombe and Lawson (2006) find countries with a higher overall institutional quality to experience a higher productivity of investment. More specifically, private investment is much more responsive to cross-country differences in economic freedom than are rates of government investment.

Figure 17

Intensity of Market Regulations and GDP per capita, 2008

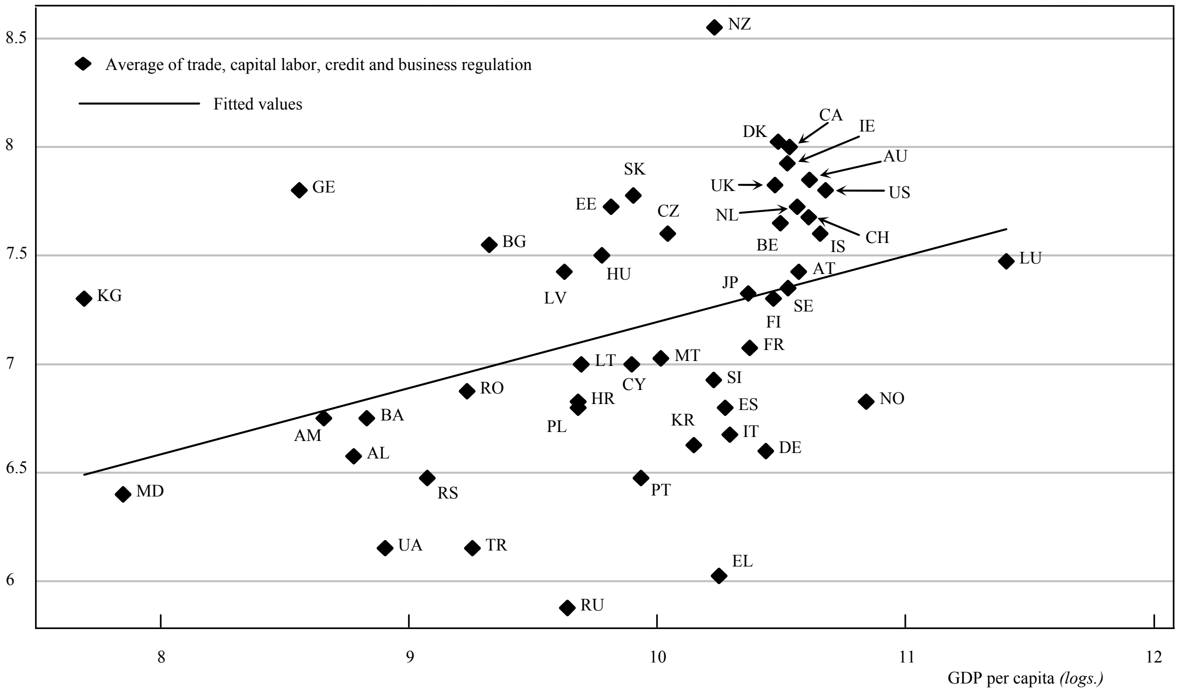
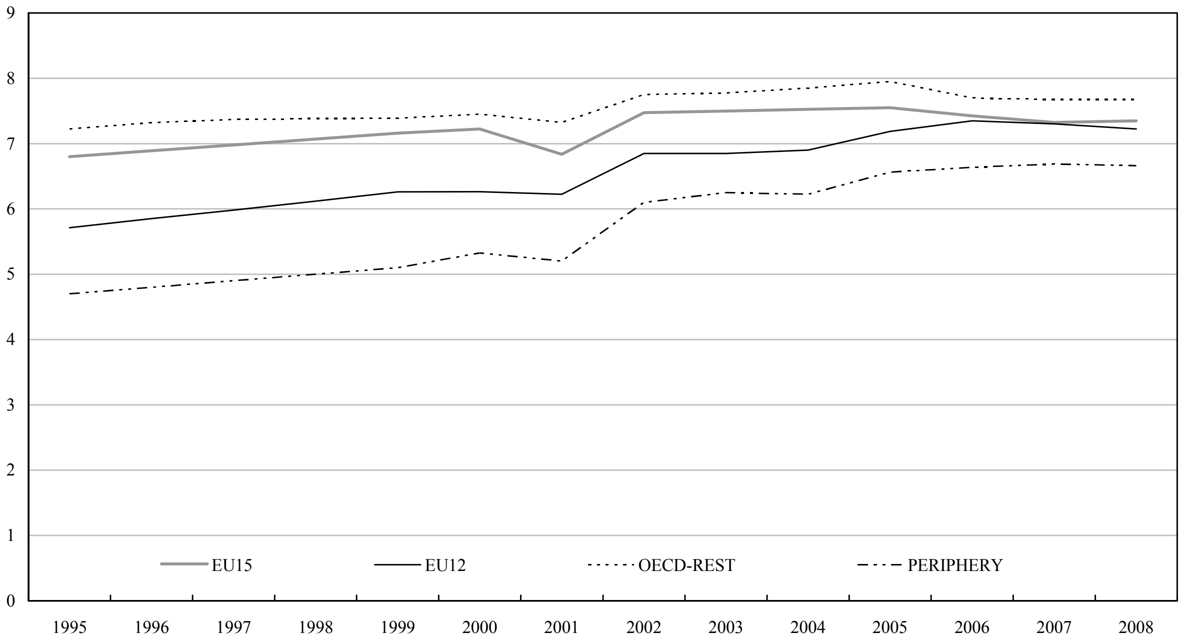


Figure 18

Median Economic Liberalization Levels in Groups of EU15, EU12 and Further OECD Countries, 1995-2008
(according to summary regulation index)



Source: WIFO calculations,based on Gwartney and Lawson (2010). Median values for the years 1996-99 derived from interpolated data.

Most recently, Braga de Macedo, Oliveira Martins and Rocha (2010) assess the possible impact of complementarities over six broad policy areas cross-country estimates in a sample of 130 countries over a time span of 13 years (1994-2006). The policy areas included are: i) trade openness, ii) business regulations, iii) freedom of capital movement, iv) openness of the domestic banking and financial system, v) property rights protection and vi) infrastructure quality. These major areas therefore resemble to some extent the policies that are considered to be growth enhancing in the present paper. Policy complementarities are captured by the standard deviation of the six aforementioned individual policy indicators, which have been standardized on a 0-100scale.[24] The authors find evidence that the variables having the strongest explanatory power are the average change of policies towards more economic liberalization and the time-averaged standard deviation of individual policy indicators, even after the inclusion of several controls. They conclude that "[t]his implies that countries where policy complementarities can unfold to a greater extent grow faster. Achieving a higher level of policy complementarity has therefore a permanent effect on growth rates". Turning to panel techniques, the introduction of (country) fixed-effects destroys the significance of the complementarities measure, indicating that the effect is driven mainly by the cross-section variance. In a simple random-effects framework, the positive impact of more coherent policies remains. Braga de Macedo, Oliveira Martins and Rocha (2010) therefore confirm the findings of a previous paper on transition economies, where the authors used different measures for complementarities (Braga de Macedo and Oliveira Martins, 2008).

In contrast to these economic complementarities between policy areas, political policy complementarities arise when the ability to gain political consent for one policy depends on the implementation of others (Orszag and Snower, 1998). This somehow parallels the famous argument of Rodrik (1998) who claims that many countries have increased social security spending and social regulation in order to compensate for higher risks due to globalization and market deregulation. On the other hand, Bergh and Karlson (2010) report evidence that high-tax countries might use a liberalization of trade as a substitute for excessive overall government size. Their results support the idea that countries with big government can use economic openness to mitigate the negative growth effects of high taxes and expenditures.

5.2 *Some empirical facts*

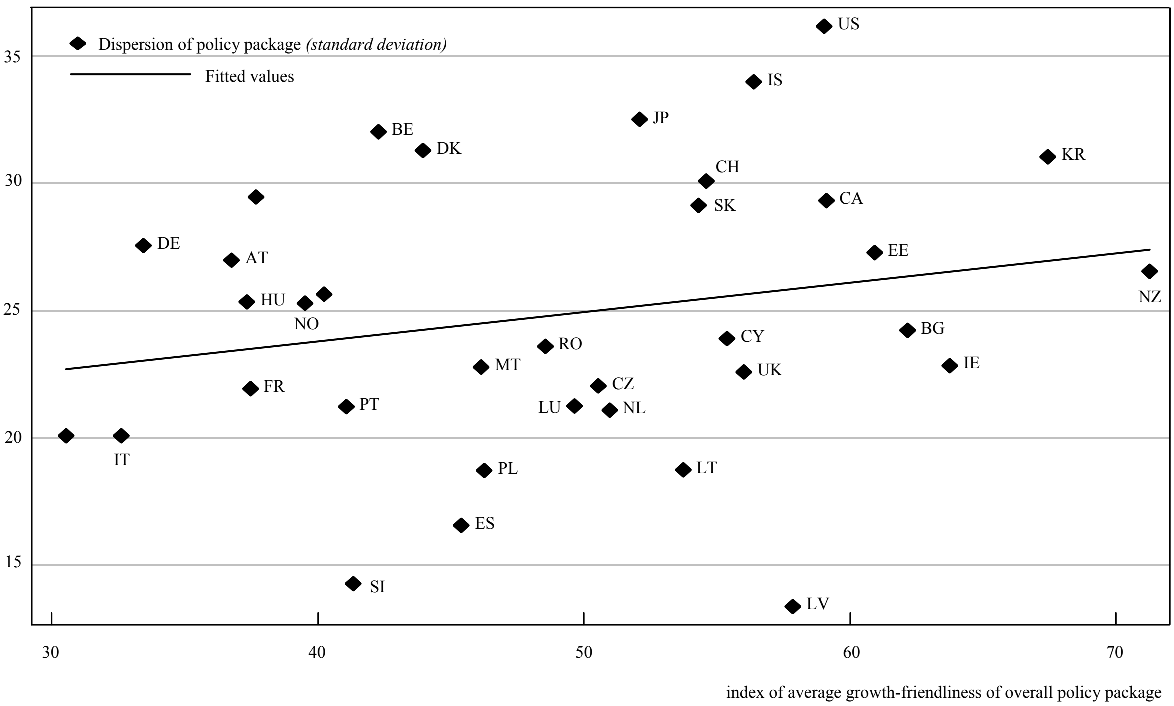
In this sub-section we will aim to investigate the existence (or absence) of complementarities between public expenditures, taxation and regulation in our sample. Note, first, that there is no single measure for complementarities, and, second, that we do not have an exact notion of the "optimal" level of productive spending or regulations. We therefore calculate a simple standardized index of the relative growth friendliness of a country's policy package as well as for the coherence/dispersion of the respective policy package, taking into account the real world range and distribution of the data in our sample. The construction of the indices assumes linearity, *i.e.*, possible non-linear relations between policy variables and economic outcomes are not reflected in the indices.

The first index is an index of the average growth friendliness of a country's policy mix, consisting of indicators for spending, taxation and regulation policies. It is constructed by measuring the growth friendliness of 13 policy indicators (see box) in relation to other countries in the sample. The resulting index is standardized on a 0-100 scale, where higher values reflect higher (average) growth friendliness.

[24] Instead of employing the Fraser Institutes measures the authors use instead the Economic Freedom index of Wall Street Journal and Heritage Foundation.

Figure 19

Policy Dispersion and Average Growth Friendliness, 2008



The second index is simply calculated as the standard deviation of the growth-friendliness index of these 13 policies. Higher values indicate more dispersion and a less coherent overall policy package. Table 7 indicates the respective values for 2008.

The average index is led by New Zealand, followed by Korea, Ireland and Bulgaria. At the bottom of the 2008 ranking we find Austria, Germany, Italy and Greece. With respect to the policy dispersion measure, the most coherent policy mix can be found in Latvia, Slovenia and Spain, while the USA, Iceland, and Japan observe the highest standard deviation of our set of 13 policy indicators. Both measures are not strongly correlated, though. Figure 19 shows that average growth friendliness and policy dispersion are not strongly connected. If anything, there is a slightly positive relation between the two variables. Simple correlation tests also reveal no significant between both indicators.

6 Summary and outlook

Are fiscal and regulation policies in Europe in line with the recommendations from the new growth literature? The present study provides an overview of the growth friendliness of fiscal and regulatory structures in a sample of developed OECD countries and EU members (EU15 and EU12). Peripheral European (transition) countries are also included, whenever respective data are available.

Based on several measures capturing the expenditure and the tax side of the budgets, as well as regulatory policies, the size and the structure of public sectors differ markedly across countries. Our analysis of regulatory regimes is based on indicators for the liberalization of international trade and capital movements, as well as domestic credit markets, labor markets and business regulations.

Table 7

Growth-friendliness Index and Policy Dispersion Index, 2008

S-code	Country	Growth Friendliness	Dispersion
NZ	New Zealand	71.3	26.6
KR	Korea	67.4	31.0
IE	Ireland	63.8	22.8
BG	Bulgaria	62.2	24.2
EE	Estonia	60.9	27.3
CA	Canada	59.1	29.3
US	United States	59.0	36.2
LV	Latvia	57.9	13.4
IS	Iceland	56.4	34.0
UK	United Kingdom	56.0	22.6
CY	Cyprus	55.4	23.9
CH	Switzerland	54.6	30.1
SK	Slovakia	54.3	29.1
LT	Lithuania	53.7	18.7
JP	Japan	52.1	32.5
NL	Netherlands	51.0	21.1
CZ	Czech Republic	50.5	22.0
LU	Luxembourg	49.6	21.2
RO	Romania	48.5	23.6
PL	Poland	46.3	18.7
MT	Malta	46.1	22.8
ES	Spain	45.4	16.5
DK	Denmark	44.0	31.3
BE	Belgium	42.3	32.0
SI	Slovenia	41.3	14.3
PT	Portugal	41.1	21.2
FI	Finland	40.2	25.7
NO	Norway	39.5	25.3
SE	Sweden	37.7	29.5
FR	France	37.5	21.9
HU	Hungary	37.3	25.3
AT	Austria	36.8	27.0
DE	Germany	33.5	27.5
IT	Italy	32.6	20.1
EL	Greece	30.5	20.1

Source: WIFO calculations.

On average, New Zealand is the least regulated country in the sample, while Greece is the most heavily regulated. Countries of the European periphery observe a bit more strict economic regulation than those of the developed countries sample. Yet, the differences have become smaller over time and in 2008 they are not very pronounced any more.

Using a standardized index of the relative growth friendliness of a country's policy package as well as for the coherence/dispersion of the respective policy mix of spending, tax and regulation policies, in 2008 the most coherent policy mix can be found in Latvia, Slovenia and Spain. The USA, Iceland, and Japan observe the least coherent policy package, as measured by the standard deviation of our set of 13 policy indicators. Average growth friendliness of public policy and the level of policy dispersion are not strongly related.

Future work will have to take a closer look at the economic and political determinants of these substantial differences in size and composition of government spending, structure and volume of taxation and the regulatory regimes. Are productive and growth-friendly spending, tax and regulation structures driven by demographic change or by income development? Empirical analyses suggest that population aging is linked to higher social expenditures (e.g., Sanz and Velazquez, 2007), but what about the economic determinants of productive spending (e.g., Shelton, 2007; Pitlik, 2009)?

REFERENCES

Afonso, A., W. Ebert, L. Schuknecht and M. Thöne (2005), "Quality of Public Finances and Growth", ECB, Working Paper, No. 438.

Afonso, A. and D. Furceri (2010), "Government Size, Composition, Volatility and Growth", *European Journal of Political Economy*, No. 26, pp. 517-32.

Aghion, P. (2008), "Higher Education and Innovation", *Perspektiven der Wirtschaftspolitik*, No. 9, Special Issue, pp. 28-45.

Aghion, P., C. Harris, P. Howitt and J. Vickers (2001), "Competition, Imitation and Growth with Step-by-step Innovation", *Review of Economic Studies*, No. 68, pp. 467-92.

Alegre, J., A. Kappeler, A. Kolev and T. Väliliä (2008), "Composition of Government Investment in Europe: Some Forensic Evidence", European Investment Bank, Paper, No. 13, pp. 22-54.

Alesina A., S. Ardagna, G. Nicoletti and F. Schiantarelli (2005), "Regulation and Investment", *Journal of the European Economic Association*, No. 3, pp. 791-825.

Angelopoulos, K., G. Economides and P. Kammas (2007), "Tax-spending Policies and Economic Growth: Theoretical Predictions and Evidence from the OECD", *European Journal of Political Economy*, No. 23, pp. 885-902.

Arnold, J. (2008), "Do Tax Structures Affect Aggregate Economic Growth? Empirical Evidence from a Panel of OECD Countries", OECD, Economics Department, Working Paper, No. 643.

Aschauer, D.A. (1989), "Is Public Expenditure Productive?", *Journal of Monetary Economics*, No. 23, pp. 177-200.

Aziz, J. and R.F. Wescott (1997), "Policy Complementarities and the Washington Consensus", IMF, Working Paper, No. WP 97/118, Washington (D.C.).

Baldacci, E., B. Clements, S. Gupta and Q. Cui (2008), "Social Spending, Human Capital, and Growth in Developing Countries", World Development, No. 36, pp. 1317-41.

Barrios, S. and A. Schaechter (2008), "The Quality of Public Finances and Economic Growth", *European Economy*, Economic Paper, No. 337.

Barro, R.J. (1990), "Government Spending in a Simple Model of Endogenous Growth", *Journal of Political Economy*, No. 98, pp. s103-17.

Barro, R.J. and X. Sala-i-Martín (1992), "Public Finance in Models of Economic Growth", *Review of Economic Studies*, No. 99, pp. 645-61.

Bassanini, A. and S. Scarpetta (2002), "Does Human Capital Matter for Growth in OECD Countries? A Pooled Mean-Group Approach", *Economics Letters*, No. 74, pp. 399-405.

Bayraktar, N. and B. Moreno-Dodson (2010), "How Can Public Spending Help You Grow? An Empirical Analysis for Developing Countries", The World Bank, Policy Research, Working Paper, No. 5367.

Beetsma, R.M.W.J. and X. Debrun (2003), "Reconciling Stability and Growth: Smart Pacts and Structural Reforms", IMF, Working Paper, No. 03/174, Washington (D.C.).

Bergh, A. and M. Henrekson (2011), "Government Size and Growth: A Survey and Interpretation of the Evidence", IFN, Working Paper, No. 858, Stockholm.

Blanchard, O., G. Dell'Ariccia and P. Mauro (2010), "Rethinking Macroeconomic Policy", IMF, Staff Position Note, No. SPN 10/03, February 12.

Blanchard, O. and J. Wolfers (2000), "The Role of Shocks and Institutions in the Rise of European Unemployment: The Aggregate Evidence", *Economic Journal*, No. 110, pp. C1-C33.

Bleaney, M., N. Gemmell and R. Kneller (2001), "Testing the Endogenous Growth Model: Public Expenditure, Taxation, and Growth over the Long Run", *Canadian Journal of Economics*, No. 34, pp. 36-57.

Bose, N., M.E. Haque and D.R. Osborn (2007), "Public Expenditure and Economic Growth: A Disaggregated Analysis for Developing Countries", *The Manchester School*, No. 75, pp. 533-56.

Braga de Macedo, J. and J. Oliveira Martins (2008), Growth, Reform Indicators and Policy Complementarities", *Economics of Transition*, No. 16, pp. 141-64.

Braga de Macedo, J., J. Oliveira Martins and B. Rocha (2010), "Policy Complementarities and Growth", paper presented at the OECD/CESIfo/Ifo Workshop *Regulation: Political Economy, Measurement and Effects on Performance*, Munich, 29-30 January.

Buchanan, J.M. (1975), *The Limits of Liberty. Between Anarchy and Leviathan*, Chicago.

Calderon, C. and A. Chong (2005), "Are Labor-market Regulations an Obstacle for Long-term Growth?", in J. Restrepo (ed.), *Labor Markets and Institutions*, Central Bank of Chile.

Campos, N.F. and F. Coricelli (2002), "Growth in Transition: What We Know, What We Don't, and What We Should", *Journal of Economic Literature*, No. 40, pp. 793-836.

Chang, R., L. Kaltani and N. Loayza (2009), "Openness Can Be Good for Growth: The Role of Policy Complementarities", *Journal of Development Economics*, No. 90, pp. 33-49.

Ciccone, A. and M. Jarocinski (2010), "Determinants of Economic Growth: Will Data Tell?", *American Economic Journal: Macroeconomics 2*, pp. 222-46.

Cincera, M., and O. Galgau (2005), "Impact of Market Entry and Exit on EU Productivity and Growth Performance", European Commission, Economic Paper, No. 222.

Crafts, N. (2009), "Transport Infrastructure Investment: Implications for Growth and Productivity", *Oxford Review of Economic Policy*, No. 25, pp. 327-43.

Dawson, J.W. (2006), "Regulation, Investment, and Growth across Countries", *Cato Journal*, No. 26, pp. 489-509.

De Mooij, R. and S. Ederveen (2008), "Corporate Tax Elasticities: A Reader's Guide to Empirical Findings", *Oxford Review of Economic Policy*, No. 24, pp. 680-97.

Djankov, S. (2009), "The Regulation of Entry: A Survey", *The World Bank Research Observer*, No. 24, pp. 183-203.

Djankov, S., C. McLiesh and R.M. Ramalho (2006), "Regulation and Growth", *Economics Letters*, No. 92, pp. 395-401.

Dowrick, S. (1996), "Estimating the Impact of Government Consumption on Growth: Growth Accounting and Endogenous Growth Models", *Empirical Economics*, No. 21, pp. 163-86.

Dutz, M.A. and A. Hayri (1999), "Does More Intense Competition Lead to Higher Growth?", The World Bank, Policy Research, Working Paper, No. 2320.

Edwards, S. (2007), "Capital Controls, Sudden Stops, and Current Account Reversals", in S. Edwards (ed.), *Capital Controls and Capital Flows in Emerging Economies: Policies, Practices and Consequences*, Chicago, pp. 73-119.

Egert, B., T. Kozluk and D. Sutherland (2009), "Infrastructure and Growth: Empirical Evidence", OECD, Economics Department, Working Paper, No. 685.

European Commission (2002), *Public Finances in EMU – 2002*, Brussels.

—————— (2010a), *Public Finances in EMU – 2010*, Brussels.

—————— (2010b), *Taxation Trends in the EU – 2010*, Brussels.

—————— (2011), *AMECO-database*. Download June, 2011: http://ec.europa.eu/economy_finance/ameco/user/serie/SelectSerie.cfm?CFID=561884&CFTOKEN=99299bfc3cb24ffd-10EF7BA1-9DF1-497B-FD17F5DB5484AD03&jsessionid=24068dd2329a5d552934

Fatás, A. and I. Mihov (2001), "Government Size and Automatic Stabilizers: International and Intranational Evidence", *Journal of International Economics*, No. 55, pp. 3-28.

Fernandes, A.M. (2008), "Structure and Performance of the Service Sector in Transition Economies", The World Bank, Development Research Group, mimeo.

Fidrmuc, J. (2003), "Economic Reform, Democracy and Growth during Post-communist Transition", *European Journal of Political Economy*, No. 19, pp. 583-604.

Fölster, S. and M. Henrekson (2001), "Growth Effects of Government Expenditure and Taxation in Rich Countries", *European Economic Review*, No. 45, pp. 1501-20.

Gemmell, N., R. Kneller and I. Sanz (2011), "The Timing and Persistence of Fiscal Policy Impacts on Growth: Evidence from OECD Countries", *Economic Journal*, No. 121, pp. F33–F58.

Ghosh, S. and A. Gregoriou (2008), "The Composition of Government Spending and Growth: Is Current or Capital Spending Better?", Oxford Economic Papers, No. 60, pp.484-516.

Griffith R., R. Harrison and H. Simpson (2006), "The Link Between Product Market Reform, Innovation and EU Macroeconomic Performance", European Economy, Economic Papers, No. 243.

Griffith-Jones, S., J.A. Ocampo and J.E. Stiglitz (2010), *Time for a Visible Hand. Lessons from the 2008 World Financial Crisis*, Oxford.

Grossman, G. and E. Helpman (1991), *Innovation and Growth in the Global Economy*, Cambridge.

Gwartney, J.D., R.G. Holcombe and R.A. Lawson (2006), "Institutions and the Impact of Investment on Growth", *Kyklos*, No. 59, pp. 255-73.

Gwartney, J.D. and R. Lawson (2010), *Economic Freedom of the World: 2009 Annual Report*, Vancouver.

Handler, H., A. Knabe, B. Koebel, M. Schratzenstaller and S. Wehke (2005), "The Impact of Public Budgets on Overall Productivity Growth", WIFO, Working Paper, No. 255.

Hayek, F.A. (1960), *The Constitution of Liberty*, Chicago.

Johannson, A., C. Heady, B. Brys and L. Vartia (2008), "Taxation and Economic Growth", OECD, Economics Department, Working Paper, No. 620.

Kamps, C. (2006), "New Estimates of Government Net Capital Stocks for 22 OECD Countries 1960-2001", IMF, Staff Papers, No. 53, pp. 120-50.

Kappeler, A. and T. Välilä (2008), "Composition of Public Investment and Fiscal Federalism: Panel Data Evidence from Europe", *European Journal of Political Economy*, No. 24, pp. 562-70.

Klein, M.W. (2005), "Capital Account Liberalization, Institutional Quality and Economic Growth: Theory and Evidence", NBER, Working Paper, No. 11112, Cambridge (MA).

Kneller, R., M. Bleaney and N. Gemmell (1999), "Fiscal Policy and Growth: Evidence from OECD Countries", *Journal of Public Economics*, No. 74, pp. 171-90.

KPMG, *Corporate and Indirect Tax Rate Survey 2010*.

Lindert, P. (2004), *Growing Public*, Cambridge.

Lucas, R. (1988), "On the Mechanics of Economic Development", *Journal of Monetary Economics*, No. 22, pp. 3-42.

Martin, P. and C. Rogers (2000), "Long-term Growth and Short-term Economic Stability", *European Economic Review*, No. 44, pp. 359-81.

Meghir, C. and D. Phillips (2008), "Labour Supply and Taxes", IZA, Discussion Paper, No. 3405.

Murphy, K.M., A. Shleifer and R.W. Vishny (1991), "The Allocation of Talent: Implications for Growth", *Quarterly Journal of Economics*, No. 106, pp. 503-30.

Musgrave, R.A. (1959), *Public Finance*, New York.

Myles, G.D. (2009), "Economic Growth and the Role of Taxation – Aggregate Data", OECD, Economics Department, Working Paper, No. 714.

Nicodème, G. and J.B. Sauner Leroy (2007), "Product Market Reforms and Productivity: A Review of the Theoretical and Empirical Literature on the Transmission Channels", *Journal of Industry, Competition and Trade*, No. 7, pp. 53-72.

Nicoletti, G. and S. Scarpetta (2003), "Regulation, Productivity and Growth: OECD Evidence", *Economic Policy*, Vol. 18, Issue 36, pp. 9-72.

Nijkamp, P. and J. Poot (2004), "Meta-analysis of the Effect of Fiscal Policies on Long-run Growth", *European Journal of Political Economy*, No. 20, pp. 91-124.

North, D. (1990), *Institutions, Institutional Change, and Economic Performance*, Cambridge.

OECD (2005a), *Modernising Government. The Way Forward*, Paris.

—————(2005b), "The Benefits Of Liberalising Product Markets And Reducing Barriers To International Trade and Investment: The Case of The United States and the European Union", OECD, Economics Department, Working Paper, No. 432, Paris.

—————(2010a), *Preparing Fiscal Consolidation*, Paris.

—————(2010b), *Revenue Statistics*. Available at: http://www.oecd-ilibrary.org/taxation/data/oecd-tax-statistics_tax-data-en

—————(2011), *Tax Database 2011*. Available at: http://www.oecd.org/document/60/0,3746,en_2649_ 34897_1942460_1_1_1_1,00.html

Orszag, M. and D. Snower (1998), "Anatomy of Policy Complementarities", *Swedish Economic Policy Review*, No. 5, pp. 303-43.

Perotti, R. (1996), "Growth, Income Distribution, and Poverty", *Journal of Economic Growth*, No. 1, pp. 149-78.

Pitlik, H. (2009), "Reforming Public Expenditure Composition: The Impact of Economic Conditions and Budget Institutions", final report to the *Jubiläumsfonds der Oesterreichischen Nationalbank*, Grant No. 12821, Vienna.

——————(2010), "Fiscal Governance and Government Investment in Europe since the 1990s", WIFO, Working Pape, No. 370.

Pitlik, H. and M. Schratzenstaller (2011), "Growth Implications of Structure and Size of Public Sectors", WIFO, Working Paper, No. 404/2011.

Posner, R.A. (1975), "The Social Costs of Monopoly and Regulation", *Journal of Political Economy*, No. 83, pp. 807-27.

Pushak, T., E. Tiongson and A. Varoudakis (2007), "Public Finance, Governance, and Growth in Transition Economies", The World Bank, Policy Research, Working Paper, No. 4255.

Ramey, G. and V. Ramey (1995), "Cross-country Evidence on the Link Between Volatility and Growth", *American Economic Review*, No. 85, pp. 1138-51.

Rodriguez, F. and D. Rodrik (2001), Trade Policy and Economic Growth: A Skeptics Guide to the Cross-national Evidence", in B. Bernanke and K. Rogoff (eds.), *NBER Macroeconomics Annual 15*, Cambridge, pp. 261-325.

Rodrik, D. (1998), "Why Do More Open Economies Have Bigger Governments?", *Journal of Political Economy*, No. 106, pp. 997-1032.

Romero-Ávila, D. and R. Strauch (2008), "Public Finances and Long-term Growth in Europe: Evidence from a Panel Data Analysis", *European Journal of Political Economy*, No. 24, pp. 172-91.

Romp, W. and J. de Haan (2007), "Public Capital and Economic Growth: A Critical Survey", *Perspektiven der Wirtschaftspolitik*, No. 8 (Special Issue), pp. 6-52.

Sanz, I. and F. Velázquez (2007), "The Role of Ageing in the Growth of Government and Social Welfare Spending in the OECD", *European Journal of Political Economy*, No. 23, pp. 917-31.

Schaltegger, C.A. and B. Torgler (2006), "Growth Effects of Public Expenditure on the State and Local Level: Evidence from a Sample of Rich Governments", *Applied Economics*, No. 38, pp. 1181-92.

Schratzenstaller, M. (2007), "Wachstumsimpulse durch die öffentliche Hand", *WIFO Monatsberichte*, No. 80, pp. 509-26.

Semmler, W., A. Greiner, B. Diallo, A. Rezai and A. Rajaram (2007), "Fiscal Policy, Public Expenditure Composition, and Growth", The World Bank, Policy Research, Working Paper, No. 4405.

Shelton, C.A. (2007), "The Size and Composition of Government Expenditure", *Journal of Public Economics*, No. 91, pp. 2230-60.

Stiglitz, J.E. (2002), *Globalization and Its Discontents*, New York.

Tanzi, V. and H. Zee (1997), "Fiscal Policy and Long-run Growth", IMF, Staff Paper, No. 44, pp. 179-209.

Task Force of the Monetary Policy Committee of the European System of Central Banks (2008), "Labour Supply and Employment in the Euro Area Countries. Developments and Challenges", ECB, Occasional Paper, No. 87.

Välilä, T., T.Kozluk and A. Mehrotra (2005), "Road on a Downhill? Trends in EU Infrastructure Investment", European Investment Bank Paper, No. 10, 1.

Wright, E.O. (2004), "Beneficial Constraints: Beneficial for Whom?", *Socio-economic Review*, No. 2, pp. 407-14.

Zagler, M. and G. Durnecker (2003), "Fiscal Policy and Economic Growth", *Journal of Economic Surveys*, No. 17, pp. 397-418.

SERVICE REGULATION AND GROWTH: EVIDENCE FROM OECD COUNTRIES

Guglielmo Barone[*] *and Federico Cingano*[*]

*We study the effects of anti-competitive service regulation by examining whether OECD countries with less anti-competitive regulation see better economic performance in manufacturing industries that use less-regulated services more intensively. Our results indicate that lower service regulation increases value added, productivity, and export growth in downstream service intensive industries. The regulation of professional services and energy provision has particularly strong negative growth effects. Our estimates are robust to accounting for alternative forms of regulation (*i.e.*, product and labour market regulation), alternative measures of financial development and a range of other specification checks.*

Do countries with less anti-competitive service regulation perform better economically? Policy makers appear to think so as regulatory barriers have fallen in many countries. And their position is generally supported by a large empirical literature looking at the effects of entry barriers, red-tape costs or legal requirements on economic performance. Much of this literature examines the effects of regulation on the performance of the regulated sector. Less is known about the impacts on downstream manufacturing activities, which is surprising as regulation affects many key service inputs.

In this paper, we study how regulation in the supply of a variety of services affects the economic performance of downstream manufacturing industries. We do so by examining whether countries with less service regulation see faster value added, productivity, and export growth in manufacturing industries using services more intensively (this methodology was pioneered for financial services by Rajan and Zingales, 1998). We measure service dependence across manufacturing industries using input-output account matrices. Our measures of service regulation are OECD indicators designed to capture anti-competitive regulatory settings for the energy sector (electricity and gas), the telecommunication and the transportation sectors and for professional services. These account for barriers to entry, for the integration between a priori competitive activities and natural monopolies (in the case of energy), and for the existence of restrictions on prices and fees, advertising or the form of business (in professional services).

Our empirical findings indicate that lower service regulation has non-negligible positive effects on the value added, productivity and export growth rates of service intensive users. To get a sense for the size of the regulation effect, consider the annual value added growth differential between an industry at the 75th percentile (Pulp, paper and printing) relative to one at the 25th percentile (Fabricated metal products) of the distribution of service dependence. Our estimates imply that this differential is 0.7-1 per cent higher in a country with average regulation at the 25th percentile (as Canada) than in a country at the 75th percentile (as France) of the distribution of service regulation. We find this effect is mainly driven by regulation in energy and in professional

---

[*] Bank of Italy. Correspondence: Bank of Italy, Research Department, via Nazionale 91, 00184 Rome, Italy; E-mail: federico.cingano@bancaditalia.it; guglielmo.barone@bancaditalia.it

We are extremely grateful to Antonio Ciccone and Alfonso Rosolia for their help and suggestions. We thank Andrea Ichino, Tullio Jappelli, Giovanni Pica, Paolo Pinotti, Fabiano Schivardi, two anonymous referees and seminar participants at the Bank of Italy, University of Bologna, University of Cagliari, CSEF Salerno and MILLS 2007 for very useful comments. We are responsible for any mistakes. The opinions expressed here are our own and do not necessarily correspond to those of the Bank of Italy.

This paper was originally published in The Economic Journal Vol. 121, Issue 555, September 2011, Blackwell Publishing Ltd., http://dx.doi.org/10.1111/j.1468-0297.2011.02433.x.

services. Also, the average effect is driven by larger economies in the sample. The results are not sensitive to how we account for other forms of regulation (*i.e.*, product and labour market regulation) and prove robust to a number of specification checks.

Our findings have important implications for the ongoing debate surrounding service deregulation. In particular, our estimates imply that the strongest gains from deregulation would come from specific policies such as the removal of conduct regulation (*i.e.*, of restrictions to price and tariff setting) by professions, or the complete separation of ownership between energy generation and other segments of the industry (the so-called "unbundling"). Both measures are among those ranking highest in the current EU competition policy agenda and in policy recommendations by international organizations.[1]

Research on the economic effects of regulation has grown in recent years, in part because of the increased availability of comparable cross-country data. Empirical work has focused mainly on the *direct* effects of regulation on the regulated sector or stage of business development. Economy-wide restrictions such as barriers to entry have been shown to hamper economy-wide entrepreneurship by stifling growth in the number of firms (Klapper *et al*., 2006), by increasing industry concentration (Fisman and Sarria-Allende, 2004), and by reducing responsiveness to global demand and technology shifts (Ciccone and Papaioannou, 2007). Sector-specific restrictions, such as those prevailing in utilities and services, have been shown to decrease investment (Alesina *et al*., 2005) and employment (Bertrand and Kramartz, 2002), and to increase prices (Martin *et al*., 2005) in the regulated sectors. Yet, regulation may also have relevant *indirect* effects on the allocation of resources among downstream industries, in particular when affecting the production of key non-tradable inputs.

In theoretical models of industry interdependence, the under-development of markets for non-tradable inputs has been shown to constrain (or even prevent) the diffusion of input-intensive technologies, thus affecting the patterns of resource allocation and international specialization (Okuno-Fujiwara, 1988; Rodriguez-Clare, 1996). Empirical research into the relationship between upstream markets development and the allocation of resources across downstream industries has, however, been largely confined to the case of finance.

Rajan and Zingales' (1998) test of the finance-growth nexus using country-industry data represents a major contribution to this literature. The authors exploit industry heterogeneity in financial dependence (*i.e.*, the need for external funds) to show that in countries with better developed financial markets, financially dependent industries experience faster value added growth than less dependent industries. Their findings, confirmed by many subsequent studies, point to financial development as one relevant determinant of the patterns of international specialization. One contribution of our work is to show that the growth effects of service regulation can be just as large. As in the case of finance studies, our main explanatory variable is obtained as the interaction of an industry characteristic (service dependence) with a country characteristic (service regulation). The coefficient for this variable measures whether countries with lower service regulation grow relatively more in industries that depend more intensively on regulated services. Following Rajan and Zingales, we use country and industry fixed-effects to deal with various concerns arising in standard growth analysis (e.g., reverse causation and omitted variables).

[1] The reduction and harmonization of legal and administrative barriers is the main goal of the recent EU Services Directive, implemented at the end of 2009 and motivated by the concern for the knock-on effects that barriers in services may trigger "given the integration of services into manufacturing". The Third Legislative Package on Energy Markets is a controversial recent set of Directives by the Commission promoting the unbundling of network operation from supply and generation in energy. Similarly, the OECD recently recommended revising the energy regulatory framework in most member countries, and indicated the liberalization of professional services as a priority policy area for six European countries (including France, Germany and Italy), and Canada (OECD, 2009, *Going for Growth*).

By highlighting the relevance of service regulation for both value added and export growth our work closely relates to a growing literature on the relevance of institutions and policies for resource allocation and comparative advantages. Recent works focused on the ability to enforce written contracts. Nunn (2007) showed that countries with better contract enforcement specialize in contract intensive industries, those for which relationship-specific investment is more important. Levchenko (2007) found these countries also tend to export goods that, by requiring a large variety or range of inputs, are more institutionally dependent. In an earlier contribution, Claessens and Laeven (2003) explored the nexus between property rights protection and growth in industries that are more intensive in intangible assets, whose returns are more exposed to the actions of competitors. Looking at labour market institutions, Caballero *et al*. (2006) found that, in countries with strong rule of law, higher job security is associated with slower adjustment to shocks and lower productivity growth. Cuñat and Melitz (2007) found that countries with light regulation of employment relationships specialize in high-volatility industries. Against this background, our results emphasize the role of regulatory settings that are on top of competition policy agendas.

Two recent papers combined indexes of service regulation with input-output coefficients to estimate the impact of regulation and productivity growth (Conway *et al*., 2006; Arnold *et al*., 2008). Differently from us, they focus on the relevance of regulation for the transfer of technology to firms behind the productivity frontier, estimated exploiting the time series relationship between productivity in frontier and non-frontier countries. Their results indicate that regulation significantly slows technology transfers, and suggest that this happens, in particular, because it increases the costs of absorbing new technologies (as ICTs). Our interest on the patterns of specialization and trade requires that we focus on different specifications and outcomes. In line with the literature of reference, we also employ a different measure of regulatory impact (both papers use the recently issued OECD Regulation Impact Indicators, see Conway and Nicoletti, 2006). As we will see, such change turns out to have relevant empirical implications.[2]

Our results indicate that service efficiency matters for growth even in a restricted sample of high-income countries, for which the relationship between financial development and growth has previously been shown to be weak (Manning, 2003). We argue that this difference can be traced to our use of value added data at constant rather than current prices. To illustrate the point we use a simple theoretical framework in which countries produce differentiated goods and lower regulation raises output in service-intensive industries by reducing the service component of production costs. In this case there are two countervailing effects of lower regulation on nominal value added of service-intensive industries: a positive effect due to higher output and a negative effect due to lower prices. Estimates of the combined effect will therefore understate the impact of service regulation on production. We find empirical support for this hypothesis: lower regulation and higher financial development reduce the growth rate of (implicit) prices relatively more in service-intensive manufacturing industries. Accordingly, we do not find any significant effects of regulation or financial development on nominal value added growth.

1 Background

In this section we introduce a simple framework relating service regulation to the costs of production in downstream industries, and illustrate why regulation might affect industry specialization using insights from the recent trade literature. We start by considering an economy

[2] Three other papers used input-output linkages to study the consequences of upstream markets inefficiencies, but focused on specific countries. Allegra *et al*. (2004) looked at competition problems (as measured by the number of antitrust cases) and exports in Italian manufactures. Faini *et al*. (2006) focused on the link between regulation of network industries and productivity growth in Germany, Italy and the UK. Arnold *et al*. (2007) showed that barriers to FDI in services slowed TFP growth by Czech manufacturing firms.

with access to two production technologies j = 1, 2 combining labour (L) and an intermediate input Z, $y_j = Z_j^{\gamma_j} L_j^{1-\gamma_j}$. We assume that industry 1 is relatively more intensive in input Z: ($\Delta\gamma = \gamma_1 - \gamma_2 > 0$). The intermediate input is a composite of different production services $x(i)$, $Z_j = \left[\int_0^1 x(i)^\sigma di\right]^{\frac{1}{\sigma}}$, where $\sigma \in (0,1)$ determines the elasticity of substitution $\chi = 1/(1-\sigma)$ between varieties. Each variety is produced using one unit of labour, priced w. The price index of the composite service can be obtained from maximization conditions as $p_z = \left[\int_0^1 p(i)^{-\frac{\sigma}{1-\sigma}} di\right]^{-\frac{1-\sigma}{\sigma}}$, where $p(i)$ is the price of the i^{th} service.

Service regulation is introduced assuming that only a fraction $\varphi \in (0,1)$ of varieties can be bought at competitive prices, while the share $(1-\varphi)$ is available in regulated markets, where inputs are sold at monopolistic prices. This assumption implies that $p(i)=w$ if $i \in (0, \varphi)$, and $p(i)=w/\sigma$ when $i \in (\varphi, 1)$ and regulation grants monopoly profits to producers of service varieties.

The equilibrium price of the composite service becomes:

$$p_z = w\left[\varphi + (1-\varphi)\sigma^{\frac{\sigma}{1-\sigma}}\right]^{-\frac{1-\sigma}{\sigma}} = w\mathrm{C}(\varphi)$$

where $\mathrm{C}'(\varphi) < 0$, $C(\varphi)=1/\sigma > 1$ if φ=0 (fully regulated services) and $C(\varphi)$=1 if φ=1 (fully competitive services). The expression above implies that, given the unit cost function $c_j = p_z^{\gamma_j} w^{1-\gamma_j}$, the relative cost in the service intensive industry can be written as a decreasing function of the fraction of deregulated markets φ:

$$c_1/c_2 = p_z^{\Delta\gamma} w^{-\Delta\gamma} = C(\varphi)^{\Delta\gamma}.$$

To see how regulation can affect the equilibrium allocation of production and trade consider first the case of a small open economy taking world relative prices of final goods $p = p_1/p_2$ as given. In this case, the condition for diversification:

$$C(\varphi)^{\Delta\gamma} = p$$

identifies a threshold level of regulation $\varphi^*(p)$ such that any country would in general be fully specialized in production. If $\varphi^* \in (0,1)$, regulatory reforms raising the share of liberalized input markets above the threshold φ^* would imply a dramatic shift in the country production structure, from full specialization in labour intensive industries to full specialization in service intensive industries.

Less extreme predictions can be obtained following the modern trade literature to think of firms within each industry as supplying varieties of imperfectly substitutable goods (see Helpman and Krugman, 1985). For simplicity, varieties will be differentiated by country of origin (as in Armington, 1969). In this case, producers of country c in industry j will face a downward sloping world-demand curve $q_{j,c} = p_{j,c}^{-\varepsilon}\Omega_{j,c}$, where $p_{j,c}$ is the domestic price, and ε>1 is the constant elasticity of substitution across varieties. The scale variable $\Omega_{j,c}$ includes the amount of domestic and foreign expenditures allocated to industry j, which can be considered exogenous to the producer. Prices are set applying a constant mark up over marginal costs ($p_j = \mu p_z^{\gamma_j} w^{1-\gamma_j}$), so that the equilibrium relative production of the service intensive variety will be an increasing function of the share of liberalized service markets φ:

$$q_{1,c}/q_{2,c} = \Theta C(\phi)^{-\varepsilon * \Delta\gamma}$$

(recall that $C'(\varphi) < 0$). The elasticity of relative production to regulation is $\varepsilon_q = -\Lambda\varepsilon$ where ε is the price elasticity of demand and $\Lambda = \Delta\gamma * [1/(1-\chi)]$ measures the impact of a change in regulation on relative prices. In this framework, service deregulation would therefore imply an increase in the service intensive industry share of total production, driven by shifts in both domestic and foreign demand. From profit maximization one can derive that relative labour productivity in the service intensive industry is also increasing in the extent of deregulation.

Notice that if the value of production is measured at current prices (*i.e.*, $r_{j,c} = p_{j,c} q_{j,c}$) the above relation becomes $r_{1,c}/r_{2,c} = \tilde{\Theta} C(\phi)^{(1-\varepsilon) * \Delta\gamma}$. Because of the counteracting effects on prices, the elasticity of relative production to regulation $\varepsilon_r = \Lambda(1-\varepsilon)$ is therefore lower when production is measured at current rather than constant prices (and tends to zero as the substitutability across varieties ε decreases). Hence, an empirically interesting implication of this framework is that detecting the effects of regulation on the structure of industrial production would be easier using real as opposed to nominal measures of value added, as they allow insulating the industry accounts from the offsetting effects of deregulation on industry prices.

The framework above suggests that the process of service liberalization many developed countries started in the early 1990s should have implied a shift in the long run composition of production towards service intensive industries.[3] In the empirical part we will check whether such reallocation reflected in industry growth differentials by testing whether service intensive industries grew more in low regulation countries relative to less intensive service users. One reason for looking at growth rates is that production reallocation across industries is likely to be a lengthy process. A second reason is that such specification eases comparison of the results with those in the financial development literature, an important benchmark when studying the consequences of service underdevelopment.

2 Data and sample

All the data needed to perform our exercise are available from the OECD.[4] Information on value added, export and employment at the country-industry-year level is obtained from the STructural ANalysis (STAN) dataset. STAN has been assembled by the OECD complementing member countries' Annual National Accounts with information from other sources, such as national business surveys and censuses. The data are classified according to the International Standard Industrial Classification (ISIC) Rev. 3 industry list; they cover 17 countries and 15 manufacturing industries.

---

[3] An alternative way to model the role of services would be thinking of regulation as limiting the number of available input varieties in a model featuring increasing returns from specialization. Rodriguez-Clare (1996), Ciccone and Matsuyama (1996) and Rodrik (1996) are examples of papers showing that, with heterogeneous industry-intensity in non-traded intermediate inputs, the long run industry composition of a small open economy will significantly vary with the amount of locally produced inputs. As in the framework presented here, this occurs because the relative cost of service-intensive industries will decrease as the intermediate sector develops.

[4] See the Data Appendix and Table 1 for detailed variable definition and sources.

2.1 Measuring service regulation

Exposure of manufacturing industries to service regulation is measured combining country-level information on service regulation and industry-level data on service dependence. Specifically, our main indicator is the weighted average

$$SERVREG_{j,c} = \sum_{s}\left(w_{j,s} X_{c,s}\right)$$

where $X_{c,s}$ is an index of service regulation in sector s and country c, and $w_{j,s}$ captures industry j dependence on regulated services.

Cross-country measures of service regulation $(X_{c,s})$ are obtained from the OECD Product Market Regulation (PMR) database. We focused on four upstream service activities: energy (electricity and gas), communication (telecommunication and postal services), transportation (air, road, rail transportation services) and professional services (including accountants, architects, engineers and legal services). For each sector, the OECD codes a large amount of basic information on regulatory settings into quantitative scores increasing in the amount of restrictions to competition (see Conway and Nicoletti, 2006). Following Alesina *et al*. (2005), we only considered those scores designed to measure ex-ante anti-competitive restrictions: barriers to entry, vertical integration and market conduct.[5] While the OECD-PMR database covers regulation in energy, communication and transports since 1975, only two observations (in 1996 and 2003) are available for professions.

Two measures of industry j dependence on service s $(w_{j,s})$ were recovered from input-output account matrices. The first measure, capturing direct dependence, is obtained as the ratio between the cost of service inputs and the value of industry output (the so-called "technical coefficients"). The second is recovered from the inverse Leontief matrix, whose coefficients account for both direct and indirect contributions of service s to the value of production in industry j.[6] In our baseline specification, service dependence will be computed based on the US input-output tables (*i.e.*, $w_{j,s} = w^{US}_{j,s}$). As in the rest of the literature following Rajan and Zingales (1998), we therefore start assuming that US input-output coefficients reflect technological differences rather than country-specific determinants, as the level of regulation itself.[7] Accordingly, the US is excluded from the sample. In the robustness section, however, we will exploit the availability of country-specific weights taken from the OECD input-output database to construct an alternative measure of service dependence not reflecting input intensities that are specific to a country or a level of regulation (Ciccone and Papaioannou, 2006). As we will see, the two approaches produce very similar results.

2.2 Alternative measures

The OECD has recently made available a measure of the relevance of service regulation (the

---

5 Entry barriers include measures distorting the structure of markets relative to a competitive outcome, as the conditions for third party access to electricity and gas transmission grids, the existence of legal limitations on the number of competitors in communications or to the number services each profession has an exclusive right to provide. Vertical integration measures whether a priori competitive activities (as electricity generation or the final supply of energy) are separated from natural monopolies such as the national grid. Finally, conduct regulation includes restrictions on prices and fees, advertising, the form of business etc. in professional services.

6 These weights thus account for potential effects of anti-competitive service regulation working through industry j linkages with other, possibly non-regulated, industries in the economy. See the Data Appendix for more information on how the direct and indirect weights are obtained from the available input-output accounts.

7 In our data, the US is the country featuring the lowest average level of service regulation for the longest time period (see the figure in the Supplementary Appendix).

Regulation Impact Indicator, *RII*) constructed in a way similar to *SERVREG*. Specifically, the *RII* is obtained as $RII_{j,c} = \sum_s w^c_{j,s} X^{RII}_{c,s}$, where $w^c_{j,s}$ are country-specific input-output coefficients and $X^{RII}_{c,s}$ are measures of service regulation from the PMR database. Service sectors *s* include energy, communication, transportation and professional services (as in our measure) and retail trade. Recent papers used the *RII* to study the relation between regulation and technology transfer (see Conway *et al*., 2006; Arnold *et al*., 2008). Despite the obvious similarities, there are several reasons to expect the *RII* would be less appropriate than *SERVREG* to study the relevance of service regulation in our framework. First, as already discussed, the Rajan and Zingales approach requires that input-output coefficients should be a measure of technological determinants of service dependence. Such condition would be hardly met using country-specific input-output coefficients as they might reflect unobserved determinants of service dependence at the country level, introducing potentially relevant sources of bias. If, in particular, country-specific weights are a combination of technological service dependence and country-specific shocks that are independent of other model determinants, then they would tend to distort the estimated coefficients towards zero (attenuation bias).[8] Second, given the focus on the relevance of services as input providers, unlike the *RII* our indicator excludes retail trade from the list of regulated services. Because it does not cover wholesale activities, the OECD measure of retail regulation is in fact based on information that is unlikely to matter for downstream performance.[9] Finally, while the index $X^{RII}_{c,s}$ accounts for all regulatory areas covered by the OECD regulation database, including for example the extent of public ownership, we focused on measures capturing ex-ante anti-competitive practices (as barriers to entry). As we will see, comparing the results obtained using the two measures confirms our concerns regarding the appropriateness of using the *RII* in our framework.

Assembling the data imposes constraints on the number of available observations: in particular, we are forced to restrict the analysis to a relatively limited growth period, starting in 1996. The reason is twofold: first indicators of regulation in professional services are available at earlier dates; second, the number of missing entries in value added data significantly increases shifting to earlier dates, due to both the reduction in the number of available countries and to changes in industry classification within each country.[10]

The main variables used in the empirical part are summarized and described in Tables 1 to 3.

3 Results

3.1 Regulation and output growth

Table 4 reports the results obtained from our baseline value added growth regression:

$$\hat{VA}_{j,c} = \alpha + \beta\ SERVREG_{j,c} + \phi\ SHARE_{j,c} + \mu_c + \mu_j + \varepsilon_{j,c}$$

where $\hat{VA}_{j,c}$ is the average (1996-2002) real value added growth in industry j a nd country c

8 On the other hand if country-specific weights respond to country-level regulation, the error in measurement could be non-classical and the direction of the bias undetermined a priori (see Ciccone and Papaioannou, 2006).

9 The retail trade indicator covers restrictions as the existence of barriers to entry in food distribution, limits to shops opening hours and price controls on products as food, pharmaceutical, tobacco and gasoline. Such retail activities have a very low relevance as input to manufactures: according to the 1997 US "use" matrix their purchase represented 0.1 per cent of manufacturing production (against 5.7 per cent of wholesale trade). Notice also that the OECD input-output matrices we use throughout the paper do not separate retail from wholesale trade, and would thus have provided an inappropriate weight for trade regulation.

10 For example, as early as in 1990 the number of observations falls by nearly 25 per cent with respect to 1996.

Table 1

Variables Definition and Sources

Variable	Definitions and Sources
Industry Level	
$w_{j,s}$	Industry dependence on service s, computed on 1997 US Input-Output accounts. It includes energy ($w_{j,ENERGY}$), telecommunications and post ($w_{j,TLCPOST}$), transports ($w_{j,TRANSP}$) and and professional services ($w_{j,PROSERV}$). Source: our calculations. See also the Data Appendix.
ED_j	Industry dependence on external finance, defined as capital expenditure minus internal funds. Source: de Serres *et al*. (2006) on Thomson Financial Worldscope database.
$LABINT_j$	Industry labor intensity measured as the ratio between employees and total assets in the US in 1996. Source: OECD STAN database (total assets are computed from investment data using the perpetual inventory method with a 15% depreciation rate).
$GROP_j$	Annual compounded growth rate of production in real terms in industry *j* in USA over the 1996-2002 period. Source: OECD STAN database.
$\hat{w}_{j,s}$	Industry dependence on service *s* net of regulation- and country-specific determinants of inputs demand. For each of the four service sectors $\hat{w}_{j,s}$ have been estimated according to the following two-steps procedure (see also Ciccone and Papaioannou, 2006): (*a*) Regress country-specific input-output coefficients $w_{j,s,c}$ on country dummies, industry dummies and industry dummies interacted with country-level regulation in sector *s*; the most deregulated country $\overline{c}$ is excluded from the regression and the estimation follows Papke and Wooldridge (1996) to account for the fact that the dependent variable is fractional. (*b*) Obtain $\hat{w}_{j,s}$ as the fitted values of $w_{j,s,c}$ when regulation is set at country $\overline{c}$ levels and country fixed effects are set to zero. Country $\overline{c}$ is set to either the USA sectors (Table 7, column 5) or Great Britain (for energy and transport), USA (*TLCPOST*) and Finland (*PROSERV*) for *ENERGY* and *TRANSP*, *TLCPOST* and *PROSERV*, respectively (Table 7, column 6).
$GLOPP_{j,s}$	Estimated world-average industry growth opportunities. For each of the four service sectors global opportunities ($GLOPP_j$) are the estimated industry value added growth over the period 1996-2002 obtained according to the following two-steps procedure (see also Ciccone and Papaioannou, 2006): (*a*) Regress $GROWTH_{j,c}$ on country dummies, industry dummies and industry dummies interacted with country-level regulation in sector *s*; the USA are excluded from the regression. (*b*) Obtain $GLOPP_j$ as the predicted values of $GROWTH_{j,c}$ for the USA.
Country Level	
$X_{c,s}$	Regulation indexes on a 0-6 scale (from least to most restrictive conditions) in 1996 in four non-manufacturing industries. $X_{c,s}$ includes $X_{c,\ ENERGY}$, $X_{c,\ TLCPOST}$, $X_{c,TRANSP}$, $X_{c,PROSERV}$ referring to energy (electricity and gas), communications (posts and telecommunications), transports (air, rail and road), professional services (legal, accounting, engineering and architects). Source: OECD Product market Regulation database. $X_{c,\ ENERGY}$ takes into account entry barriers and the degree of vertical integration in electricity and gas supply; $X_{c,\ TLCPOST}$ accounts for entry barriers in postal and telecommunications services; $X_{c,TRANSP}$ accounts for entry barriers in air, rail and road services and on vertical integration in rail; $X_{c,PROSERV}$ accounts for entry barriers and the regulation of market conduct in legal services, accounting services, engineers and architects. See also Data Appendix.
FD_c	Financial development in country *c* measured as Private Credit by Deposit Money Banks over GDP in 1996. Source: World Bank's financial development and structure database (based on IMF's Financial Statistics).
$ACCSTAN_c$	Indicator of financial disclosure in 1983. Source: Rajan and Zingales (1998).
LMR_c	Indicator of employment protection in 1988-1995. Source: Fonseca and Utrero (2005).
$COST_c$	Direct start-up costs of obtaining legal status to operate a firm as a share of per capita GDP in 1999. Source: Djankov *et al.* (2002).

Table 1 (continued)

Variables Definition and Sources

Variable	Definitions and Sources
Industry – Country Level	
$GROWTH_{j,c}$	Annual compounded growth rate of real value added in industry *j* in country *c* over the 1996-2002 period. Source: OECD STAN database.
$NGROWTH_{j,c}$	Annual compounded growth rate of nominal value added in industry *j* in country *c* over the 1996-2002 period. Source: OECD STAN database.
$DEFGROWTH_{j,c}$	Annual compounded growth rate of the value added implicit deflator in industry *j* in country *c* over the 1996-2002 period. Source: OECD STAN database.
$LPGROWTH_{j,c}$	Annual compounded growth rate of labor productivity (value added at constant prices per employee) in industry *j* in country *c* over the 1996-2002 period. Source: OECD STAN database.
$EXGROWTH_{j,c}$	Annual compounded growth rate of exports at constant prices (current exports are deflated with the value added deflator) in industry *j* in country *c* over the 1996-2002 period. Source: OECD STAN database.
$SHARE_{j,c}$	Share of industry *j* in total value added in manufacturing in country *c* in 1996. Source: OECD STAN database.
$EXSHARE_{j,c}$	Share of industry *j* in exports in manufacturing in country *c* in 1996. Source: OECD STAN database.
$LLP_{j,c}$	Natural logarithm of labor productivity (value added at constant prices per employee) in industry *j* in country *c* in 1996. Source: OECD STAN database.
$SERVREG_{j,c}$	Index of exposure of manufacturing industry *j* to regulation in four service sectors (energy, communications, transport and professional services). It is computed as $\sum_s w_{j,s} X_{c,s}$ where *s* = *ENERGY*, *TLCPOST*, *TRASP*, *PROSERV*. Source: OECD Product market Regulation database and USA 1997 Input-Output accounts.
$DSERVREG_{j,c}$	Difference between $SERVREG_{j,c}$ in 1996 and in 2002. Source: OECD Product market Regulation database and USA 1997 Input-Output accounts.
$POWN_{j,c}$	Index of exposure of manufacturing industry *j* to the degree of public ownership in three service sectors (energy, communications, transport). It is computed as $\sum_s w_{j,s} POWN_{c,s}$ where $POWN_{c,s}$ is an index measuring on a 0-6 scale (increasing with the role of public sector) the degree of public ownership in 1996 and *s* = *ENERGY*, *TLCPOST* and *TRASP*. Source: OECD Product market Regulation database and USA 1997 Input-Output accounts.
$FDIREG_{j,c}$	Index of exposure of manufacturing industry *j* to restriction to foreign investment in four service sectors. It is computed as $\sum_s w_{j,s} Z_{c,s}$ where $Z_{c,s}$ are FDI restriction indicators in electricity, telecommunications, transport and professional services. Source: Koyama and Golub (2006) and USA 1997 Input-Output accounts.

Table 2

Summary Statistics

Variable	Obs	Mean	St. Dev.	Min	Max
Industry Level					
Dependence on energy [$w_{j,ENERGY}$]	15	0.018	0.010	0.007	0.039
Dependence on communications [$w_{j,TLCPOST}$]	15	0.004	0.001	0.002	0.007
Dependence on transports [$w_{j,TRANSP}$]	15	0.030	0.014	0.011	0.063
Dependence on professional services [$w_{j,PROSERV}$]	15	0.027	0.011	0.013	0.055
External dependence [ED_j]	15	0.697	1.595	–0.450	6.200
Labor intensity [$LABINT_j$]	15	0.028	0.013	0.004	0.052
Growth opportunities [$GROP_j$]	15	0.010	0.029	–0.028	0.093
Country Level					
Regulation in energy in 1996 [$X_{c,ENERGY}$]	16	4.475	1.338	1.808	6.000
Regulation in communications in 1996 [$X_{c,TLCPOST}$]	16	2.868	1.614	0.000	5.680
Regulation in transports in 1996 [$X_{c,TRASP}$]	16	2.949	1.062	1.530	5.133
Reg. in professional services in 1996 [$X_{c,PROSERV}$]	16	2.464	1.160	0.830	4.178
Financial development [FD_c]	16	0.718	0.272	0.304	1.141
Labor market regulation [LMR_c]	16	1.359	0.491	0.300	1.933
Red tape costs [$COST_c$]	16	0.146	0.141	0.012	0.586
Financial disclosure [$ACCSTAN_c$]	16	0.647	0.122	0.420	0.810
Industry – Country Level					
Value added growth 1996-2002 (real terms) [$GROWTH_{j,c}$]	220	0.018	0.034	–0.081	0.204
Val. added gr. 1996-2002 (nominal terms) [$NGROWTH_{j,c}$]	220	0.032	0.038	–0.123	0.221
Implicit deflator growth 1996-2002 [$DEFGROWTH_{j,c}$]	220	0.014	0.030	–0.095	0.189
Labor productivity growth 1996-2002 [$LPGROWTH_{j,c}$]	220	0.025	0.026	–0.051	0.162
Export growth 1996-2002 [$EXGROWTH_{j,c}$]	205	0.050	0.050	–0.094	0.194
Value added share in 1996 [$SHARE_{j,c}$]	220	0.069	0.047	0.001	0.234
Log labor productivity in 1996 [$LLP_{j,c}$]	220	3.864	0.481	2.821	6.932
Export share in 1996 [$EXSHARE_{j,c}$]	220	0.068	0.068	0.000	0.364
Service regulation [$SERVREG_{j,c}$]	220	0.246	0.109	0.070	0.628
Change in service deregulation [$DSERVREG_{j,c}$]	220	0.080	0.054	0.001	0.291

Table 3

Correlation Between Regulation Indicators in Four Service Sectors in 1996

	Energy [$X_{c,ENERGY}$]	**Communications** [$X_{c,TLCPOST}$]	**Transports** [$X_{c,TRASP}$]	**Prof. Serv.** [$X_{c,PROSERV}$]
Energy [$X_{c,ENERGY}$]	1.000			
Communications [$X_{c,TLCPOST}$]	0.549	1.000		
Transports [$X_{c,TRASP}$]	0.801	0.541	1.000	
Professional services [$X_{c,PROSERV}$]	0.497	0.519	0.645	1.000

Table 4

Service Regulation and Growth

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline, Direct Weights	**Baseline, Indirect Weights**	**Financial Development 1**	**Financial Development 2**	**Average 1996-02 Regulation**	**Deregulation (1996-2002)**
Service regulation	–0.172*	–0.170*	–0.176**	–0.158*	–0.198**	–0.287**
[$SERVREG_{j,c}$]	*(0.069)*	*(0.072)*	*(0.068)*	*(0.071)*	*(0.075)*	*(0.080)*
Financial dev. × external dep. [$FD_c \times ED_j$]			0.010* *(0.004)*		0.011* *(0.004)*	0.009* *(0.004)*
Accounting stand. × ext. dep. [$ACCSTAN_c \times ED_j$]				0.013+ *(0.007)*		
Change in service regulation [$DSERVREG_{j,c}$]						0.320** *(0.116)*
Initial industry share [$SHARE_{j,c}$]	0.189** *(0.071)*	0.198** *(0.069)*	0.169* *(0.067)*	0.187** *(0.072)*	0.174** *(0.066)*	0.163** *(0.062)*
Constant	0.037 *(0.023)*	0.048+ *(0.025)*	0.006 *(0.019)*	–0.001 *(0.020)*	0.005 *(0.019)*	0.014 *(0.019)*
Observations	220	220	220	220	220	220
R^2	0.66	0.66	0.67	0.67	0.68	0.69

+ significant at 10 per cent; * significant at 5 per cent; ** significant at 1 per cent.

Notes:
The dependent variable is the annual compounded growth rate of real value added at the country-industry level for the period 1996-2002 ($GROWTH_{j,c}$).
$SERVREG_{j,c}$ measures exposure to service regulation at the country-industry level as a weighted average ($\Sigma_s w_{j,s} * X_{c,s}$) of country-level anti-competitive regulation indexes from the OECD-PMR databases.
Service regulation ($X_{c,s}$) is measured in 1996 except in column 5, where it is the 1996-2002 average value.
Interaction weights $w_{j,s}$ are ("direct") technical coefficients of dependence between service sector s and manufacturing industry j computed on the 1997 USA Input-Output matrix, except for column 2, where they are measured to account for both direct and indirect dependence (see the Data Appendix for computational details).
Financial development is measured as Private Credit by Deposit Money Banks over GDP in 1996 (FD_c, column 3) and as accounting standards in 1983 ($ACCSTAN_c$, column 4). It is interacted with External dependence (ED_j), an industry-level measure of reliance on external finance obtained from US firm-level data. Both interactions follow Rajan and Zingales (1998). $DSERVREG_{j,c}$ measures exposure to service deregulation obtained as $\Sigma_s w_{j,s} * \Delta X_{c,s}$, where $\Delta X = X_{1996} - X_{2002}$ is the 1996-2002 change in regulation of service s in country c.
$SHARE_{j,c}$ indicates the industry share in total value added in manufacturing in 1996.
All regressions include country- and industry-fixed effects and use (employment) weighted least squares as estimation method.
Robust standard errors are reported in parentheses.

$SHARE_{j,c}$ is the beginning-of-period value-added industry share, and μ_c and μ_j a re country- and industry-specific fixed-effects. As explained above, $SERVREG_{j,c}$ captures differences in the relevance of service regulation in country c for each manufacturing industry j. Regulation indicators are measured in 1996. There is a negative link between regulation and growth if $\beta < 0$.

The coefficient reported in column 1 of Table 4 indicates that lowering beginning-of-period anti-competitive regulation in the provision of services has a significant and positive effect on growth. One way to get a sense for the size of this effect is thinking of the annual value added growth differential between an industry with overall service-dependence $\left(D_j = \sum_s w_{j,s} \right)$ at the 75th percentile (Pulp, paper and printing) and an industry at the 25th percentile (Fabricated metal products). The coefficient estimated in column 1 implies this differential would rise by approximately 0.75 per cent if regulation were to be uniformly lowered in the four services by an amount corresponding to the difference in average regulation between the 75th (France) to the 25th (Canada) most regulated countries. For comparison, the median value added growth rate in our sample is 1.8 per cent. This finding is confirmed irrespective of which of the two available measures of industry dependence on regulated services $(w_{j,s})$ we use. This can be seen in column 2, where we replicate the previous regression using the so-called Leontief transformation of the technical coefficients, thus accounting for both direct and indirect inter-sectoral relationships. While the point estimate is unchanged, the implied effect of service deregulation would be slightly larger (about 1 per cent) in this case.[11]

A first important robustness check for the above findings consists in accounting for the well-documented empirical nexus between finance and industry growth. This is obtained augmenting the baseline specification with two alternative measures of financial development, both proposed by Rajan and Zingales (1998). Column 3 focuses on the ratio of bank credit to GDP, while column 4 considers accounting standards. In both cases, the interaction term is US industry external finance dependence. Neither of the two variables affects the relevance of service regulation. On the other hand, financial development confirms as a significant growth determinant. The coefficient estimated in column 3, for example, implies the growth differential between an industry at the 75th percentile and one at the 25th percentile of external finance dependence (Plastic products and Pulp and paper, respectively) would increase of approximately 0.2 per cent moving from a country with private credit at the 25th percentile to a country close to the 75th percentile of financial development (Norway and the Netherlands, respectively).

The last two columns in Table 4 test the robustness of our estimate to changes in the regression specification. In column 5 we focus on the relationship between industry growth and average (as opposed to initial) service regulation in 1996-2002 using initial regulation as instrument, an approach recently followed in the financial development literature. Results are slightly stronger than in previous specification. Finally, in column 6 we account for the possibility that our estimates are at least in part capturing the effects of changes in regulation occurred between 1996 and 2002. This would be the case if countries with high initial regulation implemented relatively stronger subsequent deregulation processes, and regulation has level-effects on value added. We checked for this possibility augmenting the regression with a measure ($DSERVREG = SERVREG_{96} - SERVREG_{02}$) that is increasing in the extent of deregulation. The positive and significant coefficient attracted by *DSERVREG* does in fact indicate that, holding

11 The positive coefficient we estimate on initial shares, indicating that countries tend to experience relatively faster growth in those industries they are more specialized in, is in contrast with results obtained by most of the comparable literature. While apparently puzzling, this finding can be explained by the large weight Western European countries have in our sample. The recent intense process of economic and monetary integration seems in fact to have resulted in increased industrial specialization in these countries (see Midelfart *et al*., 2003).

beginning-of-period regulation constant, value added growth in service intensive industries benefits from higher deregulation.[12] But our baseline estimate is, if anything, larger than in previous specifications.

3.2 Output and price effects

Several works adopting the Rajan-Zingales approach noticed that the empirical relevance of the finance-growth nexus is subject to strong variability depending on the countries included in the sample (Favara, 2003), and loses statistical significance as developing countries are omitted (Carlin and Mayer, 2003; Manning, 2003).[13] Building on time-series results as those in Rousseau and Wachtel (1998), one proposed explanation for this finding is that alternative financial instruments (as equity, debt, and derivative markets) may substitute for credit availability in advanced economies. But the significant coefficients we estimated in Table 4, obtained examining a sample of OECD countries, suggest we should look for a different explanation.

In a world where high-income countries tend to produce differentiated goods, one way to reconcile our findings with the literature is thinking of a possible counteracting role of prices. While we look at the growth of output (as measured by value added at constant prices), most of the existing cross-country cross-industry papers use nominal value added data. As shown at the end of Section 2, if lower regulation raises output in service-intensive industries by lowering the service component of the cost of production, then there are two countervailing effects on nominal value added: a positive effect due to higher output and a negative effect due to lower prices. Their combination will tend to weaken the relation between service underdevelopment and industry output when this is measured in nominal terms.

We explore this issue in greater detail in Table 5, estimating the effects of regulation on industry prices. We do in fact find that, among OECD countries, lower regulation and higher financial development translate into lower prices in service-intensive manufacturing industries (Table 5, columns 1 to 3). As a result, when we replicate the real value added analysis of Table 4 using nominal value added the effect becomes, as in above mentioned works, largely insignificant (Table 5, columns 4 to 6). Even so, the issue remains of why using nominal output does allow estimating significant effects when the sample includes a large share of less developed countries. According to the above argument, one possibility is that less developed countries produce more homogeneous commodities relative to advanced countries, facing a higher elasticity of demand. In this case, the counteracting effect of prices would become less and less relevant, on average, as the share of developing countries in the sample increases allowing to recover significant estimates even with nominal data.

3.3 Regulation, productivity and exports

Does lower regulation improve productive efficiency or are the estimated value added growth differentials absorbed by offsetting shifts in industry employment? Despite its relevance, the interaction between service regulation and labour productivity has so far received relatively

---

[12] To get a sense for the size of this effect, consider the comparison between a country with deregulation at the 75th percentile (e.g., Germany) and a country at the 25th percentile (e.g., Japan). Our estimates imply an annual growth gap between the industry at the 75th and the industry at the 25th percentile of service-intensity of nearly 1 per cent.

[13] Using the same dataset (UNIDO Industrial Statistics) and regression specification of Rajan and Zingales (1998) we found, for example, that their baseline estimate (0.118, with a standard deviation of 0.037, see Table 4, column 2 of Rajan and Zingales, 1998) falls to –0.004 (0.019) when the analysis is restricted to OECD countries, and to –0.021 (0.017) when further focusing on the sub-sample of developed countries we use here.

Table 5

Financial Development, Prices and Nominal Growth

	Prices			Nominal Growth		
	(1)	(2)	(3)	(4)	(5)	(6)
	Service Regulation	**Financial Development**	**Regulation and Fin. Dev.**	**Service Regulation**	**Financial Development**	**Regulation and Fin. Dev.**
Service regulation [$SERVREG_{j,c}$]	0.210^{**} *(0.072)*		0.211^{**} *(0.070)*	–0.004 *(0.078)*		–0.006 *(0.078)*
Financial dev. × external dep. [$FD_c \times ED_j$]		-0.009^{*} *(0.005)*	-0.009^{*} *(0.004)*		0.005 *(0.005)*	0.005 *(0.004)*
Initial industry share [$SHARE_{j,c}$]				0.027 *(0.049)*	0.017 *(0.048)*	0.017 *(0.050)*
Constant	0.015 *(0.014)*	0.056^{**} *(0.006)*	0.019 *(0.013)*	0.037^{*} *(0.016)*	0.036^{**} *(0.012)*	0.037^{*} *(0.016)*
Observations	220	220	220	220	220	220
R^2	0.62	0.60	0.63	0.64	0.64	0.64

$^{+}$ significant at 10 per cent; * significant at 5 per cent; ** significant at 1 per cent.

Notes:
In columns 1-3 the dependent variable is the annual compounded growth rate of value added implicit deflator at the industry-country level for the period 1996-2002 ($DEFGROWTH_{j,c}$); in columns 4-6 the dependent variable is the annual compounded growth rate of nominal value added at the industry-country level for the period 1996-2002 ($NGROWTH_{j,c}$). $SERVREG_{j,c}$ measures exposure to service regulation at the country-industry level as a weighted average ($\Sigma_s w_{j,s} {*} X_{c,s}$) of country-level anti-competitive regulation indexes from the OECD-PMR databases. Service regulation ($X_{c,s}$) is measured in 1996. Interaction weights $w_{j,s}$ are ("direct") technical coefficients of dependence between service sector *s* and manufacturing industry *j* computed on the 1997 USA Input-Output matrix. Financial development is measured as Private Credit by Deposit Money Banks over GDP in 1996 (FD_c) and is interacted with External dependence (ED_j), an industry-level measure of reliance on external finance obtained from US firm-level data. $SHARE_{j,c}$ indicates the industry share in total value added in manufacturing in 1996. All regressions include country- and industry-fixed effects and use (employment) weighted least squares as estimation method. Robust standard errors are reported in parentheses.

little empirical attention. Our cross-country and industry results indicate that service regulation has a significant impact on the growth rate of value added per worker in service intensive industries (see Table 6, panel A). This finding is robust to accounting for financial development or by changing the regression specification, as in Table 4. To get a sense for the economic relevance of the estimated coefficients, consider the annual productivity growth differential between Pulp and paper and Fabricated metal products (the two industries at the 75th and 25th percentile of the distribution of service-dependence, respectively). The coefficient in column 1 implies this growth differential is approximately 0.9 per cent larger in a low than in a high regulation country (respectively Canada and France). For comparison, the median productivity growth rate in our sample is 2.2 per cent.

Table 6

Service Regulation, Productivity and Exports

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline, Direct Weights	**Baseline, Indirect Weights**	**Financial Development 1**	**Financial Development 2**	**Average 1996-02 Regulation**	**Deregulation (1996-2002)**
Panel A: Productivity Growth						
Service regulation	-0.201^{*}	-0.218^{*}	-0.202^{*}	-0.194^{*}	-0.228^{*}	-0.280^{**}
[$SERVREG_{j,c}$]	*(0.081)*	*(0.100)*	*(0.080)*	*(0.085)*	*(0.090)*	*(0.106)*
Financial dev. × external dep.			0.009		0.010	0.008
[$FD_c \times ED_j$]			*(0.006)*		*(0.006)*	*(0.005)*
Accounting stand. × ext. dep.				0.006		
[$ACCSTAN_c \times ED_j$]				*(0.006)*		
Change in service regulation						0.228
[$DSERVREG_{j,c}$]						*(0.158)*
Initial labor productivity	0.031^{**}	0.032^{*}	0.028^{*}	0.031^{**}	0.030^{**}	0.030^{**}
[$LLP_{j,c}$]	*(0.012)*	*(0.012)*	*(0.012)*	*(0.012)*	*(0.011)*	*(0.011)*
Constant	-0.079^{+}	-0.062	-0.069	-0.082^{+}	-0.073	-0.066
	(0.047)	*(0.048)*	*(0.047)*	*(0.047)*	*(0.046)*	*(0.045)*
Observations	220	220	220	220	220	220
R^2	0.58	0.58	0.59	0.58	0.60	0.60
Panel B: Export Growth						
Service regulation	-0.213^{+}	-0.249^{*}	-0.215^{*}	-0.202^{+}	-0.242^{*}	-0.297^{*}
[$SERVREG_{j,c}$]	*(0.108)*	*(0.111)*	*(0.106)*	*(0.108)*	*(0.119)*	*(0.121)*
Financial dev. × external dep.			0.005		0.006	0.005
[$FD_c \times ED_j$]			*(0.007)*		*(0.007)*	*(0.006)*
Accounting stand. × ext. dep.				0.010		
[$ACCSTAN_c \times ED_j$]				*(0.013)*		
Change in service regulation						0.229
[$DSERVREG_{j,c}$]						*(0.179)*
Initial industry export share	-0.013	-0.007	-0.017	-0.015	-0.016	-0.024
[$EXSHARE_{j,c}$]	*(0.054)*	*(0.053)*	*(0.052)*	*(0.055)*	*(0.052)*	*(0.050)*
Constant	0.060^{**}	0.081^{**}	0.059^{**}	0.055^{**}	0.007	0.070^{**}
	(0.018)	*(0.025)*	*(0.018)*	*(0.019)*	*(0.023)*	*(0.019)*
Observations	205	205	205	205	205	205
R^2	0.72	0.72	0.72	0.72	0.72	0.72

$^{+}$ significant at 10 per cent; * significant at 5 per cent; ** significant at 1 per cent.

Notes to Table 6:
In Panel A, the dependent variable is the annual compounded growth rate of labor productivity (value added per employed worker) at the industry-country level for the period 1996-2002 ($LPGROWTH_{j,c}$).
In Panel B, the dependent variable is the annual compounded growth rate of exports at the industry-country level for the period 1996-2002 ($EXPGROWTH_{j,c}$).
$SERVREG_{j,c}$ measures exposure to service regulation at the country-industry level as a weighted average ($\Sigma_s\ w_{j,s}*X_{c,s}$) of country-level anti-competitive regulation indexes from the OECD-PMR databases.
Service regulation ($X_{c,s}$) is measured in 1996 except in column 5, where it is the 1996-2002 average value. Interaction weights $w_{j,s}$ are ("direct") technical coefficients of dependence between service sector s and manufacturing industry j computed on the 1997 USA Input-Output matrix, except for column 2, where they are measured to account for both direct and indirect dependence (see the Data Appendix for computational details).
Financial development is measured as Private Credit by Deposit Money Banks over GDP in 1996 (*FDc*, column 3) and as accounting standards in 1983 ($ACCSTAN_c$, column 4). It is interacted with External dependence (*EDj*), an industry-level measure of reliance on external finance obtained from US firm-level data. Both interactions follow Rajan and Zingales (1998).
$DSERVREG_{j,c}$ measures exposure to service deregulation obtained as $\Sigma_s\ w_{j,s}*\Delta X_{c,s}$, where $\Delta X = X_{1996} - X_{2002}$ is the 1996-2002 change in regulation of service s in country c.
$LLP_{j,c}$ indicates the log of labor productivity in 1996. $EXSHARE_{j,c}$ indicates the industry share in total exports in manufacturing in 1996.
All regressions include country- and industry-fixed effects and use (employment) weighted least squares as estimation method. Robust standard errors are reported in parentheses.

Finally, we exploited the availability of industry data on exports to explore whether the sectoral reallocation patterns implied by our value added results correspond to changes in international specialization. Results reported in panel B of Table 6 indicate that service regulation is an important determinant of comparative advantages. Throughout all the empirical specifications adopted in the previous tables we find that exports by service intensive industries tend to grow disproportionately more in countries with low levels of service regulation. The usual thought experiment yields an increase of about 1 per cent in the 25th-75th industry growth differential following a reduction in regulation.

All in all, our empirical findings point to the existence of non-negligible indirect effects of lack of competition in upstream markets for the patterns of international specialization and comparative advantages.

4 Robustness

Having established our baseline findings, we proceeded to a number of robustness checks considering the potential confounding role of regulation in other markets, the appropriateness of US weights as a measure of service dependence, the role of influential observations and the suitability of our measure of regulation impact compared to other possible measures.

4.1 The role of product and labour market regulation

We first considered the possibility that our estimates are driven by omitted country-industry shocks not captured by either country or industry fixed-effects and correlated with service regulation. If regulation is a countrywide phenomenon, our findings might in particular be capturing anti-competitive measures targeting other markets, as the labour or the product market.

We checked for this possibility augmenting the baseline specification with regulation-related variables, which have been shown to significantly affect industry growth. In columns 1 and 2 of Table 7 we accounted for country-level measures of employment protection and administrative (red-tape) barriers to entrepreneurships (Djankov *et al.*, 2002; Nicoletti and Scarpetta, 2003;

Table 7

Robustness Check

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Other Regulation Measures					**Weights**		**Influential Obs.**	
	Employee Protection	**Red Tape**	**FDI Regulation**	**Public Ownership**	**All**	**IV - US**	**IV-lowest Country**	**Most/least Dependent Industries**	**Most/least Regulated Countries**
Service regulation	–0.191**	–0.203**	–0.182**	–0.232**	–0.272**	–0.193*	–0.218*	–0.274**	–0.180*
[$SERVREG_{j,c}$]	*(0.071)*	*(0.067)*	*(0.068)*	*(0.074)*	*(0.073)*	*(0.087)*	*(0.105)*	*(0.088)*	*(0.072)*
Fin. dev. × external dep.	0.011**	0.010*	0.011**	0.010**	0.012**	0.010*	0.010*	0.008*	0.013**
[$FD_c \times ED_j$]	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*	*(0.004)*
Lab. market reg. × lab. int.	–0.400				–0.578^{+}				
[$LMR_c \times LABINT_j$]	*(0.323)*				*(0.304)*				
Red tape costs × gr. opp.		–1.449+			–1.599^{+}				
[$COST_c \times GROP_j$]		*(0.871)*			*(0.871)*				
FDI restrictions			0.879		0.807				
[$FDIREG_{j,c}$]			*(0.794)*		*(0.806)*				
Public ownership				0.084^{+}	0.059				
[$POWN_{j,c}$]				*(0.047)*	*(0.047)*				
Initial industry share	0.182**	0.135^{+}	0.167*	0.152*	0.135*	0.167*	0.165*	0.155*	0.184*
[$SHARE_{j,c}$]	*(0.067)*	*(0.069)*	*(0.067)*	*(0.063)*	*(0.066)*	*(0.068)*	*(0.068)*	*(0.071)*	*(0.073)*
Constant	0.019	0.015	–0.009	0.003	0.020	0.039	0.014	0.055*	0.035
	(0.024)	*(0.019)*	*(0.022)*	*(0.019)*	*(0.025)*	*(0.025)*	*(0.026)*	*(0.027)*	*(0.024)*
Observations	220	220	220	220	220	220	220	188	193
R^2	*0.68*	*0.68*	*0.68*	*0.68*	*0.70*	*0.67*	*0.67*	*0.67*	*0.69*

$^{+}$ significant at 10 per cent; * significant at 5 per cent; ** significant at 1 per cent.

Notes to Table 7:
The dependent variable is the annual compounded growth rate of real value added at the industry-country level for the period 1996-2002 ($GROWTH_{j,c}$). $SERVREG_{j,c}$ measures exposure to service regulation at the country-industry level as a weighted average ($\Sigma_s w_{j,s}*X_{c,s}$) of country-level anti-competitive regulation indexes from the OECD-PMR databases. Service regulation ($X_{c,s}$) is measured in 1996. Interaction weights $w_{j,s}$ are ("direct") technical coefficients of dependence between service sector *s* and manufacturing industry *j* computed on the 1997 USA Input-Output matrix.
Financial development is measured as Private Credit by Deposit Money Banks over GDP in 1996 (FD_c) and is interacted with External dependence (ED_j), an industry-level measure of reliance on external finance obtained from US firm-level data.
Labor market regulation (LMR_c) is an indicator of employment protection in 1988-95 and is interacted with labor intensity ($LABINT_j$) computed as the ratio between employees and total assets in the USA in 1996. Red tape costs ($COST_c$) are direct start-up costs of obtaining legal status to operate a firm as a share of per capita GDP in 1999; this variable is interacted with growth opportunities ($GROP_j$) measured as the growth rate of real value added growth in industry *j* in USA over the 1996-2002 period.
FDI restrictions in services ($FDIREG_{j,c}$) is an index of exposure of each manufacturing industry *j* to the degree of FDI regulation in four service sectors (electricity, telecommunications, transport and professional services). It is computed as $\sum_s w_{j,s} Z_{c,s}$ where *s = ELECTRICITY*, *TLCPOST*, *TRASP*, *PROSERV* where $Z_{c,s}$ are FDI restriction indicators on a 0-1 scale (increasing with the degree of restrictiveness). Weights $w_{j,s}$ are the technical coefficients computed on the USA 1997 Input-Output matrix (see also Data Appendix). Public ownership ($POWN_{j,c}$) is an index of exposure of industry *j* to the degree of public ownership in services. It is computed as $\sum_s w_{j,s} POWN_{c,s}$ where $POWN_{c,s}$ is an index measuring on a 0-6 scale (increasing with the role of public sector) the degree of public ownership in 1996 and *s = ENERGY*, *TLCPOST* and *TRASP*.
Columns 6 and 7 report IV estimates obtained using $\sum_s \hat{w}_{j,s} X_{c,s}$ as instrument for $SERVREG_{j,c}$. $\hat{w}_{j,s}$ is the estimated industry *j*'s dependence on service *s* net of regulation- and country-specific determinants of factor demand.
See Table 1 and Section 5 in the main text for more information on the IV approach.
Results in columns 8 and 9 are obtained removing from the sample the most and least intensive industrial users of regulated services ("Other non-metallic mineral products" and "Machinery and equipment N.E.C.") and the most and least service-regulated countries (Greece and Sweden), respectively.
$SHARE_{j,c}$ is the industry share in total value added in manufacturing in 1996.
All regressions include country- and industry-fixed effects and use (employment) weighted least squares as estimation method.
Robust standard errors are reported in parentheses.

Bassanini *et al*. 2009). Both variables are negatively related to industry growth, although the relationship is statistically significant only in the case of labour market regulation. On the other hand, the estimated impact of services regulation is unaffected. The next two columns show that our results are robust to accounting for alternative forms of regulation in services, as restrictions to foreign direct investment (column 3), or the extent of public ownerships in energy, transportation and communication services (column 4). Finally, column 5 shows robustness to accounting for all regulation variables simultaneously. The Supplementary Appendix reports further robustness checks to alternative channels highlighted by the literature on the determinants of international specialization and comparative advantages.[14]

4.2 The appropriateness of US weights

We next dealt with the possibility that using input-output weights from a benchmark country does not allow to correctly measure technological dependence on service inputs because country-specific weights differ from "true" weights by a idiosyncratic component. Such component could be unrelated to other determinants of industry growth, a case in which our estimates would be subject to standard attenuation bias, or depend on the level of regulation itself, so that using a

[14] In particular, we show our estimates are unaffected when accounting for the role of human and physical capital (as in Ciccone and Papaioannou, 2007) and property rights (Claessens and Laeaven, 2003) in value added growth regressions; and for the role of institutional quality and contract enforcement in export regressions (we used the same specifications as in Levchenko, 2007 and Nunn, 2007, respectively).

benchmark country would induce a priori ambiguous biases in the estimated coefficients (Ciccone and Papaioannou, 2006). These considerations suggest that neither choosing a different benchmark country nor using an average of input-output weights recovered from multiple sources would solve the measurement problems. An alternative procedure consists in recovering a measure of average service-dependence not reflecting input intensities specific to a country or to a level of regulation, and use it as an instrument for the benchmark-country index of service-dependence. Following Ciccone and Papaioannou (2006), one such measure was estimated for each service sector *s* in two steps. First, we regressed country-industry weights $w_{j,c}$ on country dummies, industry dummies and industry dummies interacted with country-level regulation in sector *s*, to estimate the marginal effect of regulation on industry dependence: $w_{j,c} = \mu_j + \mu_c + \delta_j X_c + \varepsilon_{j,c}$.[15] In this regression, the most deregulated country $\bar{c}$ is excluded from the sample. Second, we estimated $\hat{w}_{j,\bar{c}}$ as the fitted values of $w_{j,c}$ when regulation is set at the minimum observed value $(X_{\bar{c}})$ and country-specific averages are set to zero: $\hat{w}_j = \hat{\mu}_j + \hat{\delta}_j X_{\bar{c}}$. The fitted weights $\hat{w}_{j,\bar{c}}$ will therefore not reflect input intensities that are regulation or country-specific, and can be used as instruments for US weights in the empirical specification.

The results obtained following this procedure are reported in columns 6 and 7 of Table 7 and confirm the negative role of anti-competitive service regulation for growth. The only difference between the two columns consists in the choice of the country excluded from the service-specific first stage regressions. In column 6, we excluded the US, the country with the lowest levels of regulation from an historical perspective. In column 7, we excluded the least regulated country in each service sector in 1996 (the US for communications, the UK for energy and transportation, Finland for professional services).

4.3 *The role of influential observations*

The last two columns of Table 8 report results obtained removing from the sample the most and the least service intensive industries (Other non metallic mineral products and Machinery and equipment, respectively; column 8), and the most and the least regulated countries (Greece and Sweden, respectively; column 9). The estimated coefficient on the growth effect of service regulation is robust to both exercises.

4.4 *Alternative definitions of regulation impact*

Two recent papers used the OECD Regulation Impact Indicator (*RII*) described in Section 3 to estimate the effect of regulation on productivity growth in a time-series framework (Conway *et al.*, 2006; Arnold *et al.*, 2008). In their analyses, productivity growth in an industry is expressed as a function of regulation and of the industry "technological distance" from the frontier (*i.e.*, from the country with the highest productivity level).[16] The latter variable, a measure of the potential for technology transfer, allows estimating the speed of convergence to the productivity leader. In this context, regulation is allowed for both direct and indirect (*i.e.*, through the speed of convergence) effects on growth. Both papers find that higher regulation hinders productivity growth by slowing the speed of convergence to the technological frontier. In the sub-sample of ICT

[15] The regressions account for the fact that the dependent variable is fractional (Papke and Wooldridge, 1996).

[16] The empirical analysis moves from a first-order autoregressive distributed lag model [ADL(1,1)] where own productivity is cointegrated with frontier productivity. In the long run, this has an Error Correction Model (ECM) representation, which is the relationship estimated in the two papers.

intensive (mainly service) industries they also find evidence of direct effects of regulation on growth.

Despite the two works differ from ours in many dimensions, it is important to empirically assess the relevance and robustness of our findings against the OECD Regulation Impact Indicator. In the Supplementary Appendix we report results obtained when (*a*) the *RII* replaces *SERVREG* in our baseline specifications, and (*b*) the *RII* is added to our baseline specifications. The results suggest that the OECD indicator tends to understate the relevance of service regulation for industry growth, thus confirming our concerns regarding its appropriateness in our framework (see Section 3). On one hand, using the *RII* as main explanatory variable yields to estimate non-significant effects of regulation on two out of three of the outcomes we focus on (productivity and exports). When significant, the coefficient estimated using the *RII* implies much lower gains from deregulation with respect to what we obtained using *SERVREG*. In particular, the implied effect of a one standard deviation reduction in regulation on value added industry growth would be nearly 50 per cent lower. Finally, all estimates obtained using *SERVREG* are robust to contemporaneously adding the *RII*, whose impact on growth is not statistically significant (or even positive).[17]

5 Extensions

To further qualify the role of service regulation in the next sections we focus on two potential dimensions of heterogeneity in the estimated average coefficient: by size of the regulated market and by regulated service.

5.1 Service regulation and country size

The benefits from lower regulation might vary with the extent of the regulated market. Recent cross-country evidence by Hoekman *et al*. (2004) showed, for example, that the positive relation between entry barriers and average mark-ups in manufacturing is substantially higher in large than in small countries. In a world with imperfect competition and fixed costs of production this would happen if the level of existing regulatory barriers (e.g., licenses) is such that there is greater scope for profitable entry in larger than in smaller economies. In our setting, the positive effects of lower service regulation could therefore be stronger in countries characterized by a larger extent of demand by downstream industries.

We checked for this possibility splitting the sample in two groups of large and small OECD countries. Large countries account for nearly 90 per cent of total manufacturing employment in our data.[18] Table 8 reports the results obtained estimating alternative specifications of the value added growth regression in the two sub-samples and compares it to the average coefficient. In all cases, our evidence indicates that previous results are determined by the positive growth effects of lower

[17] The Supplementary Appendix also reports results obtained considering a third measure of regulation impact, computed to highlight the relevance of using benchmark-country (or "global") indicators of service dependence. Such measure is obtained interacting the ex-ante anti-competitive regulation index we use throughout the paper $(X_{c,s})$ with country-specific input-output weights $(w^c_{j,s})$, as in the *RII*. Using this "mixed" regulation index yields statistically significant effects on value added and productivity, but not on export growth. The implied effects of a one-standard deviation reduction in regulation is slightly higher than in the case of *RII*, but still nearly a half of what would be obtained using *SERVREG*. Finally, the estimates obtained using *SERVREG* are robust to adding the "mixed" regulation indicator, which in turn has very little statistically significance in all specifications.

[18] The sample of large countries include Canada, France, Germany, Italy, Japan, the Netherlands, Spain and the UK; small countries are Austria, Belgium, Denmark, Finland, Greece, Norway, Portugal and Sweden. The cross-country variability of our measure of service regulation is very similar in the two sub-samples (and close to the value for the whole sample).

Table 8

Service Regulation, Growth and Country Size

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Baseline			**Including Financial Development**			**Including All Controls**		
	All Countries	**Large Countries**	**Small Countries**	**All Countries**	**Large Countries**	**Small Countries**	**All Countries**	**Large Countries**	**Small Countries**
Service regulation	-0.172^{*}	-0.191^{*}	0.019	-0.176^{**}	-0.182^{*}	0.096	-0.272^{**}	-0.313^{**}	0.107
[$SERVREG_{j,c}$]	*(0.069)*	*(0.080)*	*(0.131)*	*(0.068)*	*(0.078)*	*(0.141)*	*(0.073)*	*(0.086)*	*(0.174)*
Initial industry share	0.189^{**}	0.262^{**}	0.072	0.169^{*}	0.226^{*}	0.090	0.135^{*}	0.200^{*}	0.048
[$SHARE_{j,c}$]	*(0.071)*	*(0.098)*	*(0.055)*	*(0.067)*	*(0.092)*	*(0.055)*	*(0.066)*	*(0.092)*	*(0.047)*
Constant	0.037	0.051^{*}	0.026	0.006	0.049^{*}	0.021	0.020	0.036	0.033
	(0.023)	*(0.023)*	*(0.018)*	*(0.019)*	*(0.022)*	*(0.016)*	*(0.025)*	*(0.038)*	*(0.030)*
Observations	220	113	107	220	113	107	220	113	107
R^2	0.66	0.70	0.52	0.67	0.72	0.55	0.70	0.75	0.59

$^{+}$ significant at 10 per cent; * significant at 5 per cent; ** significant at 1 per cent.

Notes:
The dependent variable is the annual compounded growth rate of real value added at the country-industry level for the period 1996-2002 ($GROWTH_{j,c}$).
SERVREGj,c measures exposure to service regulation at the country-industry level as a weighted average (Σs *wj,s*Xc,s*) of country-level anti-competitive regulation indexes from the OECD-PMR databases.
Service regulation (*Xc,s*) is measured in 1996. Interaction weights *wj,s* are ("direct") technical coefficients of dependence between service sector *s* and manufacturing industry *j* computed on the 1997 USA Input-Output matrix.
Specifications in columns 4-9 include (unreported) controls for financial development [*FDc* × *EDj*]. In columns 7-9 further account for (unreported) Labour market regulation [*LMRc* × *LABINTj*], Red tape costs [*COSTc* × *GROPj*], FDI restrictions [*FDIREGj,c*] and Public ownership [*POWNj,c*] (see Table 1 for the definition of these variables).
Large countries are Canada, France, Germany, Italy, Japan, the Netherlands, Spain and the UK; small countries include Austria, Belgium, Denmark, Finland, Greece, Norway, Portugal and Sweden.
All regressions include country- and industry-fixed effects and use (employment) weighted least squares as estimation method.
Robust standard errors are reported in parentheses.

regulation in the sub-set of larger countries, suggesting these economies should expect substantial payoffs from competition policies. For example, the coefficient estimated in column 8 implies that the annual growth differential between Pulp and paper and Fabricated metal products (the two industries at the 75th and 25th percentile of the distribution of service-dependence, respectively) would rise by nearly 1.4 per cent if regulation in a large and highly regulated country as France was lowered to the level of Canada. On the other hand, the estimates are largely insignificant in the case of smaller economies.[19]

5.2 *Sector-specific effects*

We allow for sector-specific effects focusing on the unrestricted specification:

$$\hat{V}A_{j,c} = \alpha + \sum_{s} \beta_s \left(w_{j,s} X_{c,s} \right) + \phi SHARE_{j,c} + \mu_c + \mu_j + \varepsilon_{j,c}$$

The coefficients β_s are easier interpreted recalling they represent a second derivative $\beta = \partial \hat{V}A / \partial w \partial X$. Hence, $\beta_s<0$ indicates that, other things equal, intensive users of service s fare better in those countries where the provision of such service is relatively less regulated.

Our results, reported in Table 9, point to the existence of significant sectoral heterogeneity underlying the aggregate estimates presented in previous tables. This can be seen in columns 1 to 4, where we separately considered the role of energy, professional services, communication and transportation services, respectively. All estimated coefficients are negative, but only the first two are statistically significant, a result confirmed when all regressors are jointly considered (column 5). In both cases, the implied effect of regulation is non-negligible. Consider, for example, the annual value added growth differential between an industry with an intensity in professional services at the 75th percentile (Textile and textile products) and an industry at the 25th percentile (Transport equipment). The estimated coefficient in column 5 implies this growth differential is approximately 0.8 per cent higher in a country with regulation of professions at the 25th percentile (as the UK) than in a country close to the 75th percentile (as Spain). This effect is large relative to the median industry value-added growth rates in our sample (1.8 per cent) and represents more than one-third of the observed 25th-75th difference in industry growth rates. In the case of energy, moving from a heavily regulated (e.g., Italy) to a deregulated (e.g., Finland) country would imply an even larger effect on the industry growth differential (1.4 per cent).[20]

All specifications already account for the possibility of contemporaneous effects from labour and product market regulation. In column 6, we further checked for the potential confounding role of short run shocks. This amounts to distinguishing whether low regulation induces faster growth by service intensive industries or rather facilitates downstream firms exploiting industry-level worldwide short run shocks. While still of interest, evidence in favour of the second mechanism would imply that absent these shocks, deregulation would have no effects on growth. Fisman and Sarria-Allende, (2004) raised this point in the case of finance, suggesting a test for robustness to short run shocks obtained interacting the country-level variable of interest with a direct measure of worldwide industry-specific shocks (see the Table note for a detailed description of how we

[19] In the Supplementary Appendix we show these findings extend to productivity and, although to a lesser extent, exports.

[20] Unlike the case of professional services, the OECD measure of energy regulation is available before 1996, allowing in principle to focus on a longer growth period. Unfortunately, as we go back in time the number of missing observations on the dependent variables rapidly increases, complicating the comparison of estimates. As an example, the Supplementary Appendix shows the results obtained when the specification in column 1, Table 8 is considered, and growth rates are computed starting in various years from 1980 to 1996. We always estimate negative coefficients which become statistically insignificant starting in the mid-1980s, when the number of observations becomes nearly a half with respect to those available in 1996.

Table 9

Sector-specific Effects

	(1)	(2)	(3)	(4)	(5)	(6)
	Energy Services	**Professional Services**	**Community Services**	**Transportation Services**	**All Services**	**Robs. to GLOPP**
Energy Regulation × Energy dependence	–0.482**				–0.540*	–0.530*
[$X_{c,ENERGY} \times w_{j,\ ENERGY}$]	*(0.147)*				*(0.232)*	*(0.232)*
Prof. Serv. Regulation × Prof. Serv. dependence		–0.286*			–0.254*	–0.259*
[$X_{c,PROSERV} \times w_{j,\ PROSERV}$]		*(0.124)*			*(0.118)*	*(0.114)*
Communications Regulation × Comm. dep.			–0.417		0.115	0.206
[$X_{c,TLCPOST} \times w_{j,TLCPOST}$]			*(1.193)*		*(1.147)*	*(1.100)*
Transports Regulation × Transports dependence				–0.231	0.101	0.112
[$X_{c,TRANSP} \times w_{j,\ TRANSP}$]				*(0.160)*	*(0.247)*	*(0.246)*
Energy Regulation × Global opportunities (energy)						0.038
[$X_{c,ENERGY} \times GLOPP_{j,\ ENERGY}$]						*(0.072)*
Prof. Serv. Regulation × Global opp. (prof. serv.)						–0.343**
[$X_{c,PROSERV} \times GLOPP_{j,\ PROSERV}$]						*(0.131)*
Financial dev. × external dep.	0.010*	0.011**	0.010*	0.011*	0.011**	0.011**
[$FD_c \times ED_j$]	*(0.004)*	*(0.004)*	*(0.005)*	*(0.005)*	*(0.004)*	*(0.004)*
Initial industry share	0.171*	0.156*	0.169*	0.167*	0.159*	0.182**
[$SHARE_{j,c}$]	*(0.067)*	*(0.073)*	*(0.069)*	*(0.069)*	*(0.069)*	*(0.062)*
Constant	0.004	0.014	–0.007	0.004	0.021	0.039
	(0.017)	*(0.020)*	*(0.022)*	*(0.020)*	*(0.030)*	*(0.031)*
Observations	220	220	220	220	220	220
R^2	0.69	0.68	0.67	0.67	0.70	0.71

+ significant at 10 per cent; * significant at 5 per cent; ** significant at 1 per cent.

Notes to Table 9:
The dependent variable is the annual compounded growth rate of real value added at the industry-country level for the period 1996-2002 ($GROWTH_{j,c}$).
Variables $X_{c,s} * w_{j,s}$ are interaction terms between country-level measures of regulation in energy, professional services, communications, transports in 1996 ($X_{c,s}$) and the corresponding industry-level indicators of dependence ($w_{j,s}$).
Interaction weights $w_{j,s}$ are ("direct") technical coefficients of dependence between service sector *s* and manufacturing industry *j* computed on the 1997 USA Input-Output matrix.
Financial development is measured as Private Credit by Deposit Money Banks over GDP in 1996 (FD_c) and is interacted with External dependence (ED_j) an industry-level measure of reliance on external finance obtained from US firm-level data.
$GLOPP_{j,\ ENERGY}$ and $GLOPP_{j,\ PROSERV}$ are the estimated industry value added growth in the USA. For each of the service sector *ENERGY* and *PROSERV*, global opportunities ($GLOPP_{j,s}$) are obtained according to the following two-steps procedure: (*a*) regress $GROWTH_{j,c}$ on country dummies, industry dummies and industry dummies interacted with country-level regulation in sector *s*; USA are excluded from the regression; (*b*) obtain $GLOPP_j$ as the predicted values of $GROWTH_{j,c}$ for the USA.
$SHARE_{j,c}$ indicates the industry share in total value added in manufacturing in 1996.
All regression include (unreported) controls for labor market regulation and red tape costs (see Table 1 for the definition of these variables).
All regressions also include country- and industry-fixed effects and use (employment) weighted least squares as estimation method.
Robust standard errors are reported in parentheses.

obtained such measure). The underlying idea is simple: if estimates in column 5 were to reflect short run shocks, they should be dominated by direct measures of the opportunities of expansion faced by different industries. Interestingly, our results indicate that lower regulation of professional services (but not of energy) does help accommodating short run shocks. On the other hand, however, our previous findings are unaffected and still statistically significant.[21]

Data limitations (e.g., the lack of comparable data on prices, the quality or efficiency of each of the four services, etc.) prevent a thorough analysis of the reasons why regulation is more relevant in some services than others. Interestingly, however, our results highlight the relevance of two sectors (energy and professional services) that have recently attracted increasing attention by policymakers in many developed economies.[22] Our findings can therefore be used to infer the potential growth effects of competition policies that are high in the current policy agenda: those addressing barriers to entry in energy and conduct regulation in professional services. Our estimates imply that the complete removal of the two main determinants of conduct regulation, that is *(a)* bans to comparative or price advertising and *(b)* the regulation of price and tariffs, would imply the Textiles-Transport equipment growth differential to rise by 0.3 and 0.5 percentage points, respectively.[23] As to the energy market, our findings imply the industry growth differential associated to *(a)* creating a liberalized wholesale market for electricity, *(b)* allowing third party access to the electricity and gas transmission grid, or *(c)* imposing the separation of ownership between energy production (or import) and its distribution would increase by 0.3, 0.7 and 0.9 percentage points per year, respectively.

---

[21] In the Supplementary Appendix, we show that our previous results on the aggregate effect of regulation (see Tab. 4) are also robust to accounting for a measure of global opportunities.

[22] See the European Commission "Third Legislative Package on Energy Markets" (July 2009), promoting among other things the unbundling of network operation from supply and generation in energy, and the Commission report on "Competition in Professional Services" (February 2004), urging "the reform of unjustified restrictions in the professional services sector". See also the chapter on Structural Policy Priorities in "Going for Growth" (OECD, 2009).

[23] Fore each service sector, the OECD regulation index $X_{c,s}$ is obtained as the weighted average of several sub-indexes measuring the extent of regulation in different areas (see the Data Appendix). The thought exercises reported in the text are obtained considering the change in the $X_{c,s}$ implied by the maximum possible variation of each of the sub-indexes. In the case of regulation of prices and fees in professions, for example, this would correspond to moving from having "minimum prices in all services" (as in the case of legal service in Italy) to "no regulation" (as in the case of accounting service in Canada).

6. Conclusions

Growing concerns that high levels of regulation might not reflect public interest have motivated a number of academic and policy-oriented researches aimed at evaluating the impact of regulatory barriers on the performance of regulated firms. We contribute to this debate highlighting the non-negligible indirect effects of anti-competitive regulation on downstream industries, focusing on the case of service inputs. Our results indicate that service regulation has a significant negative impact on the growth rate of value added, productivity and exports of service dependent industries. Interestingly, the impact of regulation appears to be particularly relevant in the case of those service activities (energy supply and professional services) the recent competition policy debate has been focusing on most intensively, both in Europe and in other developed countries. Also, our findings suggest the payoffs from lower service regulation would be more significant the larger the extent of the domestic market.

Our results leave several interesting questions open to future research. On one hand, the increased availability of detailed firm-level data should allow disentangling whether the aggregate growth effects we estimated here are mainly due to entry and exit of firms, to the performance of existing firms or both. On the other, it would be important to look deeper into the mechanisms underlying our findings, focusing on how regulation affects the industrial organization of services (for example, in terms of number and size of firms, of turnover rates etc), on how this shapes service market outcomes and, eventually, the patterns of international specialization and comparative advantages.

DATA APPENDIX

Country sample:

Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, the United Kingdom and the United States.

Industry sample:

"Food products, beverages and tobacco" (Isic Rev. 3 = "15-16"), "Textiles and textile products" (Isic Rev. 3 = "17-18"), "Leather, leather products and footwear" (Isic Rev. 3 = "19") , "Wood and products of wood and cork" (Isic Rev. 3 = "20") , "Pulp, paper, paper products, printing and publishing" (Isic Rev. 3 = "21-22"), "Coke, refined petroleum products and nuclear fuel" (Isic Rev. 3 = "23"), "Chemicals and chemical products" (Isic Rev. 3 = "24"), "Rubber and plastics products" (Isic Rev. 3 = "25"), "Other non-metallic mineral products" (Isic Rev. 3 = "26"), "Basic metals" (Isic Rev. 3 = "27"), "Fabricated metal products, except machinery and equipment" (Isic Rev. 3 = "28"), "Machinery and equipment, N.E.C." (Isic Rev. 3 = "29"), "Electrical and optical equipment" (Isic Rev. 3 = "30-33"), "Transport equipment" (Isic Rev. 3 = "34-35"), "Manufacturing N.E.C., recycling" (Isic Rev. 3 = "36-37").

Dependence of manufacturing industries on service inputs

Throughout most of the paper we use weights $w_{j,s}$ computed as the technical coefficients derived from the 1997 US Input-Output matrix. They are given by the elements of the matrix $T = M\,diag\,(y)^{-1}$, where M is the industry-by-industry (44 × 44) input-output matrix, y is the (44 × 1) vector of industry output. In Table 4, column 2, weights are instead computed as the product of the elements of the inverse Leontief matrix by a vector of the industry value added-to-output ratios. More specifically, let v be the (44 × 1) vector of industry value added. The inverse Leontief matrix is $F = (I - T)^{-1}$ and satisfies $\iota' = q'F$, where $q = diag\,(y)^{-1}v$ is the vector of industry value added-to-output ratios. According to the last relation the value of production in each sector (normalized to one) is decomposed in the contribution of value added produced in all the sectors (q) weighted with the (direct and indirect) measure of intersectoral dependence (F). For each industry, the relation can be written as $1_j = \sum_{k=1}^{44} q_k f_{k,j}$ with $k = 1, ..., 44$. The indirect weights used in Table 4, column 2 are given by the elements $q_k f_{k,j}$.

Data on regulation in selected non-manufacturing sectors

All the regulatory indicators range on a common (0-6) scale from least to most restrictive conditions for competition. Data are available for seven non-manufacturing sectors: electricity and gas supply, road freight, air passenger transport, rail transport, post and telecommunications and professional services (accounting, architects, engineers and legal services). For each sector, a set of sub-indexes is available covering different forms of regulation: barriers to entry, vertical integration, market structure, price regulation, conduct regulation and public ownership. See Table 1 and the main text for a description of the sub-indexes we focused on in the analysis. See Nicoletti *et al*. (1999) and Conway and Nicoletti (2006) for a complete description of the OECD-PMR database.

SUPPLEMENTARY APPENDIX

Table 10

Alternative Determinants of International Specialization and Comparative Advantages

	(1)	(2)	(3)	(4)	(5)	(6)
	Value Added Growth				**Exports**	
	Human Capital	**Physical Capital**	**Both**	**Property Rights**	**Contract Enforcement (a)**	**(b)**
Service regulation [$SERVREG_{j,c}$]	–0.154* *(0.066)*	–0.174* *(0.068)*	–0.154* *(0.067)*	–0.176* *(0.068)*	–6.786* *(3.011)*	–3.688+ *(2.032)*
Financial dev. × external dep. [$FD_c \times ED_j$]	0.007+ *(0.004)*	0.010* *(0.004)*	0.007+ *(0.004)*	0.010* *(0.004)*	0.405** *(0.144)*	0.298+ *(0.158)*
Human capital × skill intensity	0.101* *(0.048)*		0.101* *(0.048)*			
Physical capital × physical capital intensity		–0.468 *(2.885)*	0.082 *(2.790)*			
Property rights × intangible intensity				–0.001 *(0.003)*		
Quality of contract enforcement × contract intensity					0.144** *(0.048)*	
Quality of contract enforcement × institutional dependence						0.003* *(0.001)*
Initial industry share [$SHARE_{j,c}$]	0.141* *(0.064)*	0.169* *(0.067)*	0.141* *(0.064)*	0.171* *(0.067)*		
Constant	–0.789* *(0.382)*	0.023 *(0.106)*	–0.793* *(0.366)*	0.009 *(0.022)*	2.595 *(1.736)*	6.634** *(2.207)*
Observations	220	220	220	220	220	220
R^2	0.69	0.67	0.69	0.67	0.79	0.37

+ significant at 10%; * significant at 5%; ** significant at 1%.

Notes:

In cols. 1-4 the dependent variable is the annual compounded growth rate of real value added at the country-industry level for the period 1996-2002 (GROWTHj,c); in col. 5 the dependent variable is the natural logarithm of total exports in industry j from country c in 1996; in col. 6 the dependent variable is an index of export specialization given by $\left(EXPORTS_{j,c} \big/ \sum_c EXPORTS_{j,c}\right) \big/ \left(\sum_j EXPORTS_{j,c} \big/ \sum_{j,c} EXPORTS_{j,c}\right)$, where j and c represent industries and countries, respectively. SERVREGj,c measures exposure to service regulation at the country-industry level as a weighted average (Σs wj,s*Xc,s) of country-level anti-competitive regulation indexes from the OECD-PMR databases. Service regulation (Xc,s) is measured in 1996. Interaction weights wj,s are ("direct") technical coefficients of dependence between service sector s and manufacturing industry j computed on the 1997 USA Input-Output matrix. Financial development is measured as Private Credit by Deposit Money Banks over GDP in 1996 (FDc) and is interacted with External dependence (EDj) an industry-level measure of reliance on external finance obtained from USA firm-level data (see Table 1). In cols. 1 and 3, Human capital is an index of labor force quality on a (0-100) scale taken from Bosworth and Collins (2003). It is interacted with average years of schooling at the industry level in 1980 (as obtained from the US 1990 Integrated PUMS). In cols. 2 and 3, Physical capital is the physical capital-to-GDP ratio in 1980. The capital stock is calculated using the perpetual inventory method as implemented by Klenow and Rodriguez-Clare (2005). Source: Penn World Table 5.6. It is interacted with US capital-value added ratio at industry level in 1995 taken from the EUKLEMS database (http://www.euklems.net/). In col. 4, "Property rights" is an index of the protection of the private property across countries. It is interacted with an industry-level measure of intangible intensity in US industries. Both are taken from Claessens and Laeven (2003). In cols. 5 and 6, "Quality of contract enforcement" measures the extent to which agents have confidence in and abide by the rules of society (Kaufmann, Kraay and Mastruzzi, 2003). In col. 5, contract enforcement is interacted with Nunn (2007) measure of contract intensity (i.e., of the importance of relationship-specific investments). In col. 6 it is interacted with a measure of institutional dependence. Following Levchenko (2007), this is computed as the (opposite of) an Herfindahl index of intermediate input use from the U.S. Input-Output Use Table for 1997. SHAREj,c indicates the industry share in total value added in manufacturing in 1996. All regressions include country- and industry-fixed effects and use (employment) weighted least squares as estimation method. Robust standard errors are reported in parentheses.

Table 11

Alternative Measures of Regulation Impact

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Value Added Growth			**Productivity Growth**			**Export Growth**		
	ALTERN.	**SERVREG**	**BOTH**	**ALTERN.**	**SERVREG**	**BOTH**	**ALTERN.**	**SERVREG**	**BOTH**
Panel A : OECD Regulation Impact Indicator (*RII*)									
Reg. Imp. Ind. [$rii_{j,c}$]	–0.246* *(0.120)*		–0.158 *(0.111)*	–0.164 *(0.144)*		–0.043 *(0.120)*	0.199 *(0.150)*		0.352* *(0.162)*
Service reg. [$servreg_{j,c}$]		–0.176** *(0.068)*	–0.145* *(0.064)*		–0.202* *(0.080)*	–0.193* *(0.075)*		–0.215* *(0.106)*	–0.279** *(0.102)*
Implied effects	–0.009	–0.019	–	–0.006	–0.022	–	0.007	–0.023	–
Observations	220	220	220	220	220	220	205	205	205
R^2	0.67	0.67	0.67	0.57	0.59	0.59	0.71	0.72	0.76
Panel B : "Mixed" indicator of Service Regulation									
Serv. Reg. Mixed [$mixed_{j,c}$]	–0.076* *(0.036)*		–0.052 *(0.037)*	–0.086* *(0.038)*		–0.058+ *(0.035)*	–0.016 *(0.039)*		0.028 *(0.043)*
Service reg. [$servreg_{j,c}$]		–0.176** *(0.068)*	–0.123+ *(0.070)*		–0.202* *(0.080)*	–0.140* *(0.071)*		–0.215* *(0.106)*	–0.242* *(0.116)*
Implied effects	–0.011	–0.019	–	–0.012	–0.022	–	–0.002	–0.023	–
Observations	220	220	220	220	220	220	205	205	205
R^2	0.67	0.67	0.68	0.59	0.59	0.60	0.71	0.72	0.72

+ significant at 10%; * significant at 5%; ** significant at 1%.

Notes:

In cols. 1-3 the dependent variable is the annual compounded growth rate of real value added at the country-industry level for the period 1996-2002 ($GROWTH_{j,c}$). In cols. 4-6 the dependent variable is the annual compounded growth rate of labor productivity (value added per employed worker) at the industry-country level for the period 1996-2002 ($LPGROWTH_{j,c}$). In cols. 7-9 the dependent variable is the annual compounded growth rate of exports at the industry-country level for the period 1996-2002 ($EXPGROWTH_{j,c}$). $SERVREG_{j,c}$ measures exposure to service regulation at the country-industry level as a weighted average ($\Sigma_s\, w_{j,s}{}^{*}X_{c,s}$) of country-level anti-competitive regulation indexes from the OECD-PMR databases. Service regulation ($X_{c,s}$) is measured in 1996. Interaction weights $w_{j,s}$ are ("direct") technical coefficients of dependence between service sector s and manufacturing industry j computed on the 1997 USA Input-Output matrix. In Panel A the Regulation Impact Indicator (RII) is the OECD measure of the relevance of service regulation for manufacturing industries (taken from Conway and Nicoletti, 2006). In Panel B, the "Mixed" indicator of Service regulation is computed as a weighted average ($\Sigma_s\, w^c_{j,s}{}^{*}X_{c,s}$). Country-specific weights $w^c_{j,s}$ are ("direct") technical coefficients of dependence between service sector s and manufacturing industry j computed on the OECD Input-Output matrices. All regressions include (unreported) controls for financial development and for initial conditions: $SHARE_{j,c}$ in cols. 1-3, $LLP_{j,c}$ in cols. 4-6 and $EXSHARE_{j,c}$ in cols. 7-9 (see Table 1 for the definition of these variables). All regressions also include country- and industry-fixed effects and use (employment) weighted least squares as estimation method. Robust standard errors are reported in parentheses.

Table 12

Global Opportunities and Average Regulation

	(1)	(2)
	Without Other Controls	**With Other Controls**
Service regulation [$SERVREG_{j,c}$]	–0.185** *(0.068)*	–0.234** *(0.070)*
Average service regulation × global opportunities	–0.101 *(0.099)*	–0.035 *(0.123)*
Fin. dev. × external dep. [$FD_c \times ED_j$]	0.009* *(0.004)*	0.011** *(0.004)*
Initial industry share [$SHARE_{j,c}$]	0.160* *(0.067)*	0.148* *(0.068)*
Constant	0.016 *(0.020)*	0.042+ *(0.025)*
Observations	220	220
R^2	0.68	0.69

+ significant at 10%; * significant at 5%; ** significant at 1%.

Notes:

The dependent variable is the annual compounded growth rate of real value added at the industry-country level for the period 1996-2002 ($GROWTH_{j,c}$). $SERVREG_{j,c}$ measures exposure to service regulation at the country-industry level as a weighted average ($\Sigma_s w_{j,s} {}^* X_{c,s}$) of country-level anti-competitive regulation indexes from the OECD-PMR databases. Service regulation ($X_{c,s}$) is measured in 1996. Interaction weights $w_{j,s}$ are ("direct") technical coefficients of dependence between service sector s and manufacturing industry j computed on the 1997 USA Input-Output matrix. Average service regulation is the simple average of sectoral regulation given by ($X_{c,ENERGY} + X_{c,PROSERV} + X_{c,TLCPOST} + X_{c,TRANSP}$) / 4. It is interacted with an industry-level measure of global opportunities obtained according to the following two-steps procedure: (*a*) Regress $GROWTH_{j,c}$ on country dummies, industry dummies and industry dummies interacted with country-level simple average of sectoral regulation; USA are excluded from the regression. (*b*) Obtain global opportunities as the predicted values of $GROWTH_{j,c}$ for the USA. All regression include (unreported) controls for financial development, labor market regulation and red tape costs (see Tables 1, 2 and 7 for the definition of these variables). $SHARE_{j,c}$ is the industry share in total value added in manufacturing in 1996. All regressions include country- and industry-fixed effects and use (employment) weighted least squares as estimation method. Robust standard errors are reported in parentheses.

Table 13

Service Regulation and Country Size

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Value Added Growth			**Productivity Growth**			**Export Growth**		
	All Countries	**Large Countries**	**Small Countries**	**All Countries**	**Large Countries**	**Small Countries**	**All Countries**	**Large Countries**	**Small Countries**
Service reg.	–0.272**	–0.313**	0.107	–0.282**	–0.340**	0.012	–0.241*	–0.270+	–0.233
[$SERVREG_{j,c}$]	*(0.073)*	*(0.086)*	*(0.174)*	*(0.101)*	*(0.124)*	*(0.127)*	*(0.104)*	*(0.145)*	*(0.180)*
Constant	0.020	0.036	0.033	–0.070	–0.083	0.017	0.028	0.006	–0.006
	(0.025)	*(0.038)*	*(0.030)*	*(0.048)*	*(0.065)*	*(0.054)*	*(0.030)*	*(0.047)*	*(0.043)*
Observations	220	113	107	220	114	106	205	98	107
R^2	0.70	0.75	0.59	0.61	0.65	0.52	0.75	0.80	0.75

+ significant at 10%; * significant at 5%; ** significant at 1%.

Notes:

In cols. 1-3 the dependent variable is the annual compounded growth rate of real value added at the country-industry level for the period 1996-2002 ($GROWTH_{j,c}$). These columns replicate results of table 8, cols. 7-9 in the main text. In cols. 4-6 the dependent variable is the annual compounded growth rate of labor productivity (value added per employed worker) at the industry-country level for the period 1996-2002 ($LPGROWTH_{j,c}$). In cols. 7-9 the dependent variable is the annual compounded growth rate of exports at the industry-country level for the period 1996-2002 ($EXPGROWTH_{j,c}$). $SERVREG_{j,c}$ measures exposure to service regulation at the country-industry level as a weighted average ($\Sigma_s\, w_{j,s} {}^* X_{c,s}$) of country-level anti-competitive regulation indexes from the OECD-PMR databases. Service regulation ($X_{c,s}$) is measured in 1996. Interaction weights $w_{j,s}$ are ("direct") technical coefficients of dependence between service sector s and manufacturing industry j computed on the 1997 USA Input-Output matrix. All regressions include (unreported) controls for financial development [$FD_c \times ED_j$], Labour market regulation [$LMR_c \times LABINT_j$], Red tape costs [$COST_c \times GROP_j$], FDI restrictions [$FDIREG_{j,c}$], Public ownership [$POWN_{j,c}$] and the corresponding initial conditions [$SHARE_{j,c}$, $LLP_{j,c}$ and $EXSHARE_{j,c}$]. See Table 1 for the definition of these variables. The sample of large countries include Canada, France, Germany, Italy, Japan, the Netherlands, Spain and the UK while the sample of small ones include Austria, Belgium, Denmark, Finland, Greece, Norway, Portugal and Sweden. All regressions also include country- and industry-fixed effects and use (employment) weighted least squares as estimation method. Robust standard errors are reported in parentheses.

Table 14

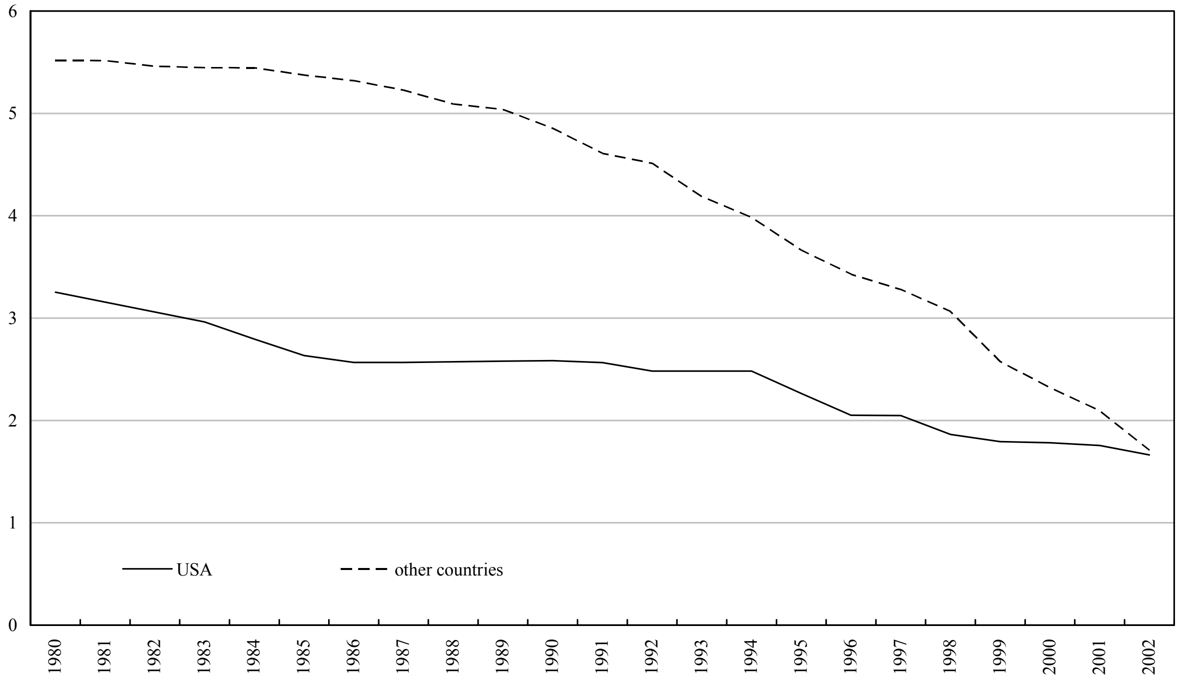
Sector-Specific Effects Over Longer Horizons: Energy

	(1)	(2)	(3)	(4)	(5)
	Initial Year: 1980	**Initial Year: 1984**	**Initial Year: 1988**	**Initial Year: 1992**	**Initial Year: 1996**
Energy Regulation × Energy dependence	–0.206	–0.210	–0.434*	–0.469**	–0.482**
[$X_{c,ENERGY} \times w_{j,\ ENERGY}$]	*(0.207)*	*(0.175)*	*(0.182)*	*(0.178)*	*(0.147)*
Observations	139	139	154	220	220
R^2	0.74	0.74	0.75	0.66	0.69

+ significant at 10%; * significant at 5%; ** significant at 1%.
Notes:
The dependent variable is the annual compounded growth rate of real value added at the industry-country level for the period 1996-2002 (GROWTHj,c). Xc,ENERGY * wj,ENERGY is an interaction term between country-level measures of regulation in energy in 1996 (Xc,ENERGY) and the corresponding industry-level indicators of dependence (wj,ENERGY). The interaction weight wj,ENERGY is the ("direct") technical coefficients of dependence between energy and manufacturing industry j computed on the 1997 USA Input-Output matrix. All regression include (unreported) controls for financial development, labor market regulation and red tape costs (see Tables 1, 2 and 7 for the definition of these variables), and the industry share in total value added in manufacturing in 1996. All regressions also include country- and industry-fixed effects and use (employment) weighted least squares as estimation method. Robust standard errors are reported in parentheses.

Figure 1

Service Regulation in USA and Other OECD Countries



Notes:
Service regulation is the simple average of the OECD measures of regulation (Xc,s) in energy, communications and transports. Other countries are: Austria, Belgium, Canada, Germany, Denmark, Finland, France, Great Britain, Greece, Italy, Japan, the Netherlands, Norway, Portugal, Spain and Sweden.

REFERENCES

Alesina, A., S. Ardagna, G. Nicoletti and F. Schiantarelli (2005), "Regulation and Investment", *Journal of the European Economic Association*, Vol. 3, No. 4, pp. 791-825.

Allegra, E., M. Forni, M. Grillo and L. Magnani (2004). "Antitrust Policy and National Growth: Some Evidence from Italy", *Giornale degli Economisti e Annali di Economia*, Vol. 63, No. 1, pp. 69-86.

Armington, P.S. (1969). "A Theory of Demand for Products Distinguished by Place of Production", IMF, Staff Paper, Vol. 16, pp. 159-78.

Arnold, J., S. Javorcik and A. Matoo (2007). "Does Services Liberalization Benefit Manufacturing Firms? Evidence from Czech Republic", The World Bank, Policy Research, Working Paper, No. 4109.

Arnold, J., G. Nicoletti and S. Scarpetta (2008), "Product Market Policies, Allocative Efficiency and Productivity: A Cross-country Analysis", OECD, Economics Department, Working Paper, No. 616.

Bassanini, A., L. Nunziata and D.Venn (2009), "Job Protection Legislation and Productivity Growth in OECD Countries", *Economic Policy*, Vol. 24, No. 58, pp. 349-402.

Bertrand, M. and F. Kramartz (2002), "Does Entry Regulation Hinder Job Creation? Evidence From The French Retail Industry", *Quarterly Journal of Economics*, Vol. 117, No. 4, pp. 1369-413.

Bosworth, B.P. and S.M. Collins (2003), "The Empirics of Growth: An Update", Brookings Papers on Economic Activity, Vol. 34 (2003-2), pp. 113-206.

Caballero, R., K. Cowan, E. Engel and A. Micco (2006), "Effective Labour Regulation and Microeconomic Flexibility", Cowles Foundation, Discussion Paper, No. 1480.

Carlin, W. and C. Mayer (2003), "Finance, Investment, and Growth", *Journal of Financial Economics*, Vol. 69, No. 1, pp. 191-226.

Ciccone, A. and K. Matsuyama (1996), "Start-up Costs and Pecuniary Externalities as Barriers to Economic Development", *Journal of Development Economics*, Vol. 49, No. 1, pp. 33-59.

Ciccone, A. and E. Papaioannou (2006), "Adjustment to Target Capital, Finance and Growth", CEPR, Discussion Paper, No. 5969.

―――――― (2007), "Red Tape and Delayed Entry", *Journal of the European Economic Association*, Vol. 5, No. 2-3, pp. 444-58.

Claessens, S. and L. Laeven (2003), "Financial Development, Property Rights, and Growth", *Journal of Finance*, Vol. 58, No. 6, pp. 2401-36.

Conway, P., D. de Rosa, G. Nicoletti and F. Steiner (2006), "Product Market Regulation and Productivity Convergence", OECD, Economic Studies, No. 43.

Conway, P. and G. Nicoletti (2006), "Product Market Regulation in the Non-manufacturing Sectors of OECD Countries: Measurement and Highlights", OECD, Economics Department, Working Paper, No. 530.

Cuñat, A. and M.J. Melitz (2007), "Volatility, Labor Market Flexibility, and the Pattern of Comparative Advantage", NBER, Working Paper, No. 13062.

de Serres, A., S. Kobayakawa, T. Sløk and L. Vartia (2006), "Regulation of Financial Systems and Economic Growth", OECD, Economics Department, Working Paper, No. 506.

Djankov, S., R. La Porta, F. Lopes-de-Silanes and A. Shleifer (2002), "The Regulation of Entry", *Quarterly Journal of Economics*, Vol. 117, No. 1, pp. 1-37.

Faini, R., J. Haskel, G. Barba Navaretti, C. Scarpa and C. Wey (2006), "Contrasting Europe's decline: Do Product Market Reforms Help?", in T. Boeri, M. Castanheira, R. Faini and V. Galasso (eds.), *Structural Reforms Without Prejudices*, Oxford University Press.

Favara, G. (2003), "An Empirical Reassessment of the Relationship Between Finance and Growth", IMF, Working Paper, No. 03/123.

Fisman, R. and V. Sarria-Allende (2004), "Regulation of Entry and the Distortion of Industrial Organization", NBER, Working Paper, No. 10929.

Fonseca, R. and N. Utrero (2005), "Financial Development, Labor and Market Regulations and Growth", Document de Recherche EPEE, No. 05-05.

Helpman, E. and P. Krugman (1985), *Market Structure and Foreign Trade*, Cambridge, The MIT Press.

Hoekman B., H. Kee and M. Olarreaga (2004), "Tariffs, Entry Regulation and Markups: Country Size Matters", *The BE Journals in Macroeconomics: Contributions to Macroeconomics*, Vol. 4, No. 1, Article 8.

Kaufmann, D., A. Kraay and M. Mastruzzi (2003), "Governance Matters III: Governance Indicators for 1996-2002", World Bank, Working Paper, No. 3106.

Klapper, L., L. Laeven and R.G. Rajan (2006), "Entry Regulation as a Barrier to Entrepreneurship", *Journal of Financial Economics*, Vol. 82, pp. 591-629.

Klenow, P.J. and A. Rodriguez-Clare (2005), "Externalities and Growth", in P. Aghion. and F. Durlauf (eds.), *The Handbook of Economic Growth*, Amsterdam: North-Holland.

Koyama, T. and S.S. Golub (2006), "OECD's FDI Regulatory Restrictiveness Index: Revision and Extension to More Economies", OECD, Working Papers on International Investment, No. 2006/4.

Levchenko, A.A. (2007), "Institutional Quality and International Trade", *Review of Economic Studies*, Vol. 74, No. 3, pp. 791-819.

Martin, R., M. Roma and I. Vansteenkiste (2005), "Regulatory Reforms in Selected EU Network Industries", ECB, Occasional Paper, No. 28.

Manning, M.J. (2003), "Finance Causes Growth: Can We Be So Sure?", *The BE Journals in Macroeconomics: Contributions to Macroeconomics*, Vol. 3, No. 1, Article 12.

Midelfart, K.H., H.G. Overman and A.J. Venables (2003), "Monetary Union and the Economic Geography of Europe", *Journal of Common Market Studies*, Vol. 41, No. 5, pp. 847-68.

Nicoletti, G. and S. Scarpetta (2003), "Regulation, Productivity and Growth: OECD Evidence", *Economic Policy*, Vol. 18, No. 36, pp. 9-72.

Nicoletti, G., S. Scarpetta and O. Boylaud (1999), "Summary Indicators of Product Market Regulation with an Extension to Employment Protection Legislation", OECD, Economics Department, Working Paper, No. 226.

Nunn, N. (2007), "Relationship-specificity, Incomplete Contracts, and the Pattern of Trade", *Quarterly Journal of Economics*, Vol. 122, No. 2, pp. 569-600.

OECD (2009), *Going for Growth*, Paris, OECD.

Okuno-Fujiwara, M. (1988), "Interdependence of Industries, Coordination Failure and Strategic Promotion of an Industry", *Journal of International Economics*, Vol. 25, Nos. 1-2, pp. 25-43.

Papke, L.E. and J.M. Wooldridge (1996), "Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates", *Journal of Applied Econometrics*, Vol. 11, No. 6, pp. 619-32.

Rajan, G.R. and L. Zingales (1998), "Financial Dependence and Growth", *American Economic Review*, Vol. 88, No. 3, pp. 559-86.

Rodrick, D. (1996), "Coordination Failures and Government Policy: A Model with Applications to East Asia and Eastern Europe", *Journal of International Economics*, Vol. 40, Nos. 1-2, pp. 1-22.

Rodriguez-Clare, A. (1996), "The Division of Labor and Economic Development", *Journal of Development Economics*, Vol. 49, No. 1, pp. 3-32.

Rousseau, P.L. and P. Wachtel (1998), "Financial Intermediation and Economic Performance: Historical Evidence from Five Industrialized Countries", *Journal of Money, Credit and Banking*, Vol. 30, No. 4, pp. 657-78.

THE MARGINAL COST OF PUBLIC FUNDS IN THE EU: THE CASE OF LABOUR VERSUS GREEN TAXES

Salvador Barrios,[*] *Jonathan Pycroft*[*] *and Bert Saveyn*[*]

One key objective of tax-based fiscal consolidations which is too often disregarded in public debate is to minimise economic distortions. This paper uses a computable general equilibrium model to gauge these potential distortions by calculating the marginal cost of public funds (MCF) for EU member states. We consider two specific tax categories which prove especially relevant in such a context: labour and green taxes. First the economic distortion provoked by labour taxes is significantly larger than for green taxes. This result suggests that a green-taxes oriented fiscal consolidation would be preferred to a labour-tax oriented one (assuming that both tax increases would yield the same tax revenues). This holds for all EU member states modelled and despite the fact that potential welfare enhancement through pollution abatement are cancelled-out. Nevertheless, this result is slightly less strong when one considers the spillover effects between countries, which are more pronounced (in relative terms) for energy taxes. This suggests that the use of energy taxes for fiscal consolidation would be more effective were there to be close coordination across EU countries. In addition the efficiency losses associated with labour taxes are also likely to be greater when labour markets are less flexible (from an efficiency-wage perspective), a result also found to a small extent for green taxes. This raises the possibility that undertaking structural reforms (especially in the labour market) would help to minimize the efficiency losses entailed by tax-driven fiscal consolidations.

Introduction

The need to restore sound fiscal balance represents a key objective of EU economic policy making in the aftermath of the financial crisis. Whenever tax increases are contemplated, the challenge for policy makers is to strike a balance between short-term recovery and long-term growth, the latter requiring supply and economic efficiency-enhancing policy measures. The need to lower the efficiency loss of tax increases is also aimed at optimising the level of extra-tax revenues obtained from it given that inappropriate tax hikes could lead to lower than expected tax revenue and would eventually require successive tax increases in order to meet fiscal policy objectives. To date, much of the policy debate has been informed by (neo) Keynesian types of models assessing the size of fiscal multipliers and potential effects of fiscal consolidation in a context of zero-bound monetary policy and impaired financial sector, see in particular Corsetti *et al.* (2010), IMF (2012) and Coenen *et al.* (2012) for recent, model-based discussions. Some additional guidance on these important issues, albeit too often disregarded in the policy debate, could be drawn from the optimal tax policy literature analysis of the potential distortionary effect of tax increases, see in particular Feldstein (1997). Accordingly, the objective for policy makers should be to minimise the distortionary effect of taxation and related adverse effects on the economic recovery since existing evidence suggests that the least distortionary a tax system is, the less detrimental its impact on growth, see in particular Arnold *et al.* (2011). The efficiency loss associated with tax increases crucially depends on the behavioural responses of economic agents

---

[*] European Commission, Directorate General Joint Research Centre, Institute for Prospective Technological studies.

The authors are thankful for comments made by Yngve Lindh, Gilles Mourre, Gaetan Nicodème, Françoise Nemry and suggestions received by participants at seminars organised by the Banca d'Italia, DG ECFIN-European Commission, Universidad Pablo de Olavide Seville, at the Joint Research Centre IPTS ECCET seminar and ECOMOD Conference 2012.

The views expressed in this paper should not necessarily be seen as those of the European Commission.

which affect the tax bases and the supply side of the economy. An appropriate metric to gauge the losses related to (and potential growth-detrimental effect of) tax increases should compare the relationship between the deadweight loss and the extra-revenue associated with a given tax increase.

In this paper we calculate more specifically the marginal cost of public funds (MCF) which proves especially useful for this purpose. This indicator is widely used in the public economics literature for the evaluation of tax reforms and public spending program requiring the transfer of resources from the private to the public sector, see in particular Dahlby (2008). Based on this measure, existing evidence suggests that the efficiency loss of tax increases vary widely across tax categories and countries and increases with the level of taxation burden in the economy, see in particular Devarajan and Robinson (2002) and Dahlby and Ferede (2011). The MCF metric is used here to gauge the cost of tax increases in the EU. To do so we make use of the computable general equilibrium model GEM-E3. One important feature of this model version is that it is calibrated using social accounting matrices derived from national account data of EuroStat. The resulting tax rates used in the simulations therefore reflect actual effective tax rates. Our analysis is carried out for all of the 24 EU member states that are specified in the model (all except for Croatia, Cyprus, Malta and Luxembourg).

We consider two specific tax categories: labour and energy taxes. Our choice of tax categories is motivated by a number of questions of special relevance in the EU context. First, we chose labour taxation because of its relatively high level in most EU countries and because it is well known to have wide-ranging effects spilling well beyond fiscal outcomes. More than any other tax category, labour taxation are directly embedded into country-specific economic and social institutions thus reflecting underlying economic structures, see Blundell *et al.* (1999). Second, green taxation links this analysis with the "double dividend" literature as it is often advocated for as potential instrument for shifting the tax systems in the current EU context in order to make taxation both more employment- and environment-friendly, see Saveyn *et al.* (2011). Because green taxes enter the indirect tax category and is in most EU countries relatively low, resorting to it is also likely to have lower detrimental effects on economic efficiency although it may have non-negligible effects onto the low-income categories of the population.[1] Green taxation may also have direct effect on energy efficiency and thus help minimize the corresponding efficiency losses to be expected from an increase in tax rates. Third we also chose these two tax categories because they could prove instrumental to implement EU-wide coordinated tax reforms despite the fact that they are generally not invoked as candidates for coordination across EU counties according to the optimal tax theory literature. In particular the so-called destination/residence principles, whereby the coordination of direct tax measure should concern primarily (cross-country) mobile production factors while indirect taxation should be collected at the country of destination (see Andersen and Sorensen, 2012, for a review). In practice in the EU however, the high degree of openness and economic integration, the high starting level of public expenditure and tax burden suggest that individual country tax policies might have non-negligible impact on EU partners, potentially influencing the outcome of fiscal consolidation strategies.

Our results show that the efficiency losses related to tax increases (as measured by the MCF) are significantly larger for the labour tax than for green taxes, the latter being represented by households′ consumption taxes on energy products. However the degree of cross-EU countries spillovers is also higher for green taxation calling for coordinated tax strategies despite the low starting level of this type of taxation. Furthermore, we show that these economic costs are also likely to be reduced with a higher degree of flexibility of the labour market, especially so in the case of labour taxes but also, although to a lower extent, for energy taxes. More generally, our

[1] See Speck (1999) for a discussion. In this paper we do not deal with inequality issues.

results tend to suggest that high burden tax categories such as labour tend to be more distortionary than low-burden tax categories lending support to the Laffer type hypothesis. As a result, EU countries might find it appropriate to shift taxation system away from high burden/highly distortionary tax categories in order to favour the growth recovery without which consolidation strategies might prove difficult to sustain in the long-run. Our results prove robust to a number of robustness checks using alternative hypotheses regarding the nature of the extra-tax revenue recycling derived from a given tax hike, the degree of cross-country interdependence in import vs. domestic production substitution and the size of labour supply elasticities.

The rest of the paper is organised as follows. In Section 1 we briefly review the existing literature on the marginal cost of public funds and present our modelling strategy. Our main results are presented in Section 2, while Section 3 provides robustness tests to check the sensitivity of our results to the main hypotheses of the model. Section 4 concludes.

1 Measuring the marginal cost of tax increases

1.1 Literature review

The existing literature provides a wide range of estimated MCF values, differentiated according to the methodology used, the tax categories and the country or region considered. A direct comparison of results across studies is rather complicated since definitions, the underlying theoretical framework and measurements are usually very different from one study to the other. Nevertheless, in order to give an impression of the magnitudes of previous MCF estimates we provide a succinct overview of possible estimates obtained using alternative methodologies.

The MCF metric is relatively straightforward: it simply indicates how many euros (or dollar) are lost in the economy to collect one extra euro (or dollar) tax revenues. As a result MCF usually value greater than one, e.g. $MCF=1+\alpha$, with α measuring the efficiency loss. On the methodological side, there are various ways of measuring the MCF. In this discussion we focus on the three main approaches to estimate the MCF econometric estimations, CGE modelling or through microsimulation.[2] Each of these methodological approaches has pros and cons. The main advantage of CGE models is to consider all potential interactions in the economy (including interactions between industrial sectors, consumers, government and the rest of the world) that determine the final welfare and tax revenue impacts of a given tax change. The drawback of this approach is that it relies on assumptions regarding the functional forms and/or elasticities of the different tax bases to the tax rate changes, however, although one must note that this limitation is not specific to the analysis of tax policy changes, however. The estimates provided by Ballard *et al.* (1985) suggested that the MCF for all taxes ranged between 1.17 and 1.56 depending on the saving and labour supply elasticity used.[3] Hansson and Stuart (1985) found a MCF between 0.67 and 4.51 for the Swedish economy although suggested that varying assumptions regarding labour supply elasticity could have substantial implication in these estimates. In a more recent paper Dixon *et al.* (2012) estimate the MCF for recent tax increases measures taken by the Finnish government in the aftermath of the global financial crisis and estimate this cost to rise up to 1.5 in the long-run. In a recent paper Auriol and Warlters (2012) compute the MCF for African countries using a CGE models with taxes on five tax bases: domestic output, exports, imports, capital and labour in the

---

2 Another strand of models concern partial equilibrium/stylised models which are also best suited to tackle specific issues in analysing the marginal cost of public funds, see Devarajan and Robinson (2002) for a review.

3 Although formally Ballard *et al.* (1985) focused on the Marginal Excess Burden, the MCF can be proxied from these calculations by simply adding 1 to the estimated MEB, see Devarajan and Robinson (2002). One should note however that with such simplification it is assumed that the income elasticity for the taxed product is zero, see Dahlby (2008, chapter 2).

formal sector. These authors show that taxes on domestic output generally have the lowest MCF (around 1.1) and taxes on capital in the formal sector had the highest MCFs (around 1.60).

Econometric estimations allow considering a wide range of countries and/or tax categories as the only limitation is on the data side. An important restriction however comes from the availability of reliable data on the effective tax bases to calculate their potential variation following a tax rate hike. A wide range of studies exist where estimates of the MCF can be derived from the tax base elasticities to tax rate changes thereby capturing the behavioural response of the tax base. For instance in a recent paper Dahlby and Ferede (2012) calculate the MCF for Canadian provinces using information derived from official data used for the tax base equalisation system in place in this country. Their estimates of the MCF of Canadian provinces concerned three tax categories: the corporate income tax, the personal income tax and the sales tax. These authors find a wide range of estimates for the MCF across provinces and potentially important interactions across tax categories ranging from a maximum of 30.6 in the case of corporate taxes to the a minimum of 1 for sales taxes. Dahlby and Ferede also find that the MCF is greatly reduced at the federal level and by considering the impact of the vertical equalisation grants between the federation and the provinces, a result in line with previous findings by Smart (2007).

Microsimulation models in turn have also been used to quantify the marginal cost of public funds to tackle the potential effects of tax reforms by strand of the population, allowing thereby a finer analysis of behavioural effect of tax changes. In particular Kleven and Kreiner (2006) showed that the estimated effects of tax hikes differed sensibly once the labour participation effects is isolated from the number of hours worked (where the extensive and intensive margin of labour supply are distinguished). This approach aims to reflect the fact that labour participation can display very large elasticities while hours-of-work elasticities can be close to zero. Kleven and Kreiner found indeed that once the participation effect was considered into the analysis (and thus once the heterogeneity in labour supply response across different categories of workers was allowed for), then the estimated marginal cost of public funds tended to rise sharply. Applying their analysis for five EU countries namely Denmark, France, Germany, Italy and the UK, Kleven and Kreiner (2006) found that the MCF in certain cases can be more than three times higher due to higher initial distortions of the tax system and higher sensitivity of the MCF to the inclusion of the extensive margin effect of labour participation.

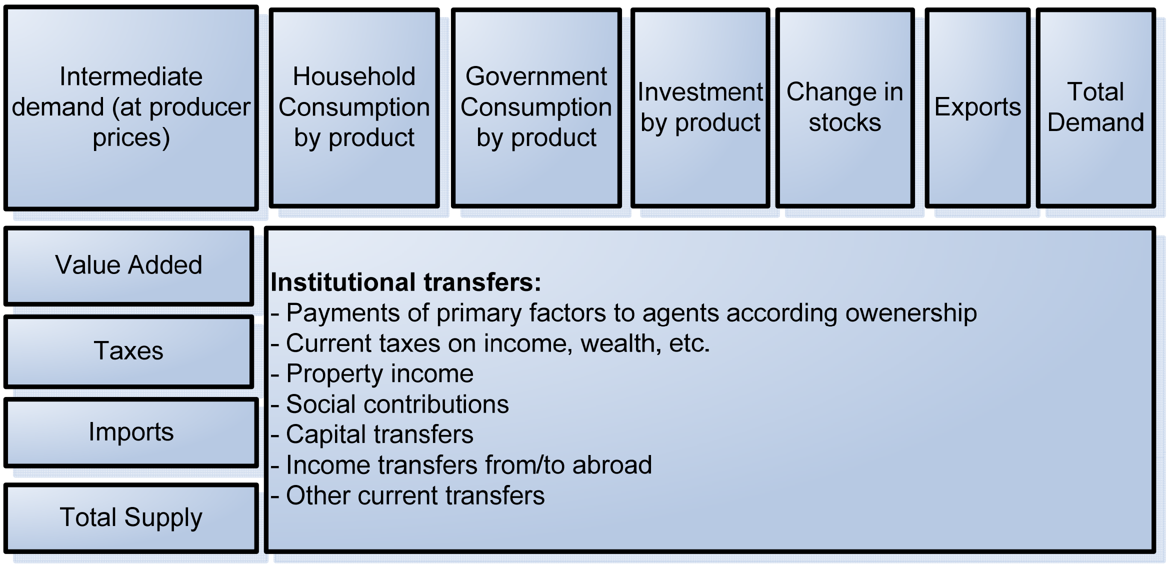
1.2 Modelling approach

In this paper we use a CGE model to quantify the welfare losses related to tax increases in the EU. As noted earlier, such an approach offers the advantage of considering altogether the different interactions in the economy, including the interactions between countries, which is particularly relevant in the EU context given the high level of integration of the EU Member States. The EU-version of the GEM-E3 model (General Equilibrium Model for Energy-Economy-Environment interactions) is a computable general equilibrium (CGE) model, which explicitly models 24 EU member states and the rest of the world. The GEM-E3 models the interactions between the economy, the energy system and the environment at country and EU level. It covers all production sectors (aggregated to 18) and institutional agents of the economy. The model computes the equilibrium prices of goods, services, labour and capital that simultaneously clear all markets under the Walras law. It formulates separately the supply or demand behaviour of the economic agents which are considered to optimise individually their objective while market derived prices guarantee global equilibrium. Further details of the model are given in the GEM-E3 Manual (European Commission, 2012).[4]

[4] For more information see also www.GEM-E3.net.

Figure 1

Social Accounting Matrix Representation as Used in GEM-E3



Source: European Commission (2012).

As discussed earlier, the use of a CGE model to calculate the MCF represents only one possible way of quantifying the welfare effect of tax increase. Such a CGE approach allows us to provide rather comprehensive approach across countries and tax categories with potentially important policy implications. Three main features of our model are especially illustrative in this respect. First, the calibration of the GEM-E3 model is based on social accounting matrices (SAMs) for 2005. As a result, the tax rates are calibrated as an effective rate, *i.e.*, the ratio between the tax revenues and the corresponding tax base for each tax category as reported in the SAMs, which provides a fairly reliable picture of the economy and the tax. The SAMs are calibrated to a base year data (2005) for each EU country built by combining input-output tables (as published by EUROSTAT) with national accounts data. Bilateral trade flows are also calibrated for each sector, taking into account trade margins and transport costs. Total demand (final and intermediate) in each country is optimally allocated between domestic and imported goods, which are assumed to be imperfect substitutes (the "Armington" assumption). Production is modelled through CES KLEM (capital, labour, energy and materials) production. Second, the GEM-E3 model offers a great level of detail regarding tax systems as it distinguishes between nine categories of government receipts, namely indirect taxes, environmental taxes, direct taxes, value added taxes, production subsidies, social security contributions, import duties, foreign transfers and government firms. These receipts are coming from product sales (*i.e.*, from branches) and from sectors (*i.e.*, agents) as described in the SAM. Unemployment benefits are part of the transfer from the government to the household sector which is a single aggregate in the SAM. We thus use observed unemployment benefit transfers to the household sector for the year 2005 which also include all other transfers related to the unemployment status (e.g., child benefit) as reported by the OECD in 2005. The latter is particular relevant to take into account the potential income loss from becoming unemployed. Third, the GEM-E3 model comprises all sectors of the economy broken down into 18 sectors while private consumption is divided among 13 durable and non-durable goods. Such level of detail allows for a consistent evaluation of the effects of tax policy changes for the different sectors of activity and economic agents. Figure 1 sketches out the main elements of these country-specific SAMs.

Though this particular CGE model does have considerable detail of taxation, one should note that the ability to fully represent the complexities of tax systems is limited. For instance, labour taxation is modelled to the representative unit of labour, which cannot incorporate the details of the (progressive) labour tax policies found in member states. Furthermore we do not aim to capture potential dynamic effects of tax changes. It is important to note also that the version of the GEM-E3 model used here includes labour market imperfections including involuntary unemployment. Due to these imperfections, employees enjoy a wage premium on the top of the wage rate that would result from non-distorted labour markets. We follow the approach of Shapiro and Stiglitz (1987) suggesting a positive correlation between wages and labour productivity (see also Blanchflower and Oswald, 1994, for empirical evidence).

The introduction of labour market imperfections has two important implications when it comes to estimating the MCF and comparing the results of labour taxes versus other tax categories. First the degree of labour market "imperfection", *i.e.*, the gap between the efficiency wage and the wage that would result from a perfect labour market where potential supply matches labour demand is likely to influence the MCF. A large wage premium should result in a greater distortive effect of labour taxation in particular. Labour market imperfections could also magnify trade-related tax spillovers effects to the extent that wages are set in some countries by partly taking into account evolutions in the main trading-partner countries (e.g., in as Belgium).

1.3 Measuring the marginal cost of public funds with GEM-E3

The measurement of welfare is central to the analysis of MCF. The welfare measure used in GEM-E3 is derived from the utility maximisation behaviour of the representative household. Here we only provide the specification of the utility function and the budget constraint, further details on the model can be found in European Commission (2012). The households receive income from their ownership of production factors (such as working time and capital), from other institutions and transfers from the rest of the world. Household expenditure is allocated between consumption, tax payment and savings. The representative household firstly decides on the allocation of its income between present and future consumption of goods and leisure. At a second stage, the household allocates its total consumption expenditure between the different consumption categories available. The consumption categories are split in non-durable consumption categories (food, culture etc.) and services from durable goods (cars, heating systems and electric appliances).

The general specification of the first stage problem, with a time separable Stone-Geary utility function, can be written as follows:

$$U_{i,t} = \sum_t \left(1+stp_{i,t}\right)^{-t}.\left(bh_{i,t}.\ln\left(HCDTOTV_{i,t}-ch_{i,t}\right)+bl_{i,t}.\ln(LJV_{i,t}-cl_{i,t})\right) \tag{1}$$

where $HCDTOTV_{i,t}$ represents the consumption of goods (in volume), $LJV_{i,t}$: the consumption of leisure, $stp_{i,t}$: the subjective discount rate of the households, or social time preference,
$ch_{i,t}$ is the subsistence quantity of consumption, $cl_{i,t}$ the subsistence quantity of leisure,
$bh_{i,t}$, $bl_{i,t}$ are the respective shares of consumption and leisure in the disposable income of the households. The maximisation is subject to the following inter-temporal budget constraint, which states that all available disposable income will be spent either now or sometime in the future:

$$\mathring{a}_t\ (1+r_{i,t})^{-t}\cdot(HCDTOT_{i,t}-PCI_{i,t}\cdot ch_{i,t}+PLJ_{i,t}\cdot LJV_{i,t}-PL_{i,t}\cdot cl_{i,t}) \tag{2}$$

where $r_{i,t}$ is the discount rate, $HCDTOT_{i,t}$ is the total private consumption, $PCI_{i,t}$ is the consumer price index, $PLJ_{i,t}$ is the price of leisure, $LTOT_{i,t}$ is the total available time to households. The non-wage income is income such as interest payments from assets, share in firms' profits, social benefits, and remittances. Based on myopic assumptions about the future, the household decides the amount of leisure that wishes to forsake in order to acquire the desired amount of income (thus also defining labour supply behaviour).

$$Welfare_i = \frac{1}{\exp(MUI)_i}\exp\left(bh_i * \ln(HCDTOT_i - ch_i) + bl_i * \ln(LJV_i - cl_i)\right)$$

where MUI is the marginal utility of income. Note that for the purposes of this version of the model, the leisure component is fixed, and therefore the changes in welfare occur only through the changes in consumption. The estimation of the MCF can be undertaken using a general equilibrium approach encompassing all the potential market effects of a given tax increase as well as the interactions between economic agents and resulting changes in the tax bases. The MCF can be calculated using the following formula:

$$MCF_{i,k} = \frac{\Delta W_{i,k}}{\Delta TR_i} \tag{3}$$

where $\Delta W_{i,k}$ is the welfare loss due to the increase of tax k in country i and is calculated as the change in consumer utility based on the indirect utility function in order to give it a monetary value. It could be conceptualised as the reduction in consumption relative to a benchmark case of no-policy change, where prices and incomes are fixed at their "no-policy-change" benchmark level. This technically corresponds to the "equivalent" variation. Alternatively, using the "compensating" variation would imply using the prices and income corresponding to "policy change" scenario. See Dahlby (2008) and Schöb (1994) for a discussion. The term ΔTR_i in equation (3) represents the corresponding change in tax collection in country i (including all tax revenues).

The MCF provides a metric for the loss in welfare (the efficiency loss) per unit of tax revenue gain. If the MCF equals one, then the tax is equivalent to a lump-sum transfer from the households to the government with no distortion. Typically, however, the MCF is greater than one such that MCF $=1+\alpha$, with α representing the cost of the distortion. This means that for every euro that goes into the government's purse, the economy pays an efficiency cost of α euros. The higher the MCF, the larger is the cost of distortion compared with the tax revenue gains.

As mentioned above, the externality modelled in GEM-E3 stems from bilateral trade relationships. A given tax policy change will affect bilateral trade flows and, thus, economic activity (*i.e.*, production and consumption). It will also impact on tax revenues via two channels: tax changes will affect both (i) relative prices of domestically produced versus foreign goods and services and (ii) disposable income through changes in price levels and purchasing power. Tax changes will also spill through the production chain: for instance countries importing intermediates from a country implementing a tax increase will face higher production costs if substitution possibilities (*i.e.*, import from alternative suppliers) are limited. Tax changes also affect demand for intermediates produced abroad. A country implementing a tax increase will thus face a competitiveness loss as well as lower purchasing power. Furthermore, partner countries may benefit on the one hand from a price-competitiveness gain if their exports are close substitutes of the goods and services produced by the tax-increasing country. On the other hand, partner countries may eventually lose if their exports are complementary to those of the tax-increasing country or if the lower economic activity in the tax-increasing country reduces its imports from the partner country.[5]

Alternatively, one can also derive a measure of the MCF where tax-related spillovers are taken into account by considering unilateral tax increases as indicated in equation (4) below:

$$MCF_{i,k} = \frac{\Delta W_j}{\Delta R_j + \sum_{j,j \neq i} \Delta R_j} + \frac{\sum_{j,j \neq i} \Delta W_j}{\Delta R_j + \sum_{j,j \neq i} \Delta R_j} \tag{4}$$

where i is the country implementing a given tax change while j are the other countries (not implementing any tax change). The second term of equation (4) represents the spillover effect which can be compared to the first term of equation (3) which represents the impact of a tax change for the country implementing it only. The average MCF for unilateral tax increases calculated as in (3) can then compared to the average value of the MCF for unilateral tax increases including the impact of unilateral tax increases on other countries welfare and tax revenues as calculated in (4).

The results presented here provide estimations of the MCF for a very small tax increase of 0.05 percentage points of the effective tax rate in 2005. The tax increase in the case of labour tax concerns total social total security contribution. In doing so, we aim at focusing on the labour "price" effect of taxation specifically. The green taxes considered here concerns an energy tax for households per petajoule of energy (which is the measure commonly used to express energy consumption by large customers groups such as countries). It is important to note that the effects of an energy tax increase on the utility level as a result of a better environmental quality due to lower CO2-emissions and other kinds of air pollution, is not taken into account here such that the resulting utility variation stems essentially from the traditional price and income effects of a price change of each product consumed by the representative consumer.

The small tax increment is intended to capture the marginal nature of the tax change. In practice the proceeds of a given tax increase are used to finance policy objectives such as an increase in public expenditure, a subsidy, or to repay public debt. As the impact of the allocation of tax proceeds is beyond the scope of this paper, the estimate of the MCF of a given tax increase is isolated by allocating the (small amount of) additional tax revenues to the rest of the world (*i.e.*, outside the EU). It is important to note also that when changing the level of taxes we fix the level of leisure to a given level. This is done in particular in order to isolate specifically the effect of labour taxes on time spent in employment and in unemployment. Given the labour market setting used, this means also that unemployment is never voluntary and thus neutralises the substitution effect of hours worked with time spent in leisure.

Table 1 provides descriptive statistics on the share of total labour taxes and energy taxes by country for the year 2005 which is used for the calibration of the model. The main source for the data is EuroStat. As one would expect, the labour taxes are substantially larger in EU countries (the simple average for labour taxes is 20.7 per cent of GDP vs. 1.4 per cent for energy taxes) although the relative dispersion of energy taxes is greater across countries (the coefficient of variation in 32.7 per cent for energy taxes vs. 25.9 per cent for labour taxes). Overall these figures also reflect the relatively large share of labour taxes in the richer EU countries.

---

[5] Andersen and Sørensen (2012) suggested recently that tax increases could also have positive side-effects on the production side since firms needed to counter-act the extra-tax burden through productivity improvement.

Table 1

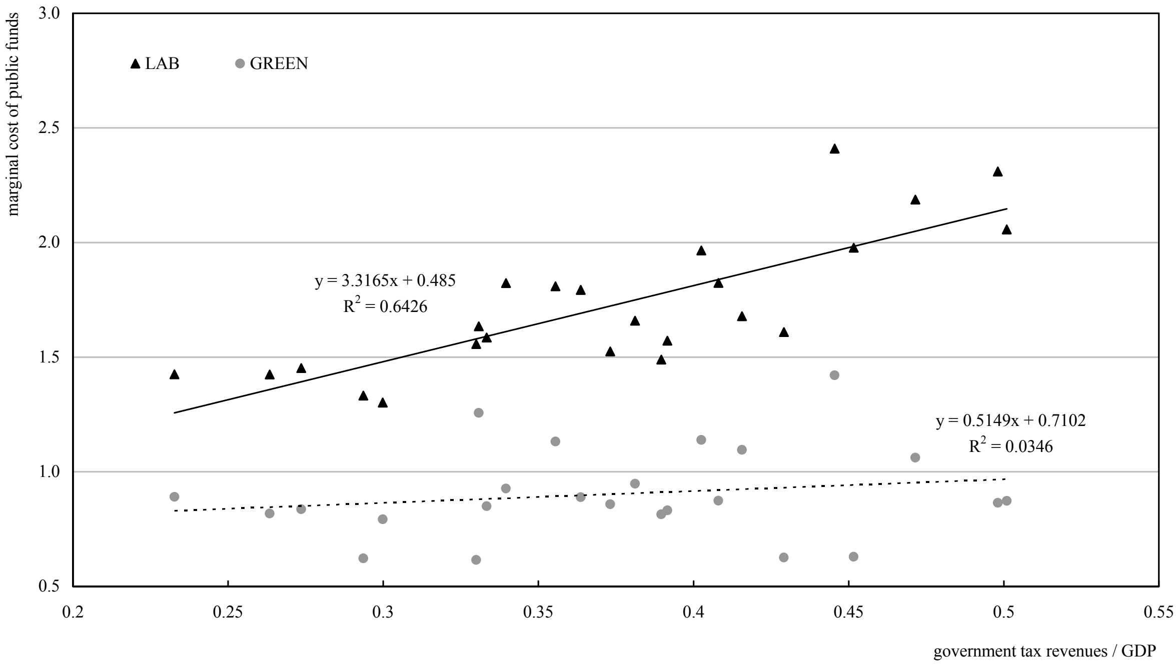
Share of Tax Revenues in GDP: Values Used for the Calibration of the GEM-E3 Model

Country	Total Tax Revenues	Labour Taxes*	Green Taxes**
Austria	40.8%	26.6%	1.5%
Belgium	45.2%	29.1%	1.0%
Bulgaria	33.0%	13.3%	2.8%
Czech Republic	39.0%	20.9%	1.7%
Denmark	49.8%	26.6%	1.4%
Estonia	30.0%	29.0%	1.0%
Finland	42.9%	16.2%	1.3%
France	44.6%	18.1%	1.3%
Germany	40.3%	19.9%	1.8%
Greece	33.3%	26.1%	1.0%
Hungary	37.3%	26.9%	2.0%
Ireland	29.4%	19.8%	0.8%
Italy	41.6%	14.7%	2.1%
Latvia	26.3%	23.7%	1.4%
Lithuania	27.4%	15.3%	0.6%
Netherlands	39.2%	14.3%	1.4%
Poland	33.1%	21.7%	1.4%
Portugal	34.0%	16.8%	1.6%
Romania	23.3%	18.1%	1.1%
Slovakia	47.1%	13.0%	1.5%
Slovenia	38.1%	29.7%	1.9%
Spain	36.4%	20.5%	1.0%
Sweden	50.1%	16.0%	1.3%
United Kingdom	35.6%	21.0%	1.7%

* Households' social security contributions + labour income tax.
** Energy taxes paid by households.

Figure 2

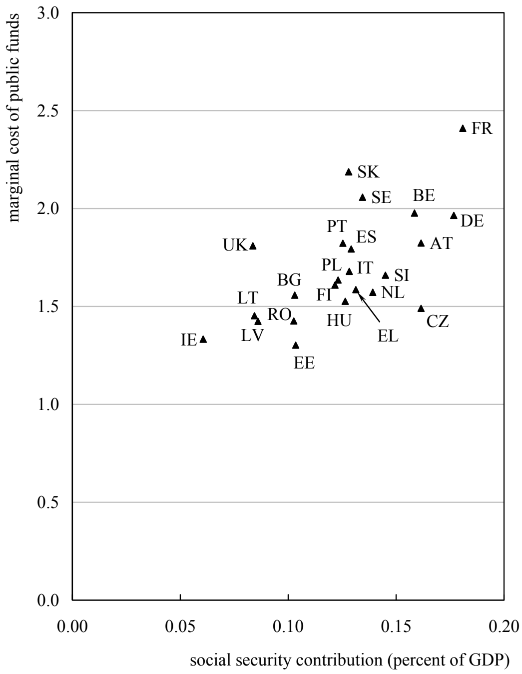
Marginal Cost of Public Funds Vs. Total Tax Revenues



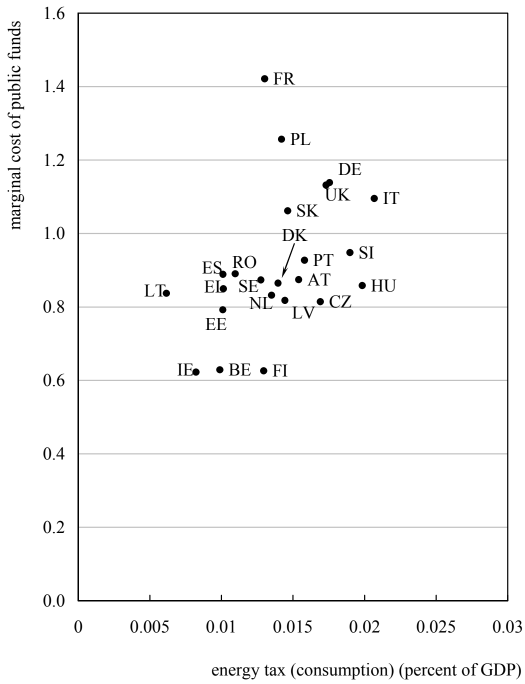
Source: GEM-E3 simulations.

Country Details for Labour and Energy Taxes

MCF Labour Vs. Labour Tax (SSC)
(percent of GDP)



MCF Green Vs. Green Tax Revenues
(percent of GDP)



2 Results

The results presented here focus firstly on the comparison of the MCF for labour and green taxes both across the EU and for individual countries, as well as the notion of tax shifting from labour to green taxes. Secondly, the impacts of each country changing their tax rates on the rest of the EU – the spillover effects – are considered. Thirdly, the investigation into the important of labour market flexibility is reported. Lastly, a robustness check on the values of the labour supply elasticity, Armington elasticities, and revenue-recycling strategies are carried out.

2.1 The marginal cost of public funds for labour versus green taxes: Individual country and EU-wide results

The MCF is calculated for each EU country introducing each tax unilaterally. The key results are reported in Table 2, which compares the GDP-weighted value for the within country MCF (corresponding to equation 1 above) for labour and energy taxes. These results show that the efficiency losses from green taxes are far smaller than for labour taxes. Considering EU-wide figures, the value for labour taxes of 1.90 implies that to raise an additional 1 euro of revenue, the average efficiency loss would be 0.90 euros. In contrast, raising an additional 1 euro of revenue from energy taxes, leads to an average efficiency loss of only 8 cents. Note that these values obtained for the MCF are broadly in line with the existing literature commented in Section 1. The result is also consistent with economic theory, which suggests that taxing relatively inelastic goods, such as energy, will result in only small distortions. This is not the case for labour if one is faced with a labour supply curve that is at least somewhat elastic. Furthermore, increased unemployment also requires additional social security payments from the government, which is also incorporated in the model. The detailed country results also bring results in line with prior expectations whereby countries with high starting level of taxation have also the highest values of the MCF. An important point to note regarding the energy taxes is that it is possible for MCF values to fall below one in some countries. This reflects the situation where a good is, in effect, under-taxed from an efficiency perspective, and raising the tax improves the overall efficiency of the economy. Tax efficiency, in this sense, is similar to the notion first put forward by Ramsey (1927), which proposed that consumption taxes for a particular good should be proportional to the inverse of the price elasticity of demand. The relative inelasticity of demand for energy taxes tends to make them good candidates for efficient taxation.

Regarding the MCF of labour taxes, there is a fair range across different countries from only 1.30 in Estonia to 2.41 in France. For the MCF of green taxes, the range is from 0.62 in Bulgaria to 1.42 in France. An important point to notice is that in every country, the MCF for labour taxes is higher than for green taxes, suggesting that all countries would see an efficiency gain from switching from labour to green taxes. These country values are compared with the total tax share of GDP in each country in Figure 2. For example, the highest potential losses from tax hikes are found for France, which has a MCF of 2.41 for labour taxes and a tax share of GDP of 44.6 per cent. Focusing firstly on labour taxes (the triangles), there is a tendency for those countries with a higher tax share of GDP to also have a higher MCF. This is consistent with the notion of the Laffer Curve, which suggests that as overall taxes rise, further taxation at the margin becomes progressively less efficient. Interestingly, this notion does not hold for green taxes where there is no clear relationship between the overall tax burden and the MCF, suggesting that (on average) green taxes are especially efficient in comparison to labour taxes for countries that have a high overall tax share. It is also interesting to note that the effect of green tax appears to be more heterogeneous across countries than labour taxes which could be explained by the original diverse taxation of energy-intensive products in EU Member States contrary to rather homogeneous factor labour. This

Table 2

Marginal Cost of Public Funds for Labour Taxes and Energy Taxes

Country	Labour Taxes	Green Taxes
Austria	1.82	0.87
Belgium	1.98	0.63
Bulgaria	1.56	0.62
Czech Republic	1.49	0.81
Germany	1.96	1.14
Denmark	2.31	0.86
Estonia	1.30	0.79
Greece	1.59	0.85
Spain	1.79	0.89
Finland	1.61	0.63
France	2.41	1.42
Hungary	1.53	0.86
Ireland	1.33	0.62
Italy	1.68	1.10
Lithuania	1.45	0.84
Latvia	1.42	0.82
Netherlands	1.57	0.83
Poland	1.63	1.26
Portugal	1.82	0.93
Romania	1.43	0.89
Sweden	2.06	0.87
Slovenia	1.66	0.95
Slovakia	2.19	1.06
United Kingdom	1.81	1.13
EU average (GDP-weighted)	*1.90*	*1.08*
Simple average	*1.73*	*0.90*
Coefficient of variation	*17.38%*	*22.21%*

point is illustrated by considering separately the values of the MCF against the initial tax burden of labour and energy tax separately in the country-specific results reported in Figure 2.

Raising tax rates in a single country primarily affects welfare in that country, but there are also spillover effects to other EU countries. Comparing the individual country results for MCF with the EU-wide results shows the extent of these spillover effects. The EU-wide MCF is calculated according to Equation 2 above. Table 3 compares the individual country MCF with the EU-wide MCF for labour taxes. The spillover effect reported here refers to the percentage of the total EU-wide MCF that is *not* accounted for in the individual country MCF. For example, for Germany the EU-wide MCF is 2.04, of which 1.96 is the individual country effect. Therefore, in percentage terms the spillover effect is 3.6 per cent of the total effect.[6] As can be seen, the spillover effects are typically modest for labour taxes. The countries with the highest percentage spillover effects (Belgium, Denmark and the Netherlands) are relatively small countries, with high trade to GDP shares. Table 4 reports the individual country and EU-wide MCFs for energy taxes and calculates the spillover effects. One difference in comparison to the comparable values for labour taxation in Table 3 is that the spillover effects, on average, represent a much higher percentage of the total EU-wide MCF. This reflects that energy-intensive goods tend to be more intensively traded than the average of the economy.

Finally one should note that the results reported in Table 1 do not allow us to say anything about the importance of each country on the magnitude of a welfare change given that the MCF measure is the ratio between this variable and the tax revenue variation. In order to check this we have calculate the share of each country in the welfare variation and the tax revenue variation of the spillover component of equation (4). These calculations indicated that some countries have a more prominent role because of their size (Germany, France and the UK are the salient cases) or because of their degree of openness to the rest of EU economies (which is the case for Belgium or the Netherlands). We also looked at the role of each separate country on the EU-wide spillovers considering separately positive and negative effects on welfare and tax revenues. As in the case of labour taxes, we again observed that the large EU countries generate most of the spillovers although here some relatively small albeit open countries tend to play a bigger role (e.g., Belgium and the Netherlands in particular). The sign of the spillover effect was predominantly negative, thus suggesting that, *ceteris paribus*, a tax increase in a given country deteriorates the overall EU economic efficiency.

More generally, our results suggest overwhelmingly that should tax increases be considered in EU countries, energy taxes represent a better candidate than labour taxes. One possible reason for this could be that labour taxes have a bearing on labour supply and production levels. Green taxes in turn only impact on consumption and only indirectly on labour supply (through the level of post-tax increases level of income). In a second best world, a new distortion balances other distortions and the equalisation of the MCF across tax categories suggests that energy is relatively under-taxed compared to labour taxes, at least in the EU countries considered here. This result is not necessarily surprising given that the MCF is known to increase linearly with the level of taxation, see Dahlby (2008) such that it is generally a better option to increase low- burden tax rates rather than increasing tax rates which are already at a high level. Our investigation of the cross-country spillovers on energy taxes provides more nuanced results, however. Adopting the view of a benevolent EU-tax policy makers would certainly advocate for increasing the green rather than the labour tax, although the advantage of the former over the latter becomes less important once cross-country spillovers are considered. Indeed our analysis shows that these spillovers are potentially more important for energy rather than for labour taxes. This result in a way illustrates the theoretical finding by Bovenberg and De Moij (1994) who showed that the

---

6 The calculation is: (2.04 – 1.96) / 2.04 = 3.6%.

Table 3

MCF of Labour Taxes: Country Vs EU-wide Effects

Country	Country-level MCF	EU-level MCF	Spillover Effect*
Austria	1.82	1.91	4.30%
Belgium	1.98	2.29	13.52%
Bulgaria	1.56	1.59	1.77%
Czech Republic	1.49	1.50	0.97%
Germany	1.96	2.04	3.63%
Denmark	2.31	2.56	9.69%
Estonia	1.30	1.36	4.20%
Greece	1.59	1.60	0.88%
Spain	1.79	1.84	2.37%
Finland	1.61	1.66	2.77%
France	2.41	2.50	3.71%
Hungary	1.53	1.58	3.71%
Ireland	1.33	1.41	5.27%
Italy	1.68	1.68	–0.19%
Lithuania	1.45	1.49	2.47%
Latvia	1.42	1.49	4.27%
Netherlands	1.57	1.69	7.00%
Poland	1.63	1.63	–0.36%
Portugal	1.82	1.93	5.34%
Romania	1.43	1.42	–0.56%
Sweden	2.06	2.15	4.37%
Slovenia	1.66	1.78	6.80%
Slovakia	2.19	2.22	1.46%
United Kingdom	1.81	1.86	2.76%
EU (GDP-weighted)	*1.90*	*1.97*	*3.49%*
Simple average	*1.73*	*1.80*	*4.04%*
Coefficient of variation	*17.38%*	*18.99%*	

* Calculated as the percentage of the second term in the right hand side of equation (2) divided by the MCPF measured for the EU. The change in the labour tax concerns total social security contribution paid by the employers and the employees. The tax increase is equal to 0.05 percentage points.

Table 4

The MCF of Green Taxes: Country Vs EU-wide Effects

Country	Country-level MCF	EU-level MCF	Spillover Effect* *(percent of total MCF)*
Austria	0.87	1.07	18.3%
Belgium	0.63	0.87	27.9%
Bulgaria	0.62	0.64	4.6%
Czech Republic	0.81	0.87	6.5%
Germany	1.14	1.24	8.2%
Denmark	0.86	0.93	6.5%
Estonia	0.79	0.92	13.5%
Greece	0.85	0.90	5.5%
Spain	0.89	0.98	9.5%
Finland	0.63	0.70	10.6%
France	1.42	1.54	7.7%
Hungary	0.86	1.01	14.6%
Ireland	0.62	0.88	29.5%
Italy	1.10	1.14	3.6%
Lithuania	0.84	0.95	11.8%
Latvia	0.82	0.84	2.1%
Netherlands	0.83	0.97	14.4%
Poland	1.26	1.27	1.1%
Portugal	0.93	1.06	12.9%
Romania	0.89	0.95	6.0%
Sweden	0.87	0.95	8.0%
Slovenia	0.95	1.10	13.7%
Slovakia	1.06	1.17	9.5%
United Kingdom	1.13	1.17	3.6%
EU (GDP-weighted)	*1.08*	*1.17*	*7.8%*
Simple average	*0.90*	*1.00*	*10.2%*
Coefficient of variation	*22.21%*	*19.02%*	

* Calculated as the percentage of the second term in the right hand side of equation (2) divided by the MCPF measured for the EU. The change in the Energy tax concerns the energy consumption by households (in real terms). The tax increase is equal to 0.05 percentage points.

optimal level of environmental taxes lied below the Pigouvian level once tax interactions were considered. Our results show similarly that when countries' interactions are considered the advantage of raising green versus labour taxes is reduced although green taxes increases remain a better option than labour tax increases thus suggesting that potential tax shifting between labour and energy taxes would yield significant benefits in terms of economic efficiency.

2.2 The role of labour market flexibility

The degree of labour market flexibility reflects the extent to which a change in wages resulting from a tax increase affects the supply of labour. By altering the degree of labour market flexibility, we address the question of whether the real wage reflects the marginal product of labour or whether wage rigidity, linked to labour market imperfection, hinders such an adjustment (see, in particular, Boeters and Savard, 2011, for a review of the literature, and Hutton and Ruocco, 1999, for an example of analysis of the impact of tax changes with efficiency wage in a CGE model). In the labour market setting adopted here, the tax change will not be fully reflected in the real wage because of the existence of a wage premium of certain categories of workers. In such a setting the interaction between the tax system and the labour market setting can be non-negligible, especially, though not exclusively, when considering labour tax changes.[7] The version of GEM-E3 used in this paper includes a labour market setting consistent with the efficiency wage theory of Shapiro and Stiglitz (1987). This theory posits, firstly, that the productivity of labour has a positive correlation with wages leading firms to offer a wage premium, and secondly, that this wage premium increases with lower employment. In periods of high unemployment firms have less need to offer high wages to attract more productive workers or to increase productivity of existing workers. The wage setting in such model is given by the following expression:

$$w \cdot \frac{PCI}{\overline{PCI}} = \overline{w}\overline{r} + e + \frac{e}{q} \cdot \left[\left(\frac{b}{u} \right)^{eg} + r \right] \tag{3}$$

where PCI is the consumer price index and eg an adjustment parameter to reflect the different labour market flexibility conditions that prevail in each country, b is the quit from job rate, u is the actual unemployment rate, r is the interest rate, w is the wage rate, e is the disutility from working (for the "shirker" e=0) and q measures the efficiency of the workforce, see European Commission (2012) for more details on the derivation of equation (3). In this equation, the degree of labour market flexibility in the model is captured in the parameter eg, which can be adjusted. A higher eg indicates a higher degree of labour market flexibility, *i.e.*, according to equation (3) the higher the transmission of the quit rate and the lower the impact of unemployment changes on the real wage level.[8] Re-running the model with different values of eg allow an investigation into the

---

[7] Note that in our model there is only one representative individual and only one tax rate for each tax category based on the calibration using the data contained in the SAMs. For the labour market in particular we thus consider only one country average effective tax rate for each tax category. Therefore the progressivity of tax systems is not accounted for. Studies tend to show that the labour tax progressivity can have non-trivial effects on labour supply and therefore on the MCF (see in particular Lockwood and Manning, 1993).

[8] There is arguably no specific reason for choosing a specific value for eg against another one, as the highly stylised representation of the labour market used in the version of GEM-E3 allow us to say little about whether this is convenient or not. One could argue, for instance, that since the eg parameter should represent as closely as possible the degree of flexibility of the labour market, country-specific values should be set in accordance to "estimated", e.g., by the labour market literature. In fact, this is only partly true in the labour market setting outlines in Appendix 2, given that, while the parameter eg is set at an *ad hoc* value, the level of unemployment used is taken from observed data. Instead of trying to stick to some *ad hoc* country-specific measure of labour market flexibility, we chose instead to keep the same value of this parameter across countries and rather to check whether the MCF estimates change when the degree of flexibility is higher or lower than in our benchmark cases, without inferring too much about whether this degree of flexibility reflects the reality of EU countries labour markets. In adopting this approach, we are therefore more interested in the change in the value of the MCF on average across EU countries rather than on whether the country-specific degrees of "flexibility" are correctly reflected.

impact of labour market flexibility on the MCF.[9] Our high flexibility scenario involved doubling *eg*, whereas our low flexibility scenario involved halving *eg*. These are large hypothetical changes in order to allow us to explore the responsive of the MCF values without being intended to reflect possible policy changes affecting the labour market. Table 5 shows the results for the high and low labour market flexibility cases for the labour tax MCF and green tax MCF respectively for the EU as a whole. These results clearly shows a large impact on the MCF for labour taxes, with a less flexible labour market raising the EU average MCF (GDP-weighted) by 33.6 per cent to 2.54 and a more flexible labour market reducing it by 13.6 per cent to 1.64. These results should not come as a surprise given that labour market flexibility affects directly the way the change in wage costs is transmitted to the employment level, such as from a marginal rise in labour taxes. Nevertheless, the results do demonstrate the importance of labour market flexibility for the MCF of labour taxes. By contrast, the effect on the MCF of energy taxes is much less pronounced. On average, the MCF rises by less than 5 per cent under less flexible labour market conditions and is reduced by just over 3 per cent under more flexible conditions. The country-specific results are shown in Tables 13 and 14 in the Appendix. These show some interesting features, however given that in some cases the efficiency wage assumption does not fully capture the degree and nature of the rigidity of each specific labour market, we feel that the country-specific results should be interpreted with care. For example, Spain barely experiences a change in its MCF while this country is known to have especially distorted labour market, whereas other large countries, especially France and Germany, show large fluctuations in the MCF for labour taxes.

3 Robustness checks

We provide a number of additional results to the analysis carried out above in order to verify their robustness to alternative assumptions regarding the values of the labour supply elasticities, which may ultimately affect the number of hours worked in our model where time worked is chosen against leisure or unemployment. In addition, given that we consider EU economies, which are closely linked together through international trade, we also provide alternative estimates of the MCF depending on the degree of substitution between domestic production and imported goods. This is done by specifying alternative assumption regarding the Armington elasticities. Finally we also consider alternative hypotheses regarding the recycling of the extra-tax revenues yielded from the marginal tax increases in order to check whether our central benchmark case (*i.e.*, through a direct income transfer to the rest of the world) does not influence our results.

In order to investigate the impact of the labour supply elasticities on the MCF values, we replaced the labour supply elasticities with values from the literature, where available, and average values otherwise. Specifically, we took the values for labour supply elasticity from Evers *et al.* (2008). This study reports estimates of labour supply elasticity for selected countries for men and women separately. We took these values and weighted them by gender share in the workforce to give an overall value using Eurostat data for 2005. This gave us estimates for France, Sweden, Germany, Italy and the Netherlands. Two further countries, UK and Finland, have values for women only. Using the average ratio of the elasticity of men to women, we further completed the missing estimates for the overall elasticity in these two countries. For the rest of the EU, we took an average of these values. We then recalibrated our model to have these labour supply elasticities, and re-ran the simulations to calculate the MCF for labour and energy taxes. The values of the base labour supply elasticities are compared with those used in this robustness check in Table 6.

As can be seen from Table 7, the average, GDP-weighted MCF is lower when using these elasticities – the individual country average falls from 1.90 to 1.62 and the EU-wide average falls

[9] Note that, in this case, the values for *ef* must be recalibrated.

Table 5

The Marginal Cost of Public Funds and Labour Market Flexibility: The Case of Labour Tax

	MCF, Benchmark Case	Less Flexible Labour Market	More Flexible Labour Market
Labour Taxes	**1.90**	**2.54**	**1.64**
EU average (GDP-weighted)			
percent change vs. benchmark		*+33.6%*	*–13.6%*
Green Taxes	**1.08**	**1.13**	**1.04**
EU average (GDP-weighted)			
percent change vs. benchmark		*+4.6%*	*–3.3%*

Table 6

Labour-supply Elasticities: Base Vs. Robustness-check Values

Country	Base L-supply Elasticity Values*	New L-supply Elasticity Values*
Austria	0.520	0.346
Belgium	0.761	0.346
Bulgaria	0.474	0.346
Czech Republic	0.405	0.346
Germany	0.611	0.024
Denmark	0.814	0.346
Estonia	0.511	0.346
Greece	0.646	0.346
Spain	0.820	0.346
Finland	0.709	0.019
France	0.657	0.179
Hungary	0.533	0.346
Ireland	0.471	0.346
Italy	0.481	1.173
Lithuania	0.685	0.346
Latvia	0.691	0.346
Netherlands	0.521	0.554
Poland	0.577	0.346
Portugal	1.154	0.346
Romania	0.601	0.346
Sweden	0.670	0.389
Slovenia	0.778	0.346
Slovakia	0.532	0.346
United Kingdom	0.816	0.085

* Base values calculated from GEM-E3 model; new values derived from Evers *et al.* (2008, see text above).

Table 7

MCF with Different Labour-supply Elasticities: Labour Taxes

Country	Country-level MCF		EU-level MCF	
	Base L-supply Elasticity	New L-supply Elasticity	Base L-supply Elasticity	New L-supply Elasticity
Austria	1.82	1.69	1.91	1.72
Belgium	1.98	1.59	2.29	1.68
Bulgaria	1.56	1.60	1.59	1.62
Czech rep.	1.49	1.51	1.50	1.51
Germany	1.96	1.32	2.04	1.24
Denmark	2.31	1.66	2.56	1.72
Estonia	1.30	1.31	1.36	1.34
Greece	1.59	1.47	1.60	1.47
Spain	1.79	1.88	1.84	1.86
Finland	1.61	1.51	1.66	1.44
France	2.41	1.75	2.50	1.73
Hungary	1.53	1.48	1.58	1.50
Ireland	1.33	1.35	1.41	1.41
Italy	1.68	1.96	1.68	2.01
Lithuania	1.45	1.51	1.49	1.49
Latvia	1.42	1.42	1.49	1.44
Netherlands	1.57	1.48	1.69	1.62
Poland	1.63	1.61	1.63	1.58
Portugal	1.82	1.61	1.93	1.62
Romania	1.43	1.52	1.42	1.48
Sweden	2.06	1.82	2.15	1.86
Slovenia	1.66	1.56	1.78	1.60
Slovakia	2.19	2.29	2.22	2.27
United Kingdom	1.81	1.51	1.86	1.52
EU (GDP-weighted)	*1.90*	*1.62*	*1.97*	*1.61*
Simple average	*1.73*	*1.60*	*1.80*	*1.61*
Coefficient of variation	*17.4%*	*13.7%*	*19.0%*	*13.9%*

from 1.97 to 1.61. Note that the net spillover effects are near-zero when using the new elasticities. Nevertheless, the pattern is quite closely related to the base case with a correlation coefficient for the individual country values of 0.58. In the case of energy taxes, shown in Table 8, the GDP-weighted values for the EU also fall from 1.08 to 1.01 for individual country MCF, and from 1.17 to 1.06 for the EU-wide MCF. The values for MCF closely reflect the base values with a correlation coefficient of 0.97 for the individual country MCFs. Considering both Table 7 and 8, one notes that the relative size of the MCF for labour and energy taxes tells the same story as our base case, strongly suggested that our main result – that energy taxes are generally less distortionary than labour taxes – is robust to these new specifications.

As noted, an important feature of our CGE model, GEM-E3, is the modelling on international trade. The price sensitivity of these trade flows is determined primarily by the trade elasticities in the model. These elasticities are always somewhat uncertain, and therefore, it is good practice to test the robustness of our results against alternative values. Four extra model runs are carried out for each tax type and the MCF re-estimated. These are (i) increased then (ii) decreased import (Armington) elasticities, and then (iii) increased then (iv) decreased export elasticities. Tables 9 and 10 show the EU average results (GDP-weighted). The values reported as "base trade elasticity" are the benchmark results (as reported in Table 2). One can detect a minor tendency for higher trade elasticities to cause higher MCF estimates. However, the main observation is that the value of the trade elasticities have little impact on the MCF, and so the conclusions are robust to such changes.

As explained in Section 2, the calculation of the MCF involves implementing a marginal increase in the tax rate. Our preferred methodology for dealing with the extra revenue raised is to give it to the rest of the world, so there is no domestic benefit from additional government spending. Nevertheless, it is sensible to try an alternative closure of the model in order to assess whether this choice unduly influences our results. With this in mind, we ran the model with the additional revenues being returned to household by means of a lump sum transfer. This was run for both labour and energy taxes, with the results being reported in Tables 11 and 12. Note that in this case, the MCF values obtained are not one plus the distortion ($1 + \alpha$), but simply the distortion itself (α), as the 1 extra-tax revenues is transferred back to households already. In order for the results tables to be comparable to the earlier values, a one has been added to the MCF estimates obtained. Evidently, the different closure rule results in a smaller MCF for labour taxes, both at the individual country and the EU-wide levels. Otherwise, the variation across countries is similar to the standard values; the correlation coefficient for the individual country MCFs is 0.68. Regarding the MCF for green taxes, again the different closure rule reduces the estimates. However as for labour taxes, the variation across countries is similar with a correlation coefficient for the individual country MCFs of 0.80. From this robustness check, we can clearly see that our main result holds – that the MCF for labour is considerably higher than for green taxes. The magnitude of the MCF in this specification is lower. We choose to rely more on our standard estimates, because with this closure, the measurement of the MCF is altered as one must now take into account the benefits from additional spending.

4 Conclusions

Our research provides some useful evidence for EU countries that are considering how to approach fiscal consolidation. Firstly, the modelling work makes a strong case that the economic distortions caused by labour taxes are greater than for green taxes. This is an important consideration when seeking to promote economic recovery. Assuming that the revenue yield would be the same, relying on energy taxation to raise revenues, rather than labour taxation, would be expected to be more efficient for the economy as a whole. This result holds for all EU member

Table 8

MCF with Different Labour-supply Elasticities: Green Taxes

Country	Country-level MCF		EU-level MCF	
	Base L-supply Elasticity	New L-supply Elasticity	Base L-supply Elasticity	New L-supply Elasticity
Austria	0.87	0.81	1.07	0.97
Belgium	0.63	0.59	0.87	0.78
Bulgaria	0.62	0.61	0.64	0.64
Czech Republic	0.81	0.80	0.87	0.85
Germany	1.14	0.99	1.24	1.01
Denmark	0.86	0.87	0.93	0.89
Estonia	0.79	0.81	0.92	0.98
Greece	0.85	0.84	0.90	0.87
Spain	0.89	0.91	0.98	0.99
Finland	0.63	0.68	0.70	0.71
France	1.42	1.26	1.54	1.32
Hungary	0.86	0.81	1.01	0.95
Ireland	0.62	0.57	0.88	0.81
Italy	1.10	1.11	1.14	1.17
Lithuania	0.84	0.85	0.95	0.79
Latvia	0.82	0.85	0.84	0.85
Netherlands	0.83	0.76	0.97	0.91
Poland	1.26	1.25	1.27	1.25
Portugal	0.93	0.89	1.06	0.98
Romania	0.89	0.91	0.95	0.95
Sweden	0.87	0.84	0.95	0.87
Slovenia	0.95	0.91	1.10	1.06
Slovakia	1.06	1.05	1.17	1.12
United Kingdom	1.13	1.08	1.17	1.10
EU (GDP-weighted)	*1.08*	*1.01*	*1.17*	*1.06*
Simple average	*0.90*	*0.88*	*1.00*	*0.95*
Coefficient of variation	*22.2%*	*20.5%*	*19.0%*	*17.2%*

Table 9

MCF with Different Trade Elasticities: Labour Taxes
(EU averages)

	Country-level MCF			EU-level MCF		
	High Trade Elasticity	Base Trade Elasticity	Low Trade Elasticity	High Trade Elasticity	Base Trade Elasticity	Low Trade Elasticity
Different import elasticities	1.91	1.90	1.88	1.97	1.97	1.96
Different export elasticities	1.90	1.90	1.89	1.97	1.97	1.96

Table 10

MCF with Different Trade Elasticities: Green Taxes
(EU averages)

	Country-level MCF			EU-level MCF		
	High Trade Elasticity	Base Trade Elasticity	Low Trade Elasticity	High Trade Elasticity	Base Trade Elasticity	Low Trade Elasticity
Different import elasticities	1.10	1.08	1.05	1.17	1.17	1.16
Different export elasticities	1.09	1.08	1.07	1.17	1.17	1.17

states modelled and despite the fact that potential welfare-enhancing effect of pollution abatement are cancelled out in our model.

Nevertheless, further investigation showed that this result is somewhat less strong when one considers the spillover effects between countries, as these are more pronounced (in relative terms) for green taxes. This suggests that close coordination across EU countries would be beneficial, especially in the case of green taxation. Another key result from our research is that the flexibility of the labour market has important effects on the level of distortion: more flexible labour markets are associated with lower distortions. As one would expect, the effect is more pronounced for labour taxes, though there is also some effect for green taxes. The implication is that were EU countries to undertake structural reforms (especially in the labour market), this would help to minimise the efficiency losses from tax-driven fiscal consolidations. A final consideration, not addressed in the current paper, is the progressivity of the different tax types, which would be an interesting avenue to explore in future research.

Table 11

MCF of Labour Taxes: Alternative Tax Recycling

Country	Country-level MCF		EU-level MCF	
	Standard Closure Rule	Alternative closure rule (with 1 Added)	Standard Closure Rule	Alternative Closure Rule (with 1 Added)
Austria	1.82	1.39	1.91	1.49
Belgium	1.98	1.28	2.29	1.48
Bulgaria	1.56	1.32	1.59	1.37
Czech Republic	1.49	1.29	1.50	1.38
Germany	1.96	1.64	2.04	1.75
Denmark	2.31	1.41	2.56	1.52
Estonia	1.30	1.18	1.36	1.24
Greece	1.59	1.48	1.60	1.51
Spain	1.79	1.40	1.84	1.46
Finland	1.61	1.36	1.66	1.41
France	2.41	1.78	2.50	1.87
Hungary	1.53	1.31	1.58	1.40
Ireland	1.33	1.14	1.41	1.19
Italy	1.68	1.38	1.68	1.42
Lithuania	1.45	1.21	1.49	1.29
Latvia	1.42	1.25	1.49	1.31
Netherlands	1.57	1.15	1.69	1.29
Poland	1.63	1.37	1.63	1.43
Portugal	1.82	1.45	1.93	1.56
Romania	1.43	1.37	1.42	1.42
Sweden	2.06	1.41	2.15	1.48
Slovenia	1.66	1.37	1.78	1.48
Slovakia	2.19	1.34	2.22	1.43
United Kingdom	1.81	1.37	1.86	1.41
EU (GDP-weighted)	*1.90*	*1.48*	*1.97*	*1.56*
Simple average	*1.73*	*1.36*	*1.80*	*1.44*
Coefficient of variation	*17.4%*	*10.4%*	*19.0%*	*10.2%*

Table 12

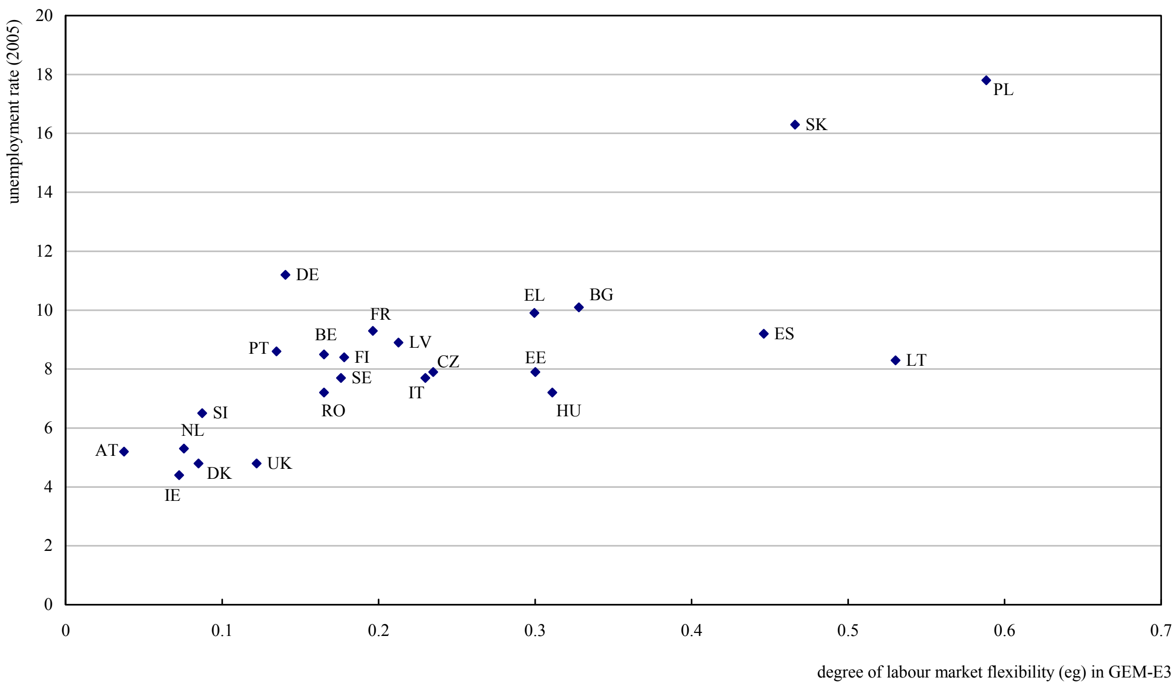
MCF of Energy Taxes: Alternative Tax Recycling

Country	Country-level MCF		EU-level MCF	
	Standard Closure Rule	Alternative closure rule (with 1 Added)	Standard Closure Rule	Alternative Closure Rule (with 1 Added)
Austria	0.87	0.70	1.07	0.85
Belgium	0.63	0.55	0.87	0.73
Bulgaria	0.62	0.58	0.64	0.66
Czech Republic	0.81	0.72	0.87	0.84
Germany	1.14	0.97	1.24	1.10
Denmark	0.86	0.86	0.93	0.90
Estonia	0.79	0.79	0.92	0.86
Greece	0.85	0.79	0.90	0.84
Spain	0.89	0.73	0.98	0.82
Finland	0.63	0.71	0.70	0.78
France	1.42	1.05	1.54	1.17
Hungary	0.86	0.75	1.01	0.88
Ireland	0.62	0.61	0.88	0.77
Italy	1.10	0.89	1.14	0.96
Lithuania	0.84	0.70	0.95	0.78
Latvia	0.82	0.74	0.84	0.74
Netherlands	0.83	0.65	0.97	0.80
Poland	1.26	1.01	1.27	1.10
Portugal	0.93	0.71	1.06	0.81
Romania	0.89	0.85	0.95	0.93
Sweden	0.87	0.77	0.95	0.82
Slovenia	0.95	0.83	1.10	0.94
Slovakia	1.06	0.58	1.17	0.68
United Kingdom	1.13	0.89	1.17	0.92
EU (GDP-weighted)	*1.08*	*0.88*	*1.17*	*0.97*
Simple average	*0.90*	*0.77*	*1.00*	*0.86*
Coefficient of variation	*22.2%*	*17.3%*	*19.0%*	*14.7%*

APPENDIX

Figure 3

Labour Market Flexibility in GEM-E3 and Actual Unemployment Rates, 2005



Sources. GEM-E3 calibration and Ameco (European Commission, DG ECFIN).

Table 13

MCF and Labour-market Flexibility: The Case of Labour Tax

EU Results			
	MCF, Benchmark Case	**Less Flexible Labour Market**	**More Flexible Labour Market**
EU	1.9	2.54	1.64
percent of change vs. benchmark		*33.60%*	*–13.60%*
Country Results			
Country	**MCF, Benchmark Case**	**Less Flexible Labour Market**	**More Flexible Labour Market**
Austria	1.82	2.41	1.6
Belgium	1.98	2.98	1.64
Bulgaria	1.56	1.51	1.6
Czech Republic	1.49	1.63	1.42
Germany	1.96	3.07	1.56
Denmark	2.31	4.85	1.75
Estonia	1.3	1.29	1.33
Greece	1.59	1.77	1.43
Spain	1.79	1.8	1.8
Finland	1.61	1.77	1.52
France	2.41	3.64	1.91
Hungary	1.53	1.7	1.43
Ireland	1.33	1.27	1.38
Italy	1.68	1.92	1.52
Lithuania	1.45	1.44	1.47
Latvia	1.42	1.44	1.41
Netherlands	1.57	2.43	1.31
Poland	1.63	1.78	1.53
Portugal	1.82	2.05	1.66
Romania	1.43	1.4	1.46
Sweden	2.06	2.57	1.79
Slovenia	1.66	1.84	1.55
Slovakia	2.19	2.3	2.13
United Kingdom	1.81	2	1.66

Table 14

MCF and Labour-market Flexibility: The Case of Green Taxes

EU Results			
	MCF, Benchmark Case	**Less Flexible Labour Market**	**More Flexible Labour Market**
EU	1.08	1.13	1.04
percent of change vs. benchmark		*4.60%*	*–3.30%*
Country Results			
	MCF, Benchmark Case	**Less Flexible Labour Market**	**More Flexible Labour Market**
Austria	0.87	0.88	0.87
Belgium	0.63	0.61	0.65
Bulgaria	0.62	0.61	0.64
Czech Republic	0.81	0.82	0.82
Germany	1.14	1.24	1.07
Denmark	0.86	0.87	0.88
Estonia	0.79	0.81	0.93
Greece	0.85	0.87	0.84
Spain	0.89	0.86	0.92
Finland	0.63	0.61	0.65
France	1.42	1.55	1.33
Hungary	0.86	0.87	0.85
Ireland	0.62	0.59	0.65
Italy	1.1	1.13	1.07
Lithuania	0.84	0.87	0.88
Latvia	0.82	0.83	1.02
Netherlands	0.83	0.85	0.82
Poland	1.26	1.29	1.23
Portugal	0.93	0.93	0.91
Romania	0.89	0.86	0.91
Sweden	0.87	0.88	0.84
Slovenia	0.95	0.96	0.94
Slovakia	1.06	1.06	1.06
United Kingdom	1.13	1.16	1.11

REFERENCES

Andersen, T.M. and A. Sørensen, (2012), "Globalization, Tax Distortions and Public Sector Retrenchment", *Scandinavian Journal of Economics*, Vol. 114, Issue 2, pp. 409-39.

Annabi, N. (2003), "Modeling Labor Markets in CGE Models Endogenous Labor Supply, Unions and Efficiency Wages", Centre Interuniversitaire sur le Risque, les Politiques Economiques et l'Emploi (CIPREE).

Auriol, E. and M. Warlters (2012), "The Marginal Cost of Public Funds and Tax Reform in Africa", *Journal of Development Economics*, Vol. 97, No. 1, pp. 58-72.

Blanchflower, D.G. and A.J. Oswald (1994), *The Wage Curve*, The MIT Press, Cambridge (MA).

Boeters, S. and L. Savard (2011), "The Labour Market in CGE Models", forthcoming chapter of P. Dixon and D. Jorgenson (eds.), *The Handbook of CGE Modelling*.

Campbell, C. and J.M. Orszag (1998), "A Model of the Wage Curve", *Economic Letters*, Vol. 59, Issue 1, pp. 119-25.

Cnossen, S. (2003), "How Much Tax Coordination in the European Union?", *International Tax and Public Finances*, No. 10, pp. 625-49.

Coenen, G., C. Erceg, C. Freedman, D. Furceri, M. Kumhof, R. Lalonde, D. Laxton, J. Linde, A. Mourougane, D. Muir, S. Mursula, C. de Resende, J. Roberts, W. Roeger, S. Snudden, M. Trabandt and J. in't Veld (2012), "Effects of Fiscal Stimulus in Structural Models", *American Economic Journal: Macroeconomics*, Vol. 4, No. 1, pp. 22-68.

Corsetti, G., K. Küster, A. Meier and G.J. Müller (2010), "Debt Consolidation and Fiscal Stabilization of Deep Recessions", *American Economic Review*, Vol. 100, No. 2, pp. 41-45.

Dahlby, B. (2008), *The Marginal Cost of Public Funds. Theory and applications*, The MIT Press, Cambridge (MA).

Dahlby, B. and E. Ferede (2012), "The Effects of Tax Rate Changes on Tax Bases and the Marginal Cost of Public Funds for Canadian Provincial Governments", *International Tax and Public Finance*, Vol. 19, pp. 844-83.

DeLong, J.B., and L.H. Summers (2012), "Fiscal Policy in a Depressed Economy", Brookings Papers on Economic Activity, Spring.

Devarajan, S. and S. Robinson (2002), "The Influence of Computable General Equilibrium Models on Policy", International Food Policy Research Institute, Trade and Macroeconomics Division, Discussion Paper, No. 98, Washington (D.C.).

Dixon, P., J. Honkatukia and M. Rimmer (2012), "The Marginal Cost of Funds from Different Taxes in Finland", GTAP Resource #3877, available at: https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=3877

Evers, M., R. De Mooij, D. Van Vuuren (2008), "The Wage Elasticity of Labour Supply: A Synthesis of Empirical Estimates", *De Economist*, Vol. 156, No. 1, pp. 25-43.

European Commission (2011), "Smarter Energy Taxation for the EU: Proposal for a Revision of the Energy Taxation Directive", COM(2011) 168/3.

——————(2012), "GEM-E3 Model Manual", Institute for Prospective Technological Studies, Joint Research Centre, Seville, available at: http://ipts.jrc.ec.europa.eu/activities/energy-and-transport/gem-e3/

Hansson, I. and C. Stuart (1985), "Tax Revenue and the Marginal Cost of Public Funds in Sweden", *Journal of Public Economics*, Vol. 27, No. 3, pp. 331-53.

IMF (2012), *World Economic Outlook, Coping with High Debt and Sluggish Growth*, International Monetary Fund, October, Washington (D.C.).

Lockwood, B. and A. Manning (1993), "Wage Setting and the Tax System: Theory and Evidence for the United Kingdom", *Journal of Public Economics*, No. 52, pp. 1-29.

Phelps, E.S. (1994), Low-wage Employment Subsidies Versus the Welfare State, *American Economic Review*, Vol. 30, No. 3, pp. 1476-90.

Ramsey, F.P. (1927), "A Contribution to the Theory of Taxation", *Economic Journal*, Vol. 37, No. 145, pp. 47-61.

Robichaud, V. (2001), "Calculating Equivalent and Compensating Variations in CGE Models," PEP Training Material, available at: http://www.pep-net.org/fileadmin/medias/pdf/ev3.pdf

Saveyn, B., D. Van Regemorter and J.C. Ciscar (2011), "Economic Analysis of the Climate Pledges of the Copenhagen Accord for the EU and Other Major Countries", *Energy Economics*, No. 33, pp. S33-S40.

Schöb, R. (1994), "On Marginal Cost and Marginal Benefits of Public Funds", *Public Finances /Finances Publiques*, Vol. 49, No. 1, pp. 87-106.

Shapiro, C. and J.E. Stiglitz (1987), "Equilibrium Unemployment as a Worker Discipline Device", *American Economic Review*, Vol. 74, No. 3, pp. 433-44.

Sørensen, P.B. (2001), "Tax Coordination in the European Union: What Are the Issues?", *Swedish Economic Policy Review*, No. 8, pp. 143-95.

Speck, S. (1999), "Energy and Carbon Taxes and Their Distributional Implications", *Energy Policy*, Vol. 27, No. 11, p. 659-67.

QUALITY OF GOVERNMENT AND LIVING STANDARDS: ADJUSTING FOR THE EFFICIENCY OF PUBLIC SPENDING

Francesco Grigoli[*] *and Eduardo Ley*[**]

*It is generally acknowledged that the government's output is difficult to define and its value is hard to measure. The practical solution, adopted by national accounts systems, is to equate output to input costs. However, several studies estimate significant inefficiencies in government activities (*i.e.*, same output could be achieved with less inputs), implying that inputs are not a good approximation for outputs. If taken seriously, the next logical step is to purge from GDP the fraction of government inputs that is wasted. As differences in the quality of the public sector have a direct impact on citizens' effective consumption of public and private goods and services, we must take them into account when computing a measure of living standards. We illustrate such a correction computing corrected per capita GDPs on the basis of two studies that estimate efficiency scores for several dimensions of government activities. We show that the correction could be significant, and rankings of living standards could be re-ordered as a result.*

1 Introduction

> "*Citizens, especially poor people, who ultimately consume the education and health services generated by the public system are the clients. They have a direct relationship with frontline service providers, such as teachers in public schools and health care workers in public health facilities – the short route of accountability. Crucially, however, the service providers generally have no direct accountability to the consumers, unlike in a market transaction. Instead, they are accountable only to the government that employs them. The accountability route from consumers to service providers is therefore through the government – the long route. To hold service providers accountable for the quantity and quality of services provided, citizens must act through the government a process that is difficult for poor people especially because they can seldom organize themselves and be heard by policy makers. Moreover, the government rarely has enough information or indeed the mechanisms to improve service provider performance*". Global Monitoring Report, *World Bank*, 2011; p. 74.

Despite its acknowledged shortcomings, GDP per capita is still the most commonly used summary indicator of living standards. Much of the policy advice provided by international organizations is based on macroeconomic magnitudes as shares of GDP, and framed on cross-country comparisons of per capita GDP. However, what GDP does actually measure may differ significantly across countries for several reasons. We focus here on a particular source for this heterogeneity: the quality of public spending. Broadly speaking, the "quality of public spending" refers to the government's effectiveness in transforming resources into socially valuable outputs.

---

[*] IMF, Fiscal Affairs Department. E-mail address: fgrigoli@imf.org

The views expressed in this presentation are those of the authors and should not be attributed to the IMF, the World Bank, their Executive Boards, or their management.

We thank Richard Allen, Benedict Clements, Jesus González-García, Steve Knack, Nick Manning, Pierre Pestieau, Mona Prasad, Baoping Shang and Vito Tanzi for their insightful comments. We also acknowledge the suggestions made by the participants in the 14th Bank of Italy Workshop on Public Finance and in the 12th Oxford CSAE Conference 2012: *Economic Development in Africa*.

[**] The World Bank. E-mail Address: eley@worldbank.org

The opening quote highlights the disconnect between spending and value when the discipline of market transactions is missing.

Everywhere around the world, non-market government accounts for a big share of GDP[1] and yet it is poorly measured – namely the value to users is assumed to equal the producer's cost. Such a framework is deficient because it does not allow for changes in the amount of output produced per unit of input, that is, changes in productivity (for a recent review of this issue, see Atkinson *et al.*, 2005). It also assumes that these inputs are fully used. To put it another way, standard national accounting assumes that government activities are on the best practice frontier. When this is not the case, there is an overstatement of national production. This, in turn, could result in misleading conclusions, particularly in cross-country comparisons, given that the size, scope, and performance of public sectors vary so widely.

Moreover, in the national accounts, this attributed non-market (government and non-profit sectors) "value added" is further allocated to the household sector as "actual consumption". As Deaton and Heston (2008) put it: "[...] there are many countries around the world where government-provided health and education is inefficient, sometimes involving mass absenteeism by teachers and health workers [...] so that such 'actual' consumption is anything but actual. To count the salaries of AWOL[2] government employees as 'actual' benefits to consumers adds statistical insult to original injury". This "statistical insult" logically follows from the United Nations System of National Accounts (SNA) framework once "waste" is classified as income – since national income must be either consumed or saved. Absent teachers and health care workers are all too common in many low-income countries (Chaudhury and Hammer, 2004; Kremer *et al.*, 2005; Chaudhury *et al.*, 2006; and World Bank, 2004). Beyond straight absenteeism, which is an extreme case, generally there are significant cross-country differences in the quality of public sector services. World Bank (2011) reports that in India, even though most children of primary-school age are enrolled in school, 35 per cent of them cannot read a simple paragraph and 41 per cent cannot do a simple subtraction.

It must be acknowledged, nonetheless, that for many of government's non-market services, the output is difficult to define, and without market prices the value of output is hard to measure. It is because of this that the practical solution adopted in the SNA is to equate output to input costs. This choice may be more adequate when using GDP to measure economic activity or factor employment than when using GDP to measure living standards.

Moving beyond this state of affairs, there are two alternative approaches. One is to try to find indicators for both output quantities and prices for direct measurement of some public outputs, as recommended in SNA 93 (but yet to be broadly implemented). The other is to correct the input costs to account for productive inefficiency, namely to purge from GDP the fraction of these inputs that is wasted. We focus here on the nature of this correction. As the differences in the quality of the public sector have a direct impact on citizens' effective consumption of public and private goods and services, it seems natural to take them into account when computing a measure of living standards.

To illustrate, in a recent study, Afonso *et al.* (2010) compute public sector efficiency scores for a group of countries and conclude that "[...] the highest-ranking country uses one-third of the inputs as the bottom ranking one to attain a certain public sector performance score. The average input scores suggest that countries could use around 45 per cent less resources to attain the same

1 Note that public expenditure (which includes transfers) is a different concept than the public sector's contribution to GDP (which excludes transfers). For instance, in France, in 2003, while the former amounted to 54 per cent of GDP, the latter was a smaller 16 per cent of GDP as social transfers (including pensions) are a substantial share of French public spending (see, e.g., Lequiller and Blades, 2006).

2 AWOL is an acronym meaning: "absent without official leave".

outcomes if they were fully efficient". In this paper, we take such a statement to its logical conclusion. Once we acknowledge that the same output could be achieved with less inputs, output value cannot be equated to input costs. In other words, *waste* should not belong in the living-standards indicator – it still remains a cost of government but it must be purged from the *value* of government services. As noted, this adjustment is especially relevant for cross-country comparisons.

The remainder of this paper is structured as follows. Section 2 discusses the measurement of living standards and the measurement of waste. Section 3 illustrates the empirical size this correction for waste would entail, and Section 4 concludes.

2 Measuring living standards

Per capita Gross National Income (GNI)[3] is the statistic that defines who is who in development rankings. The World Bank uses it to classify economies in groups. For a country to be eligible for international development assistance[4] (e.g., services which include grants and low-cost loans), it must satisfy two criteria, one of which is the relative poverty defined as GNI per capita below an established threshold that is updated annually. The cutoff for fiscal year 2011 is a 2009 GNI per capita of US$1,165. Likewise, to be eligible for International Bank for Reconstruction and Development (IBRD) lending, in 2011, a country must have a 2009 GNI per capita of between US$1,165 and US$6,885.[5]

While, under general circumstances, the GDP may be a suitable indicator for tracking economic activity for a given country over time,[6] its shortcomings in measuring economic welfare are well known. As it is often pointed out, GDP does not, for example, capture differences in leisure or in longevity; it does not reflect differences in inequality or in poverty; and it does not take into account the effect of economic activity on the environment. This has led to alternative attempts to enlarge the concept of GDP, one of the earliest being the "Measure of Economic Welfare" developed by Nordhaus and Tobin (1973). The recent *Report by the Commission on the Measurement of Economic Performance and Social Progress* prepared for the French government by Stiglitz *et al.* (2010) presents an insightful up-to-date summary of the issues.[7] Some of the report's main recommendations include (i) using net income- or consumption-based measures, together with wealth, rather than gross production-based aggregates, (ii) to broaden measures to non-market activities, and (iii) to consider a dashboard of indicators for the quality of life, environment, and sustainability. In addition, in the context of the public sector, many government activities (e.g., police, defense, sanitation services, road maintenance) are intermediate inputs[8] for production activities rather than genuine final outputs. Government services used by firms are called "instrumental expenditures" in Nordhaus and Tobin (1973). Similarly, in the private sphere, commuting to work would also be an "instrumental expenditure". These *instrumental expenditures* should be appropriately deducted from the aggregate measure of net income. Several government

---

3 Gross National Income (GNI) differs from Gross Domestic Product (GDP) by the net factor income of nationals (net primary income from rest of the world). Adding official transfers and remittances (net current transfers from the rest of the world) we obtain Gross National Disposable Income (GNDI). All the issues that we raise pertaining to the measurement of GDP apply to the measurement of GNI.

4 The International Development Association (IDA) is the part of the World Bank that helps the world's poorest countries. It currently provides the world's poorest 79 countries with interest-free loans and grants.

5 See http://data.worldbank.org/about/country-classifications.

6 Nonetheless, for new issues posed by the growth of services at the expense of manufacturing, see Abraham (2005).

7 See also Dasgupta (2001).

8 See Hicks and Hicks (1939) for a summary of the early debate on what ought to be included in the national income (which, at the time, was a considered a welfare concept rather than a production concept as in the SNA).

functions that provide public goods – e.g., justice and defense – are arguably better classified as instrumental expenditures rather than goods and services for final household consumption notwithstanding the importance of these several issues, we restrict ourselves here to the SNA framework where GDP is taken as a measure of production, not welfare. We also ignore the issue of netting out "instrumental expenditures" from output.

In this context, as noted, the standard practice is to equate the value of government outputs to its cost, notwithstanding the SNA 93 proposal to estimate government outputs directly. The value added that, say, public education contributes to GDP is based on the wage bill and other costs of providing education, such as outlays for utilities and school supplies.[9] Similarly for public health, the wage bill of doctors, nurses and other medical staff and medical supplies measures largely comprises its value added. Thus, in the (pre-93) SNA used almost everywhere, non-market output, by definition, equals total costs. Yet the same costs support widely different levels of public output, depending on the quality of the public sector.

Atkinson *et al.* (2005, p. 12) state some of the reasons behind current SNA practice: "Wide use of the convention that (output = input) reflects the difficulties in making alternative estimates. Simply stated, there are two major problems: (a) in the case of collective services such as defense or public administration, it is hard to identify the exact nature of the output, and (b) in the case of services supplied to individuals, such as health or education, it is hard to place a value on these services, as there is no market transaction".

Murray (2010) also observes that studies of the government's *production* activities, and their implications for the measurement of living standards, have long been ignored. He writes: "Looking back it is depressing that progress in understanding the production of public services has been so slow. In the market sector there is a long tradition of studying production functions, demand for inputs, average and marginal cost functions, elasticities of supply, productivity, and technical progress. The non-market sector has gone largely unnoticed. In part this can be explained by general difficulties in measuring the output of services, whether public or private. But in part it must be explained by a completely different perspective on public and private services. Resource use for the production of public services has not been regarded as inputs into a production process, but as an end in itself, in the form of public consumption. Consequently, *the production activity in the government sector has not been recognized*" (our italics.)

The simple point that we make in this paper is that once it is recognized that the effectiveness of the government's "production function" varies significantly across countries, the simple convention of equating output value to input cost must be revisited. Thus, if we learn that the same output could be achieved with less inputs, it is more appropriate to credit GDP or GNI with the *required* inputs rather than with the *actual* inputs that include waste.[10] While perceptions of government effectiveness vary widely among countries as, e.g., the World Bank's Governance indicators attests (Kaufmann *et al.*, 2009), getting reliable measures of government actual effectiveness is a challenging task as we shall discuss below.

In physics, *efficiency* is defined as the ratio of useful work done to total energy expended, and the same general idea is associated with the term when discussing production. Economists simply replace "useful work" by "outputs" and "energy" by "inputs". Technical efficiency means the adequate use of the available resources in order to obtain the maximum product. Why focus on

---

9 Note that value added is defined as payments to factors (labor and capital) and profits. Profits are assumed to be zero in the non-commercial public sector. As for the return to capital, in the current SNA used by most countries, public capital is attributed a *net* return of zero – *i.e.*, the return from public capital is equated to its depreciation rate. This lack of a net return measure in the SNA is not due to a belief that the net return is actually zero, but to the difficulties of estimating the return.

10 Among others, Prichett (2000), and Keefer and Knack (2007) have called attention to the quality of public investment where spending often may not translate into genuine asset-building. See also Tanzi and Davoodi (1997) and Gupta *et al.* (2011).

technical efficiency and not other concepts of efficiency, such as price or allocative efficiency? Do we have enough evidence on public sector inefficiency to make the appropriate corrections?

The reason why we focus on technical efficiency in this preliminary inquiry is twofold. First, it corresponds to the concept of waste. Productive inefficiency implies that some inputs are wasted as more could have been produced with available inputs.[11] In the case of allocative inefficiency, there could be a different allocation of resources that would make everyone better off but we cannot say that necessarily some resources are unused – although they are certainly not aligned with social preferences. Second, measuring technical inefficiency is easier and less controversial than measuring allocative inefficiency. To measure technical inefficiency, there are parametric and non-parametric methods allowing for construction of a best practice frontier. Inefficiency is then measured by the distance between this frontier and the actual input-output combination being assessed.[12]

Indicators (or rather ranges of indicators) of inefficiency exist for the overall public sector and for specific activities such as education, healthcare, transportation, and other sectors. However, they are far from being uncontroversial. Sources of controversy include: omission of inputs and/or outputs, temporal lags needed to observe variations in the output indicators, choice of measures of outputs, and mixing outputs with outcomes. For example, many social and macroeconomic indicators impact health status beyond government spending (Spinks and Hollingsworth, 2009, and Joumard *et al.*, 2010) and they should be taken into account. Most of the output indicators available show autocorrelation and changes in inputs typically take time to materialize into outputs' variations. Also, there is a trend towards using outcome rather than output indicators for measuring the performance of the public sector. In health and education, efficiency studies have moved away from outputs (e.g., number of pre-natal interventions) to outcomes (e.g., infant mortality rates). When cross-country analyses are involved, however, it must be acknowledged that differences in outcomes are explained not only by differences in public sector outputs but also differences in other environmental factors outside the public sector (e.g., culture, nutrition habits).

Empirical efficiency measurement methods first construct a reference technology based on observed input-output combinations, using econometric or linear programming methods. Next, they assess the distance of actual input-output combinations from the best-practice frontier. These distances, properly scaled, are called *efficiency measures* or scores. An input-based efficiency measure informs us on the extent it is possible to reduce the amount of the inputs without reducing the level of output. Thus, an efficiency score, say, of 0.8 means that using best practices observed elsewhere, 80 per cent of the inputs would suffice to produce the same output.

We base our corrections to GDP on the efficiency scores estimated in two papers: Afonso *et al.* (2010) for several indicators referred to a set of 24 countries, and Evans *et al.* (2000) focusing on health, for 191 countries based on WHO data. These studies employ techniques similar to those used in other studies, such as Gupta and Verhoeven (2001), Clements (2002), Carcillo *et al.* (2007), and Joumard *et al.* (2010).

- Afonso *et al.* (2010) compute public sector performance and efficiency indicators (as performance weighted by the relevant expenditure needed to achieve it) for 24 EU and emerging economies. Using DEA, they conclude that on average countries could use 45 per cent less resources to attain the same outcomes, and deliver an additional third of the fully efficient

[11] A related concept is "productive public spending" (see IMF, 1995), however this deals with the contribution of spending to capital formation, accumulation and its depreciation.

[12] While technical efficiency focuses on "doing things right", allocative efficiency focuses on the harder question of "doing the right things".

output if they were on the efficiency frontier. The study included an analysis of the efficiency of education and health spending that we use here.

- Evans *et al.* (2000) estimate health efficiency scores for the 1993-97 period for 191 countries, based on WHO data, using stochastic frontier methods. Two health outcomes measures are identified: the disability adjusted life expectancy (DALE) and a composite index of DALE, dispersion of child survival rate, responsiveness of the health care system, inequities in responsiveness, and fairness of financial contribution. The input measures are health expenditure and years of schooling with the addition of country fixed effects. Because of its large country coverage, this study is useful for illustrating the impact of the type of correction that we are discussing here.

We must note that ideally, we would like to base our corrections on input-based technical-efficiency studies that deal exclusively with inputs and outputs, and do not bring outcomes into the analysis. The reason is that public sector outputs interact with other factors to produce outcomes, and here cross-country hetereogenity can play an important role driving cross-country differences in outcomes. Unfortunately, we have found no technical-efficiency studies covering a broad sample of countries that restrict themselves to input-output analysis. In particular, these two studies deal with a mix of outputs and outcomes. The results reported here should thus be seen as illustrative. Furthermore, it should be underscored that the level of "waste" that is identified for each particular country varies significantly across studies, which implies that any associated measures of GDP adjusting for this waste will also differ.

3 Corrected GDP

Let y_i be country i's per capita GDP (or GNI):

$$y_i = g_i + x_i$$

where g_i is the government's value added (*i.e.*, its contribution to national income), and x_i is the contribution of the non-government sector. If country i had an overall efficiency score of ε_i for the public sector, then the corrected per-capita GDP is given by:

$$\widetilde{y}_i = \varepsilon_i g_i + x_i$$

Arguably, $\widetilde{y}_i$ is a better measure of living standards, as it removes the waste, $(1-\varepsilon)g_i$, from y_i – and, consequently, from household consumption. Note that this correction is not needed for the private x_i as its value is assessed directly by the consumers in their market transactions.

This correction may be carried out in a more disaggregated way when efficiency scores for different government functions are available. For illustrative purposes, we shall first use the efficiency scores estimated in Afonso *et al.* (2010), rescaled to lie in [0,1]. In their paper, they estimate public sector efficiency indicators for different categories – *i.e.*, administration, human capital, health, distribution, stability, and economic performance. We focus here on the ones corresponding to the functional categories of health and education.

Let ε_i^h and $\varepsilon_i^\varepsilon$ be the corresponding (rescaled) efficiency scores, and let H_i and E_i be country i's public expenditure in health and education (h_i and e_i as percentages of GDP). If the fraction $\omega_i^h = (1-\varepsilon_i^h)$ of resources is wasted, then:

$$\widetilde{H}_i = (Health\ Expenditures - Waste = (1-\omega_i^h)H_i$$

is the corrected estimate of the contribution of public health services to GDP. Similarly, with education we have public waste equal to $\omega_i^\varepsilon\ E_i$, and effective expenditures of $\widetilde{E}_i = (1-\omega_i^\varepsilon)E_i$.

Next we purge $\omega_i^h H_i$ and $\omega_i^\varepsilon E_i$ from GDP using the average (1998-2002) functional shares reported in Table 1 of Afonso *et al.* (2010).[13]

Table 1 shows the percentage-of-GDP losses due to public waste in education and health – *i.e.*, $\omega^\varepsilon e_i$ and $\omega^h h_i$. Overall, the size of the correction is quite remarkable; the average loss amounts to 4.1 percentage points of GDP, while averages for education and health are 1.5 and 2.6. Given an average spending of 4.6 per cent of GDP on education and 4.0 per cent of GDP on health, this means that 32.6 and 65.0 per cent of the inputs are wasted in the respective sectors. Note that the best-practice frontier that is used as reference to compute the efficiency scores is constructed on the basis of this set of 24 countries. Increasing the reference group to a larger set of countries can only make these efficiency scores worse, as the reference technology becomes richer.

Figure 1 plots the GDP losses against the corresponding per capita GDPs. For this set of countries, there is no strong discernible pattern, as the points scatter rather uniformly over the plot area. Perhaps it could be argued that the range of correction sizes increases with the level of income – the lower envelope of the scatter slopes negatively while the upper envelope slopes positively.

Another matter of interest is whether the per-capita-GDP ranking is altered at all due to the correction (*i.e.*, whether any country changes relative position). This re-ordering happens in 9 occasions out of the 24 countries. In the scatter plot (Figure 1), the candidates are pairs of countries where one is almost vertically on top of each other, but slightly to the right, and where the vertical (correction) distance is substantial. For example, Korea overtakes Cyprus; Cyprus, in turn, almost catches up with Greece, Brazil overtakes Lithuania, and Poland overtakes Estonia.

We turn now to the WHO study by Evans *et al.* (2000) covering health in both advanced and developing economies. The average GDP loss is 0.9 percentage points (the median is 0.8 per cent of GDP). This is lower than the estimate in Table 1 for health, reflecting the lower level of health spending in the wider country dataset used in the WHO study. The losses are uniformly distributed over the per-capita-GDP range. Baldacci *et al.* (2008) find that in countries suffering from poor governance, the positive effects of increased spending on education is reduced, and those of higher health spending can be completely negated. Rajkumar and Swaroop (2008) also show that, in a context of low quality of governance, increased expenditures in health and education are not reflected in improved social outcomes. Given the high correlation between income and governance, poorer countries tend to have more ineffective governments. At the same time, they tend to spend less on health. The combined effect is a broadly uniform distribution of waste, as Figure 2 shows.

While we recognize that inefficiency scores are sector-specific, we perform a "virtual experiment" by asking what would be the implications if these inefficiencies applied, on average, throughout all public-sector activities. What would be the extent of the "missing" GDP? Figure 3 shows the distribution of the correction vs. per capita GDP and technical efficiency scores. Technical efficiency is positively correlated with per capita GDP. As before, the correction is roughly uniformly distributed across the range of per capita GDP. The effects of lower efficiency scores and lower spending broadly compensate for each other. Thus, poorer countries with more ineffective government also spend a smaller share of GDP in public services, so any correction of the sort discussed here is going to be small. The scatter of technical efficiency vs. total waste displays an upper envelope: the estimated waste is bounded by the efficiency score.

[13] Note that the percent correction is a linear operation and, thus, can be applied either to components and ratios. If, e.g., we are considering per-capita GDP, then $\widetilde{y}_i = \frac{\widetilde{Y}_i}{N_i} = \frac{\varepsilon_i Y_i}{N_i} = \varepsilon_i y_i$.

Table 1

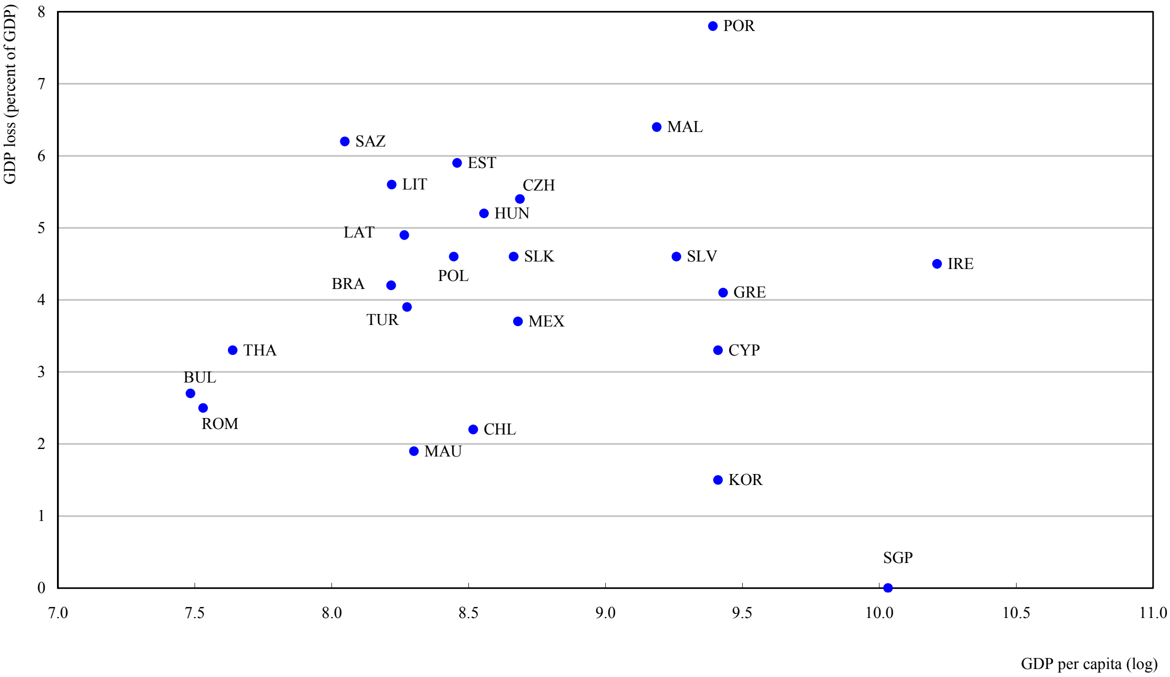
GDP Losses Associated with Wasted Public Resources
(averages 1998-2002, percent of GDP)

Country	Education	Health	Sum
Brazil	2.2	2.0	4.2
Bulgaria	0.1	2.6	2.7
Chile	1.2	1.0	2.2
Cyprus	2.2	1.1	3.3
Czech Republic	0.6	4.8	5.4
Estonia	2.8	3.0	5.9
Greece	0.5	3.6	4.1
Hungary	1.3	3.9	5.2
Ireland	1.0	3.5	4.5
Korea, Rep.	0.5	1.0	1.5
Latvia	2.8	2.1	4.9
Lithuania	2.5	3.1	5.6
Malta	1.7	4.7	6.4
Mauritius	1.2	0.7	1.9
Mexico	2.4	1.2	3.7
Poland	1.8	2.8	4.6
Portugal	3.1	4.8	7.8
Romania	0.0	2.5	2.5
Singapore	0.0	0.0	0.0
Slovak Republic	0.8	3.8	4.6
Slovenia	0.0	4.6	4.6
South Africa	3.7	2.5	6.2
Thailand	2.3	1.0	3.3
Turkey	1.2	2.6	3.9
Average	1.5	2.6	4.1

Source: Authors' calculations based on efficiency scores in Afonso *et al.* (2010).

Figure 1

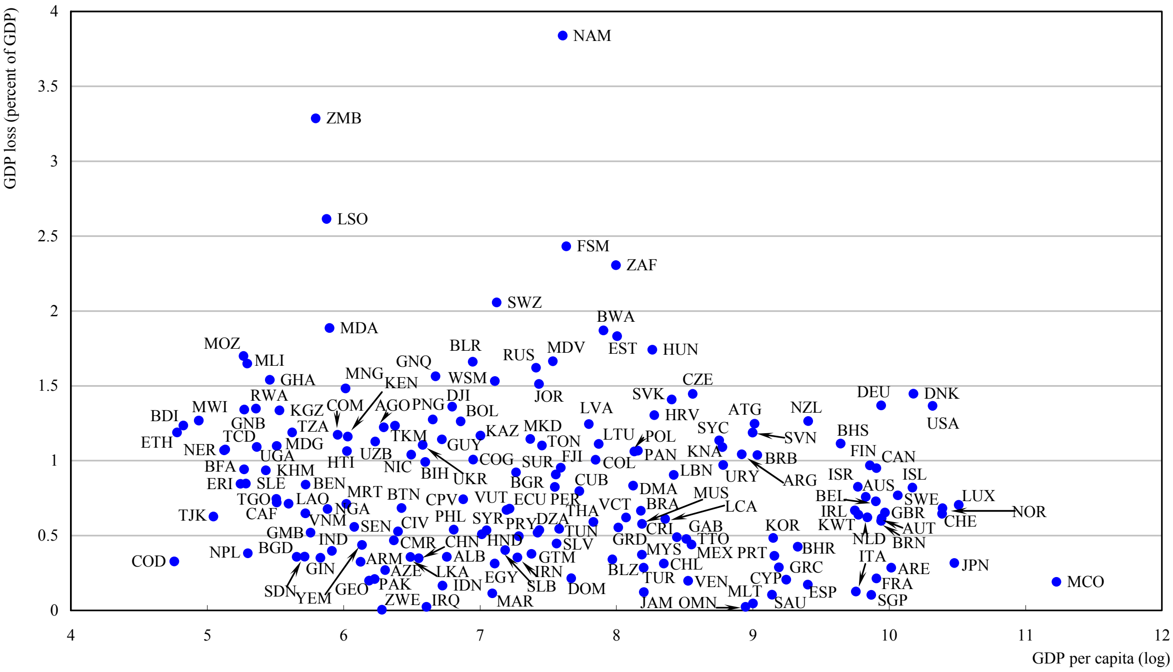
GDP Loss Due to Health and Education Waste *vs.* Per Capita GDP



Source: Authors' calculations based on efficiency scores in Afonso *et al.* (2010).

Figure 2

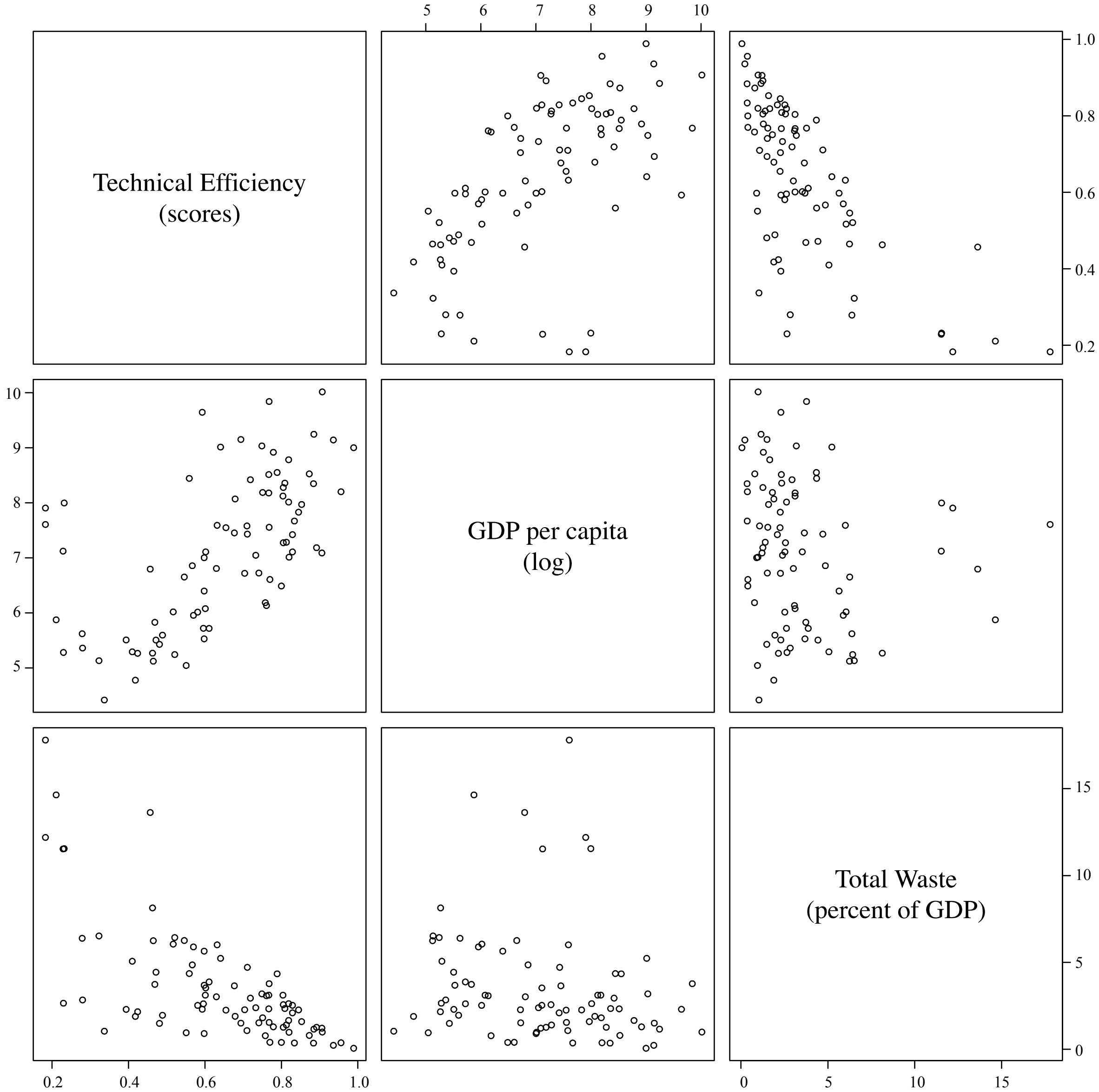
GDP Loss Due to Health Waste *vs.* Per Capita GDP



Source: Authors' calculations based on efficiency scores in Evans *et al.* (2000).

Figure 3

Technical Efficiency Scores, per Capita GDP, and Total Loss



Source: Authors' calculations based on efficiency scores in Evans *et al.* (2000).

Finally, we turn our attention to the country rankings of living standards, the GNI per capita computed using the World Bank's Atlas methodology.[14] As noted, this is the measure that the World Bank uses for classifying countries in income groups, as well as to set lending eligibilities.

---

[14] The Atlas method converts countries GNI in US dollars applying the Atlas conversion factor. This consists of a three-year average of exchange rates to smooth effects of transitory exchange rate fluctuations, adjusted for the difference between the rate of inflation in the country and that in a number of developed countries. For more details see: http://data.worldbank.org/indicator/NY.GNP.PCAP.CD.

What is the effect on the ranking of the corrections that we are discussing here? Let us consider the correction based on the health efficiency scores of Evans *et al.* (2000) applied to the value added of public administration and defense for the 2009 GNI. The result is a re-ordered country ranking where 51 countries out of 93 change their relative positions. Since the value added variable is available only for non-developed countries, we perform the same correction on the wage bill – to cover a larger set of countries. The portion of reordered countries is still higher than 50 per cent, as 59 of 116 countries are repositioned. In both corrections, about 70 per cent of the reordering happens in the lower half of the original ranking and the average shift is approximately equal to two positions.

How does this relate to governance indicators? There are several governance indicators available, all of which are highly correlated. The broadest coverage set is probably the Worldwide Governance Indicators (WGI) by Kaufmann *et al.* (2009). This database draws together information on perceptions of governance from a wide variety of sources, and organizes them into six clusters corresponding to the six broad dimensions of governance. These are voice and accountability, political stability and absence of violence/terrorism, government effectiveness, regulatory quality, rule of law, and control of corruption. Other very important sources of governance indicators are Freedom House and Transparency International.

The indicator "Government Effectiveness" attempts to capture perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies.[15] Figure 4 plots the "government effectiveness" WGI against technical efficiency scores, GDP loss due to health waste, and per capita GDP. The WGI is positively correlated with GDP per capita, and, as a result, with the efficiency scores. Its relationship with estimated waste is less clear-cut. The biggest waste is associated with intermediate values of the government effectiveness indicator. Waste is biggest in inefficient countries that spend significant resources on health. Otherwise, waste is limited in inefficient countries that do not allocate significant resources to health spending.

4 Concluding remarks

We have argued here that the current practice of estimating the value of the government's non-market output by its input costs is not only unsatisfactory but also misleading in cross-country comparisons of living standards. Since differences in the quality of the public sector have an impact on the population's effective consumption and welfare, they must be taken into account in comparisons of living standards. We have performed illustrative corrections of the input costs to account for productive inefficiency, thus purging from GDP the fraction of these inputs that is wasted.

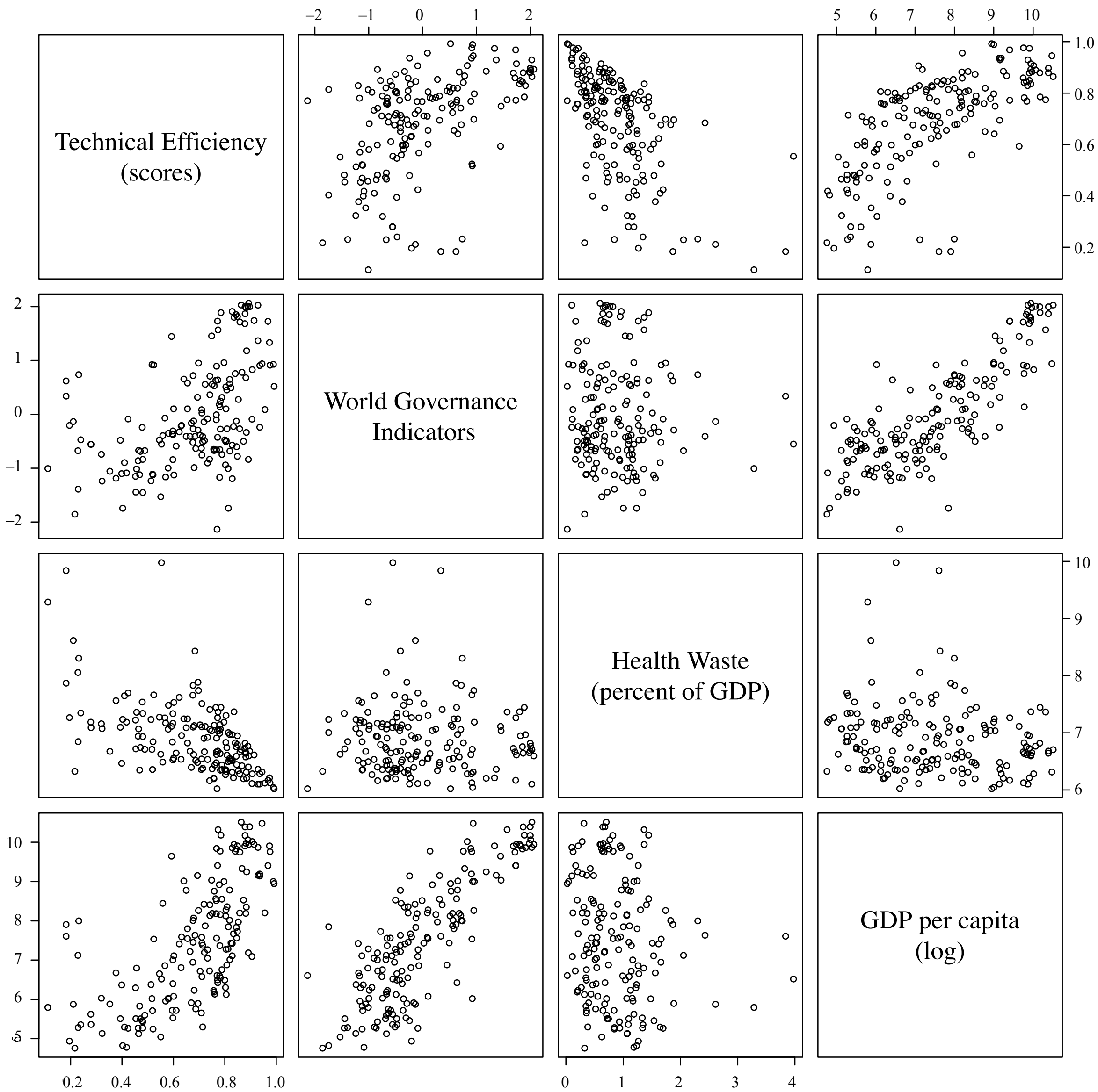
Our results suggest that the magnitude of the correction could be significant. When correcting for inefficiencies in the health and education sectors, the average loss for a set of 24 EU member states and emerging economies amounts to 4.1 percentage points of GDP. Sector-specific averages for education and health are 1.5 and 2.6 percentage points of GDP, implying that 32.6 and 65.0 per cent of the inputs are wasted in the respective sectors. These corrections are reflected in the GDP-per-capita ranking, which gets reshuffled in 9 cases out of 24. In a hypothetical scenario where the inefficiency of the health sector is assumed to be representative of the public sector as a whole, the rank reordering would affect about 50 per cent of the 93 countries in the sample, with 70 per cent of it happening in the lower half of the original ranking. These results, however, should

---

[15] See Kaufmann *et al.* (2010) for details on methodology, data sources, and interpretation of the indicators.

Figure 4

Technical Efficiency Scores, WGI's Government Effectiveness, GDP Loss Due to Health Waste, and Per Capita GDP



Source: Authors' calculations based on efficiency scores in Evans *et al.* (2000).

be interpreted with caution, as the purpose of this paper is to call attention to the issue, rather than to provide fine-tuned waste estimates.

A natural way forward involves finding indicators for both output quantities and prices for direct measurement of some public outputs. This is recommended in SNA 93 but has yet to be implemented in most countries. Moreover, in recent times there has been an increased interest in outcomes-based performance monitoring and evaluation of government activities (see Stiglitz *et al.*, 2010). As argued also in Atkinson (2005), it will be important to measure not only public sector outputs but also outcomes, as the latter are what ultimately affect welfare. A step in this direction is suggested by Abraham and Mackie (2006) for the US, with the creation of "satellite" accounts in specific areas as education and health. These extend the accounting of the nation's productive inputs and outputs, thereby taking into account specific aspects of non-market activities.

REFERENCES

Abraham, K.G. and C. Mackie (2006), "A Framework for Non-market Accounting", *A New Architecture for the US National Accounts*, University of Chicago Press.

Afonso, A., L. Schuknecht and V. Tanzi (2010), "Public Sector Efficiency: Evidence for New EU Member States and Emerging Markets", *Applied Economics*, Vol. 42, No. 17, pp. 2147-64.

Atkinson, A.B. (2005), "Measurement of Government Output and Productivity for the National Accounts", Palgrave MacMillan, Basingstoke.

Baldacci, E., B. Clements, S. Gupta and Q. Cui (2008), "Social Spending, Human Capital, and Growth in Developing Countries", *World Development*, Vol. 36, No. 8, pp. 1317-41.

Chaudhury, N. and J.S. Hammer (2004), "Ghost Doctors: Absenteeism in Rural Bangladeshi Health Facilities", *World Bank Economic Review*, Oxford University Press, Vol. 18, No. 3, pp. 423-41.

———, M. Kremer, K. Muralidharan and F. Halsey Rogers (2006), "Missing in Action: Teacher and Health Worker Absence in Developing Countries", *Journal of Economic Perspectives*, Vol. 20, No. 1, pp. 91-116.

Carcillo, S., V. Gunnarsson and M. Verhoeven (2007), "Education and Health in G7 Countries: Achieving Better Outcomes with Less Spending", IMF, Working Paper, No. 07/263, Washington (D.C.).

Clements, B. (2002), "How Efficient is Education Spending in Europe?", *European Review of Economics and Finance*, Vol. 1, No. 1, pp. 3-26.

Dasgupta, P.S. (2001), *Human Well-being and the Natural Environment*, Oxford University Press.

Deaton, A. and A. Heston (2008), "Understanding PPPs and PPP-based National Accounts", NBER, Working Paper, No. 14499.

Evans, D.B., A. Tandon, C.L. Murray and J.A. Lauer (2000), "Comparative Efficiency of National Health Systems in Producing Health: An Analysis of 191 Countries", World Health Organization, GPE, Discussion Paper, No. 29, Geneva.

Gupta, S. and M. Verhoeven (2001), "The Efficiency of Government Expenditure Experiences from Africa", *Journal of Policy Modelling*, Vol. 23, pp. 433-67.

Gupta, S., A. Kangur, C. Papageorgiou and A. Wane (2011), "Efficiency-adjusted Public Capital and Growth", IMF, Working Paper, No. 11/217, Washington (D.C.).

Hicks, J.R. and U.K. Hicks (1939), "Public Finance in the National Income", *Review of Economic Studies*, Vol. 6, No. 2, pp. 147-55.

International Monetary Fund (1995), "Guidelines for Fiscal Adjustment", Fiscal Affairs Department, Pamphlet Series, No. 49, Washington (D.C.).

Joumard, I., C. André and C. Nicq (2010), "Health Care Systems: Efficiency and Institutions", OECD, Economics Department, Working Paper, No. 769.

Kaufmann, D., A. Kraay and M. Mastruzzi (2009), "Governance Matters VIII: Aggregate and Individual Governance Indicators for 1996-2008", The World Bank, Policy Research, Working Paper, No. 4978, Washington (D.C.).

——— (2010), "The Worldwide Governance Indicators: Methodology and Analytical Issues", The World Bank, Policy Research, Working Paper, No. 5430, Washington (D.C.).

Keefer, P. and S. Knack (2007), "Boondoggles, Rent-seeking, and Political Checks and Balances: Public Investment under Unaccountable Governments", *The Review of Economics and Statistics*, Vol. 89, No. 3, pp. 566-72.

Kremer, M., K. Muralidharan, N. Chaudhury, J. Hammer and F. Halsey Rogers (2005), "Teacher Absence in India: A Snapshot", *Journal of the European Economic Association*, Vol. 3, No. 2, pp. 658-67.

Lequiller, F. and D. Blades (2006), *Understanding National Accounts*, OECD, Paris.

Murray, R. (2010), "Measuring Output from the Public Sector: A Critical Examination of the Atkinson Review", *Review of Income and Wealth*, Vol. 56, No. 2, pp. 413-23.

Nordhaus, W.D. and J. Tobin (1971), "Is Growth Obsolete?", Cowles Foundation, Discussion Paper, No. 319, Yale University.

Pritchett, L. (2000), "The Tyranny of Concepts: CUDIE (Cumulated, Depreciated, Investment Effort) Is Not Capital", *Journal of Economic Growth*, Vol. 5, pp. 361-84.

Rajkumar, A.S. and V. Swaroop (2008), "Public Spending and Outcomes: Does Governance Matter?", *Journal of Development Economics*, Vol. 86, pp. 96-111.

Spinks, J. and B. Hollingsworth (2009), "Cross-country Comparisons of Technical Efficiency of Health Production: A Demonstration of Pitfalls", *Applied Economics*, Vol. 41, pp. 417-27.

Stiglitz, J.E., A. Sen and J.P. Fitoussi (2010), "Report of the Commission on the Measurement of Economic Performance and Social Progress", available at http://www.stiglitz-sen-fitoussi.fr/en/index.htm. (The first part was also published by The New Press as a stand-alone paperback in 2010 with the title *Mismeasuring Our Lives: Why GDP Doesn't Add Up*).

Tanzi, V. and H. Davoodi (1997), "Corruption, Public Investment and Growth", IMF, Working Paper, No. 97/139, International Monetary Fund.

World Bank (2004), "World Development Report 2004: Making Services Work for the Poor", Washington (D.C.).

——— (2011), *Global Monitoring Report 2011: Improving the Odds of Achieving the MDGs*, Chapter 3: "Linking Spending and Outcomes: Some Lessons from Impact Evaluations in Education and Health", Washington (D.C.).

INCOME TAXATION, TRANSFERS AND LABOUR SUPPLY AT THE EXTENSIVE MARGIN

Péter Benczúr[*], *Gábor Kátay*[**], *Áron Kiss*[***] *and Olivér M. Rácz*[****]

This paper estimates the effect of income taxation on labour supply at the extensive margin, i.e.*, the labour force participation. We extend existing structural form methodologies by considering the effect of both taxes and transfers. Non-labour income contains the (hypothetical) transfer amount someone gets when out of work, while the wage is replaced by the sum of net wages and the amount of lost transfers due to taking up a job (gains to work, GTW). Using data from the Hungarian Household Budget Survey (HKF), we find that participation probabilities are strongly influenced by transfers and the GTW, particularly for low-income groups and the elderly. Moreover, the same change in the net wage leads to a much larger change in the GTW for low earners, making them even more responsive to wages and taxation. Our parametric estimates can be readily utilized in welfare evaluations, or microsimulation analyses of tax and transfer reforms.*

1 Introduction

This paper presents a unified parametric approach to estimate the impact of taxes and transfers on the participation decision (the extensive margin of labour supply). In our framework, participation probabilities are determined by the comparison of disposable income in and out of the labour force, consisting of the (often non-observed) amount of transfers and non-labour income an individual gets if not working and the gains to work (GTW; change in disposable income if accepting a job offer, the sum of net wages and lost transfers). Identification is achieved by utilizing a multitude of tax and transfer reforms. Unlike in the existing literature, our results allow a general assessment of the efficiency and effectiveness of government interventions into the labour market, and more importantly, a micro-based prediction of the impact of tax and welfare reforms.

There is a multitude of existing studies which establish that taxes and the welfare system influence the participation decision. There is, however, a notable heterogeneity in terms of implied elasticity measures. Arrufat and Zabalza (1986) do a cross section estimation on the U.K. General Household Survey dataset, and find a participation elasticity (the change in the probability of being active in response of a unitary shock in net wages) of 1.41 for married women. Dickert *et al.* (1995), conducting a cross-section estimation on the Survey of Income and Program Participation (SIPP) to analyse a large expansion of the Earned Income Tax Credit (EITC) in the U.S., find an elasticity of $\eta = 0.2$ for single parents. Eissa and Liebman (1996) follow a program evaluation methodology (difference in differences) using the Current Population Survey to analyse the same episode of EITC expansion. They find that single mothers increased their participation rate by 2.8 percentage points relative to single women without children. Kimmel and Kniesner (1998) adopt a panel estimation on SIPP, and find elasticities of [0.6; 2.4; 1.8; 1.1] for single men, single women, wives and husbands respectively. Finally, Aaberge *et al.* (1999) follow a cross section

[*] Magyar Nemzeti Bank and Central European University. E-mail: benczurp@mnb.hu

[**] Magyar Nemzeti Bank. E-mail: katayg@mnb.hu

[***] Magyar Nemzeti Bank. E-mail: kissa@mnb.hu

[****] Magyar Nemzeti Bank. E-mail: raczo@mnb.hu

This version: November 2012. We thank Gábor Kézdi, Mihály Szoboszlai, Mirco Tonin, participants at MKE, Atiner 2011, the 2012 Budapest Workshop on Macroeconomic Policy Research, and at seminars at Magyar Nemzeti Bank, Ceska Narodni Banka, the Finance Ministry of France and Central European University for comments and suggestions. All remaining errors are ours. The views expressed here are of the authors and do not necessarily reflect the official view of MNB.

estimation based on the Survey on Household Income and Wealth (Italy), and obtain average elasticities for men and women as [0.04; 0.65] respectively.

From our point of view, these findings have important shortcomings. First, most of them focus on special subgroups and tend to follow a reduced form approach (program evaluation methodology, see Moffitt, 2002, for a review). Though such approaches are capable of precisely estimating the impact of a particular tax or transfer reform episode, they are not suitable for evaluating the impact of future (hypothetical) scenarios. There is also a substantial heterogeneity in the way after-tax wages are controlled for (if at all). Meyer and Rosenbaum (2001) is an example of a structural approach, but is not suitable for simulations either: wages are proxied, so the results do not imply a wage elasticity.

Second, the existing literature usually focuses on either taxes or transfers. Though the meta-analysis of Chetty *et al.* (2012) provides a "new consensus estimate" of extensive margin elasticities of 0.25, this result still does not necessarily control for the entire tax and transfer system. As argued by Blundell (2012), it is important to take taxes and transfers into account simultaneously and combine them into effective tax wedges. Besides influencing non-labour income (income at zero hours worked), transfers also show characteristics resembling both marginal and average tax rates. Suppose that a certain benefit is means tested with a gradual phaseout. For example, every extra income earned as wage reduces transfers by 20 per cent. In that case, it is equivalent to a 20 per cent extra marginal tax rate. Once the individual has lost all of this means tested benefit, lost transfers become similar to an average tax rate: the total amount of lost transfers decreases the payoff from work, just like the average tax rate does.

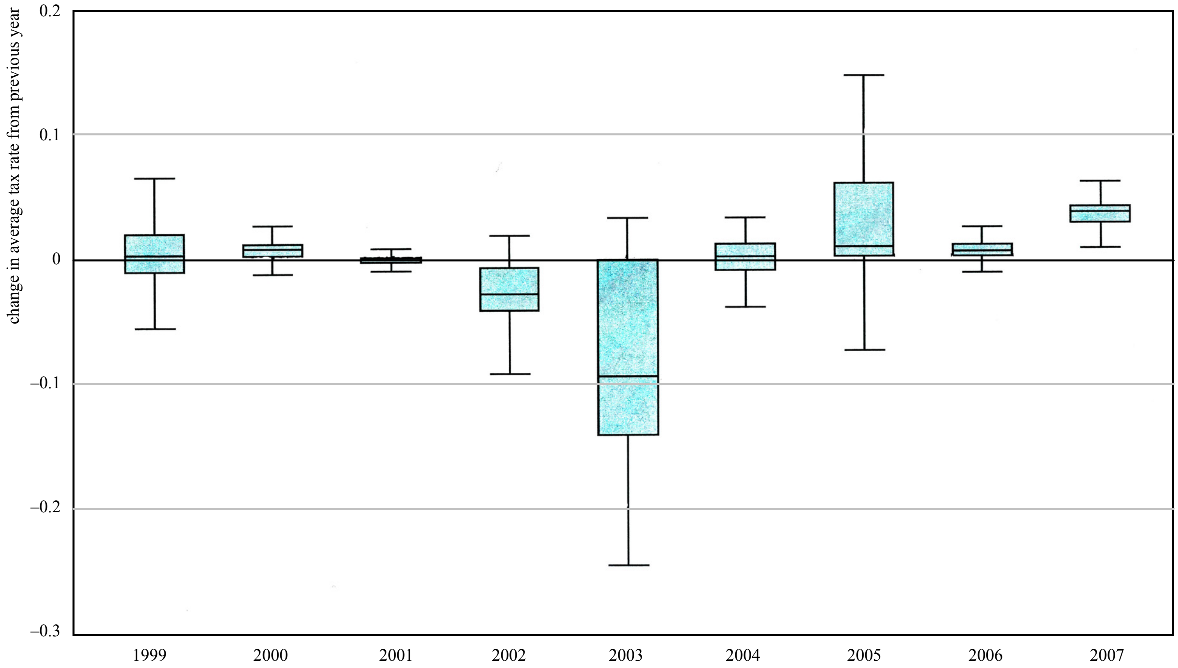
One major reason for the lack of structural studies is that it is not obvious how to incorporate all the relevant features of the tax and transfer system into a theory-based framework of labour supply. This paper presents an extension of the standard labour supply model that can incorporate both the marginal and participation tax rate aspect of transfers, but at the expense of constraining the participation decision to a fixed job size. Jobs usually have a fixed minimum size (half-time, or in some cases even full-time), which implies that an interior solution at a too low number of hours might also be practically infeasible. In that case, the labour supply choice of individuals is determined by the average tax rate at her initial gross monthly earnings and the total amount of transfers. The overall summary measure in this case is the gains to work, which consists of the net wage (for the fixed size of the job) minus the amount of lost transfers.

We carry out our estimation on the Hungarian Household Budget Survey (HKF), containing detailed income and consumption measures of individuals for the years 1998-2008. Numerous policy measures on both income tax rates and transfers adopted during this period provide enough cross-sectional and time variation for the estimation of the elasticity of participation probabilities with respect to gains to work. Figure1 show how individuals' average tax rates would have changed if their real income remained unchanged over time. It is seen that minor income tax changes occurred every year and major changed occurred in 1999 and between 2002 and 2005. The right graph show that tax changes affected lower income earners to a greater extent. As for the transfers, Figure 1 illustrates the impact of various transfer reforms on the Hungarian participation rate. The simple decomposition exercise of Kátay and Nobilis (2009) clearly demonstrates that transfer changes do impact the participation rate, providing us with sufficient exogenous variation in transfers to identify our specification.

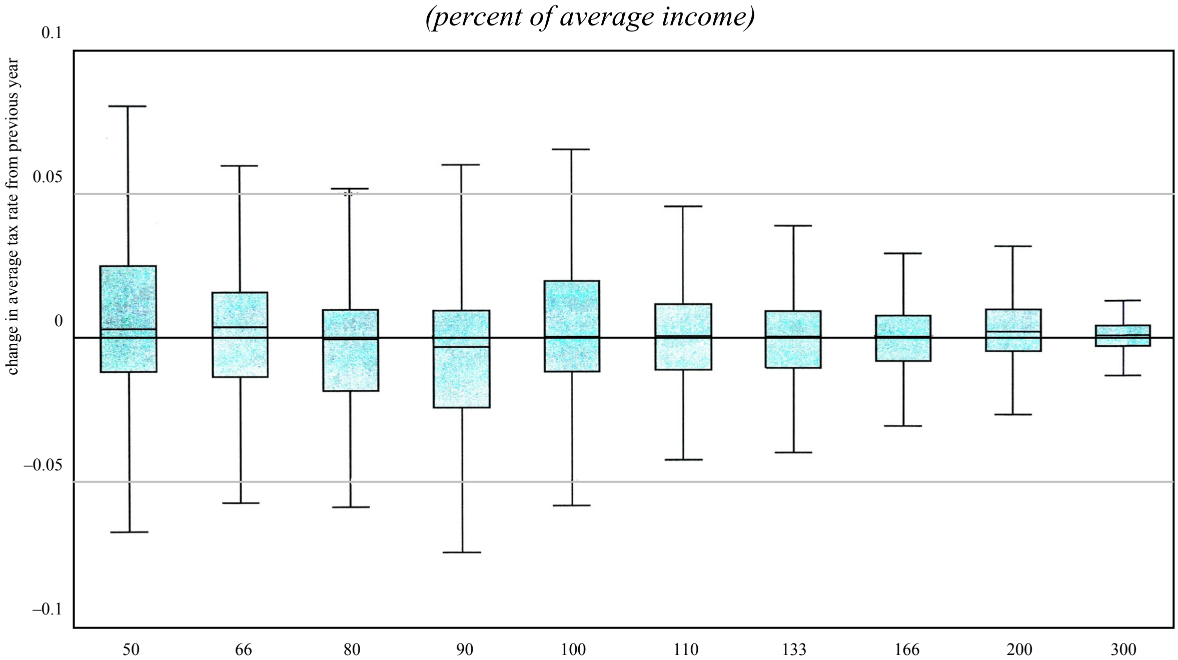
The underlying theory – presented in Section 2 – leads to a structural probit equation which relates participation probabilities to gains to work from a full time job, the total amount of non-labour income (including the hypothetical amount of transfers one gets or would get at zero hours worked) and other individual characteristics. The unobserved hypothetical amount of transfers are backed up using individual characteristics and the welfare system's details for every given year.

Figure 1

Variation in the Changes in Average Tax Rates (ATR)*
a) Yearly Changes in ATR



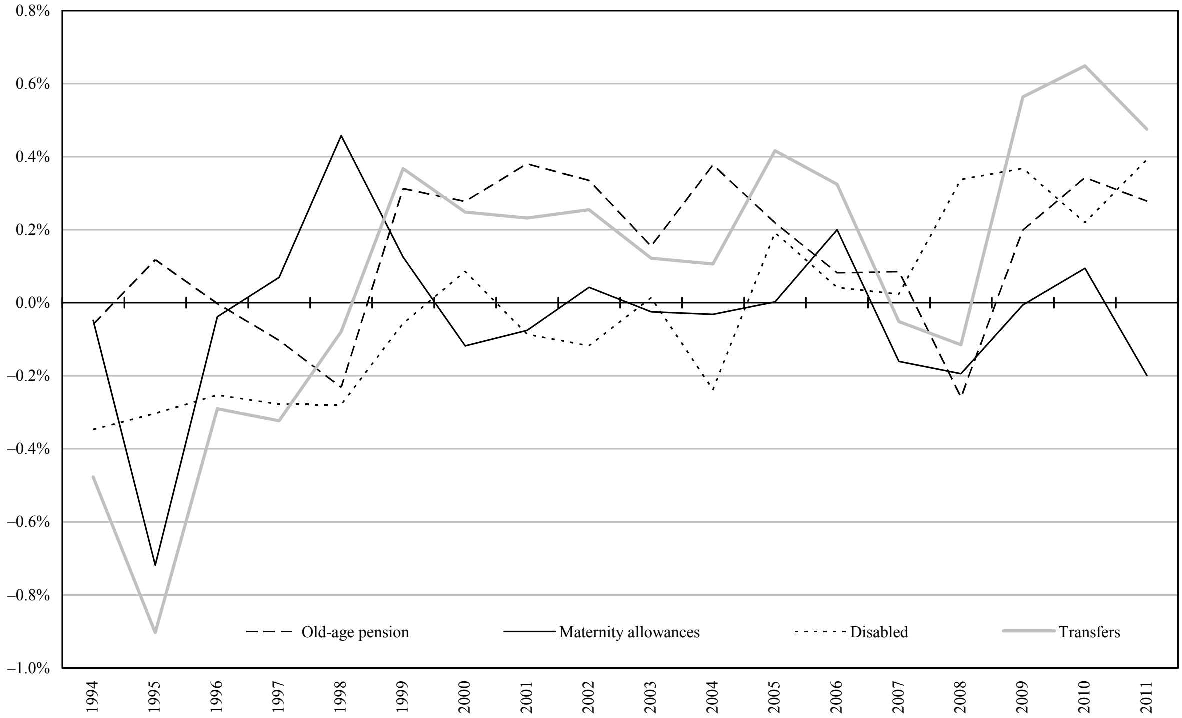
b) Changes in ATR by Income Categories
(percent of average income)



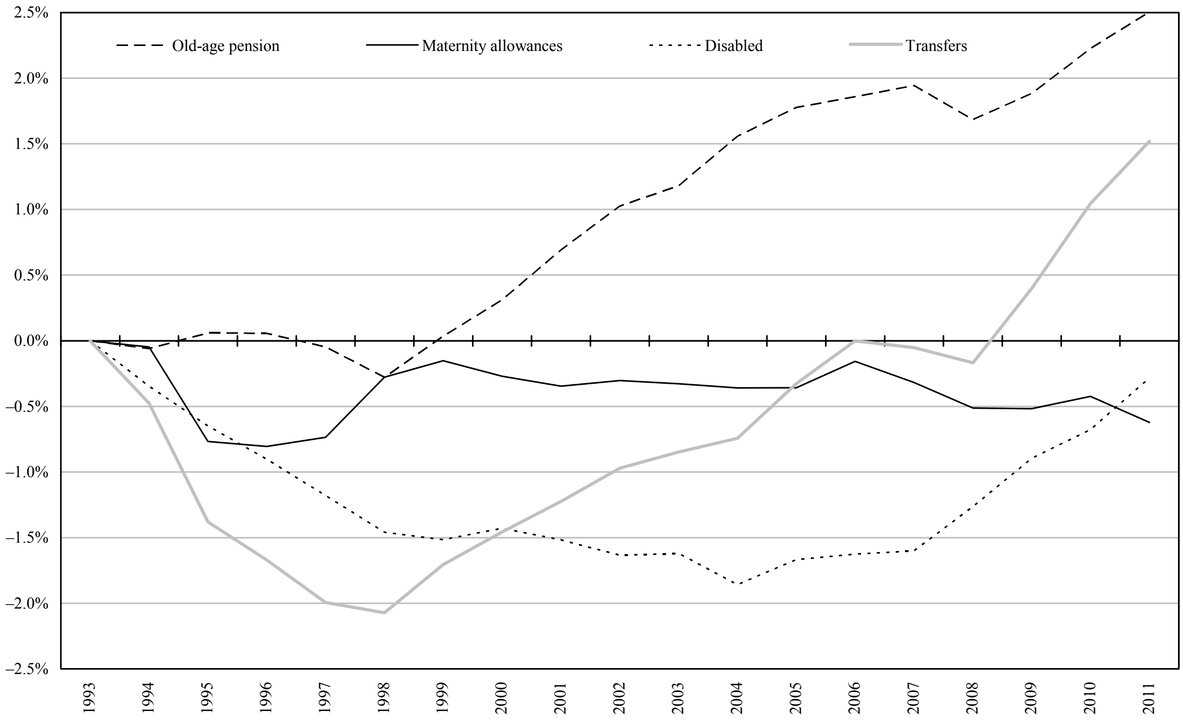
Graphs show the yearly changes in average tax rates between 1998 and 2008 for the individuals observed in 2008, assuming that their real income did not change during this period. Outside values are excluded in both graphs.
Source: Household Budget Surve and own calculations.

Figure 2

Decomposition of the Aggregate Participation Rate*
a) Year-to-year Percentage Point Changes



(b) Cumulative Changes



Source: Kátay and Nobilis (2009), updated.

The estimation process – described in Section 3 – follows the often used three step procedure, as, e.g., in Kimmel and Kniesner (1998). The key element of the identification is the careful choice of labour demand shifters, *i.e.*, the variables which have no (or negligible) impact on labour supply directly, but strongly impact the wage and hence impacts activity indirectly. In Section 4, we argue that county dummies and (once we control for individuals' lifecycle position with a large set of dummy variables) individuals' age are such variables.

Section 5 presents the estimation results. We find that a single equation can already explain a large heterogeneity of individual responsiveness to taxes and transfers: there are large differences among subgroups, driven partly by a composition effect, and partly by a different share of lost transfers in the GTW. The most responsive subgroups are low-skilled, (married) women at child-bearing age and elders, while prime-age higher educated individuals are practically unresponsive to tax and transfer changes at the extensive margin. As argued for example by Kátay (2009), Hungary's labour participation deficit compared to other EU members is mostly due to these special groups.

2 Theory

2.1 The underlying theory

The usual approach is to define the reservation wage, which is the threshold for accepting a job offer. Let us start from a standard utility maximization problem:

$$\begin{cases} \max \dfrac{c^{1-\psi}-1}{1-\psi} + \chi \dfrac{(1-l)^{1-\phi}-1}{1-\phi} \\ \text{s.t.}: c + w(1-l) = w + T, \end{cases}$$

where c is consumption, l is labour, w is the wage, and T denotes transfers and other non-labour income. The total time endowment is normalized to *l*, so leisure is $1-l$. The optimality condition can be written as:

$$\chi(1-l)^{-\phi} = wc^{-\psi}$$

The reservation wage corresponds to the case where $1-l^{*}=1$. Then $c=T$, so:

$$\chi = w_{res}T^{-\psi}$$

defines the reservation wage. The participation decision is then determined by $w \geq w_{res}$, or, in logs:

$$\log w \geq \log \chi + \psi \log T$$

Finally, we expand $\log \chi_i$ as $Z_i A' + \varepsilon_i$, where Z_i is a vector of observable individual characteristics and $\varepsilon_i \sim N(0, \sigma^2)$:

$$\log w_i - Z_i A' - \psi \log T_i \geq \varepsilon_i$$

The probability of someone working given a wage offer w_i, non-labour income T_i and individual characteristics Z_i is then:

$$P = \Phi\left(\frac{\log w_i - Z_i A' - \psi \log T_i}{\sigma}\right) = \Phi\left(\gamma \log w_i + Z_i \alpha' - \overline{\psi} \log T_i\right) \qquad (1)$$

yielding the standard structural probit specification.[1]

The next step is to add taxes and transfers. One the one hand, we have to modify the wage rate by the effective tax rate (marginal rate, at zero labour income), including taxes, social contributions, and the phaseout of social transfers (if applicable). On the other hand, there are certain transfers which get lost immediately at taking up any job. In such a case, there is a discrete downward jump in T for any nonzero hours worked. One could try to redefine the reservation wage similarly to before, as the level that could still induce an epsilon amount of work. This is, however, not feasible: from Roy's identity, the welfare gain from a marginal wage increase is the same as the income gain from the extra income due to the higher wage. But there is no such income gain at zero hours worked, so the income equivalent gain is zero, while there is a nonzero income loss due to the drop in T. In other words, the reservation wage is infinite (this can also be established formally by total differentiation).

Instead, we redefine the reservation wage by constraining the participation decision to a fixed "job size"l^* – in our empirical specification, it will be a full time job.[2] The reservation wage is thus set by the following comparison:

• Do not work: then $c = T, 1 - l = 1$, welfare is $\dfrac{T^{1-\psi} - 1}{1-\psi}$

• Work l^* : then $c = T - \Delta T + wl^*,\ 1 - l = 1 - l^*$, welfare is $\dfrac{\left(T - \Delta T + wl^*\right)^{1-\psi} - 1}{1-\psi} + \chi \dfrac{\left(1 - l^*\right)^{1-\phi} - 1}{1-\phi}$

Introducing the notation $W = wl^* - \Delta T$ (gains to work, GTW), the comparison becomes:

$$\frac{(T+W)^{1-\psi} - 1}{1-\psi} + \chi \frac{\left(1-l^*\right)^{1-\phi} - 1}{1-\phi} \geq \frac{T^{1-\psi} - 1}{1-\psi}$$

$$\frac{(T+W)^{1-\psi} - 1}{1-\psi} - \frac{T^{1-\psi} - 1}{1-\psi} \geq -\chi \frac{\left(1-l^*\right)^{1-\phi} - 1}{1-\phi} \tag{2}$$

One can also give a simple graphical representation (see Figure 2): draw the indifference curve going through $(C = T, l = 0)$, find the point of this curve where $l = l^*$, and connect this with point $(C = T - \Delta T, l = 0)$. Its slope is then the reservation wage: at such a wage level, the individual is just indifferent between not working and getting the full amount of transfers $(C = T, l = 0)$, or working l^* hours and getting only $T - \Delta T$ as transfers $(C = T - \Delta T + W, l = l^*)$.

To derive a formal expression for the probability of being active (the analogue of (1)), let us linearize the left hand side of (2):

$$\frac{(T+W)^{1-\psi} - 1}{1-\psi} - \frac{T^{1-\psi} - 1}{1-\psi} \approx WT^{-\psi},$$

1 One could repeat the same exercise using a growth-consistent utility function of the form $\dfrac{(\mathrm{c} \cdot \exp(\mathrm{f}(1-1)))^{1-\psi} - 1}{1-\psi}$. Assuming that $f(1-1) = \dfrac{(1-l)^{1-\phi} - 1}{1-\phi}$, we would get an almost identical probit equation, with an extra constraint of $\gamma = \overline{\psi}$.

2 Once working, an individual may decide to work more than l^*. We assume, however, that it is not known in advance whether there would be opportunities for overtime or performance bonuses, so the activity decision is determined by the base salary.

so the comparison becomes:

$$WT^{-\psi} \geq \chi \underbrace{\frac{1-\left(1-l^{*}\right)^{1-\phi}}{1-\phi}}_{Q} = \chi Q$$

The individual works if:

$$\log W - \psi \log T - \log \chi - \log Q \geq \varepsilon$$

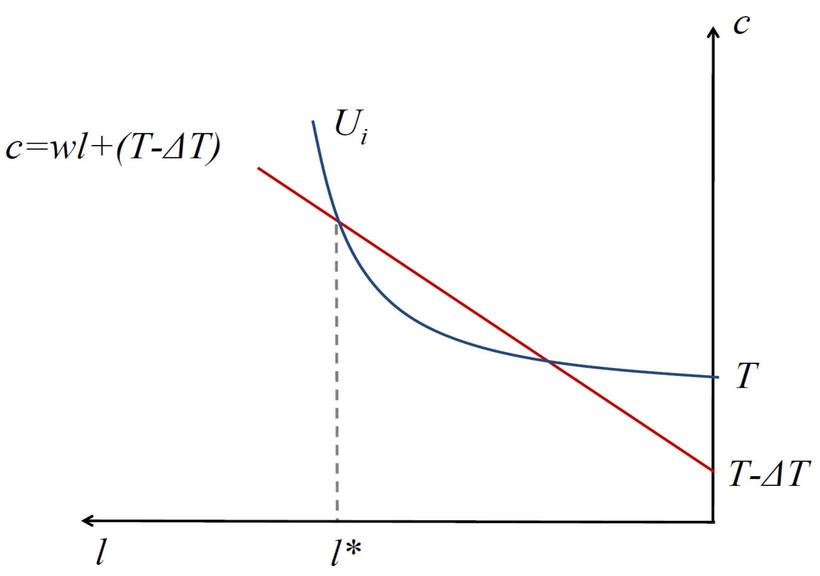
yielding again a structural probit of the form:

$$P = \Phi\left(\gamma \log W_i + Z_i \alpha' - \overline{\psi} \log T_i\right) \tag{3}$$

Let us compare the two structural probit equations (1) and (3). First, W_i in (3) represents the gains to work (from a full time job): $W_i = w_i l^* - \Delta T$, as opposed to the net wage w_i. Second, T_i is the hypothetical amount of transfers one gets (or would get) at zero hours worked.

Figure 3

The Reservation Wage
when There Is a Discrete Drop in Transfers



From a practical point of view, T is not directly observable for the employed, since they get T- ΔT; while ΔT is not observed for the inactive, since they get T. Using individual characteristics and the welfare system's details (for every given year), however, one can back up T and ΔT. This essentially requires a microsimulation tool. For those who work, we determine T based on their characteristics and welfare regulations for the given year, and then obtain $\Delta T = T - T_{obs}$. For those who do not work, we determine ΔT by again applying welfare rules, while $T = T_{obs}$.

3 Econometric issues

Here we closely follow Kimmel and Kniesner (1998), up to a certain point. We want to estimate a structural probit equation:

$$P(\text{employed/active}) = \Phi\left(\gamma \log W_i + Z_i \alpha' - \overline{\psi} \log T_i\right)$$

where $W_i = w_i l^* - \Delta T_i$. Here the vector Z_i contains individual characteristics which shift the labour supply of an individual. As usual in the literature on participation, there is a missing data issue: the wage is unavailable for those who do not work. The solution is to use a predicted W for the inactives: run

$$\log W_i = X_i\beta' + \mu_i$$

for the employed, and use the predicted wage $\hat{W} = X_i\hat{\beta}$ for the unemployed. Here the vector X_i contains individual characteristics which are relevant for defining an individual's wage. Note that the two vectors X_i and Z_i may overlap, but there can be elements in each of them which are excluded from the other set. This regression, however, is run on a nonrandom sample, since the employment and the W error terms might be correlated. The solution is thus to adopt a Heckman-type correction, yielding a three step procedure.

In variant A, we thus adopt the following procedure:

1) Run a reduced form probit:

$$P(\text{employed}) = \Phi\left(X_i\beta'_{RF} + Z_i\alpha'_{RF} - \psi_{RF}\log T_i\right)$$

2) Use the inverse Mills ratio $\lambda(x) = \dfrac{\phi(x)}{\Phi(x)}$ as a correction in the log GTW regression:

$$\log W_i = X_i\beta' + \delta\lambda\left(X_i\hat{\beta}'_{RF} + Z_i\hat{\alpha}'_{RF} - \hat{\psi}_{RF}\log T_i\right) + \mu_i$$

3) Use the predicted log GTW $lo\hat{g}W_i = X_i\hat{\beta}'$ in the structural probit equation:

$$P(\text{employed/active}) = \Phi\left(\gamma\, lo\hat{g}\, W_i + Z_i\alpha' - \overline{\psi}\log T_i\right)$$

Notice that here $X \supseteq Z$, since there is practically no observable characteristics which would not be related to transfer measures, which are there in $\log W$.

In variant B, we slightly modify the previous procedure:

1) Run a reduced form probit

$$P(\text{employed}) = \Phi\left(X_i\beta'_{RF} + Z_i\alpha'_{RF} - \psi_{RF}\log T_i\right)$$

2) Use the inverse Mills ratio $\lambda_i(x) = \dfrac{\phi(x)}{\Phi(x)}$ as a correction in the *wage* (more precisely: monthly income) regression:

$$\log w_i = X_i\beta' + \delta\lambda\left(X_i\hat{\beta}'_{RF} + Z_i\hat{\alpha}'_{RF} - \hat{\psi}_{RF}\log T_i\right) + \mu_i$$

3) If W_i is also lognormal with some mean and a variance σ_W^2, then one can show that:

$$E\left(log(W_i)\middle|X_i, Z_i\right) = log\left(E\left(W_i\middle|X_i, Z_i\right)\right) - \frac{1}{2}\sigma_W^2 = log\left(e^{X_i\beta_1 + \frac{1}{2}\sigma_1^2} - \Delta T_i\right) - \kappa$$

Thus we can use the predicted log wage $lo\hat{g}\, w_i = X_i\hat{\beta}'$, add the standard error correction for lognormals, exponentiate, subtract ΔT_i and take logs again to obtain the predicted log GTW for the structural probit equation:

$$P(\text{employed/active}) = \Phi\left(\gamma\, lo\hat{g}\, W_i + Z_i\alpha' - \overline{\psi}\log T_i\right)$$

Four remarks are in order. The first is regarding endogeneity and measurement error of the gains-to-work variable. In the structural probit, $\log W$ can be endogenous, since the wage error term can be correlated with the participation decision error term. Moreover, $\log W$ can also contain measurement error: in case of an individual working only for some part of the year, her reported wage is less than the true annual wage. Alternatively, unreported wage income can also lead to a mismeasurement of wages. Notice, however, that we are in fact running an IV-probit in step 3, which offers a remedy to both of these problems (as long as there are variables in X_i which are excluded form Z_i, an issue we address in the data section).

The second issue is whether the selection correction is identified only through a functional form assumption. This is indeed the case when $X \supseteq Z$ in the wage equation, since the inverse Mills ratio is then just a nonlinear reshuffling of the right hand side variables in the wage equation (variant A). On the other hand, the inverse Mills ratio does contain additional variation if $X \not\supseteq Z$, which is the case in Variant B. This means that we are free from the functional-form criticism in Variant B, but it applies for the wage equation in Variant A. In that case, however, there is no alternative: if a variable impacts the participation equation directly, it is also likely to impact the GTW ($\log W$) at least through the change in transfers term ΔT. For the structural probit equation (3) however, we are again on safe grounds: though the predicted $\log W$ contains the variables X, Z and their nonlinear combinations (in the inverse Mills ratio), X is excluded from the structural equation, so we are identifying γ from variations both in X and the inverse Mills ratio. In other words, the key element of the identification method is the existence of controls for labour demand included in X_i and excluded form Z_i.

Third, the use of generated regressors in the third stage calls for an adjustment of standard errors. Usual Heckman correction implementations do incorporate necessary corrections for the second but not for the third step. In practice, such a correction often leads to minor changes; hence it is common to ignore the issue (Kimmel and Kniesner, 1998, also follow this route). As one alternative, one could implement a full-blown correction of the third step standard errors, along the lines of Fernandez *et al.*. We instead opted for bootstrapping the standard errors, which should be more robust in case of noisy data or misspecification problems.[3]

Finally, there is a tradeoff between adopting Variant A or B. The latter would seem more appealing, since it allows for $X \not\supseteq Z$, hence even the wage equation is free from functional form criticisms. The drawback, however, is that nothing guarantees that our estimated $\hat{W}_i = e^{X_i\hat{\beta}_1 + \frac{1}{2}\hat{\sigma}_1^2} - \Delta T_i$ is positive, causing a nonrandom sample selection issue in our third step. One could produce better second stage regressions for $\log w_i$, taking for example the impact of the minimum wage into account.[4] That would mean, however, a Tobit-type truncated regression in the second stage, making our procedure even more complicated and potentially four-step. For this reason, we proceed only with Variant A; also recalling that although the wage equation is subject to a functional form criticism, is is much less of an issue in the structural probit equation.

Since our "wage" measure in the structural estimation is the GTW, the calculation of regular wage elasticities requires one more step. The structural probit gives us a log GTW coefficient γ. Since the probit is a nonlinear function, one has to evaluate it at a certain vector Z and $\log T$ to obtain the marginal impact of a percentage change in the GTW. Even then, however, it is still the

---

3 In particular, our reported standard errors are calculated as the standard deviation of the point estimates from the three-step estimation procedure performed on 200 bootstrapped random samples (with replacement, and of the same size as the estimation sample).

4 It was indeed the case in our sample that the predicted wage was too low for the low-skilled, where the minimum wage is often binding, making their predicted $\hat{W}_i$ negative.

impact of a change in W, not w.

To obtain the impact of the wage itself, note that:

$$\frac{\partial \log(w-\Delta T)}{\partial \log w}=\frac{\partial \log\left(e^{\log w}-\Delta T\right)}{\partial \log w}=\frac{e^{\log w}}{e^{\log w}-\Delta T}=\frac{w}{w-\Delta T}$$

So:

$$\frac{\partial \Phi}{\partial \log w}=\frac{\partial \Phi}{\partial \log W}\frac{\partial \log W}{\partial \log w}=\frac{\partial \Phi}{\partial \log W}\frac{w}{w-\Delta T} \tag{4}$$

Notice that the marginal effect of $\log W$ gets magnified if $w - \Delta T = w$; which is the case for transfer-dependent people (low skill, around retirement, etc.).

4 Data

We use data from the Hungarian Household Budget Survey (HKF), years 1998-2008. This is in principle a rotating panel database with a one-third renewing part every year, but it is very difficult to make the actual connections between consecutive waves. For this reason, we only use it as a pooled cross-section. The dataset contains detailed income and consumption measures of broadly 25,000 individuals per year.

The key challenge is to define the counterfactual transfers: First, how much would someone who is currently working receive in transfers if that individual is laid off? Second, how much would someone who is currently inactive lose if that individual takes up a full time job? Calculating these measures requires the detailed coding of the full transfer system, basically a microsimulation model. We detail the major tax expenditure and cash transfer items in the Appendix. With one exception, the database contained all the relevant information to deduct the counterfactual transfer entitlements or losses of each individual. The exception was the work history of individuals, on which certain transfers depend (for example, eligibility to the more generous maternity support schedule GYED). To resolve this issue, we used a predicted value based on the Labour Force Survey database (a conditional expectation based on observable characteristics).

The main left hand side variable was labour force participation,[5] though we also ran the same estimations with employment. All wage variables (w and W) refer to annual net wage income calculated from the gross wages reported by survey participants. The right hand side measures form two major groups: labour-supply shifters (Z_i) and wage equation controls ($X_i \backslash Z_i$). Following MaCurdy (1985), MaCurdy (1987), and Kimmel and Kniesner (1998), labour-supply shifters contain personal and family characteristics, while the vector X_{it} includes variables which determine the market wage (labour demand shifters). In particular, the first group consists of the following variables: log of non-labour income, education dummies, household head, mother with infant (<3 years old), attending full-time education, household size (number of persons), pensioner, family status (husband, wife, child, single, divorced,...), age-group dummies (15-24, 25-49, 50-) and year dummies. The second group contains county dummies, and interactions of age and age square with education.

One needs to justify the choice for variables in $X_i \backslash Z_i$, since those variables serve both as instruments for treating endogeneity and measurement error issues about our wage measure (see the first remark at the end of Section 3), and also as a source of additional variation to identify the

5 It is the "most typical" status for the given year, self-reported by survey respondents. Unemployment is defined along the ILO classification.

Table 1

Main Results

	(A) Estimation Results			
	Participation (1)		Employment (2)	
	Coeff.	**Std. Err.**	**Coeff.**	**Std. Err.**
gains to work	0.820	*0.099*	0.761	*0.089*
non-labour income	–0.844	*0.110*	–0.702	*0.098*
	(B) Conditional Marginal Effects			
	dy/dx	**Std. Err.**	**dy/dx**	**Std. Err.**
gains to work	0.290	*0.028*	0.301	*0.031*
non-labour income	–0.298	*0.030*	–0.277	*0.035*
net wage	0.395	*0.038*	0.410	*0.042*
transfer	–0.136	*0.013*	–0.137	*0.015*

Notes: Three-step estimates, as described in the paper. Standard errors are bootstrapped with 200 replications. Structural probit equation includes: log of gains to work, log of non-labour income, mother with infant (less then three years-old), full time student, education dummies (less then elementary school, elementary school, vocational, secondary education, tertiary education), age-group dummies (15-24, 25-49, ≥50), pensioner, gender, head of household dummy, household size, family status dummies (single, married living together, married living separately, widow(er), divorced), household membership status dummies (husband, wife, companion, single parent, child, ascendant, other relation, non-relation, single), year dummies. Controls included in the reduced-form probit and the wage equation which are missing from the structural probit are: county dummies, interaction of age and age square with education dummies.
Source: Household Budget Survey database, 1998-2008.

parameter γ (remark two of the same section). In our view, county dummies represent regional differences in economic conditions, which has an indirect effect on activity (through different wages) but no direct effect (two individuals with identical individual characteristics and wage but living in different regions should exhibit the same attitude towards economic activity). For the interaction of age and age square with education, our argument is the following. Age has two main effects on the likelihood of activity: one is through an impact on the lifecycle position (student, prime age and nearing retirement), and another through increased experience (an upward sloping relationship between age and wages). The first effect is a labour-supply shifter, which we capture by a large set of dummies that controls for individuals' lifecycle position, such as age-group, family status (single, married, divorced...), attending full-time education, mother with infant and others. On top of that, we argue that an extra year has a negligible impact on labour supply directly, but it strongly impacts the wage and hence impacts activity indirectly (a labour demand shifter).

5 Results

This section reports and discusses our empirical results. We focus mostly on the participation margin: with employment, we only report the results of the main specification but no detailed conditional marginal effects by subgroups (they are available upon request). The main parameters of interest are the coefficient of gains to work and non-labour income (always in logs). Table 1 displays our baseline results, following the econometric methodology of Variant A. Panel A reports the estimates for the structural probit equation (3). Most point estimates have the expected sign and are significant. A higher GTW increases the probability of being active, while non-labour income has the opposite effect (both are in logs). From the additional controls (unreported but available upon request), education has a mixed but insignificant effect. Being a household head or having a larger family increases the probability of being active, while being a mother with small children, full-time student or pensioner decreases it. Age has the usual hump-shaped effect on activity. The results are quite similar when the left hand side variable is employment.

Since the probit function is nonlinear, the point estimates in Panel A are not indicative about the conditional marginal effect of variables of interest on activity. Panel B displays these numbers, evaluated at the sample means. Numbers here are already semi-elasticities: a 10 per cent increase in the GTW leads to a 2.9 per cent increase in the probability of being active. As explained by equation (4), the same increase in the net wage (as opposed to the net wage minus transfers) leads to a potentially larger effect. The difference is quite substantial at the sample mean, as the effect is about 36 per cent higher. The opposite happens with non-labour income: transfers are only part of them, so a 10 per cent change in transfers implies a smaller increase in non-labour income.

The conditional marginal effects presented in Table1 are not directly comparable to the "consensus" 0.25 value of aggregate net wage elasticity reported by Chetty *et al.* (2012): these marginal effects indicate the effect of one percent increase in net wage on the "average individual's" probability of being active (or on the participation rate) in percentage points, as opposed to the elasticity measures in Chetty *et al.* (2012) indicating the percentage change in total employment to the same shock. To produce the equivalent of the exercise by Chetty *et al.* (2012), one needs to increase the net wage of all individuals by one percent and look at its employment effect. The resulting 0.28 per cent increase in total employment implies an elasticity of 0.28, quite in line with the consensus.

Next we look at the conditional marginal effects by subgroups to see how much they differ from each other. Table 1 presents two variants, a full and a restricted sample estimate. The full sample means that all observations are included (as in Table 2), but the marginal effects are evaluated at a subgroup-specific mean. The restricted sample means that the entire estimation procedure is carried out only on the subsample at hand, so even the structural probit estimates can be different.

Notice that the net wage (or even the GTW) elasticity of activity is highly different across the three educational groups even in the full sample estimation case, when the only reason is a different conditional mean of the subgroups. The probit estimates somewhat differ between the full and the restricted sample, though the latter is also much less precisely estimated. Still, the conditional marginal effects are quite similar. This result is noteworthy, as it means that one can explain the heterogeneity of participation elasticities without an underlying difference in the utility functions (*i.e.*, the parameters γ and $\overline{\psi}$ in equation (3)).

If those two parameters are common across individuals, than labour supply elasticities at the intensive margin are also common: one can show that for a fixed income share $W/(W+T)$ and expenditure share (1 $\alpha = c/(c + w(1-\bar{l})$, the impact of a change in the net wage or transfers is the same on the hours worked decision of every individual. This homogeneity is however partial, since individuals with different gross wages (productivity) or transfers (non-labour income in general) will have different income and expenditure shares. When there is no non-labour income (T=0), this homogeneity becomes even more complete, as the labour supply elasticity depends only on common parameters and original hours worked ($\alpha = \bar{l}$). So if individuals differ in their characteristics but their original hours worked is the same, so is their intensive margin labour supply elasticity. If utility is linear in consumption ($\psi = 0$), then the elasticity (of leisure) to net wages is common across all individuals (full homogeneity).

Table 2

Probit Estimates and Conditional Marginal Effects by Subgroups

		Full Sample (1)		Restricted Sample (2)	
		dy/dx	**Std. Err.**	**dy/dx**	**Std. Err.**
elementary school or less	gains to work (probit)	0.820	*0.099*	0.583	*0.082*
	non-labour income (probit)	–0.844	*0.110*	–0.639	*0.111*
	gains to work	0.212	*0.064*	0.175	*0.085*
	non-labour income	–0.218	*0.068*	–0.192	*0.101*
	net wage	0.294	*0.089*	0.275	*0.133*
	transfer	–0.093	*0.028*	–0.109	*0.053*
secondary education	gains to work (probit)	0.820	*0.099*	0.710	*0.151*
	non-labour income (probit)	–0.844	*0.110*	–0.715	*0.165*
	gains to work	0.219	*0.022*	0.213	*0.031*
	non-labour income	–0.225	*0.024*	–0.214	*0.034*
	net wage	0.310	*0.031*	0.286	*0.041*
	transfer	–0.118	*0.012*	–0.098	*0.014*
tertiary education	gains to work (probit)	0.820	*0.099*	0.915	*0.323*
	non-labour income (probit)	–0.844	*0.110*	–0.856	*0.326*
	gains to work	0.110	*0.012*	0.130	*0.029*
	non-labour income	–0.113	*0.012*	–0.121	*0.031*
	net wage	0.139	*0.015*	0.156	*0.035*
	transfer	–0.045	*0.005*	–0.043	*0.010*

Notes: Column (1) reports probit estimates and conditional marginal effects computed from the estimation on the full sample and evaluated at the subgroup-specific mean values of the covariates. Column (2) reports similar marginal effects, but computed from the estimations on the restricted samples.

Table 2 further explores the prime-age sample, checking whether education status also matters there. The low overall elasticity of this age group splits into a sizeable elasticity for the "elementary school or less" group (a group which is also highly welfare dependent) and a smaller but still significant number for prime-age individuals with secondary education. Estimations suggest that prime-age higher educated individuals are inelastic to tax and transfer changes at the extensive margin. The restricted samples yield similar though smaller differences, both for structural probit parameters and conditional marginal effects.

Table 3 displays the conditional marginal effects for the two remaining main welfare dependent social groups, the elderly and women of child-bearing age. The group of age above 50 exhibits a very substantial elasticity – this partly explains the large gap between the elasticity of the entire population and the prime-age group. This finding is quite important, as it shows that taxes and transfers have a strong impact on activity around retirement age, and that the tax and social insurance system can contribute to the large activity gap of the elderly in Hungary. Women at child-bearing age show a smaller wage elasticity, though they are still more responsive than the overall prime-age group. This is also true about the impact of transfers.

Table 3

Probit Estimates and Conditional Marginal Effects by Subgroups, Prime-age Subsample

		Full Sample (1)		Restricted Sample (2)	
		dy/dx	**Std. Err.**	**dy/dx**	**Std. Err.**
full prime-age sample	gains to work (probit)	0.820	*0.099*	0.646	*0.122*
	non-labour income (probit)	–0.844	*0.110*	–0.620	*0.129*
	gains to work	0.088	*0.010*	0.086	*0.008*
	non-labour income	–0.091	*0.010*	–0.083	*0.008*
	net wage	0.127	*0.014*	0.124	*0.011*
	transfer	–0.054	*0.006*	–0.051	*0.005*
prime-age, elementary school or less	gains to work (probit)	0.820	*0.099*	0.323	*0.164*
	non-labour income (probit)	–0.844	*0.110*	–0.299	*0.185*
	gains to work	0.249	*0.025*	0.109	*0.051*
	non-labour income	–0.256	*0.026*	–0.101	*0.058*
	net wage	0.409	*0.040*	0.180	*0.085*
	transfer	–0.194	*0.019*	–0.084	*0.041*
prime-age, secondary education	gains to work (probit)	0.820	*0.099*	0.403	*0.182*
	non-labour income (probit)	–0.844	*0.110*	–0.364	*0.192*
	gains to work	0.081	*0.008*	0.057	*0.017*
	non-labour income	–0.084	*0.008*	–0.051	*0.019*
	net wage	0.122	*0.012*	0.084	*0.025*
	transfer	–0.054	*0.005*	–0.036	*0.011*
prime-age, tertiary education	gains to work (probit)	0.820	*0.099*	–0.206	*0.420*
	non-labour income (probit)	–0.844	*0.110*	0.217	*0.400*
	gains to work	0.038	*0.003*	–0.019	*0.041*
	non-labour income	–0.039	*0.003*	0.020	*0.040*
	net wage	0.050	*0.004*	–0.023	*0.051*
	transfer	–0.019	*0.001*	0.008	*0.017*

Notes: Column (1) reports probit estimates and conditional marginal effects computed from the estimation on the full sample and evaluated at the subgroup-specific mean values of the covariates. Column (2) reports similar marginal effects, but computed from the estimations on the restricted samples.

Finally, Table 3 also report results for the usual classification by sex and marital status. Consistently with most of the previous empirical findings, women are, in general, more responsive to tax and transfer changes than men. Married women, the group mostly studied in the literature exhibits the highest marginal elasticity, while married men seem to be the less responsive group.

In summary, we have found that wages, taxes and transfers have a large impact on the participation decision, particularly for elders, the low-skilled, married women and women at child-bearing age. Moreover, these differences can be largely explained by different group characteristics, leading to different conditional marginal effects of the same structural probit estimates, and also to a different multiplication of a net wage change into the change in the GTW.

Table 4

Conditional Marginal Effects by Selected Subgroups

		dy/dx	**std. err.**
elder (>=50)	gains to work	0.311	*0.052*
	non-labour income	–0.320	*0.057*
	net wage	0.392	*0.065*
	transfer	–0.103	*0.017*
women at child-bearing age (25-49)	gains to work	0.146	*0.013*
	non-labour income	–0.151	*0.014*
	net wage	0.231	*0.021*
	transfer	–0.108	*0.010*
prime-age, single men	gains to work	0.069	*0.008*
	non-labour income	–0.071	*0.009*
	net wage	0.096	*0.012*
	transfer	–0.038	*0.005*
prime-age, single women	gains to work	0.113	*0.013*
	non-labour income	–0.116	*0.013*
	net wage	0.168	*0.019*
	transfer	–0.076	*0.008*
prime-age, married men	gains to work	0.028	*0.003*
	non-labour income	–0.029	*0.004*
	net wage	0.039	*0.005*
	transfer	–0.016	*0.002*
prime-age, married women	gains to work	0.183	*0.016*
	non-labour income	–0.189	*0.017*
	net wage	0.290	*0.025*
	transfer	–0.133	*0.012*

We now demonstrate how our results can be utilized for the simulation of the labour supply (participation) effect of a personal income tax and transfer reform. The main step is to calculate the probability of being active for a given hypothetical wage, tax and transfer system. First we obtain the pre- and post-reform aftertax wage income of everyone in our sample, using predicted wages. Then we calculate the pre- and post-reform hypothetical “zero hours worked” transfer level for everyone, and construct the log of the GTW (log$W$) before and after the reform.

Equipped with these, we form:

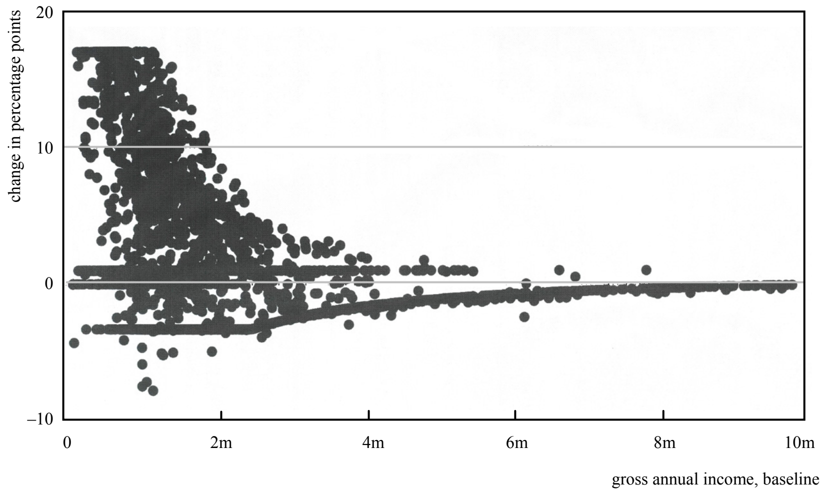
$$\Phi\left(\hat{\gamma}\log W_i + Z_i\hat{\alpha}' - \hat{\bar{\psi}}\log T_i\right)$$

before and after the reform. The change in its value is the change in the probability of individual i being active. Finally, we add up the probabilities in the sample (weighted) to get an estimate for the change in the aggregate activity rate. This gives us the shift of the labour supply curve: in equilibrium, labour demand might be downward sloping so the equilibrium wage may change, offsetting partly the change in labour supply.

As an illustration, we fed the main changes of the Hungarian personal income tax and transfer system of 2012 into this framework. The particular measures are the following: the complete elimination of the employee tax credit (ETC) scheme, a 27 per cent reduction in the tax rate (from 20.3 to 16 per cent) below the average monthly income of 202,000 HUF, and a 1 percentage point increase in the social contribution rate. As illustrated by Figure 3, these changes have a very heterogeneous effect on the average tax rate of taxpayers: the abolishment of the ETC pushes up the average tax rate for low earners, for which they are partly compensated by the cut in the tax rate. Medium earners, who were not or at most partially eligible for the ETC gain by a reduction in their tax rate. High earners also gain a little due to the reduction in the tax rate on their first 202,000 HUF income per month. Finally, there is a common loss from increased social contributions.

Figure 4

Change of Average Effective Tax Rates (AETR)



As a result, aggregate activity decreases by 0.97 per cent, from which the elimination of the ETC is responsible for 2.09 per cent,[6] the increase in social contributions leads to another 0.34 per cent reduction, which are partly offset by an increase of 1.51 per cent due to the rate cut.[7] Overall, this illustrates both the usefulness of our parametric approach for assessing the impact of tax and transfer reforms, and the economic significance of our parameter estimates.

6 Conclusion

This paper presents a first (at least to our knowledge) structural form estimation of labour supply at the extensive margin that simultaneously takes into account taxes and transfers. We show that one has to modify the net wage by deducting the amount of lost transfers to get the measure which determines the participation decision (the gains to work). This implies, however, that the same change in the net wage leads to a very different change in the GTW if lost transfers are a different share of the net wage.

We find that a single equation can already explain a large heterogeneity of individual responsiveness to taxes and transfers: there are large differences among subgroups, driven partly by

---

6 There is a subtle issue here: under the Hungarian tax code, a large part of social transfers are also affected by personal income taxes and the ETC. Consequently, the elimination of the ETC also decreases the net value of many social transfers. Thanks to our integrated treatment of taxes and transfers, we can take this into account in our calculation. Without the corresponding cut in the net value of transfers, there would be an even more substantial reduction in participation.

7 The sum of the effects of these measures may differ to the total effect due to interactions.

a composition effect, and partly by a different share of lost transfers in the GTW. These highly responsive subgroups are exactly the ones who are mostly responsible for Hungary's low participation rate (low-skilled, women at child-bearing age, elders), implying that a reform of the tax and transfer system can be a powerful tool to boost employment.

Our results directly lend themselves to reform simulations. We demonstrated how our model can be utilized to calculate the labour supply shift of a complex personal income tax reform. In related work (Benczúr *et al.*, 2012), we build a model where this labour supply block is expanded by an intensive margin adjustment (based on a combination of Bakos *et al.*, 2008; and Á. Kiss and Mosberger, 2011), and then it is embedded in a small general equilibrium macro model. With such a fully fledged model, we were able to evaluate at depth the 2011-12 Hungarian tax and transfer reforms as well (Benczúr *et al.*, 2011).

APPENDIX
SUMMARY OF CASH TRANSFERS AND TAX EXPENDITURES TAKEN INTO ACCOUNT IN THE ESTIMATION

This Appendix summarizes the basic features of tax expenditures and the cash transfers and tax expenditures taken into account in the estimation. In particular, we discuss child care (family) benefits and unemployment (welfare) benefits. We treated old-age and disability benefits as exogenous and, accordingly, did not include these benefits in the summary. This rests on the assumption that if an individual is entitled for these benefits (due to age or health status), we will observe that he/she is a recipient. This looks like a natural assumption in the case of disability benefits. In the case of old-age benefits, this treatment is justified by the fact that during the sample period old-age pension recipients were allowed to work without any penalty. Thus they did not face a choice between pensions and earnings.

1 Tax expenditures in the PIT

a) *Employee tax credit (adójóváírás),*[8] *ETC* is a non-refundable tax credit applying to wage income. The ETC was modest in size until its expansion in 2002. During the period 2003-11 it made the minimum wage nearly PIT-free. The ETC was phased out in most years at a rate of 9 per cent in an income range around the average wage. Until its abolishment in 2012, its exact parameters were adjusted each year.
b) *Family tax credit (családi adókedvezmény).* The Hungarian PIT has been an individual-based (as opposed to a family-based) tax system during the sample period. One of the parents can deduct the family tax credit from his or her tax payment (or both can share the credit) based on the number of children in the household. Starting in 2006, families with one or two children were not eligible for the tax credit (until the tax credit was expanded in 2011).
c) *Other tax credits* were abundant in the tax code until 2006; since then they have been gradually eliminated. We use information in the Household Budget Survey to assess the tax credits each individual can take advantage of.
d) *Tax base issues*. During the sample period, insurance-based benefits were generally treated as wage income by the tax code while universal benefits were tax exempt. During the years 2007-10 pension income constituted part of the tax base although it was not taxed itself (it pushed other incomes into the upper tax bracket). Benefits 2c and 2d were treated similarly during the whole sample period.

2 Family benefits

a) *Maternity benefit (TGYÁS)* is an insurance-based benefit that mothers are entitled to receive for 5 months around child-birth. Its condition is current employment (at the time of applying for the benefit). The monthly benefit is equal to 70 per cent of past monthly wage. The recipient may not engage in paid work while receiving this benefit. No couple can receive two of benefits 2a-d at the same time.
b) *Child-care benefit I (GYED)* is an insurance-based benefit that one of the parents is entitled to receive until the second birthday of the youngest child. Its condition is at least 12 months of

[8] There is considerable heterogeneity in the official and scientific publications regarding the English translation of the various benefits. In this table we chose to use the simplest English translations that reflect the nature of the given benefit; we included the official Hungarian designations so that the benefits can easily be identified.

employment in the 24 months before the child is born. The monthly benefit is equal to 70 per cent of past monthly wage but it may not exceed 140 per cent of the minimum wage. The recipient may not engage in paid work while receiving this benefit. No couple can receive two of benefits 2a-d at the same time.

c) *Child-care benefit II (GYES)* is not conditional on employment (social insurance) history. One of the parents is entitled to receive the benefit until the third birthday of the youngest child. The benefit is pegged to the so-called “minimum pension benefit”, equal to HUF 28500 (around 40 per cent of the minimum wage) in 2008. Recipients are restricted from working full time in the first year of this benefit. (The employment restrictions were loosened for the second and third year during the period of study.) No couple can receive two of benefits 2a-d at the same time.

d) *Child-care benefit III (GYET):* A parent is entitled to this benefit if he or she raises at least 3 children until the 8th birthday of the youngest child, independently of employment (social insurance) history. The benefit is pegged to the ’minimum pension benefit’ (see 2c). Recipients of this benefit are restricted from working full time. No couple can receive two of benefits 2a-d at the same time.

e) *Family supplement (sometimes called “family allowance”; családi pótlék)* is a universal benefit all families with children are entitled to receive. The sum of the benefit depends on the number of children, whether there are twins among the children, and whether any of the children is chronically ill. It was equal to HUF 12,200 (around 18 per cent of the minimum wage) for a family with one child in 2008.

3 Unemployment benefits

a) *Unemployment benefit I (1998-2005: munkanélküli járadék; 2006-: álláskeresési járadék):* Individuals who lost their jobs are eligible for the insurance-based unemployment benefit (renamed as “job-seekers’ benefit” in 2006). Its maximum duration was shortened from 12 months to 9 months in 2000. Until 2006 it was equal to 65 per cent of the previous wage (capped at 180 per cent of the “minimum pension benefit”, see 2c). After 2006 it had two phases. The first phase lasted 3 months, during which the recipient received 60 per cent of his/her past wage (capped at 120 per cent of the minimum wage). The second phase lasted 6 months, during which the benefit was equal to 60 per cent of the minimum wage. (If the individual did not have a full employment history in the four years before the job loss, the duration of the benefit could be shorter. The second phase was abolished in 2012.)

b) *Unemployment benefit II (2003-05: álláskeresést ösztönzõ juttatás; 2006-: álláskeresési segély):* Established in 2003, this was a fixed-sum benefit for individuals whose unemployment benefit I expired but still did not find a job. It was conditional on cooperation with the local unemployment administration. Between 2003-05 the benefit lasted a maximum of 6 months; it was reduced to 3 months in 2006. From that year onwards the benefit was equal to 40 per cent of the minimum wage. (It was abolished in 2012).

c) *Pre-retirement unemployment benefit (Nyugdíj elõtti álláskeresési segély):* Individuals are entitled for this insurance-based benefit (which used to be a sub-case of benefit 3b after 2006) if they lose their job in the five years before the statutory pension age. The benefit is equal to 40 per cent of the minimum wage. The benefit payment is suspended if the individual finds employment.

d) *Regular social benefit (1998-2000: jövedelempótló támogatás; 2001-: rendszeres szociális segély)* is a a welfare benefit individuals can receive if they are not eligible to any other unemployment (or disability or child-care) benefit (any more). For most of the sample period it was means-tested. The details of the means-testing changed in 2006. After 2006 the benefit supplemented a family's income to 90 per cent of the "minimum pension benefit" per consumption unit but could not exceed the net minimum wage. (Its predecessor in the years 1998-2000 was a fixed-sum transfer and it was succeeded by a fixed-sum transfer in 2010.)

REFERENCES

Aaberge, R., U. Colombino and S. Strom (1999), "Labour Supply in Italy: An Empirical Analysis of Joint Household Decisions, with Taxes and Quantity Constraints", *Journal of Applied Econometrics*, Vol. 14, No. 4, pp. 403-22.

Arrufat, J.L. and A. Zabalza (1986), "Female Labor Supply with Taxation, Random Preferences, and Optimization Errors", *Econometrica*, Vol. 54, No. 1, pp. 47-63.

Bakos, P., P. Benczúr and D. Benedek (2008), "The Elasticity of Taxable Income: Estimates and Flat Tax Predictions Using the Hungarian Tax Changes in 2005", Magyar Nemzeti Bank, Working Paper, No. 2008/7.

Benczúr, P., G. Kátay and Áron Kiss (2012), "Assessing Changes of the Hungarian Tax and Transfer System: A General Equilibrium Microsimulation Approach", Magyar Nemzeti Bank, Working Paper, forthcoming.

Benczúr, P., G. Kátay, Áron Kiss, B. Reizer and M. Szoboszlai (2011), "Analysis of Changes in the Tax and Transfer System with a Behavioural Microsimulation Model", *Magyar Nemzeti Bank Bulletin*, Vol. 6, No. 3, pp. 15-27.

Blundell, R. (2012), "Tax Policy Reform: The Role Of Empirical Evidence", *Journal of the European Economic Association*, Vol. 10, No. 1, pp. 43-77.

Chetty, R., A. Guren, D. Manoli and A. Weber (2012), "Does Indivisible Labor Explain the Difference Between Micro and Macro Elasticities? A Meta-analysis of Extensive Margin Elasticities", in *NBER Macroeconomics Annual 2012*, Vol. 27, NBER Chapters, pp. 1-51, National Bureau of Economic Research.

Dickert, S., S. Houser and J.K. Scholz (1995), "The Earned Income Tax Credit and Transfer Programs: A Study of Labor Market and Program Participation", in J.M. Poterba (ed.), *Tax Policy and the Economy*, Vol. 9, NBER chapters, pp. 1-50, National Bureau of Economic Research.

Eissa, N. and J.B. Liebman (1996), "Labor Supply Response to the Earned Income Tax Credit", *The Quarterly Journal of Economics*, Vol. 111, No. 2, pp. 605-37.

Fernandez, A.I., J.M. Rodriguez-Poo and S. Sperlich (2001), "A Note on the Parametric Three Step Estimator in Structural Labor Supply Models", *Economics Letters*, Vol. 74, No. 1, pp. 31-41.

Kimmel, J. and T.J. Kniesner (1998), "New Evidence on Labor Supply: Employment Versus Hours Elasticities by Sex and Marital Status", *Journal of Monetary Economics*, Vol. 42, No. 2, pp. 289-301.

Kiss, Á. and P. Mosberger (2011), "The Elasticity of Taxable Income of High Earners: Evidence from Hungary", technical report.

Kátay, G. (ed.) (2009), "Az alacsony aktivitás és foglalkoztatottság okai és következményei Magyarországon [Causes and Consequences of Low Participation and Employment in Hungary]", Magyar Nemzeti Bank, Occasional Paper, No. 79.

Kátay, G. and B. Nobilis (2009), "Driving Forces Behind Changes in the Aggregate Labour Force Participation in Hungary", Magyar Nemzeti Bank, Working Paper, No. 2009/5.

MaCurdy, T.E. (1985), "Interpreting Empirical Models of Labor Supply in an Intertemporal Framework with Uncertainty", in J.J. Heckman and B. Singer (eds.), *Longitudinal Analysis of Labor Market Data*, NBER chapters, pp. 111-55, Cambridge University Press.

MaCurdy, T.E. (1987), "A Framework for Relating Microeconomic and Macroeconomic Evidence of Intertemporal Substitution", in T.F. Bewley (ed.), *Advances in Econometrics*, acts of the *5th World Congress of the Econometric Society*, Vol. 2, pp. 149-76, Cambridge University Press.

Meyer, B.D. and D.T. Rosenbaum (2001), "Welfare, The Earned Income Tax Credit, and the Labor Supply of Single Mothers", *The Quarterly Journal of Economics*, Vol. 116, No. 3, pp. 1063-114.

Moffitt, R.A. (2002), "Welfare Programs and Labor Supply", in A.J. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics*, Vol. 4, Chap. 34, pp. 2393-430, Elsevier.

AN EVALUATION OF THE 1997 FISCAL DECENTRALIZATION REFORM IN MEXICO: THE CASE OF THE HEALTH SECTOR

André Martínez Fritscher[*] *and Carolina Rodríguez Zamora*[**]

This paper studies the impact of the health decentralization of funds and responsibilities that took place in Mexico in 1997 on state level health outcomes. It renders two main results. First, the magnitude of transfers from the federal government to states failed to take into account state-specific needs; instead, transfers were mainly determined by the pre-reform health expenditures of the federal government in each state. Second, decentralization did not boost the advances in health outcomes already achieved under the centralized health sector regime. We conclude by discussing potential reasons for the results found in this paper.

1 Introduction

Fiscal decentralization has been part of the reform agenda in many developing countries for the last two decades. Theoretically, state and local fiscal autonomy is founded on the idea that public policy decisions by lower tiers of governments would bring about more efficient outcomes in the provision of public goods (Oates, 1972). It is argued that sub-national governments are better able to identify the needs and preferences of citizens. Under fiscal decentralization, taxpayers are closer to authorities, allowing them to better demand transparency, accountability, and efficiency in the use of public resources. As a result, decentralization is expected to generate economic growth and improvements in the welfare of the population.[1] Having these positive effects in mind, Mexico undertook a profound reform in the 1990s to modify the expenditure responsibilities of the federation and state governments. The main aim of the reform was to transfer financial resources and responsibilities to state and local governments for the provision of specific public goods. By 1998, five earmarked funds were created (one for basic education, one for health services, one for social infrastructure, one for municipal strengthening, and one for multiple destinations);[2] these were financed through federal transfers to sub-national governments.[3]

This paper focuses on one of these earmarked funds: the Health Services Fund[4] (FASSA, for its acronym in Spanish). Particularly, we analyze the consequences that such fund had over the health of the population according to specific health outcomes. We present results for infant mortality rate at the state level, a broadly used health indicator; but our results are robust to the use of other health indicators. The reform entitled the states to organize, control, coordinate, evaluate, and monitor the supply of health services, facilities and medical attention for the non-insured

[*] Research Economics Division, Banco de México, 5 de Mayo No. 18, Col. Centro, México, D.F., 06059, México. E-mail: amartinez@banxico.org.mx

[**] Corresponding author, Research Economics Division, Banco de México, 5 de Mayo No. 18, Col. Centro, México, D.F., 06059, México. Phone: +52 55 5237 2553, fax +52 55 52372571. E-mail: carolina.rodriguez@banxico.org.mx

The views expressed in this paper are those of the authors and do not necessarily reflect those of Banco de México. Karla Arroyo, David Camposeco, Daniel Leff and Mauricio Pérez provided superb research assistance. All errors are our own. We also thank seminar participants at Banco de México and at the 2011 Southern Economic Association Conference held in Washington (D.C.).

[1] However, the outcomes of fiscal federalism may be the opposite if political economy considerations are included in the analysis (Prud'homme, 2004 and Weingast, 2009).

[2] In 1999 two more funds were added: one for public safety and the other one for technological and adult education.

[3] It is important to address that the reform focused on changing the expenditures assignments between states and federation but it did not modify tax collection responsibilities among tiers of governments. Federal government is still responsible for collecting more than 90 per cent of the public revenue of the country, but after unconditional and earmarked federal transfers, sub-national governments spend around 50 per cent of the public expenditure in Mexico.

[4] In Spanish, Fondo de Aportaciones para Servicios de Salud (FASSA).

population[5] in the following areas: maternity care; visual and hearing health; nutrition; epidemiology; among others. In this context, FASSA's aim was to endow states with resources to meet the new health responsibilities that came with the decentralization of the sector. Decentralization meant that the medical attention of the non-insured (and therefore more vulnerable) population would now become the responsibility of state health authorities. Likewise, states were responsible for the administration of state hospitals for primary health care that used to be operated by the federal Ministry of Health (MofH hereafter) before the reform. One particular feature of the decentralization reform is that during the first years of its implementation, the amount of funds received by the states from FASSA was similar to what the federal MofH used to spend for non-insured population, via Ramo12, in each state before the reform took place. Another interesting feature is that the allocation of FASSA among states did not respond to the particular health needs of each state. These two facts, besides being clearly surprising, allow us to identify the impact on health indicators when health budget is spent by state governments rather than by the federal one.

We explore whether the decentralization of health provision in Mexico can account for the improvements of state level health indicators experienced in the last twenty years. First, we discuss whether the institutional arrangement of health decentralization is appropriate to maximize the impact of each peso spent. For instance, the Law of Fiscal Coordination determines a formula that specifies the factors used to calculate the share of FASSA assigned to each state, but does not present the weights given to each factor. Even more importantly, the factors determining what every state receives do not include health needs or rewards to those states that are spending efficiently. In order to address these issues, we present regressions that analyze the determinants of FASSA. Surprisingly, we find that the money spent by the federal government in each state in 1997, that is, the year before the reform was implemented, is the strongest predictor of what each state receives from the FASSA in any given year. We also found that health outcome variables, like infant mortality rate and deaths by infectious and parasitic diseases, do not show stable or significant coefficients. Medical resources are, in general, statistically insignificant, contrary to what the formula of FASSA stipulates. Population is the variable that more consistently shows a negative sign. We also perform similar regressions to look at the determinants of the non-insured health expenditure made by the federal government (Ramo12) before the reform. The results are very similar to the FASSA regressions and we conclude that the most important determinant driving health expenditure is the expenditure made in prior to decentralization.

The second part of our empirical strategy studies whether transferring health resources from the federal government to states has an effect on the infant mortality rate. For this purpose, we rely on different empirical exercises. We first compare FASSA to the federal budget on health, *i.e.*, Ramo12, by estimating the effect each budget had over the infant mortality rate for the years after the reform and for the years before the reform, respectively. This allows us to make a comparison between how state governments performed between 1998 and 2003 relative to how the federal government did between 1993 and 1997. The former exercise is an important comparison because the decentralization reform consisted in transfers of resources and responsibilities from the federal to state governments. We find no significant difference between the efficiency of Ramo12 and that of FASSA. Perhaps one reason we do not find a significant effect is that some states did very well whereas others underperformed, neutralizing the gains when averaging across states. Thus, in our second set of estimates, we test whether states that received more FASSA resources observed better health outcomes than those that received less resources when comparing the years after the reform with the years before the reform. Again, we find no significant difference between the high FASSA group relative to the low FASSA group. In another set of estimations that do not use infant

5 The non-insured is the fraction of the population that is not covered by an insurance mechanism; however they can access health care services at less than full-cost prices in Ministry of Health and state health facilities (OECD, 2005, pp. 29-30).

mortality rate but fetal death rate[6] and that take as control group that fraction of the population that is insured, we find that the fetal death rate among the non-insured population did not have a significant change after 1997 when compared to the fetal death rate in the insured population. However if we compare the expenditure efficiency (as measured by the effect of health expenditure on the infant mortality rate) for the non-insured with that of the insured population, we find that the former became more efficient after the decentralization reform. Thus, excluding the last specification, the evidence suggests that the decentralization of the health sector did not have an effect on the well-being of the population.

This paper has four main contributions. The first two are empirical ones. In the first place, this is the first work studying the effects of decentralizing the health sector in Mexico as well as the determinants of the distribution of health funds across states. Second, to the best of our knowledge, this paper is the only one that compares the efficiency in the provision of health services between the federal and state governments in two different federalist settings: centralized and decentralized. The other two contributions are related to the methodology. First, our identification strategy allows us to overcome some problems of endogeneity between decentralization and health outcomes, an issue seldom discussed in the literature. Finally, our measure of health decentralization is the actual health expenditure made by the state governments (from federal transfers), which, we consider, is a cleaner way to analyze efficiency issues relative to previous literature as we will discuss below.

The results of the present work may give important lessons about the conditions under which fiscal decentralization maximizes its impact on people's welfare. We argue that successful decentralization may be related to some necessary conditions: revenue collection decentralization, the strengthening of transparency and accountability of state governments, and improving institutional checks and balances.

The structure of the paper is as follows. Next section reviews previous literature related to health decentralization. The third section discusses briefly some characteristics of the health system in Mexico and the evolution of the main health indicators in the last two decades. The fourth section presents a description of the process of health decentralization and an analysis of how FASSA is allocated between states. The fifth part describes our empirical strategy followed by the analysis of the effects of decentralization on the infant mortality and fetal death rates. Finally, the paper concludes by discussing some lessons and plausible explanations for the (lack of) results of decentralization.

2 Literature review

Previous work on health decentralization has already pointed out the pros and cons of health provision by local state governments (see Asfaw *et al.*, 2007 and Robalino *et al.*, 2001 for a summary of these arguments). Among the advantages of decentralization the following can be listed: a) local authorities may have access to better information on local circumstances, needs and preferences of citizens; b) information is used more promptly and cuts costs without procedures that require central authorization, thereby enabling a more flexible operation of local governments; and c) it can also promote transparency, accountability, efficiency and community's participation. On the other hand, decentralization may hinder welfare gains due to: a) diseconomies of scale; b) lack of capacity, skills and information of local authorities on how to implement public policies; c) inability to collect own revenue to provide public goods; d) lack of interest from local elites in

[6] In this case, we did not use the infant mortality rate because we cannot divide it between non-insured and insured population. Due to the way fetal deaths are registered, it is possible to construct a fetal death rate for non-insured and insured population. The way we construct these rates is explained in detail in Section 5.

community's needs (capture of rents if there is no transparency and accountability); and e) implementation and coordination problems with national policies across regions.

Notwithstanding the importance of the topic, the empirical evidence on the consequences of decentralization is scarce. In the particular case of the health sector, previous literature has found that a more decentralized health sector is associated with a lower infant mortality rate, results that are opposite to our findings. Countries covered in this literature include India (Asfaw *et al.*, 2007), Argentina (Habibi *et al.*, 2001), China (Uchimura and Jütting, 2007), Canada (Jiménez Rubio, 2011), Spain (Cantarero and Pascual, 2008), Colombia (Soto *et al.*, 2011) and others included in a cross country study (Robalino *et al.*, 2001).

Nevertheless, this empirical research on the effects of decentralization has not provided compelling answers. First, it has had difficulties finding data on health spending by local governments. For instance, Asfaw *et al.* (2007), Robalino *et al.* (2001), Habibi *et al.* (2001), and Uchimura and Jütting (2007) use the proportion of total public expenditure or revenue that is spent or collected by provincial or sub-national governments as a measure of decentralization, even if such resources are used in sectors different than health. This indicator of decentralization clearly fails to deliver credible evidence about the real impact of decentralization in particular sectors, such as the health sector. Moreover, it is common that countries differ in the spheres that are decentralized. For instance, a country may have high local fiscal autonomy in many spheres but health, or it may be that the only type of decentralized expenditure is health (see Jiménez Rubio, 2011), which may lead to an identification problem of the relationship between health decentralization and outcomes. The only works that tackle this issue are Cantarero and Pascual (2008), Jiménez Rubio (2011), and, Soto *et al.* (2011) as they use a health specific decentralization indicator.

An additional issue of just using the percentage of health decentralized resources is that the estimations do not control for the level of health expenditure. This may lead to obtain biased estimates due to omitted variable issues if the share of sub-national resources is correlated to the level of health expenditure – Jiménez Rubio (2011) is an exception. In the absence of health expenditure in the econometric estimation, the results that find a negative relationship between decentralization and infant mortality rate may be capturing the effect of higher health expenditure (see, for instance, Joumard *et al.* (2008), which shows a positive effect of health expenditure on outcomes).[7]

Our paper solves both shortcomings by using the actual money spent by state governments in the health sector from transfers of the federal government as measure of health decentralization, which represents a high portion of health expenditure for non-insured population (around 80 per cent between 1997 and 2003).

Moreover, following Jiménez Rubio (2011), we consider it is important to control for other types of health expenditure (private, federal and social security institutions) that may be also driving health outcomes. The absence of these controls could confound the actual effect of greater local and state government's health expenditures. In order to deal with this issue the econometric estimation presented in Section V controls for a variety of health expenditure made by private and public institutions.

Methodologically, this paper deals with the issue of reverse causality between infant mortality rate and decentralization, a topic seldom discussed in the health decentralization literature. An advantage of this paper is that, for the case of Mexico, there is little evidence to

---

7 See also Mosca (2006) and Akin *et al.* (2005), which study the determinants of local health expenditures in Switzerland and Uganda, respectively.

support the hypothesis that the state assignment of decentralized resources is driven by health status, which allows us to have a clean identification strategy.

Finally, to the best of our knowledge, our paper distinguishes itself from previous literature as it is the only one that evaluates the effects on health of a reform that decentralized health provision from the federal government to state government. Therefore, we directly explore whether health state provision had better effects than the provision made by the federal government before the reform. In other words, we depart from the existent literature on health decentralization (which explores whether the degree of decentralization improves health outcomes) using a methodology that allows us to compare explicitly the performance of the health expenditure made by the federal government and state governments.

3 Mexican health system

3.1 Health institutions

The Mexican public health system is highly fragmented, with health services being provided by several institutions. Each institution is different in whether they provide care for the insured or non-insured population. "The insured receive care for free from providers belonging to their social insurance institution […][The] uninsured population, although not covered by an insurance mechanism, can still access health care services at markedly less than full-cost prices in publicly financed Ministry of Health and state health facilities"(OECD, 2005, pp. 29 and 30). Workers in the formal labor market and their families are covered by a set of social security institutions. Basically there are three types of public health insurance institutions: i) the Mexican Social Security Institute (IMSS for its Spanish acronym) provides services to 40 per cent of the population (private formal salaried workers and their families); ii) the Institute of Social Security and Services for Government Workers (ISSSTE) covers 9 per cent of the population (federal government workers and some state workers); and iii) others, which include social security systems for workers of the state-owned oil company (Petróleos Mexicanos, PEMEX), the Navy, the Army, among others, covering around 2 per cent of the population. These institutions are financed through tripartite contributions by the federal government (subsidies), the employer and, employees. Each institution has and operates its own set of clinics and hospitals and employs salaried doctors. The provision of health services is mandatory and there are no cost sharing mechanisms (OECD, 2005).

The responsibility to provide health care to those who do not have access to the social security system (less than half of the population) is shared by the MofH and state governments' health services. The rates charged for health services depend on the patient's income and varies among hospitals and states. The benefits include the provision of primary, secondary and tertiary care, as well as preventive and curative services, but services are subject to the availability of resources. Besides the rates charged, (a small portion of the non-insured expenditure) financing comes from the federal budget (Ramo12[8] and FASSA) and states' own resources (*participaciones*[9] and other own state income). In addition, numerous programs have been implemented in order to improve the access of non-insured and poor people to basic health services.

Finally, a minority of the population (around 3 per cent) has private health insurance (half

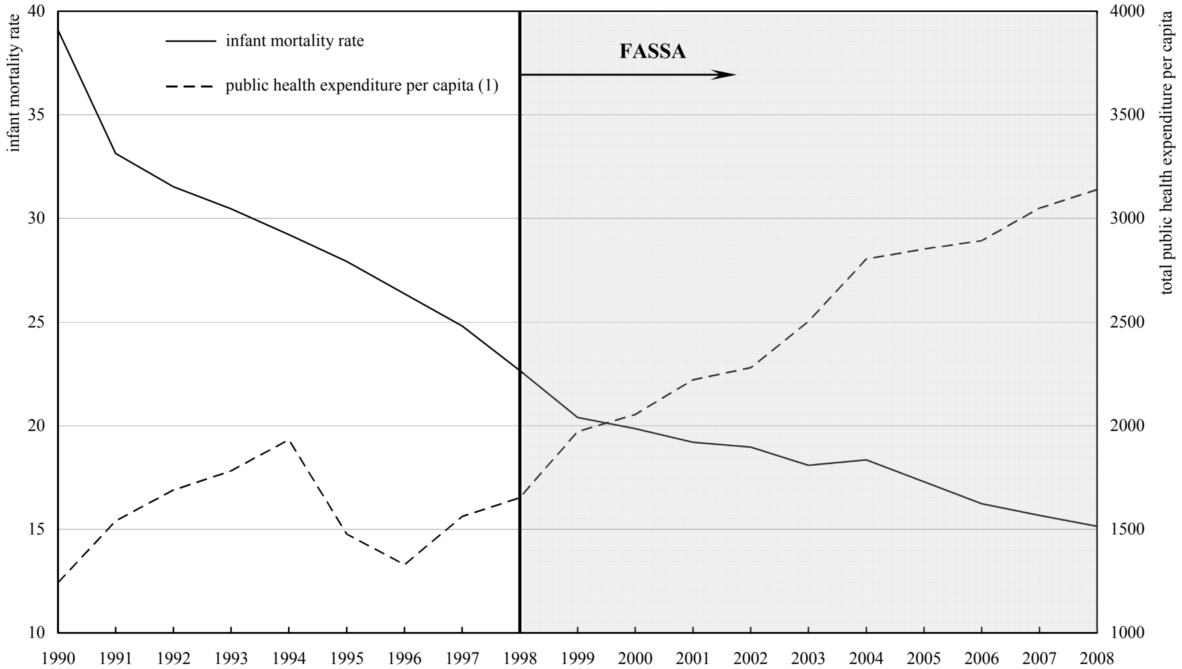
---

8 Ramo12 is the federal budget assigned for the provision of health services for the non-insured population. It includes the MofH budget, the health component of Oportunidades (an anti-poverty program based on conditional cash transfers), resources for public health programs and some resources for the Seguro Popular, the National Health Institutes and other large hospitals run by the federal government. IMSS-Oportunidades was previously financed through Ramo12 but these resources were directly transferred to the IMSS budget.

9 *Participaciones* are non-earmarked funds transferred from the federal government to state and local governments.

Figure 1

Infant Mortality Rate and Public Health Expenditure Per Capita, 1990-2008



(1) Units expressed in 2010 pesos.
Source: Own elaboration with data from SINAIS.

are financed by employers), which can be deduced from taxable income. There are two main types of private health policies: more than 97 per cent of the private insured population is covered through catastrophic medical insurance policies (*gastos médicos mayores*) for hospital expenses and various treatments for defined diagnoses; the remaining 3 per cent of the insured population on private institution has coverage through Products by Specialized Health Insurance Institutions (ISES), which is a "health care system that assumes or shares both the financial risks and delivery risks associated with providing comprehensive medical services to insured, usually in return for a fixed, prepaid fee" (OECD, 2005, p. 39). ISES offer full health coverage through private providers.

3.2 *Health financing: amounts and evolution*

Mexico spent 6.4 per cent of its GDP in health in 2009, up from 3.1 per cent in 1990. As of 2009, 48 per cent of the financing of the Mexican health system is public (up from 40 per cent in 1990).[10] As Figure 1 shows, the per capita public health expenditure more than doubled between 1990 and 2008. However, total and public health expenditure in Mexico is still the lowest among OECD countries, which on average spent 8.9 per cent of GDP in 2008. Most of the health expenditure in the OECD countries is financed by the public sector (72 per cent).

10 Private health expenditure is mostly (92.3 per cent) done in the form of out-of-pocket payments. Within out-of-pocket expenditures, only a minor fraction is due to public sector cost sharing schemes. Most of the out-of-pocket is spent in the private sector. Just to have a perspective, OECD countries spend around 18.9 per cent of the total expenditure in out-of-pocket payments, versus almost 50 per cent in Mexico.

Covering around half of the population, social security institutions (IMSS, ISSSTE and PEMEX) were responsible of more than 80 per cent of the public health expenditure in 1993 and around two thirds in 2003. In 1993, Ramo12 represented 13.02 per cent of the overall public expenditure on health (0.33 per cent of GDP)[11] and in 2003 its participation decreased to 9.17 per cent of total health expenditure (0.26 per cent of GDP). While state governments (without FASSA)[12] had a share of around 5 per cent of health public expenditure[13] in 2003, FASSA represented about 16.8 per cent of the public health expenditure (0.47 per cent of GDP).

The growth in public health expenditure came along with a deeper penetration of health services in Mexico. Coverage has improved in the last years, as physicians per 1000 people went from 1.06 in 1990 to 1.44 in 2003 and nurses per 1000 increased from 1.55 to 1.76 between 1990 and 2003. Medical consultations also showed an important increase: in 1990, there were 1195 consultations per 1000 people; 13 years later, this indicator grew to 1726. Although these numbers show improvements over the last decade, Mexico still has one of the lowest health coverage among OECD countries.[14]

The expansion in health resources was translated into important progress in health status over the last twenty years. For instance, life expectancy at birth in 2008 was 75 years, up from 70 years in 1990; infant mortality rate went from 39 deaths per 1000 live births (see Figure 1) in 1990 to 15.2 deaths. As these numbers suggest, Mexico experienced great improvements in health but there is still some gap with respect to OECD countries.[15]

Historically, regional differences in health indicators have been important but the progress observed in the last years favored poor states as they have closed the gap. For instance, the state with the highest infant mortality rate in 1990 was Chiapas with 60.72 and Federal District had the lowest (22.36). Thirteen years later, Guerrero had the highest infant mortality rate (25.89) and Nuevo León had the lowest (12.44).

In spite of the recent achievements in health, Mexico still faces important challenges (OECD, 2005). The government has limited economic resources to deal with the demographic and epidemiological (from infectious to degenerative diseases) transition that will increase the demand for health care in the near future. An institutional reform is needed to avoid the current fragmentation of the various social security structures which has led to an inefficient provision of health care as well as to overcome the disparities in health expenditure among several dimensions such as: across states, between social security institutions and the non-insured population, and between federal and state governments. Moreover, it is important to minimize the out-of-pocket expenditure and to increase infrastructure and equipment investment in the sector (Gómez Dantés and Ortiz, 2004).

4 Decentralization and FASSA

4.1 Evolution of Health Decentralization in Mexico

In the last three decades, Mexico undertook two waves of health decentralization, mainly for the coverage of non-insured population. The first wave was in the 1980s but it was not generalized

---

11 For the calculations before 1998, it is noteworthy that there is no available data for state governments' expenditure.

12 Those resources come from own state resources and non-earmarked transfers from the Federation to states.

13 State governments made an effort equivalent to 8 per cent of the all public sector effort in 2008.

14 According to OECD data, Mexico had 2 doctors per 1000 population in 2008 and the OECD average was 3. The number of nurses per 1000 population averaged almost 9 in the OECD countries; Mexico had 2.4 nurses. Finally, doctor consultations per capita in Mexico were 2.8 compared to 7.1 among OECD countries.

15 OECD life expectancy is 79 years old and infant mortality rate is 4.6 deaths per 1000 live births.

since only 14 states[16] out of 32 signed the agreement with the federal government. Although the program included the transfer of responsibilities to states for the operation of some hospitals and administrative tasks and the consolidation of the services provided by IMSS-Coplamar[17] and the MofH, the spending decisions, regulation and policy formulation remained controlled by the MofH (see Cabrero and Martínez Vázquez, 2000 and Merino, 2003). According to Birn (1999), the provision of health services and health outcomes from this attempt of decentralization were not different between the signers and non-signers of the health decentralization agreement of the '80s.

After some minor decentralizing actions during the administration of President Salinas (1988-94) (see Merino, 2003), a comprehensive decentralization reform was launched in 1996 as part of an important strategy of the Health Sector Reform Program 1995-2000. Centralism in the sector was seen as a cause of several problems such as low efficiency in the allocation of resources; lack of clarity in the responsibilities of each tier of government, excessive bureaucracy, inertia and inequality in the distribution of resources among states and absence of coordination between IMSS-Solidaridad,[18] the MofH and state health authorities (Merino, 2003). In order to tackle these issues, the reform defined clearly the health responsibilities of federal and state governments.[19] The federal government transferred operative functions, along with human, physical and monetary resources to states, thereby providing them with greater autonomy. Former employees of the federal MofH became part of state health units. Although the reform of the 1990s was deeper than the one implemented in the 1980s, Merino (2003) argues that the implementation of health decentralization was uniform across states without taking into account differences in administrative capacity, willingness to take the transfer of responsibilities or characteristics of population, services and geography, among others.

In order to meet their new responsibilities, states were endowed with FASSA, a fund that was created along with others in the context of a federalist reform in 1997. FASSA is a fund that transfers federal resources to states for health provision; it must be spent exclusively on health services for the non-insured population. FASSA represents the main source of financing for states as 77 and 64 per cent of the states' health expenditures came from this federal fund in 1998 and 2009, respectively.[20] Although FASSA is distributed among states according to criteria such as health infrastructure, health service workers, the budget assigned the previous year and a component that is aimed to equalize health accessibility,[21] the law does not set the weight of each component or the total amount allocated to the fund. Hence, the law does not establish a clear criterion for its distribution, allowing discretionary decisions by legislators and the federal government. Further, the resources obtained by every state were based on the amount originally

---

16 Tlaxcala, Nuevo León, Guerrero, Jalisco, Baja California Sur, Morelos, Tabasco, Querétaro, Sonora, Colima, Estado de México, Guanajuato, Aguascalientes and Quintana Roo. Note that, on average, these states are more industrialized, have less population dispersion, and have fewer nutrition, health and education problems.

17 Coplamar stands for "General Coordination of the National Plan for Depressed Zones and Marginalized Groups", which was a social programs implemented in the seventies.

18 This is a poverty program implemented during the Presidency of Salinas (1988-94).

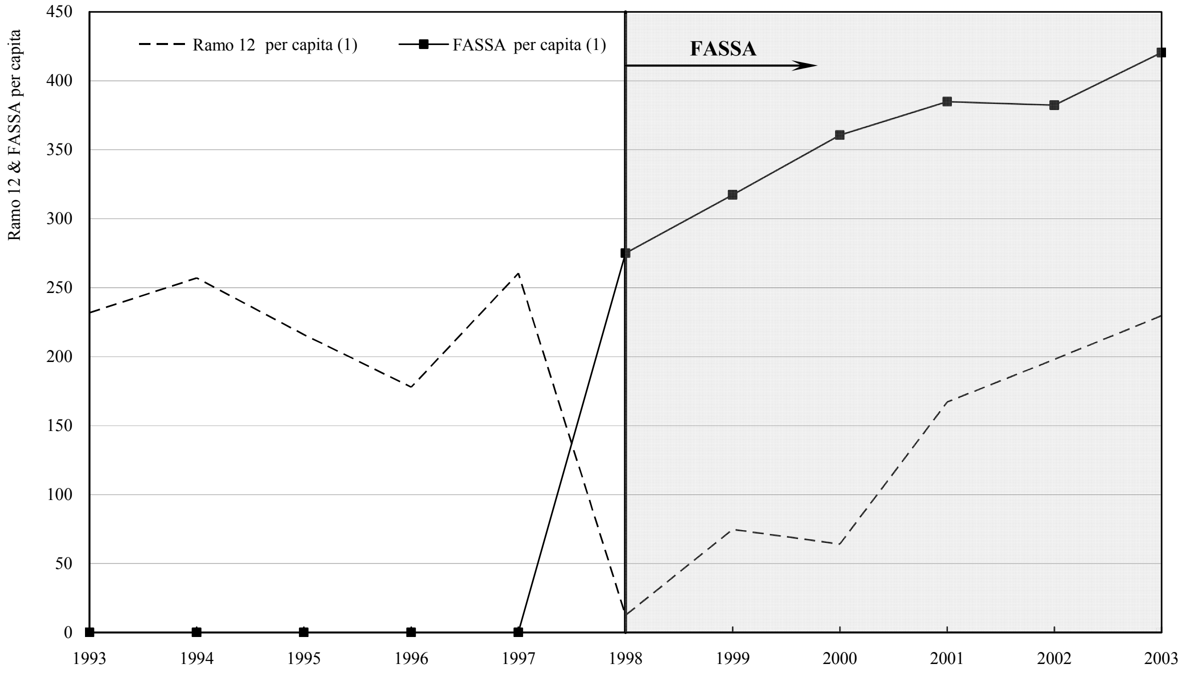
19 Articles 3rd, 13th and 18th of the Health Law establish the responsibilities of both levels of governments. In short, states are in charge of the organization and operation of health establishments and services, prevention of contagious diseases, maternity child care, nutrition, visual and auditive health, among others. The federation, in turn, operates most of the secondary and tertiary hospitals; designs health regulation and policies; watches the use of economic resources, deals with labor relations of the non-insured system, and takes mayor investment decisions.

20 Merino (2003) considers that the high dependence of states on transfers is not optimal for health provision as they have little flexibility to make adjustments to respond to their needs. Moreover, states may limit their health expenditures if they believe that a higher effort would be seen as a lower need for resources and thus less transfers from the federal government.

21 This component receives the remaining of the total budget of FASSA, which represents a low share. For instance, in 2001 its allocation was of only 100 million pesos when the overall FASSA budget was around 25,000 million pesos. The distribution of this component among states has a formula established in the Law and depends on the non-insured population, mortality, marginalization and federal budget (article 31 of the Fiscal Coordination Law). This is the only formula for FASSA in the Law.

Figure 2

Ramo12 and FASSA Per Capita
(national level)



(1) Units expressed in 2010 pesos.
Source: Own elaboration with data from SINAIS and the Ministry of Health.

agreed between the federal government and states in 1997 (Sour *et al.*, 2004), which depended on the expenditure made by the Ministry of Health before decentralization (Merino, 2003).

In fact, FASSA allocation between states in its first year of operation (1998) was very similar to the allocation of the MofH budget in 1997. Later, during the first years of the reform, federal expenditure to states was reduced considerably (see Figure 2). In 1997 MofH distributed resources to states equivalent to 0.34 per cent of GDP while in 1998 the number dropped to 0.02 per cent with 14 states not receiving any resources. In contrast, FASSA budget in 1998 was equal to 0.39 per cent of GDP. We next show the MofH budget for each state in 1997 is a good predictor of FASSA in any given year, suggesting that the fund has a strong inertial component.[22]

4.2 What explains FASSA allocation among states?

In this section we provide some empirical evidence on the determinants of expenditure allocation among states for the non-insured population (Ramo12 before 1998 and FASSA after 1997). First, we present the descriptive statistics of this exercise. After which we proceed to describe the empirical strategy and its results.

22 After 2004, the nature of FASSA changed because it was used by the federal government to finance the operation of a program called Popular Insurance (Seguro Popular) under different expenditure rules. For this reason the analysis of this paper stops in that year.

Table 1

Summary Statistics

Variables	Panel A – Ramo12 (1993-97)				Panel B – FASSA (1998-2003)			
	Mean	Std. Dev.	Min.	Max	Mean	Std. Dev.	Min.	Max
Ramo12	278.77	116.93	100.82	724.83	-	-	-	-
Ramo12 from 1992	253.94	100.6	108.34	583.68	-	-	-	-
FASSA	-	-	-	-	438.36	176.95	178.79	1034.61
Ramo12 from 1997	-	-	-	-	310.96	119.53	173.37	724.83
Infant Mortality Rate	27.51	4.89	16.59	40.87	19.56	3.97	12.44	32.86
DIP	0.25	0.09	0.12	0.73	0.2	0.06	0.09	0.42
DNIP	1.36	1.05	0.51	10.18	1.41	0.61	0.64	3.83
PUP	0.47	0.15	0.15	0.78	0.5	0.14	0.22	0.8
Pop	2.86	2.44	0.35	12.11	3.09	2.63	0.41	13.59
GSP	66.12	31.61	26.76	185.65	76.36	36.35	28.46	213.92
Number of observations	160				192			

Note: The definition and units of the variables are in Table 2.

4.2.1 Descriptive statistics

Table 1 shows the descriptive statistics for the two set of regressions: Ramo12 (1993-97) and FASSA (1998-2003) in per capita terms. The definition, corresponding acronym, units of measure and source for each of these variables is included in Table 2. We use one-year lagged covariates because health budget is allocated at the end of the previous year, when legislators approve the federal budget.

The dependent variables, Ramo12 and FASSA, are on average 279 and 438 pesos per person, respectively (see Table 1). The potential explanatory variables for the non-insured population are some proxies for health needs, resources, and socioeconomic variables. First, we include the infant mortality rate (the sample average is of 27.6 and 19.6 deaths of children younger than 1 year per 1000 live births in the pre and post reform years) and the infectious and parasitic mortality rate which is denoted as DIP_{it} (0.25 and 0.2 deaths per 1000 inhabitants, respectively).[23]

Second, according the Law of Fiscal Coordination, FASSA allocation should be partly determined by the physical and medical infrastructure available in each state. In order to control for these elements, we include total number of doctors assigned for the non-insured population in each state per 1000 non-insured individuals which is represented as $DNIP_{it}$ (1.36 and 1.41 doctors

[23] We also collected other variables like deaths by maternal causes, fetal deaths, deaths by conditions originated in the perinatal period, deaths by diabetes, and deaths by nutritional deficiencies, among others. We do not include these variables as regressors because many of them are highly correlated. However, the results are robust to the use of one specific variable instead of another.

Table 2

Definition of Variables

Variable	Definition	Units	Source
DIP_{it}	Deaths by infectious and parasitic diseases for state *i* and year *t*	Per 1000 inhabitants by state	Ministry of Health
$DNIP_{it}$	Doctors for non-insured population for state *i* and year *t*	Per 1000 inhabitants non-insured	SINAIS
DP_{it}	Population Density for state *i* and year *t*	Inhabitants per Km^2	INEGI
$FASSA_{it}$	Health services fund for state *i* and year *t*	Thousand pesos per capita	Ministry of Health
GSP_{it}	Gross state product for state *i* and year *t*	Thousand pesos per capita (2nd half dec 2010=100)	INEGI
$HBPS_{it}$	Hospital beds in the private health sector for state *i* and year *t*	Per 1000 inhabitants by state	SINAIS
$HEEP_{it}$	Health services expenditure from public institutions (IMSS, ISSSTE,PEMEX) for state *i* and year *t*	Thousand pesos per capita (2nd half dec 2010=100)	Ministry of Health
$I(t>1997)$	Is an indicator function that takes the value of zero before the reform was implemented and one after the reform	N.A.	N.A.
IMR_{it}	Natural logarithm of the infant mortality rate for state *i* and year *t*	Number of deaths of children less than one year old per 1000 live births by state	UN Millennium Development Goals
$IMR_{Biased,\ it}$	Natural logarithm of the infant mortality rate for state *i* and year *t*	Per 1000 live births by state	SINAIS
$IMR_{Ratio,\ it}$	$\log(IMR_{it}) - \log(IMR_{Biased,\ it})$	N.A.	N.A.
FDR_{ijt}	Natural logarithm of fetal deaths for state *i*, year *t*, and group *j* divided by population in state *i*, year *t*, and group *j*[1]	Per 100 insured or non-insured population	INEGI
$Ramo12_{it}$	Federal government directly spend on health services for state *i* and year *t*	Thousand pesos per capita	SINAIS
Pop_{it}	Total population for state *i* and year *t*	Total number of inhabitants per state	CONAPO
$PSCR_{it}$	Percentage of students who completed primary school in 6 years for state *i* and year *t*	Percentage	UN Millennium Development Goals
PUP_{it}	Proportion of non-insured population for state *i* and *t*	Between zero and one	Ministry of Health
THE_{ijt}	Total health expenditure for state *i*, year *t* and group *j* divided by population for state *i*, year *t* and group *j*[1]	Thousand pesos per insured or non-insured population	Ministry of Health

[1] *j* is insured or non-insured group.

Sources: National Population Council (CONAPO), Bureau of Health Information in Mexico (SINAIS), National Institute of Statistics, Geography and Informatics of Mexico (INEGI) and United Nations (UN) Millennium Development Goals Statistics.

before and after 1998).[24] Third, we also include socioeconomic variables such as the annual gross state product per capita (GSP_{it}); the ratio of the non-insured population over the total population, denoted as PUP (47 and 50 per cent), and total population, represented as Pop_{it} (2.9 and 3.1 millions).[25]

Finally, according to the Law of Fiscal Coordination, the allocation of FASSA also depends on the resources received in the previous year. In fact, when the FASSA started to operate, the allocation of such resources among states crucially depended on what the federal government directly spent on each state in 1997 through centralized resources, *i.e.*, Ramo12. This means that as of today, the allocation of FASSA between states still depends on what each state received in 1997 from Ramo12. For this reason, we add the amount of resources that each state received in 1997 through Ramo12 as a regressor. On average, this variable is 311 pesos per capita. Following the same logic in Ramo12 per capita regressions, we include Ramo12 per capita in 1992 (the state average of this variable was of 254 pesos per capita).

4.2.2 Health expenditure 1993-2003

Our empirical strategy aims to unveil the key determinants of the state allocation of non-insured health expenditure: Ramo12 for the previous years of the reform of 1997 and FASSA for the 1998-2003 period in order to check if there was a change in the criteria of assignation once decentralization took place.

For each period (before and after 1997), we run two sets of regressions on state level data. The first one is a pooled data approach, in which we regress per capita FASSA (and Ramo12) flows received by state i in year t in constant pesos, on a set of covariates that presumably determines the amount of resources that each state receives in a specific year. We include year dummies to the specification to control for aggregate time effects. In this estimation, we add a time-invariant regressor: the federal budget on health in 1997 (in 1992 for Ramo12 specifications) because we want to see how important this inertial component is for FASSA allocation, as some authors have suggested. We also include a state fixed effects estimations (removing the Ramo12 per capita component) in order to check whether our results hold under this alternative specification. The second set of estimations are cross section regressions for 1998 and 2003 (results are consistent for 1999, 2000, 2001 and 2002) as we are interested to analyze the criteria of individual years of the Federal Congress in the assignment of FASSA for all the period. We also run a similar set of regressions for the Ramo12 per capita before the reform (between 1993 and 1997) as we want to analyze whether its allocation is correlated to variables that indirectly could be affecting FASSA.

4.2.3 Results

The results for the determinants of FASSA and Ramo12 per capita are shown in Table 3. The results show a strong inertial component for health expenditure, as the coefficient of Ramo12 of 1992 and 1997 is significant at 1 per cent level (specifications 1, 3, 5, 7, 9 and 11). For instance, specification 7 shows that for every peso per capita that every state received from FASSA in 1997,

---

[24] We also try other variables including the number of non-insured medical offices and appointments; number of dentists, number of nurses, and number of hospital beds of the Ministry of Health. As before, we do not include these variables as regressors because many of them are highly correlated. However, the results are robust to the inclusion of one of these variables instead of the one included in the specification.

[25] Education was also included in some specifications and the results remain unchanged.

Table 3

Ramo12 and FASSA Determinants

Independent Variables	Panel A – (1993-97)						Panel B – (1998-2003)					
	Dependent Variable is Ramo12 Per Capita						Dependent Variable is FASSA Per Capita					
	Panel Data		Cross Section				Panel Data		Cross Section			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	1993 to 1997	1993 to 1997	1993	1993	1997	1997	1998 to 2003	1998 to 2003	1998	1998	2003	2003
$Ramo12_{i\ 1992}$	0.963***	-	1.083***	-	0.909***	-	-	-	-	-	-	-
	(0.0461)	-	*(0.0582)*	-	*(0.122)*	-	-	-	-	-	-	-
$Ramo12_{i\ 1997}$	-	-	-	-	-	-	1.329***	-	1.401***	-	1.279***	-
	-	-	-	-	-	-	*(0.111)*	-	*(0.0891)*	-	*(0.191)*	-
IMR_{it-1}	0.354	–19.88	–0.571	17.87*	–2.163	11.77	–1.837	–18.11***	–3.637	8.341	7.228	13.52
	(2.638)	*(12.84)*	*(1.81)*	*(8.862)*	*(9.687)*	*(7.554)*	*(2.585)*	*(5.353*	*(2.185)*	*(11.3)*	*(5.234)*	*(15.47)*
DIP_{it-1}	78.12	134.7	56.57	–62.07	–170.6	–419.6	–172.7*	252.6	73.6	–351.5	–319.6**	–639.5
	(51)	*(136.1)*	*(47.42)*	*(158.2)*	*(163.2)*	*(258)*	*(92.49)*	*(211.4)*	*(59.97)*	*(314.7)*	*(146.1)*	*(413.5)*
$DNIP_{it-1}$	6.022	20.72	6.939	64.36	37.94**	89.14*	–7.265	9.011**	–13.19***	22.78	–4.81	145.8**
	(6.964)	*(25.97)*	*(6.42)*	*(43.65)*	*(14.15)*	*(45.2)*	*(4.723)*	*(3.655)*	*(4.709)*	*(23.18)*	*(27.9)*	*(64.1)*
PUP_{it-1}	–4.082	2.082**	–37.91	–126	207	120.9	–89.59	241.2	–208.7*	–310.8	–152.1	80.43
	(61.76)	*(880.3)*	*(64.34)*	*(240.1)*	*(197.7)*	*(288.9)*	*(92.46)*	*(326.7)*	*(105.2)*	*(479.3)*	*(135.8)*	*(389.8)*
Pop_{it-1}	–5.171***	77.86	–3.230*	–24.71**	–5.894*	–21.49**	–4.383*	–45.06	2.182	–34.39**	–7.270*	–30.19*
	(1.874)	*(46.21)*	*(1.731)*	*(11.63)*	*(3.274)*	*(10.04)*	*(2.541)*	*(35.56)*	*(2.172)*	*(16.13)*	*(3.923)*	*(15.43)*
GSP_{it}	–0.0607	0.378	–0.540**	0.201	0.643	0.956	–1.046**	0.986	–1.179**	–0.361	–0.889	–0.329
	(0.2)	*(0.763)*	*(0.224)*	*(0.875)*	*(0.437)*	*(0.926)*	*(0.432)*	*(0.63)*	*(0.45)*	*(1.684)*	*(0.734)*	*(1.752)*
Constant	17.36	–293.3	63.33	–235.6	1.177	–71.49	274.0***	715.1***	184.2***	467.3	218.4*	251.2
	(66.89)	*(401.1)*	*(52.5)*	*(232.8)*	*(158.2)*	*(144.4)*	*(61.69)*	*(202.3)*	*(49.51)*	*(309.4)*	*(117.2)*	*(330.9)*
Year Indicators	Yes	Yes	-	-	-	-	Yes	Yes	-	-	-	-
Fixed Effects	No	Yes	-	-	-	-	No	Yes	-	-	-	-
R^2	0.878	0.565	0.958	0.5	0.847	0.502	0.934	0.805	0.978	0.351	0.923	0.528
Observations	160	160	32	32	32	32	192	192	32	32	32	32

Panel data estimations show state cluster robust standard errors in parentheses & cross section estimations show robust standard errors in parentheses.
Note: The definition and units of the variables are in Table 2. Significance interpretation is as follows: *** p<0.01, ** p<0.05, * p<0.10.

it will get from FASSA 1.33 pesos on average in the 1998-2003 period. The effect is statistically significant at 1 per cent level.

This result remains unchanged in the cross section specifications (3, 5, 9 and 11): the inertial component is crucial for the allocation of health public expenditure for the non-insured population. Probably this result should not be a surprise because there is persistence on health outcomes and resources over time and the initial allocation of expenditure might be capturing the effect of initial outcomes. However, we believe that health outcomes (such as infant mortality rate) should matter independently in how health expenditure was allocated in past years, even if that allocation depended on past health indicators. In this sense, we do not find consistency in the signs and significance of the different potential explanatory variables (even though they are explicitly contained in FASSA's formula) across the different regressions. This result suggests that legislators assign health budget exclusively taking into account the previous year's allocation but no other health fundamentals. The only variable that seems to be consistent in the significance and magnitude is Pop_{it-1}. The sign is negative, implying that more populous states obtained lower health transfers. It could be thought that this sign is due to its correlation with other variables. For instance, it is plausible that a state with high mortality has restricted access to health facilities that are negatively correlated to DP_{it-1}. However, discarding Pop_{it-1} as an explanatory variable does not change our results.

In particular, IMR_{t-1} and DIP_{t-1} yield no significant estimates in most of the cases. In some specifications they even have an opposite expected sign. The result would indicate that states with high health needs would receive fewer resources from FASSA, suggesting a regressive distribution allocation of the health budget.

With respect to the variable related to medical infrastructure ($DNIP_{it-1}$), the coefficient is positive for Ramo12 per capita but only the regressions for 1997 (columns 5 and 6) are significant. Interestingly, for FASSA per capita regressions without Ramo12 per capita for 1997 included, the results for medical infrastructure are positive and significant for the fixed effects and 2003 regressions (columns 8 and 12), which could be related to the FASSA allocation formula stated in the Law of Fiscal Coordination.

Finally, in few specifications, state GDP shows a negative and significant coefficient, indicating that there is some redistributive element in FASSA. However, this result is not consistent across the different specifications. It is surprising that the proportion of non-insured population is not significant because it is precisely the population that should be targeted by non-insured expenditure (either Ramo12 or FASSA).

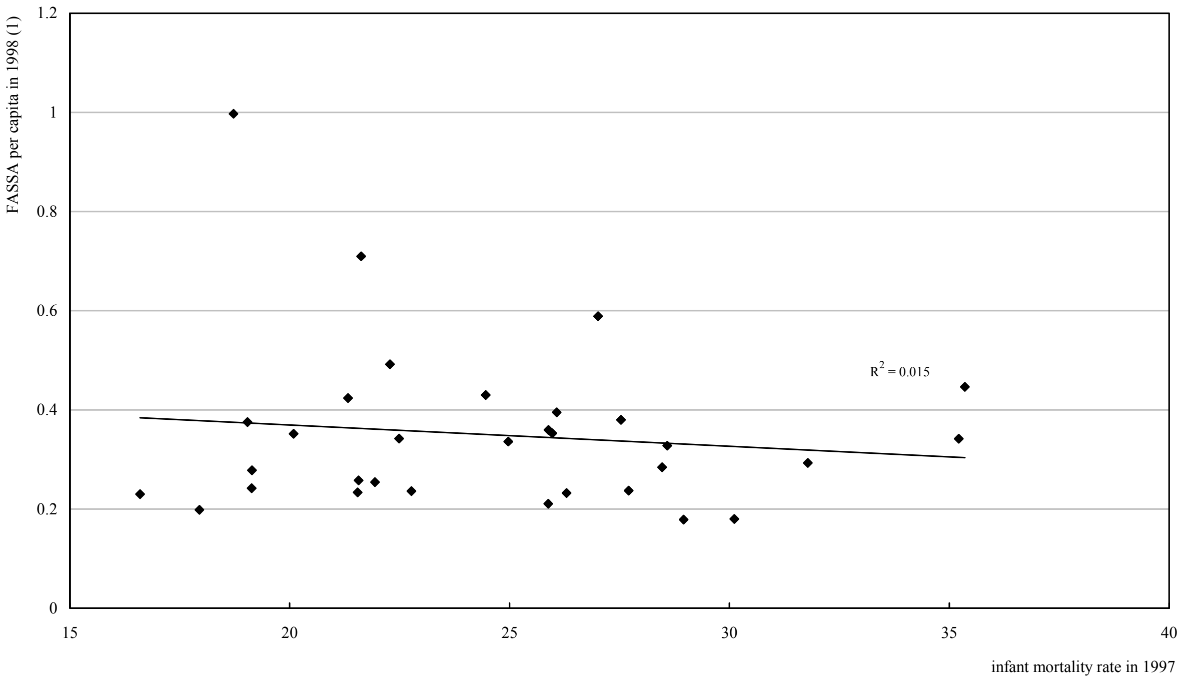
In sum, the results indicate that health outcomes (and other variables) do not determine how the resources are allocated. Our regressions suggest that the most important determinant of state non-insured expenditure is the past allocation. This finding is critical for our empirical strategy for the consequences of decentralization, as we do not have any evidence that FASSA is endogenously allocated as a result of health outcomes. So we are confident that, in particular, infant mortality rate is exogenous to how FASSA is determined (see Figure 3).

5 Does decentralization of resources for health services improve state-level health outcomes?

In this section we test, through different estimation procedures and specifications, whether the decentralization of resources for health services improve state-level health outcomes. First, we test whether state health outcomes improved in the years after the implementation of FASSA relative to how Ramo12 did in the years previous the reform. We find no significant difference

Figure 3

Infant Mortality Rate in 1997 vs. FASSA Per Capita in 1998



$$\underset{}{FASSA\ Per\ Capita_{1998}} = \underset{(0.158)}{0.455} - \underset{(0.006)}{0.004}\ IMR_{1997}$$

(1) Units expressed in 2010 pesos.
Note: Standard errors are shown in parenthesis.
Source: Own elaboration with data from the Ministry of Health and UN Millennium Development Goals.

between the effectiveness of Ramo12 and FASSA. Second, we test whether states that received more FASSA resources observed better health outcomes than low FASSA states after the reform. Again, we find no significant difference. Third, we test whether there is a difference between state health outcomes of the uninsured relative to the insured population after the implementation of the reform. Since Ramo12 and FASSA focus on the non-insured population, we took the insured population as a control group. We find, as before, no significant difference between health improvements observed after the implementation of the reform among the treatment and control groups. Finally, focusing on expenditure amounts, we test whether FASSA and Ramo12, which focus on the non-insured population, between the years before the reform (1993-97) and the years after the reform was implemented (1998-2003) is more efficient than the health expenditure for the insured population.

Contrary to all previous results, we find that in fact FASSA and Ramo12 together are more effective than the IMSS, ISSSTE or PEMEX in reducing fetal deaths.

5.1 Summary statistics

Before presenting the final results, we briefly summarize the main variables used in this section. In Table 4 we show the summary statistics of these variables used by pooling the data from

Table 4

Summary Statistics 1993-2003

Variables	Mean	Std. Dev.	Min	Max
DP	266.3	1003	4.78	5920
Fetal death rate	0.262	0.121	0.026	0.783
Log(Fetal deaths)	–1.47	0.566	–3.666	–0.244
Fetal death rate for the non-insured population	0.304	0.130	0.035	0.783
Log(Fetal deaths) for the non-insured population	–1.307	0.538	–3.352	–0.244
Fetal death rate for the insured population	0.220	0.094	0.026	0.522
Log(Fetal deaths) for the insured population	–1.634	0.547	–3.666	–0.65
GSP	71.7	34.61	26.75	213.9
HBPS	0.297	0.132	0.082	0.832
HEEP	2.663	1.03	1.173	9.384
Log(infant mortality rate)	3.11	0.255	2.521	3.71
PSCR	85.52	9.185	43.42	99.16
PUP	0.49	0.148	0.148	0.798
Ramo12	0.19	0.144	0	0.725
THE	1.805	1.196	0.167	9.384
Log(*THE*)	0.355	0.736	–1.792	2.239
THE for the non-insured population	0.946	0.567	0.167	3.356
Log(*THE*) for the non-insured population	–0.218	0.577	–1.792	1.211
THE for the insured population	2.664	1.031	1.173	9.384
Log(*THE*) for the insured population	0.928	0.305	0.16	2.239

Total number of observations is 352 for all variables with exception of total health expenditure, fetal deaths and its logarithmic function which have 704 observations due the distinction between non-insured and insured population.
Note: The definition and units of the variables are in Table 2.

1993 through 2003. We follow the literature using as our preferred health status variable, infant mortality rate (deaths of babies younger than 1 year old divided by life births). According to summary statistics, the natural log of the infant mortality rate is on average 3.11, that is, approximately 22 infant deaths per thousand births among all states and years. There are various reasons we focus on IMR_{it} as our main dependent variable. Infant mortality rate is a good health outcome measure as it reflects health attention to sensitive care groups of population (children and pregnant women); it is also known that it responds rapidly to changes in the health systems (Jiménez Rubio, 2011); it is better measured than other indicators such as life expectancy; and is correlated with many other health indicators (Joumard *et al.*, 2008; and Jiménez Rubio, 2011). The other variable we use as measure of state health status is total fetal death rate. As shown in Table 4, the natural log of total fetal deaths (FDR_{it}) averages –1.470, that is, about 0.26 fetal deaths per thousand individuals. The main advantage of this variable relative to IMR_{it} is that we can obtain the fetal death rate for non-insured and insured population, respectively. According to summary statistics, for the non-insured population fetal death rate averages around 0.30 fetal deaths per thousand non-insured individuals. For insured population, there are on average 0.22 fetal deaths per thousand insured persons.

Continuing with the variables summarized in Table 4, $Ramo12_{it}$ is on average 190 pesos per capita between 1993 and 2003. The variable $FASSA_{it}$ averages 438 pesos per capita for the years after its implementation (see Table 1). Gross state product per capita (GSP_{it}) in constant pesos is on average 71,707 pesos. Population density (PD_{it}) is around 266 persons per squared kilometer on average.

The average expenditure by IMSS, ISSSTE and PEMEX is 2663 pesos per eligible person ($HEEP_{it}$). The proportion of uninsured population (PUP_{it}) over the total population per state is on average 0.49. The primary school completion rate ($PSCR_{it}$), a measure of schooling, is on average 85 per cent. We do not observe out-of-pocket expenditure on health services by the population for years before 1998. However, on average, there are 0.29 hospital beds in the private sector per 1000 inhabitants ($HBPS_{it}$).

5.2 *What was the impact on state health outcomes of FASSA relative to Ramo12?*

In this section we test whether state health outcomes improved in the years after the implementation of FASSA relative to how Ramo12 did in the years previous the reform. This is a way to test whether decentralizing resources from the federal to the state government improved the health of the population. Recall that before 1998 the resources for health services were channeled through Ramo12 and the federal government was responsible of their use in each state. After 1997, FASSA was created to channel those same health resources to states and now state governments are responsible of the administration of such budget. The empirical specification is the following:

$$\begin{aligned} IMR_{it} = \alpha + \beta_1 I\ (t > 1997) + \beta_2\ (Ramo12_{it}) + \beta_2\ [I(t > 1997) * (Ramo12_{it})] + \\ + \beta_4[I(t > 1997) * (FASSA_{it})] + X_{it}B_5 + c_i + u_{it} \end{aligned} \tag{1}$$

$$I = 1, \ldots 32 \qquad\qquad t = 1, \ldots 11$$

In equation (1), IMR_{it} is the natural logarithm of the infant mortality rate in state i and year t; $I(t>1997)$ is an indicator function that takes value zero for the years before the reform was implemented and one after the reform; $Ramo12_{it}$ is the amount of resources per capita directly spent by the federal government for health services in state i and year t; $FASSA_{it}$ is the amount of decentralized resources per capita for health services provision in state i and year t after 1997; X_{it} refers to a vector of control variables which are described below; c_i denotes the state fixed effect which is assumed to be arbitrarily correlated with the regressors; and u_{it} denotes the idiosyncratic error for state i in year t. There are 32 states in Mexico and the analysis covers eleven years, from 1993 through 2003.

Notice that $FASSA_{it}$ enters only as an interaction with the reform-years indicator, *i.e.*, $I(t>1997)$. This is because FASSA was implemented in 1998 and thus it takes value zero for years before 1998. In contrast, $Ramo12_{it}$ operates both before and after the decentralization reform. $Ramo12_{it}$ appears by itself and as interaction with the reform-years indicator. Also, notice that β_2 is the effect of $Ramo12_{it}$ over the IMR_{it} in the years before the reform and β_4 is the effect of $FASSA_{it}$ on the IMR_{it} in the years after the reform. Thus, our interest is in β_4–β_2. We expect this difference to be negative. However, we also need this difference to be significant to be able to conclude that the decentralization improved health outcome of the population. If β_4–β_2 turns out to be not significant, even if it has the correct sign, it implies that there is no significant difference between what central government was doing with the money and what state governments do with the same resources.

Equation (1) also permits us to test whether the money spent on health services by state governments improves the IMR relative to the money spent by the federal government for the same purpose but considering both effects in the years after 1997, that is, after the decentralization reform took place. In this case our interest is in $\beta_4-(\beta_2+\beta_3)$. If this difference is negative it implies that FASSA is more efficient than Ramo12. However, regardless of the sign, if $\beta_4-(\beta_2+\beta_3)$ is not significant, we can only say that there is no difference between the two funds after the reform.

There are other variables besides $FASSA_{it}$ and $Ramo12_{it}$ that could explain the IMR_{it}. For this reason, we include different control variables in the specification equation (X_{it}). We include gross state product per capita (GSP_{it}) to control for level of income. We also try to control for the average distance between health facilities and the inhabitants by including population density (PD_{it}) as control variable. As mentioned above, there are three main public institutions in charge of providing health services to eligible population: IMSS, ISSSTE and PEMEX. The expenditures made by these institutions could also be contributing to the decrease of the IMR_{it}. We added the per insured person expenditure made by these institutions in health services provision and name the variable $HEEP_{it}$. Another control variable we include is percentage of uninsured population (PUP_{it}) in each state and in each year. This variable is a proxy of the necessities of health services for non-insured population in each state. We control for the primary school completion rate per state, $PSCR_{it}$, as a measure of schooling. Finally, we do not observe the out-of-pocket expenditure on health services by the population for years before 1998. Of course, these expenses could also be improving the health outcomes of the population. Therefore, we proxy this variable with the number of hospital beds per 1000 inhabitants in the private health sector, *i.e.*, $HBPS_{it}$.

We estimated equation (1) by fixed-effects panel estimation method, correcting standard errors for cluster effects of states.

Results from estimating equation (1) are in Table 5. The second column contains the estimates of the coefficients of specification (1) with fixed effects but without control variables.[26] Results indicate that an increase by one thousand pesos per capita in $FASSA_{it}$ decreases IMR_{it} in 39.4 per cent whereas an increase by the same amount in $Ramo12_{it}$ before 1997 decreases IMR_{it} in 33.7 per cent (and both effects are statistically significant at the 1 per cent level). Recall that average $FASSA_{it}$ is 438 pesos, thus if it increases to 1438, an increase of 228 per cent, the infant mortality decreases 39.4 per cent. For the case of $Ramo12_{it}$ an increase from its average of 278 pesos per capita between 1993 and 1998 to 1278 pesos, a 1000 pesos increase or a 359 per cent increase, the infant mortality decreases by 33.7 per cent. The difference between the two semi-elasticities is $\beta_4-\beta_2 = -0.394 - (0.337) = -0.057$, but not statistically significant. This implies that $FASSA_{it}$ and $Ramo12_{it}$ are indistinguishable.

In column (3) we estimate the same specification as before but we added control variables. Results are similar as those in column (2), that is, there is no significant difference between how $Ramo12_{it}$ did before the decentralization reform and how $FASSA_{it}$ did after its implementation. However, the difference is positive and equal to 0.0129, which implies that the semi-elasticity related to $FASSA_{it}$ is 1 percentage points higher than the corresponding for $Ramo12_{it}$. In column (4) and (5) we show the results from estimating equation (1) when we include a time trend and year indicators, respectively. In both cases, $\beta_4-\beta_2$ is negative, as expected, though not statistically different from zero. Notice that increasing $Ramo12_{it}$ and $FASSA_{it}$ by 1000 pesos decreases the IMR_{it} by 1.8 and 6.8 per cent, respectively, but neither coefficient is statistically significant (column 5).

Using the results in Table 5, we also compare $Ramo12_{it}$ and $FASSA_{it}$ with each other but in the years after the reform. In other words, we test whether $\beta_4-(\beta_2+\beta_3)$ is different from zero. In all

26 Results in column (1) were included to compare the R^2 from equation (1) without including fixed effects and when including such effects. In such case the R^2 is 0.474. We also regress IMR on time dummies only and on fixed effects only. The corresponding R^2's are 0.539 and 0.452, respectively.

Table 5

Fixed Effects Panel Estimated Coefficients

Independent Variables		Log Infant Mortality Rate					Log Fetal Death Rate
		(1)	(2)	(3)	(4)	(5)	(6)
b_1	*I(t>1997)*	–0.239***	–0.228***	–0.189***	–0.0807***	–0.074***	0.0864*
		(0.047)	*(0.025)*	*(0.022)*	*(0.013)*	*(0.012)*	*(0.0427)*
b_2	$Ramo12_{it}$	0.201	–0.337***	–0.353***	–0.061	–0.018	0.123
		(0.252)	*(0.096)*	*(0.078)*	*(0.052)*	*(0.06)*	*(0.304)*
b_3	$Ramo12_{it}$ * *I(t>1997)*	0.0387	0.006	0.088	0.093	0.014	–0.549
		(0.324)	*(0.131)*	*(0.126)*	*(0.059)*	*(0.065)*	*(0.499)*
b_4	$FASSA_{it}$ * *I(t>1997)*	–0.177	–0.394***	–0.340***	–0.097*	–0.068	–0.129
		(0.152)	*(0.055)*	*(0.08)*	*(0.05)*	*(0.056)*	*(0.203)*
	Time Trend	-	-	-	–0.047***	-	-
		-	-	-	*(0.002)*	-	-
	GSP_{it}	-	-	–0.003***	–0.0005	–0.0006	0.00119
		-	-	*(0.0007)*	*(0.0003)*	*(0.0004)*	*(0.0015)*
	DP_{it}	-	-	0.0002	0.0001	0.0003***	0.00346***
		-	-	*(0.0004)*	*(0.0001)*	*(0.0001)*	*(0.00124)*
	$HEEP_{it}$	-	-	0.073***	0.036**	0.027*	0.0802*
		-	-	*(0.01)*	*(0.013)*	*(0.013)*	*(0.0412)*
	PUP_{it}	-	-	–1.712***	–0.159	–0.318*	–0.894
		-	-	*(0.209)*	*(0.147)*	*(0.182)*	*(0.71)*
	$PSCR_{it}$	-	-	–0.005***	0.001**	0.002**	0.00515
		-	-	*(0.001)*	*(0.0009)*	*(0.0008)*	*(0.00498)*
	$HBPS_{it}$	-	-	0.061	0.069	0.055	0.0156
		-	-	*(0.075)*	*(0.042)*	*(0.04)*	*(0.116)*
	Constant	3.243***	3.393***	4.596***	3.402***	3.012***	1.021
		(0.072)	*(0.028)*	*(0.142)*	*(0.124)*	*(0.133)*	*(0.637)*
	Year Indicators	No	No	No	No	Yes	Yes
	Fixed effects	No	Yes	Yes	Yes	Yes	Yes
	$b_4 - b_2$	–0.378	–0.056	0.012	–0.036	–0.05	–0.252
	$Prob > F_1$	0.061	0.494	0.825	0.298	0.181	0.0671
	$b_4 - (b_2 + b_3)$	–0.417	–0.063	–0.076	–0.13	–0.064	0.298
	$Prob > F_2$	0.201	0.535	0.544	0.034	0.276	0.487
	Number of Groups	-	32	32	32	32	32
	Number of Observations	352	352	352	352	352	352
	R^2	0.474	0.872	0.936	0.973	0.983	0.316
	R^2 Overall	-	0.401	0.003	0.103	0.005	0.0869
	R^2 Between	-	0.0292	0.458	0.187	0.154	0.0923

Panel data estimations show state cluster robust standard errors in parentheses.
Note: The definition and units of the variables are in Table 2. Significance interpretation is as follows: *** p<0.01, ** p<0.05, * p<0.10.

five columns, except for column (4), it is the case that $FASSA_{it}$ is not significantly different from $Ramo12_{it}$ after the reform was implemented. However, notice that such difference is negative in all five cases. According to results in column (5), when we added year indicators and control variables, the difference is 0.064 which implies that $FASSA_{it}$ decreases IMR_{it} relative to $Ramo12_{it}$ when comparing them after 1998.

From Table 5 it is also possible to compare $Ramo12_{it}$ performance in the years after the reform with the years before the reform, coefficient β_2 captures this difference. This coefficient is positive in all four columns, but fails to be statistically significant. This implies that there is no difference between $Ramo12_{it}$ nowadays compared to before the reform. In accordance to column (5), the coefficient is 0.014. This means that one thousand pesos increase in $Ramo12_{it}$ after the reform took place decreases in 1.42 per cent the IMR_{it} compared to the effect of $Ramo12_{it}$ in the years before the reform took place.

Finally, another coefficient of interest from Table 5, is the one associated to the decentralization reform, $I(t>1997)$. Notice that in all five columns this coefficient is negative and statistically significant at 1 per cent level. This coefficient is capturing the fact that over time the IMR_{it} is decreasing between 1993-97 and 1998-2003. The magnitude of the coefficient decreases when we include either a time trend or year fixed effects.

Results presented in Table 5 are robust to different measures of health well-being, specifically, infant mortality rate for children less than 5 years old, child deaths by respiratory diseases per 1000 births, child deaths by intestinal diseases per 1000 inhabitants, and fetal death rate per 1000 inhabitants. Results from estimating equation (1) using as the dependent variable the fetal death rate are shown in column 6 of Table 5. Notice results are the same as before, $\beta_4 - \beta_2$ is negative, although significant at 10 per cent level.

5.3 *What was the impact of decentralization on health outcomes in states that received more resources from FASSA?*

The lack of significance of the previous results is evidence that, in general, decentralization of responsibilities and funds from federal to state authorities regarding state health services provision did not significantly improve the well-being of the population. Although the sign of the coefficients of interest are negative, their magnitudes are rather small. However, perhaps states that received more resources from FASSA did a better job than states that received fewer resources.

In this section we follow a difference in difference approach which will enable us to address the following question: Did states that receive more FASSA get better health outcomes than states that received less FASSA after the reform? Ideally, we would like to have an experiment with one group of states that were treated with health decentralization and other set of control states that were not submitted to the institutional change, and compare the performance of both groups after the reform was implemented. However, as previously discussed, all states received FASSA funds. Thus, we perform a pseudo experiment. We divide the states into two groups according to FASSA transfers per capita received in the first year of the reform (1998).[27] We called the first group high FASSA states[28] (or treated group) and are those that are above the median of the 32 states. The low

[27] The range of the distribution of FASSA per capita is large as the descriptive statistics point out. The median of FASSA per capita in 1998 was 332 pesos of 2010 and the mean was 350 pesos, with the maximum value being 997 pesos and the minimum 179 pesos. The coefficient of variation (standard deviation/mean) is 0.48. The average FASSA per capita for the high group is 458 pesos and for the low group is 242 pesos.

[28] Baja California Sur, Colima, Campeche, Quintana Roo, Guerrero, Nayarit, Aguascalientes, Durango, Tabasco, Sonora, Tlaxcala, Tamaulipas, Yucatán, Morelos, Chiapas, and Querétaro.

FASSA states group (or control group) are the remaining states. We estimate a set of difference in difference regressions with the following simple framework:

$$IMR_{it} = \alpha + \beta_1(H_i) + \beta_2 I\,(t > 1997) + \beta_3\,[I(t > 1997) * (H_i)] + X_{it}B_4 + c_i + u_{it} \qquad (2)$$

$$i = 1, \ldots 32 \qquad t = 1, \ldots 11$$

In this specification the dependent variable refers to the natural log of the infant mortality rate; H_{it} is an indicator function that takes the value of one if the state i belongs to the high FASSA group and zero if it belongs to the low FASSA group; $I(t>1997)$ is also an indicator function defined as before; and the variable multiplied by β_3 is an interaction term between the previous variables. This is the coefficient of interest because it is the difference in difference effect on health of the reform on the treated states (high FASSA) relative to the control group (low FASSA). X_{it} refers to the same vector of control variables as before; c_i denotes the state fixed effect which is assumed to be arbitrarily correlated with the regressors; and u_{it} denotes the idiosyncratic error for state i in year t. Also, in some specifications we also include state fixed effects, a time trend common to all states, and year fixed effects, just as before.

The interpretation of the coefficients of interest is as follows: α refers to the health indicator average of low FASSA group before the intervention; β_1 is the difference in the average of the dependent variable of the high and low FASSA groups before 1998; and β_2 is the change in the average for the control group (low FASSA) after the reform relative to the pre reform period. Finally, β_3 captures the difference of health indicator average between high and low FASSA states after the decentralization relative to the difference between high and low FASSA states in the years prior to decentralization. We expect this last coefficient to be negative, but also significant. If it turns out to be not significant, then we cannot conclude that there is a difference between the control and treatment group due to the decentralization.

Before presenting our results, it is worth pointing out that our identification strategy requires that per capita FASSA assignment in 1998, and thus our classification of states according to FASSA, to be exogenous and not correlated to the error term conditioned on the variables included in the right hand side of equation (2). For instance, if FASSA is assigned to states according to their health indicators, that is, states with worse health indicators receive more FASSA, our classification of states according to FASSA would not be exogenous. Table 6 shows the average of both groups for a variety of health indicators and other controls in 1997, the previous year to the reform. Last column indicates the *p*-value for the *t*-test of differences in means between both groups. With the exception of two of our shown variables, it is not possible to reject the hypothesis that the difference in means is statistically different from zero. Given the classification of the groups and the persistency of FASSA per capita as a function of the allocation of Ramo12 per capita in 1997, it is not a surprise that such variables are the only ones that are significantly different from zero at 1 per cent level. This result suggests that the initial allocation of FASSA and its classification were not determined by health indicators, as one would expect.

Table 7 shows the results of the estimating equation (2) between 1993 and 2003. The difference-in-difference coefficient (β_3) is negative but not significant in any of the regressions. Although the direction of the coefficient indicates that states receiving more FASSA had lower infant mortality rate after the reform than low FASSA states, this coefficient is statistically not different from zero. Thus, the results suggest that there is no significant difference in health indicators between the treated and control states after the reform relative to the years previous to the introduction of FASSA. The very small magnitude of the coefficient provides further assurance that decentralizing resources did not have an impact on health indicators for states which received more resources relative to those states who received fewer resources from FASSA. According to the results in column (4), which include control variables and a time trend, the coefficient associated to the high FASSA (β_1) states is negative and statistically significant. This implies that

Table 6

Mean Comparison Between Low and High FASSA States
(null hypothesis: high FASSA mean – low FASSA mean = 0)

	Year	High FASSA per capita mean	Low FASSA per capita mean	*p*-value
FASSA	1998	457.66	242.18	0
*Ramo*12	1997	392.51	229.4	0
HBPS	1998	0.21	0.29	0.04
DP(1)	1997	77.46	451.1	0.3
Log (infant mortality rate)	1997	3.2	3.17	0.71
Infant mortality rate	1997	24.88	24.28	0.73
GSP	1997	65993	68259	0.85
PSCR	1997	86.96	87.4	0.89
PUP	1997	0.49	0.49	0.9
HEEP	1997	2343	2330	0.97
Number of observations		16	16	

(1) Population density of the Low FASSA group in 1997 (451.10) seems to be quite bigger than the High FASSA counterpart; this difference is mainly explained because Distrito Federal belongs to the Low FASSA group. Alone in 1997 Distrito Federal had a population density of 5786.15 habitants per square kilometer. By excluding Distrito Federal from the Low FASSA group the new population density mean would be 95.43 and the new *p*-value would be 0.6531.

previous to the reform, high FASSA states had a mortality rate 34 per cent lower than low FASSA states. This suggests that FASSA was not assigned accordingly to health necessities by states. Finally, β_2 is significantly negative (–0.080) reflecting the downward trend of infant mortality in control states.

Results presented in Table 7 are robust to different measures of health well-being, as the ones used for robustness in Table 5; results are also robust to excluding states around the median. For example, we pick only the 10 states with the highest and the 10 with the lowest FASSA and the results do not change (column 6). We also run the same specification with the top and bottom six FASSA states and results remain.

5.4 *What was the impact of decentralization on the health outcomes of the non-insured population relative to the insured population?*

So far we have not found evidence that health decentralization significantly improved the infant mortality rate, used as a proxy of the health conditions of the population. In this section we present two more empirical exercises. As mentioned before, all the states received FASSA funds, so in that sense, all states were treated, that is, all states were affected by the reform. However, recall that FASSA and Ramo12 have a target population: those who have no insurance. Thus there

Table 7

Difference in Difference Estimated Coefficients (Pseudo Experiment)

Independent Variables	Log Infant Mortality Rate						Log Fetal Death Rate
	(1)	(2)	(3)	(4)	(5)	(6) [1]	(7)
I(High FASSA group)	0.021 *(0.057)*	–0.264*** *(0.007)*	–0.573*** *(0.106)*	–0.348*** *(0.047)*	–0.407*** *(0.054)*	–0.393*** *(0.065)*	–0.275 *(0.274)*
I(t>1997)	–0.341*** *(0.007)*	–0.341*** *(0.008)*	–0.255*** *(0.024)*	–0.080*** *(0.012)*	- -	–0.066*** *(0.017)*	- -
*I(High FASSA group) * I(t>1997)*	–0.007 *(0.013)*	–0.007 *(0.013)*	–0.022 *(0.018)*	–0.003 *(0.012)*	–0.002 *(0.011)*	0.007 *(0.016)*	–0.147 *(0.097)*
GSP_{it}	- -	- -	–3.875*** *(0.765)*	–0.548 *(0.396)*	–0.587 *(0.431)*	–0.864* *(0.446)*	0.213 *(1.474)*
$HEEP_{it}$	- -	- -	0.036** *(0.018)*	0.028** *(0.011)*	0.023* *(0.012)*	0.026* *(0.013)*	0.038 *(0.044)*
$Ramo12_{it}$	- -	- -	–0.152** *(0.069)*	0.045 *(0.039)*	0.036 *(0.033)*	0.101** *(0.04)*	–0.332 *(0.34)*
$PSCR_{it}$	- -	- -	–0.005*** *(0.001)*	0.002*** *(0.001)*	0.002*** *(0.001)*	0.003** *(0.001)*	0.007 *(0.006)*
DP_{it}	- -	- -	0.001 *(0.001)*	0.000** *(0)*	0.000*** *(0)*	0 *(0)*	0.003*** *(0.001)*
PUP_{it}	- -	- -	–2.107*** *(0.241)*	–0.166 *(0.162)*	–0.367* *(0.205)*	–0.316 *(0.205)*	–0.894 *(0.663)*
$HBPS_{it}$	- -	- -	0.019 *(0.088)*	0.076 *(0.046)*	0.048 *(0.046)*	0.153** *(0.059)*	–0.016 *(0.121)*
Time Trend	- -	- -	- -	–0.049*** *(0.002)*	- -	–0.049*** *(0.003)*	- -
Constant	3.288*** *(0.044)*	3.448*** *(0.004)*	5.114*** *(0.123)*	3.585*** *(0.12)*	2.962*** *(0.155)*	3.621*** *(0.136)*	–4.986*** *(0.52)*
Year Indicators	No	No	No	No	Yes	No	Yes
Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	352	352	352	352	352	220	352
Number of Groups	32	32	32	32	32	20	32
R^2	0.457	0.904	0.958	0.985	0.991	0.983	0.948

[1] Only for Top 10 and Bottom 10 FASSA states.
Note: The definition and units of the variables are in Table 2. Significance interpretation is as follows: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.
Panel data estimations show state cluster robust standard errors in parentheses.

is a fraction of the population in each state that was not affected by the reform, namely, those who had already health coverage. Taking advantage of this fact, we perform two exercises in which we consider the non-insured population as the treatment group and the insured population as the control group. Under this assumption, we are able to compare the performance of both groups for the years before (1993-97) and after (1998-2003) the reform was implemented.

To compare these groups we need to observe the infant mortality rate for each group. However, official statistics do not include IMR by insurance status, nor is there available data that permit us to construct the IMR for the insured and the uninsured population, respectively. Therefore, we rely on another health outcome: fetal deaths. This variable is part of *Estadísticas Vitales* published by INEGI. It is based on the information contained in Fetal Death Certificates. The main advantage of this variable is that it permits us to classify fetal deaths into our two groups of interest, according to whether the mother has insurance or not.

On the one hand, women who reported being beneficiary of either IMSS, ISSSTE, PEMEX, SEDENA,[29] SEMAR[30] or other institution are considered as insured. On the other hand, women who reported not having insurance are considered as non-insured.[31]

Using this data we construct the fetal deaths rate (FDR_{ijt}) defined as the number of fetal deaths occurred in state i, for group j, in year t as a fraction of the total population in state i which belongs to group j, in year t. In this case, j is equal to 1 for the non-insured population and equal to 2 for the insured population. Another advantage of this health outcome is that, similar to IMR, it responds relatively quickly to improvements in health provision. Moreover, this measure continues to be closely related to maternal health, one of the responsibilities transferred to states in the reform.

Nonetheless, FDR_{ijt} has one important problem. It tends to be biased because not all fetal deaths are reported to the corresponding authorities. Therefore not all fetal deaths have their corresponding certificate. This problem is more evident in poor, less educated and more disperse states, as well as states with a high proportion of uninsured population and less administrative capacity to register deaths. By controlling for some of these variables we take care for part of this bias. However, we do not observe other drivers of the bias. We have available two different series for the IMR, one that is biased (IMR_{Biased}) and one not (which corresponds to our IMR measure used along this study). We use the difference between these two series to approximate the bias in our FDR measure. By including this difference as a regressor, we try to control for the FDR bias we observe.

In a first exercise, we analyze whether the non-insured population had greater improvements in health outcomes after decentralization relative to the insured population. The identification strategy behind this specification is that the health provision decentralization was implemented for the benefit of non-insured people, leaving insured people unaffected. We expect that non-insured population observed improvements in fetal death rate relative to the insured population after the reform.

Our identification strategy requires that the distribution of people between the uninsured and insured cohorts is exogenous, *i.e.*, that insured population is almost the same as non-insured population but the treatment itself. There are many reasons we can think of that these two groups are not similar. However, Figure 4 graphs the national version of FDR_{ijt} per insurance eligibility group. As we would expect, insured population has a lower FDR than the one for non-insured

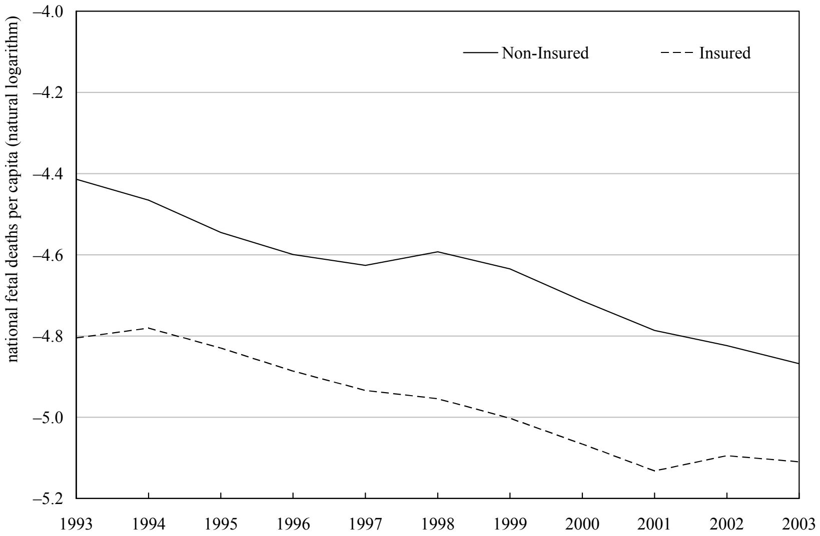
29 SEDENA stands for Secretaría de la Defensa Nacional, that is, Ministry of National Defense.

30 SEMAR stands for Secretaría de Marina, that is, Mexican Navy.

31 Those who reported insurance institution as unknown or not specified were excluded from the estimation. Nevertheless, as we will see in the results, classifying this group as insured or non-insured makes no significant difference in the results.

Figure 4

National Fetal Deaths Per Capita (Natural Logarithm) By Insurance Elegibility Group



Note: The insured fetal deaths per capita accounts for the fetal deaths of mothers who reported having some kind of medical insurance (*i.e.*, IMSS, ISSSTE, PEMEX, SEDENA, SEMAR or other institutions). Whereas the non-insured fetal deaths per capita accounts for the Fetal Deaths of mothers who reported not having any kind of medical insurance.
Source: Own elaboration with data from INEGI.

population. Second, from the graph it is also clear that both groups had very similar trends, particularly in the years before the reform took place. This is perhaps enough for our difference in difference approach to be credible. After 1997, the insured population continued with no particular changes whereas the non-insured population observed a small increase in 1998 to later show a steady decrease along the following years.

Another important assumption behind our identification strategy is that the composition of groups does not change over time, particularly as the result of decentralization. However, the insurance status depends on whether the person works in the formal or informal sector. Therefore, most people do not choose whether to have insurance or not, but in which sector of the labor market to work. Moreover, health services for non-insured people tend to be worse than health services for insured people.

We perform a difference in difference approach with fixed effects. The equation to regress is as follows:

$$FDR_{ijt} = \alpha + \beta_1 T_{ij} + \beta_2 I(t>1997) + \beta_3[I(t>1997) * (T_j)] + X_{ijt}B_4 + c_i + u_{ijt} \qquad (3)$$

$i = 1, \ldots 32$ $\quad j =$ Non-insured population, Insured population $\quad t = 1, \ldots 11$

In this case, FDR_{ijt} is the natural log of the fetal death rate for state i, group j, in year t. T_{ij} is equal to one for the non-insured population in state i, and zero otherwise. Finally, $I(t>1997)$ is defined as before. Our interest focuses on the coefficient that accompanies the interaction the latter two variables: β_3. This coefficient is the difference in difference effect of the reform on FDR_{ijt} for the non-insured population relative to the control group, that is, the insured population. We expect this coefficient to be negative and significant. If it is only negative but not significant, we cannot conclude that the reform had an impact on the treatment group relative to the control group. As before, c_i denotes the state fixed effect which is assumed to be arbitrarily correlated with the regressors; and u_{it} denotes the idiosyncratic error for state i in year t.

The vector of control variables, X_{it}, is the same as in previous exercises, except for two differences. First, total health from public institutions per capita, THE_{ijt}, is equal to FASSA and Ramo12 expenditures for non-insured population, that is when j=1, and equal to the sum of the health expenses by IMSS, ISSSTE and PEMEX for insured population (j=2).[32] Second, since our dependent variable is most probably biased, we add $log(IMR)–log(IMR_{Biased})$ as an additional variable to control for the possible bias contained in the data.[33] As already mentioned, the assumption behind this inclusion is that the bias observed in *FDR* is the same as the bias observed in IMR. Our IMR measure does not have this problem because corresponding authorities already corrected the statistics from this bias. However, such bias can be observed at the national level, if we compare our measure of IMR, available at the Millennium Development Goals Statistics published by United Nations, and what we denote IMR_{Biased}, published by the Bureau of Health Statistics of Mexico, SINAIS.

Results of the difference in difference regressions are shown in Table 8. Columns from (1) to (4) were included to keep the table comparable with previous exercises. According to the results in column (5), which include year indicators and control variables, the coefficient β_3 is negative (–0.0269) but it is not significant. This result suggests that average FDR_{ijt} after the decentralization reform took place relative to previous years, is 0.026 lower for the treatment group relative to the control group, however, it is not statistically different from zero. According to the same set of results, β_1 suggests that fetal deaths rate for the non-insured is significantly higher (0.621) than the insured population in the years before the reform and the coefficient is statistically significant at 1 per cent level. Moreover, β_2 suggests that the fetal deaths rate for the insured population decreased (–0.162) after the reform relative to previous years, and the coefficient is statistically significant at 5 per cent level. In column (6) and (7) we run the same specification as in column (5); however, in column (6) we included those fetal deaths in which the insurance status was not specified as if they were part of the insured population group, and in column (7) those fetal deaths were instead included in the non-insured population group. In both cases, β_3 is negative and not significant. These columns are included to check whether omitting the unknown or unspecified insurance status fetal deaths makes a difference for our results. Concluding, we found no significant difference between the non-insured and the insured population when comparing the mean FDR_{ijt} after the reform relative to previous years.

In a second exercise we continue exploiting our identification strategy and study whether there are differences in expenditure efficiency for insured and non-insured population, respectively, after the reform was implemented relative to previous years.

Fortunately, we are able to measure the efficiency of the expenditure for each of the two groups, because we also have detailed data on health expenditures made by various public health institutions. This information is summarized in the variable THE_{ijt} explained above. In equation notation this variable is:

$$THEijt = \begin{cases} Ramo\ 12 + FASSA_{it} \quad if \ \ j = Non-insured \ \ population \\ IMSS + ISSSTE + PEMEX_{it} \quad if \ \ j = Insured \ \ population \end{cases}$$

Therefore, we study whether the change in the elasticity of FDR_{ijt} with respect to total health expenditure for the non-insured population between 1998-2003 and 1993-97 is different from the change in the same elasticity for the insured population. The equation to estimate is the following:

[32] We do not have data about health expenditure realized by other health institutions, for example, private institutions. Nevertheless, IMSS, ISSSTE and PEMEX provide health coverage to more than 95 per cent of the insured population.

[33] Results are not significantly different if we do not include this difference as control variable. Results are available upon request.

Table 8

Difference in Difference Estimated Coefficients

Independent Variables	**Log Fetal Death Rate**						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
I(Non-insured)	0.478^{***}	0.478^{***}	0.534^{***}	0.587^{***}	0.621^{***}	0.512^{***}	0.743^{***}
	(0.0423)	*(0.0433)*	*(0.0952)*	*(0.0907)*	*(0.1)*	*(0.125)*	*(0.0902)*
I(t>1997)	-0.340^{***}	-0.227^{***}	-0.150^{**}	–0.0196	-0.162^{**}	-0.213^{***}	-0.174^{**}
	(0.045)	*(0.0492)*	*(0.0602)*	*(0.0517)*	*(0.0709)*	*(0.0632)*	*(0.0647)*
*I(Non-insured)*I(t>1997)*	0.0175	0.0175	0.000165	–0.0169	–0.0269	–0.0294	–0.0581
	(0.0311)	*(0.0318)*	*(0.0452)*	*(0.0454)*	*(0.0469)*	*(0.0499)*	*(0.0439)*
$HEEP_{it}$	-	-	0.0405	0.0795	0.104	0.148	0.104
	-	-	*(0.0692)*	*(0.0684)*	*(0.0798)*	*(0.0971)*	*(0.0725)*
$PSCR_{it}$	-	-	-0.00494^{*}	0.0039	0.00305	0.00331	0.00285
	-	-	*(0.00269)*	*(0.00337)*	*(0.00361)*	*(0.00331)*	*(0.00332)*
DP_{it}	-	-	0.00250^{***}	0.00251^{***}	0.00258^{**}	0.00253^{**}	0.00280^{***}
	-	-	*(0.00072)*	*(0.000896)*	*(0.000976)*	*(0.000927)*	*(0.000948)*
$HBPS_{it}$	-	-	–0.0606	0.00357	–0.0354	–0.0805	–0.0847
	-	-	*(0.0841)*	*(0.102)*	*(0.13)*	*(0.116)*	*(0.123)*
GSP_{it}	-	-	$-3.72e{-}05^{**}$	–1.77E–06	1.59E–05	1.09E–05	1.20E–05
	-	-	*(1.62E–05)*	*(1.43E–05)*	*(1.60E–05)*	*(1.66E–05)*	*(1.54E–05)*
PUP_{it}	-	-	-2.471^{***}	–0.637	-1.281^{*}	–1.103	-1.498^{**}
	-	-	*(0.577)*	*(0.652)*	*(0.751)*	*(0.716)*	*(0.691)*
$IMR_{Ratio,\ it}$	-0.823^{***}	0.0189	–0.115	–0.157	–0.155	-0.187^{*}	–0.157
	(0.141)	*(0.105)*	*(0.111)*	*(0.107)*	*(0.115)*	*(0.101)*	*(0.111)*
Trend	-	-	-	-0.0252^{***}	-	-	-
	-	-	-	*(0.00437)*	-	-	-
Constant	-1.172^{***}	-1.225^{***}	–0.00731	-1.611^{***}	-1.588^{***}	-1.493^{***}	-1.491^{***}
	(0.0811)	*(0.0565)*	*(0.23)*	*(0.397)*	*(0.389)*	*(0.351)*	*(0.356)*
Year Indicators	No	No	No	No	Yes	Yes	Yes
Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	704	704	704	704	704	704	704
R^2	0.638	0.886	0.896	0.901	0.904	0.894	0.916

Panel data estimations show state cluster robust standard errors in parentheses.
Note: The definition and units of the variables are in Table 2. Significance interpretation is as follows: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

$$FDR_{ijt} = \alpha + \beta_1 T_{ij} + \beta_2 I\,(t > 1997) + \beta_3\,[I(t > 1997) * T_{ij}] + \beta_4 \log(THE_{ijt}) + \beta_5[\log(THE_{ijt}) * \\ * I(t > 1997)] + \beta_6[\log(THE_{ijt}) * T_{ij}] + \beta_7[\log(THE_{ijt}) * I(t > 1997) * T_{ij}] + X_{it}B_8 + c_i + u_{ijt} \quad (4)$$

$i = 1, \ldots 32$ j = Non-insured population, Insured population $t = 1, \ldots 11$

Equation (4) is just an extension of equation (3) where we interact $log(THE_{ijt})$ with the decentralization reform indicator, the treatment indicator and with both indicators together. As in previous exercise, FDR_{ijt} is the natural log of the fetal death rate for state i, group j, in year t; T_{ij} is equal to one for the non-insured population in state i, and zero otherwise; $I(t>1997)$ is decentralization reform indicator; c_i denotes the state fixed effect which is assumed to be arbitrarily correlated with the regressors; and u_{it} denotes the idiosyncratic error for state i in year t. The vector of control variables, X_{it}, is the same as in the previous exercise, that is, includes all controls discussed before plus THE_{ijt} and $log(IMR)$–$log(IMR_{Biased})$.

In this case, the coefficient of interest is β_7. This coefficient compares the elasticity of the fetal death rate with respect to total health expenditure after the reform relative to years previous the reform for the non-insured population relative to the insured population. We expect this coefficient to be negative and significant. In other words, we expect health expenditure for non-insured population to have a greater impact in reducing fetal death rate after the reform relative to the control group.

Results for the difference in difference regressions are shown in Table 9. We again include columns (1) through (4) just to keep all tables comparable. Results in column (5) are the more general since they include control variables and year indicators. According to such results, which include control variables and year indicators, the coefficient β_7 is negative (–0.192) and significant at the 10 per cent level. It implies that the difference in elasticities from 1998-2003 and 1993-97 is 0.192 lower for the non-insured population relative to insured population. In other words, if health expenditure increases 1 per cent for both groups and both periods, the FDR exhibits a larger fall by 0.19 per cent for the non-insured population relative to the insured population. Contrary to our previous results, the health expenditure for the non-insured population, through Ramo12 and FASSA, is significantly more effective after the reform took place than the health expenditure for the insured population. This is perhaps an indication that the health production function in general is convex. Thus, further reductions of the FDR are more costly in the insured sector, for which the FDR is already low, compared to the non-insured sector. Another possible explanation is that when analyzing the performance of Ramo12 and FASSA expenditure together, they do much better than each by their own. Understanding what is explaining the obtained result certainly is an interesting line of future research.

This result can be explained by the fact that the elasticity of FDR with respect to THE did not improve for the insured group from 1993-97 to 1998-2003, that is, coefficient β_5 is 0.0322 and it is not statistically significant. This is in accordance with the implicit assumption that the insured population group was not affected by the decentralization reform. Moreover, for the non-insured group that same elasticity improved after the reform, *i.e.*, $\beta_5+\beta_7$, is –0.16 and it is statistically significant at 5 per cent level. This is because the elasticity of FDR with respect to THE for the period 1998-2003 is 0.02 and not significant, whereas the same elasticity for the period 1998-2003 is 0.184 and statistically significant at 1 per cent level (therefore, 0.18–0.02=–0.16). Although this implies that the reform did improve the health well-being of the population, notice that these elasticities are positive. In other words, increasing Ramo12 before the reform by 1 per cent increased the FDR by 0.18 per cent and increasing Ramo12+FASSA by 1 per cent for the years after the reform increased the FDR by 0.02 per cent although we cannot distinguish this effect from zero. This is thus in accordance to our results from previous sections.

Table 9

Health Expenditure Efficiency Comparison: Estimated Coefficients

Independent Variables		Log Fetal Deaths Rate						
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
b_1	*I(Non-insured)*	0.510*** (0.167)	0.484*** (0.133)	0.452*** (0.125)	0.516*** (0.125)	0.538*** (0.144)	0.351** (0.163)	0.642*** (0.118)
b_2	*I(t*>1997*)*	–0.215 (0.177)	–0.189 (0.127)	–0.195** (0.0926)	–0.0414 (0.0923)	–0.194* (0.102)	–0.220** (0.0924)	–0.228** (0.0995)
b_3	*I(J=Non-insured) * I(t*>1997*)*	–0.127 (0.168)	–0.114 (0.128)	–0.018 (0.0822)	–0.0623 (0.084)	–0.063 (0.0885)	–0.109 (0.0724)	–0.0764 (0.0877)
b_4	THE_{ijt}	–0.0022 (0.156)	–0.104 (0.137)	–0.13 (0.122)	–0.0694 (0.125)	–0.0511 (0.147)	–0.105 (0.169)	–0.078 (0.115)
b_5	THE_{ijt} * *I(t*>1997*)*	–0.123 (0.163)	–0.0254 (0.132)	0.0668 (0.0711)	0.0287 (0.0737)	0.0322 (0.0788)	0.0126 (0.0685)	0.0587 (0.0764)
b_6	THE_{ijt} * *I(J=Non-insured)*	0.0637 (0.155)	0.269* (0.137)	0.277** (0.121)	0.241* (0.123)	0.235* (0.134)	0.339** (0.151)	0.270** (0.11)
b_7	THE_{ijt} * *I(Non-insured) * I(t*>1997*)*	–0.187 (0.177)	–0.231 (0.156)	–0.260** (0.0987)	–0.199* (0.102)	–0.192* (0.102)	–0.161 (0.1)	–0.235** (0.098)
	$PSCR_{it}$	- -	- -	–0.00615** (0.00245)	0.00214 (0.00299)	0.00195 (0.00332)	0.00237 (0.00295)	0.0016 (0.00299)
	DP_{it}	- -	- -	0.0019*** (0.000584)	0.002** (0.000823)	0.00215** (0.0009)	0.0021** (0.000842)	0.00231** (0.00088)
	$HBPS_{it}$	- -	- -	0.00316 (0.0925)	0.0562 (0.114)	–0.00393 (0.141)	–0.042 (0.127)	–0.0484 (0.132)
	GSP_{it}	- -	- -	–2.64e–05* (1.51E–05)	6.32E–06 (1.42E–05)	1.82E–05 (1.54E–05)	1.33E–05 (1.62E–05)	1.36E–05 (1.47E–05)
	PUP_{it}	- -	- -	–1.957*** (0.505)	–0.273 (0.548)	–0.83 (0.672)	–0.559 (0.619)	–1.039 (0.633)
	$IMR_{Ratio,it}$	–0.826*** (0.142)	–0.0343 (0.108)	–0.127 (0.113)	–0.167 (0.109)	–0.17 (0.114)	–0.202* (0.0996)	–0.172 (0.11)
	Time Trend	- -	- -	- -	–0.0234*** (0.00403)	- -	- -	- -
	Constant	–1.168*** (0.169)	–1.124*** (0.119)	0.056 (0.279)	–1.466*** (0.394)	–1.467*** (0.396)	–1.339*** (0.361)	–1.324*** (0.349)
	Year Indicators	No	No	No	No	Yes	Yes	Yes
	Fixed Effects	No	Yes	Yes	Yes	Yes	Yes	Yes
	$b_4 + b_5 + b_6 + b_7$ $Prob > F_1$	–0.248 0.000732	–0.0906 0.199	–0.0462 0.554	0.00152 0.984	0.0242 0.766	0.0862 0.282	0.0156 0.837
	$b_4 + b_6$ $Prob > F_2$	0.0615 0.373	0.165 0.00267	0.147 0.00761	0.172 1.83E–03	0.184 2.89E–03	0.235 0.0000985	0.192 0.000778
	$b_5 + b_7$ $Prob > F_3$	–0.31 0.0000336	–0.256 2.47E–03	–0.193 8.70E–03	–0.17 0.0165	–0.16 0.0228	–0.148 0.0348	–0.176 0.00915
	$b_4 + b_5$ $Prob > F_4$	–0.126 0.605	–0.129 0.0506	–0.0631 0.432	–0.0406 0.648	–0.0189 0.861	–0.0922 0.445	–0.0193 0.821
	Number of Observations	704	704	704	704	704	704	704
	R^2	0.649	0.893	0.9	0.905	0.907	0.9	0.92

Panel data estimations show state cluster robust standard errors in parentheses.
Note: The definition and units of the variables are in Table 2. Significance interpretation is as follows: *** p<0.01, ** p<0.05, * p<0.10.

Just as in the previous exercise, column (6) and (7) are the same specification with the only difference being related to the dependent variable: in column (6) fetal death certificates with insurance status not specified were classified as in the insured population group; and in column (7) those same fetal deaths were classified in the non-insured population group. In both cases, β_7 is negative, however, it is not significantly different from zero in column (6). This is accordance to the hypothesis that those fetal deaths with unspecified insurance status are in fact non-insured because the magnitude of the coefficient β_7 in column (6) decreases sufficiently to become insignificant; and the magnitude of the same coefficient but in column (7) increases and becomes significant at 5 per cent level. As before, these columns are included to check that omitting the unknown or unspecified insurance status fetal deaths makes no significant difference for our results.

6 Conclusions

The results presented in this paper suggest that health decentralization in Mexico did not have the desired effects on state-level health outcomes. We did not find strong evidence that expenditure after the reform can explain improvements in health indicators, such as the child mortality or the fetal death rates. In particular, we did not find that the effectiveness of FASSA expenditure was higher than the impact of Ramo12 previous to the reform. Nevertheless, our exercises also suggest that the non-insured population had better outcomes derived from the reform than insured population. These results contrasts to what the policy makers that implemented the reform intended as well as what the classical theory of federalism would predict.

We believe that the results observed in Mexico may have obeyed to different factors that are worth exploring in future extensions of this paper. First, the reform was implemented from one year to the next and it is possible that states lacked the capacity to meet their new responsibilities immediately and neither were they able to administer the economic resources associated to health provision (Merino, 2003). The reforms may take some time in order to be effectively implemented as governments learn to operate and spend efficiently. A second hypothesis is that the institutional framework in which health was decentralized did not provide states with the incentives to provide better services to people. As we discussed in the text, the allocation of FASSA among states is rather unclear and it does not depend on the own state effort or health results. A merit-based system, in which future FASSA allocations depend on state's own contributions and the efficiency with which each state used its resources in previous years, could have helped to boost the impact of health expenditure. In this sense, a study of the effects of the Seguro Popular (which is partially financed by FASSA) would contribute to the discussion since the rules and uses of decentralized resources for that program are better defined. A third explanation is related to checks and balances that states have when spending public resources, the capacity of the taxpayers to know how efficiently their money is being spent and the availability of mechanisms for accountability. We think that these three potential explanations are not exclusive and certainly complement the results of the paper.

REFERENCES

Akin, J., P. Hutchinson and S. Koleman (2005), "Decentralization and Government Provision of Public Goods: The Public Health Sector in Uganda", *The Journal of Development Studies*, No. 41, pp. 1417-43, November.

Asfaw, A., K. Frohberg, K.S. James and J. Jutting (2007), "Modeling the Impact of Fiscal Decentralization on Health Outcomes: Empirical Evidence from India", *Journal of Developing Areas*, No. 41, pp. 17-35, Fall.

Birn, A. (1999), "Federalist Flirtations: The Politics and Executions of Health Services Decentralization for the Non-insured Population in Mexico, 1985-1995", *Journal of Public Health Policy*, No. 20, pp. 81-108, March.

Cabrero, E. and J. Martínez-Vázquez (2000), "Assignment of Spending Responsibilities and Service Delivery", in M. Giugale and S.B. Webb (eds.), *Achievements and Challenges of Fiscal Decentralization: Lessons from Mexico*, The World Bank, pp. 139-74, Washington (D.C.).

Cantarero, D. and M. Pascual (2008), "Analysing the Impact of Fiscal Decentralization on Health Outcomes: Empirical Evidence from Spain", *Applied Economics Letters*, No. 15, pp. 109-11, February.

Gómez Dantés, O. and M. Ortiz (2004), "Seguro popular de salud: Siete perspectivas", *Salud Pública de México*, No. 46, pp. 585-88, November-December.

Habibi, N., C. Huang, D. Miranda, V. Murillo, G. Ranis, M. Sarka and F. Stewart (2001), "Decentralization in Argentina", Yale Economic Growth Center, Discussion Paper, No. 825.

Jiménez Rubio, D. (2011), "The Impact of Decentralization of Health Services on Health Outcomes: Evidence from Canada", *Applied Economics*, No. 43, pp. 3907-17, October.

Joumard, I., A. Christophe, C. Nicq and O. Chatal (2008), "Health Status Determinants: Lifestyle, Environment, Health Care Resources and Efficiency", OECD, Economics Department, Working Paper, No. 627.

Merino, G. (2003), "Descentralización del sistema de salud en el contexto del federalismo", in F. Marie Knaul and G. Nigenda (eds.), *Caleidoscopio de la Salud, de la investigación a las políticas, y de las políticas a la acción*, Fundación Mexicana para la Salud, pp. 195-207, México (D.F.).

Mosca, I., L. Crivelli and M. Filippini (2006), "Federalism and Regional Health Care Expenditures: An Empirical Analysis for the Swiss Cantons", *Health Economics*, No. 15, pp. 535-41, May.

Oates, W. (1972), "Fiscal Federalism", *Journal of Public Economics*, No. 2, pp. 188-91, April.

OECD (2005), *Reviews of the Health System*, OCDE, México (D.F.).

Prud'homme, R. (2004), "On the Dangers of Decentralization", The World Bank, Policy Research, Working Paper, No. 1252.

Robalino, D., O. Picazo and A. Voetberg (2001), "Does Fiscal Decentralization Improve Health Outcomes? Evidence from a Cross-country Analysis", The World Bank, Policy Research, Working Paper, No. 2565.

Soto, V.E., M.I. Farfan and V. Lorant (2011), "Fiscal Decentralization and Infant Mortality Rates: The Colombian Case", Working Paper.

Sour, L., S. Robles, I. Ortega, O. Avelino and M. Peña (2004), "Diagnóstico jurídico y presupuestario del Ramo 33", CIDE, México (D.F.).

Uchimura, H. and J. Jütting (2007), "Fiscal Decentralization, Chinese Style: Good for Health Outcomes?", *World Development*, No. 37, pp. 1926-34, December.

Weingast, B. (2009), "Second Generation Fiscal Federalism: The Implications of Fiscal Incentives", *Journal of Urban Economics*, No. 65, pp. 279-93, May.

COMMENTS ON SESSION 3
TAXATION, REGULATIONS AND PUBLIC SERVICES

Stefan Bach[*]

Comments on "How Costly Are the Public Sector Inefficiencies? An Integrated Framework for Its Assessment" by Jorge Onrubia-Fernández and A. Jesús Sánchez-Fuentes

Summary

The paper provides a theoretical framework to analyse public sector performance. Two equivalent measures of social welfare changes are proposed, obtained from the cost function, and directly from the production function. Applications to empirical analysis are discussed.

Comments

The efficiency issues of public spending are increasingly on the political agenda against the outstanding budgetary imbalances in many countries. It is helpful to provide and enlarge theoretical models to assess public sector inefficiencies in terms of social welfare. The latter implies not only budgetary savings but also indirect monetary gains, e.g., from better education and health. The authors discuss goods and services that are *excludable*, unlike pure public goods. It would be helpful to extend the analysis on the character of pure public goods such as defense, social security, etc. Financing issues could also be discussed. Excludable goods and services would allow for user fees covering the "private" character, whereas distortionary taxes are required to finance the mere public good impact such as redistribution or positive externalities. A further critical topic is the assumption of the exogenous degree of efficiency. Actually, organizational issues or rent-seeking behavior of politicians and public administration play an important role in public sector reform.

Transaction costs of implementing public sector reforms could be substantial with respect to the devaluation of existing capital and protection of trust/grandfathering, which provokes compensation requirements to the losers and thus reduces the welfare benefits from the reform. In a more dynamic setting, collective decision-making as well as the lack of competition and "creative destruction" in public sector performance and reform might be considered. Thus, one could distinguish between technical efficiency and economic efficiency in a narrower sense, which is largely addressed in the study, and a wider scope of dynamic and political efficiency.

Measurement and application issues regard the availability of information on production and cost functions, including organizational slacks. This would require raising internal information from public authorities. An alternative would be benchmark comparisons between different jurisdictions or countries, which have their own shortcomings. Demand functions on public goods could be derived from specific surveys, or by estimates from existing surveys and from political decision making and voting. Social welfare functions could be used to operationalize political programs.

[*] DIW (Deutsches Institut für Wirtschaftsforschung / German Institute for Economic Research), Berlin.
E-mail: sbach@diw.de

Comments on "The Quality of Government and Living Standards" by Francesco Grigoli and Eduardo Ley

Summary

The study analyses the potential impact of public waste on national income and living standards in international comparison. Illustrative calculations based on scores from different studies are used to demonstrate the significant impact, which could imply a re-ordering of cross-country rankings on living standards.

Comments

The illustrative calculations reveal the economic importance of public waste in macroeconomic terms. However, the reliability of the efficiency scores is contentious. This would require scrutinizing public sector efficiency more detailed. Moreover, an implementation within national accounting is intricate. This would introduce a normative element of output valuation that goes beyond simple accounting. Similar corrections could also be applied to externalities of the private sector, such as environmental pollution, market failure, or inequality.

Anyway, it is meritorious to point out that public waste reduces real income and living standards. Larger disparities between countries or regions should be considered within the pertinent comparisons. Finally, this is another topic of criticism to GDP as an indicator of economic performance or even welfare, which should be part of the discussion following the Stiglitz-Sen-Fitoussi report. Therefore, items of public waste could be included into complementary satellite information attached to GDP compilations and rankings. This would, however, require measurable and reliable indicators of public waste in international comparison, and thus call for more detailed data from the public administration as well as output indicators on public goods such as health, education levels, etc.

Comments on "An Evaluation of the 1997 Fiscal Decentralization Reform in Mexico: The Case of the Health Sector" by André Martínez Fritscher and Carolina Rodríguez Zamora

Summary

The paper provides an ex-post evaluation of the decentralization reform of health funds and responsibilities in Mexico in 1997. It aims to identify the impact of decentralization on health indicators, as there were no changes in the regional distribution of funds after reform. The authors found no significant effects on infant mortality rate at the state level by a comparison of outcomes before and after reform, further differentiated by state groups with different endowments. Moreover, as a natural experiment, the insured population is used as a control group, which indicates some increased efficiency of the program. The authors discuss reasons of the reform's meager results. In particular, they argue that it took some time to become effectively, and that there were no incentives for state governments to provide better services.

Comments

This paper is a fine impact assessment study, which aims to identify the impact of decentralization on public sector outcome at the example of public health care in Mexico. With respect to the empirical specification one might question whether the outcome measures are too rough. Child mortality of fetal death rate seems to be a rather specific indicator, although important especially for low developed regions. Actually, the long-term impact, e.g. from medical prevention and rehabilitation would be interesting if measurable. Moreover it would be challenging to exploit the heterogeneity within the states, e.g., by measuring rural vs. urban areas, or the share of indigenous population. Finally, further reasons for ineffectiveness could be analyzed, such as organizational issues, or incentives for service provision before and after the reform. This would, however, require case studies or expert interviews on the implementation of the reform in single states.

COMMENTS ON SESSION 3
TAXATION, REGULATIONS AND PUBLIC SERVICES

Sergio Clavijo[*]

Comments on "Service Regulation and Growth: Evidence from OECD Countries" by Guglielmo Barone and Federico Cingano

Barone and Cingano argue that anti-competitive regulations go against growth in provision of services like energy, telecom and transportation in OECD countries. The authors also argue that such anti-competitive regulations impair price reductions in those services that would, otherwise, benefit consumers at large.

This lack of growth in service provision and the slow transmission of price reduction is due to three main factors, according to the authors. In the first place, setting regulation of prices and tariffs is a very complex issue, where even knowledgeable regulators tend to err. In the second place, by forcing "unbundling" of investments between generators and distributors, most economies loose opportunities to exploit economies of scale and scope in such services. Finally, the authors also argue that such excessive regulations hamper productivity gains at the inter-industry level, which is the main focus of their analysis.

This is very well crafted paper, where macro- and micro-analysis are carefully entangled and explained. In my opinion, the main conclusions against over-regulation in the service sectors could as well be extended to the health sector, where regulators have also requested "unbundling" of investments between the insurance component and the hospitals components, losing "economies of scale-scope", as explained before.

However, such conclusions seem to me a bit "counter-intuitive" when applied to the financial sector, where the recent financial global crises tells us that the lack of proper regulation prompted a severe and long-lasting mortgage and derivative crises. For instance, the Dodd-Frank Act in the United States and the Basle III regulations seem to be on the right track of strengthening regulations in order to avoid future "systemic risks".

Regarding their econometric work, their "working-horse" regressions focuses on the Real Value-Added Growth for the 1996-2002 period for OECD countries, as in equation (1):

$$VA_{j,c} = B_0 + B_1\ SERVREG + B_2\ SHARE + U_c + U_j + Error_{j,c} \tag{1}$$

where one of the main hypotheses has to do with finding $B_1<0$; in this case the argument is that higher regulation would imply lower growth in the provision of such services. Interestingly, the authors find statistical support, in a cross-country panel of a fixed-effect model, to argue that the rule of law (strong institutions) would permit that firms operate better in a deregulated framework, where markets conditions would benefit consumers.

Although the paper does not focus on emerging markets, let me suggest the authors to extend their analysis to those countries, since there seems to be a historical cycle regarding the regulation of services. In my experience as civil servant in Colombia, I have noticed that in many less developed economies the State moves late in regulating the provision of services. Hence, in order to catch up historically, then they move to the point of setting an over-regulatory framework which, indeed, might end up causing a lot of the problems stated here by Barone and Cingano.

---

[*] Director of ANIF, Colombia.

Comments on "Growth Implications of Structure and Size of Public Sectors" by Hans Pitlik and Margit Schratzenstaller

The main message of Pitlik and Schratzenstaller, in their interesting paper about structure of public sectors, is that there is not such a thing as "one-size-fits-all" both regarding public sector structure in promoting growth and concerning the topics of taxes and expenditure.

The authors analyze the "friendliness" indicators of growth for EU-12-15 and OECD countries and find, in the spirit of "endogenous growth models", that tax/expenditure composition is much more important that the size of revenue collections of outlays.

The authors take dispersion in the growth "friendliness" index as evidence of lack of policy coherence. Consequently, Pitlik and Schratzenstaller call for pursuing complementary policies to gain coherence, finding that over-regulation seems to play a role in growth stagnation (as in the case of Greece), while deregulation apparently promotes growth (as in the case of New Zealand).

On the issue of productive vs. unproductive expenditure, the authors explain that this continues to be an open debate matter. On the operative side, you could always argue about increasing expenditures in the "meritory ones" (education and health), while in the case of the "golden fiscal rules" you could as well argue that fixed capital formation is good to propel sustainable growth in the near future.

Let me suggest to the authors the adoption of an explicit theoretical framework in order to better organize this kind of discussion. For instance, the adoption of a model would allow the authors to better cast their hypothesis about growth promotion/retardyness, especially since productive/unproductive definitions are rather arbitrary. The second suggestion I offer is to include in their analysis cases of *ex ante*/*ex post* responses to the current European crisis, which I reckon could easily be introduced, given the complete research they have already conducted regarding both tax and expenditure structures.

Finally, let me pose two questions. How is it that well positioned countries such as Spain and USA ("friendliness index") have experienced so much macroeconomic pain recently (2010-12), lagging behind in the growth field and facing high fiscal tension? This is an example of how useful an analysis of *ex ante*/*ex post* experiences could be. My last query has to do with deepening their analysis with regard to the "effective tax burden", because clearly nominal or marginal rates do not tell the whole story regarding tax collections. On the expenditure side, it would be vital to include the impact of the so-called "contingent liabilities", which will significantly alter current expenditure structures, as discussed in previous fiscal workshops of the Banca d'Italia.

COMMENTS ON SESSION 3
TAXATION, REGULATIONS AND PUBLIC SERVICES

Yngve Lindh[*]

The papers presented in this session provide interesting insights in the current debate on taxation. The two papers I will comment on are related to each other as they both analyses aspects of how of tax systems affect employment and economic growth. While the paper by Peter Benczur, Gabor Katay, Aron Kiss and Oliver Racs concentrates on the tax system and its interaction with transfers in one country, in this case Hungary, the paper by Bert Saveyn, Jonathan Pycroft and Salvador Barrios highlights the importance to take into account cross-country spillovers when analysing effects of tax changes in single countries.

1 Income taxation, transfers and labour supply at the extensive margin

The paper by Benczur, Katay, Kiss and Racs delves into a very relevant issue: The effects of reforms in taxes and transfers on labour market participation. This issue is highly topical in many countries. Related to the economic and fiscal crises in the Euro Area, structural reforms that have significant positive impact on employment and growth are search for high and low. Reforms that improve labour supply are obvious examples of growth-friendly policies, at least in the longer term. And more generally, reshuffling tax systems to make them more economically efficient is a good example of reforms that could be used in the current situation to boost growth.[1]

This issue is not least relevant for Hungary, a country with one of the lowest labour market participation rates in the European Union. As the authors point out this has been an obstacle for convergence to higher income-levels after Hungary joined the EU in 2004. Some types of individuals have particularly low rates: women in child bearing ages, elderly and low skilled.

In my own country, Sweden, there has in recent years been a strong focus on the joint effects on participation in the labour market from a substantially increased Earned Income Tax Credit (EITC), together with reforms of the unemployment and sick leave insurances. An assessment is that these reforms will have a significant positive long-run effect on employment even if there are uncertainties around how large they will be in a longer perspective.[2]

In the Hungarian paper, effects on labour market participation of changes in taxes and transfers are estimated for different types of households and individuals. Related to this, it would be informative to get a bit more details about the Hungarian reforms in this area under the relevant time period and also how these reforms affect the calculated disposable income variable. The introduction of the flat tax in Hungary is mention, but not much more. For instance, reforms in unemployment benefit systems have been important in some countries. Is this also the case in Hungary? And, if this is the case, are these reforms included in the dataset?

Generally, the results in the paper for the different types and households and individuals seem reasonable. Weak groups in the labour market are more sensitive to changes in taxes and transfers for their decision to participate in the labour market. The only results that are a bit surprising are those related to the education level. Elementary, secondary and tertiary school backgrounds are related to weak effects of tax and transfer changes, while a vocational training background is related to stronger effects. Is there a rational for this difference?

---

[*] Finance Ministry, Stockholm.

[1] See, for instance, Å. Johansson, C. Heady, J. Arnold, B. Brys and L. Varia (2009), *Taxation and Economic Growth*, OECD.

[2] Swedish Fiscal Policy Council (2011), *Swedish Fiscal Policy*, pp. 222-23, Stockholm.

The features of the disposable income variable I mentioned earlier have also implications for the estimation of elasticities. In a recent Swedish study[3] the authors look into the effects on labour supply of the recent Earned Income Tax Benefit (EITB) reform in Sweden. From this study it is concluded that the effects are significantly positive. However, the results are found uncertain because there is too little variation in treatment between different individuals and that there are underlying trends in participation/employment that co-varies with the tax credit in ways that are hard to control. A question is if such estimation problems also could be relevant in the Hungarian study?

A last issue is that reforms in tax, transfer and benefit systems could have effects on the equilibrium wage level and consequently on labour demand. It seems that the effects of these types of reforms on participation and employment could go through both supply and demand channels.

In the end of the paper the elasticities found at the micro level are used to calculate the aggregated effect of recent Hungarian reforms. The result is unfortunately not encouraging. In its latest Economic Survey of Hungary by the OECD,[4] the Organisation also warns that the recent reforms in Hungary potentially can have negative effects on the participation rate, especially for low-income earners. This really shows how highly policy-relevant the work by Benczur, Katay, Kiss and Razc is.

2 The cost of tax increases in the EU

Not least in the wake of the global economic and financial crises governments need to implement tax systems that are growth-friendly. This is a complex issue in the European Union where economies are deeply interdependent. Bert Saveyn, Jonathan Pycroft and Salvador Barrios have in their paper chosen to gauge the size of potential cross-country spillover effects from tax changes by calculating the marginal cost of increases in labour taxes and energy taxes. The authors also analyses the role of labour market rigidities for the sizes of tax distortions.

The first question put by the authors is which types of tax reforms will promote growth in European countries. A second question is which types of taxes should or should not be coordinated at the European level.

The authors main contribution is that they take into account "spillover" effects when analyzing tax distortions, which they also claim has been ignored in earlier literature. Labour and energy taxes are in the focus of the analyses and this choice is well motivated in the paper. However, in the tax literature property taxes and broad based taxes on consumption are often seen as taxes which are least distortive, *i.e.*, most growth-friendly.[5]

A few questions on the analytical framework:

- is there empirical evidence that R&D expenditure is a good proxy for technological progress? There has been some criticism that this "input measure" is a rather blunt approximation;
- cross-border shopping is not included in the analysis. Could that potentially be of importance? What do we know empirically?
- the possibility to vary labour market imperfections are built into the model used by the authors through a parameter, e.g., in equation 2.8 in Appendix 2. A question is if this parameter has an economic interpretation. Would it be possible to, as an alternative, use an index reflecting degrees of imperfections in the labour market in different countries?

---

[3] K. Edman, C.Y. Liang, E. Mörk and H. Selin (2012). "Evaluation of the Swedish Earned Income Tax Credit", IFAU, Uppsala.

[4] OECD (2012), *Economic Survey – Hungary*, March.

[5] See, for instance, Å. Johansson, C. Heady, J. Arnold, B. Brys and L. Varia (2009), *Taxation and Economic Growth*, OECD.

I also believe it would be fine, as a reader, to get more explicit descriptions about channels and mechanisms in the model leading to the spillover effects.

Most empirical results in the study seem plausible. First, distortions of income tax increases are higher in high tax countries compared in low tax countries. Second, “spillover” effects of income taxes are small, but larger in small open economies; third, it is really plausible that large countries have important roles in inducing “spillover” effects. Fourth, energy taxes has small direct effects, but relatively large “spillover” effects and last, distortions increases with labour market rigidities.

However, a less intuitive result is described by the statement: “A low degree of flexibility would result in lower welfare losses as wages adjust less to lower labour demand”. This result is probably true in the short term, but in the longer run there would be negative effects on employment (hysteresis effects) and on production resulting over time in lower welfare. This puts a question mark on the time horizon of the used model.

My concluding remarks are, first, that analyses of effects of tax changes in a coordinated European perspective, taking into account spillover effects, really is interesting and a promising strand of research. Second, a more detailed description of spillover channels and mechanisms given by the used model would be welcomed. And last, it would be interesting to see analyses of a broader set of taxes and their effects by the use of the presented analytical framework.

Session 4

POLICIES TO PROMOTE SUSTAINABLE GROWTH

FISCAL CONSOLIDATION NEEDS AND IMPLICATIONS FOR GROWTH

Douglas Sutherland[*]

Public debt in the OECD area passed annual GDP in 2011 and is still rising. For many countries, just stabilising debt - let alone bringing it down to a more sustainable level – is a major challenge. The debt overhangs can affect growth through channels such as raising the cost of capital. The main focus of this paper however is the implications for growth both in the short term and in the long term of reducing debt levels. Consolidation needs are large and most of the reduction in debt will need to come from improvements in the primary balance. In the short term, the pace of consolidation needs to balance consolidation requirements with the effects of fiscal retrenchment on aggregate demand. The trade-off will depend on the choice of fiscal instrument and on the ability of monetary policy to accommodate consolidation. However, other things being equal, a slow consolidation will ultimately require more effort to meet a fixed debt target. In this context, consolidation should aim to use instruments that are friendly to long-term growth. There is scope to improve budgetary positions by reforming transfer systems, raising the efficiency of public services, eliminating certain tax expenditures and collecting additional revenues from less distortionary tax bases.

Introduction

1 Public debt in the OECD area passed annual GDP in 2011 and is still rising. For many countries, just stabilising debt – let alone bringing it down to a more sustainable level – is a major challenge. Concerns about debt sustainability have manifested themselves in the euro area debt crisis, but could spread beyond that area.

2 Both high debt levels and efforts to reduce them can affect growth. The debt overhangs can affect growth through channels such as raising the cost of capital and increasing the burden of distortionary taxation. The main focus of this paper however is on the implications of reducing debt levels for growth both in the short term and in the long term. In the short term, the trade-off between macroeconomic stabilisation and consolidation creates a particular challenge, especially in an environment when many countries need to implement fiscal consolidation more-or-less simultaneously and with policy interest rates close to the zero lower bound giving little scope for monetary policy to accommodate fiscal consolidation. In this context, fiscal consolidation needs to be carefully designed, notably in the choice of policy instruments which will affect the trade-off not only with short-term but also long-term growth.

3 The rest of the paper is organised as follows: after a brief review of the lead up to the current debt debacle, the second section looks at the impact of high debt on economic growth and establishes consolidation needs, relying principally on fiscal gap calculations, and considers the factors likely to influence debt dynamics; the next section discusses the combined challenge of consolidation and macroeconomic stabilisation. This section also discusses the short-term impact through the multiplier effects of different instruments, with pension reform representing an extreme case of little initial impact but potentially large long-term impact on fiscal sustainability; the

---

[*] OECD Economics Department

The views expressed in this paper are those of the author and do not necessarily represent those of the OECD or its member countries. Comments from Sergey Vlasov and other participants at the workshop and secretarial assistance from Lyn Urmston are gratefully acknowledged. The paper draws on Elmeskov and Sutherland (2012).

following section discusses available policy instruments and their implications for long-term growth. A final section concludes.

The size of debt overhangs

4 Debt levels in the OECD have trended upwards since the early 1970s, with countries often insufficiently ambitious in bringing debt levels down during expansions. Indeed, during the upswing that preceded the recent crisis, underlying deficits were not reduced much, such that debt levels were not brought down, notably in Greece, the United Kingdom and the United States. In some cases, declines in revenue shares during the expansion suggest that governments were engaging in a pro-cyclical easing of fiscal policy – something which has been a consistent feature of policy in some European countries since the early 1970s (Égert, 2010). The impact of lower interest rates and in some cases lower debt on debt servicing and the apparent strength of revenues seduced some governments into cutting taxes and relaxing control over spending. Indeed, new estimates of underlying budget balances that adjust not only for the effect of the economic cycle but also take account of asset price effects on revenues suggest significantly weaker balances as a share of GDP in a number of countries, notably Ireland and Spain (Price and Dang, 2011). As such, when fiscal positions appeared to improve before the financial crisis, they often gave an impression that was too flattering. And in retrospect, given the weaknesses in financial sector prudential policy, fiscal positions were insufficiently robust given the scale of the liabilities and contingent liabilities that some governments had to assume during the crisis.

5 What sets the crisis apart is how widespread and rapid the build-up of debt has been, making the need for fiscal consolidation pressing for most OECD countries. The automatic stabilisers played a role with spending on unemployment benefits surging and tax revenues evaporating. Tax revenues were further dented by asset price movements, which had boosted revenues in the pre-crisis period. Spending further jumped due to support packages and assuming various liabilities. In addition, a downward level shift in potential output as an effect of the crisis effectively meant that prevailing levels of spending became inconsistent with pre-existing tax rates and implied a need to tighten just to stand still. For the OECD as a whole, gross government debt is expected to rise to unprecedented levels, exceeding 100 per cent of GDP for the first time in 2011 (Figure 1). In Japan, this ratio has risen to over 200 per cent of GDP. Even in some low-debt countries gross debt increased quite strongly. Only Norway and Switzerland have bucked the trend, reducing debt levels.

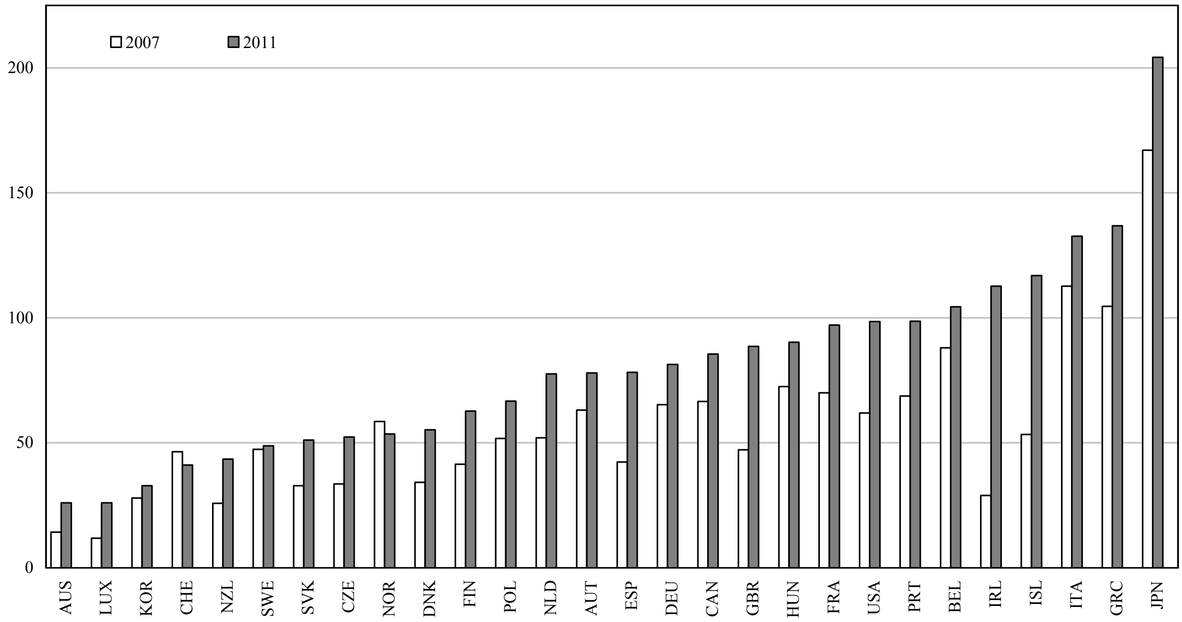
6 In emerging market economies, less debt build-up occurred over the crisis and debt levels are often more favourable than in many OECD countries, not least because high growth rates tend to ease debt dynamics. Nonetheless, in a number of countries debt levels are not negligible. In Brazil and India, debt levels were around 65 per cent of GDP at the end of 2010. Fiscal consolidation is underway in both countries and Brazil is already running a relatively large primary surplus. For India, consolidation will be difficult due to large spending pressures and possibly weaker revenue growth. In China, the official debt burden was low at 19 per cent of GDP in 2010. However, off budget sub-central government and state enterprise debt could potentially raise total debt well over one third of GDP at the end of 2010, with contingent liabilities in the financial sector of uncertain magnitude and the on-going push to provide affordable housing potentially adding to debt.

Consequences of high debt levels for growth

7 High public debt levels may have adverse effects on growth. Higher debt loads could affect output by raising the costs of capital or more speculatively through higher distortionary taxes,

Figure 1

Gross Government Financial Liabilities
(percent of GDP)



Source: OECD Economic Outlook 89 Database.

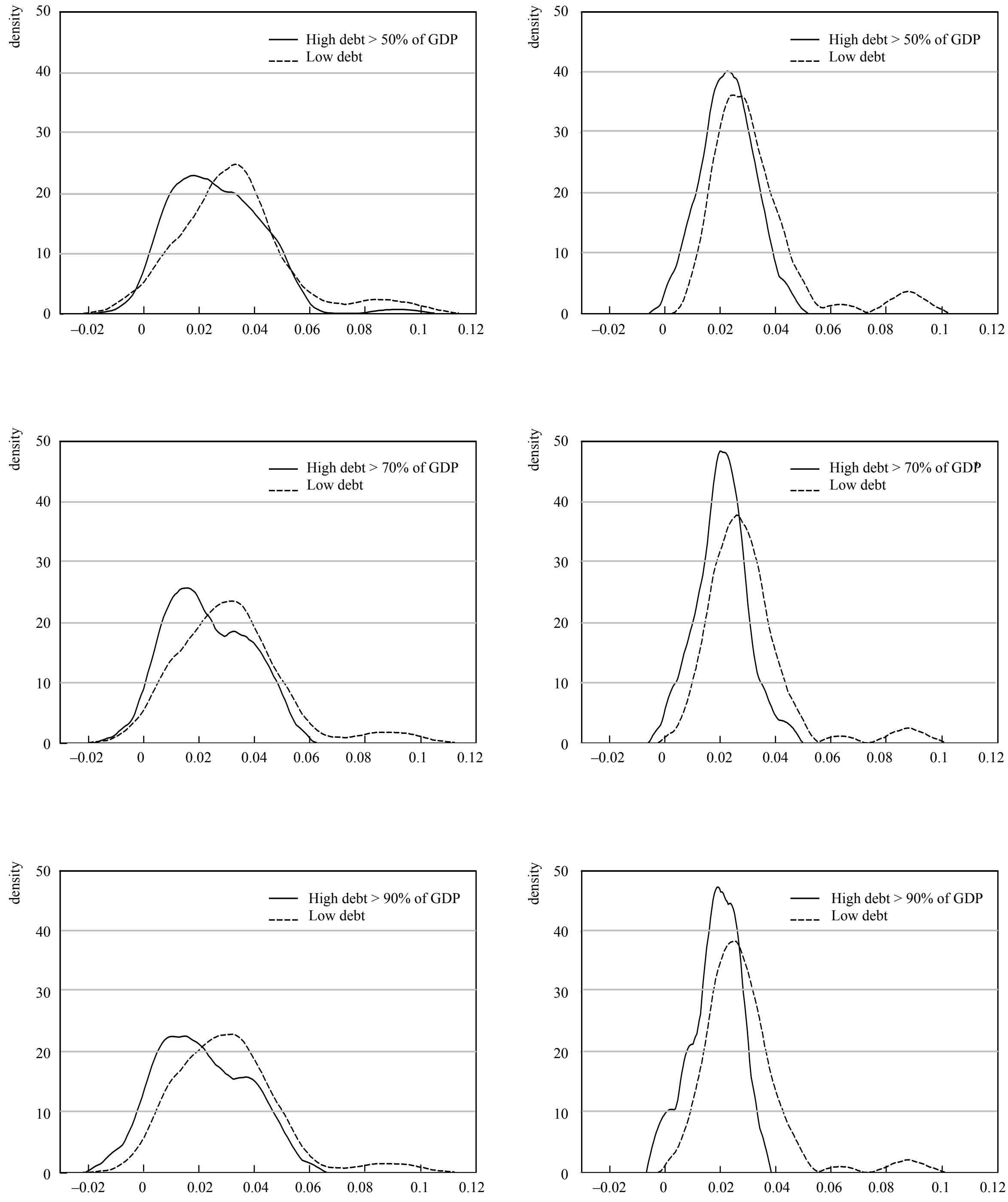
inflation or greater volatility in policy. Cournède (2010) demonstrated the potential impact of higher corporate financing costs, which may be a consequence of not only a normalisation of the artificially low risk premia that prevailed before the crisis but also of crowding out due to higher government issuance of debt. A higher cost of capital is likely to reduce the capital-to-labour ratio and hence productivity. Using the assumptions embodied in the OECD's medium-term baseline and a production function with three factors (labour, business sector capital and oil), the calculations suggest that the level of GDP in the long run would fall by just over 2 per cent in the United States and 2.6 per cent in the euro area for a normalisation of interest rates following the crisis, which would entail a real interest rate shock of around one percentage point in both the United States and the euro area. If higher government debt does lead to crowding out, with the real interest rate shock rising by around an additional percentage point, then the fall in GDP could be more substantial, with the level of output falling by around 5 per cent in both the United States and euro area.

8 The effects of higher costs of capital on the intensity of capital in production should essentially lead to a level shift in potential output and therefore to growth rate effects over some finite period only. More long-lasting effects on economic growth could arise to the extent higher costs of capital lead to reduced investment in research and development. More speculative and uncertain combinations of OECD research suggests that if the fall in potential output by 3 per cent as a result of lower capital intensity were combined with the above higher cost of capital, then the stock of R&D could fall by 5.4 per cent, which would reduce long-run total factor productivity (TFP) by 0.7 per cent, based on an estimated long-run elasticity (Guellec *et al.*, 2004). In practice, evidence on TFP growth in OECD countries before and after past crises suggests that experience is very heterogeneous (Haugh *et al.* 2009). Since impacts of debt via R&D should be expected to accrue via TFP, this underlines the need to treat the calculations with care.

Figure 2

Growth Conditional on Past Debt Levels

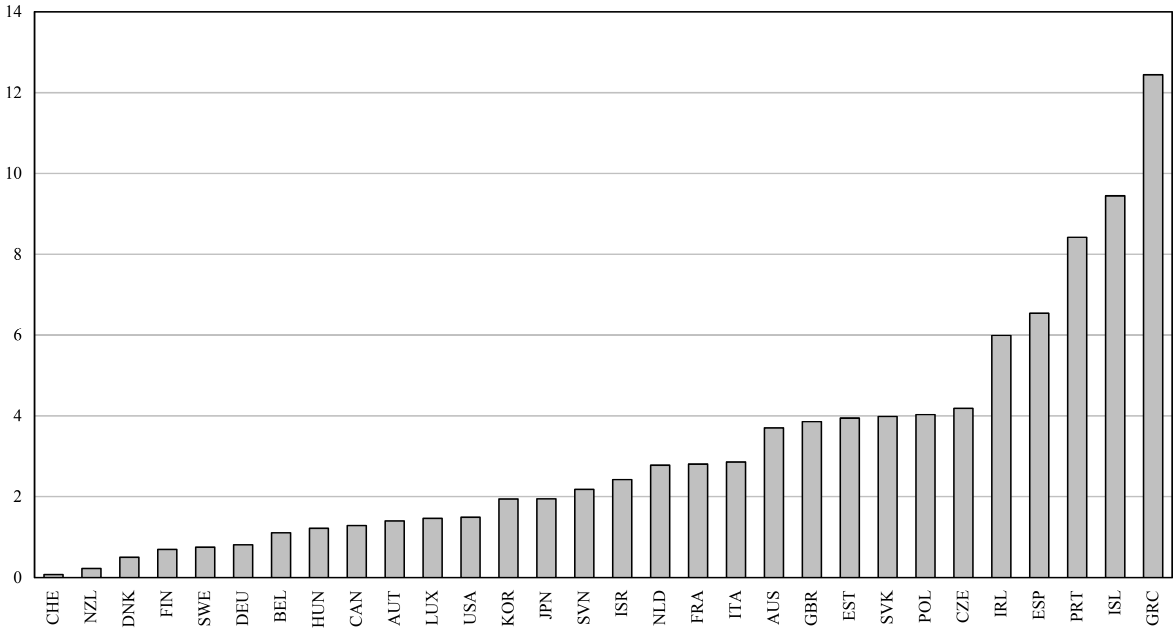
(left hand panel: growth in the following 5 years; right hand panel: growth in the following 10 years; top panel: debt threshold 50 per cent of GDP; middle panel: debt threshold 70 per cent of GDP; bottom panel: debt threshold 90 per cent of GDP)



Note: The distributions are kernel densities for growth rates in the subsequent 5 and 10 years when growth rates are above and below the given threshold.

Figure 3

Cumulative Fiscal Tightening Between the Deficit Trough and 2012
(change in underlying primary balance, percent of GDP)



Source: OECD Economic Outlook 90 Database.

9 Empirical work has identified various thresholds in the relationship between public debt and growth. For example, Reinhart and Rogoff (2010) found that growth rates in both developed and developing countries where the public debt to GDP ratio exceeds 90 per cent are about 1 per cent lower than in the less indebted countries (Cecchetti *et al.,* 2011 find a similar threshold effect). In a similar vein, Caner *et al*. (2010) found a threshold effect on growth rates at 77 per cent of GDP for a large sample of countries, with the threshold being lower for emerging markets, and Kumar and Woo (2010) found that a 10 percentage point increase in debt reduces annual real per capita GDP growth by 0.2 percentage points per year, with the effect being smaller for advanced economies and some evidence for non-linearity beyond a debt/GDP ratio of 90 per cent of GDP.

10 Indeed, fitting density functions to growth rates of OECD countries suggests that growth is typically lower in periods that follow years of high debt (Figure 3). This is more obvious when looking at growth rates over a short window of 5 years, where some of the effect may reflect that high debt is followed by consolidation with negative effects on the cycle. However, the effect appears to persist over 10 years when cyclical effects of consolidation should matter less. Even so, the relationship could be spurious to some degree given the secular tendency for debt levels to drift up and growth rates to trend down which may account for some of the relationship. Moreover, causality may be less than clear with, for example, less well managed countries likely to have both high debt and low growth.

11 In sum, high debt levels are likely to have negative impacts on growth. Hence, there are good reasons for many countries to reduce their debt overhangs, including creating room to react to future shocks. Reducing debt in turn has implications for growth both in the short and long term, with the scale of the necessary adjustment likely to give some indication of how painful fiscal consolidation will be. We turn to this issue in the next section.

Size of adjustment

12 Facing large debt overhangs, many countries have already started fiscal consolidation, which has implications for economic growth in the short term. In some cases, notably for those countries most under pressure from the bond markets, the on-going and announced tightening is substantial, rapid and unusually correlated by historical comparison (Figure 3). Between the trough (measured by the underlying primary balance) following the onset of the crisis, which was 2009 for most countries, and the projected value for 2012, five countries are expected to tighten by more than 5 per cent of GDP (Greece, Iceland, Ireland, Portugal and Spain). In 11 other countries, underlying primary balances are expected to have tightened by more than 2 per cent of GDP. Recent policy announcement imply that these numbers would be larger if recalculated today.

13 Additional fiscal consolidation will be required beyond 2012. Recent OECD work has assessed these post-2012 needs, both in terms of stabilising debt over the medium term and also meeting prudent long-term debt targets. The consolidation requirements to stabilise debt (OECD, 2011c), are based on stylised assumptions about a sustained and gradual annual tightening of the underlying primary balance by 0.5 per cent of GDP until debt stabilization is reached. The long-term fiscal gaps on the other hand make an alternative stylised assumption that the tightening will be implemented immediately and sustained until 2050 to meet a specific debt target (Merola and Sutherland, 2011). Both sets of assumptions ignore the implications for output, which will obviously be important (discussed below). Both approaches come to similar conclusions on the need for consolidation, but here we concentrate on the long-term fiscal gap calculations, which will be used later in the paper to illustrate consolidation options.

Fiscal gaps

14 The fiscal gap shows the *immediate* and *permanent* improvement in the underlying primary balance that is required to ensure that debt meets a target at a certain point in time, based on a simplified model of the economy and a number of assumptions about growth, interest rates, inflation and underlying fiscal policy (see Appendix). [1] The presentation of the results below typically reports the fiscal gaps for ensuring gross financial liabilities is 50 per cent of GDP in 2050 (Box 2). This is intended to be illustrative and not normative. Indeed, different debt targets will be appropriate for different countries. For example, a low gross debt target may be less compelling for countries with large government financial asset holdings. In other cases, the public has demonstrated a preference for very low levels of debt. Countries with large implicit liabilities due to a large financial sector may wish to err on the side of caution. Although the 50 per cent target is arbitrary it may nonetheless be supported by some arguments. Thus, empirical estimation suggests that changes in the functioning of the economy occur around debt levels of 70-80 per cent of GDP. For example, interest rate effects of debt seem to become more pronounced (Egert, 2010), offsetting saving responses to discretionary policy changes become more powerful (Röhn, 2010) and, as illustrated above, trend growth seems to suffer. Building in a safety margin to avoid exceeding the 70-80 per cent levels in a downturn may suggest aiming for 50 per cent or thereabout during normal times. In any case, over a very long period such as up to 2050, the size of fiscal gap does not depend strongly on the particular target debt level (see opposite).

---

1 Following a severe economic dislocation, estimating potential output and thereby the underlying primary balance represents a challenge. While the fiscal gap simulations do not directly assess uncertainties about potential output, the variety of simulations reported below reveal how varying different parameters affect the fiscal gap calculations.

Box 2
DEBT OBJECTIVES

Various choices have to be made in setting a debt target:

The target can be based on either gross or net debt/financial liabilities. Gross financial liabilities are a visible headline indicator and typically the measure used in empirical analysis. Net financial liabilities are in principle more appropriate when considering long-term sustainability, though government net worth, which also takes into account non-financial assets (the public capital stock), may be the more appropriate when also considering inter-generational issues. However, there are serious problems due to lack of comparability across countries, particularly when valuing government non-financial assets. Furthermore, government assets may not be easily used to offset liabilities, at least in the short term. For example, it may not be advisable to privatise public enterprises operating in sectors with significant market failures or when financial markets could not easily absorb large asset sales. There may also be asymmetries across levels of government and with social security funds between the holding of assets and liabilities.

The scope of the public sector can vary. For example, the debt target may affect only the central government, general government or an even wider definition, including for instance, public enterprises. The choice can make a sizeable difference. In the United Kingdom, recent whole of government accounts estimated net liabilities to be 84.5 per cent of GDP in 2009-10, whereas the national accounts net liabilities measure was 52.8 per cent of GDP (HM Treasury, 2011).

The target should address the effect of ageing on entitlement spending (ageing is not the primary driver of health spending but is used as a catch-all label here). The appropriate degree of consolidation will need to take into account the impact of ageing-related spending. Ageing-related spending pressures stem from two factors. First, in many OECD countries spending ramps up with the demographic transition as the post-war "baby boomers" move into retirement. As this transition is either already happening or is imminent, the policy options are limited. In this light, the "hump" in spending may need to be absorbed and adds to the consolidation requirement. A second, uncertain but potentially huge or even infinite, ageing effect on spending stems from longevity, which has been more or less steadily rising for more than 150 years across OECD countries. In this case, the appropriate response is to reform pension and other benefit systems, such as long-term care, rather than to attempt to pre-save to finance the rising ageing-related spending. Attempting to pre-save for future increase in longevity rather than adjusting pension and other programmes would be unfair across generations and would be difficult in light of uncertainty concerning the development of longevity.

More generally, the target should also consider inter-generational fairness. Pay-as-you-go pension systems present an obvious example of a transfer of resources between generations. Likewise, "excessive" deficits can transfer liabilities to future generations. In other cases, investment can create assets which will be enjoyed by future generations. As such, the degree of consolidation will need to consider the source of the transfer between generations and how much of a burden it is fair to pass onto future generations.

15 The fiscal gaps should be seen as giving a common metric for assessing the need for fiscal consolidation rather than being normative about how such a consolidation should be implemented. When the fiscal gap is large, it would be difficult to implement such a large consolidation effort immediately. Furthermore, sustaining the fiscal policy tightening, even seemingly modest ones, over very long periods may also present a considerable challenge. Finally, as the fiscal gaps are based on meeting arbitrary debt targets in 2050, the evolution of gross debt is unlikely to be stable as a share of GDP at the end of the simulation. In some cases, for example, the fiscal gap will involve substantial undershooting of the debt target early in the simulation, masking pressures on public finances that will continue to mount beyond the end of the simulation.

Baseline simulation

16 The baseline simulation shows the immediate tightening of the underlying primary balance in 2013 needed to ensure that gross financial liabilities are 50 per cent of GDP in 2050. The baseline assumes that pension, health and long-term care spending is constant as a share of GDP and, as such, the fiscal gaps present the minimum that is required to meet consolidation needs in the case when pensions and health schemes are reformed to alleviate any upward pressure on spending or when other spending categories are curtailed and taxes raised to accommodate such spending pressures (simulations incorporating spending pressures emanating from pensions, health and long-term care are presented below).

17 Fiscal gaps differ across countries mainly because of large differences in underlying deficits at the starting point and to some extent due to differences in the level of initial debt (Table 4 in the Appendix). Countries already undertaking large fiscal consolidations (Greece, Iceland, Portugal and Spain) generally face moderate fiscal gaps on the assumption that the present large improvements in underlying primary balances are maintained. Countries where underlying deficits are expected to remain substantial in 2012 face much larger fiscal gaps. For example, the fiscal gaps for Japan, the United States, the United Kingdom and New Zealand exceed 5 per cent of GDP. On the other hand, a number of countries – Korea, Luxembourg, Sweden and Switzerland – do not face any additional tightening requirements to meet the debt target. It may seem ironic that euro area countries with relatively modest fiscal gaps are the victims of a virulent debt crisis whereas other countries with much larger fiscal gaps enjoy very low bond yields at present. This partly reflects concerns about potential needs for intervention in euro area banking systems, but also that euro area debt essentially corresponds to foreign currency denominated debt for the individual country. Lately, pressures may also have reflected increased concerns about the integrity of the euro area more generally.

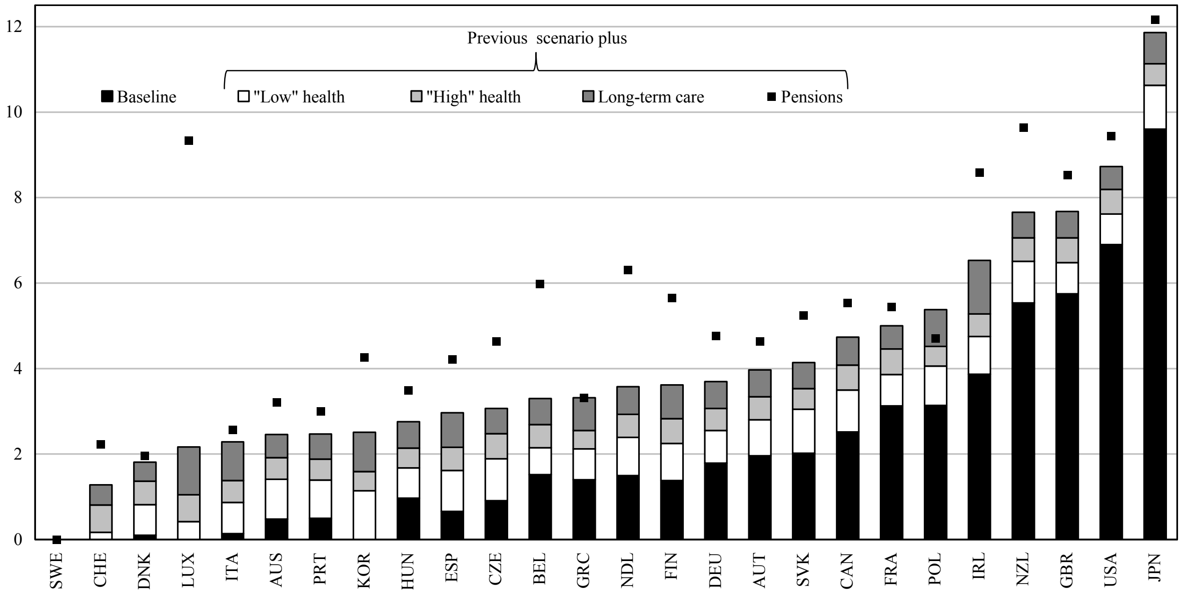
18 When spending pressures projected to arise from health and long-term care and pensions are included, all countries, with the exception of Sweden, will require significant additional fiscal consolidation.

- In the case of health care spending, higher levels of spending are not necessarily undesirable, but financing higher spending can create difficulties (Hall and Jones, 2007). Two different sets of health care spending projections are used (Oliveira-Martins and de la Maisonneuve, 2006). The average projected increases in health and long-term care spending by 2050 are 3½ per cent of GDP in a low spending scenario, when it is assumed that spending increases above those related to demographic change and to a unitary income elasticity will gradually fade, and around 6 per cent of GDP in a high spending one. As the projected increases are relatively similar across countries, because health spending is not primarily driven by demographics but rather to a large extent by expected supply developments, the impact on the fiscal gaps does not vary much across countries. Nonetheless, the fiscal gaps rise over 1.5 per cent of GDP in Canada, the

Figure 4

Fiscal Gaps, Baseline and with Health and Long-term Care Spending and Pensions

(immediate rise in the underlying primary balance needed to bring gross financial liabilities to 50 per cent of GDP in 2050, percent of GDP)



Note: "Low" health assumes policy action curbs health spending growth. "High" health is the additional cost pressure in the absence of these policy actions.

Czech Republic, Japan, New Zealand and Switzerland when greater cost pressures affect health spending (Figure 5).

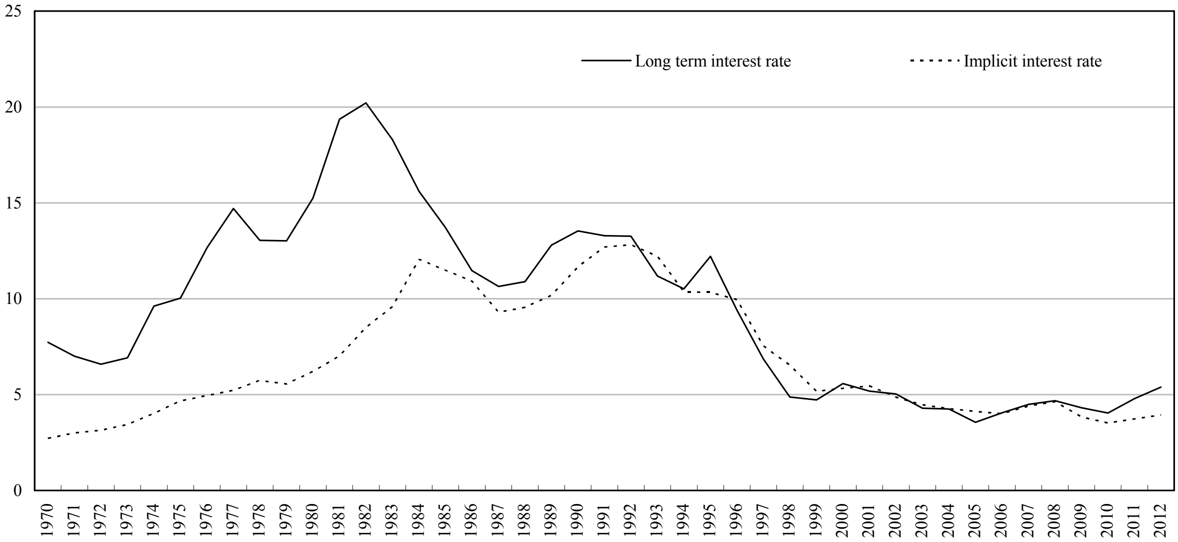
- Including pension spending alters radically the fiscal gaps for many countries relative to the baseline scenario (Figure 4).[2] The fiscal gaps of the countries facing the largest pension problems, such as Luxembourg, Belgium and the Netherlands underscore that meeting these challenges would be better addressed by reform rather than pre-saving. In some cases, such as Greece and Spain, reforms to the pension systems in 2010, which are incorporated in the projections, have addressed significant pressures emanating from this source. In Sweden and Poland, the notionally-defined contribution pension system means that no additional or even less tightening is required to meet a gross financial liabilities debt target of 50 per cent of GDP in 2050.

19 The fiscal gaps do not change markedly relative to the baseline if alternative debt targets are used. This occurs because even relatively small changes to underlying fiscal positions add up when maintained for 40 years. It is the same effect that lies behind initial debt levels having an only modest effect on fiscal gaps compared with initial deficit levels. Taking government financial assets into consideration may indicate that fiscal positions are in relatively better shape, notably for Japan. In other cases, such as in Finland, the large net asset position reflects pre-funding for pension spending.

[2] The pension projections are based on OECD (2011a). For Greece and Spain, estimates of the impact of reforms in 2010 and a change in the law in 2011, respectively, are used. For the United States, estimates from CBO (2011) are used. For most European countries, public sector occupational schemes are included. This is not the case for Canada and Japan. The path of projected public pension spending is phased in so that the spending profile follows the profile of the old-age dependency ratio.

Figure 5

Borrowing Rates in Italy
(percent)



Source: OECD Economic Outlook 90 Database.

Debt dynamics

20 How will the debt overhang be worked off? A review of episodes of declining debt since the early 1970s suggests that improvements in the primary balance are more consistently important in reducing debt, though at times interest rate and growth dynamics can help.[3] One possible decomposition of past debt developments shows the difference between the inertial contributions of debt dynamics on the one hand and the more direct policy lever of the primary balance on the other (Table 1). When debt has been falling in recent decades this has been typically accompanied by the primary balance having a negative effect on debt. The real interest rate and real growth rate effects often offset one another. That said, in some countries during the 1970s, negative real interest rates had an effect allowing them to run larger primary deficits.

The effects of stronger productivity growth

21 Going forward, debt dynamics can be influenced by stronger productivity growth. To illustrate this, simple calculations reveal the effect of productivity growth on debt levels over a 10 year period (Table 2). Extending the calculation beyond the medium term would have a larger impact. Nonetheless, for the countries with the largest fiscal gaps, while productivity gains would help, the fiscal challenge remains large. In these calculations, interest rates are assumed not to change, although they would likely rise with a boost in productivity, thereby undoing some of the potential gains. On the other hand, if government spending did not rise fully in line with GDP, the gains from higher growth could be substantial by improving the underlying primary balance.

---

[3] In earlier periods of very high debt, overhangs were worked off by rapid growth, primary balances and negative real returns, helped in some cases by financial repression (see below). For example, Hall and Sargent (2011) estimate that the debt reduction as a per cent of GDP in the United States between 1945 and 1974 was mainly the result of high growth and primary surpluses with about one-fifth of the reduction stemming from negative real returns due principally to high inflation.

Table 1

Episodes of Falling Debt: The Contribution of the Primary Balance, Inflation and Growth

Country	Episode	Total Change in Gross Financial Liabilities, *(percent of GDP)*	Change After Accounting for Valuation Effects	of which:		
				Primary Balance	Real Growth	Real Interest
Australia	1996-2008	–27.7	0.0	–24.0	–11.6	17.6
Belgium	1994-2007	–52.6	0.0	–64.0	–37.8	63.9
Canada	1971-1976	–11.7	0.0	5.7	–10.9	–1.1
	1997-2000	–19.6	0.0	–21.7	–17.1	25.5
	2002-2007	–16.1	0.0	–13.6	–11.6	17.1
Denmark	1985-1989	–12.5	0.0	–31.7	–7.5	21.3
	1994-2007	–58.0	0.0	–41.3	–23.0	34.5
France	1999-2001	–6.0	0.0	–3.4	–5.8	6.8
Germany	1999-2001	–2.4	0.0	–5.0	–4.0	8.8
Italy	1999-2003	–15.7	0.0	–16.6	–9.2	14.6
Japan	1988-1991	–13.6	0.0	–11.2	–14.0	10.0
Spain	1999-2007	–33.2	0.0	–19.8	–19.3	3.2
Sweden	1985-1990	–24.6	0.0	–25.8	–9.4	13.4
	1997-2003	–23.6	0.0	–17.9	–15.9	18.7
United Kingdom	1972-1976	–20.1	0.0	8.6	–6.9	–12.8
	1978-1981	–11.8	0.0	4.5	–1.5	–4.8
	1985-1990	–18.3	0.0	–9.1	–9.1	11.9
	1999-2001	–12.2	0.0	–12.1	–4.6	5.4
United States	1972-1974	–5.4	0.0	–1.5	–4.5	–0.7
	1976-1979	–3.5	0.0	–0.2	–5.4	–0.2
	1994-2001	–17.4	0.0	–15.6	–18.5	24.7

Note: the decomposition is based on the relationship: $d_t - d_{t-1} = \frac{r_t}{1+g_t} d_{t-1} - \frac{g_t}{1+g_t} d_{t-1} - pb_t$, where d is the debt as a ratio of GDP, r is the real interest rate, g is the real growth rate and pb is the primary balance as a ratio of GDP.

Table 2

The Effect of Higher Productivity on the Real Growth Effect
(reduction in initial debt stock as per cent of GDP after 10 years with growth in the baseline (OECD Economic Outlook 89 medium term baseline) and with growth rates raised by 0.25 and 0.5 basis points)

Country	**Initial Debt Level** *(percent of GDP)*	**Real Growth Effect** *(percent of GDP)*		
		Baseline	**+ 0.25 Basis Points**	**+ 0.5 Basis Points**
Australia	31	7.2	7.5	8.0
Austria	82	12.9	14.2	15.6
Belgium	100	13.5	15.2	16.9
Canada	88	15.2	16.5	17.9
Czech Republic	51	10.7	11.3	12.0
Denmark	60	8.2	9.2	10.2
Estonia	19	4.1	4.3	4.6
Finland	66	12.1	13.1	14.1
France	100	16.0	17.6	19.2
Germany	87	9.5	11.1	12.7
Greece	159	35.8	37.7	39.9
Hungary	81	11.5	12.9	14.2
Iceland	120	26.7	28.1	29.8
Ireland	126	34.8	35.9	37.4
Isreal	70	19.5	20.2	21.1
Italy	128	14.7	17.1	19.4
Japan	219	32.8	36.4	40.0
Korea	33	7.5	7.9	8.4
Luxembourg	24	5.9	6.1	6.4
Netherlands	75	8.5	9.9	11.3
New Zealand	52	10.8	11.5	12.2
Norway	51	11.8	12.4	13.1
Poland	66	10.7	11.7	12.8
Portugal	116	26.7	28.1	29.6
Slovak Republic	51	10.2	10.9	11.6
Slovenia	56	6.8	7.9	8.9
Spain	75	17.8	18.7	19.7
Sweden	41	6.9	7.5	8.2
Switzerland	37	6.0	6.5	7.1
United Kingdom	93	17.1	18.5	19.9
United States	107	22.3	23.7	25.2

Inflation and interest rates

22 One possible way to deal with a high debt level is to erode it through higher inflation, but this is likely to be accompanied by drawbacks. Higher inflation is most likely to have an effect in an environment when debt is non-indexed, maturity is relatively long and rollover requirements are low, given that interest rates are likely to respond to higher inflation rates.[4] Even in this case, simulations presented in the *OECD Economic Outlook* 89 show that the contribution of inflation to reducing debt is modest (OECD, 2011c). For a standard country with debt around 100 per cent of GDP and an average maturity structure, 1 percentage point on inflation would typically reduce the debt ratio by some 5-6 percentage points assuming the interest rate on new borrowing rose in tandem with inflation. Getting debt to even lower levels would correspondingly require higher permanent inflation rates. The drawbacks of such an approach to reducing debt would be felt principally through the negative growth effects of higher rates of inflation, some of which may accrue through associated higher price volatility as well as distortions created through interactions with the tax and benefit system (Edey, 1994).

23 For higher inflation to make a marked dent in debt levels, some form of financial repression would probably be needed to ensure interest rates remain low relative to inflation.[5] Following the end of World War II until the beginning of the 1980s, financial repression often played a role in reducing the huge stocks of debt accumulated during the war. Reinhart and Sbracia (2011) estimate that financial repression contributed to a "liquidation effect" which, for example, amounted to a reduction of Italian government debt of around 5 per cent annually. Figure 5 presents suggestive evidence of financial repression during the 1970s, particularly after mid-decade when inflation was no longer surging, during which a large wedge existed between the yield on 10 year government bonds and the effective interest rate the government was paying on debt. While financial repression may be one avenue to liquidate debt there are adverse consequences. For example, Jonung (2011) argues that the imbalances which developed as a cause of financial repression contributed directly to financial crises in the Nordic countries in the late 1980s and early 1990s.

Dynamics of adjustment

24 The previous section suggested that relying on favourable debt dynamics to address the debt overhang may not be a viable option. Hence, improvements in the primary balance are called for. The pace of consolidation needs to balance consolidation requirements with the effects of fiscal retrenchment on aggregate demand. Ideally, in the short term, the pace should depend on the state of the public finances, the strength of the recovery, the ability of monetary policy to cushion the demand effects of fiscal tightening, and the need to signal a credible commitment to fiscal consolidation. However, there are significant uncertainties surrounding several of these factors, which make gauging the appropriate pace of consolidation complicated. These uncertainties would argue for a consolidation strategy that could be implemented flexibly, capable of adjusting the speed and intensity as new information becomes available. Moreover, it argues for implementation that initially favours policies with comparatively low multipliers and reforms that underpin credibility, but have little negative effect on demand in the short run. For example, pension reforms can have large effects on long-term sustainability and may have little negative effect in the short

---

4 Aizenman and Marion (2009) show for the United States that the maturity structure of publically-held debt is shorter than in the post-war period, reducing the incentive to use inflation to reduce the debt overhang. On the other hand, a larger share of debt is held by foreigners, which pulls in the opposite direction.

5 Financial repression includes directed lending to government by captive domestic lenders, caps on interest rates, regulation of cross-border capital movements and a tighter connection between government and the operation of banks.

term. Indeed, insofar as postponed retirement reduces the need for future pensioners to save for retirement there could in principle even be a positive effect.

The pace of consolidation

25 Given high government debt-to-GDP ratios, some countries run the risk of unsustainable debt dynamics developing, especially if financing costs spike because of lack of credibility. While interest rates on government debt remain relatively low in many countries, debt levels in the wake of the crisis are significantly higher, implying latent upward pressure on borrowing costs. When interest rates are linked to government debt levels, this can tilt the case towards earlier consolidation. Even moderate delays may incur high costs with the development of particularly adverse debt dynamics (Corsetti *et al.*, 2011). On average for the OECD, interest payments accounted for around 2.5 per cent of GDP in 2007, but higher debt levels coupled with a normalisation of interest rates could push up interest payments to over 4 per cent of GDP in 2026 (OECD, 2011c). Thus, in countries which are particularly exposed to a financial market reaction the extent of consolidation may need to be larger and the pace faster than may be optimal if the main concern was the strength of the recovery.

26 With policy rates low in many countries, and the zero lower bound still an important constraint, monetary policy is unlikely to be able to offer much support, arguing for a gradual phasing in of consolidation measures. As economies recover, monetary policy is less likely to be constrained by the zero bound and thus the pace of consolidation could be increased. Another argument for slower consolidation may arise when governments consolidate simultaneously; the implications for output are more severe due to international spillovers. Simulations reported in OECD (2009) suggest that multipliers increase by a factor of ¼-½ in major OECD regions when they consolidate jointly as opposed to individually.

27 The impact of fiscal consolidation on economic activity will depend on the size and time profile of the fiscal multipliers (Barrell *et al.* 2012). Differences across countries are largely related to the size and openness of the economy, the size of the public sector, the degree of dependence of consumption on current income and also the flexibility of the economy. The multipliers in the NiGEM model tend to be largest for government consumption, whereas tax impulses tend to have lower multipliers than spending. The differences in multipliers across instruments suggest that the sequencing of fiscal consolidations could start with tax increases before cutting government spending, though political economy considerations may suggest otherwise. Beyond the model-based multipliers, pension reform that delay retirement may, as argued above, have particularly attractive features.

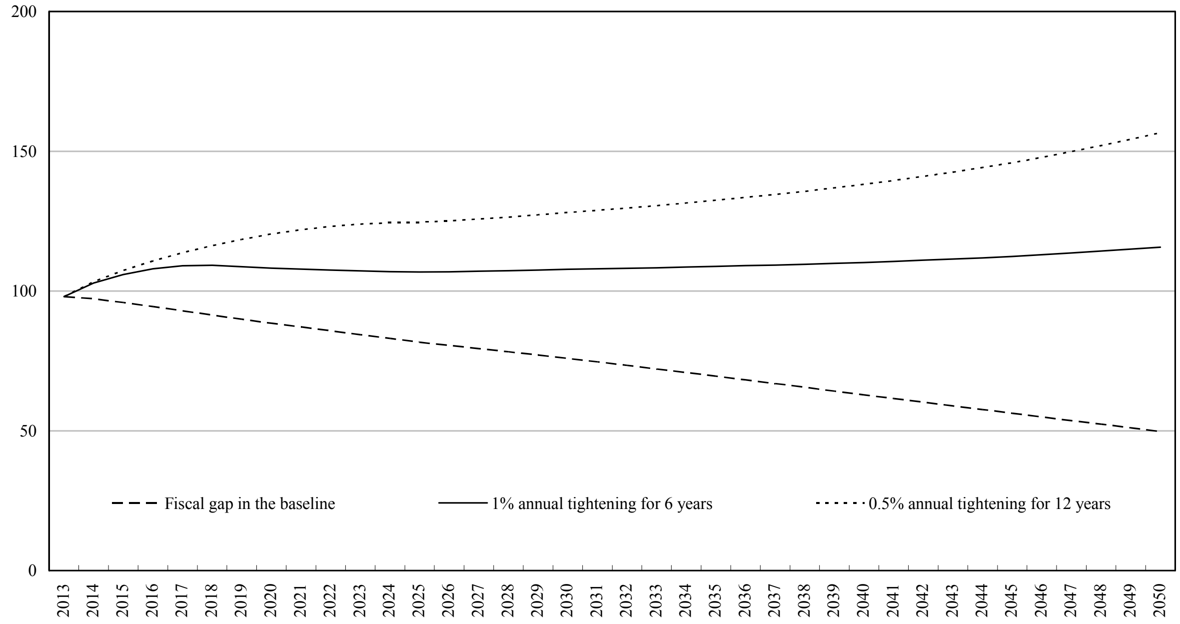
Consequences of gradual and delayed consolidation needs

28 When the state of fiscal policy doesn't dictate the pace of consolidation, more gradual tightening may minimise the short term pain but require a larger overall amount of consolidation. Simulations for the United States, using the long-run model behind the fiscal gaps and therefore assuming no impact of consolidation on output, shows that gradual tightening could allow adverse debt dynamics to develop (Figure 6). Thus, too slow a consolidation may require further fiscal tightening to bring debt down to prudent levels. This arises because debt levels above a threshold of around 75 per cent of GDP are assumed to incur a higher risk premium of four basis points for each additional percentage point of debt (Egert, 2010). Using the model, fiscal gap calculations examining the consequences of a short delay to fiscal consolidation generally find that for most countries this has little effect on the necessary tightening, as long as the subsequent consolidation is large, as implied by the fiscal gap. However, for countries where actual debt is high or current

Figure 6

The Pace of Fiscal Tightening

(evolution of gross financial liabilities for the United States when the underlying primary balance is tightened so that debt is 50 per cent of GDP in 2050 and the consequences of phasing in the same tightening more gradually, percent of GDP)



deficit levels imply a particularly rapid run-up in debt, such as New Zealand, the United Kingdom, the United States and Japan, even a short delay would visibly increase the required tightening of the underlying primary balance to reach prudent debt levels.

Long-term growth and choice of instruments

29 The scale of consolidation needs suggests that consolidation should aim to use instruments that are friendly to long-term growth. In addition, supporting structural reforms can help, both through their implied effects on primary budget balances and to the extent higher growth is beneficial for debt dynamics. As concerns the primary balance and the respective contributions from lower spending and higher revenues, the "optimal" size of government is not known. However, the marginal net social costs - including the excess burden of taxation – of additional public spending are usually thought to increase more than proportionately with the additional taxation needed to finance spending. Hence, given the current high level of public spending in many OECD countries and the future spending pressures due to population ageing, a large part of consolidation probably should consist of cuts in public spending and addressing drivers of future spending pressures. In countries where spending is low, greater emphasis may have to be put on revenue measures.

30 Given that spending cuts are largely unavoidable, a key question is how to maximise the positive and minimise the negative impacts on long-run growth, while at the same time considering other policy objectives such as equity concerns. In some cases, rethinking how distributional goals are achieved may offer scope to reduce transfers while encouraging greater labour force

Table 3

Quantifying the Contribution of Various Policy Instruments to Fiscal Consolidation

(percent of GDP)

	AUS	AUT	BEL	CAN	CHE	CZE	DEU	DNK	ESP	FRA	FIN	GBR	GRC	HUN	ISL
1. Social transfers															
A. Family benefits	0.5	0.7	0.6	-	-	0.1	-	1.4	-	1.1	0.9	1.3	-	1.4	1.0
B. Disability benefits	-	0.3	0.2	-	0.5	0.5	-	1.3	0.6	-	0.9	0.3	-	0.6	-
2. Pensions															
A. Eliminate tax breaks	2.7	0.1	0.1	2.0		0.1	0.8		0.2	0.0	0.1	1.2			1.0
3. Health care															
A. Increase efficiency	0.5	1.8	2.1	2.5	0.5	1.3	1.3	2.8	1.6	1.3	2.5	3.7	3.9	1.7	1.9
4. Education															
A. Increase efficiency in primary and secondary education	0.4	0.4	0.5	0.2	0.2	0.2	0.4	0.6	0.2		0.2	0.2		0.3	1.1
B. Introduce or raise tuition fees for tertiary education	-	0.4	0.4	-	0.4	0.3	0.4	0.4	0.2	0.3	0.4	-	0.4	0.4	0.3
5. Government wage bill															
A. Restore public-private sector pay relativities	-	0.3	0.6	-	-	0.4	0.2	2.0	1.0	-	0.5	1.8	-	-	-
6. Reduce subsidies as share of GDP to OECD average	-	2.3	0.8	-	2.4	0.7	-	1.2	-	0.2	-	-	-	-	0.4
7. Broaden VAT base	0.6	-	1.4	-	-	-	0.4	-	1.4	1.4	0.1	1.8	2.0	0.1	0.8
8. Introduce or increase taxes on immovable property	-	0.8	0.6	-	0.9	0.8	0.6	-	0.3	-	0.5	-	0.8	0.7	-
9. Environmental taxes															
A. Cut GHG emissions to 20 per cent below 1990 levels via an emission trading system with full permit auctioning	4.2	1.8	1.8	2.5	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	

Table 3 (continued)

Quantifying the Contribution of Various Policy Instruments to Fiscal Consolidation
(percent of GDP)

	IRL	ITA	JPN	KOR	LUX	MEX	NLD	NZL	NOR	POL	PRT	SVK	SWE	TUR	USA
1. Social transfers															
A. Family benefits	0.7	-	-	-	1.2	-	0.1	1.1	0.9	-	-	-	1.4	-	-
B. Disability benefits	-	-	-	-	0.1	-	0.8	0.7	1.8	0.6	0.3	-	1.3	-	-
2. Pensions															
A. Eliminate tax breaks	1.2	0.0	0.7		0.5	0.2			0.6	0.2	0.1	0.2			0.8
3. Health care															
A. Increase efficiency	4.8	1.1	0.8	0.6	2.0	0.7	2.7	2.6	1.5	1.5	1.0	2.7	2.7	1.5	2.7
4. Education															
A. Increase efficiency in primary and secondary education	0.3	0.4	0.2	-	0.5	-	0.3	0.3	0.8	0.2	0.1	0.2	0.5	-	0.8
B. Introduce or raise tuition fees for tertiary education	0.3	0.2	-	-	0.4	0.1	0.2	-	0.4	0.1	0.1	-	0.4	0.4	-
5. Government wage bill															
A. Restore public-private sector pay relativities	0.9	1.1	0.6	-	0.8	-	0.3	0.9	-	2.2	-	0.8	0.7	-	0.5
6. Reduce subsidies as share of GDP to OECD average	-	-	-	-	0.2	-	0.1	-	0.7	-	-	0.2	0.1	-	-
7. Broaden VAT base	0.4	2.6	-	-	-	2..5	-	-	0.2	1.4	1.2	0.6	-	3.3	
8. Introduce or increase taxes on immovable property	0.2	0.4	-	0.0	0.9	0.8	0.4	-	0.7	-	0.3	0.6	0.2	0.9	-
9. Environmental															
A. Cut GHG emissions to 20 per cent below 1990 levels via an ETS with full permit auctioning	1.8	1.8	1.2		1.8		1.8	4.2		1.8	1.8	1.8	1.8		2.2

Notes:
An empty cell indicates that no information was available. Cells with a dash indicate that no savings are available from this source.
Estimates for family benefits are based on reducing the figure reported in the OECD Socex Database to the unweighted OECD average as a per cent of GDP.
Estimates for disability benefits are based on reducing the figure reported in the OECD Socex Database to the unweighted OECD average as a per cent of GDP.
The elimination of tax breaks for retirement is based on data for 2007 from OECD (2011), Pensions at a Glance.
Health care efficiency estimates are from Joumard *et al.* (2010).
Education efficiency estimates are based on Sutherland et al. (2007) updated to 2007 spending figures.
Tuition fees for tertiary education are based on raising direct household expenditure for tertiary education institutions to the unweighted average of those countries where households spend on this category.
Government wage relativities are based on returning the government to private sector wage ratio in the early 2000s.
Estimates for subsidies are based on reducing national account data for 2009 to the unweighted OECD average.
The figures for broadening VAT base assume collection efficiency rises to the unweighted OECD average.
The figures for immovable property are based on the unweighted average for 2008 from the Revenue Statistics.
Revenues from greenhouse gas emissions are based on de Serres *et al.* (2010).

participation. In other cases, scope to minimise costs exists by aiming to improve both allocative efficiency (better use of resources) and technical efficiency (maximising output for a given level of inputs). In most OECD countries, fiscal consolidation will also entail revenue reforms. There is scope to increase revenue by base broadening measures, particularly targeting so-called tax expenditures. When marginal rates need to go up, orientating measures towards those tax bases that have less distortionary effects can help to make fiscal consolidation on the revenue side less costly to long-term output. Finally, taxation of negative externalities may improve both welfare and public budgets.

Instrument options

Social transfers

31 Reforms in a number of countries have aimed to transform social transfers so that vulnerable groups are protected while encouraging greater labour force attachment. This includes, for example, reforming previously unconditional unemployment benefit systems and re-orientating child and family benefits towards employment-conditional measures such as child-care support. In other cases, some transfers, such as disability benefits, have been prone to misuse. Measures which address inflows into disability rolls can be effective in reducing spending while encouraging greater labour force participation. If such measures allowed high spending countries to move towards the current cross-country average spending ratio on family and disability benefits, countries could enjoy savings of over 0.5 per cent of GDP on average and up to almost 3 per cent of GDP in some countries (Table 3), while boosting long-term output.

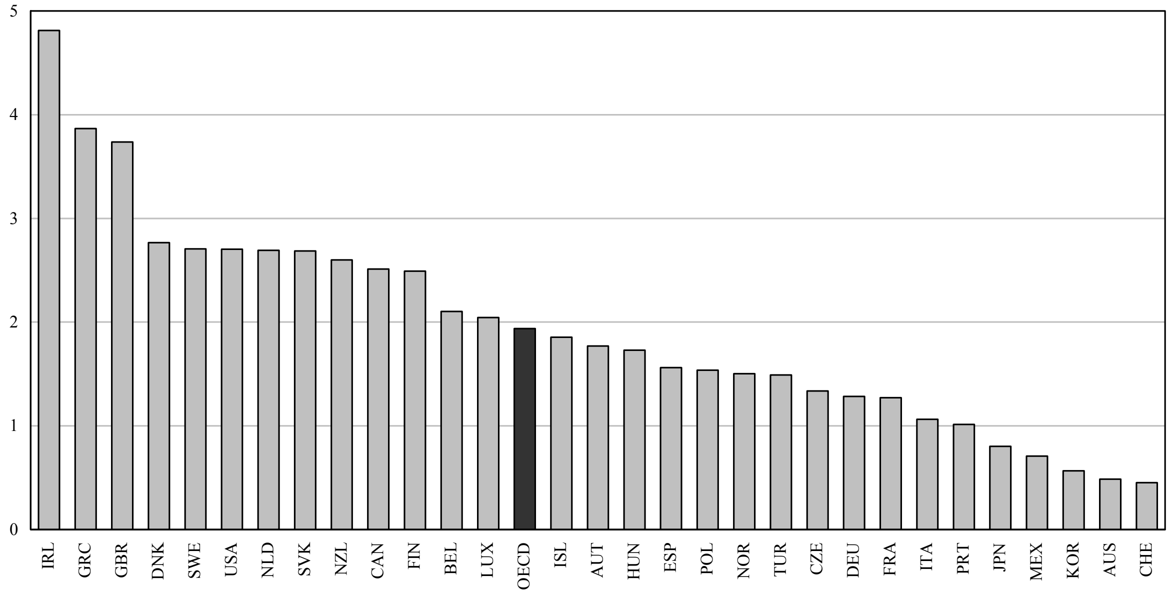
Greater efficiency

32 Work by the OECD has examined the opportunities to improve the efficiency in service delivery for health and education (similar savings are likely to be available in other spending programmes, Hagemann, 2011). These are important spending programmes accounting for about a quarter of government spending or on average across OECD countries around 10 per cent of GDP between them.

- No "one-size-fits-all" exists for health, in the sense that no "model" of health care delivery seems to be universally more cost efficient than other "models". However, within each "model" countries achieve widely divergent degrees of cost efficiency, suggesting that optimisation at the margin rather than a switch of model is the best way to achieve savings. Indeed, adopting best practice policies could see potential efficiency gains in the region of 2 per cent of GDP on average by 2017 (Joumard *et al.*, 2010), thereby allowing savings to be made without compromising service delivery (Figure 7, Table 3).
- For primary and secondary education, schools adopting best practice measures could realise important savings, up to around 1 per cent of GDP in some cases (Sutherland *et al.*, 2007). The estimates for school savings are based on benchmarking individual school performance against the best performing schools with similar student populations and resources (using data envelopment analysis). The implications of reducing inefficiency are then translated into aggregate resource savings by the implied possible reduction in staffing costs (Figure 8, Table 3).

Figure 7

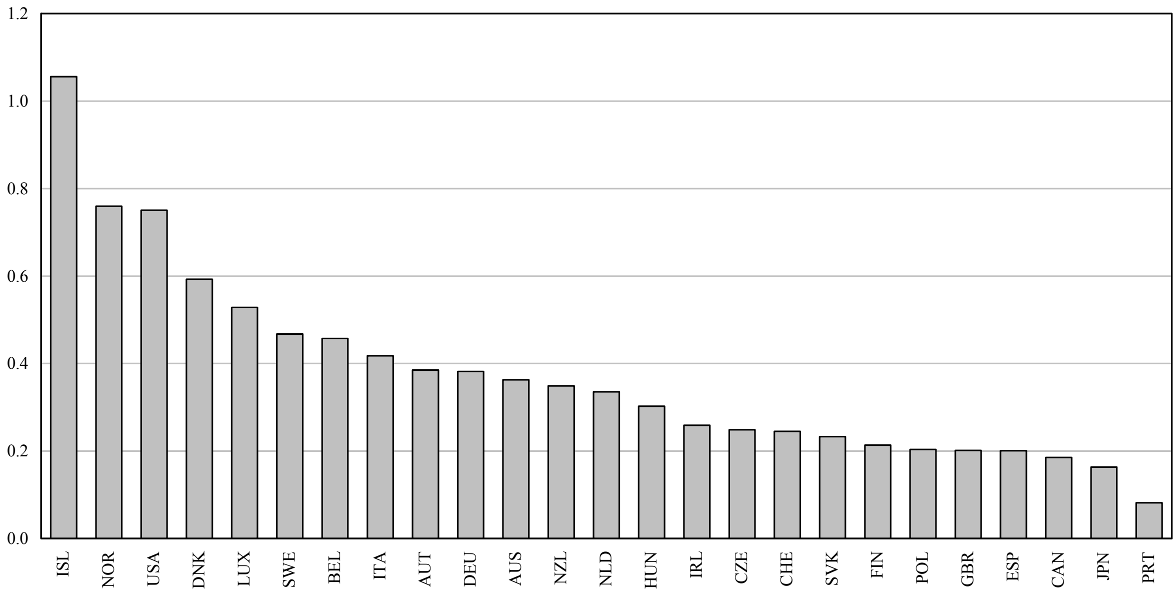
Potential Savings from Greater Efficiency in Public Health Care Spending
(percent of 2017 GDP)



Note: Potential savings represent the difference between a no-reform scenario and a scenario where countries would exploit efficiency gains. The no-reform scenario assumes that between 2007 and 2017 life expectancy and spending increase at the same pace as over the previous 10 years and that the mix between public and private spending remains constant over time.
Source: Joumard *et al.* (2010b).

Figure 8

Potential Savings from Greater Efficiency in Primary and Secondary Education Spending
(percent of GDP)



Source: Sutherland *et al.* (2007).

Government wages

33 Important gains can be achieved through management and pay reforms, and reducing the public sector wage bill is a candidate for fiscal consolidation in many countries. On average, the general government wage bill is close to 10 per cent of GDP and accounts for roughly one quarter of overall spending. Indeed, there are countries where a large public-private sector wage gap has developed over time. Restoring the wage relativities in the early 2000s could yield significant savings in a number of countries (Table 3). Ireland and Hungary have demonstrated recently that substantial cuts in public sector wages can be implemented if there is an urgent need for consolidation and a case arising from public-private pay relativities. That said, comparing public and private remuneration levels poses serious challenges, and requires valuation of working conditions and non-wage remuneration, such as defined benefit pension schemes. The ultimate test of adequacy is likely to be the difficulty or ease of recruitment into and retention in the civil service. From this perspective, budgetary savings achievable through reductions in the government wage bill should best be the outcome of a thorough review rather than across-the-board or arbitrary cuts in pay.

Subsidies

34 Subsidy reduction should rank high on the policy agenda as many subsidies may have surpassed their initial intended objective and may now have adverse economic effects. The elimination of subsidies (as defined in the national accounts), to the average for the OECD could yield sizeable savings in a number of countries (Table 3). Furthermore, by reducing the distortions they create, cutting subsidies offers the potential to boost growth.

Tuition fees

35 Close to a quarter of public spending on education is to support tertiary education, including tuition-free attendance in many countries, especially in continental Europe. A large share of returns to publicly-funded tertiary education accrue to individuals rather than to society (Blöndal *et al.*, 2002), and although some of the private returns are reduced by progressive taxes continued generous public support for higher education can be questioned. This is more so given the greater prevalence of tertiary education among middle and upper income households. The introduction or increase of tuition fees may also improve educational outcomes, by making schools more responsive to market demands, with long-term gains to human capital, the quality of labour supply, the economy's rate of potential growth, and overall fairness. Introducing or raising tuition fees to the average spending in countries that use tuition fees could yield additional revenues of around 0.4 per cent of GDP (Table 3). Concerns that such reforms would reduce enrolment by students from poor backgrounds could to a large extent be addressed by loan programmes with repayment conditional on subsequent income level.

Tax expenditures

36 All OECD governments use tax expenditures to promote a range of policy objectives. The scope of tax expenditures varies greatly across OECD countries, but they account for very substantial revenue leakages in some cases. Not all tax expenditures are undesirable, though, as some improve equity-efficiency trade-offs, like the case of earned income tax credits. Many, however, are distorting, poorly targeted, and contribute to a lack of transparency. In some cases, estimates of the revenues forgone by a tax expenditure can exceed a percentage point of GDP and the aggregate impact of all tax expenditures is likely to exceed several percentage points of GDP in

most OECD countries. Typically, the most costly tax expenditures are those aimed at boosting retirement savings, promoting homeownership, health insurance and charitable giving (OECD, 2010a).

37 Two examples reveal the potential importance for consolidation of reforming tax expenditures in personal income tax:

- Tax-favoured treatment of saving for retirement is found to boost retirement savings *per se*, but there is scant evidence that it raises aggregate private saving. Instead, such tax breaks result in a reallocation of saving from non-tax preferred to tax-preferred vehicles, while causing substantial revenue leakages, which may even reduce aggregate national saving. Phasing out such incentives could yield 1.7 per cent of GDP or more in additional revenues on average across a sample of OECD countries (Antolin *et al.*, 2004).
- Preferential tax treatment of owner-occupied housing is one of the costliest tax preferences in many OECD countries. The most important source of housing-related revenue leakages arises from the tax exemption granted to the implicit rental income of the owner-occupied home. Whereas the owner of a residence that is rented pays tax on the rental payments (less interest and operational costs), the implicit rental income of the owner-occupant is tax-exempt in the vast majority of member countries, except in the Netherlands, Sweden and Switzerland.[6] Despite the exclusion of the implicit rental income, some countries nevertheless allow the deductibility of mortgage interest, as well as property taxes (normally paid at the sub-national level). In addition, many countries provide favourable treatment to long-term capital gains from the sale of owner-occupied housing, adding further to the post-tax attractiveness of investment in housing. Thus, by removing a bias favourable to owner-occupied housing, reform could not only increase revenue but also improve the allocation of capital, boosting growth.

38 There are also important tax expenditures in indirect taxation. While VAT is widely recognised as an efficient and buoyant revenue source, its revenue potential is not fully used. Indeed, with the exception of New Zealand, a substantial portion of potential revenue is foregone in most countries due to a combination of reduced VAT rates, a narrow base, and low compliance (Figure 9). There is thus considerable scope for boosting revenue through VAT reforms (Table 3). Direct fiscal consolidation aside, broadening the base and reducing the number of rates offer scope to improve administration and compliance, by reducing complexity and countering political pressure for additional low rates. A more effective means to meet distributional objectives may be to target compensatory increased cash transfers or refundable tax credits to compensate low-income households.

39 Financial services are typically exempted from the VAT, largely due to technical difficulties in determining the precise tax base for margin-based services (*i.e.*, intermediation). Since much of VAT paid by financial service providers on inputs is non-recoverable, the sector's VAT exemption causes a number of economic distortions that result in more household consumption of financial services, and less use of and greater self-provision of financial services by businesses. However, the evolution of accounting methods and information systems has reduced the technical obstacles to imposing VAT on financial services considerably (OECD, 2010b). Moreover, following the recent financial crisis, there is increased interest among governments in both raising revenue from financial institutions and reducing moral hazard in the financial services sector via new taxes on financial services or (elements of) balance sheets.

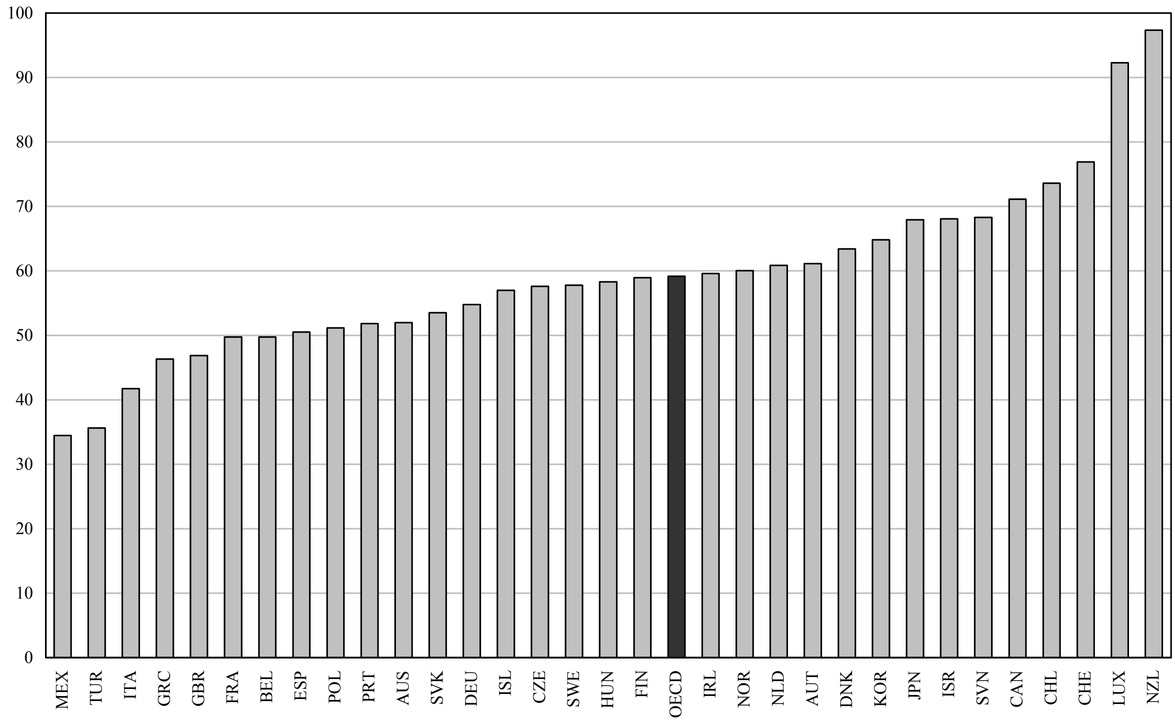
---

6 In the Netherlands and Switzerland, however, taxable imputed rentals are very low, which combined with mortgage interest deductibility acts to reduce personal income tax revenues significantly.

Figure 9

Value Added Tax Performance: The VAT Revenue Ratio

(average 2007-08, percent)



Note: The VAT revenue ratio measures the difference between the VAT revenue actually collected and what would theoretically be raised if VAT was applied at the standard rate to the entire potential tax base in a "pure" VAT regime and all revenue was collected: The VAT revenue ratio equals VAT Revenue/(Consumption * Standard VAT rate)*100.
Source: OECD (2011), Consumption Tax Trends 2010: VAT/GST and Excise Rates, Trends and Administration Issues.

Less distortionary tax bases

40 When tax rates need to be raised, some taxes are natural candidates for fiscal consolidation programmes both from an efficiency and revenue-raising perspective. The efficiency costs of taxes on immobile property are lower than on consumption or income, but represent a small share of overall tax revenue in many OECD countries.[7] Where they are low or non-existent, corrective taxes such as so-called "sin" taxes that can help deter harmful behaviours (*e.g.* alcohol and tobacco consumption), or taxes on polluting activities or consumption (*e.g.* fossil fuels) can improve welfare while boosting revenues.

41 Environmental taxes hold the promise of both boosting revenue and helping to achieve environmental objectives by discouraging pollution. While some countries raise considerable revenues from such taxes, reaching 4 per cent of GDP in Denmark and the Netherlands in 2008, their yield is relatively low in several countries, notably Canada, New Zealand and the United States. Nonetheless, imposing a tax on carbon emissions or auctioning tradable emission rights to contain greenhouse gas emissions has become more widespread. For example, the European Union has auctioned permits as part of the Emission Trading Scheme. Despite such

---

7 In most countries, property taxes are a main source of finance for sub-national governments, posing potentially challenging fiscal federalism problems should national property taxes be introduced or raised.

developments, many countries maintain differences in taxation depending on fuel type that run counter to estimates of environmental externalities. From a fiscal consolidation perspective, greenhouse gas levies consistent with international action to stabilise atmospheric concentrations of greenhouse gases by 2020, could generate around 2 per cent of GDP (de Serres *et al.*, 2010) (Table 3).

Summing up potential for primary balance adjustment

42 The potential contributions of spending and revenue measures to fiscal consolidation reported in Table 3 could inform a choice of where potential may exist to make savings or increase revenues. Even without being able to quantify all the possible measures across countries, and not taking into account any dynamic effects, the cumulative potential cuts in spending (benchmarked using the OECD average or estimates of potential efficiency gains) and increases in taxation (benchmarked using the OECD average) are sizeable. On average across countries, budget enhancements could reach around 7 per cent of GDP, with the larger part available on the spending side. Given that there are measures that are difficult to quantify this is a lower estimate. Furthermore, the potential tends to be somewhat greater in the English-speaking countries which generally face the larger consolidation needs. A large share of the savings in spending would come from reaping efficiency gains, which are likely to take some time to emerge. On the revenue side, relatively large opportunities exist for the greater use of environmental taxes and the broadening of income and indirect tax bases.

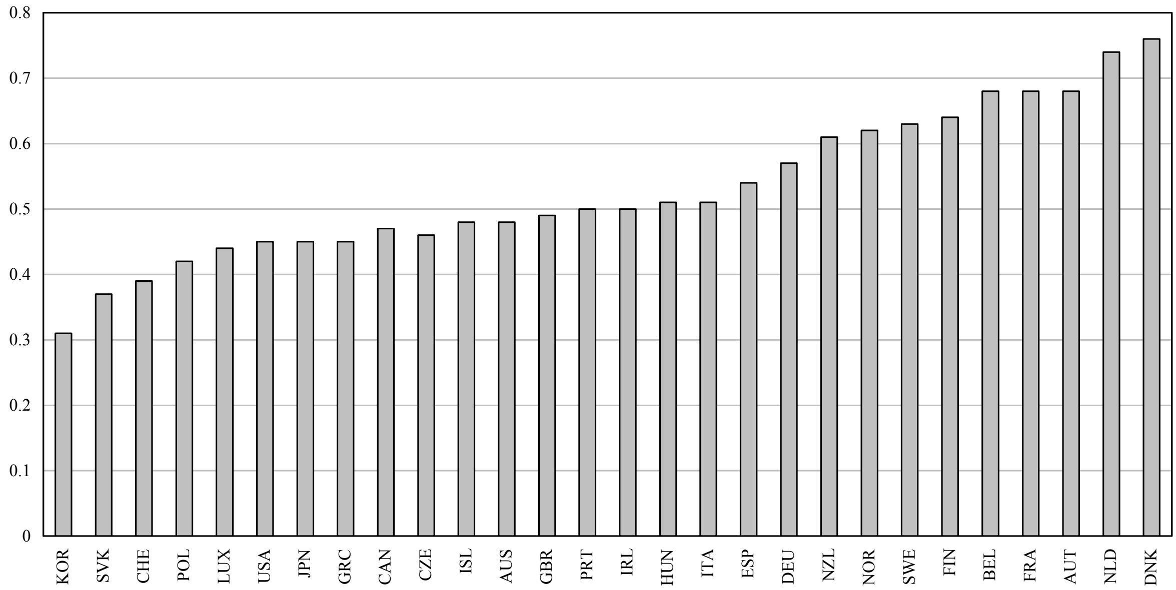
Supporting reforms

43 In a number of cases supporting reforms could assist fiscal consolidation. Aside from their direct budgetary impact, as discussed above, reforms to pension systems that delay retirement and increase labour force participation will boost revenues and thereby reduce long-run budget pressures. Reforms that link retirement age to gains in longevity would thus help cushioning budgets against future changes in longevity. More generally, growth-enhancing structural policy reform may support fiscal consolidation. This is most obvious when reforms, such as retirement reforms, lead to a higher sustainable employment level because such a change will have a permanent impact on the primary balance (Figure 10). The size of the effect will depend on the taxes levied on the additional income and consumption created as well as on whether the reform in question has any direct budgetary impact. The latter will be the case, for example, when additional spending on active labour market policy boosts aggregate spending or cutbacks on unemployment benefit duration reduces it. But many structural reforms have little direct impact on budgets while at the same time boosting employment levels, such as in the case of product market reforms that boost competition.

44 The effects of productivity-enhancing structural reforms on public budgets are less clear. Higher productivity in the private sector will tend to boost revenues but also spending unless public/private wage relativities change or transfer income replacement ratios are altered. Hence, the effect on the primary budget balance may be muted. However, to the extent higher productivity growth is not matched by a corresponding increase in real interest rates debt dynamics will be favourably affected. Such an effect is particularly likely for individual countries participating in a monetary union since the general structure of interest rates is unlikely to be strongly affected by structural reform in an individual country while at the same time higher growth may lead to a narrowing of risk premia.

Figure 10

Effect of 1 Per Cent Higher Potential Employment on the Primary Balance
(percent of GDP)



Source: OECD Economic Outlook 88 database; and OECD calculations.

Conclusions

45 Overall, the link between economic growth and the post-crisis debt overhang is complicated. On the one hand, high debt seems to be associated with lower growth. But, on the other hand, fiscal consolidation may weaken growth both in the near term and over a longer horizon. Realistically, debt problems are so serious in many countries that consolidation has the potential to hamper growth strongly.

46 In the short run, consolidation may weaken demand and monetary policy may not be able to compensate for such effects for some time to come. This argues for phasing in consolidation. Appropriate and clear fiscal objectives together with institutions that ensure accountability may help to preserve credibility in the process. However, to maintain credibility it may also be necessary to take some action up-front, in which case instruments with small short-term multipliers may be given some weight. This may involve some political economy risk, to the extent it skews consolidation towards inappropriate instruments. Slow consolidation may also entail a price insofar as it involves higher debt and thereby higher interest rates.

47 In the longer run, effects of consolidation on growth will depend on the choice of instruments. Some instruments are available that will have limited detrimental impacts on growth and little or no conflict with other policy objectives. Notably, increasing spending efficiency, reforming unsustainable pension systems, putting prices on environmental externalities and maximising the benefits of structural reforms could make sizeable contributions to consolidation. In addition, reviewing tax and benefit systems more generally could help identify how policy objectives could be achieved at lower cost and where support is less justified.

APPENDIX
FISCAL GAPS

48 The underlying model used to calculate fiscal gaps is deliberately simple (Merola and Sutherland, 2011). It builds on the assumptions underlying the *Economic Outlook* medium-term baseline on potential output growth, output gaps, interest and inflation rates until 2025. Between 2025 and 2050, GDP growth is determined by the growth rate of potential, which is driven by demographic developments and assumptions about productivity growth. The fiscal side of the model assumes that revenues adjusted for the cycle remain a constant share of GDP and, in the baseline, primary spending is also a constant share of GDP.

49 For any long-run fiscal projections, GDP growth, interest rates and inflation together with the fiscal assumptions determine long-run sustainability (Table 4). In the country models the main assumptions are as follows:

- GDP growth in the long term is driven by potential output. One of the main components of potential output that is varying over time is working age population growth, which is based on cohort data from long-term demographic projections. GDP growth is then determined by participation rates and employment and labour productivity growth. The latter is assumed to converge to 1.75 per cent by 2035 at the latest. The simulations ignore possible impacts of fiscal policy and debt developments on output.
- Interest rates on government borrowing are partly determined by monetary policy. The return of output to potential is accompanied by a normalisation of interest rates, such that the risk-free rate is at its estimated natural rate by 2025. Inflation converges to the monetary authorities' target, typically 2 per cent annually. Interest payments are determined by the stock of debt and an interest rate that is based on a mix of long and short-term rates, with the long-term rate including a premium of 4 basis points for each percentage point of financial liabilities in excess of 75 per cent of GDP. Japan is assumed to remain unusual, with the very high share of domestic financing keeping the risk premium at only 1 basis point for each percentage point of financial liabilities in excess of 75 per cent of GDP.
- The other major assumptions concern fiscal policy. In the baseline, underlying revenues and primary spending are constant as shares of GDP, though the automatic stabilisers operate while the economy moves back to potential. In some scenarios, ageing-related spending is added to underlying spending to highlight the fiscal pressures coming from population ageing. For health care, given that only a relatively small portion of the projected increase is ageing-related, additional spending is phased in linearly over the projection horizon.

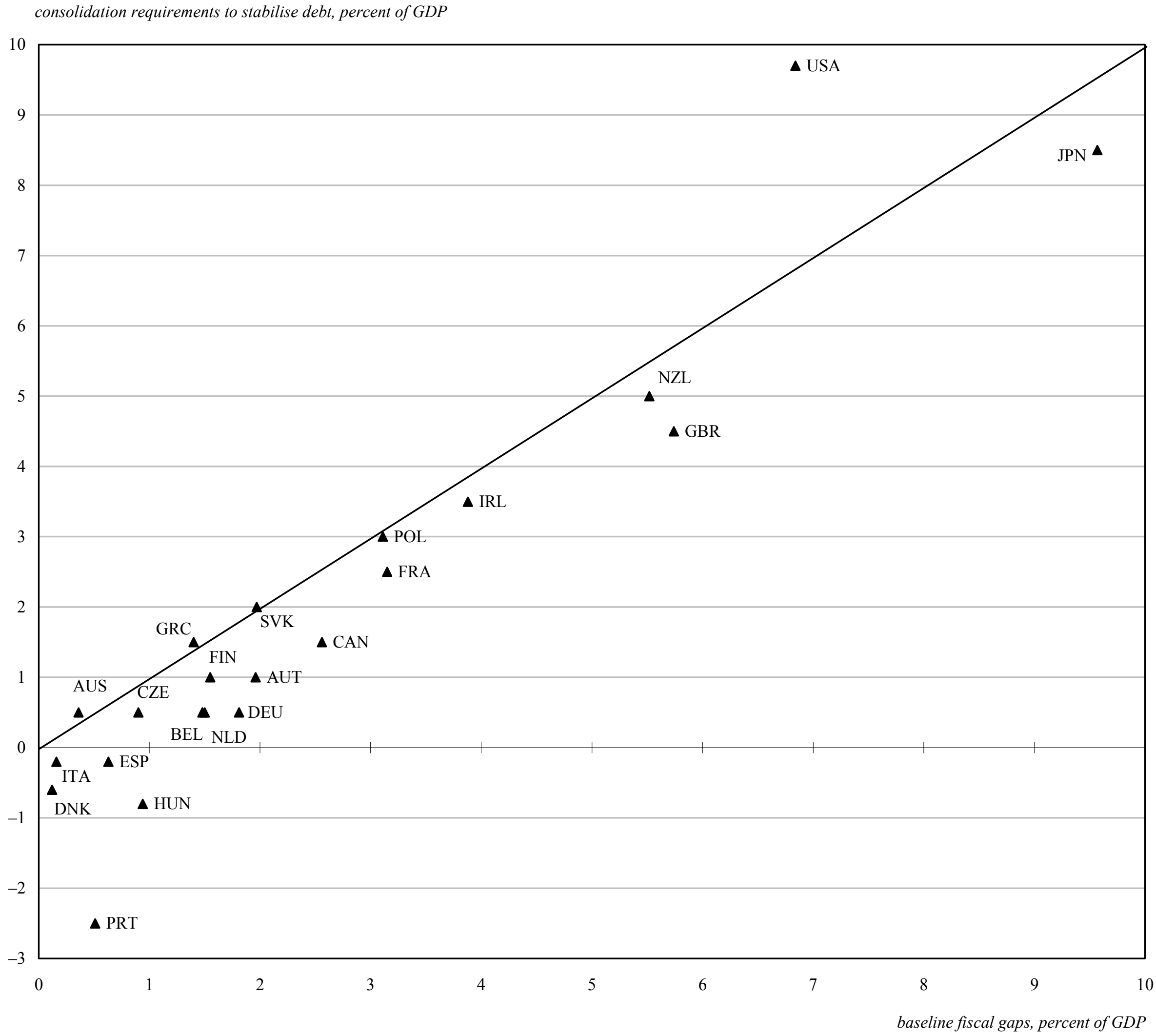
50 The fiscal gaps are distinct from recent work by the OECD that has assessed the consolidation requirements to stabilise debt (OECD, 2011c). These requirements are based on stylised assumptions about a sustained gradual annual tightening of the underlying primary balance by 0.5 per cent of GDP until debt stabilization is reached. The fiscal gaps on the other hand make the alternative stylised assumption that the tightening will be implemented immediately and sustained until 2050 to meet a specific debt target. Both sets of assumptions ignore the implications for output, which will obviously be important.

51 Overall the two approaches produce similar rankings of consolidation needs across counties (Figure 11). The two approaches differ in three ways. First the time path of consolidation is different. Second, the final debt level is different. Third, the time horizon is different. The first and third differences in particular pull in opposite directions for the two approaches. The combined effect of the differences leads to the additional tightening to bring debt down to 50 per cent of GDP in 2050 being typically not much greater than the gradual fiscal tightening needed after 2012 to stabilise debt levels. In general, the immediate consolidation assumed by the fiscal gap calculations is sufficient to bring debt dynamics under control more quickly which combined with the

assumption that the fiscal tightening is permanent over a longer time horizon will see debt levels gradually fall for the rest of the simulation. The estimates of the amount of consolidation needed to stabilise debt are particularly large for the United States and Japan and the gradual tightening takes considerably longer to stabilise debt. As a higher interest premium for each percentage point of debt above 75 per cent of GDP is assumed for the United States than Japan, the consequences of the gradual tightening for adverse debt dynamics are more severe, which explains why the relationship with the fiscal gap estimates differs from the other countries. If countries do not need to consolidate to meet the terminal debt target, such as in the case of Sweden, no fiscal gap is calculated and the country is excluded from the figure.

Figure 11

Relation Between Fiscal Gaps and Consolidation Requirements



Source: OECD (2011c), OECD Economic Outlook 89.

Table 4

Key Assumptions in the Baseline Simulation

Country	Starting Point, 2012		Average Over Simulation	
	Gross Debt *(percent of GDP)*	**Underlying Primary Balance** *(percent of GDP)*	**Effective Interest Rate**	**Nominal GDP Growth**
Australia	31	0.6	6.9	4.8
Austria	82	0.1	4.4	3.5
Belgium	100	0.9	4.7	3.8
Canada	88	–1.8	4.9	4.2
Czech Republic	51	0.3	4.4	4.2
Denmark	60	0.8	5.0	3.5
Finland	66	0.8	4.2	3.9
France	100	–0.6	4.1	3.6
Germany	87	0.6	4.3	3.0
Greece	159	3.5	5.5	3.4
Hungary	81	1.1	5.8	3.2
Ireland	126	–0.4	4.7	4.3
Italy	128	3.3	4.6	3.1
Japan	219	–4.2	3.0	2.2
Korea	33	0.5	5.6	2.4
Luxembourg	24	2.0	4.5	4.9
Netherlands	75	0.0	4.3	3.5
New Zealand	52	–4.0	5.8	4.3
Poland	66	–1.5	5.3	3.2
Portugal	116	3.5	4.6	3.1
Slovak Republic	51	–1.7	5.1	2.8
Spain	75	0.5	4.2	3.5
Sweden	41	2.6	4.7	4.0
Switzerland	37	1.2	2.9	2.9
United Kingdom	93	–3.0	4.6	4.1
United States	107	–5.8	4.6	4.3

REFERENCES

Aizenman, J. and N. Marion (2009), "Using Inflation to Erode the U.S. Public Debt", NBER, Working Paper, No. 15562.

Antolin, P., A. de Serres and C. de la Maisonneuve (2004), "Long-term Budgetary Implications of Tax-favoured Retirement Plans", OECD, Economics Department, Working Paper, No. 393.

Barrell, R., D. Holland and I. Hurst (2012), "Fiscal Consolidation: Part 4. Fiscal Multipliers and Fiscal Consolidation", OECD, Economics Department, Working Paper, forthcoming.

Blöndal, S., S. Field and N. Girouard (2002), "Investment in Human Capital through Post-compulsory Education and Training: Selected Efficiency Aspects", OECD, Economics Department, Working Paper, No. 333.

Caner, M., T. Grennes and F. Köhler-Geib (2010), "Finding the Tipping Point – When Sovereign Debt Turns Bad", The World Bank, Policy Research, Working Paper, No. 5391.

CBO (2011), *2011 Long-term Budget Outlook*, Congressional Budget Office, Washington (D.C.).

Cecchetti, S., M. Mohanty and F. Zampolli (2011), "The Real Effects of Debt", Bank for International Settlements, Working Paper, No. 352.

Corsetti, G., K. Küster, A. Meier and G. Müller (2011), "Sovereign Risks and the Effects of Fiscal Retrenchments in Deep Recessions", Federal Reserve Bank of Philadelphia, Research Department, Working Paper, No. 11-43.

Cournède, B. (2010), "Gauging the Impact of Higher Capital and Oil Costs on Potential Output", OECD, Economics Department, Working Paper, No. 789.

de Serres, A., F. Murtin and G. Nicoletti (2010), "A Framework for Assessing Green Growth Policies", OECD, Economics Department, Working Paper, No. 774.

Edey, M. (1994), "Costs and Benefits of Moving from Low Inflation to Price Stability", OECD, Economic Studies, No. 23.

Égert, B. (2010), "Fiscal Policy Reaction to the Cycle in the OECD: Pro- or Counter-cyclical?", OECD, Economics Department, Working Paper, No. 763.

Elmeskov, J. and D. Sutherland (2012), "Post-crisis Debt Overhang: Growth Implications Across Countries", paper prepared for the Reserve Bank of India Second International Research Conference 2012 *Monetary Policy, Sovereign Debt and Financial Stability: The New Trilemma*.

Guellec, D. and B. van Pottelsberghe de la Potterie (2004), "From R&D to Productivity Growth: Do the Institutional Settings and the Source of Funds of R&D Matter?", *Oxford Bulletin of Economics and Statistics*, Vol. 66, No. 3.

Hagemann, R. (2012), "Fiscal Consolidation: Part 6. Policy Instruments for Fiscal Consolidation", OECD, Economics Department, Working Paper, forthcoming.

Hall, G. and T. Sargent (2010), "Interest Rate Risk and Other Determinants of Post-WWII U.S. Government Debt/GDP Dynamics", NBER, Working Paper, No. 15702.

Hall, R. and C. Jones (2007) "The Value of Life and the Rise in Health Spending", *Quarterly Journal of Economics*, Vol. 122, No.1, pp. 39-72.

Haugh, D., P. Ollivaud and D. Turner (2009), "Macroeconomic Consequences of Banking Crises in OECD Countries", OECD, Economics Department, Working Paper, No. 683.

HM Treasury (2011), *Whole of Government Accounts: Unaudited Summary Report for the Year Ended 31 March 2010*, HM Treasury, London.

Jonung, L. (2010), "Lessons from the Nordic Financial Crisis", paper prepared for the *AEA Meeting* in Denver.

Joumard, I., C. André and C. Nicq (2010a), "Health Care Systems: Efficiency and Institutions", OECD, Economics Department, Working Paper, No. 769.

Joumard, I., P. Hoeller, C. André and C. Nicq (2010b), *Health Care Systems: Efficiency and Policy Settings*, OECD Publishing.

Kumar, M.S. and J. Woo (2010), "Public Debt and Growth", International Monetary Fund, Working Paper, No. WP/10/174.

Merola, R. and D. Sutherland (2012), "Fiscal Consolidation: Part 3. Long-run Projections and Fiscal Gap Calculations", OECD, Economics Department, Working Paper, No. 934.

OECD (2003), *Transforming Disability into Ability: Policies to Promote Work and Income Security for Disabled People*, OECD Publishing.

——— (2009), *OECD Economic Outlook: Interim Report*, OECD Publishing.

——— (2010a), *Choosing a Broad Base – Low Rate Approach to Taxation*, Tax Policy Study No. 19, OECD Publishing.

——— (2010b), *Going for Growth*, OECD Publishing.

——— (2011), *Consumption Tax Trends 2010: VAT/GST and Excise Rates, Trends and Administration Issues*, OECD, Paris.

——— (2011a), *Pensions at a Glance: Retirement Systems in OECD and G20 Countries*, OECD Publishing.

——— (2011c), *OECD Economic Outlook* 89, OECD Publishing.

Oliveira-Martins, J. and C. de la Maisonneuve (2006), "The Drivers of Public Expenditure on Health and Long-term Care: An Integrated Approach", OECD, Economic Studies, No. 43, February.

Price, R. and T. Dang (2011), "Adjusting Fiscal Balances for Asset Price Cycles", OECD, Economics Department, Working Paper, No. 868.

Röhn, O. (2010), "New Evidence on the Private Saving Offset and Ricardian Equivalence", OECD, Economics Department, Working Paper, No. 762.

Reinhart, C. and K. Rogoff (2010), "Growth in a Time of Debt", *American Economic Review*, Vol. 100, No. 2.

Reinhart, C. and M. Belen Sbrancia (2011), "The Liquidation of Government Debt", NBER, Working Paper, No. 16893.

Sutherland, D., R. Price, I. Joumard and C. Nicq (2007), "Performance and Indicators for Public Spending Efficiency in Primary and Secondary Education", OECD, Economics Department, Working Paper, No. 546.

HUMAN CAPITAL FORMATION IN ARGENTINA: CONTRIBUTION TO GROSS DOMESTIC PRODUCT

Ernesto Rezk,[*] Maria De los Ángeles Mignon[*] and Agustin Ramello De la Vega[*]

1 Introduction

The influence of human capital formation upon countries' gross domestic product and its long run growth path was always a matter of interest both for theorists on growth theory as well as for policy makers involved in the design of fiscal growth and development policies. The idea of the inclusion of human capital in production functions had already been considered by Uzawa (1965) and Lucas (1988) in their two sector endogenous growth models; in one sector, the final production stemmed from the combination of physical and human capital whereas in the other production and human capital accumulation were derived from human capital use alone.

Lucas theoretical contribution (1988, 1990) also dealt with externality features, by suggesting that investment in human capital not only enhanced individuals' earning abilities but might also generate an external effect that raised the aggregate level of productivity and served in turn to explaining countries' long run income diversity. Contemporaneously, Romer (1990) also highlighted the importance of human capital by putting forward his well-known $I + D$ and growth model in which the underlying research technology only depended on labour or human capital.

In assessing the hypothesis of human capital as "engine of growth", Frenkel and Razin (1996) carried out their analysis based on a classical textbook endogenous growth model including both physical and human capital and showed that the long-run growth rate was always positively related to the human capital saving rate but positively, negatively related or totally unrelated at all to the physical capital saving rate, this depending on the value taken by the reciprocal of the intertemporal elasticity of substitution in consumption; the above verification led Frenkel and Razin to defend public policies targeted at raising the human capital saving rate on grounds that they would directly impact on the economy's long-run growth rate.

In attempting to ascertain the role played by human capital, the influential paper by Mankiw, Romer and Weil (1992), focused on the empirics of economic exogenous growth and brought about a revaluation of the traditional Solow-Swan Model (SSM) by showing that the latter's predictions were somehow consistent with their own econometric evidences. Even though the SSM rightly predicted the directions of the effects of saving and population growth upon income, they found that estimates of parameters fell short of being satisfactory as they clearly overstated the size of the coefficient on physical capital compared to the actual capital share of one third usually assumed in the formulation of the Cobb Douglas production functions.

This empirical lack of consistency was dealt with by Mankiw *et al.* by building what they called an "Augmented Solow Model" which explicitly included human capital in the production function; the resulting log equation, holding now that real per capita income depended on population growth as well as on physical and human capital accumulation had, according to the econometric results, a much better performance as the human capital variable turned out to be significant, the size of the physical capital coefficient fell in line with it expected actual value and the fit of the equation improved compared to the regression in which human capital was omitted whereas the restriction that all three coefficients (on population growth and on propensities to accumulate physical and human capital) summed to zero was not rejected.

Following the line drawn by the above mentioned contributions, this research paper aims at assessing the impact of the Investment in Education in Argentina (as one of components of Human Capital Formation) upon the Gross Domestic Product, therefore the Augmented Solow Model is used as underlying the theoretical framework. It is worth point out here that the empirical developments due to Mankiw *et al.*, given difficulties found in computing the variable, resorted to a proxy for the propensity to invest in human capital accumulation consisting in taking the percentage of the working age population enrolled in secondary school; in connection to this, one main contribution of this paper resides in furthering the empirical treatment of the "augmented SSM" on the following three accounts: a) the possibility is investigated of finding better representations for the average propensity to invest in human capital other than the one above mentioned, b) missing components, such as the opportunity costs incurred by parents and students are added to all government and educational levels' budgetary expenditures and c) a methodology is developed for the measurement of the stock of human capital in order that the variable be available to be used, in a second stage to this project, in place of the rate of human capital accumulation.

Furthermore, and given the widespread admission that valuable empirical and policy implications may arise from including human capital, the Augmented SSM econometric performance is assessed by resorting to cointegration and error correction models and innovation accounting involving impulse response function and variance decomposition analysis.

A worth stressing point is that the advance on methodological aspects relating data treatment and measurement, as well as the results from the carried out econometric estimation of equations, are expected to serve as inputs for the second stage in which the inclusion of human capital will be assessed in the frame of endogenous growth models.

In line with objectives held above, the rest of the paper is organized as follows: Section 2 summarizes the theoretical treatment given by Uzawa (1965) and Lucas (1988) to the inclusion of human capital in endogenous growth models as well as a review of the Mankiw, Romer and Weil's Augmented Growth Model (1992); in Section 3 a methodological alternative is introduced and applied to the Argentine economic scenario, for computing both the average propensity to invest in human capital and its stock; Section 4 presents a synthetic review of stylized facts that highlights the joint performance –in the period considered- of gross domestic product and human capital; Section 5 presents the econometric estimation for Argentina of the Augmented Growth Model's parameters by using an Error Correction Model as well as the evaluation of results with tools of innovation accounting; section 6 concludes.

2 Human capital inclusion in economic growth models[1]

Theoretical contributions aimed at stressing the role of human capital in models of economic growth, and at empirically assessing its real impact upon long-run growth path, are ample and can be traced back close in time to the moment when the classical Solow-Swan Growth Model came into being.[2] Three of these contributions were selected to be reviewed: in the first two, Uzawa and Lucas, resorted to an endogenous growth model in which they included human capital whereas in the third one Mankiw, Romer and Weil (1992) extended the Solow-Swan Model by adding what they deemed to be the omitted variable; that is, human capital accumulation.

[1] This section builds on papers by Uzawa (1965), Lucas (1988) and Mankiw *et al.* (1992) and on Heijdra and van der Ploeg (2002), ch. 14, and Sala-i-Martín (1994), ch. 8.

[2] Suffice it in this connection to mention Schultz's communication (1961) on the impact of labour quality improvement upon the pattern of economic growth.

2.1 The Uzawa-Lucas Model

AK endogenous growth models including physical (K) and human capital (H) were founded on the assumption that both were similar goods, obtained with the same technology and able to be produced and accumulated out of not consumed units of production; as a consequence of this, the following two relationships between stocks of both capital variants were seen to hold implying that a temporal reduction in K (and in K/H ratio) would be made up by getting a part of H immediately converted in K:

$$\frac{K}{H} = \frac{\alpha}{1-\alpha} \tag{1}$$

$$H = K\,\frac{1-\alpha}{\alpha} \tag{2}$$

where $0<\alpha<1$ stood for the physical capital's share in the production function.

Simple and practical as it might appear, this unrealistic assumption was challenged by Uzawa by suggesting that technological knowledge could only be raised by devoting resources to this end, following a pattern of allocation conducive to optimum growth within the framework of a two sector aggregative growth model whose main features were intuitively simple. Uzawa started by drawing the productive sector represented by the production function (3) below, in which physical capital and labour used for final goods production combined and yielded a homogenous output which could be either instantaneously consumed or devoted to enhancing the stock of physical capital:

$$Y(t) = f\big(K(t), A(t)L_P(t)\big) \tag{3}$$

and where $A(t)$ stood for the state of technological knowledge at any time t[3] and L_P labour used in the production of final goods.

The second sector, broadly defined as "the educational sector", employed only labour and its impact diffused over the economy via the enhancement of labour efficiency $(\dot{A}L(t))$; Uzawa made the rate of change of labour efficiency to depend on non increasing marginal returns[4] and the ratio between labour employed by the educational sector and total labour force:

$$\frac{\dot{A}_L(t)}{A_L(t)} = \phi\left[\frac{L_E(t)}{L(t)}\right] \tag{4}$$

In interpreting expression in (4) it should be noticed that, for Uzawa, the larger the change in labour efficiency, the larger the amount of labour devoted to the educational sector (L_E) which, in the context of an inelastically supplied labour force growing at a rate n, amounted to meeting the restriction imposed by the identity (5):

$$L_E(t) + L_P(t) = L(t) \tag{5}$$

The rest of the model formulation was completed by traditionally stating the rate of physical capital accumulation as the difference between the positive annual rates of aggregate investment and of capital stock depreciation:[5]

---

3 For Uzawa, changes in technological knowledge were exclusively embodied in labour and therefore labour efficiency's increases did not depend on the amount of employed physical capital.

4 Non increasing marginal returns to labour meant that $\phi'(s) \geq 0$ and $\phi''(s) \leq 0$ for all $0 \leq s \leq 1$.

5 It must be noted that while Uzawa used this equation to define the rate of capital accumulation, both Lucas and Mankiw *et al.* used a similar formulation to express the net investment in physical capital or, in other words, the capital accumulation (see equation 33 below).

$$\frac{\dot{K}(t)}{K(t)} = I(t) - \delta_K K(t) \tag{6}$$

and by introducing a linear utility function whereby the optimum time path was characterized in terms of the discounted sum of per capita consumption:

$$U(0) = \int \frac{C(t)}{L(t)} e^{-\rho t}\, dt \tag{7}$$

In relation to the treatment of human capital in endogenous growth models, the main thrust represented by Uzawa's contribution was however extended by Lucas, at least on the following three accounts: a) while Uzawa broadly regarded $A_L(t)$ as embodying educational activities, health and provision and building of public goods, Lucas modified the idea by interpreting $A_L(t)$ as human capital; b) based on empirical evidence[6] showing that individual earnings were consistent with a linear knowledge production function, Lucas rejected the assumption of diminishing returns to knowledge accumulation implied by expression (4) and put forward in change a modified expression (4') for the human capital accumulation function in which $\phi_E > 0$ was now a parameter:

$$\frac{\dot{H}(t)}{H(t)} = \varphi_E \left[\frac{L_E(t)}{L(t)}\right] \tag{4'}$$

Expression (4') rested on Rosen's theory, applied to each finite-lived individual and extended by Lucas to the same technology applied to an entire infinitely-lived representative household; that is, individuals' acquired human capital were somehow transferred to next generations.[7]

The third change consisted in Lucas' introduction of a curved intertemporal utility function for the representative infinitely lived household, in place of the linear function (7), as expressed now in (7'):

$$U(0) = \int \frac{C(t)^{1-\theta} - 1}{1 - \theta} e^{-(\rho - n)t}\, dt \tag{7}$$

in which θ stood for the reciprocal of consumption's intertemporal elasticity of substitution. As known, $\theta = 1/\sigma$ is a constant that measures the degree of concavity of the utility function (7') its value in turn implying that the larger θ the greater the interest in smoothing consumption over time.

With the modifications introduced by Lucas (shown by equations (4') and (7')) the model development, and its resolution, followed endogenous growth models' standard procedures by incorporating the ensuing per capita equations[8] for physical and human capital accumulation in which the simplifying assumption of similar depreciation rates was used:[9]

$$\dot{k} = A(t)k(t)^{\alpha} \left(u\, h(t)\right)^{1-\alpha} - c(t) - (\delta + n)\, k(t) \tag{8}$$

$$\dot{h} = \varphi_E (1 - u)\, h(t) - (\delta + n)\, h(t) \tag{9}$$

---

6 Rosen (1976).

7 The assumptions that individuals' capital formation followed the pattern depicted by 4' and that the initial level each family member began with was proportional to the level already accumulated by the family's older members led Lucas (1988, p. 19) to assert that human capital accumulation was a social activity with no counterpart in physical capital accumulation

8 Equations (8) and (9) were derived from accumulation equations $\dot{K}$ and $\dot{H}$ divided by L, making next $k = K/L$ and $h = H/L$, taking derivatives with respect to time in order to obtain $\dot{k} = \dot{K}/L - nk$ and $\dot{h} = \dot{H}/L - nh$ and replacing $\dot{K}/L$ and $\dot{H}/L$ for their equivalents in per capita accumulation equations.

9 Similar to the effect caused by δ, increases in the population's rate of growth *(n)* dwindle the available per capita physical and human capital stock.

Thus, while u stood for the proportion of total human capital used for the production of final goods, $(1–u)$ indicated in turn the effort devoted to human capital accumulation.[10] Let it be noticed that, if L were normalized to unity in (5), L_E and L_P would respectively equal to $(1–u)$ and u^2.[11]

In line with the usual procedure, Uzawa-Lucas made individuals to choose temporal trajectories for consumption and stocks of physical and human capital that maximized the utility function already introduced; that is, equation (7') was maximized subject to non leisure time individuals devoted to each of the two sectors (time constraint 10) and the accumulation restrictions 8 and 9, as represented by the Hamiltonian in (11), including now two state variables (k and h) and two control variables (c and u):

$$h(t) = u\,h(t) + (1 - u)\,h(t) \tag{10}$$

$$\begin{aligned}\mathcal{H}(t) = e^{-(\rho-n)t}\frac{C(t)^{1-\theta}-1}{1-\theta} + \eta_K[A(t)k(t)^{\alpha}\big(u\,h(t)\big)^{1-\alpha} - c(t)\\ -(\delta + n)\,k(t)] + \eta_H\,[\varphi_E(1-u)h(t) - (\delta+n)h(t)]\end{aligned} \tag{11}$$

where the co-state variables $\eta_K(t)$ and $\eta_H(t)$ respectively stood for shadow prices of per capita investment in physical and human capital $k(t)$ and $h(t)$. The corresponding first order conditions, resulting from the derivation of the Hamiltonian with respect to control and state variables, and the transversality conditions, were:[12]

$$e^{-(\rho-n)t}C(t)^{-\theta} = \eta_K(t) \tag{12}$$

$$\eta_K(t)\,A(t)\,k(t)^{\alpha}\,(1-\alpha)u^{-\alpha}\,h(t)^{1-\alpha} = \eta_H(t)\varphi_E\,h(t) \tag{13}$$

$$-\dot{\eta}_K(t) = \eta_K(t)\left(A(t)\alpha\,k(t)^{\alpha-1}\big(u\,h(t)\big)^{1-\alpha} - (\delta + n)\right) \tag{14}$$

$$-\dot{\eta}_H(t) = \eta_K(t)(A(t)k(t)^{\alpha}\,u^{1-\alpha}\,(1-\alpha)h(t)^{-\alpha}) + \eta_H(t)\big(\varphi_E(1-u) - (\delta+n)\big) \tag{15}$$

$$\lim_{t\to\infty}\eta_K(t)k(t) = \lim_{t\to\infty}\eta_H(t)h(t) = 0 \tag{16}$$

What first order conditions were stating was that produced output must on the margin be equally valuable in its uses, either as consumption or investment goods (12), while at the same time individuals' non leisure time must also be equally valuable in its uses, namely, physical and human capital accumulation (13). Finally, first order conditions (14) and (15) reflected the fundamental principle of valuation of the perfect competition institutional setting whereby the rate of return on different assets (in this case physical and human capital) must also be equalized. In Lucas' words, "…equations (4') and (12)-(16) implicitly describe the optimal evolution of $k(t)$ and $h(t)$ from an initial mix of these two kinds of capital".[13]

By taking logarithms and derivatives with respect to time in (12), and replacing $\eta_K(t)$ by its expression in (14), the resulting consumption dynamic equation was obtained that placed the

---

10 Although physical capital may not straightforwardly be ruled out as an input for the production of human capital, the accumulation equation (9) reflects Uzawa-Lucas assumption that only human capital is used to enhancing human capital stock.

11 What Lucas called effective workforce in production (or skill-weighted man hours devoted to current production) was precisely $N(t) = uH(t)$, or $N(t) = L_PH(t)$, were L is being normalized to unity.

12 As known, equal to 0 first order conditions are required for derivatives of the Hamiltonian with respect to control variables whereas for Hamiltonian's derivatives with respect to state variables first order conditions must equal the negative of shadow prices' derivatives with respect to time.

13 Lucas (1988), p. 21.

consumption growth rate in terms of the model's variables:[14]

$$\frac{\dot{c}}{c} = \gamma_C = \frac{1}{\theta}\left(A(t)\alpha k(t)^{\alpha-1}\big(uh(t)\big)^{1-\alpha} - (\delta + \rho)\right) \tag{17}$$

In accompanying Lucas' solution for steady state values of variables c, k and h,[15] it is easily verifiable that by passing to the left hand side of equation (17) all constant terms, and taking logarithms and derivatives with respect to time, the resulting expression will fall in line with the steady state underlying principle asserting that all variables (in this case physical and human capital) must exhibit an equal and constant growth rate:

$$0 = (\alpha - 1)\gamma_k^* + (1-\alpha)\gamma_h^* \qquad or \qquad \gamma_k^* = \gamma_h^* \tag{18}$$

By dividing next for k the equation for physical capital accumulation (8), and passing to the right hand side all steady state constant terms, equation (19) was obtained:

$$\frac{c}{k} = A(t)(u^*)^{1-\alpha}\left[\left(\frac{h(t)}{k(t)}\right)^*\right]^{1-\alpha} - (\delta + n) - \gamma_k^* \tag{19}$$

from which (20) was straigthforwardly assumed to follow:[16]

$$\gamma_k^* = \gamma_h^* = \gamma_c^* \tag{20}$$

Finally, by taking logarithms of the production function for final goods (y), and derivatives with respect to time, the rate of growth of final output would be depicted by the ensuing expression (21):

$$\gamma_y = \alpha\,\gamma_k + (1-\alpha)\gamma_u + (1-\alpha)\gamma_h \tag{21}$$

which for steady state growth rate values, and given that $\gamma_u^* = 0$, permitted also to include γ_y^* in expression (22):

$$\gamma_y^* = \gamma_k^* = \gamma_h^* = \gamma_c^* \tag{22}$$

Thus far, growth rates in (22), apart from including γ_h^*, did not add any other relevant element to the already traditional conclusion of endogenous growth models; that is, in the steady state all variables grow at a similar constant rate. It is therefore important to show in what Lucas-Uzawa Model's rates differ from those yielded by other endogenous growth models (as, for instance, the AK Model) which did not explicitly include human capital stock and accumulation.

The matter raised in the above paragraph is easily dealt with by following a few simple mathematical steps whereby both sides of the first order condition (13) are multiplied by u and appropriately cancelling where required:

$$\eta_K(t)A(t)k(t)^{\alpha}(1-\alpha)u^{1-\alpha}\,h(t)^{-\alpha} = \eta_H(t)\varphi_E\,u \tag{23}$$

In taking next logarithms and derivatives with respect to time, the expression turned into (24) showing equality of shadow prices' growth rates:[17]

$$\gamma_{\eta h}^* = \gamma_{\eta k}^* \tag{24}$$

---

[14] As can be seen, the rate of growth of consumption was, in the Uzawa-Model, also function of the physical capital marginal product; nevertheless, the latter not only depends now on the stock of physical capital but also on the share of human capital stock used for the production of final goods.

[15] As the amount of human capital devoted to final goods production was a positive constant of the total stock h, the steady state value of u^* is also fixed and its rate of growth equal to 0.

[16] A constant quotient $(k/h)^*$ means that –in the steady state- both capital stocks grow at the same rate; as the same should apply to $(c/k)^*$, growth rates for consumption and physical capital will necessarily be equal and similar to the rate of growth of human capital.

[17] In obtaining (24) it must be remembered that all steady state terms in (23) were constant, except the two shadow prices.

The left hand side of (23) is identical to the first term in the right hand side of (15). Consequently, substituting it in the first order condition and cancelling terms, the steady state rate of growth of shadow price η_H is brought out:

$$\left(-\frac{\dot{\eta}_H}{\eta_H}\right)^* = -\gamma^*_{\eta h} = \varphi_E - \delta - n \tag{25}$$

By taking next logarithms and derivatives with respect to time of the first order condition (12), the ensuing equation results:

$$\frac{\dot{c}}{c} = \gamma_C = \frac{1}{\theta}\left(-\frac{\dot{\eta}_K}{\eta_K} - (\rho - n)\right) = \frac{1}{\theta}(\varphi_E - \delta - \rho) \tag{26}$$

and given that all variables must have, in the steady state, an equal rate of growth:

$$\gamma^*_y = \gamma^*_k = \gamma^*_h = \gamma^*_c = \frac{1}{\theta}(\varphi_E - \delta - \rho) \tag{27}$$

As can be seen, conversely to AK Models in which the rate of growth was affected by the production function's exogenous productivity constant, the long-run economic growth here is affected by the educational sector's productivity parameter φ_E. Needless to say, this result rests on Lucas' assumption that only human capital was used by the educative sector to producing human capital (equation (9)) and that there existed a linear knowledge production function (expression in (4')).

On the other side, feasibility of (27) will depend on the relationship between the intertemporal substitution elasticity, represented by $1/\theta$ and the productivity constant φ_E ; in this connection, expression in 4' suggested that if the entire non leisure time were devoted to human capital production (that is, if u=0) φ_E would be the maximum attainable γ_h , therefore (27) would stand if and only if and this would require in turn would the following upper limit to be placed upon the intertemporal elasticity of substitution:

$$\frac{1}{\theta} < \frac{\varphi_E}{\varphi_E - \rho} \tag{28}$$

Although not considered in the carried out review, it is important however to point out that Lucas stressed also out the possibility of knowledge having a positive external effect upon productivity, apart from the effects of and individual's on his own productivity, what he modeled as follows:

$$N(t)c(t) + \dot{K}(t) = A(t)k(t)^{\alpha}\big(u(t)h(t)N(t)\big)^{1-\alpha}h_a(t)^{\zeta} \tag{29}$$

In the above formulation the net national product (left hand side member) is still seen to depend on the levels of capital and labour inputs and on the level of a constant $A(t)$ technology, but also on the term $h_a(t)^{\zeta}$ intended to capture what Lucas called possible external effects of human capital.[18]

2.2 *The Augmented Solow Model*

In the very influential paper by Mankiw, Romer and Weil (1992), one of the outstanding features was its empirical success in revaluing Solow's Model by econometrically proving that their predictions were in principle consistent with evidence;[19] thus, while estimated coefficients' signs rightly predicted the direction of effects of investing in physical capital, and of population

---

[18] As stressed by Heijdra and van der Ploeg (2002, ch. 14, p. 463), in so doing Lucas aimed at reinforcing the notion that the formation of human capital was, in part, a social activity.

[19] In Mankiw *et al.*'s words, "…the Solow model gave the right answers to the questions it was designed to address".

growth, they failed in correctly predicting magnitudes. The matter of the assumedly failure of countries' income per capita convergence was also empirically analyzed and restated in the paper as the authors concluded that – instead of convergence – the Solow Model should rather be viewed as implying that countries would reach in general different steady states.[20]

The response to the deemed high influence of saving and population growth had to be sought, as explained below, at the exclusion of human capital from the traditional Solow Model which resulted in disproportionate larger but biased variables' estimated regression coefficients, as physical capital accumulation and population growth failed to reflect that part of their impact upon income was due to the omitted human capital variable.

The introduction of human capital within the traditional Solow Model permitted not only to solve the mentioned inconsistencies, arising when this variable, was omitted but also to use the model with greater confidence on its predictive potential. In this regard, and as is shown in the coming sections, the possibility of drawing empirically sound evidences from the model's testing enhances its policy implications with respect to the cost-benefit analysis of devoting tax revenue to human capital formation.

In presenting the augmented Solow Model, the equation (30) shows how the Cobb Douglas production function looks like after the omitted variable is included alongside physical capital:

$$Y(t) \;=\; K(t)^{\alpha} H(t)^{\beta} [A(t) L(t)]^{1-\alpha-\beta} \tag{30}$$

$K(t)$, $H(t)$ and $L(t)$ represent now the stocks of physical and human capital and labour availability respectively, $A(t)$ the technological level, $[A(t)\,L(t)]$ the effective units of labour[21] and α, β, and $(1–\alpha–\beta)$ the respective factor shares.[22] Similar to the original Solow-Swan Model, Mankiw *et al.* consider logarithmic labour and technology functions whose exogenous growth rates are respectively n and g:

$$L(t) = L(0)\, e^{nt} \tag{31}$$

$$A(t) = A(0)\, e^{gt} \tag{32}$$

The inclusion of human capital makes the model to consider now not only what determines the evolution of physical capital stock but also that of human capital, as the two ensuing capital accumulation equations show:

$$\dot{k}(t) \;=\; s_k y(t) \;-\; (n + g + \delta)\, k(t) \tag{33}$$

$$\dot{h}(t) \;=\; s_h y(t) \;-\; (n + g + \delta)\, h(t) \tag{34}$$

obtained by making $y=Y/AL$, $k=K/AL$, and $h=H/AL$ and s_k and s_h respectively standing for the fraction of income invested in physical and human capital.[23]

As in the traditional Solow-Swan Model, decreasing returns to scale entail that the economy will converge to a steady state in which $\dot{k}(t) = \dot{h}(t) = 0$ and $k(t) = k^*$ and $h(t) = h^*$; consequently, by using the production function in (30) and capital accumulation equations in (33) and (34), the following two expressions are obtained:

[20] In connection to this argument, the point was emphasized that – when differences in saving and population growth rates were taking into consideration – convergence was seen to exist at a rate in line with the model's prediction.

[21] The effective units of labour grow at the compound rate $(n+g)$.

[22] In stating that $\alpha+\beta<1$, Mankiw, Romer and Weil keep Solow's assumption of decreasing returns to physical and human capital, although the assumption that $\alpha+\beta=1$ is also critically discussed in the paper.

[23] Equations (33) and (34) do not only imply that both types of capital have the same depreciation rate but also that one unit of consumption can costlessly be changed into either a unit of physical or human capital, which notoriously differ from the assumptions upheld in the Lucas-Uzawa model.

$$k* = \left(\frac{s_k^{1-\beta} s_h^{\beta}}{n+g+\delta}\right)^{\frac{1}{1-\alpha-\beta}} \tag{35}$$

$$h^* = \left(\frac{s_k^{\alpha} s_h^{1-\alpha}}{n+g+\delta}\right)^{\frac{1}{1-\alpha-\beta}} \tag{36}$$

By substituting (35) and (36) into the Cobb Douglas production function (30), and taking logarithms, the estimable expression in (37) standing for per capita income along the balanced growth path is achieved:[24]

$$\ln\left(\frac{Y(t)}{L(t)}\right) = \ln A(0) + gt - \frac{\alpha+\beta}{1-\alpha-\beta} ln(n+g+\delta) + \frac{\alpha}{1-\alpha-\beta} ln(s_k) + \frac{\beta}{1-\alpha-\beta} ln(s_h) \tag{37}$$

It is worth emphasizing that although coefficients are still predicted as function of factor shares, the above expression is better fitted to explaining cross-country income differences, owing to the fact that human capital accumulation now accompanies population growth and physical capital accumulation. In this regard, Mankiw *et al.* pointed out in the first place that, even if $ln(s_k)$ were independent of other variables in the right hand side of expression (37), its coefficient would still be greater than in the classical Solow Model without human capital; since higher saving would lead to higher income, this would, in turn, lead to a higher steady-state level of human capital even if s_h remained unchanged, the implication being that the inclusion of human capital accumulation enlarged the impact of physical capital accumulation. Moreover, the coefficient on $ln(n+g+\delta)$ is, in absolute value, greater than $ln(s_k)$'s coefficient reflecting the fact that high population growth lowers income per capita as physical and human capital stocks need now to be spread over more individuals.

Mankiw, Romer and Weil also suggest an alternative way, stemming from the combination of (37) and the steady-state level of variable h in (36), whereby the impact of human capital upon per capita income can be highlighted. As can be seen below, the resulting equation renders now income per capita as a function of the propensity to accumulate physical capital, the population growth rate and the level of human capital:[25]

$$\ln\left(\frac{Y(t)}{L(t)}\right) = \ln A(0) + gt + \frac{\alpha}{1-\alpha} ln(s_k) - \frac{\alpha}{1-\alpha} ln(n+g+\delta) + \frac{\beta}{1-\beta} ln(h^*) \tag{38}$$

As there exist now two variants for the Augmented Solow Model's econometric estimation; that is, one in which the rate of human capital accumulation is resorted to and another including the level of human capital, Mankiw *et al.* aimed at empirically sorting out the posed testing dilemma by suggesting to verify – in the first place – whether human capital's available data corresponded to (s_h) or to (h) a matter that, for Argentina, will be dealt with in the next section.

---

[24] The point is worth mentioning that, for Mankiw *et al.*, $lnA(0)$ also reflects, apart from technology, other features such as resource endowments or institutions, therefore the term is better depicted as being equal to $\alpha+\varepsilon$ where α is a constant and ε stands for a country's specific shock.

[25] It is easily noticed that the structure of 38 is practically similar to the traditional Solow-Swan equation without human capital in which the latter is part of the error term. Since saving and population growth rates influence h^*, human capital should be expected to be positively correlated with the saving rate and negatively in turn with population growth. In reason of this Mankiw *et al.* suggested that omission of the term on h^*, in Solow's Model, biased coefficients on saving and population growth.

3 Methodologies for computing the human capital stock and the average propensity to invest in human capital

3.1 *Preliminary ideas*

Even acknowledging the difference between this paper's aims and those in articles which explicitly refer to economic growth, such as Mankiw *et al.*, the construction of a variable that clearly serves the purpose of capturing the effect of human capital upon gross domestic product contributes to enriching future empirical results. Simple as it was, Mankiw's proxy did what it was intended to do, but it fell short from unveiling the policy effectiveness of budgetary efforts directed to human capital creation, therefore more accurate measures are in order.

Even by restricting to a single narrow variant of human capital, i.e. investment in education, Mankiw *et al.* acknowledged from the outset the "practical difficulties" involved in the variable's measurement, particularly if the model's second alternative (involving human capital level) were aimed at for econometric estimation. On grounds therefore of statistical feasibility, the first alternative was resorted to by using a proxy for the rate of human-capital accumulation (s_h) which simply approximated the percentage of the working-age population actually enrolled in secondary school; however, the authors pointed out that the measure was not free from flaws, at least on the following four accounts: primary and university education were not included, the input of teachers was also ignored, students' forgone earnings and their variation with the level of human capital investment were not considered, and the proxy resulted from two data series respectively embodying the eligible population (12 to 17 years) and the working age population of school age (15 to 19 years) that clearly covered different age ranges. Needless to say, these flaws did not impede that a one sector model were used; the mentioned omissions and inconsistencies would be however a bounding restriction should a proper production function for human capital were included.

In the light of the above comments, efforts in the rest of the section are oriented to describing components of investment in human capital and to computing both the variables better representing in Argentina the level of per capita human capital (h) and the rate of human capital accumulation (s_h); while the former is required for the estimation of equation (38), the latter, whose new computed value seeks to avert the criticisms Mankiw *et al.* placed upon their proxy variable, is in turn used for testing equation (37).

3.2 *An alternative estimation of* s_h *and* h

From the outset, the specification of what "investment in human capital" will mean or include is crucial as, despite that much has been said and written in this matter, the need of counting with an econometrically practical variable and the scarceness of available data imposed always severe constraints. In this connection, the following principles governed the methodology followed to achieving variables standing for human capital in Argentina:

i) Notwithstanding the relevance of activities in the form of health and construction and maintenance of public goods, whose importance as components of human capital was particularly stressed by Uzawa as they resulted in an improvement of labour efficiency, difficulties involved in gathering data[26] and jointly dealing with all of them advise to focus only in investment in education.[27]

[26] This was particularly true for health expenditures as major modifications underwent by the system during the period considered made very difficult to obtain statistical series while at the same time benefits rendered by their inclusion were scant.

[27] Technical knowledge (derived from investment in education), must be built upon an inherited social capital, should it be expected to *(continues)*

ii) In correctly ascertaining the real value of the variable, the opportunity costs of investment in education; that is, the forgone income of working age students, should be determined and added to the actual budgetary resources component. The importance of opportunity costs in empirical work has repeatedly been noticed, as was the case in Kendrick's calculations (1976). Maintenance costs of university students, borne by parents, must also be taken into account as a component of opportunity costs.

iii) Investment in education is an all inclusive term, therefore primary, secondary and higher education, as well as science and technology, are also encompassed.[28]

iv) Budgetary expenditures in the field of Culture are excluded on grounds that they generally yield consumption rather than productive goods.

v) In a country like Argentina, characterized by a federal institutional setting in which investment in education spreads over the three government levels, the variable's right assessment calls for national, provincial and municipal spending in education to be altogether considered.[29]

In order to meet the preceding general guidelines, the variable standing for investment in human capital is built considering the following methodological principles:

a) Educational expenditure is an overall item including actual budgetary outlays in basic education (primary and secondary levels) and higher education (tertiary and university studies) of all the three government levels: central government, provinces and municipalities.

b) National and subnational spending in science and technology is also included, inasmuch as they aim at raising productivity by helping to develop the current state of the applied scientific knowledge and productive techniques.

c) Minima legal wages are used to approximately computing opportunity costs on the following two grounds: they by definition represent households' cost of basic needs whereas they also serve as a proxy for incomes earned for working age students still no having completed their higher studies.
Nevertheless, secondary students' maintenance costs borne by parents are not added in opportunity costs, the idea being that households customarily support children up to the age of eighteen. By the same token, not forgone incomes are suppose to exist in the case of secondary students under fifteen as labour regulations and practical limitations are more strictly applied upon this particular age range. This explains the decision not to compute forgone earnings for secondary students under fifteen while only a minimum legal wage was taken for students above this age.

d) Contrariwise to what is asserted in the preceding paragraph, higher education and university students are expected to somehow support themselves, therefore the following three cases may be considered: a) they work full time and bear their maintenance costs; b) they work part time but their parents still bear their maintenance cost and c) they do not work at all and therefore, apart from forgone incomes, their maintenance cost is also borne by their parents. These three categories serves to explain the opportunity cost structure that follows: a double legal minimum wage is assigned to the percentage of higher education students who, according to statistical information drawn from household and university surveys, do not work; in this case, one

improving the country's productivity matrix. In this context the expression embodies elements such as institutions, values and social and collective behaviour.

28 Expenditures devoted to different university's careers are not made explicit at this stage. Given that disciplines can have different marginal impacts on gross domestic product this could be a further step in future investigations.

29 As of the nineties, primary and secondary education became in Argentina a provincial budgetary responsibility, the national government performing thereafter a subsidiary role through annual transfers sent to the subnational level (based on the so-called Ley del Financiamiento Educativo 26075). The national government keeps in change the responsibility of wholly financing national public universities whereas spending in science and technology is a shared commitment, though mostly funded by the central government.

minimum wage accounts for forgone earnings and the other for students' maintenance costs borne by parents. For the percentage of students having a job but still receiving economic support from their families, no forgone incomes are assumed and only one minimum wage is computed in order to reflect maintenance's costs. For students that work and defray their own expenses no opportunity costs are assigned.

e) Outlays in a)-b) above stand for the investment in human capital restricted to budgetary expenditures in Education, Science and Technology. By including c)-d) an augmented version of human capital investment is obtained which also includes opportunity costs. By dividing both variants of investment in Education by gross domestic product, average propensities to invest in human capital result.

f) In building up series for human capital stock the conventional assumption is upheld that actual educational investments, similarly to physical capital, are subject to an annual depreciation rate of 10 per cent.[30] The reason for using a single depreciation rate for both capital assets not only responds to computational simplification, but also seeks to reduce the loss of degrees of freedom: should more than one depreciation rate be used, more parameters will have to be estimated and the data constraint binds tighter.

g) All variables are in real terms, deflated by CPI series (see sources in Annex 1).

3.3 Variables' specification

Once components of investment in human capital are completely assessed and included, both variants of the average propensity to invest in human capital are computed; nevertheless, only the variant "average propensity to invest in human capital (inclusive of opportunity costs)" is used in the econometric estimation. The variable's computed values (with and without opportunity costs) are shown in Annex 1, whereas that its performance over time is depicted by figures in next section, in which stylized facts related to human capital performance in Argentina are considered.

As for human capital stock (H), the annual value of the variable includes the preceding years' still not depreciated investment together with the year's actual not amortized investment (e.g., if 1998's human capital stock is to be computed, 90 per cent of the year's investment is included plus the remaining not amortized investments from previous periods). The value of variable human capital stock (H) needs not be confused with the variable (h) in equation (38), representing per capita human capital stock.[31] The variants included here are in line with different forms of regarding human capital (with and without opportunity costs) and their graphical evolution is considered in the next Section. Although computing both H and h appears like a major step in fathoming with some of postulated questions, the econometric use of these variables has not proven fruitful in the present step of the investigation. Nevertheless, a better performance is expected from theirs being used in an endogenous model, where a human capital production function is included.

4 Stylized facts concerning the evolution of gross domestic product and human capital formation

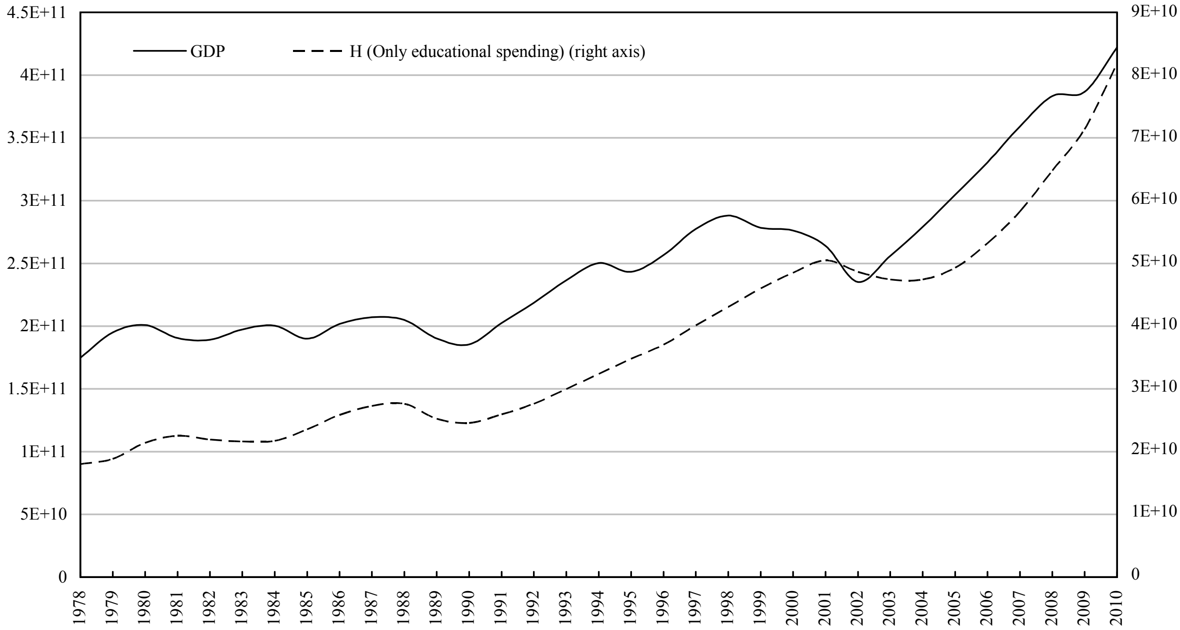
In analyzing Figure 1 below, tracing the evolution, as of 1978, of the Argentine gross domestic product and human capital stock, this having been computed as explained in Section 3

[30] It is obvious that this simplifying assumption does not rule out alternatives; thus, while Mankiw *et al.* prefer a longer amortization period (a smaller depreciation rate), the argument may also be defended that amortization need not be linear but decreasing.

[31] See the value for h (inclusive of opportunity cost) in Annex 1, quoted as HOCPC.

Figure 1

Series Stand for Gross Domestic Product and Human Capital Stock at Current Prices



and including only budgetary outlays in Education,[32] an immediate feature deserving being stressed is the direct correlation found between both series. A more careful inspection of the figures, however, sheds light on the matter of causation closely, which, in turn, is related to objectives motivating this research. As can be seen, the gap between GDP and H shrinks in time in coincidence with the working of the so-called "Ley de Financiamiento Educativo", whereby educational spending should be gradually increased until it reaches a determined percentage of GDP. One important preliminary conclusion, verified below by the econometric results and running counter to what it would have been expected, is that GDP clearly hauled human capital formation (represented here by investment in education), with little evidence of the reverse causation order significantly taking place.

A conclusion somehow similar to the one just arrived upon in the previous diagram can be drawn when GDP and H's growth rates are jointly assessed, as in the following Figure 2: strikingly, except for a few periods in which both growth rates exhibited the same pattern, there seems not to be a particular positive correlation between the series' respective maxima and minima values; thus, growth rates, rather than coinciding, behave differently in a large part of the period considered and it is also noted that when both have a decline – as in the period 1983-2002) the fall is more deeper in the case of the gross domestic product growth rate. In line with what the cointegration analysis will show in Section 5, bad or good performances of the overall Argentine growth rate seem to be based in factors no considered here and it can hardly be argued that investment in education significantly counted as one of them.

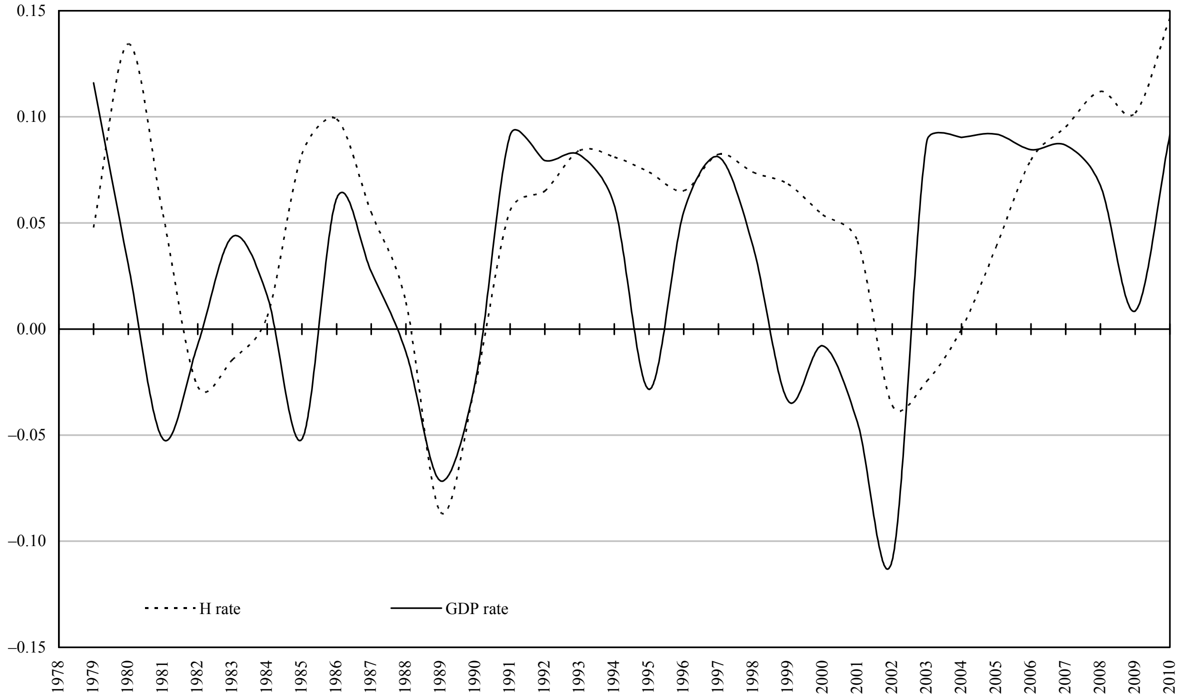
It is therefore important to point out that, however expected the evolution of human capital stock following the path traced by GDP (mainly due to the form in which the variable was computed), hopes that H would somehow behave as a GDP's growing factor or stabilizer can be hardly fed from evidences in the figures shown.

---

32 That is, investment in education is here computed exclusive of opportunity costs.

Figure 2

Gross Domestic Product and Human Capital Stock Growth Rates Derived from the Respective Series in Current Prices



The next diagram, in which the ratios of investment in education over gross domestic product and over the previous period's gross domestic product (the lag of the same variable) are respectively plotted, not only enriches the analysis of the real impact of human capital upon product but also help in reasserting conclusions derived in the preceding paragraphs by introducing an element that has so far not been considered. The steady increase of H throughout the whole period (see Figures 1 and 3) is seen to be practically accompanied by a similar performance of ratios H/GDP shown in Figure 3, except for some isolated cyclical decreases the latter underwent; since ratios stand for human capital stock per unit of product, it is possible to argue that the nature, quality and efficacy of human investment (measured as outlays for education) fell short of what was expected in terms of their product enhancing capacity and that may in turn explain why an incremental product-investment in education relationship failed to prevail.

Suffice it to mention that the second ratio was aimed at ascertaining whether human capital formation had a lagged impact upon product; needless to say, this hypothesis could not either being proven as the similar pattern exhibited by dashed line ruled out chances of a clearer relationship and higher impact between variables stemming from taking policy variables' lagged values.

The conclusion obtained from the graphs in Figure 3 is still more evident when the plots of product and average propensity to invest in human capital growth rates, shown in Figure 4, are carefully observed. Even though the former (already shown in Figure 2) shares its cyclical behaviour with S_h, ups and downs of the average propensity to invest in education and technology's growth rate were by far much more marked, and yet this did not seem to have had a definite weight upon the evolution of the product's growth rate, let alone the fact that their performance run counter in several time spans during the period analyzed.

Figure 3

Ratios of Human Capital Stock/Gross Domestic Product in 1993 Prices
(figures in left and right vertical axis respectively stand for the value of the ratio and the human capital stock)

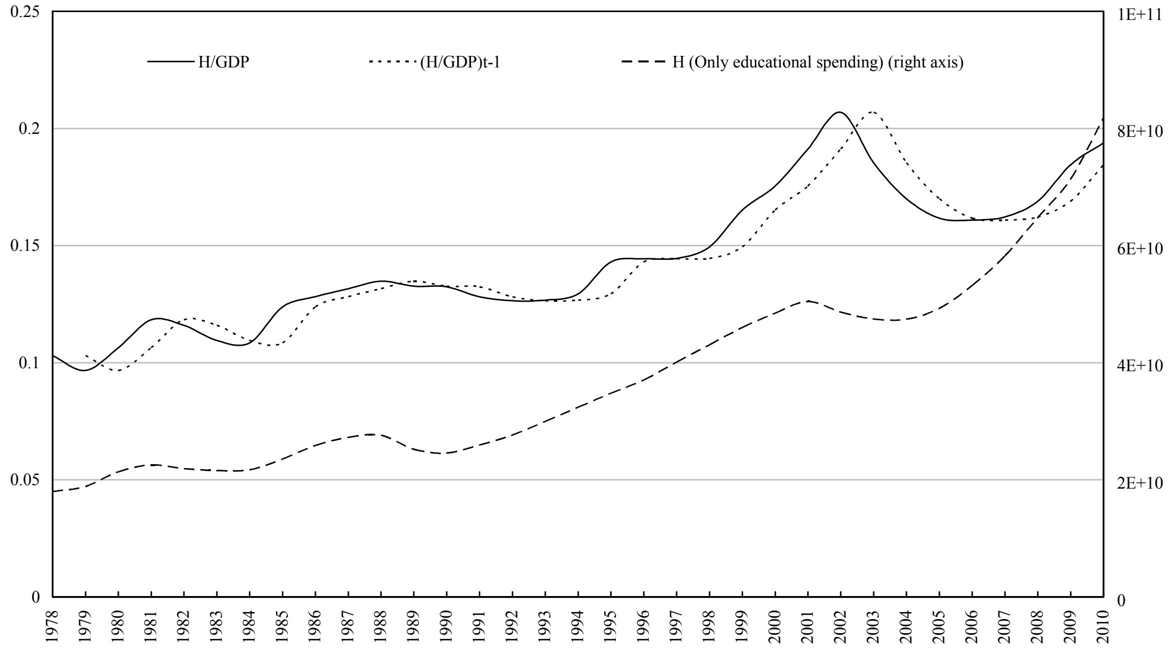


Figure 4

Rates of Growth of the Average Propensity to Invest in Human Capital (When Opportunity Costs are Not Considered) and of Gross Domestic Product

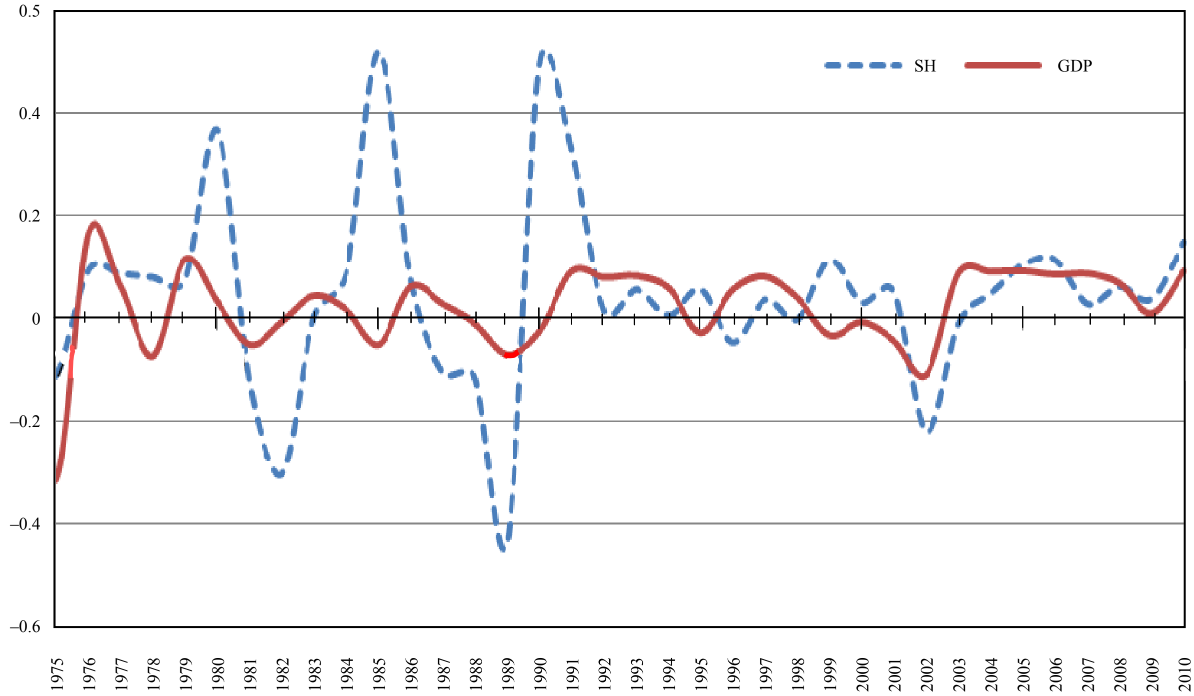
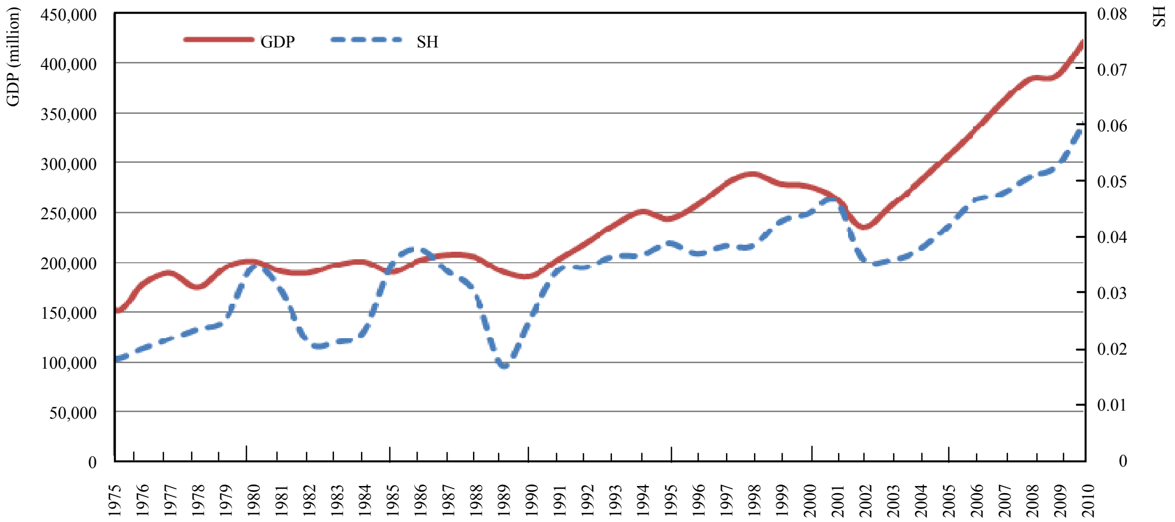


Figure 5

Series for Gross Domestic Product and Average Propensity to Invest in Human Capital



When the gross domestic product and average propensity to invest in human capital series are plotted together, as in Figure 5, their evolution did not seem to offer explanations different to what has so far been presented: for the first part of the period, S_h exhibited a marked cyclical behavior not accompanied by the steady low growth path of product while the stable increase of S_h as of 2003, for reasons given above, did not seem to have produced any particular incremental effect upon product but rather the other way round.

The performance in the period 1991-98 is however worth mentioning as it seems to have been the only case in which human capital formation exerted any incremental effect upon product; this situation was also reflected in Figure 4, as can be easily noticed when the behaviour of product and average propensity rates of growth is observed.

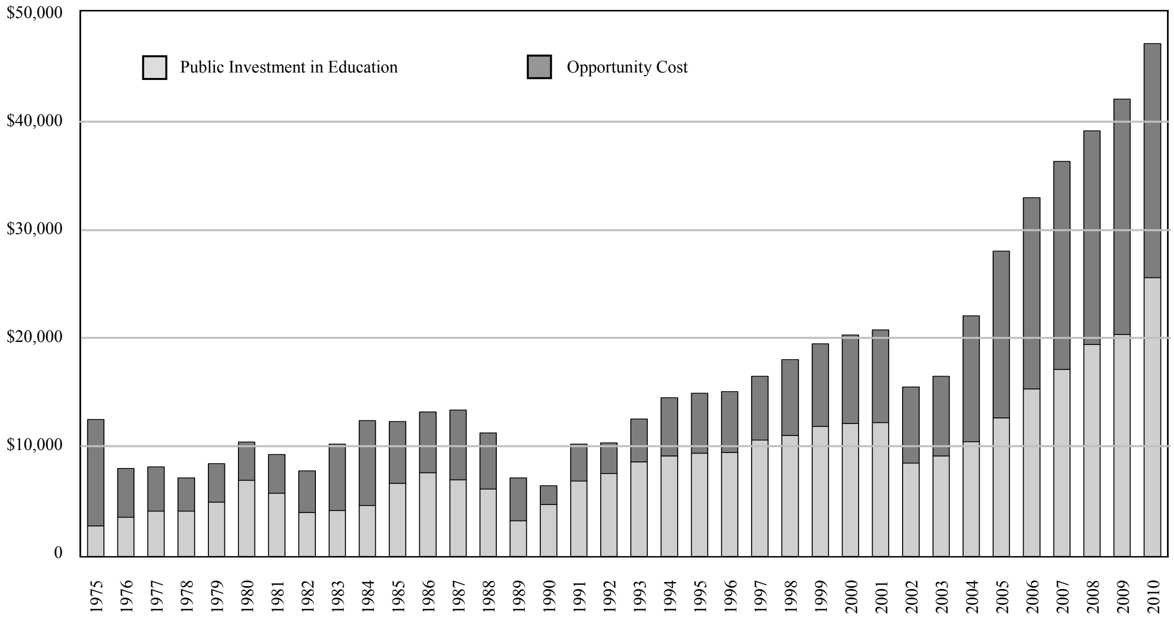
The presentation of stylized facts is completed with the analysis of the following figures in which the overall concept of investment in education, embodying budgetary outlays as well as the opportunity costs (as defined in the preceding section) is considered. In the first place, the graph in bars of Figure 6 showing the evolution of the actual investment in education aims at highlighting how their two components evolved throughout the period.

The first worth pointing out evidence shown by Figure 6 is that students' forgone earnings and maintenance costs supported by parents have been an important component of the overall investment in education all throughout the period; in this regard, the very magnitude of opportunity costs as a representation of the burden implied for the society as whole poses a question whose answer falls well beyond this paper's reach but that seems anyhow worth ascertaining in terms of cost benefit analysis.

The second evidence yielded by Figure 6 is that opportunity costs' percentage share within investment in education was not stable but underwent significant variations throughout the years. The explanation for that must be sought at the form opportunity costs were computed; that is, in terms of minima legal wages. It is therefore clear that opportunity costs' share of investment in education was straightaway conditioned by updating opportunities of minima legal wages by the government.

Figure 6

Public Investment in Education (Budgetary Outlays) and Opportunity Costs (Forgone Incomes and Maintenance Costs Borne by Households)
(million pesos of 1993)



Finally, the evolution of the overall average propensity to invest in education (Figure 7) is in turn split in order to show its two components' actual weight. Although bars in Figure 7 are expected to follow the pattern set by the investment in education in Figure 6, figures for S_h permit in turn to add some additional comments that shed light on human capital performance in Argentina during the period considered. In the first place, the evolution of both the overall average propensity to invest in education as well as components' share did not appear to follow a definite pace, conversely to what by and large happened as of the nineties.

However, one interesting feature revealed in the bar Figure 7 is that, apart from the positive effect of parliamentary mandated increases in education outlays, which subsequently raised the percentage participation of investment in education to gross domestic product, the opportunity cost component grew steadily as of the nineties to the extent that its participation ranged between 40 per cent and 45 per cent of the overall average propensity to invest.

5 Econometric estimation for Argentina of an Error Correction Model

5.1 Theoretical aspects of the Error Correction Model

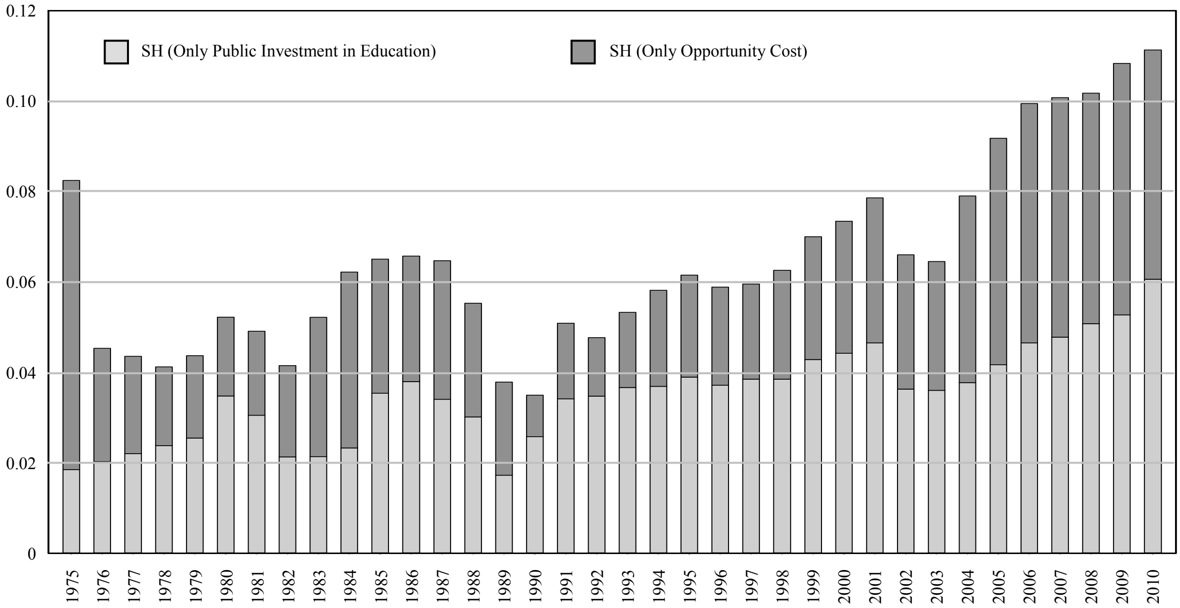
As known, an error correction model responds to the following structure:

$$\Delta X_t = \pi_0 + \pi X_{t-1} + \sum_{i=1}^{p} \pi_i \Delta X_{t-i} + \phi D_t + \varepsilon_t \tag{39}$$

where ΔX_t stands for a $(n \times 1)$ vector representing the set of endogenous variables, π_0 is a constant terms vector included in the VAR, ΔX_{t-1} stands in turn for the "i periods" lagged vector

Figure 7

Average Propensity to Invest in Education



of variables while the dummies vector D aims at capturing the model's structural break points. The term πX_{t-1} is important in so far as it differences the ECM from a VAR in differences by incorporating information contained in variables in levels; matrix π results from the product of matrices α' and β', the first embodying speed of adjustment parameters to short term changes respect of long run (or equilibrium) relations whereas the second one holds cointegration coefficients by means of which a linear combination of order one integrated variables comes up to be stationary. Thus, equation (39) can be similarly represented by the following expression:

$$\Delta X_t = \pi_0 + \alpha\beta' X_{t-1} + \sum_{i=1}^{p} \pi_i \Delta X_{t-i} + \phi D_t + \varepsilon_t \tag{40}$$

The rank of matrix $\pi = \alpha\beta'$ suffices to determine the number of cointegration equations: if it were zero, the matrix would be null $(\pi=0)$ and the model would be stated in terms of a VAR(p) in differences; if there is, on the contrary, a complete Rank matrix, all variables will be stationary, as a stationary variable cannot be equaled to a non-stationary one (in this case integrated of order one).

When the rank of π is r, (for $0<r<n$), there will be r cointegration equations, β will be now a $(n\times r)$ matrix, and product $\beta' X_{t-1}$ generates stationary variables that will stand for the short-run disequilibria with respect to each of the long-run relations. Matrix α also $(n\times r)$ holds the parameters determining the adjustment speed *vis-à-vis* these disequilibria.

The Johansen Methodology permits to calculate the rank of π by means of a Dickey-Fuller multivariate proof,[33] from which characteristic roots are obtained; the amount of distinct-from-0 roots will indicate the rank of π and the amount of linearly independent cointegration equations.

---

[33] When having the expression $X_t = A_1 X_{t-1} + \varepsilon_t$, in which X is a vector, the Dickey-Fuller Proof permits to check whether the matrix π in $\Delta X_t = (A_1 - \mathrm{I})\, X_{t-1} + \varepsilon_t$, or in $\Delta X_t = \pi\, X_{t-1} + \varepsilon_t$, is null or not.

Trace and Maximum Eigen Value Statistics are used to identify the number of statistically different from zero roots: while the former one test the null hypothesis that the number of linearly independent cointegration equations is equal to or smaller than r, as against the alternative of greater than r, the second test is used to check the null hypothesis that the number of cointegration equations is r as against the alternative $r+1$.

It is expected to find, for the Augmented Solow Model, only one long run relation representing equation (37) above, from which all produced disequilibria will force variables to move till they newly reach equilibrium, both by means of long run effects included in the error correction term and through the VAR's short run effects.

5.2 Econometric estimation of the ECM for the Argentine case

The assessment of the impact of human capital upon the Argentine per capita gross domestic product is carried out for the period 1975-2010. Diverse data sources were resorted to in order to construct the series necessary for the econometric estimation of variables' coefficient, whose detail is referred to in Annex 1.

As will be shown below, variables in levels are not stationary (that is, not $\sim I(0)$), which can bring about the problem of spurious correlation and its undesired effects. Despite the fact that some controversy still exists in the literature as to whether to discard non stationary variables in time series regression, other solutions are at hand to deal with the problem,[34] as is the case of cointegration and the error correction model developed in the preceding section and used in this paper for estimating the previously introduced equation (37):

$$ln\left(\frac{Y(t)}{L(t)}\right) = \ln A(0) - \frac{\alpha + \beta}{1-\alpha-\beta}\ln(n+g+\delta) + \frac{\alpha}{1-\alpha-\beta}\ln(s_k) + \frac{\beta}{1-\alpha-\beta}ln(s_h) \qquad (37)$$

Variables used in order to estimate the model are described below:

$ln[Y(t)/L(t)]$**,** indicating the log of per capita (or per effective labour unit) income (hereafter quoted as GDPPC);

$ln(n+g+\delta)$**,** standing for the log of the sum of population and knowledge rates of growth plus the depreciation rate (hereafter quoted as NDG). As it is obvious, the coefficient must be negative since the effect of a raise in the first two rates – by increasing both the population and the number of effective units of labour – will be a smaller per capita o per worker income.

$ln(s_k)$, $ln(s_h)$, respectively showing the log of the propensity to invest in physical (SK) or human capital. As in the previous case, their positive coefficients will indicate the expansive effect exerted by higher propensities. As said, s_h admits the two variants: actual expenditures in education over gross domestic product (SH) and actual expenditures in education plus opportunity costs over gross domestic product (SHOC); the statistical software EViews was used to obtain the econometric results of regression equation (37) shown above.

As the estimation process requires, in the first place, the order of integration of series used to be determined, Table 1 shows results of unit root tests;[35] as can be noticed, variables are not stationary.

---

[34] In particular, the risk of spurious regressions disappears if a lineal combination of non stationary series happens to be stationary or $I(0)$. As Rezk and Irace (2008) pointed out, the economic significance is in this case no minor as the existence of cointegrated series indicate in turn a long run equilibrium relation among the variables.

[35] All variables are in logs and the amount of lags used for the Dickey-Fuller Test was automatically determined by Schwarz Information Criterium. The human capital stock (h^*) was not used in this case as its first difference turned out to be not stationary (see tests in Annex 3, Table 7).

Table 1

Augmented Dickey-Fuller and Phillips-Perron Test

Levels	ADF	PP
GDPPC	–1.301158	–1.277143
NDG	0.450308	0.180568
SK	–2.979997	–2.501809
SHOC	–1.334194	–1.320521
SH	–1.471598	–1.120892

First Diff.	ADF*	PP*
Δ*GDPPC*	–6.328994	–6.328994
Δ*NDG*	–5.232868	–3.616207
Δ*SK*	–5.060128	–4.931142
Δ*SHOC*	–5.350248	–6.77113
Δ*SH*	–6.315386	–10.48566

* In all cases, the null hypothesis is rejected for/at 1 per cent significance level.

Table 2

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Max-Eigen Statistic	0.05 Critical Value
None*	0.819268	98.19719	54.07904	61.58658	28.58808
At most 1**	0.390939	36.61061	35.19275	17.8501	22.29962
At most 2	0.26849	18.76051	20.26184	11.25518	15.8921
At most 3	0.188184	7.505336	9.164546	7.505336	9.164546

* The hypothesis of no integration equations is rejected for/at a 5 per cent significance level, as against the alternative of one equation (Max-Eigen Vaule) or more than one (Trace).
** The Trace Test rejects the null hypothesis of one cointegration equation as against the "more than one" alternative whereas the Max-Eigen Test does not reject the null hypothesis of one cointegration equation as against the "two cointegration equations" alternative.

Johansen Cointegration Method was resorted to for the ECM estimation, including respectively a constant term (both in the cointegration equation and the VAR) and a dummy for year 2002, when the country incurred in default of its external debt.[36] Given the constraint imposed by the scarce data availability, variables were in turn allowed only one lag.

Results for the Johansen Test are shown in Table 2 above, in which the computed Trace and Maximum Eigen statistics were compared with their respective critical levels for a significance level of 0.05.

Even though the Trace Statistic seems to suggest that two long run equilibrium relations exist, only one is pointed out by the Max-Eigen Statistic, therefore the ECM is finally estimated with one long run relation.[37]

---

36 The dummy variable previously included for 1989, the year of hyperinflation, but it was discarded as it turned to be not significantly different from 0 in all cases.

37 On the basis of results yielded by the Trace Statistic, the error correction model was also estimated for two cointegration equations but, in one case, results were scarcely significant.

Table 3

The Cointegration Equation

	GDPPC		**C**	**NDG**	**SK**	**SHOC**
Coefficient	1	=	8.47	–0.33	0.34	0.17
t-statistic			*(31.14)*	*(–6.88)*	*(11.17)*	*(5.08)*

Cointegration coefficients (β matrix in equation (40)) turned out to be significant and held also the expected signs; after coefficients are normalized, and taking GDPPC as the dependent variable, the expression representing equation (37) of the Augmented Solow Model is obtained, as shown in Table 3.

Deviations with respect to this long run equilibrium relationship are stationary as shown by Table 4 in which the hypothesis of unit root is rejected at a 1 per cent significance level both by the Augmented Dickey-Fuller and Phillip-Perron Tests:[38]

Table 4

Augmented Dickey-Fuller and Phillip-Perron Tests

Level	**ADF Stat.***	**PP Stat.***
Coint. Eq.	–2.710985	–7.881003

* Unit root rejected at 1 per cent.

As short-run disequilibria are incorporated by the Error Correction Model via an "error correction vector", their actual impact upon endogenous variables is in turn determined by coefficients included in matrix α (equation (40)) standing for the adjustment speed. Table 5 shows the estimation's outcome.

The first column standing for matrix α, vector in this case, reveals that the speed of adjustment coefficients for both the gross domestic product and variable NDG are significantly different from zero, thus confirming the existence of an error correction vector. The following four columns (*endogenous variables*) represent matrix π_1 corresponding to the endogenous variables' first lag; finally, the last two columns stand for the vector of constant terms and dummies for year 2002 respectively.

Once the econometric estimation of coefficients is performed, tools provided by "innovation accounting" allow to assess the used model's adequacy, therefore the consideration of some impulse response functions (complete graphical detail in Annex 2) is accompanied by a variance decomposition analysis.[39] In the first place, Figure 8 highlights variables' response to a positive innovation in the average propensity to invest in human capital (inclusive of opportunity costs).

[38] The amount of lags used resulted from the Schwarz Information Criterium (see Table 8 in Annex 3).

[39] For IRF and Variance Decomposition the Cholesky Decomposition was resorted to with the following imposed variable ordering: SHOC-SK-NDG-GDPPC.

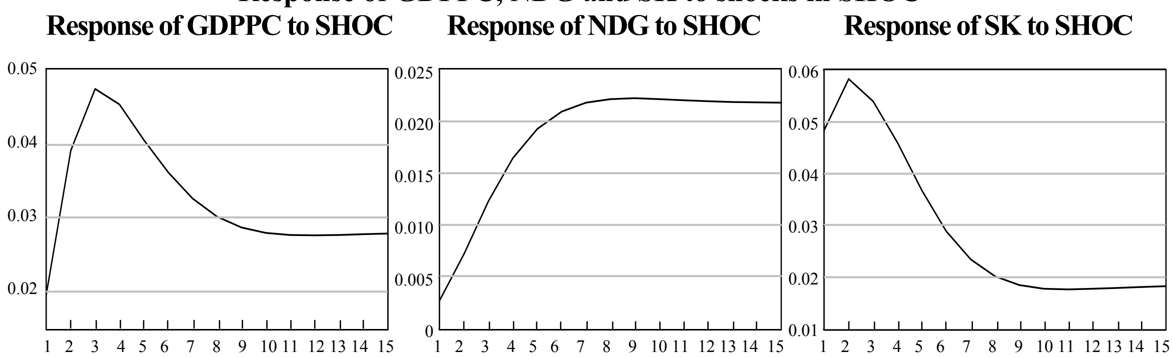
Table 5

	(α_i)	Endogenous Variables				Exogenous	
	$C.Eq_{t-1}$	$\Delta GDPPC_{t-1}$	ΔNDG_{t-1}	ΔSK_{t-1}	ΔSH_{t-1}	C	$D02$
$\Delta GDPPC_t$	**–1.002** *(–6.13)*	**0.3** *(2.13)*	–1.3 *(–0.19)*	–0.18 *(–1.79)*	0.027 *(0.34)*	0.009 *(0.54)*	–0.19 *(–2.84)*
ΔNDG_t	**–0.011** *(–5.70)*	**0.008** *(4.96)*	**0.76** *(9.56)*	–0.003 *(–2.74)*	0.0003 *(0.37)*	–0.0004 *(–2.06)*	–0.001 *(–1.22)*
ΔSK_t	0.49 *(1.51)*	0.23 *(0.83)*	–24.43 *(–1.84)*	0.19 *(0.92)*	0.088 *(0.57)*	–0.033 *(–1.08)*	–0.34 *(–2.49)*
ΔSH_t	0.46 *(1.38)*	**0.97** *(3.41)*	–7.29 *(–0.53)*	**–0.4** *(–1.95)*	0.24 *(1.52)*	0.003 *(0.08)*	–0.2 *(–1.37)*

* t-statistics in parentheses.

Figure 8

Response to Cholesky One S.D. Innovations
Response of GDPPC, NDG and SK to shocks in SHOC[40]



As expected, gross domestic product's response to increases in the rate of expenditure in Education is positive and particularly greater in the first periods following the shock. The positive reaction of *NDG vis-à-vis* a *SHOC* innovation may be indicating that rises in the rate of investment in education somehow leads to more units of effective labour, as Mankiw *et al.* (1992) stated it in the Augmented Solow Model. Finally, *SK* also reacts in a positive way to a sudden rise in SHOC, which seems to suggest a sort of complementarity feature between both productive factors; nevertheless, the feature reverts when the impact on SH of a shock in SK (Figure 9) is considered, as in this sequence the negative response seems to indicate substitutability between both factors which deserves at least a further analysis.

[40] The used software EViews does not graphically show confidence intervals for the impulse response functions.

Figure 9

Response to Cholesky One S.D. Innovations
Response of SHOC to shocks in NDG, GDPPC and SK

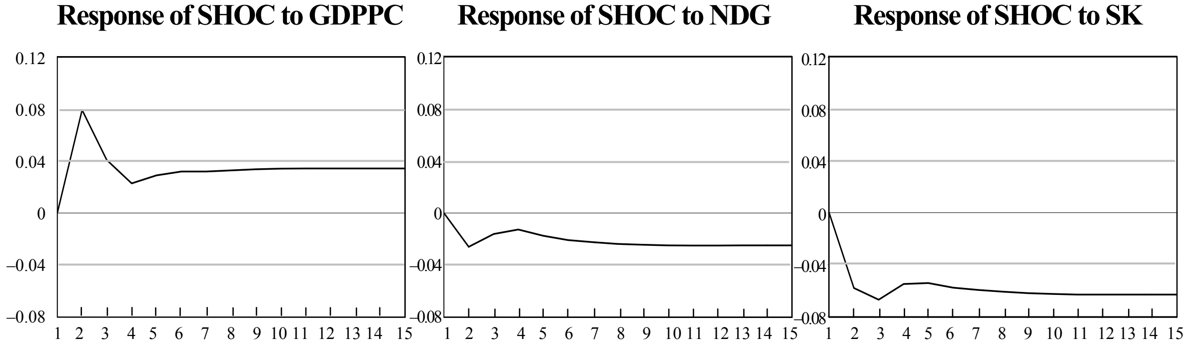
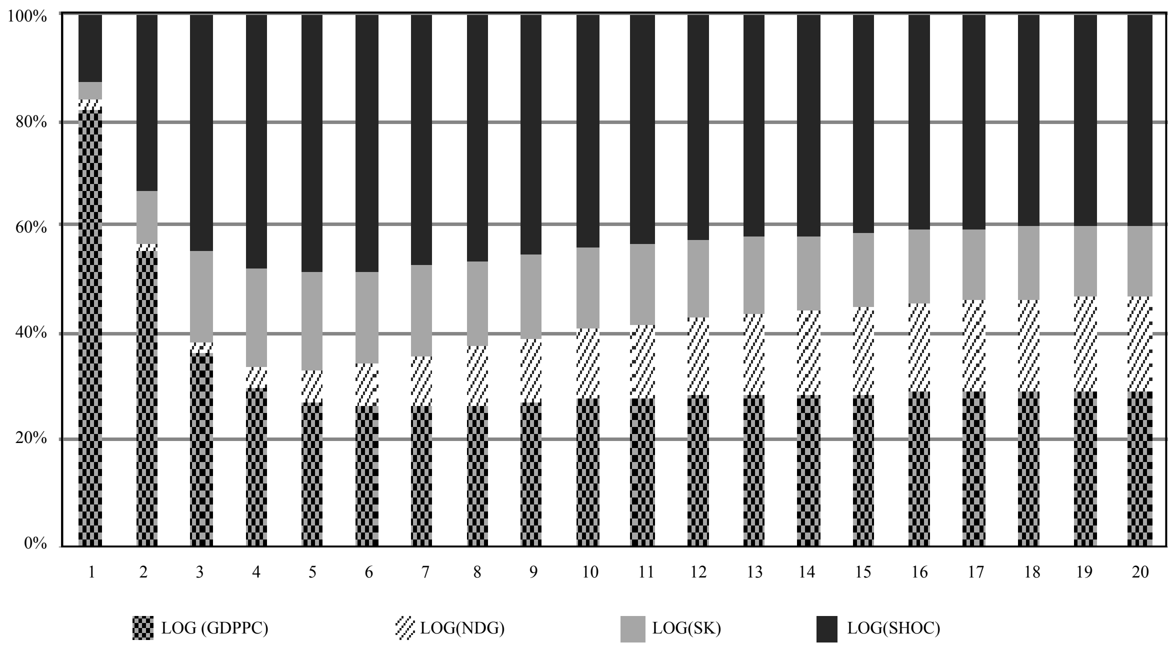


Figure 10

Variance Decomposition of Log of Gross Domestic Product Per Capita (GDPPC)



In completing the analysis of graphs in Figure 9, the expected positive response of SH to innovations in GDPPC reflects not only the common sense perception that societies will raise their demand for human capital formation as income per capita increases but, and for the Argentine case, the parliamentary decision that budgetary spending in Education should gradually reach 6 per cent points of GDP.

The recourse to variance decomposition permits in turn to ascertain the extent to which more relevant variables' total variance is explained by their own variance as compared to explanation given by other variables' variance. In order to illustrate the preceding statement, Figure 10 is used to show variance decomposition in the case of GDPPC.[41]

In spite that both impulse response functions and variance decomposition reveal that human capital investment (measured here as investment in Education) somehow impact upon gross domestic product, values for the Granger Causality Test for the GDPPC equation (shown in Table 6) run counter the preceding evidence since, only for the case of average propensity to invest in physical capital, the hypothesis that the s_k does not cause gross domestic product is rejected at/for a 10 per cent significance level (that is, SK Granger Causes GDPPC), whereas the non-causality hypothesis cannot be rejected for the rest of variables. Therefore, for the case expected to entail policy implications, preliminary results show that SHOC does not Granger-Cause GDPPC.[42]

Table 6

VEC Granger Causality/Block Exogeneity Wald Test

Dependent variable: D(GDPPC)

Excluded	**Chi-sq**	**df**	**Prob.**
D(NDG)	0.039624	1	0.8422
D(SK)	3.230270	1	0.0723
D(SHOC)	0.168026	1	0.6819
All	3.601399	3	0.3078

On the other side, it can be noticed that both GDPPC and SK Granger Caused SHOC (see Annex 3, Table 10), which is not an unexpected outcome regarding GDPPC, due to the already quoted parliamentary acts mandating that educational spending should gradually reach a percentage of product.

In conclusion, cointegration analysis and the error correction model enabled the empirical study to be carried out even though the involved variables were not stationary, and permitted also to verify the existence of a long run equilibrium relationship between gross domestic product and average propensities to invest in physical and human capital and population growth rate.

Furthermore, the error correction model with one lag permitted to find short run relations the most notable being the one between product and s_h which, conversely to what was expected and suggested by the Augmented Solow Model, indicated inverse causality; that is, from product towards SHOC but not from the latter to the former variable.

Nevertheless, impulse response functions as well as variance decomposition analysis do show a human capital participation or impact upon the trajectory of product due to the incorporation of a cointegration equation in the model.

[41] Fort the rest of variables, variance decomposition is shown in graphs of Annex 4.

[42] Although model included only one lag, Granger-causality was not reverted when it was allowed to include a larger amount of lags (see Granger-Causality Test in Annex 3, Table 10).

It is also worth pointing out, as a final comment, that residuals are normally distributed and that no heteroskedasticity was found when the joint test was performed; some point problems of autocorrelation were however detected. Test results are shown by Tables 11, 12 and 13 in Annex 3.

6 Preliminary conclusions

The proposed methodology allowed a new way of computing the series of marginal propensity to invest in human capital and of human capital stock in Argentina, which were later used in estimating the key equations of the Augmented Solow Model. One key aspect of the new methodology was that the variable standing for human capital formation (represented by Investment in Education) also included opportunity costs.

Given the econometric problems caused by variables' non stationarity feature, usual estimation procedures were discarded and alternative approaches, such as cointegration and the Error Correction Model, including lags and dummies, were resorted to. Results identified cointegration equations denoting in turn the existence of long run equilibrium relations among variables; in this connection, variables' coefficients showed the expected signs and were, in all cases, significantly different from zero.

Econometric estimates also exceeded the usual tests for specific problems. Traces of autocorrelation found in some of estimations remains as a point to be dealt with, although at this stage they did not affect results' soundness.

The Granger causality test did not indicate the expected sequence of causality between the average propensity to invest in human capital and the gross domestic product, but it did it in the opposite direction; that is, a change in human capital investment measured as public expenditure on education plus the opportunity cost, did not necessarily cause Argentine GDP to experience – in contemporaneous or subsequent periods – variations of the same sign.

Econometric results showing that per capita gross domestic product caused average propensity to invest in human capital, but not the other way round as suggested by empirical findings of the Augmented Solow Model, had also been sufficiently backed by the evidence yielded by stylized facts, which showed that in Argentina (and particularly as of 2003 when the Financiamiento Educativo law was enacted) investment in education was practically a function of income.

It follows from the above that although the formation of human capital (in part represented here by Expenditures in Education) grew substantially during the study period, there seemed not to exist a clear relationship between the characteristics and effectiveness of spending programmes and the needs of the country's productive technological matrix.

Innovation accounting tools, which include impulse response functions and variance decomposition analysis, were used in order to assess the adequacy of the model. VAR impulse response functions highlighting the response of GDP to shocks in average propensities to invest in physical and human capital appeared to be significantly different from zero, particularly in the early years following the innovations, in spite that what resulted from Granger Causality Tests.

Variance decomposition that shows the proportion of the movements in the sequence of a variable that is caused by its own shocks, versus shocks to the other variables, also yielded consistent results. Suffice it to point out here that despite different orderings imposed to the respective variables in Choleski decomposition, impulse response functions and variance decomposition yielded relatively similar results.

Bearing in mind the original objective of studying the link between human capital formation (represented here as investment in education) and economic growth, and of empirically assessing whether human capital helped enhancing the Argentine gross domestic product, it can be preliminary stated, in the light of commented results, that either it did not or it did it in a minor magnitude.

Although reasons for that were not sufficiently considered in the present study, it might be suggested that the nature, structure and design of current fiscal policies were in this field nor efficient neither efficacious to achieving human capital's greater contribution to product.

ANNEX 1

Argentine Macroeconomic Series

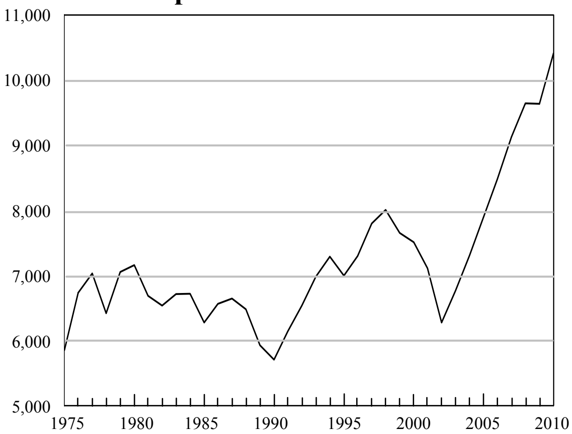
YEAR	GDPPC	NDG	SK	SH	SHOC	HOCPC
1970	7715.9264	0.1253	0.1997	0.0190	0.0470	
1971	7785.5126	0.1264	0.2105	0.0203	0.0526	
1972	7761.8169	0.1271	0.2100	0.0217	0.0526	
1973	9059.4481	0.1273	0.1641	0.0230	0.0558	
1974	8677.9784	0.1271	0.1752	0.0207	0.0693	
1975	5843.8484	0.1264	0.2564	0.0184	0.0825	
1976	6724.6234	0.1257	0.2422	0.0202	0.0453	
1977	7024.7289	0.1253	0.2761	0.0219	0.0435	
1978	6408.1012	0.1250	0.2600	0.0237	0.0412	1647.4577
1979	7045.6543	0.1250	0.2488	0.0254	0.0436	1563.9050
1980	7151.8987	0.1251	0.2544	0.0347	0.0522	1546.5807
1981	6678.3653	0.1253	0.2281	0.0304	0.0491	1491.0020
1982	6528.7571	0.1254	0.1840	0.0212	0.0414	1392.5059
1983	6709.6411	0.1254	0.1770	0.0213	0.0521	1388.3534
1984	6712.7158	0.1252	0.1671	0.0232	0.0622	1466.9411
1985	6270.0715	0.1250	0.1514	0.0353	0.0651	1540.4293
1986	6558.0232	0.1249	0.1583	0.0379	0.0657	1623.7858
1987	6636.7852	0.1249	0.1741	0.0340	0.0647	1692.6675
1988	6469.4241	0.1247	0.1698	0.0301	0.0553	1680.7746
1989	5921.0672	0.1244	0.1434	0.0171	0.0378	1544.7039
1990	5695.0102	0.1240	0.1222	0.0257	0.0349	1402.9336
1991	6130.8974	0.1237	0.1456	0.0341	0.0508	1377.9094
1992	6529.2661	0.1235	0.1789	0.0347	0.0476	1349.4917
1993	6972.9608	0.1232	0.1906	0.0366	0.0532	1379.4880
1994	7286.3332	0.1228	0.2047	0.0368	0.0581	1460.0144
1995	6992.3066	0.1224	0.1831	0.0389	0.0615	1543.1614
1996	7291.4349	0.1220	0.1889	0.0371	0.0589	1623.7531
1997	7792.3407	0.1216	0.2056	0.0385	0.0595	1733.5769
1998	8002.2277	0.1213	0.2110	0.0384	0.0626	1864.8330
1999	7647.7994	0.1209	0.1908	0.0428	0.0700	1999.7494
2000	7507.9856	0.1206	0.1792	0.0442	0.0735	2116.5662
2001	7105.0514	0.1201	0.1581	0.0465	0.0786	2216.1276
2002	6270.3354	0.1197	0.1128	0.0363	0.0660	2161.5160
2003	6760.6361	0.1194	0.1432	0.0360	0.0645	2124.0208
2004	7302.3836	0.1194	0.1765	0.0377	0.0790	2212.4454
2005	7897.0342	0.1196	0.1984	0.0417	0.0919	2418.6531
2006	8481.6154	0.1198	0.2161	0.0465	0.0995	2700.1878
2007	9126.0902	0.1199	0.2260	0.0478	0.1009	3010.3028
2008	9647.4593	0.1199	0.2309	0.0507	0.1018	3332.1291
2009	9635.2292	0.1198	0.2057	0.0527	0.1085	3663.9004
2010	10418.0894	0.1196	0.1822	0.0606	0.1114	4049.2467

Sources:
Gross Domestic Product: National Institute of Statistics and Censuses.
Gross Investment in Physical Capital: ECLAC STATS, Argentine Direction of National Accounts.
Consolidated budgetary educational expenditure and spending in science and technology: Direction for the Analysis of Public Spending and Social Programmes, Ministry of Economy of Argentina.
Consumer Price Index: National Institute of Statistics and Censuses and Statistics Direction, Province of San Luis, Argentina.
Population: National Institute of Statistics and Censuses.
Working Age Population: ECLAC, ILO.
Legal Minimum Wage: Ministry of Labour and Social Security.
Population enrolled in primary and secondary school and in universities: UNESCO.
Percentage of working age population over population in school age (secondary level): Argentine National Censuses.
University students having (not having) jobs and defraying (not defraying) their career and maintenance costs: Permanent Household Survey and information provided by the National Universities of Córdoba and La Plata.

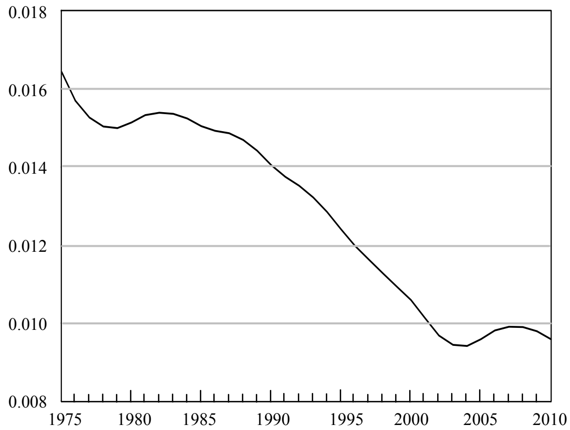
ANNEX 2

Argentina – Graphs in Levels of Macroeconomics Series

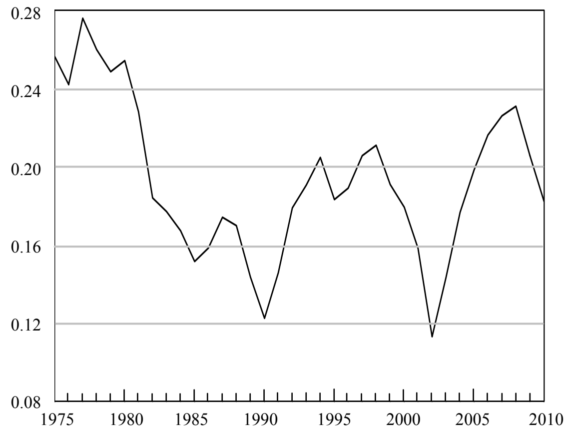
Per Capita Gross Domestic Product



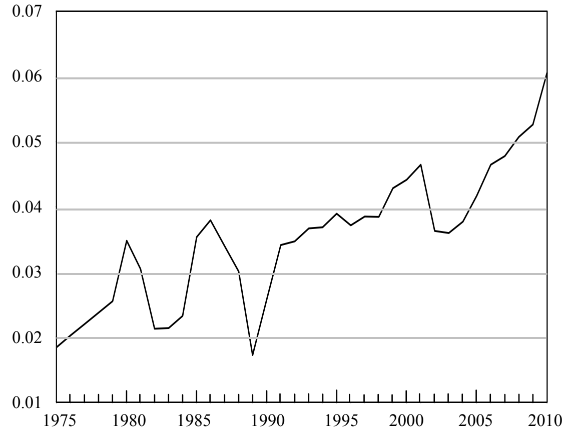
NDG



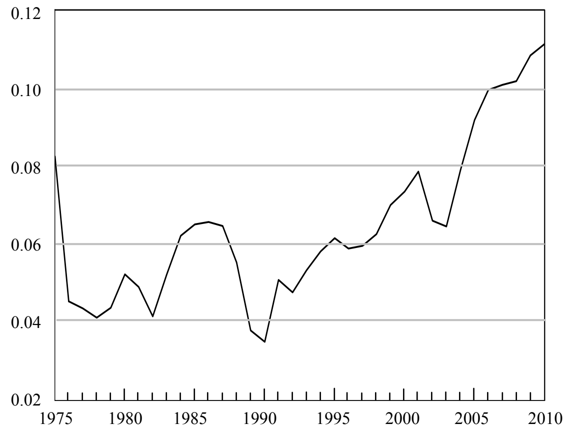
Average Propensity to Invest in Physical Capital



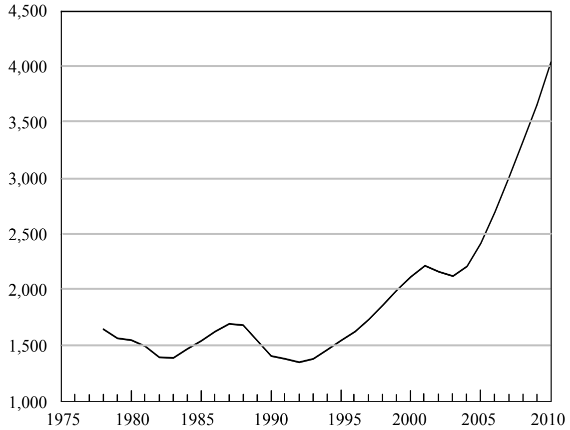
Average Propensity to Invest in Human Capital
(only public investment)



Average Propensity to Invest in Human Capital
(inclusive of opportunity costs)



Per Capita Human Capital Stock
(inclusive of opportunity costs)



ANNEX 3
ECONOMETRIC TESTS

Table 7

Unit root test for Per Capita Human Capital Stock
Augmented Dickey-Fuller test for Per Capital Human Capital Stock
(inclusive of opportunity cost)

Null Hypothesis: HOCPC has a unit root
Exogenous: Constant
Lag Length: 2 (Automatic – based on SIC, maxlag=8)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		2.489400	1.0000
Test critical values:	1% level	–3.670170	
	5% level	–2.963972	
	10% level	–2.621007	

* MacKinnon (1996) one-sided *p*-values.

Augmented Dickey-Fuller test for First Difference of Per Capital Human Capital Stock
(inclusive of opportunity cost)

Null Hypothesis: D(HOCPC) has a unit root
Exogenous: None
Lag Length: 1 (Automatic – based on SIC, maxlag=8)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		–0.688838	0.4100
Test critical values:	1% level	–2.644302	
	5% level	–1.952473	
	10% level	–1.610211	

* MacKinnon (1996) one-sided *p*-values.

Augmented Dickey-Fuller test for First Second of Per Capital Human Capital Stock
(inclusive of opportunity cost)

Null Hypothesis: D(HOCPC,2) has a unit root
Exogenous: None
Lag Length: 0 (Automatic – based on SIC, maxlag=8)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		–3.724724	0.0005
Test critical values:	1% level	–2.644302	
	5% level	–1.952473	
	10% level	–1.610211	

* MacKinnon (1996) one-sided *p*-values.

Table 8

Unit Root Test for Cointegration Equation Residuals
Augmented Dickey-Fuller Test for CEq residuals

Null Hypothesis: CE has a unit root
Exogenous: None
Lag Length: 1 (Automatic – based on SIC, maxlag=9)

		***t*-Statistic**	**Prob.***
Augmented Dickey-Fuller test statistic		–2.710985	0.0082
Test critical values:	1% level	–2.634731	
	5% level	–1.951000	
	10% level	–1.610907	

* MacKinnon (1996) one-sided *p*-values.

Phillips-Perron Test for CEq residuals

Null Hypothesis: CE has a unit root
Exogenous: None
Bandwidth: 3 (Newey-West automatic) using Bartlett kernel

		Adj. *t*-Stat.	**Prob.***
Phillips-Perron test statistic		–7.881003	0.0000
Test critical values:	1% level	–2.632688	
	5% level	–1.950687	
	10% level	–1.611059	

* MacKinnon (1996) one-sided *p*-values.

Residual variance (no correction)	0.003609
HAC corrected variance (Bartlett kernel)	0.007622

Table 9

ECM Tests

Error Correction	D(LOG(GDPPC))	D(LOG(NDG))	D(LOG(SK))	D(LOG(SHOC))
R^2	0.620383	0.840351	0.324847	0.510517
Adj. R^2	0.541842	0.807320	0.185161	0.409244
Sum sq. resids	0.113939	0.001847	0.478804	0.497416
S.E. equation	0.062681	0.007981	0.128493	0.130967
F-statistic	7.898804	25.44145	2.325541	5.041027
Log likelihood	52.51918	126.7142	26.67790	25.99146
Akaike AIC	–2.528843	–6.650789	–1.093217	–1.055081
Schwarz SC	–2.220937	–6.342882	–0.785310	–0.747175
Mean dependent	0.005077	–0.016042	0.001096	0.013190
S.D. dependent	0.092604	0.018182	0.142346	0.170395
Determinant resid covariance (dof adj.)		3.50E-11		
Determinant resid covariance		1.47E-11		
Log likelihood		244.6247		
Akaike information criterion		–11.81249		
Schwarz criterion		–10.40491		

Table 10

Granger Causality Test

VEC Granger Causality/Block Exogeneity WaldTests
Date: 08/07/12 Time: 10:08
Sample: 1975 2010
Included observations: 36

Dependent variable: D(GDPPC)

Excluded	Chi-sq.	Df	Prob.
D(NDG)	0.039624	1	0.8422
D(SK)	3.230270	1	0.0723
D(SHOC)	0.168026	1	0.6819
All	3.601399	3	0.3078

Dependent variable: D(NDG)

Excluded	Chi-sq.	Df	Prob.
D(GDPPC)	15.96479	1	0.0001
D(SK)	3.704779	1	0.0543
D(SHOC)	0.118192	1	0.7310
All	17.88140	3	0.0005

Dependent variable: D(SK)

Excluded	Chi-sq.	Df	Prob.
D(GDPPC)	0.713625	1	0.3982
D(NDG)	3.141984	1	0.0763
D(SHOC)	0.222395	1	0.6372
All	3.313999	3	0.3457

Dependent variable: D(SHOC)

Excluded	Chi-sq.	Df	Prob.
D(GDPPC)	12.48611	1	0.0004
D(NDG)	0.476580	1	0.4900
D(SK)	4.269784	1	0.0388
All	13.63916	3	0.0034

Table 11

Normality Test

VEC Residual Normality Tests
Orthogonalization: Cholesky (Lutkepohl)
Null Hypothesis: residuals are multivariate normal
Date: 08/15/12 Time: 14:47
Sample: 1975 2010
Included observations: 36

Component	Skewness	Chi-sq.	df	Prob.
1	0.046288	0.012856	1	0.9097
2	–0.209914	0.264384	1	0.6071
3	0.397369	0.947411	1	0.3304
4	–0.202659	0.246425	1	0.6196
Joint		1.471075	4	0.8318

Component	Kurtosis	Chi-sq.	df	Prob.
1	2.942689	0.004927	1	0.9440
2	3.165719	0.041194	1	0.8392
3	2.975202	0.000922	1	0.9758
4	3.010935	0.000179	1	0.9893
Joint		0.047223	4	0.9997

Component	Jarque-Bera	Df	Prob.
1	0.017783	2	0.9911
2	0.305579	2	0.8583
3	0.948333	2	0.6224
4	0.246604	2	0.8840
Joint	1.518298	8	0.9924

Table 12

Heteroskedasticity Test

VEC Residual Heteroskedasticity Tests: Includes Cross Terms
Date: 08/15/12 Time: 14:48
Sample: 1975 2010
Included observations: 36

Joint test:					
Chi-sq	Df	Prob.			
221.1420	210	0.2854			
Individual components:					
Dependent	R^2	$F(21,14)$	**Prob.**	**Chi-sq.(21)**	**Prob.**
res1*res1	0.386844	0.420604	0.9645	13.92637	0.8727
res2*res2	0.506948	0.685455	0.7890	18.25012	0.6331
res3*res3	0.798639	2.644144	0.0330	28.75102	0.1201
res4*res4	0.460111	0.568154	0.8828	16.56399	0.7372
res2*res1	0.392246	0.430269	0.9608	14.12086	0.8644
res3*res1	0.689491	1.480342	0.2275	24.82166	0.2550
res3*res2	0.594404	0.977004	0.5318	21.39853	0.4348
res4*res1	0.472714	0.597669	0.8610	17.01770	0.7100
res4*res2	0.673155	1.373037	0.2744	24.23358	0.2819
res4*res3	0.568793	0.879383	0.6154	20.47656	0.4913

Table 13

Autocorrelation Test

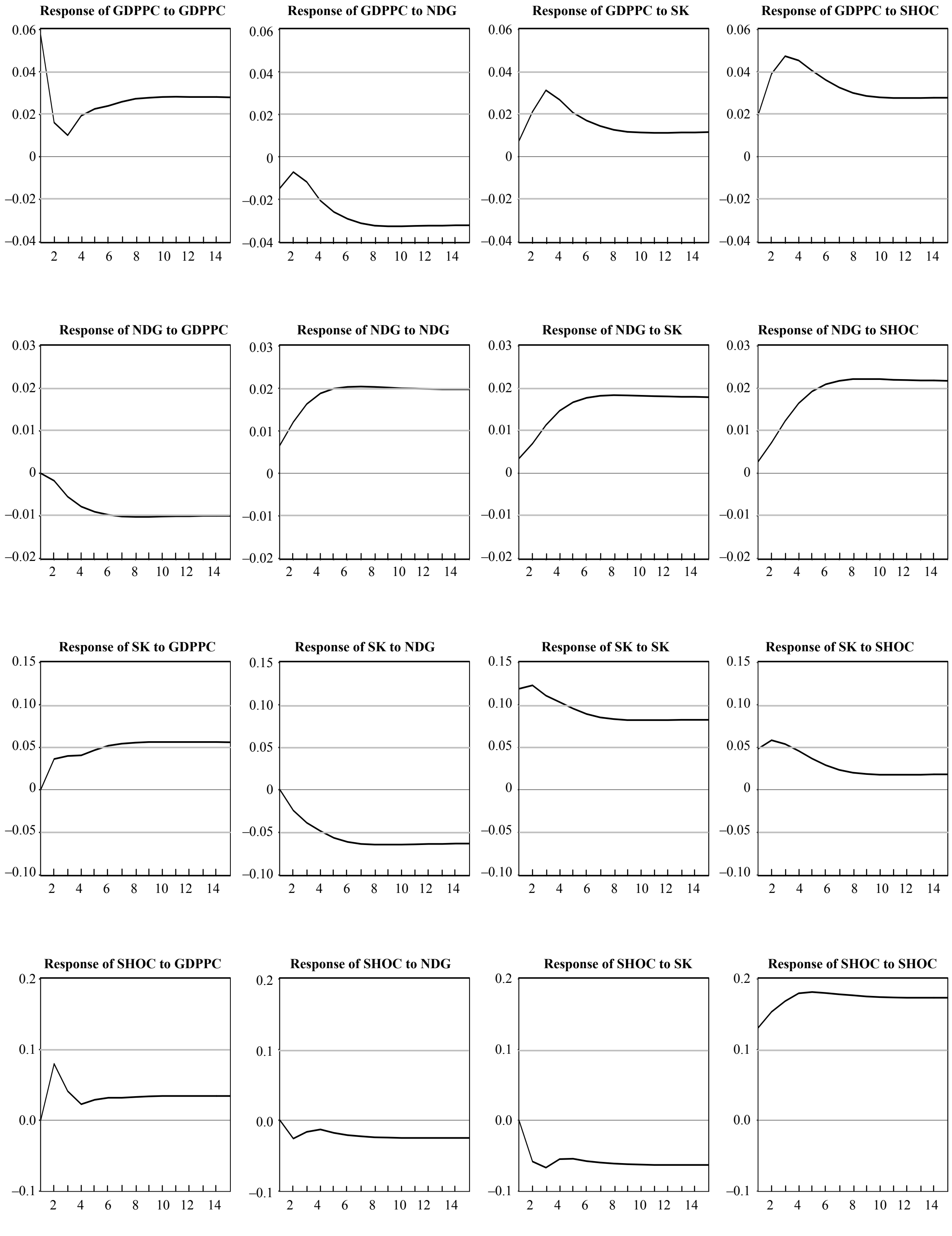
VEC Residual Serial Correlation LM Tests
Null Hypothesis: no serial correlation at lag order h
Date: 08/15/12 Time: 14:50
Sample: 1975 2010
Included observations: 36

Lags	**LM-Stat.**	**Prob.**
1	51.97732	0.0000
2	25.25147	0.0655
3	22.28609	0.1342
4	12.98071	0.6742
5	14.12149	0.5897
6	28.47514	0.0277
7	24.30961	0.0830
8	17.28722	0.3673

Probs from chi-square with 16 df.

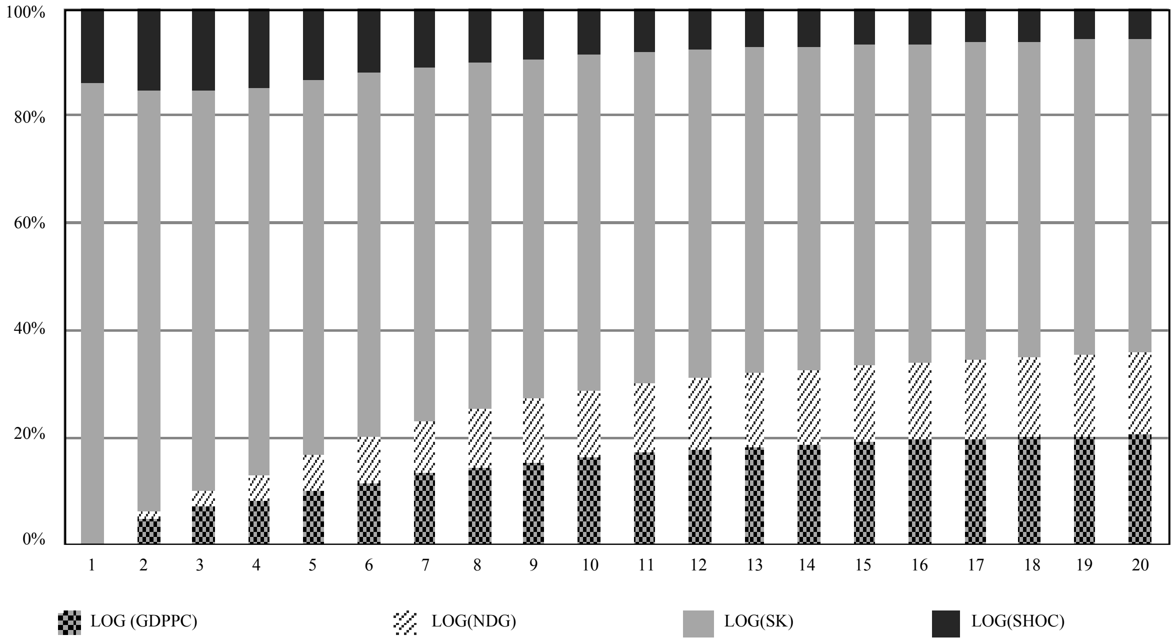
ANNEX 4
IMPULSE RESPONSE FUNCTION AND VARIANCE DECOMPOSITION

Impulse Response Functions
Response to Cholesky One S.D. Innovations



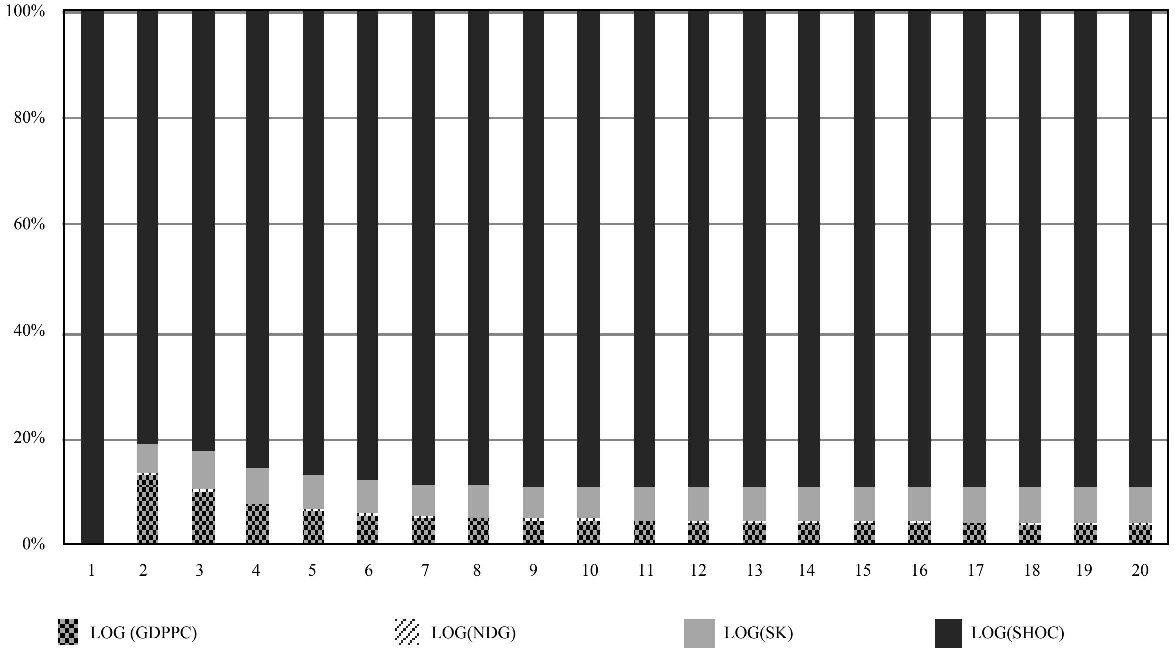
Variance Decomposition of Average Propensity to Invest in Physical Capital

SK



Variance Decomposition of Average Propensity to Invest in Human Capital
(inclusive of opportunity costs)

SHOC



REFERENCES

Arrow, K.J. (1962), "The Economic Implications of Learning by Doing", *Review of Economic Studies*, No. 29, pp. 155-73.

Enders, W. (1995), *Applied Econometric Series*, John Wiley and Sons, Inc. (USA).

Frenkel, J.A. and A. Razin (1996), *Fiscal Policies and Growth in the World Economy*, The MIT Press, Cambridge (MA).

Heijdra, A.J. and F. van der Ploeg (2002), *The Foundations of Modern Macroeconomics*, Oxford University Press (U.K.).

Lucas, R.E. Jr. (1988), "On the Mechanics of Economic Development", *Journal of Monetary Economics*, No. 22, pp. 3-42.

Mankiw, N.G., D. Romer and D.N. Weil (1992), "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics*, Vol. 107, No. 2, pp. 407-37.

Rezk, E. and M. Irace (2008), "Efectos de los shocks fiscales y monetarios en Argentina: Un enfoque de cointegración", paper presented at the *43rd Annual Meeting of the Argentine Association of Political Economy*, Córdoba (AR), November.

Romer, P.M. (1990), "Endogenous Technological Change", *Journal of Political Economy*, Vol. 98, No. 5-II.

Rosen, S. (1976), "A Theory of Life Earnings", *Journal of Political Economy*, No. 84, pp. 545-67.

Sala-i-Martín, X. (1994), *Apuntes de Crecimiento Económico*, Antoni Bosch, Barcelona.

Schultz, T.W. (1961), "Investment in Human Capital", *American Economic Review*, No. 51, pp. 5-6, March.

Solow, R.M. (1956), "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, Vol. 70, No. 1, pp. 65-94.

Swan, T.M. (1956), "Economic Growth and Capital Accumulation", *Economic Record*, No. 66, pp. 344-61.

Uzawa, H. (1965), "Optimal Technical Change in an Aggregative Model of Economic Growth", *International Economic Review*, No. 6, pp. 18-31.

PUBLIC DEBT AND ECONOMIC GROWTH:
A QUICK LOOK AT THRESHOLD EFFECTS

Balázs Égert[*]

1 Introduction

The 2007-08 financial and economic crisis principally caused by the collapse of the US subprime market triggered economic recession in many countries. Governments and central banks of the developed world swiftly reacted by implementing substantial fiscal and monetary policy easing, coupled with State aid to the troubled financial sector. These actions no doubt helped contain the Great Recession but pro-cyclical discretionary fiscal expansion and the banking sector bail-outs led to an unprecedented rise in public debt-to-GDP ratios. Against this backdrop, Reinhart and Rogoff (2010) argued that an excessively high public debt (as a share of GDP) hampers economic activity. On the basis of descriptive statistics, they showed that there was a tipping point at 90 per cent of GDP: economic growth slows down sharply if the debt-to-GDP ratio exceeds 90 per cent of GDP. A number of recent papers investigated this issue and used more advanced statistical methods to analyse the non-linear negative relation between growth and public debt. Indeed, Cecchetti *et al.* (2011) find a threshold of about 85 per cent of GDP. Kumar and Woo (2010), Checherita and Rother (2010) and Baum *et al.* (2012) confirm the 90 per cent threshold.

The ambition of this note is to take a quick look at how robust the 90 per cent threshold is. In doing so, we use a subset of a variant of the Reinhart-Rogoff dataset. We estimate the bivariate relationship between growth and debt (and lagged debt) in a two-regime threshold model for a variety of thresholds. We also perform a robustness check of the 90 per cent threshold by jackknifing the sample, *i.e.*, dropping one country from the sample at a time. We find that the threshold may be different from 90 per cent, that it varies a lot whether we use contemporaneous or lagged debt and that the negative impact of debt on growth is sensitive to outlier observations.

2 Data and estimation issues

The main evidence in Reinhart and Rogoff (2010) is based on a sample of 20 industrialised countries for the period from 1946 to 2009. For this reason, we use in this note this subset of the Reinhart and Rogoff dataset. Reinhart and Rogoff (2010) do not give the sources of the data they use in their paper. But data on central government debt can be obtained from the data appendix of Reinhart and Rogoff (2011). Real GDP growth rates are available for a number of countries for the same time period from the Barro-Ursúa macroeconomic dataset (Barro and Ursúa, 2011). Matching these two datasets helps us reproduce the Reinhart and Rogoff dataset. The difference between their data and our dataset is that our data does not include Ireland but contains data for Switzerland. A marginal difference is that our dataset ends in 2010, while the data used in Reinhart and Rogoff (2010) stops in 2009. Table 1 below gives the differences.

Our estimation approach involves two steps. First, we estimate the linear bivariate relation between growth and debt (equation 1) and then go on to estimate threshold models (equation 2) with tipping points at 10, 15, 20, …, 90, 95, 100 per cent, …, 180 per cent of GDP).

$$\Delta y_t = \alpha + \beta \, debt_t + \varepsilon_t \tag{1}$$

[*] OECD Economics Department. E-mail: balazs.egert@oecd.org
The usual disclaimer applies.

Table 1

Data Coverage: Reinhart and Rogoff (2010) Versus the Dataset Used in the Paper

Country	Reinhart and Rogoff (2010)	Our Dataset, Which Uses Data from Reinhart and Rogoff (2011) for the Level of Central Government Debt and Barro and Ursúa (2012) for Real GDP Growth
Australia	1902-2009	1861-2009
Austria	1880-2009	1880-2009
Belgium	1835-2009	1847-2009
Canada	1925-2009	1871-2009
Denmark	1880-2009	1880-2009
Finland	1913-2009	1914-2009
France	1880-2009	1880-2009
Germany	1880-2009	1880-2009
Greece	1884-2009	1848-2009
Ireland	1949-2009	-
Italy	1880-2009	1862-2009
Japan	1885-2009	1872-2009
Netherlands	1880-2009	1814-2009
New Zealand	1932-2009	1831-2009
Norway	1880-2009	1880-2009
Portugal	1851-2009	1851-2009
Spain	1850-2009	1850-2009
Sweden	1880-2009	1801-2009
Switzerland	-	1880-2009
United Kingdom	1830-2009	1831-2009
USA	1790-2009	1791-2009

$$\Delta y_t = \begin{cases} \alpha_1 + \beta_1 \cdot debt_t + \varepsilon_t & if \quad debt < T \\ \alpha_2 + \beta_2 \cdot debt_t + \varepsilon_t & if \quad debt \geq T \end{cases} \tag{2}$$

where Δy is annual real GDP growth, *debt* stands for the central government debt-to-GDP ratio and T is the value of the debt threshold (10, 15, 20, …, 90, 95, 100, …, 180 per cent of GDP). Equations (1) and (2) are estimated for a pooled panel and with country fixed effects and for contemporaneous and lagged debt. Finally, equation (2) for the debt threshold equalling 90 per cent is jackknifed: one country is dropped from the sample at a time.

3 Estimation results

Linear bivariate panel regressions show a negative link between growth and public debt but this effect does not seem to be statistically significant (Table 2). When imposing a threshold of 90 per cent of public debt, the estimation results show that the contemporaneous relation between growth and debt is strongly negative if public debt is lower than 90 per cent of GDP, whereas the relation breaks down above that threshold. Carrying out the estimations using alternative threshold values (from 10 to 180 per cent of GDP by steps of 5 per cent) does not change this picture: the coefficient estimates are never statistically significant in the upper regime (in which observed debt is above the debt threshold). In addition, in the range from 25 to 55 per cent of GDP, the coefficient estimates are not different from zero in any of the two regimes. Let us now pick the model from the many estimated models, which seems to fit best the underlying data. The models, which minimise the Schwarz and Akaike information criteria and for which the adjusted R-squared is the highest, are the ones with threshold values of 170 and 175 per cent of GDP. These results basically imply an almost linear relationship given that most observations for public debt are below these thresholds.

To check the robustness of the results, we re-estimated the same models using lagged public debt as a right-hand side variable (Table 2). The results are markedly different. First, for the 90 per cent threshold, the coefficient estimates are not only negative but also statistically significant in both regimes, even though they are very similar in size. Second, for turning points higher than 135 per cent of GDP, the coefficient estimate in the upper regime becomes insignificant. Finally, the threshold is at 20 per cent of GDP for the model for which the information criteria are the lowest and the adjusted R-squared the highest. This is quite different from the 170-175 per cent threshold finding. In addition, there is a positive relation between debt and growth below 20 per cent and it becomes negative only above this threshold.

In a second step, we jackknife the sample for the model with a 90 per cent debt threshold. Table 3 shows the sensitivity of the results to specific countries. In particular, if the Netherlands is taken out from the sample, the coefficient on contemporaneous debt becomes negative and statistically significant in the upper regime, *i.e.* when public debt exceeds 90 per cent of GDP. When lagged debt is used as a right-hand side variable, the results are more robust in terms of statistical significance. In all cases, the coefficients remain negative and significant in both regimes. Nevertheless, it is worth mentioning that the variability in the size of the coefficient estimates (measured by the range between the lowest and highest coefficient estimate) is considerable higher in the upper regime than in the lower regime.

4 Conclusions

The ambition of this note was to provide a quick robustness check with regard to the 90 per cent threshold. Using a subset of a variant of the Reinhart-Rogoff dataset including industrialised countries for 1946 to 2010, we found that the non-linear effect linking growth and public debt is not particularly robust. First, whether there is a strong negative link between growth and debt above 90 per cent and how large it is depends on model specification and the inclusion of specific countries in the sample. Second, a simple model selection shows that the 90 per cent threshold may be considerably lower or higher, depending again on model specification.

Table 2

Estimation Results for Alternative Thresholds, 1946-2010

	Debt				Lagged Debt			
	Pooled Panel		Country Fixed Effects		Pooled Panel		Country Fixed Effects	
Linear model: $\Delta y_t = \alpha + \beta\, debt_t + \varepsilon_t$								
β	–0.007		–0.010		–0.007		–0.010	
Threshold Model: $\Delta y_t = \begin{cases} \alpha_1 + \beta_1 \cdot debt_t + \varepsilon_t & if \quad debt < T \\ \alpha_2 + \beta_2 \cdot debt_t + \varepsilon_t & if \quad debt \geq T \end{cases}$								
T	β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2
10% of GDP	0.142**	–0.004	0.132*	–0.007	0.094**	–0.011**	0.086*	–0.013**
15% of GDP	0.084*	–0.002	0.067	–0.007	0.045	–0.010**	0.032	–0.012**
20% of GDP	0.083**	0.001	0.075*	–0.004	**0.064****	**–0.006***	**0.058****	**–0.009****
25% of GDP	0.050	–0.001	0.047	–0.005	0.028*	–0.008**	0.026*	–0.011**
30% of GDP	0.028	–0.002	0.031	–0.006	0.003	–0.011**	0.004	–0.012**
35% of GDP	0.014	–0.004	0.015	–0.007	–0.014	–0.013**	–0.017	–0.015**
40% of GDP	0.004	–0.005	0.005	–0.007	–0.014	–0.013**	–0.016	–0.014**
45% of GDP	–0.004	–0.006	0.000	–0.008	–0.019**	–0.014**	–0.019*	–0.015**
50% of GDP	–0.009	–0.007	–0.007	–0.009	–0.024**	–0.014**	–0.024**	–0.016**
55% of GDP	–0.012	–0.007	–0.014	–0.010	–0.017**	–0.013**	–0.019**	–0.015**
60% of GDP	–0.019*	–0.008	–0.021*	–0.011	–0.024**	–0.014**	–0.026**	–0.016**
65% of GDP	–0.025**	–0.008	–0.026**	–0.012	–0.027**	–0.014**	–0.027**	–0.016**
70% of GDP	–0.025**	–0.008	–0.027**	–0.011	–0.027**	–0.013**	–0.027**	–0.015**
75% of GDP	–0.023**	–0.007	–0.024**	–0.010	–0.027**	–0.013**	–0.026**	–0.015**
80% of GDP	–0.022**	–0.006	–0.024**	–0.010	–0.019**	–0.012**	–0.017**	–0.014**
85% of GDP	–0.024**	–0.005	–0.026**	–0.010	–0.019**	–0.012**	–0.017**	–0.014**
90% of GDP	–0.023**	–0.005	–0.026**	–0.009	–0.020**	–0.012**	–0.019**	–0.014**
95% of GDP	–0.023**	–0.004	–0.027**	–0.008	–0.021**	–0.011**	–0.020**	–0.014**
100% of GDP	–0.023**	–0.003	–0.029**	–0.007	–0.020**	–0.011**	–0.021**	–0.014**
105% of GDP	–0.023**	–0.001	–0.030**	–0.006	–0.020**	–0.010**	–0.022**	–0.013**
110% of GDP	–0.020**	0.002	–0.028**	–0.002	–0.018**	–0.010**	–0.022**	–0.011**
115% of GDP	–0.020**	0.003	–0.029**	0.000	–0.016**	–0.010**	–0.020**	–0.012**
120% of GDP	–0.015**	0.002	–0.024**	0.000	–0.016**	–0.010**	–0.020**	–0.011**
125% of GDP	–0.017**	0.005	–0.026**	0.003	–0.017**	–0.009*	–0.021**	–0.009**
130% of GDP	–0.017**	0.007	–0.026**	0.005	–0.017**	–0.008*	–0.021**	–0.009*
135% of GDP	–0.018**	0.009	–0.027**	0.007	–0.017**	–0.007	–0.021**	–0.009*
140% of GDP	–0.018**	0.011	–0.027**	0.009	–0.017**	–0.007	–0.021**	–0.008
145% of GDP	–0.018**	0.011	–0.027**	0.009	–0.017**	–0.007	–0.021**	–0.008
150% of GDP	–0.018**	0.012	–0.025**	0.010	–0.017**	–0.006	–0.020**	–0.008
155% of GDP	–0.017**	0.013	–0.025**	0.011	–0.016**	–0.007	–0.019**	–0.008
160% of GDP	–0.017**	0.013	–0.025**	0.011	–0.016**	–0.007	–0.019**	–0.008
165% of GDP	–0.016**	0.018	–0.023**	0.017	–0.015**	–0.006	–0.019**	–0.006
170% of GDP	**–0.016****	**0.020**	**–0.024****	**0.020**	–0.016**	–0.004	–0.020**	–0.003
175% of GDP	**–0.016****	**0.020**	**–0.024****	**0.020**	–0.016**	–0.004	–0.020**	–0.003
180% of GDP	–0.015**	0.024	–0.022**	0.024	–0.015**	–0.004	–0.018**	–0.004

* and ** denote statistical significance at the 10 and 5 per cent levels, respectively. Shaded cells indicate the models which minimise the Schwarz and Akaike information criteria and for which the adjusted R^2 are the highest.

Table 3

Estimation Results for the Jackknifed Sample (Debt Threshold-90% of GDP), 1946-2010

	Debt								Lagged Debt							
	Pooled Panel				Country Fixed Effects				Pooled Panel				Country Fixed Effects			
Threshold Model: $\Delta y_t = \begin{cases} \alpha_1 + \beta_1 \cdot debt_t + \varepsilon_t & if \quad debt < T \\ \alpha_2 + \beta_2 \cdot debt_t + \varepsilon_t & if \quad debt \geq T \end{cases}$, T=90%																
Country Excluded	β_1		β_2		β_1		β_2		β_1		β_2		β_1		β_2	
AUS	–0.025	**	–0.005		–0.028	**	–0.010		–0.025	**	–0.011	**	–0.026	**	–0.015	**
AUT	–0.020	**	–0.003		–0.023	**	–0.008		–0.020	**	–0.008	**	–0.020	**	–0.012	**
BEL	–0.024	**	–0.007		–0.026	**	–0.010		–0.024	**	–0.009	**	–0.024	**	–0.012	**
CAN	–0.022	**	–0.004		–0.027	**	–0.009		–0.022	**	–0.009	**	–0.024	**	–0.013	**
DNK	–0.022	**	–0.005		–0.026	**	–0.009		–0.023	**	–0.010	**	–0.024	**	–0.013	**
FIN	–0.023	**	–0.005		–0.027	**	–0.009		–0.024	**	–0.010	**	–0.025	**	–0.013	**
FRA	–0.021	**	–0.004		–0.024	**	–0.008		–0.021	**	–0.009	**	–0.021	**	–0.012	**
DEU	–0.022	**	–0.004		–0.024	**	–0.008		–0.022	**	–0.009	**	–0.021	**	–0.012	**
GRC	–0.021	**	–0.003		–0.023	**	–0.007		–0.020	**	–0.008	**	–0.021	**	–0.011	**
ITA	–0.024	**	–0.002		–0.026	**	–0.005		–0.024	**	–0.008	**	–0.024	**	–0.010	**
JPN	–0.017	**	0.000		–0.018	**	–0.003		–0.017	**	–0.006	*	–0.015	**	–0.007	*
NLD	–0.028	**	–0.017	**	–0.030	**	–0.024	**	–0.025	**	–0.014	**	–0.024	**	–0.017	**
NZL	–0.023	**	–0.003		–0.029	**	–0.008		–0.021	**	–0.009	**	–0.023	**	–0.013	**
NOR	–0.024	**	–0.005		–0.028	**	–0.010		–0.023	**	–0.010	**	–0.025	**	–0.013	**
PRT	–0.020	**	–0.004		–0.023	**	–0.008		–0.020	**	–0.009	**	–0.020	**	–0.012	**
ESP	–0.022	**	–0.004		–0.026	**	–0.009		–0.022	**	–0.009	**	–0.023	**	–0.013	**
SWE	–0.023	**	–0.005		–0.027	**	–0.009		–0.023	**	–0.010	**	–0.024	**	–0.013	**
GBR	–0.022	**	–0.002		–0.026	**	–0.009		–0.022	**	–0.010	**	–0.024	**	–0.015	**
USA	–0.023	**	–0.003		–0.028	**	–0.008		–0.022	**	–0.009	**	–0.023	**	–0.012	**
CHE	–0.027	**	–0.006		–0.029	**	–0.010		–0.026	**	–0.011	**	–0.024	**	–0.013	**
MIN	**–0.028**		**–0.017**		**–0.030**		**–0.024**		**–0.026**		**–0.014**		**–0.026**		**–0.017**	
MAX	**–0.017**		**0.000**		**–0.018**		**–0.003**		**–0.017**		**–0.006**		**–0.015**		**–0.007**	

* and ** denote statistical significance at the 10 and 5 per cent levels, respectively.

REFERENCES

Barro, R.J. and J.F. Ursúa (2012), "Barro Ursúa Macroeconomic Dataset", Harvard University.

Baum, A., C. Checherita-Westphal and P. Rother (2012), "Debt and Growth: New Evidence for the Euro Area", ECB, mimeo.

Cecchetti, S., M. Mohanty and F. Zampolli (2011), "The Real Effects Of Debt", Bank for International Settlements, Working Paper, No. 352.

Checherita, C. and P. Rother (2010), "The Impact of High and Growing Government Debt on Economic Growth: An Empirical Investigation for the Euro Area", European Central Bank, Working Paper, No. 1237.

Kumar, M.S. and J. Woo (2010), "Public Debt and Growth", IMF, Working Paper, No. 10/174.

Reinhart, C.M. and K.S. Rogoff (2010), "Growth in a Time of debt", *American Economic Review*, Vol. 100, No. 2, pp. 573-78.

―――――(2011), "From Financial Crash to Debt Crisis", *American Economic Review*, Vol. 101, No. 5, pp. 1676-706.

FISCAL POLICIES ENHANCING GROWTH IN EUROPE: CAN WE APPLY COMMON REMEDIES TO DIFFERENT COUNTRIES?

Carine Bouthevillain[*] *and Gilles Dufrénot*[**]

We provide evidence of heterogeneous reactions of the growth rates in the European Union countries to changes in taxes and public expenditure, when the governments' budget constraint is taken into account. Direct taxation exerts a much more damaging effect on the growth rate of the European emerging countries than on the most industrialized countries'. Indirect taxes are not inconsistent with growth in the latter, while they are harmful in the former. Increases in human capital expenditure stimulate growth in the low-growth countries, while welfare and sovereign spending are efficient for growth in the economies that grow rapidly.

1 Introduction

There is a widespread view in the European economic policy circles that in order to get out of the current economic depression while respecting at the same time the sustainability of their public finances, the European Union countries should implement common fiscal policies. Some people even suggest a fiscal federalism, by comparison with the United States, where a federal budget can be operated to conduct countercyclical policies. This paper argues against such a view. We provide evidence of great heterogeneities among the EU countries regarding the fiscal/growth relationship. We conclude that similar policies cannot work in a similar way in countries that are still experiencing a catch-up dynamics and which experienced the highest growth rates over the last 10 years (the most recent members of the Central and Eastern Europe and some countries such as Spain, Portugal and Ireland) and in countries whose growth rates have been lower (the older members). This conclusion is valid whether or not we consider demographic differences between the countries. For example, welfare and social spending, usually considered in the literature as non-productive, stimulate growth in countries with fast growth, but are harmful on the growth rates of low-growth countries. Tax cuts have stronger positive effects on the growth rate of the emerging economies than on the growth rate of the most industrialized countries. Increases in social security contributions inhibit the growth rates of the low-growth countries, but stimulate the growth rate of countries that are growing fast.

Such differences would not necessarily appear if we tried to link growth to fiscal variables by using aggregate indicators of spending and taxation, for instance, the ratio of total spending out of GDP, or the ratio of total taxes over GDP. Differences among the countries appear once we consider that the European countries face the joint problem of attempting to boost growth while simultaneously tracking a sustainable level of their public finance. This double choice is motivated by the fact that the fiscal policies are coordinated through the implementation of a Stability and Growth Pact which imposes restrictions on the levels of debt and deficits. Therefore, the fiscal/growth link rests on the structure of spending and taxes. For instance, suppose that governments decide to increase welfare spending to conduct countercyclical policies. To guard control on the sustainability of public deficits, they may decide to finance this increase by higher direct or indirect taxes, or to offset the increase in welfare expenditure by a decrease in other spending. In this case, the impact on growth will be different, as compared with a situation in

[*] Banque de France, DGEI-DERIE, Service des relations européennes (49-1487), 31 rue Croix des Petits Champs, 75001, Paris. Tel: +33 1 42 92 42 92, E-mail: carine.bouthevillain@banque-france.fr

[**] Banque de France, CEPII and Aix-Marseille Université(AMSE) , Château Lafarge, Route des Milles, 13290 Les Milles. Tel: +33 4 42 93 59 60, E-mail: gilles.dufrenot@univ-amu.fr

which the structure of other spending and taxes are left unchanged, which would mean that higher welfare expenditure would result into a higher deficit. The argument that is put forward in this paper is that the effects of similar fiscal policies can differ across countries, because the economic growths react not only to the amounts of expenditure and revenue, but also to the structure of expenditure and revenue when a budgetary measure is adopted.

In previous papers of the literature, some authors relate the growth rate of the European countries to the structure of taxes and public spending, but they assume that the links are the same across countries. Afonso and Furceri (2010) find that a rise the following components of taxes and expenditures negatively affect growth: indirect taxes, social contributions, subsidies. An important contribution of their paper is the finding that the disaggregated components impact growth when changes occur in both their size and volatility. Nikos (2009) examines whether a reallocation of the components of public spending and revenues in 14 EU countries have enhanced their economic growth between 1990 and 2006. He concludes that government outlays on education, social protection and defense, as well as public expenditures on infrastructures, are growth-enhancing, while distorsionary[1] taxes depress growth. Furceri and Karras (2009) show that increases in social security contributions and in taxes on goods and services have had a larger negative effect on growth in the European countries between 1965 and 2003, than increases in income taxes.[2]

In contrast to these studies, we take in consideration the issue of heterogeneity. Quantile regression analysis provides a useful empirical framework within which we explore the idea of heterogeneous reactions of growth to fiscal policies in the European Union. In such a framework, it is possible to justify that the impact of changes in public spending and taxes varies across the conditional distribution of growth. This is an alternative methodology to the usual ones, either focusing on mean effects, or treating the issue of fiscal/growth heterogeneity by splitting the data into separated groups of countries.[3] Our approach differs from the previous empirical papers in two ways.

The first difference concerns the empirical methodology. Although there is a considerable literature on the fiscal policy/growth relationship, a relatively small amount of this literature is concerned with heterogeneous reactions in terms of growth to the same fiscal policies.[4] To our knowledge, three exceptions are Bassanini and Scapetta (2001), Arnold (2011) and Gemmel *et al.* (2011). These authors use the pooled mean group (PMG) and mean group (MG) estimators introduced by Pesaran *et al*. (1999). Although these estimators are useful in accounting for different slopes across the countries of a panel, the cost of using them is a reduction of the degrees of freedom. Indeed, they are based on an average of the estimates from individual countries regressions. We instead use an estimator that keeps the pooled dimension of the panel while allowing at the same time to deal with the diversity of reactions across the countries: a quantile regression estimator. One advantage is to consider the entire panel and to distinguish the countries by their location in the conditional distribution of growth. Instead of estimating models for conditional means functions, we consider a full range of other conditional quantile functions.

---

1 Following the definition given by Kneller *et al.* (1999), distortionary taxes are those which affect the investment decisions of agents (with respect to physical and / or human capital), creating tax wedges and hence distorting the steady-state rate of growth. Non-distortionary taxation does not affect the saving / investment decisions because of the assumed nature of the preference function, and hence has no effect on the rate of growth.

2 There are other examples of papers linking growth to the composition of expenditure and tax structure in other industrialized countries, among which Lee and Gordon (2005), Angelopoulos *et al*. (2007), Gemmel *et al*. (2011).

3 See, for instance Angelopoulos *et al*. (2007), Arnold (2008), Arnold and Schwellnus (2008), Bleany *et al*. (2001), Gemmel *et al*. (2011), Lee and Gordon (2005), Peretto (2003, 2007), Vartia (2008).

4 In a survey of the growth empirics, Eberhart and Teal (2011) note that the possibility of heterogeneous parameters is ignored by a vast majority of studies.

Secondly, unlike many previous studies, we do not only consider estimates of fiscal/growth regressions based on the growth rate of the GDP per capita, but also the growth rate of the real GDP itself in a context where the European governments search to avoid a rise in the burden of public debt. Our approach is motivated as follows. Fiscal policy usually has several objectives. The first is equity. Taxation and expenditure are considered in terms of their ability to impact fairly personal incomes. In this case, using the growth rate of GDP per capita (or a multidimensional welfare indicator) as the endogenous variable is convenient. Previous papers examining the impact of fiscal measures on per capita growth implicitly assume that a higher growth of the real GDP translates into a higher standard of living within and across individuals, on average (but this is an assumption that would need to be proved, since average effects mask potential changes in income distribution). A second concern of fiscal policy is efficiency. This can be defined as the way in which expenditure and taxes "deliver" in terms of the growth rate of the real GDP. For instance, finding a negative elasticity of the latter with respect to welfare expenditure, or direct taxation, might be interpreted as the existence of waste in the public sector inducing inefficiencies in the global activity (the channels yielding such inefficiencies are for instance a lower productivity of the labor supply, a higher wage reservation level, a reduced competitiveness of firms, etc). In this paper, we adopt the interpretation in terms of both efficiency and equity. We thereby consider both the growth rate of the real GDP and the growth rate of per capita GDP.

The remainder of the paper is structured as follows. Section 2 describes the theoretical underpinnings, while section 3 presents the empirical methodology and data. Section 4 contains our estimation results and our comments. Finally, Section 5 concludes.

2 Theoretical underpinnings

In this section, we briefly explain the theoretical framework underlying our empirical equations. Although, we do not present the theoretical models formally, this is important to motivate the choice of our variables as well as some restrictions imposed on some coefficients of our equations. Our empirical framework relies upon two different strands of the theoretical literature on the fiscal/growth link. One is the correlation between growth and the composition of public spending and taxes. The second concerns the effects of fiscal policy on growth with respect to how a public spending or deficit is financed.

2.1 Linking growth to the structure of taxes and the composition of expenditure

The Lisbon strategy puts an emphasis on the efficiency of fiscal policy on the European countries' growth rate. Indeed, the EU member States agreed on improving the contribution of public spending to growth by directing public expenditure towards growth-enhancing investment in both physical and human capital.[5] Besides, in a report published in 2011, the European Commission points to several challenges of tax policy, among which the potential to make the tax structure more growth friendly.[6]

These issues cannot be examined within the first-generation endogenous growth models linking fiscal policy to growth. Indeed, as pointed by Agell *et al.* (2007) and Myles (2000), when the growth effects are apprehended by considering aggregate measures of tax burden and public expenditure, these models only capture the role of government size. In the second-generation

[5] Wierts (2005) discusses some aspects of redirecting public expenditure under the Lisbon experience.

[6] The report published on October 2011 was entitled "Tax reform in EU member States 201: tax policy challenges for economic growth and fiscal sustainability".

models of endogenous growth the share of different categories of public expenditure and taxes is explicitly taken into account. A fairly simple approach consists in separating public spending between productive and non-productive components and distinguishing between distortionary and non-distortionary taxes. An important conclusion of the papers is that different spend-tax combinations yield different effects on growth. For instance, productive expenditure financed by non-distortionary taxes have a higher effect on growth than when they are financed by distortionary taxes. Another approach, widely used in the growth literature to identify the effects of fiscal policy, is to consider a fine disaggregation of public spending and taxes. On the spending side, it is usual to consider a functional disaggregation of government expenditure: spending on health, education infrastructure, defense, recreation, social protection, etc. On the revenue side, the decomposition of taxes is generally between personal income taxes, corporate income taxes, direct and indirect taxes, taxation of capital gains, etc. A motivation for doing this is that determining the direction of the response of growth to changes in the fiscal variables requires somewhat careful judgment on the transmission channels, for instance through their influence on private production, human capital accumulation, on productivity, or through the diffusion of innovations and network externalities.[7]

There are several findings in the theoretical literature regarding the direction of the different fiscal components on growth. Recommendations for tax policy and government spending from the view of endogenous growth models do not lead to consensual conclusions. In general, the conclusion of the endogenous growth models is that the mechanisms through which the different components of taxes and spending influence growth are diverse, thereby implying that the question of composition of government spending and tax structure on growth remains an open question. For instance, some models support the idea that income taxes are detrimental for growth through the decline of the rate of capital accumulation (see, for instance, Lucas 1990), Easterly and Rebelo 1993). This leads to the policy recommendation that reducing taxes on capital income could lead to increases in growth. However, in some other models, a positive impact on long-run growth of changes in income tax is shown to exist when these taxes are used to finance public services (see Rivas 2003). Another example, government consumption spending has been shown to affect growth alternatively negatively or positively depending upon whether public goods and services enter the households' utility function or whether they enter as inputs in the production function (see Barro 1990, Turnovsky and Fisher 1995, Dhont and Heylen 2008). Myles (2000) provides a literature review of the diversity of theoretical models analyzing tax incidence and its influence on growth. The channels through which taxation can affect growth are many: the elasticity of substitution between capital and labor in production, households' preferences over consumption in different periods of life over the life-cycle, the relationship between capital and the non-taxable factors, the share of physical capital in human capital, the way in which taxes affect risky assets, the proportion of wealth invested in foreign assets, etc.

Given the great variety of theoretical models, the diversity of their predictions regarding the effects of fiscal variable changes on growth, it is unlikely that the same model would illustrate the case of all the EU countries. Further, the balance between the various items of taxes and expenditure vary in each country and across time depending upon the juncture and their priorities. Our aim here is not to test a particular theory. The above brief review of the theoretical literature is useful to shed some light on the fact that, given the wide range of predictions from the theoretical models, imposing a priori common parameters across countries would be restrictive and may result in non robust conclusions.

When the purpose is to test the fiscal policy/growth relationship from the view of the endogenous growth model, whichever the theoretical apparatus, the empirical relationship is very

[7] See, among others, Zeng and Zhang (2002), Zagler and Durnecker (2003), Blankenau and Simpson (2004), Linneman and schabert (2003), Greiner *et al.* (2005), Agenor and Yilmaz (2011), Peretto (2003, 2007), Semmler *et al.* (2007), Gosh and Gregoriou (2008).

often a linear equation between the growth rate of GDP per capita and the different items of taxes and public spending, for given control variables describing the economic environment. We modify the standard equation by taking into account the distributional heterogeneity of fiscal policy effects on growth. The specific form employed in this paper is the following:

$$\gamma_{it}(\theta) = \sum_{j=1}^{J} \alpha_{1j}(\theta)\Delta\gamma_{it-j} + \sum_{k=1}^{K} \alpha_2(\theta)\Delta F_{it-k} + \sum_{l=1}^{L} \alpha_3(\theta)\Delta X_{it-l} + \upsilon_{it}(\theta) \quad (1)$$

where i indicates a country, t is year, γ is the growth rate of real GDP, F is a vector of fiscal variables, X is a vector of control variables, $\beta_1(\theta)$, $\beta_2(\theta)$, $\alpha_1(\theta)$, $\alpha_2(\theta)$ are vectors of coefficients to be estimated, $\alpha_{1j}(\theta)$, is a lagged coefficient and υ_{it} is an error term.

Equation (1) provides a useful way to deal with the issues discussed above and to confront the predictions of the theoretical models with the experience of the European countries by considering the percentiles of the conditional distribution of the growth rates. The θ^{th} percentile is assumed to vary between 0 and 100 per cent. The idea is to obtain the value of the estimate of the parameter vectors which best fits the impact of the fiscal variables at various points along the conditional distribution of growth. This approach permits a flexibility to capture heterogeneity. Indeed, since we are considering a pooled panel, the percentiles do not only refer to countries but also allow time variation and therefore possible non-monotonic effects of the components of taxes and expenditure on growth. Finally, finding different coefficients according to the percentiles is a way of showing that fiscal changes in the European countries may result in multiple equilibriums both in terms of transitional growth and long-run growth.

2.2 *Deficit financing and growth*

Given the importance of the government budget constraint in the theoretical models, the influence of a given component of public spending on growth depends on how an increase in this component is financed. A government considering new spending programs must decide on how to raise the necessary revenue. A financing of productive public spending by higher direct taxes will not necessarily results in a positive impact on long-run growth, because of the negative effect of the taxes on the returns of capital. Also, as far as we consider the structure of taxes and the composition of public spending, the final decision is the result of different trade-offs. For instance, cuts in labor income or capital tax might be compensated by increases in indirect taxes; or a government can search to balance an increase in productive expenditure by a decrease in non-productive expenditure. Another example is that any change in a given spending or tax can be decided by maintaining a continuously balance budget, by keeping a constant share of expenditure and taxes in GDP, or alternatively by allowing a higher or lower fiscal balance. Taxes and public policies are thus restricted by the budget constraint.

The implication is that, different financing methods have different effects on the economic growth. In his seminal papers, Harberger (1964a, 1964b) showed that the mix of direct and indirect taxes in a growth-accounting framework has a negligible effect on growth. Mendoza *et al.* (1997) show that this conjecture can be supported within the framework of an endogenous growth model. In standard endogenous growth models, expansionary fiscal spending stimulates economic growth provided that they are financed by lump-sum taxes or by non-distorsionary taxes (see, Devereux and Love 1995, Palivos and Yip 1995). This finding is, however, challenged by Pelloni and Waldman (2000). The authors find that a small amount of capital taxation can increase the growth rate. Barro (1990), Blankeneau and Simpson (2004) show that spending funded by distortionary taxes such as taxes on capital or labor income has a non-monotonic effects: increases in productive spending is growth-enhancing in the short-run, but growth-depressing in the long-run. There are conflicting views in the theoretical literature about the growth implications of a financing of public spending by public debt. Some authors conclude that the effect is unambiguously positive (for instance Greiner and Semmler 2000). Others reject this finding (Minea and Villieu 2010).

As pointed by Easterly *et al.* (2007), irrespective of the theoretical framework, it is likely that the combination of fiscal variables needed to obtain a positive impact on growth vary across countries and across time depending on a number of structural factors: the initial level of debt, the composition of revenues and taxes, fiscal institutions, different public finance constraints, etc. Again, the issue of heterogeneity is at stake.

What this implies in our case is the following. The government budget constraint can be written by considering the various components of the vector of fiscal variable F as follows:

$$\sum_{n=1}^{N} rev_{it}^{n} - \sum_{m=1}^{M} \exp_{it}^{m} + b_{it} = 0, \quad i=1,..,I \text{ and } t=1,\dots,T \tag{2}$$

where *exp* means expenditure and *rev* stands for revenue. We consider M components of public spending and N components of fiscal taxes. b is the budget surplus. As shown in a paper by Bleany *et al.* (1995), not taking into account this constraint when examining the fiscal policy/growth link yields strong biases in growth equations. Further, since the different components of the fiscal vector are linked through the budget constraint, considering all them in equation (1) yield inefficient estimates due to co-linearity between the variables. Some of them must be omitted. The omitted variables are interpreted as the financing instruments. To show this, consider for instance that we separate the taxes into distorsionary (*DIST*) and lump-sum (*LUMP*) taxes and that spending are categorized as productive (*PROD*) and non productive (*NPROD*). Equation (2) can be rewritten as follows:

$$DIST_{it} + LUMP_{it} - PROD_{it} - NPROD_{it} + b_{it} = 0 \tag{3a}$$

Assume that the omitted variable is the distorsionary tax. Then (3a) implies that:

$$[(DIST_{it} = -(LUMP_{it} - PROD)]_{it} - NPROD_{it} + b_{it}) \tag{3b}$$

In the general case, we decompose the vector F into two sub-vectors vectors F_1 and F_2 containing respectively the omitted and non-omitted variables. The constraint (3b) implies that $F_1 = -F_2$. Equation (1) can thus be rewritten as follows:

$$\gamma_{it}(\theta) = \sum_{j=1}^{J} \alpha_{1j}(\theta)\,\gamma_{i,t-j} + \sum_{k=1}^{K} \{\alpha_{22}(\theta) - \alpha_{21}(\theta)\}_{k}\, \Delta F_{2,it-k} + \sum_{l=1}^{L} \alpha_{3l}(\theta)\, X_{i,t-l} + \upsilon_{it} \tag{4}$$

Therefore, the coefficients of the fiscal variables are interpreted as follows. They indicate how changes in given fiscal variables, offset by changes in omitted fiscal variables, affect the economic growth. Equation (4) is retained as our benchmark equation for testing the fiscal policy/growth link.

3 The econometric methodology and data

3.1 Quantile regressions

Equation (4) can be rewritten in matrix form as follows:

$$Y_{it}(\theta) = X'_{it}\beta(\theta) + \upsilon_{it}(\theta), \qquad i = 1, \dots, N \text{ and } t = 1, \dots, T \tag{5}$$

where X is the vector of explanatory variables, $\beta(\theta)$ is the vector of coefficients and Y is the endogenous variable. We apply a double-quantile regression to equation (5).

Before turning to the estimation, some discussion about the methodology of quantile regression is warranted. The idea is to model the percentiles of the conditional distribution of the growth rate as functions of the explanatory variables. In a situation of heterogeneous responses of the endogenous variable to changes in the explanatory variables the standard linear estimators (OLS, GLS, GMM, etc.) are not suited. Indeed, those methodologies focus on the estimation of a conditional mean function and conditional dispersion of the endogenous variable around its mean. So doing, one assumes that the conditional mean summarizes the behavior of all the observations in the endogenous variables. This approach is good as far as we consider that the fluctuations of Y_{it} around its conditional mean are erratic or "accidental". However, when the reaction of the endogenous variables to its covariates are assumed to vary across the sample, the standard methodologies do not fully account of the diversity of reaction across the distribution of Y_{it}. In this case, we need alternative estimators.

In panel data methodologies, a now widely used approach consists in using estimators but that allow slope variations across individuals and/or time. In a recent paper, Gemmel *et al.* (2011) use Pesaran *et al.* (1999)'s pooled mean group (PMG) and mean group (MG) estimators to study the impact of fiscal policy on growth using a panel of 17 OECD countries from 1970 to 2004. Although these estimators are useful to account for different slopes across the countries of a panel, the cost of using them is a reduction of the degrees of freedom. Indeed, they are based on an average of the estimates from individual country regressions either for the short or for the long-run coefficients. Quantile estimators avoid this caveat since growth estimators conditional on fiscal policy variables, for given control variables, are obtained by considering the entire sample and by distinguishing the countries and the years according to their location in the conditional distributional of growth. Quantile estimator allows a greater flexibility by allowing all the parameters in a regression to vary across the distribution.

Let $F(y)$ be the probability distribution function of Y. The θ^{th} percentile of Y is defined as the smallest y satisfying $F(y) \geq \theta$. In a regression context, it can be shown that the finding of θ amounts to estimating β such that:

$$\hat{\beta}(\theta) = \arg\min\left\{\sum_{i=1}^{T} H(\theta, \upsilon_t)\right\}, \qquad H(\theta, \upsilon_t) = \theta\upsilon_t^{+} + (1-\theta)\upsilon_t^{-} \tag{6}$$

where v_t^{+} is the vector of residuals with positive value and 0 otherwise, v_t^{-} is the vector of negative residuals and 0 otherwise. We thus have as many estimators of β as values of $\theta \in (0,1)$. Therefore, a quantile regression leads to estimate β by changing the "representative" individual. The latter can be the "mean" (as in OLS), the median ($\theta = 0.5$) or any other percentile.

Basset and Koenker (1978) derive the asymptotic normality results for the quantile regression and show that:

$$\sqrt{T}\left(\hat{\beta}_\theta - \beta_\theta\right) \approx N\left(0, \theta(1-\theta)s(\theta)^2 J^{-1}\right) \tag{7}$$

$$J = \lim_{T \to \infty}\left(X' X / T\right) \tag{8}$$

$$s(\theta) = 1 / f(F^{-1}(\theta)) \tag{9}$$

While the estimation of β is quite simple and requires the use of simplex algorithms (see Koenker and d'Orey, 1987), the estimate of the standard error of the estimated parameters is more complicated since it requires the estimation of the unknown probability distribution function of the endogenous variable and its derivative. The latter are required in order to estimate the quantile density function $s(\theta)$, also called sparsity function. Here, the coefficient covariance matrix is computed using bootstrap resampling and the sparsity function is estimated by using a kernel density estimator as proposed by Powell (1984) and Buschinsky (1994).

All the variables in the right-hand side of equation (4) are purged from reverse causality (endogeneity) by using the double-stage quantile regression proposed by Kim and Muller (2004). They show that the double-quantile estimator is consistent for finite samples.[8] In order to obtain efficient estimates, we however depart from these authors by bootstrapping the standard errors of our estimated coefficients at the second step. Indeed, we are working with the EU countries and our data are contaminated by country cross-correlation. In order to avoid problem of inefficient estimation, we prefer a direct method of estimating the covariance matrix of the estimates by employing a bootstrapping technique (residual bootstrap).

Possible effects running from growth to control and fiscal variables are taken into account in the first step by instrumenting as fully as possible for those variables. We use the logarithm of per capita GDP, the lags of the growth rates of per capita GDP, the difference between the long and short-term interest rate, the ratio of labor force to population as well as lags of the explanatory variables themselves. In addition, endogenous relationships are avoided by not considering the contemporaneous effects of the fiscal variables.

Finally, in each regression, unobserved heterogeneity is taken into account through country fixed effects.

3.2 Data

Our dataset cover 22 countries of the European Union from 2000 to 2010.[9] Our motivation for considering the recent ten years is the following. The current members of the EU are composed of three groups of countries regarding the date of adhesion. 15 were members before the 2000s, 10 countries entered the Union in the early 2000's (in 2004) and 2 in 2007. We consider as many countries as possible and not limit our attention to EU 15. With the exception of Romania and Bulgaria whose adhesion is very recent, we therefore consider the other countries. Luxembourg has a very high GDP therefore may appear as an "outlier". To avoid a strong influence on our results, we drop it from the panel. We also do not include Cyprus and Malta for problem of data availability. This leaves us with 22 countries. Regarding the choice of the time period, we restrict years from 2000 to 2010. We begin after the introduction of the euro, since after 1999, a new institutional framework for fiscal policy was set up (Stability Growth Pact, multilateral surveillance) intended to reinforce the coordination of national fiscal policies. For the countries which joined the EU in 2004, they also had to change the conduct of their fiscal policy at least 4 to 5 years before their adhesion (the Maastricht conditions were entry requirements). Therefore, our aim is to see whether, the adoption of a common fiscal framework makes taxation and expenditure measures become growth-enhancing or growth-reducing in a similar way across countries, or whether their impact on growth have still been different across countries.

In our pooled data, an individual observation describes a country and a year, which we call "an episode" of growth rate of real GDP. Our fiscal variables are taken from the functional

---

8 Other methodologies have been proposed in the literature to deal with endogeneity bias in quantile regressions. For instance, Chernozukov and Hansen (2006, 2008) have suggested an instrumental variable quantile regression estimator. However, the latter is computationally demanding when applied to our case since it is based on a grid search procedure on the coefficients of all the variables which are suspected to be endogenous. Their method is well suited to models where there are few endogenous variables among the explanatory variables of a regression.

9 The countries are the following: Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, the Netherlands, Norway, Poland, Portugal, Slovenia, Spain, Slovak Republic, Sweden and the United Kingdom.

classification of public administration expenditure (COFOG) as set by the OECD and by considering the disaggregated taxes. This yields the fiscal categories described in Table 1. The set of endogenous and conditioning variables includes those described in Section 2.2. The GDP, the long-run and short-run interest rates, as well as the inflation and unemployment rates are from the OECD statistics. Private investment is measured by the gross capital formation of corporations and comes from the European Commission AMECO dataset. Data on labor markets were obtained from the OECD: employment, working-age population, population, hours worked per employee, labor force (the latter are used as instruments in our regressions).

All the variables in the regressions are in logarithm, except the budget surplus, the inflation rate and the interest rate term structure defined as the long-run minus the short-run interest rates. The fiscal variables are measured as share of GDP.

4 The results

4.1 The conditional distribution of growth episodes

We consider both measures of the growth rate: the simple growth rate of the real GDP and the growth rate of per capita GDP. The policy recommendations regarding the design of tax structure and composition of expenditure in the EU are usually made by considering the real GDP growth (not deflated by the population size). The second indicator, the growth rate of the real GDP per capita, is helpful for evaluating how economic growth feeds into welfare (a rough measure of income distribution).

In order to contrast the different growth episodes with each others, we first examined how the countries and years are shared across the main percentiles of the conditional distribution of the growth rate of the real GDP. In this view, we ran different regressions corresponding to different percentiles from the 10^{th} to 90^{th} percentiles (each percentile estimate is obtained using the pooled panel). Then, we examined the regressions for which the coefficients measuring the impact of the fiscal variables on growth were quite similar. We computed the fitted value of growth and consider that two fitted values belonged to the same group if they were obtained from regressions in which the coefficients of the fiscal variables were quite similar in magnitude. Again, we stress that this classification is made *after* running quantile regressions with the pooled data. In terms of the growth impact of changes in taxation and expenditure, we observed that the estimated coefficients of the explanatory variables were rather similar for four “subgroups” of percentiles as described in Table 2a.

At the left-hand side of the distribution, below the 40th percentile, the real GDP growth rate is less than or equal to 2.6 per cent. This corresponds to times of crisis. Indeed, the intervals up to the 40th percentiles contain the data for all the countries corresponding to the years 2008 and 2009. In addition, these intervals also include the growth episodes of the most ancient members of the EU corresponding to the years 2002, 2003 and 2010. The percentiles up to the 40th are therefore refereed as low growth episodes in times of crisis. At the higher end of the distribution, above the 70th percentile, the real GDP growth is driven by a catch-up dynamics. Indeed, the group of years and countries is made of the new member states between 2002 and 2007 (Central and Eastern Europe) and some former member countries belonging geographically to the periphery of Europe, for instance Ireland, Portugal, Spain in the earlier 2000’s. Their growth rate is greater than 4.3 per cent per annum. There is a broad consensus in the literature that these countries’ very fast growth was an illustration of a catch-up dynamics to the standard of living of the richest members of the EU from 2000 onwards. We therefore consider the percentiles above the 70th as illustrating transitional growth rate. Then we have medium growth episodes, between 2.6 and 3.3 per cent (from the 40th to the 50th percentile) and high growth not corresponding to transitional growth, between 3.3 and 4.3 per cent, (from the 50th to the 70th percentiles).

Table 1

Classification of Fiscal Variables

Theoretical Classification	Classification in the Data Source
Taxes	
Direct taxation	Direct taxes on business Direct taxes on households
Payroll taxes	Social security contributions received by governments
Indirect taxation	Taxes on production and imports
Other government revenues	General Government total receipts minus direct and indirect taxation
Expenditure	
Sovereign expenditure	Defence expenditure Security expenditure Economic affairs General public service expenditure
Human capital	Education expenditure Health expenditure Social security expenditure Recreation and culture Environment
Other expenditure	General government total disbursements minus productive and unproductive expenditure
Budget surplus	Government total revenues minus Government total disbursements

Table 2a

Classification of Countries According to the Results of Quantile Regressions
(real GDP growth)

Low-growth episodes: <2.6% [0^{th}-40^{th}]	All countries Most ancient members	Crisis episodes (2008-09) Years: 2002, 2003, 2010
Medium low growth episodes: 2.6%-3.3% Medium high growth episodes: 3.3%-4.3% [50^{th}-70^{th}]	Most ancient members Most ancient members New members	2000, 2001 and 2004 to 2007 2000, 2001 and 2004 to 2007 2000, 2001, 2010
High-growth episodes: >4.3% [70^{th}-100^{th}]	New member states Periphery	Period 2002 to 2007 (catch-up growth) Early 2000's

An interesting feature of the data is that the more industrialized members of the EU move in the distribution over different years (all the intervals of the different percentiles are "visited"), which is not the case of the Central and Eastern emerging countries. For the latter we indeed have few observations between the 40^{th} and 70^{th} percentiles, which could be explained by the fact that they are still converging to the other countries and therefore they experience a higher growth rate (catch-up dynamics).

Comparing the cases of two leading economies of the EU, France and Germany, we observe an unhooking of the former with regard to the latter from 2006 onwards. Indeed, from Table 2b, it is seen that France's growth rates systematically lies in lower percentile intervals.

For purpose of comparison, a classification was also done by considering the regressions with the growth rate of per capita GDP. The conditional distribution of per capita GDP growth led us to classify the growth episodes in three intervals. The first group was composed of countries and years for which the conditional growth rate is below 3.2 per cent (which correspond to the following interval of percentiles: [0^{th} – 40^{th}]), the second group for countries and years for which the growth rate lies between 3.2 and 5 per cent (the interval of percentiles is [40^{th} – 70^{th}]) and finally the third group consisted of countries and years characterized by a growth rate above 5 per cent in the interval [70^{th} – 100^{th}].

4.2 Tax and expenditure effects on real GDP growth across percentiles

Tables 3 till 5 report the estimation results of equation (4). The reported coefficients are cumulative sums over the two years following the initial changes in taxation and expenditure. We report the cumulative sum of the coefficients over the two years. This corresponds to the length of time usually required for changes in investment to fully affect growth in Europe. Further, we assume that the implementation of fiscal policy requires a delay before impacting the economy and that short-run effect are completely dissipates after two years.

Table 2b

Classification of Growth Episodes Across Quantile Intervals for France and Germany

(real GDP growth)

France		Germany	
2000	70^{th}-80^{th}	2000	50^{th}-60^{th}
2001	30^{th}-40^{th}	2001	30^{th}-40^{th}
2002	20^{th}-30^{th}	2002	10^{th}-20^{th}
2003	40^{th}-50^{th}	2003	20^{th}-30^{th}
2004	40^{th}-50^{th}	2004	30^{th}-40^{th}
2005	30^{th}-40^{th}	2005	20^{th}-30^{th}
2006	50^{th}-60^{th}	2006	70^{th}-80^{th}
2007	40^{th}-50^{th}	2007	60^{th}-70^{th}
2008	10^{th}-20^{th}	2008	20^{th}-30^{th}
2009	0^{th}-10^{th}	2009	0^{th}-10^{th}
2010	20^{th}-30^{th}	2010	60^{th}-70^{th}

The different coefficients must be interpreted in light of our discussion in Section 2.2. In Table 3, first regression, the coefficients indicate the effect on growth of changes in the different variables (two years after the initial change) when these changes are accompanied by changes of similar amount in welfare expenditure. For instance the estimate –0.05 of direct taxation says that a 1 per cent increase in direct taxation, used to finance a 1 per cent increase in welfare expenditure, reduces growth by 0.05 per cent two years after the initial change in direct taxation. In Table 4, second regression, the coefficients measure the impact of changes in the variables on growth, when there are changes of equal amount in the budget surplus. For instance, the coefficient –0.09 of social security contributions says that a 1 per cent increase in social security spending, entirely reflected in the budget balance (which means that neither other spending, nor taxes are modified) reduces growth by 0.09 per cent, two years after the initial change. All the regressions in Tables 3 till 6 must be interpreted in a similar way.

We report the results of the regressions based on the 40^{th}, 50^{th}, 60^{th} and 70^{th} percentiles. The reader must keep in mind that for the different choices of percentiles, we do not split the data into different sub-samples. We use the whole pooled observations. The difference with the classical “mean-based” estimations is that, instead of the conditional mean, the representative observation to which the others are compared is the reported percentile.

Instead of commenting on all the estimated coefficients, we focus on the variables related to the ongoing debate in Europe on the fiscal tools that are viewed as growth-enhancing instruments: the improvement of competitiveness on the labor cost which may imply a reform of the social security systems, optimal taxation and in particular the trade-off between direct and indirect taxes, the rationalization of public expenditure by reducing unproductive public spending, fiscal devaluation.

Table 3

Growth Equation. Two-stage Quantile Regression with Bootstrapped Standard Errors

(t-ratios in parentheses)

Omitted Variable	**Welfare Expenditure**				**Direct Taxation**			
	0.4	0.5	0.6	0.7	0.4	0.5	0.6	0.7
Constant	–2.61***	–2.87***	–0.60***	–0.45**	–4.74***	–3.17***	–2.96***	–1.30***
	(–4.51)	*(–5.04)*	*(–5.58)*	*(–2.49)*	*(–4.70)*	*(–4.57)*	*(–5.61)*	*(–3.58)*
Business investment	0.04**	0.04**	–0.007	0.07***	0.09***	0.005	–0.025	0.008
	(2.34)	*(2.13)*	*(–0.26)*	*(2.99)*	*(4.94)*	*(0.185)*	*(–0.84)*	*(0.39)*
Employment growth	–0.15	0.14	0.04	0.003	–0.009	–0.09	0.03	0.004
	(–1.51)	*(1.51)*	*(0.55)*	*(0.03)*	*(–0.108)*	*(–0.97)*	*(0.39)*	*(0.053)*
Hum. capital expenditure	–0.02	–0.04	–0.003	–0.148***	–0.005	–0.07	–0.28***	–0.32***
	(–0.44)	*(–1.06)*	*(–0.09)*	*(–4.60)*	*(–0.54)*	*(–0.73)*	*(–4.06)*	*(–3.41)*
Welfare expenditure	-	-	-	-	–0.02	0.074	0.11*	0.102
					(–0.42)	*(1.27)*	*(2.20)*	*(1.58)*
Sovereign expenditure	–0.08	–0.004	0.36*	0.24	–0.04	–0.001	0.16**	0.122***
	(–0.37)	*(–0.02)*	*(1.88)*	*(1.08)*	*(–0.66)*	*(–0.02)*	*(2.55)*	*(2.10)*
Direct taxation	–0.05***	–0.08***	–0.27***	–0.11	-	-	-	-
	(–3.20)	*(–4.66)*	*(–6.16)*	*(–1.50)*				
Soc. Sec. contributions	0.22	–1.66***	0.08	–0.04	–0.69*	–1.51***	0.05	0.06
	(1.08)	*(–3.56)*	*(0.53)*	*(–0.176)*	*(–1.87)*	*(–4.16)*	*(0.28)*	*(0.21)*
Indirect taxation	–1.22***	0.34	0.16	–0.11	–1.59***	0.297	–1.31***	–0.45***
	(–4.34)	*(0.70)*	*(1.09)*	*(–0.51)*	*(–4.84)*	*(1.17)*	*(–5.15)*	*(–2.33)*
Other taxes	–0.03	0.05	–0.19	–0.46***	–0.16	0.04	–0.12	–0.18
	(0.35)	*(0.29)*	*(–1.25)*	*(–2.64)*	*(–0.96)*	*(0.18)*	*(–0.70)*	*(–0.96)*
Budget surplus	0.03	–0.01	–0.20***	–0.407***	–0.03	–0.19**	–0.19***	–0.32***
	(0.35)	*(–0.22)*	*(–3.73)*	*(–4.98)*	*(–0.36)*	*(–1.99)*	*(–2.92)*	*(–3.63)*
Pseudo R^2	0.62	0.61	0.66	0.56	0.63	0.58	0.65	0.58

Note: *, **, *** mean statistical significance at 10, 5 and 1 per cent respectively.

Table 4

Growth Equation. Two-stage Quantile Regression with Bootstrapped Standard Errors

(t-*ratios in parentheses)*

Omitted Variable	Indirect Taxes				Budget Surplus			
	0.4	0.5	0.6	0.7	0.4	0.5	0.6	0.7
Constant	–2.59***	–3.90***	–0.54***	–0.75***	–2.87***	–3.85***	–0.57***	–1.00***
	(–3.38)	*(–5.00)*	*(–5.48)*	*(–4.94)*	*(–4.35)*	*(–5.43)*	*(–4.82)*	*(–6.29)*
Business investment	0.09***	0.05**	–0.025	0.008	0.08***	0.04	–0.004	–0.02
	(4.94)	*(2.29)*	*(–0.842)*	*(0.39)*	*(4.15)*	*(1.59)*	*(–0.15)*	*(–1.19)*
Employment growth	–0.009	0.06	0.03	0.004	–0.02	–0.009	–0.003	–0.06
	(–0.108)	*(0.70)*	*(0.39)*	*(0.05)*	*(–0.28)*	*(–0.11)*	*(–0.04)*	*(–0.64)*
Hum. capital expenditure	0.22*	0.38***	–0.001	–0.38***	0.27***	0.34***	–0.05	–0.24***
	(1.91)	*(2.97)*	*(–0.014)*	*(–4.56)*	*(2.64)*	*(3.25)*	*(–0.58)*	*(–2.78)*
Welfare expenditure	–0.21***	–0.24***	–0.03	0.13**	–0.24***	–0.21***	–0.07	–0.15*
	(–2.84)	*(–3.39)*	*(–0.51)*	*(2.05)*	*(–3.49)*	*(–3.62)*	*(–0.83)*	*(–1.86)*
Sovereign expenditure	–0.07	–0.15*	0.05	0.15***	–0.07	–0.11	0.138**	0.31***
	(–1.29)	*(–1.81)*	*(0.84)*	*(2.77)*	*(–1.24)*	*(–1.60)*	*(2.41)*	*(6.09)*
Direct taxation	–0.09***	–0.10***	–0.26***	–0.14**	–0.09***	–0.09***	–0.24***	–0.19***
	(–4.84)	*(–5.28)*	*(–5.15)*	*(–2.33)*	*(–4.96)*	*(–5.03)*	*(–4.63)*	*(–3.53)*
Soc. Sec. contributions	–1.31***	–2.06***	0.05	–0.31	–1.45***	–2.01***	0.51**	0.63**
	(–3.29)	*(–4.88)*	*(0.28)*	*(–1.36)*	*(–4.24)*	*(–5.25)*	*(2.03)*	*(2.28)*
Indirect taxation	-	-	-	-	0.14	0.17	0.33	1.05***
					(0.60)	*(0.74)*	*(1.52)*	*(4.10)*
Other taxes	–0.165	0.01	–0.12	–0.06	–0.02	0.05	0.02	0.28
	(–0.96)	*(0.05)*	*(–0.71)*	*(–0.31)*	*(–0.12)*	*(0.28)*	*(0.09)*	*(1.34)*
Budget surplus	–0.03	–0.02	–0.195***	–0.32***	-	-	-	-
	(–0.36)	*(–0.21)*	*(–2.93)*	*(–3.63)*				
Pseudo R^2	0.63	0.61	0.65	0.58	0.62	0.62	0.65	0.58

Note: *, **, *** mean statistical significance at 10, 5 and 1 per cent respectively.

Table 5

Growth Equation. Two-stage Quantile Regression with Bootstrapped Standard Errors
(t-ratios in parentheses)

Omitted Variable	Sovereign Expenditure			
	0.4	0.5	0.6	0.7
Constant	–3.37*** *(–4.74)*	–3.05*** *(–5.53)*	–0.76 *(–1.60)*	–0.53*** *(–3.08)*
Business investment	0.05** *(2.54)*	0.05** *(2.24)*	–0.01 *(–0.43)*	0.06*** *(3.01)*
Employment growth	–0.07 *(–0.76)*	0.06 *(0.65)*	0.09 *(1.26)*	0.002 *(0.02)*
Hum. capital expenditure	0.10 *(1.27)*	–0.09 *(–0.52)*	0.07 *(0.97)*	–0.15* *(–1.80)*
Welfare expenditure	–0.134* *(–1.98)*	0.03 *(0.24)*	–0.06 *(–0.77)*	0.03 *(0.37)*
Sovereign expenditure	-	-	-	-
Direct taxation	–0.077*** *(–3.90)*	–0.10*** *(–5.31)*	–0.29* *(–1.74)*	–0.15** *(–2.34)*
Soc. Sec. contributions	0.35 *(1.54)*	–1.97* *(–1.82)*	0.28 *(0.99)*	0.16 *(0.67)*
Indirect taxation	–1.59*** *(–4.51)*	0.61 *(0.49)*	0.40* *(1.70)*	–0.04 *(–0.17)*
Other taxes	0.12 *(0.65)*	0.01 *(0.05)*	–0.03 *(–0.16)*	–0.55*** *(–2.81)*
Budget surplus	–0.03 *(–0.40)*	–0.017 *(–0.21)*	–0.18** *(–2.50)*	–0.32*** *(–3.65)*
Pseudo R^2	0.63	0.61	0.66	0.57

Note: *, **, *** mean statistical significance at 10, 5 and 1 per cent respectively.

4.2.1 Social security contributions

Social security contributions have the strongest influence among the different fiscal variables (greatest coefficients) but their effect on growth varies across percentiles. Their expected total effect is ambiguous. Indeed, on the one side, they have a negative impact on growth (due to higher labor costs). On the other side, they may have a positive impact due to second round demand effects (in Europe, higher income transfers are usually the counterpart of higher social security spending). From the tables, we see that augmenting social security taxes had the potential for reducing growth during medium low growth episodes. Indeed, we recall from Table 2a, that the 40^{th} and 50^{th} percentiles correspond to medium low growth episodes (growth rates between 2.6 and 3.3 per cent). In Tables 3 till 5 it is seen that we obtain a negative and statistically significant coefficient of social security contributions for these two percentiles. Therefore, in the European economies that have been growing moderately (with a real growth rate between 2.6 and 3.3 per cent), increases in social security contributions have been detrimental for growth. This happened when the increases in social security contributions were not followed by any changes in public spending or taxes (we obtain negative coefficients in Table 4, when the budget surplus is the omitted variable), or when the governments decided to compensate the increase in social security contributions by lower direct or indirect taxes (see the negative coefficients in Table 3, when the omitted variable is direct taxation and in Table 4, when indirect taxation is omitted from the regressions).

In countries with a fast growth rate, we find that the total impact on growth of an increase in social security contribution has been positive (always for the 60th percentile and sometimes for the 70th percentile), though they are not found to be significantly related to growth, except when the omitted variable is the budget surplus (Table 4). Therefore, the estimates suggest that in the European emerging countries (whose growth episodes are located in the percentiles above the 60th), the negative growth effects of social security revenues are cancelled out by their positive demand effects.

Therefore, from these results, we can conjecture that a reduction in the employers and employees' contribution to social security would have no effect in the fast growing countries (Southern and Eastern European countries) , while they may be growth-enhancing in those countries experiencing a moderate growth (the industrialized countries). For instance, if the governments in Hungary, Poland, or Spain would like to raise growth by improving the competitiveness on labor costs and decide to reduce the social contribution revenues, this policy would be ineffective on growth. But, it would work in countries like Sweden, Germany or UK. One reason may be that in the latter countries social security contributions account for a high proportion of the total labor costs. Another reason is that in these countries, the supply effects of a reduction in social security contribution more than outweigh the negative demand effects (since the contribution finances unemployment benefits). In the eastern European countries social benefits are rather financed by taxes.

4.2.2 Direct and indirect taxation

We first consider the growth impact of a mix between direct and indirect taxation, looking at the respective coefficients of these variables in Tables 3 and 4 when the other variable is omitted from the regression. In Table 3, the coefficients corresponding to the line "indirect taxation" and columns 6 till 9 measure the impact on growth of a shift from indirect to direct taxation. In Table 4, the coefficients in the line labeled "direct taxation" and columns 1 till 4 indicate the impact on growth of a shift from direct to indirect taxation. Indirect taxes can be considered as taxes on consumption, while direct taxes are taxes on production (labor and capital revenues). It is seen that a shift from direct to indirect taxes (Table 4), that is a fall of the former followed by an increase in the latter, is growth-augmenting. Indeed the estimated coefficients are negative, thereby indicating that growth moves in the opposite direction of direct taxes. Table 3 yields a similar conclusion if one considers instead a shift from indirect to direct taxation (higher direct taxes substituted for lower indirect taxes). However, the impact of direct taxation in Table 4 is much smaller than the impact of indirect taxation in Table 3 (compare the coefficients for the different percentiles). This suggests that a reduction of direct taxes compensated by higher indirect taxes is more efficient for growth than a decrease in indirect taxes followed by an increase in direct taxes. One reason may be that direct taxation is more distortionary than indirect taxation.

Now, what happens if the governments rely on either one or the other form of taxation (when none of them is considered as an omitted variable)? Higher indirect taxes reduce growth mainly in times of crises or during low-growth episodes (in Tables 3,4,5, we find a statistically significant coefficient for the 40th percentile, while the coefficient is often non-significant for the other percentiles). Recall that, in Table 2a, the 40th percentile refers growth rates less than 2.6 per cent and includes years of crisis. Higher direct taxes significantly reduce growth in all the countries. But the negative effect is stronger in the fastest growth countries (compare the coefficients in Tables 3 till 5 between the 40th, 50th percentiles and the 60th, 70th percentiles). Therefore, increases in direct taxation have been more detrimental for the economies which were experiencing a catching-up dynamics.

4.2.3 The impact of public spending

On the expenditure side, our results point to different effects of sovereign and welfare expenditure across the percentiles and the way in which they affect growth depends upon the financing variables in the government budget constraint. When an increase in welfare or sovereign expenditure is financed by an equivalent increase in taxes (direct or indirect), these expenditures boost growth in the countries that are growing fast (the coefficients for the 60th and/or 70th percentiles are positive and statistically significant in Tables 3 and 4), but they are be neutral or even detrimental in the countries with a low growth rate (we obtain negative coefficients for the 40th and 50th percentiles in Tables 3 and 4). The coefficients of welfare expenditure are statistically significant when higher welfare spending is financed by higher indirect taxes). The reported coefficients capture the influence on growth of recreation, culture, and environment spending, social security benefits, sovereign spending. Our results suggest the following interpretation. Although the literature usually classifies these spending as unproductive, they may have a demand effect on growth that cancel out the negative effects of the accompanying tax increases, specifically in the European emerging countries that experience a catch-up growth.

Table 5 also suggests that welfare expenditures have usually no significant impact on growth, if a trade-off is made with other spending items, for instance sovereign expenditure. Finally, if a government raises welfare expenditures and maintain the other spending and taxes at their current level, the increase results in a negative impact irrespective in all countries (in Table 4, when the omitted variable is budget surplus, the coefficients of welfare expenditure is negative for all the percentiles and statistically significant in most cases). In the same context (no changes in the structure of taxes and spending), sovereign expenditures appear to have significant positive effects on growth only in those countries growing fast (the estimated coefficient are statistically significant for the 60th and 70th percentiles).

The empirical evidence regarding the growth effect of human capital spending (health and education expenditure) is mixed. These expenditures, when their coefficient is statistically significant, contribute positively to economic growth in times of crisis or during low-growth episodes in the richest European countries (see the coefficients in Table 4 for 40th and 50th percentiles). However, any increase in this category of spending reduces growth during high-growth episodes (see the coefficients, in Tables 3 till 5, for the 60th and 70th percentile). The positive sign is intuitive, since such expenditure is expected to enhance labor productivity. The negative sign reflects the fact that, in the European emerging economies, educational and health expenditure seem to have been inefficient in generating a positive growth rate, which could be explained by a weaker linkage between public education and wealth outlays. As reported in the literature, there may be several causes of ineffective human capital spending, among which the inefficient role of institutions and governance in mediating the nexus between social spending indicators and growth. Incorporating institutions indicators as additional control variables in the model would be interesting in assessing the negative link. We let this for a further study.

An important policy consequence of our findings is that we would be unable to draw recommendations regarding the composition of public expenditure in the EU countries in connection with growth, without considering two groups of countries, namely the most ancient members and the recent members that are still in a catch-up growth process. For instance, the usual suggestion of reducing welfare expenditure would be a good thing for growth efficiency in the industrialized countries, but would have doubtful effects on growth in the emerging countries. A reallocation of welfare expenditure to sovereign expenditure (which mean reducing the former while increasing the latter) would be a good thing in the low-growth European countries, but would certainly not be a mean of enhancing growth in the countries with a fast growth rate (as is seen in Table5, the coefficient of welfare expenditure, when sovereign expenditure is the omitted variable, carries a statistically negative sign only for the 40th percentile).

4.2.4 Fiscal devaluation

The principle of a fiscal devaluation is to reduce social security contributions (essentially payroll tax) and to increase in VAT. Such a policy is expected to work through both a demand channel and a supply channel by inciting firms to reduce their prices more or less in proportion to the decrease in unit labor costs. Our results lead mixed conclusions. The estimations suggest that such a policy could lead to a sizeable positive effect on growth, but only in the countries that experience a low growth rate (the most industrialized countries of Europe, like France, the UK, Germany, Finland, etc). Conversely, the impact would be neutral for growth in the emerging high-growth countries (see Table 4, the coefficients in the regressions where indirect taxation are the omitted variable. They are negative and statistically significant for the 40th and 50th percentiles, but non-significant for the 60th and 70th percentiles). Therefore, a transfer of fiscal revenues from payroll taxes to indirect taxes can either drive growth downwards or boost it. In the most industrialized countries (a majority of which have their growth episodes located below the median), one may expect the shift in the tax schedule to result in a higher growth. One reason may be that, in the EU, when growth is low, the price channel (domestic goods are sold at a reduced price) plays more intensively than the tax channel on domestic demand (the elasticity of domestic demand with respect to relative prices may be higher than the elasticity with respect to indirect taxes). Conversely, a reason why a measure like a fiscal devaluation would be neutral in the emerging EU countries facing a fast growth rate may be that the fall consumption fall following the rise in indirect taxation outweigh its increase due to higher real wage.

4.3 Impact of fiscal policy on per capita growth rate under alternative financing hypotheses

We now test the robustness of the above results to different changes in the specification. First, we consider the growth rate of per capita GDP as has been done in previous papers. We are no longer reasoning from a growth efficiency point of view, but we want to see whether different fiscal policies can raise or jeopardize the growth rate of the standard of livings across years and countries. As said before, working with per capita growth rate means that we assume that a shift in GDP modifies the average income per individuals.

We further add one additional lag to the explanatory variables since the annual macroeconomic programs transmitted by the countries to the EU Commission are evaluated over a period of three years. We also consider an alternative classification of spending. As shown in Table 6, we now consider three groups of expenditure: social spending, economic and sovereign expenditure, and, other public expenditure. Direct taxation now incorporates a third component, namely other government revenues. These include for instance taxes on property transactions. Another difference with the preceding section is that, instead of omitting variables from our specifications one by one, we also consider the case where several fiscal variables are omitted. Finally, we add inflation and the initial growth rate of per capita GDP to the list of control variables.

Tables 7a and 7b report the results for the 25th, 50th and 75th quantiles. The reported coefficients are cumulative sums over the three years following the initial changes in taxation and expenditure. Regression (1) assumes that changes in taxes and expenditure are fully reflected by changes in the budget surplus. In regression (2), it is assumed that changes in taxation and public spending are not entirely reflected in budget deficit/surplus, because the government modifies the structure of spending by modifying social expenditure. Similar interpretations apply to regressions (3) till (5).

Table 6

An Alternative Classification of Fiscal Variables

Theoretical Classification	Classification in the Data Source
Direct taxation	Direct taxes on business Direct taxes on households Other direct taxes (total direct taxes minus direct taxation on business and households)
Indirect taxation	Taxes on production and imports
Other government revenues	General Government total receipts minus direct and indirect taxation
Economic and sovereign expenditure	Defense expenditure Security expenditure Education expenditure Health expenditure General public service expenditure Economic affairs expenditure
Social expenditure	Expenditure on recreation and culture Social security and welfare expenditure
Other expenditure	General government total disbursements minus productive and unproductive expenditure
Budget surplus	Government total revenues minus Government total disbursements

Table 7a

Growth Equation (per capita). Two-stage Quantile Regression with Bootstrapped Standard Errors

(t-ratios in parentheses)

Regression No.	**(1)**			**(2)**			**(3)**		
Omitted Variable	**Budget Surplus**			**Budget Surplus and Social Expenditure**			**Budget Surplus, Indirect Taxes and Social Expenditure**		
	0.25	0.50	0.75	0.75	0.50	0.75	0.25	0.50	0.75
Constant	0.17	0.02	0.27**	–0.009	–0.09	–0.11	0.11	–0.08	0.10
	(1.51)	*(0.18)*	*(2.39)*	*(–0.09)*	*(–0.94)*	*(–1.29)*	*(1.29)*	*(–1.00)*	*(1.32)*
Growth (–1)	0.12	0.35***	0.20*	0.28**	–0.05	0.05	0.23**	0.24**	0.07
	(1.07)	*(3.34)*	*(1.94)*	*(2.84)*	*(–0.66)*	*(0.62)*	*(2.24)*	*(2.60)*	*(0.90)*
Inflation	–1.48***	–1.87***	–1.39***	–1.14***	–1.13***	–1.09***	–1.69***	–0.80***	–0.87***
	(–4.63)	*(–6.84)*	*(–4.35)*	*(–3.93)*	*(–4.88)*	*(–4.12)*	*(–5.72)*	*(–3.11)*	*(–3.83)*
Business investment	0.74***	0.45**	0.71***	0.44**	0.37**	0.33**	0.92***	0.29	0.29*
	(3.56)	*(2.14)*	*(3.43)*	*(2.17)*	*(2.11)*	*(2.02)*	*(4.58)*	*(1.48)*	*(1.70)*
Employment growth	0.58**	0.54**	0.51**	0.72***	0.70***	0.62***	0.43*	0.69**	0.68***
	(2.51)	*(2.32)*	*(2.07)*	*(2.82)*	*(3.26)*	*(2.89)*	*(1.73)*	*(2.94)*	*(3.26)*
Direct taxation	–0.39	–0.15	–0.66	–0.19	–0.84**	–0.79**	–1.21***	–0.42	–0.99**
	(–0.85)	*(–0.32)*	*(–1.55)*	*(–0.43)*	*(–2.27)*	*(–2.18)*	*(–2.71)*	*(–0.99)*	*(–2.58)*
Indirect taxation	–1.05	–0.52	–1.46**	–0.77	1.28*	1.37**	-	-	-
	(–1.54)	*(–0.76)*	*(–2.34)*	*(–1.21)*	*(2.24)*	*(2.51)*			
Other taxation	0.10	0.30	–0.05	0.21	–0.19	0.11	0.11	0.33	–0.30
	(0.24)	*(0.94)*	*(–0.14)*	*(0.55)*	*(–0.63)*	*(0.34)*	*(0.28)*	*(0.92)*	*(–0.93)*
Economic and sovereign expenditure	0.07	0.52**	–0.26	0.43*	0.32*	0.51**	–0.26	0.45**	0.28
	(0.28)	*(2.11)*	*(–1.07)*	*(1.91)*	*(1.88)*	*(2.68)*	*(–1.11)*	*(2.09)*	*(1.49)*
Social expenditure	–0.65	–0.77*	–0.67**	-	-	-	-	-	-
	(–0.65)	*(–1.89)*	*(–2.20)*						
Budget surplus	-	-	-	-	-	-	-	-	-
Pseudo R^2	0.53	0.38	0.53	0.39	0.45	0.45	0.66	0.57	0.45

Note: *, **, *** mean statistical significance at 10, 5 and 1 per cent respectively.

Table 7b

Growth Equation (per capita)
Two-stage Quantile Regression with Bootstrapped Standard Errors
(t-ratios in parentheses)

Regression No.	**(4)**			**(5)**		
Omitted Variable	**Budget Surplus and Indirect Taxes**			**Indirect Taxes, Other Taxes and Social Expenditure**		
	0.25	0.50	0.75	0.25	0.50	0.75
Constant	0.10 *(1.12)*	–0.08 *(–0.97)*	0.09 *(1.13)*	0.02 *(0.31)*	–0.03 *(–0.42)*	0.11 *(1.47)*
Growth (–1)	0.04 *(0.39)*	0.22** *(2.37)*	0.06 *(0.65)*	–0.09 *(–0.82)*	0.07 *(0.72)*	–0.07 *(–0.69)*
Inflation	–1.60*** *(–5.18)*	–1.34*** *(–5.04)*	–1.25*** *(–5.26)*	–1.79*** *(–5.30)*	–0.74** *(–2.32)*	–1.44*** *(–5.81)*
Business investment	0.96*** *(4.93)*	0.37* *(1.83)*	0.36** *(2.00)*	1.09*** *(5.19)*	0.31 *(1.49)*	0.51** *(2.53)*
Employment growth	0.76*** *(3.18)*	0.79*** *(3.25)*	0.62** *(2.87)*	0.62** *(2.56)*	1.01*** *(4.79)*	0.39* *(1.80)*
Direct taxation	–0.97** *(–2.19)*	–0.39 *(–0.87)*	–0.88** *(–2.29)*	–0.84* *(–1.77)*	–0.54 *(–1.19)*	–1.12** *(–2.50)*
Indirect taxation	-	-	-	-	-	-
Other taxation	0.04 *(0.12)*	0.46 *(1.16)*	–0.34 *(–1.02)*	-	-	-
Economic and sovereign expenditure	0.18 *(0.64)*	0.68** *(2.62)*	0.63*** *(2.85)*	–0.14 *(–0.55)*	0.37 *(1.59)*	0.24 *(1.18)*
Social expenditure	–0.94** *(–2.45)*	–0.62 *(1.63)*	–0.60* *(–1.89)*	-	-	-
Budget surplus	-	-	-	–0.04 *(–0.18)*	–0.04 *(–0.21)*	0.36* *(1.69)*
Pseudo R^2	0.54	0.40	0.44	0.54	0.41	0.46

Note: *, **, *** mean statistical significance at 10, 5 and 1 per cent respectively.

We begin with a brief comment of the results obtained for the conditioning variables (the variables other than the fiscal variables). Their coefficients have the expected signs. Both the business investment ratio and the employment growth enter the regressions with a positive sign and they are mostly statistically significant, irrespective of the quantiles. This seems better than in our previous regression where the ratio of business investment to GDP was positive and statistically significant for the low-growth countries only and the growth rate of the employment rate was rarely significant across the different regressions. Inflation negatively affects per capita GDP growth, which is not a surprised given that price stability has been set up as a prerequisite for sustainable growth in the EU.

Higher direct taxation significantly reduces growth if a country is experiencing either a low-growth or a high-growth rate (in Regressions 3 till 5) while the effect is statistically insignificant for middle-growth countries. Therefore, an increase in direct taxation financed by an equivalent decrease in indirect taxes, social expenditure, or which results in a higher budget surplus, is growth-reducing when growth is below 1.14 or above 3 per cent (these are the average growth rates in the intervals of percentiles shown in Table 2c). In Tables 3 till 5, we see that the coefficients of direct taxations are statistically negatively significant for the 25^{th} and 75^{th} quantiles. When indirect taxation is excluded from the list of omitted variables (regressions 1 and 2), higher direct taxes are growth-reducing only for the high-growth countries (with a growth rate above 3 per cent).

The regressions also report that sometimes, higher indirect taxes can have a negative growth effect in the low-growth economies but a positive effect in the high-growth countries (Regression 2). A reduction of public deficit by higher indirect taxes, or the financing of additional social spending by a higher indirect taxation has several theoretical effects. In principle, deficits and indirect taxes imply a shift in growth in opposite directions. The effect of the former is either positive or negative depending upon whether one observes strong or weak Keynesian multipliers (this depends upon crowding out effects, Barro-Ricardo effects, etc). Indirect taxes are expected to be growth-reducing. The total impact is thus either positive or negative depending upon the effects which is predominant. If we look at Regression 2, it seems that the taxation effect is larger in low-growth countries, while the negative effects of higher budget surpluses dominates in high-growth economies. Therefore, an indirect taxation used to finance social expenditure has the benefit of shifting growth upward if an economy evolves on its transition growth path to its long-run per capita GDP level. Otherwise, once the transition phase is achieved, indirect taxation is likely to result in a lower growth. This finding can be explained by our previous observation that social spending are growth-enhancing in the European emerging countries, but growth-reducing in the industrialized countries (see Section 4.2.4).

Interestingly, the results report a positive effect on growth of economic and sovereign expenditure in high-growth countries, while they are neutral for the group of low-growth countries. Indeed in Regressions 1, 2 and 4, we obtain statistically significant positive coefficients for the median and the 75^{th} quantile only. Economic and sovereign expenditure are therefore beneficial for per capita growth above 3 per cent, when the initial composition of taxes and spending remains unchanged (Regression 1), when their increase is substituted for social expenditure (Regression 2), or even if they are partially financed by higher indirect taxation (Regression 4).

Finally, we can see that social expenditure, when included in the list of explanatory variables, has a negative effect on growth irrespective of the quantile (Regression 4). This contrasts with our findings in the preceding section, since we saw that such spending had strong demand effects in the fast-growth countries.

5 Conclusion

Can we apply common fiscal policies in Europe to boost growth in Europe? The answer seems to be negative.

While using taxes and public spending to foster growth, the EU governments also use their fiscal policy to keep their finance sustainable. Our results cast some doubts on a widespread idea in the policy circles according to which a higher growth rate in the EU could be achieved with the same fiscal mix in all member countries. Against this view, the quantile estimates strongly illustrate heterogeneous reactions across the EU economies.

In light of our findings, we favor the idea of distinguishing among the ancient member countries and the recent emerging countries which adhered to the EU in the early 2000's. On the differences discussed in this paper, social security spending, direct taxation, welfare and sovereign expenditure and human capital expenditure have strikingly different effects on the growth rate of the real GDPs. Increases in human capital spending are growth-enhancing in the industrialized EU countries, while welfare and sovereign expenditure play a more important role in fostering growth in the emerging economies. Direct taxation exerts a much more detrimental impact in the countries that are growing rapidly than in those that experiment a slow growth. When the growth rate is considered in per capita terms, indirect taxes appear to exert an asymmetric effect on the EU economies: they are harmful in the low-growth countries, but not inconsistent with a stronger growth dynamics in the economies that grow rapidly. Direct taxation is growth-enhancing if an economy has either a slow or fast growth rate. Direct taxes are neutral at moderate growth rates.

One implication of the above results is that, in analyzing the fiscal policies which could act friendly to growth in the EU, using average fiscal multipliers could be of very little use. One needs to consider the different growth impacts in times of crises and normal times and to acknowledge the different ways in which the same policies can affect the growth rates in different countries. This rules out the use of a single fiscal/growth model for the EU economies.

REFERENCES

Agell, J., T. Lindle and H. Ohlsson (1997), "Growth and the Public Sector: A Critical Review Essay", *European Journal of Political Economy*, No. 13, pp. 33-52.

Agenor, R. and D. Yilmaz (2011), "The Tyranny of Rules: Fiscal Discipline, Productive Spending and Growth in a Perfect Foresight Model", *Journal of Policy Reform*, No. 14, pp. 69-99.

Angelopoulos, K., G. Economides and P. Kammas (2007), "Tax Spending Policies and Economic Growth: Theoretical Predictions and Evidence from the OECD", *European Journal of Political Economy*, No. 23, pp. 885-902.

Arnold, J. (2008), "Do Tax Structures Affect Aggregate Economic Growth? Empirical Evidence from a Panel of OECD Countries", OECD, Economics Department, Working Paper, No. 643.

Arnold, J. and C. Schwellnus (2008), "Do Corporate Taxes Reduce Productivity and Investment at the Firm Level? Cross-country Evidence from the Amadeus Dataset", OECD, Economics Department, Working Paper, No. 641.

Barro, R. (1990), "Government Spending in a Simple Model of Endogenous Growth", *Journal of Political Economy*, No. 98, pp. 103-25.

Bassanini, A. and S. Scarpetta (2001), "The Driving Forces of Economic Growth: Panel Data Evidence for the OECD Countries", OECD, Economic Studies, No. 10.

Basset, G. and R. Koenker (1978), "Regression Quantiles", *Econometrica*, No. 46, pp. 33-50.

Blankenau, W. and N. Simpson (2004), "Public Education Expenditures and Growth", *Journal of Development Economics*, No. 73, pp. 583-605.

Bleany, M., N. Gemmel and R. Kneller (2001), "Testing the Endogenous Growth Model: Public Expenditure, Taxation and Growth over the Long-run", *Canadian Journal of Economics*, No. 34, pp. 36-57.

Buchinsky, M. (1994), "Changes in U.S. Wage Structure 1963-1987: An Application of Quantile Regression", *Econometrica*, No. 62, pp. 405-58.

Chernozhukov, V. and C. Hansen (2006), "Instrumental Quantile Regression Inference for Structural and Treatment Effect Models", *Journal of Econometrics*, No. 132, pp. 491-525.

Devereux, M. and D. Love (1995), "The Dynamic Effects of Government Spending Policies in a Two-sector Endogenous Growth Model", *Journal of Money, Credit and Banking*, No. 27, pp. 232-56.

Dhint, T. and F. Heylen (2008), "Why Do Europeans Work (Much) Less? It Is Taxes and Government Spending", *Economic Enquiry*, No. 46, pp. 197-207.

Easterly, W. and S. Rebelo (1993), "Fiscal Policy and Economic Growth: An Empirical Investigation", *Journal of Monetary Economics*, No. 32, pp. 417-458.

Easterly, W., T. Irwin and L. Serven (2007), "Walking Up the Down Escalator: Public Investment and Fiscal Stability", *The World Bank Research Observer*, No. 23, pp. 37-56.

Eberhardt, M. and F. Teal (2011), "Econometrics for Grumblers: A New Look at the Literature on Cross-country Growth Empirics", *Journal of Economic Surveys*, No. 25, pp. 109-55.

Furceri, D. and G. Karras (2009), "Taxes and Growth in Europe", *South Eastern Europe Journal of Economics*, No. 7, pp. 181-204.

Gemmell, N., R. Kneller and I. Sanz (2011), "The Timing and Persistence of Fiscal Policy Impacts on Growth: Evidence from OECD Countries", *Economic Journal*, No. 121, pp. F33-F58.

Gosh, S. and A. Gregoriou (2008), "The Composition of Government Spending and Growth: Is Current or Capital Spending Better?", Oxford Economic Papers, No. 60, pp. 484-516.

Greiner, A. and W. Semmler (2000), "Endogenous Growth, Government Debt and Budgetary Regimes", *Journal of Macroeconomics*, No. 22, pp. 363-84.

Greiner, A., W. Semmler and G. Gong (2005), *The Forces of Economic Growth: A Time Series Perspective*, Princeton University Press.

Harberger, A. (1964a), "The Measurement of Waste", *American Economic Review*, Vol. LIV, pp. 58-76.

—————(1964b), "Taxation, Resource Allocation and Welfare", in NBER and the Brookings Institution (eds.), *The Role of Direct and Indirect Taxes in the Federal Reserve System*, Princeton University Press.

Kim, T.H. and C. Muller (2004), "Two-stage Quantile Regression when the First Stage is Based on Quantile Regression", *Econometrics Journal*, No. 7, p. 218-31.

Koenker, R. and V. d'Orey (1987), "Computing Regression Quantiles", *Applied Statistics*, No. 36, p. 383-93.

Lee, Y. and R. Gordon (2005), "Tax Structure and Economic Growth", *Journal of Public Economics*, No. 89, pp. 1027-43.

Linneman, L. and A. Schabert (2003), "Fiscal Policy and Economic Growth", *Journal of Money, Credit and Banking*, No. 35, pp. 911-29.

Lucas, R. (1990), "Supply-side Economics: An Analytical Review", *Oxford Economic Papers*, No. 42, pp. 293-316.

Mendoza, E.G., G.M. Milesi-Ferretti and P. Asea (1997), "On the Ineffectiveness of Tax Policy in Altering Long-run Growth: Harberger's Superneutrality Hypothesis", *Journal of Public Economics*, No. 66, 99-126.

Minea, A. and P. Villieu (2010), "Endogenous Growth, Government Debt and Budgetary Regimes: A Corrigendum", *Journal of Macroeconomics*, No. 32, pp. 709-11.

Myles, G. (2008), "How Do Taxes Affect Investment and Productivity? Industry Level Analysis of OECD Countries", OECD, Economics Department, Working Paper, No. 656.

Nikos, B. (2009), "Fiscal Policy and Economic Growth: Empirical Evidence from EU Countries", MPRA Paper, No. 19174, University Library of Munich, Germany.

Palivos, T. and C. Yip (1995), "Government Expenditure Financing in an Endogenous Growth Model: A Comparison", *Journal of Money, Credit and Banking*, No. 27, pp. 1159-78.

Pelloni, A. and R. Waldmann (2000), "Can Waste Improve Welfare?", *Journal of Public Economics*, No. 77, pp. 45-79.

Peretto, P. (2003), "Fiscal Policy and Long-run Growth in R&D-based Models with Endogenous Market Structure", *Journal of Economic Growth*, No. 8, pp. 325-47.

—————(2007), "Schumpeterian Growth with Productive Public Spending and Distortionary Taxation", *Review of Development Economics*, No. 11, pp. 699-722.

Pesaran, H., Y. Shin and R. Smith (1999), "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels", *Journal of the American Economic Association*, No. 94, pp. 621-34.

Powell, J. (1984), "Least Absolute Deviation Estimation for the Censored Regression Model", *Journal of Econometrics*, No. 25, pp. 303-25.

Rivas, L. (2003), "Income Taxes, Spending Composition and Long-run Growth", *European Economic Review*, No. 47, pp. 477-503.

Turnovsky, S. and W. Fisher (1995), "The Composition of Government Expenditure and Its Consequences for Macroeconomic Performance", *Journal of Economic Dynamic and Control*, No. 19, pp. 747-86.

Semmler, W., A. Greiner, B. Diallo, A. Rezai and A. Rajaran (2007), "Fiscal Policy, Public Expenditure Composition and Growth. Theory and Empirics", The World Bank, Policy Research, Working Paper, No. 4405.

Zagler, M. and G. Durnecker (2003), "Fiscal Policy and Economic Growth", *Journal of Economic Surveys*, No. 17, p. 397-418.

Zeng, J. and J. Zhang (2002), "Long-run Growth Effects of Taxation in a Non-scale Growth Model with Innovation", *Economics Letters*, No. 75, pp. 391-403.

AN ASSESSMENT OF STABILITY AND GROWTH PACT REFORM PROPOSALS IN A SMALL-SCALE MACRO FRAMEWORK

Jérôme Creel[*]*, Paul Hubert*[**] *and Francesco Saraceno*[**]

This paper contributes to the debate on fiscal governance for the European Monetary Union. We simulate a small scale macroeconomic model with forward looking agents, augmented with a public finance block. We account for both positive (output stabilization) and negative (via risk premia) effects of debt and deficit. By the appropriate choice of the exogenous fiscal variables in the fiscal block, we replicate the working of the rule embedded in the so-called "fiscal compact". We compare this rule with the Maastricht 3 per cent deficit limit (status quo), and with an "investment" rule leaving room for public investment. We evaluate the performance in terms of output and inflation during a fiscal consolidation, as well as following demand and supply shocks at the steady state. All rules guarantee long run sustainability. The investment rule emerges robustly as the one guaranteeing the lowest output loss, followed by the status quo. The "fiscal compact" rule appears to be the most recessionary and deflationary.

1 Introduction

This paper assesses the macroeconomic impact of a number of fiscal rules that have been or could be implemented by countries belonging to the European Union. The European fiscal crisis, and the ensuing requirement to reduce public debt levels, paved the way for a set of reforms of the European fiscal rules. On March 2nd 2012, 25 of the 27 EU countries (the UK and the Czech Republic did not sign) adopted the *Treaty on Stability, Coordination and Governance in the Economic and Monetary Union,* that is currently (November 2012) under ratification. This so-called "Fiscal Compact" complemented the provisions of the Maastricht Treaty and of the Stability and Growth Pact (SGP): the limit of public deficit at 3 per cent of GDP has been supplemented with a limit on structural deficit at 0.5 per cent of GDP, and an average yearly reduction by 1/20th of the difference between the debt to GDP ratio and the 60 per cent of GDP Maastricht limit. The limit on structural deficit goes beyond the 3 per cent Maastricht provision, in that it aims at introducing balanced budget constraints at the Constitutional level of each euro zone member state.

It is somewhat paradoxical that rules aimed at constraining the capacity of governments to run countercyclical policies are discussed precisely after the worldwide financial crisis required large public deficits to dampen shocks ensuing from market failures. Thus, the introduction of the above-mentioned rules raises the question of their incidence on the usual objectives of economic policies, namely the output gap and the inflation rate.

---

[*] OFCE/Sciences Po and ESCP Europe.

[**] OFCE/Sciences Po.

Corresponding author: francesco.saraceno@sciences-po.fr. Address: OFCE, 69 Quai d'Orsay, 75007 Paris, France.

A first version of this paper was presented at the International POLHIA Conference *Rethinking Economic Policies in a Landscape of Heterogeneous Agents*, October 13-15, 2011, Catholic University, Milan, Italy. We thank the audience for useful comments as well as participants at conferences and seminars at AFSE (Paris), Euroframe (Kiel), OFCE (Paris), CRIEF (Poitiers), the Developments in Economic Theory and Policy Conference (Bilbao), University of Grenoble and the Villa Mondragone International Economic Seminar (Rome). This project has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under Socio-economic Sciences and Humanities, grant agreement no. 225408 (POLHIA). We are also grateful to the referees and the editor of this Journal for their comments which helped improve the paper. The remaining errors are our own responsibility.

The contribution of this paper is to simulate the macroeconomic effects of the adoption of these rules in a structural small scale New-Keynesian model, in which we introduce a public finance block and a yield curve embedding risk premia. We aim at shifting the attention back from the objective of *fiscal* stabilization to the one of *macroeconomic* stabilization. As the proposed rules stand, public deficit and debt are not instruments to smooth the cycle. European authorities – governments, the ECB, or the Commission – seem to consider them as objectives of policy action rather than what they should be, namely *instruments* for obtaining the final objective of stabilizing output gap and inflation. This reversal of targets and instruments is equivalent to *a priori* denying any role to macroeconomic (in particular fiscal) policy. With this exercise we intend to account for the negative impact of excessive deficit and debt, while emphasizing their role as instruments for attaining the *final objective* of aggregate welfare maximization.

Medium or large scale New-Keynesian models have often been used to assess the impact of fiscal policy on real GDP and inflation rates. Coenen *et al*. (2012) for instance review the fiscal properties of nine dynamic stochastic general equilibrium (DSGE) models in which Keynesian features like price and wage rigidities are introduced. Most models use rule-of-thumb fiscal rules by which taxes respond to deficits or debts (as in the seminal specification of Barro, 1986). Hence, public finance sustainability is always met. In this study and in contrast with these models, we specify the fiscal rules which governments have to abide by.

While in a number of occasions fiscal rules have not been respected, we assume governments to follow the fiscal rules which have been decided at the EU level. We wish to investigate the real consequences of sticking to the rules. We assess these consequences under two different assumptions regarding the initial levels of public deficits and debts. In the first scenario – initial deficits and debts are at their current level, *i.e.*, above their steady-state values – we evaluate the effect of fiscal consolidation under the regimes corresponding to each fiscal rule. In the second scenario we assume that the economy starts at steady-state, and we investigate in our small-scale model the different consequences of supply or demand shocks under the different fiscal rule regimes. Most standard DSGE models deal with the second scenario, while the scenario of fiscal consolidation is specific to our paper. In both cases, our value-added stems from the comparison of the specific EU fiscal rules.

We assess the macroeconomic impact of the fiscal rules on four economies that we take as representative of the euro zone: a large (relatively) low-debt economy (France), a small high-debt one (Belgium), a large high-debt one (Italy) and a small-low debt one (the Netherlands). The size of nations – large or small – relates to the size of their fiscal multiplier. The four countries also differ in terms of the size and sign of their primary structural balance: France and the Netherlands have a large deficit, whereas Belgium has a low one, and Italy holds a large surplus.

We simulate the effect of the rules on the level and variability of the output gap, the inflation rate and the structural deficit, and the impact on the level of public debt. This is done in a framework in which on the one hand, the evolution of deficit is countercyclical, but on the other hand, excessive debt feeds back into the economy through increasing risk premia. Among the nine large-scale DSGE models reported in Coenen *et al.* (2012), only one, the, European Commission's QUEST III, embeds a government debt risk premium. Finally, we simulate the different rules over a 20-year horizon, consistently with the target of the one twentieth debt reduction rule.

The rules we simulate are (a) the Fiscal Compact, with its balanced (at 0.5 per cent of GDP) structural budget and the 1/20th yearly debt reduction rule; (b) the 3 per cent total deficit cap (status quo). We also evaluate the effect of (c) adopting an investment rule in the vein of the UK golden rule of public finances, that imposes budget balance over the cycle only for current spending, while allowing public investment to be financed through debt.[1] The simulations are carried out starting from a structural New-Keynesian model, where the IS and Phillips curves have hybrid specifications with backward and forward expectation terms. Moreover, our specification of the economy also takes into account the nonlinearity of the risk premium and the zero lower bound.

It is worth emphasizing that the macroeconomic framework is partly biased against the use of an investment rule, because we rule out the endogeneity of potential output, which could be positively affected by public investment. On the opposite, the investigation is partly biased in favor of the Fiscal Compact because we simulate the less restrictive rule among the two embedded in the Fiscal Compact.

Results are manifold. First, the adoption of the rules produces a short-run recession, even in a small country with a low multiplier and relatively low initial public debt like the Netherlands. Second, recessions sometimes foster deflation. Although we do not model deflation differently from inflation in this framework, the former is very difficult to reverse in presence of a binding fiscal constraint and of a zero lower bound for the interest rate (Woodford, 2001). Third, the investment rule performs better than the other two rules: recessions are shorter and milder; hence the average loss of output over a 20-year horizon is smaller, all the more so when the fiscal multiplier is large. Fourth, this result is strongly robust to changes in the parameters' values. Fifth, when the economy is hit by demand and supply shocks at the steady state, none of the rules emerges as superior in coping with them.

The paper is structured as follows: Section 2 introduces and discusses our model. In Section 3 we simulate the two scenarios of a fiscal consolidation and of different types of shocks starting from the steady state, and give a quantitative assessment of the macroeconomic performance for the different fiscal rules. Section 4 is devoted to a robustness check of the main results, and Section 5 concludes.

2 An augmented New-Keynesian model

The economy is characterized by a standard framework with the aggregate demand side described by a dynamic IS curve and the aggregate supply side by a hybrid Phillips curve, in the vein of Clarida *et al.* (1999). By hybrid, we mean that expectations are forward and backward-looking. In order to study the different fiscal rules, we add to this core a public finance block to simulate the differences between the rules. To take into account the effect of debt and deficit on private agents' behavior, we explicitly model the equations for government and central bank interest rates.

2.1 The model

The AD bloc is described by a dynamic hybrid IS curve, detailing the determinants of the output gap x_t, *i.e.*, the percentage difference between real GDP (y) and potential GDP ($\bar{y}$):

$$x_t = \alpha_1 \cdot x_{t-1} + (1-\alpha_1) \cdot E_t x_{t+1} + \alpha_2 \cdot (r_t - E_t \pi_{t+1} - rr^*) + \alpha_3 \cdot (dsp_t - \overline{dsp}) + \varepsilon_t^d \quad (1)$$

where α_1 stands for the incidence of backward-expectations on demand behavior, r_t and π_t are the nominal long-term interest rate and the inflation rate respectively, both in percent; rr^* is the long-term real interest rate in percent; dsp_t is structural primary balance (*i.e.*, deficit net of interest

---

[1] Introduced in the 1997, the UK golden rule of public finance excludes public investment from the budget limits over the cycle. See Buiter (2001) for a comprehensive discussion.

payments and of cyclical components) as a percentage of GDP, and we define as the fiscal impulse, or fiscal stimulus, its deviation from the steady-state value *(dsp-* $\overline{dsp}$*)*. α_2<0 and α_3>0 are parameters.

The introduction of the fiscal impulse in the expectational IS curve stems from the linearised Euler equation of a closed economy with consumption and government expenditure. Considering a simplified budget constraint for the government, with transfers and proportional taxation, public expenditure has to be considered as net of cyclical components and interest payments; its deviation from steady state is therefore captured by the deviation of the structural primary balance.[2] The introduction of backward-looking expectations in the Euler equation, on the other hand, has an empirical justification (see, e.g., Fuhrer and Rudebusch, 2004).

The aggregate supply block is represented by a standard hybrid Phillips curve, where λ_1 captures the incidence of backward-expectations on supply behavior; λ_2 is the elasticity of inflation to the output gap and is a positive parameter:

$$\pi_t = \lambda_1 \cdot \pi_{t-1} + (1-\lambda_1) \cdot E_t \pi_{t+1} + \lambda_2 \cdot x_t + \varepsilon_t^s \quad (2)$$

The third equation describes the behavior of nominal government bonds' interest rates r_t along the yield curve, where i stands for central bank nominal interest rate, and γ represents the risk premium associated with upwards debt variation over the target b^*:

$$1 + r_t = (1 + i_t) \cdot [1 + (\gamma \cdot \max(0, b_t - b^*)] + \varepsilon_t^f \quad (3)$$

Monetary policy is described through a usual Taylor rule. The central banker sets the nominal interest rate in response to expected future inflation and current output gap. We explicitly introduce a close-to-zero bound on the nominal rate (here at 0.25 per cent):

$$i_t = \max(0.25,\ rr^* + E_t \pi_{t+1} + \Phi_1 \cdot (E_t \pi_{t+1} - \pi^*) + \Phi_2 \cdot x_t + \varepsilon_t^m) \quad (4)$$

In equations (1) to (4) the error terms ε capture exogenous shocks. Hence ε^d and ε^s represent a demand and a supply shock respectively.

We develop the public finance block to enable the introduction of different fiscal rules in the model. Total government deficit can be decomposed into a cyclical component and a structural component, all expressed as a percentage of GDP:

$$dt_t \equiv dc_t + ds_t \quad (5)$$

As commonly assumed in the literature (see, e.g., Buti *et al.*, 1998; and Girouard and André, 2005), the cyclical component, or cyclical deficit, depends linearly on the output gap, hence characterizing automatic stabilizers:

$$dc_t = \psi_1 \cdot x_t \quad (6)$$

The structural deficit is by construction the sum of interest payments ip and structural primary deficit dsp, interpreted as the discretionary part of fiscal policy:

$$ds_t = ip_t + dsp_t \quad (7)$$

[2] The government budget constraint can be written as $T(Y_t) + P_t G_t + (1 + i_{t-1}) B_t = \tau Y_t + B_{t-1}$, where we interpret T and τY as the components of public deficit related to automatic stabilizers (net transfers, affected by the business cycle, and proportional taxation). Under the assumption that the tax rate is given, there is no discretionary tax policy. Thus, G is the (real) amount of discretionary expenditure of the government, net of net transfers and interest payments, and its variation translates into variations of the primary structural balance.

Public debt, expressed in percentage of GDP, follows the usual law of motion, where everything else equal, a higher nominal growth rate mechanically reduces the debt to GDP ratio:

$$b_t = \frac{b_{t-1}}{1+\pi_t + x_t + \overline{y}} + dt_t \tag{8}$$

2.2 *Fiscal rules*

The medium-to-long term performance of European economies depends on the macroeconomic governance tools put in place by the EU. Three main options are before policy makers: (a) a *status quo* where the ratio of public deficit to GDP must be maintained below the 3 per cent limit. (b) The "investment rule" that allows to finance an increase in net public assets by public debt issuance. (c) The "Fiscal Compact", that embeds the double requirement of a balanced (at 0.5 per cent of GDP) structural deficit and a constant rate of reduction of debt bringing it asymptotically to the 60 per cent-of-GDP ratio (*i.e.*, a 5 per cent reduction per year of the difference between the current debt and its reference level). These rules differ on the criteria and on the type of constraints imposed to countries. Specifically, each rule imposes different constraints on the choice of endogenous and exogenous variables in the fiscal block of the model.

a) For the status quo, we impose that total deficit is exogenously given at 3 per cent:

$$dt = 3$$

We assume in other words that countries use the entire margin given by the rule, and never breach it. The other fiscal variables adapt to this exogenous constraint.

b) The "investment rule" allows increasing public investment inv^g, expressed in percentage of GDP, up to a threshold equal to the inflation depreciation of steady-state debt. Thus, all else equal, the "investment rule" keeps the debt-to-GDP ratio constant. Higher investment may produce higher net interest charges; the rule forces the government to compensate them with a lower cyclically-adjusted primary deficit, *i.e.*, with lower current expenditures $dcur$, also expressed in percentage of GDP. The "investment rule" is described as follows:

$$dsp = inv^g + dcur$$

$$inv^g = \pi \, \overline{b}$$

$$dcur = -\,\delta\cdot(inv^g + (ip - \overline{ip})) + (1 - \delta)\cdot dcur_{t-1}$$

where the last equation assumes that the current surplus needed to finance interest payments is spread over a $1/\delta$ year period. δ represents the smoothing of expenditure over future periods, and it may have a strong impact on the restrictiveness of the rule. In the benchmark simulations below, we set δ=1 with all the cost of consolidation borne in the current period, whereas in alternative simulations, we show the effect of setting δ equal to 0.5, 0.2 and 0.1 (spreading over 2, 5 and 10 years respectively). Note that this is a severe version of the rule, first because investment is accepted only up to the limit that keeps the debt ratio on a stationary path; second, because public investment has no impact on potential growth (that we assume exogenous and constant), so that it is analytically equivalent to current spending. This puts us in a "worst-case scenario", in which we artificially shut off the long-run positive effects of the investment rule.

c) The Fiscal Compact has two arms. As regards the debt reduction advocated by the Treaty, the exogenous variable is the yearly change in the debt ratio, supposed to be reduced each year by 5 per cent of the difference with its reference rate (60 per cent). In order to simulate this rule, we need to make three assumptions, not explicit in the Treaty. First, we assume the rule to be symmetric around its reference level of 60 per cent; second, we assume it to be asymptotic, as debt is reduced of 5 per cent of the difference between the ratio in the previous period and the reference level.[3] Finally, we assume that the debt reduction is net of the cyclical balance. Taken together, these three assumptions allow convergence to the Maastricht steady state. Moreover, the third assumption designs a mild version of the debt reduction rule, which minimizes its recessionary impact. Hence, the one twentieth rule runs as follows:

$$ds_t = -0.05 \cdot (b_{t-1} - \overline{b}) + (\pi_t + x_t + \overline{y}) \cdot b_t$$

$$\Rightarrow$$

$$dsp_t = -0.05 \cdot (b_{t-1} - \overline{b}) + (\pi_t + x_t + \overline{y}) \cdot b_t - ip_t$$

The structural balance can be decomposed into the surplus needed to reduce debt by one twentieth of its difference to its steady state value, and the room for maneuver obtained from debt depreciation.

The second arm of the Fiscal Compact concerns the limit to structural deficit. The Treaty states that general government budgets shall be balanced or in surplus, a criterion that "*shall be deemed to be respected if the annual structural balance of the general government is at its country-specific medium-term objective, as defined in the revised Stability and Growth Pact, with a lower limit of a structural deficit of 0.5 per cent of the gross domestic product at market prices*". This amounts to simulating the model with structural deficit exogenously constrained at $ds_t = 0.5$.

The Fiscal Compact implicitly assumes that once the 60 per cent debt threshold is attained the structural balance rule becomes binding. This would imply that the debt ratio keeps decreasing until it stabilizes at 10.5 per cent of GDP, converging to a steady state different from the other rules. More substantially, whether the one twentieth rule or the structural balance rule is more binding depends on the nominal growth of the economy and the level of debt. If the nominal growth rate g (with $g = x_t + \overline{y} + \pi$) is above 5 per cent, then the structural balance rule is always more restrictive. If the nominal growth rate g is below 5 per cent, the level of debt under which the structural balance rule is more restrictive is 120 per cent of GDP for g=3%, 82 for g=2% and 50 for g=0%. Therefore by deciding to focus on the one twentieth rule which is consistent with the Maastricht steady-state, we can reasonably argue that countries follow the least restrictive arm of the Fiscal Compact.

2.3 The steady-state

We use a Newton algorithm to compute the simultaneous solution for the equations of the model for every period, and compute a numerical simulation of the trajectory of the model's solution. The solution technique is described in Juillard (1996).

The model has a steady state with a potential real growth rate y^* of the economy exogenously set at 3 per cent, in accordance with the underlying hypotheses of the European Union Treaty. The real natural interest rate rr^* also equals 3 per cent, the debt target b^* is 60 per cent and the inflation target π^* is 2 per cent, for a nominal growth rate in steady state of 5 per cent. At the steady-state, public deficit is therefore equal to interest payments ($\overline{dt} = \overline{ip} = 3\%$), and primary structural balance is achieved ($\overline{dsp} = 0$).

[3] The letter of the Treaty is ambiguous, (TSCG, 2012; and Whelan, 2012) and it is usually associated with the requirement to reach the level of 60 per cent in 20 years. Nevertheless discussions with Commission officials and economists lead to interpret the rule as asymptotic convergence.

Table 1

Steady State Values for Endogenous Variables

$\overline{x}$	0
$\overline{\pi}$	π^*
$\overline{r}$	$rr^* + \pi^*$
$\overline{i}$	$rr^* + \pi^*$
$\overline{dt}$	3
$\overline{dc}$	0
$\overline{ds}$	3
$\overline{ip}$	3
$\overline{dsp}$	0
$\overline{b}$	b^*

The three fiscal rules that we assess make the economy converge to the Maastricht steady state, both in the scenario of fiscal consolidation from current debt and deficit levels, and in the scenario of an economy at steady state which is hit by supply and demand shocks.

2.4 Calibration

The output gap and inflation rate in the expectational IS and Phillips curve equations are introduced with both forward and backward components (α_1 = 0.4 and λ_1 = 0.5). For the IS-augmented curve, this seems to be a reasonable hypothesis considering the average results by Fuhrer and Rudebusch (2004) over a wide range of estimations. Estimations by Goodhart and Hofmann (2005), however, point to a relatively lower incidence of forward-looking expectations for the US and Euro area economies, which would put α_1 in the range of [0.2, 0.4]. The parameters of the expectations-augmented-Phillips curve are more controversial (and estimations are more numerous). Galí *et al.* (2005) and Goodhart and Hofmann (2005) find that the coefficient on lagged inflation is rather modest (around 0.2-0.3). Rudd and Whelan (2006), on the contrary, conclude that the forward-looking component is not significant, and a recent evaluation drawing on survey-based expectations concludes that the hybrid Phillips curve (with a backward component) outperforms the New-Keynesian Phillips curve with no inflation persistence, finding that the forward-looking coefficient is close to $\lambda_1 = 0.5$ (see Paloviita, 2008). We decide to follow this road, which is agnostic with respect to a debate that is yet unsettled.

Table 2 reports the parameters in the simulations. The coefficient value of the incidence of the output gap in the hybrid Phillips curve is close to Paloviita's (2008) estimate. The parameters in the monetary rule are taken from Taylor (1993). The targets are consistent with the Maastricht Treaty and the Stability and Growth Pact's requirements, and with the model's steady state. We introduce two different values for the coefficient of the fiscal impulse in the expectational IS equation, in order to take into account the larger external leakage of domestic fiscal policy in a small open economy. It is worth noticing that even for large countries the fiscal multiplier in this calibration is significantly smaller than recent estimates (e.g., IMF, 2012) and is in line with the modeling literature for the euro zone (Smets and Wouters, 2003; Dieppe *et al.*, 2005; Adolfson *et al.*, 2007; Coenen *et al.*, 2008; Christoffel *et al.*, 2009; Ratto *et al.*, 2009; Cogan *et al.*, 2010; Gelain, 2010; and Cwik and Wieland, 2011). Our choice of the fiscal multiplier is well below the value that risks triggering the vicious circle of austerity and economic contraction that some European peripheral countries have been experiencing since early 2010s. In other words, our estimates of the output cost of fiscal consolidation, are based on a conservative fiscal multiplier, and therefore can be interpreted as a lower bound.

The theoretical and empirical uncertainty about many of these parameters (especially α_1 and λ_1) requires thorough robustness checks. The results of Monte Carlo simulations are reported in Section 4.

Table 2

Calibration Parameter Values

α_1	0.4
α_2	–0.2
α_3	0.8 (large country / 0.2 (small country)
λ_1	0.5
λ_2	0.2
γ	0.02
Φ_1	0.5
Φ_2	0.5
Δ	1
ψ_1	–0.5
y^*	3%
r^*	3%
b^*	60%
π^*	2%
discount rate	0.95 [=1/1.05]

3 Simulations

To our knowledge, there are very few examples of papers attempting at the evaluation of different fiscal rules in the EU context. Most recent papers dealing with this issue focus on one type of rule, like an expenditure rule (e.g., Hauptmeier *et al.*, 2011), whereas those which study different rules use the classification by Kopits and Symansky (1998) (see, e.g., Creel and Saraceno, 2010; and Schuknecht *et al.*, 2011). In contrast, Creel *et al.* (2012) performed a comparison between various fiscal rules within a simple estimation exercise in the vein of Eichengreen and Wyplosz (1998) and Monperrus-Veroni and Saraceno (2005). These exercises start from a simple reduced form VAR system and the estimation results are the basis for a counterfactual assessment of the effect of alternative fiscal rules. While not exempt from a number of methodological problems, the paper by Eichengreen and Wyplosz and the followers using a similar methodology retained a remarkable interest because they give a measure of the magnitude of costs and benefits of the SGP and of other rules. Our analysis completes these results; instead of relying on an estimated model, it builds on a theoretical model, and the differences among countries are given by the value of the fiscal multiplier in the output gap equation and by the initial conditions of public finance variables.

We first discuss the application of the different fiscal rules to a consolidation occurring in the four countries starting from current conditions; and then we examine the case of supply and demand shocks hitting an economy at the Maastricht steady state.

3.1 Fiscal consolidation

The economy starts from 2011 levels of deficit and debt, and is tracked for a time span of 20 years. We decided to focus on fiscal consolidation abstracting from the *initial* size of the output gap and inflation which, as a consequence, in the simulations are set at their steady state values

(0 for the output gap and the 2 per cent central bank target for inflation).[4] Initial debts and deficits for the four countries under study are 2011 OECD figures. They are reported in Table 3 below. France and Italy are larger countries than Belgium and the Netherlands; hence, by assumption, the fiscal multiplier is equal to 0.8 for the former and 0.2 for the latter.

Figures 1 and 2 show output gap and inflation, together with interest rates and the public finance variables, for France. The figures for the other countries are qualitatively similar and are presented in the Appendix.

Table 3

Initial Debt and Deficit Values, 2011

Country	Initial Debt	Initial Structural Primary Deficit	Fiscal Multiplier*
France	86	1.45	0.8
Italy	120	–2.34	0.8
Belgium	98	0.78	0.2
Netherlands	65	2.53	0.2

* Authors' assumption.
Source: OECD.

The economy starts outside the steady state equilibrium to capture the effects of a fiscal consolidation. The initial impulse stems from how fiscal rules applied in period one constrain the primary structural deficit which therefore impacts the economy. For instance, in the case of the status quo, the initial impulse brings total deficit back to 3 per cent of GDP at period one when the rule is set up. Before discussing the outcome of each rule, it is worth pointing out two things. First, all the rules yield long run convergence of output gap, inflation, and public finance variables, towards their steady state levels. Furthermore, debt dynamics are comparable: the debt ratio steadily decreases albeit at different rates. The second feature that is common to all the rules is the deep recession induced by fiscal consolidation in the short run, which may even be deflationary and results in a sharp drop of interest rates.

Looking at the rules in detail, the Fiscal Compact yields the larger initial drop of output (Figure 1, upper panel), which causes deflation in the medium run (lower panel). The status quo's output drop is larger than the one of the investment rule, whereas inflation dynamics are quite similar for these two rules. On the other hand, the long run reduction of debt is more substantial with the Fiscal Compact than for the other rules (Figure 2, upper-left panel). The central bank interest rate drops below two percent, and as a consequence interest payments are lower than in the two other rules. This in turn yields faster debt reduction in the medium to long run.

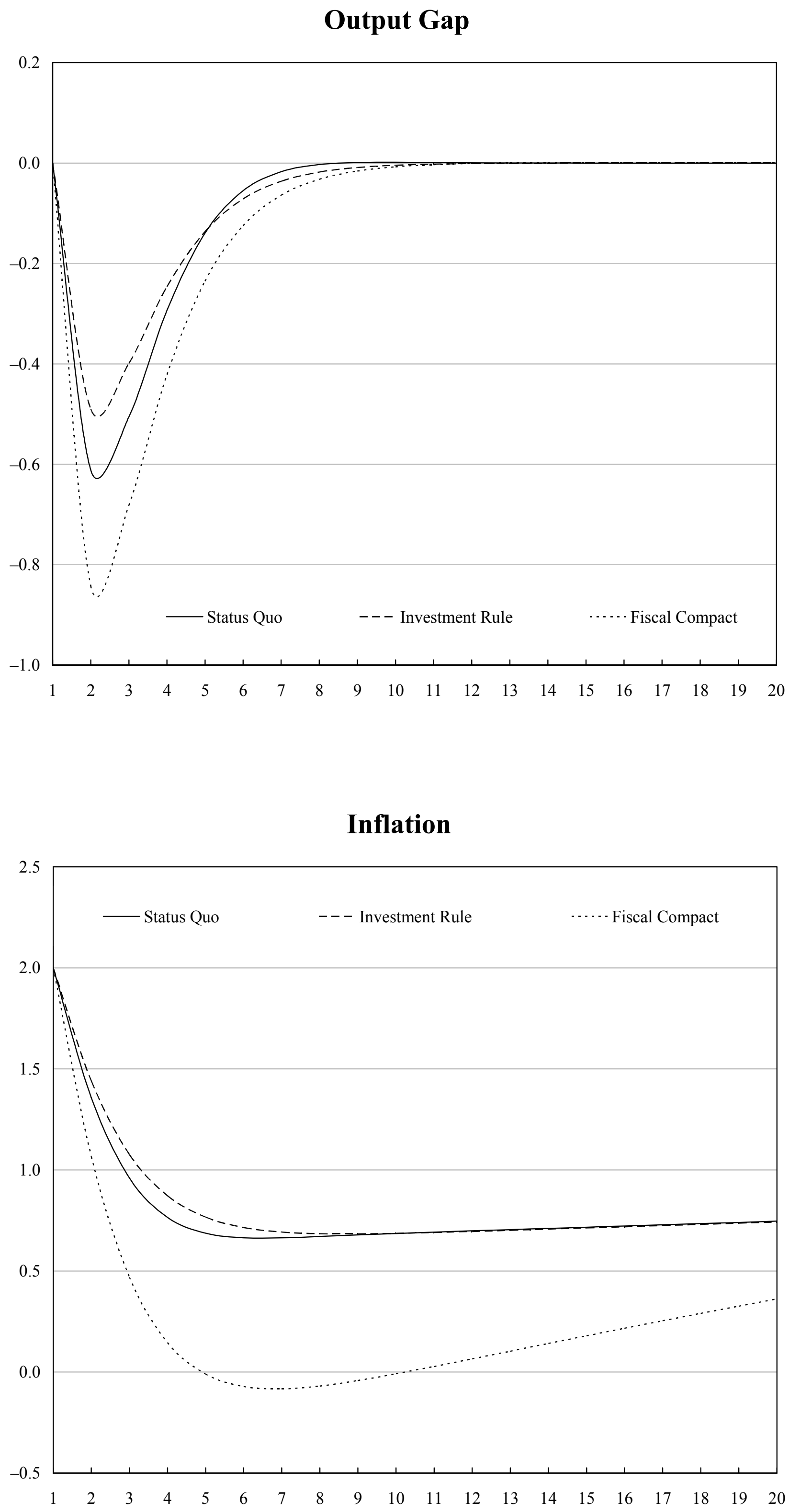
To compare the different rules, we computed for each country (*i.e.*, with different initial public finances values) the average of the discounted variables of interest (assuming a discount rate of 5 per cent). They are reported in Table 4. The table shows that for the four countries the average

---

4 If we began with the current values of the (negative) output gap and inflation, the initial drop of output would be larger, and the interest rate would hit the zero lower bound earlier.

Figure 1

Fiscal Consolidation Under Different Rules: France



loss of output is lower in the case of the investment rule. For the small countries the status quo also minimizes discounted cumulative loss, which can be explained by the assumption of a smaller fiscal multiplier. In addition, the investment rule is associated with lower output variability for all countries except the Netherlands. As can be guessed from Figures 1 and 2, this can most probably be attributed to the lesser recessionary impact in the early phase of the consolidation process. In all cases, the visual impression of Figure 1 for France is confirmed, and the Fiscal Compact fares worse than the other rules.

As regards inflation, the investment rule yields a lower inflation gap to its target on average, and the status quo exhibits lower variability. As expected, on the other hand, the Fiscal Compact yields substantially lower debt levels at t=20. One additional remark refers to the application of the Fiscal Compact in Italy. Setting the γ parameter on the risk premium in the government bonds' interest rates equation to 0.02 as for other simulations prevents the economy to converge back to the steady-state, possibly because of the high initial level of debt. The

Figure 2

Fiscal Consolidation Under Different Rules: France (Fiscal Variables)

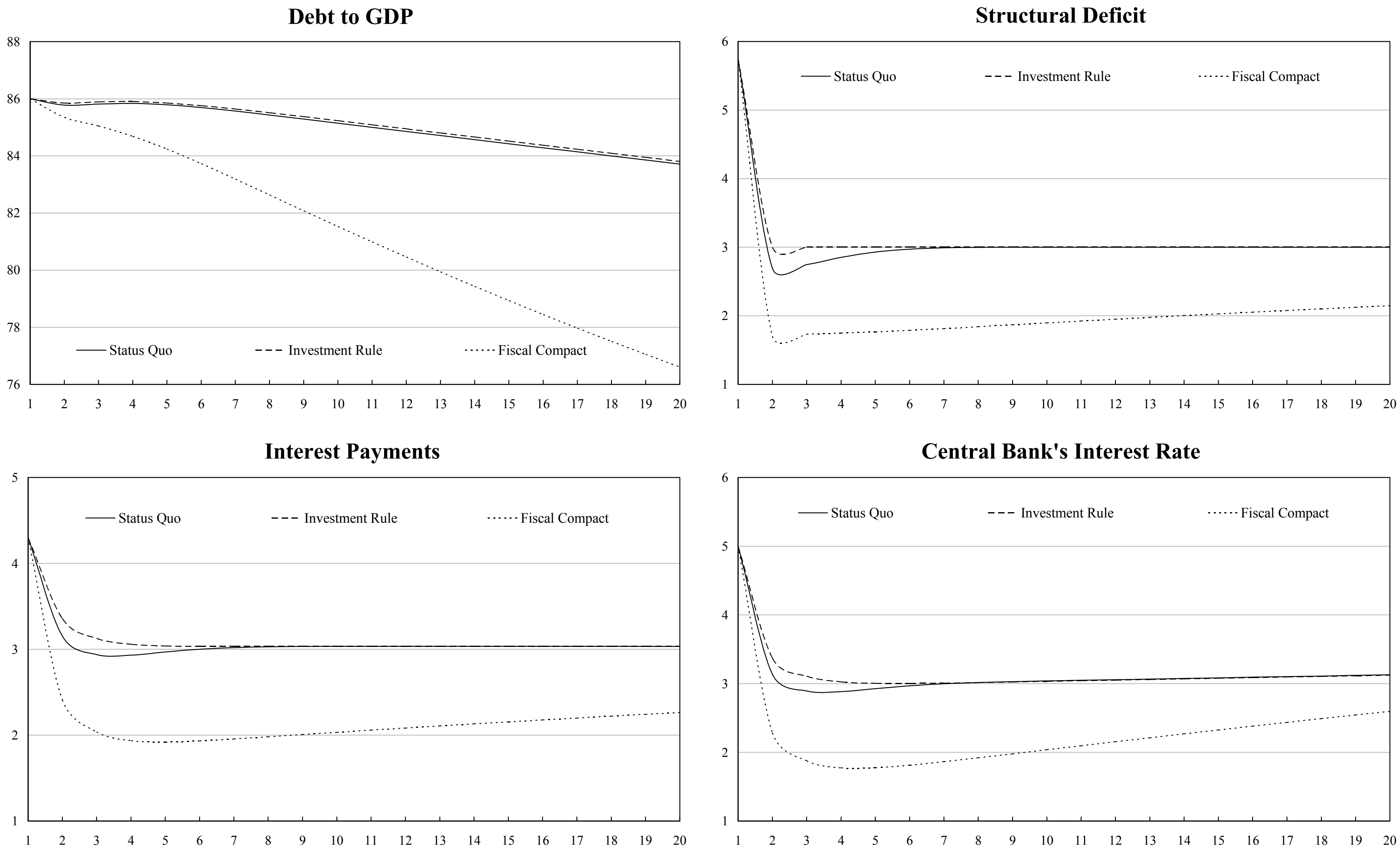


Table 4

Discounted Average Values of the Rules for 20 Years

France

	Status Quo	Inv. Rule	Fiscal Compact
mean(x)	–0.07	–0.06	–0.10
s.d.(x)	0.16	0.13	0.22
mean(π)	0.54	0.56	0.20
s.d.(π)	0.39	0.40	0.46
mean(ds)	1.97	2.00	1.36
s.d.(ds)	0.93	0.95	0.99
b (t=20)	83.71	83.80	76.60

Italy

	Status Quo	Inv. Rule	Fiscal Compact*
mean(x)	–0.13	–0.11	–0.15
s.d.(x)	0.30	0.25	0.32
mean(π)	–0.09	–0.07	–0.18
s.d.(π)	0.53	0.53	0.57
mean(ds)	1.84	1.90	0.55
s.d.(ds)	0.55	0.61	0.71
b (t=20)	120.59	120.75	90.21

Belgium

	Status Quo	Inv. Rule	Fiscal Compact
mean(x)	–0.07	–0.07	–0.10
s.d.(x)	0.13	0.12	0.19
mean(π)	0.38	0.38	0.11
s.d.(π)	0.44	0.45	0.50
mean(ds)	1.96	2.00	1.13
s.d.(ds)	0.92	0.94	1.01
b (t=20)	93.83	94.20	80.40

Netherlands

	Status Quo	Inv. Rule	Fiscal Compact
mean(x)	–0.01	–0.01	–0.02
s.d.(x)	0.02	0.02	0.03
mean(π)	1.11	1.11	1.07
s.d.(π)	0.35	0.36	0.34
mean(ds)	2.00	2.00	1.89
s.d.(ds)	0.95	0.96	0.95
b (t=20)	63.86	63.92	62.30

* The fiscal compact enables convergence back to the steady-state in Italy only if gamma = 0. The simple average values are presented in Table 9 in the Appendix.

convergence in the Italian case thus required to set γ to zero. Even in this case, with no market penalty for large debt, the Fiscal Compact yields a larger output loss than alternative rules.

Setting aside the investment rule, which is currently not an option in the policy debate, we can observe that the status quo performs considerably better than the 5 per cent debt reduction rule in terms of macroeconomic performance.

To conclude, for all possible initial situations (large and small countries; high and low initial debt), the model yields the unequivocal result that implementing the investment rule would minimize the average loss of output, and would also prove less deflationary than the different EU fiscal rules. Among these, the status quo is largely to be preferred if we use the output gap as a metrics, while the debt reduction rule is less inflationary and yields faster debt reduction. The simulations show that relatively larger structural deficits are not necessarily inconsistent with output stabilization and public finances sustainability. Because of depressed growth, debt ratios may actually decrease less than actually planned during fiscal consolidation.

Table 5

Response to Demand and Supply Shocks Starting from Steady State
(average discounted values over 20 years)

Small Countries – Fiscal Multiplier = 0.2

Negative Demand Shock			
	Status Quo	**Inv. Rule**	**Fiscal Compact**
mean(x)	–0.05	–0.05	–0.05
s.d.(x)	0.21	0.19	0.19
mean(π)	1.14	1.13	1.12
s.d.(π)	0.29	0.29	0.29
mean(ds)	1.84	1.87	1.83
s.d.(ds)	0.50	0.54	0.53
b (t=20)	61.27	61.53	60.93

Positive Supply Shock			
	Status Quo	**Inv. Rule**	**Fiscal Compact**
mean(x)	0.07	0.07	0.07
s.d.(x)	0.14	0.12	0.12
mean(π)	1.12	1.11	1.11
s.d.(π)	0.30	0.29	0.29
mean(ds)	1.91	1.87	1.87
s.d.(ds)	0.59	0.54	0.53
b (t=20)	60.39	60.03	60.01

Large Countries – Fiscal Multiplier = 0.8

Negative Demand Shock			
	Status Quo	**Inv. Rule**	**Fiscal Compact**
mean(x)	–0.04	–0.03	–0.03
s.d.(x)	0.19	0.14	0.13
mean(π)	1.17	1.18	1.16
s.d.(π)	0.31	0.32	0.32
mean(ds)	1.85	1.87	1.85
s.d.(ds)	0.51	0.54	0.53
b (t=20)	60.90	60.92	60.65

Positive Supply Shock			
	Status Quo	**Inv. Rule**	**Fiscal Compact**
mean(x)	0.10	0.08	0.08
s.d.(x)	0.23	0.17	0.17
mean(π)	1.19	1.15	1.16
s.d.(π)	0.31	0.30	0.30
mean(ds)	1.92	1.87	1.88
s.d.(ds)	0.61	0.54	0.53
b (t=20)	59.62	59.50	59.64

3.2 Supply and demand shocks at the steady-state

The previous section dealt with the performance of the different rules during a fiscal consolidation process, starting from high debt ratios. Our next question is how these rules would affect the dynamics of the economy if it were hit by a demand shock (in the output gap equation) or by a supply shock (in the Phillips curve equation) when at the steady state. Both shocks are temporary shocks with the value of each exogenous variable namely ε^d and ε^s being equal to minus one during one period. The results are summarized in Table 5, where we distinguish between "small" countries (with a low fiscal multiplier) and "large" ones (with a large fiscal multiplier).

The table shows first that the differences between the fiscal rules are very marginal, a result that is not surprising given that we are studying adjustments close to the steady state. When the economy is hit by demand and supply shocks at the steady state, none of the rules emerges as

Table 6

Smoothing Over the Business Cycle

Investment Rule				
Fiscal Consolidation – France				
	1y	**2y**	**5y**	**10y**
mean(x)	–0.06	–0.04	0.01	0.09
s.d.(x)	0.13	0.08	0.21	0.55
mean(π)	0.56	0.57	0.63	0.74
s.d.(π)	0.40	0.43	0.52	0.80
mean(ds)	2.00	2.03	2.14	2.32
s.d.(ds)	0.95	1.00	1.23	1.84
b (t=20)	83.80	83.76	83.53	82.99

Average discounted values over 20 years. δ equals 0.5, 0.2 and 0.1 for 2, 5 and 10 years respectively.

superior. The status quo seems slightly worse than the two others regarding the variance of output, while the Fiscal Compact appears better, at the margin, regarding the debt level. The reader should bear in mind, however, that as this is a rule designed to come back to the debt reference level of 60 per cent of GDP, its rationale seems quite limited for policymakers when debt is close to the steady state.

Although not surprising, the outcome of these simulations is important. Two interpretations are possible. On the one hand, the Maastricht rule – the status quo – is not worse than alternative rules, which vindicates the claim that in normal times this rule gives sufficient fiscal margins for maneuver (see, e.g., Buti and Giudice, 2002). On the other hand, the rule is not superior to the two others despite the fact that the simulation takes place exactly at the Maastricht steady state. The lack of enforcement of the Maastricht rule by EU governments has certainly had to do with the costly convergence path that we described in the previous section as well as with the absence of relative advantage of this rule at the steady state.

3.3 The lower bound of the investment rule

It is worth recalling that our simulations are partly biased against the use of the investment rule, since we rule out the endogeneity of potential output, which could be positively affected by public investment. Indeed, we consider the negative effect of public investment on output and public debt – through the interest rate and risk-premia. The crowding-out effect of public debt and deficit *via* interest rates (implicitly) on capital accumulation and (explicitly) on output is included in the model, in contrast with the probable effects of public investment as education, health or infrastructures on the potential of the economy.

Moreover, we deliberately set the smoothing parameter of expenditures δ in the investment rule to 1 which is equivalent to assuming that interest payments are not spread over many years but financed by a current surplus. Relaxing this assumption and smoothing the financing of interest payments and the cost of consolidation over different periods of time yields very different outcomes. Table 6 shows the macroeconomic performance of the investment rule in France for different values of δ. As expected, smoothing the consolidation over several years reduces the restrictiveness of the rule: the output loss is smaller over 2 years or even turns to an output gain over 5 or 10 years. The inflation gap to the target is smaller when the smoothing horizon increases. On the other hand, output and inflation volatility increases. This is not the more interesting result, however. One would expect that a more gradual financing of interest charges, driven by higher and more persistent deficits, would come at the expense of a higher debt ratio. However, there is no such tradeoff: the debt ratio at t=20 is smaller, for all three longer horizons, than in the 1-year case. This result mirrors the situation where debt ratios may actually decrease less than actually planned because of depressed growth during fiscal consolidation; here debt ratios may actually decrease more rapidly thanks to preserved growth when fiscal consolidation is smoothed.

4 Robustness

The results of our simulations show that the investment rule fares better in terms of output performance than the two other rules in the fiscal consolidation scenario. That results was obtained with a particular set of parameter values, as described in Section 2.4. While these values are all reasonable, we need to check for the robustness of this result, performing a Monte Carlo experiment over the space of the most relevant parameters. The objective is to make sure that the comparison between the three rules has not been dependent on the particular set of parameter chosen in Table 2.

We investigate the most representative parameters, *i.e.*, the ones capturing the degree of backward looking expectations in the IS and Phillips curves (α_1 and λ_1 respectively); the impact of real interest rates (α_2) and of the fiscal impulse (α_3) on the output gap (IS curve); the impact of the output gap on inflation in the Phillips curve (λ_2); the risk premium in the government bonds' interest rates equation (γ), and the initial levels of debt (b_{init}) and structural primary deficit (dsp_{init}).

The simulation is conducted as follows:

a) we make random draws of the parameters, within a certain range chosen to be consistent with most of the existing literature;
b) for each draw, we simulate the model for the three rules and select the run only if they all converge;
c) we record the average of discounted output gap and inflation values for each rule, and each parameter draw, over 20 periods.

The range of the 8 parameters random draws is reported in Table 7. We ran 11,000 simulations, and for about 96 per cent of them (10,591), the solution algorithm converged for the three rules. Non-convergence was most of the time due to the Fiscal Compact rule and to high values of γ the parameter capturing the risk premium in the government bonds' interest rates equation.

The 10591 converging iterations form our dataset. In Table 8, we report the descriptive statistics for the average of discounted output gap and inflation over the twenty years following the adoption of each of the three rules.

The results are remarkably stable and insensitive to large changes in parameters. The standard deviation of the average of the discounted output gap and inflation is higher for the one twentieth rule than for the two other rules. This confirms that the debt reduction rule, even if it converges, is more sensitive than the others to parameter variations.

Turning at the analysis of the results, we show that the investment rule fares significantly better than the others (the difference is significantly different from zero). The investment rule always provides the lowest output loss and inflation gap *vis-à-vis* the inflation target. Were the investment rule applied during the consolidation process, then the cost in terms of output gap would be of approximately one half lower than for the Fiscal Compact rule, over the parameters range. The sensitivity analysis run with this Monte Carlo experiment therefore confirms that the result according to which the investment rule outperforms the others in term of output loss is strongly robust to large parameters changes.

Table 7

Parameter Ranges for the Monte Carlo

Parameter	Range
α_1	[0.1 , 0.8]
α_2	[–0.9 , –0.1]
α_3	[0.2 , 0.8]
λ_1	[0.2 , 0.8]
λ_2	[0.1 , 0.5]
γ	[0 , 0.03]
b_{init}	[60 , 100]
dsp_{init}	[–1 , 4]

5 Conclusion

This paper evaluates the macroeconomic impact of a set of different fiscal rules that were, will, or might be implemented in Europe. We simulate a small-scale New Keynesian model with both forward- and backward expectations. The calibration draws on the existing literature and on the 2011 values of public finance data of 4 eurozone countries which we take as representative of the different types of eurozone member states. The three fiscal rules are: the status quo 3 per cent limit on public deficit, a debt reduction scheme and an investment rule in the vein of the UK golden rule of public finances.

Table 8

Monte Carlo Simulation

	Output Gap			Inflation		
	Status Quo	Inv. Rule	Fiscal Compact	Status Quo	Inv. Rule	Fiscal Compact
mean	–0.037	–0.035	–0.059	0.819	0.826	0.658
s.d.	0.040	0.039	0.052	0.256	0.253	0.349
min	–0.407	–0.399	–0.500	0.132	0.133	–0.278
max	0.000	0.000	0.000	1.246	1.246	1.246

Average over the 10591 simulations of the discounted sum of output gap and inflation.

We focus on two different scenarios. The first involves assessing the path followed by the four economies under each fiscal rule under fiscal consolidation from 2011 debt and deficit levels, towards the Maastricht steady state. The second assesses the impact of demand and supply shocks affecting the economy at the steady state.

The main results are first that abiding by the rules produces in all cases a short-run recession, even in a country with a small fiscal multiplier and a low initial public debt like the Netherlands. Second, during a consolidation phase, the investment rule performs better than the other rules: the recession is milder and shorter, thus leading to a substantially lower average loss of output over a 20-year horizon. Third, if the economy is hit by a demand or supply shock at the steady state, none of the rules emerges as superior in coping with them.. Finally, the Fiscal Compact, with its constant debt reduction rule, generally imposes large costs to the economy, while not necessarily performing better in terms of public finances' sustainability. These results are robust to parameters changes.

This leads to a general concluding remark. The Fiscal Compact requires a constant debt reduction, together with a "semi-balanced" (at 0.5 per cent) structural deficit. This implies that, once the target level of 60 per cent is reached, the debt ratio will continue to decrease, led by the structural deficit balance. Our results show that these rules are extremely costly, in terms of output loss, if compared to the investment rule or even the status quo. Such a drastic consolidation strategy embedded into EU constitutional laws threatens future macroeconomic performances of eurozone countries.

APPENDIX

Table 9

Simple Average Values of the Rules for 20 Years Fiscal Consolidation Scenario

France

	Status Quo	**Inv. Rule**	**Fiscal Compact**
mean(x)	–0.08	–0.07	–0.12
s.d.(x)	0.18	0.14	0.24
mean(p)	0.82	0.84	0.27
s.d.(p)	0.32	0.33	0.48
mean(ds)	3.10	3.14	2.11
s.d.(ds)	0.63	0.61	0.87
b (t=20)	83.71	83.80	76.60

Italy

	Status Quo	**Inv. Rule**	**Fiscal Compact***
mean(x)	–0.15	–0.13	–0.18
s.d.(x)	0.34	0.28	0.37
mean(p)	–0.26	–0.23	–0.35
s.d.(p)	0.62	0.63	0.66
mean(ds)	2.96	3.03	0.91
s.d.(ds)	0.23	0.15	0.78
b (t=20)	120.59	120.75	90.21

Belgium

	Status Quo	**Inv. Rule**	**Fiscal Compact**
mean(x)	–0.08	–0.08	–0.12
s.d.(x)	0.15	0.14	0.22
mean(p)	0.53	0.53	0.11
s.d.(p)	0.43	0.43	0.55
mean(ds)	3.09	3.13	1.77
s.d.(ds)	0.61	0.60	0.95
b (t=20)	93.83	94.20	80.40

Netherlands

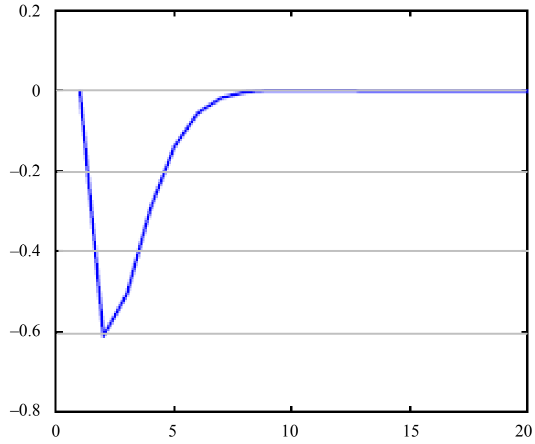
	Status Quo	**Inv. Rule**	**Fiscal Compact**
mean(x)	–0.01	–0.01	–0.02
s.d.(x)	0.02	0.02	0.03
mean(p)	1.78	1.78	1.72
s.d.(p)	0.06	0.07	0.09
mean(ds)	3.13	3.14	2.97
s.d.(ds)	0.62	0.62	0.66
b (t=20)	63.86	63.92	62.30

Average discounted values over 20 years.
* The fiscal compact enables convergence back to the steady-state in Italy only if gamma = 0.

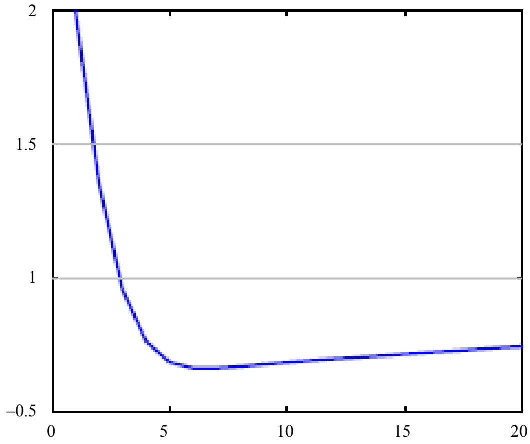
FRANCE

Status Quo

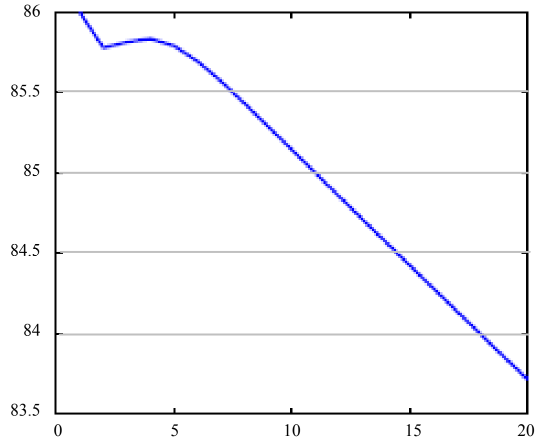
Output Gap



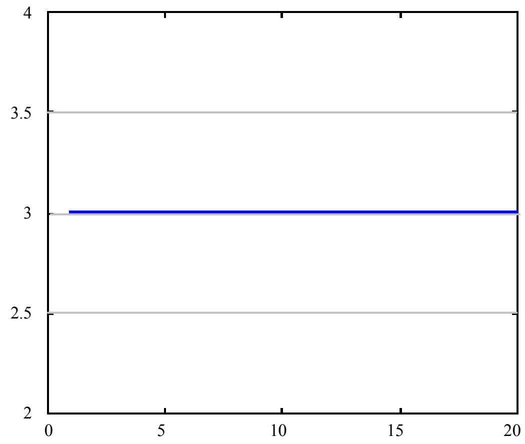
Inflation



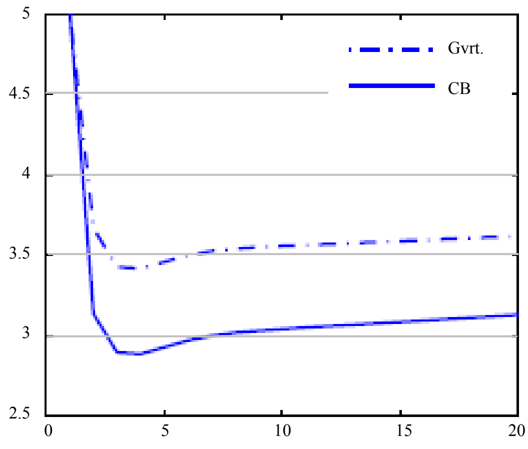
Debt to GDP



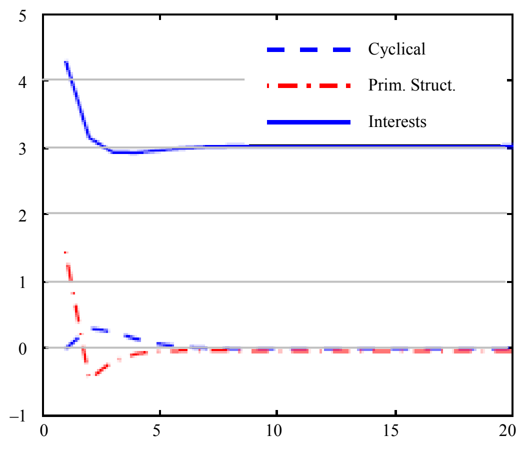
Total Budget Deficit



Nominal Interest Rates

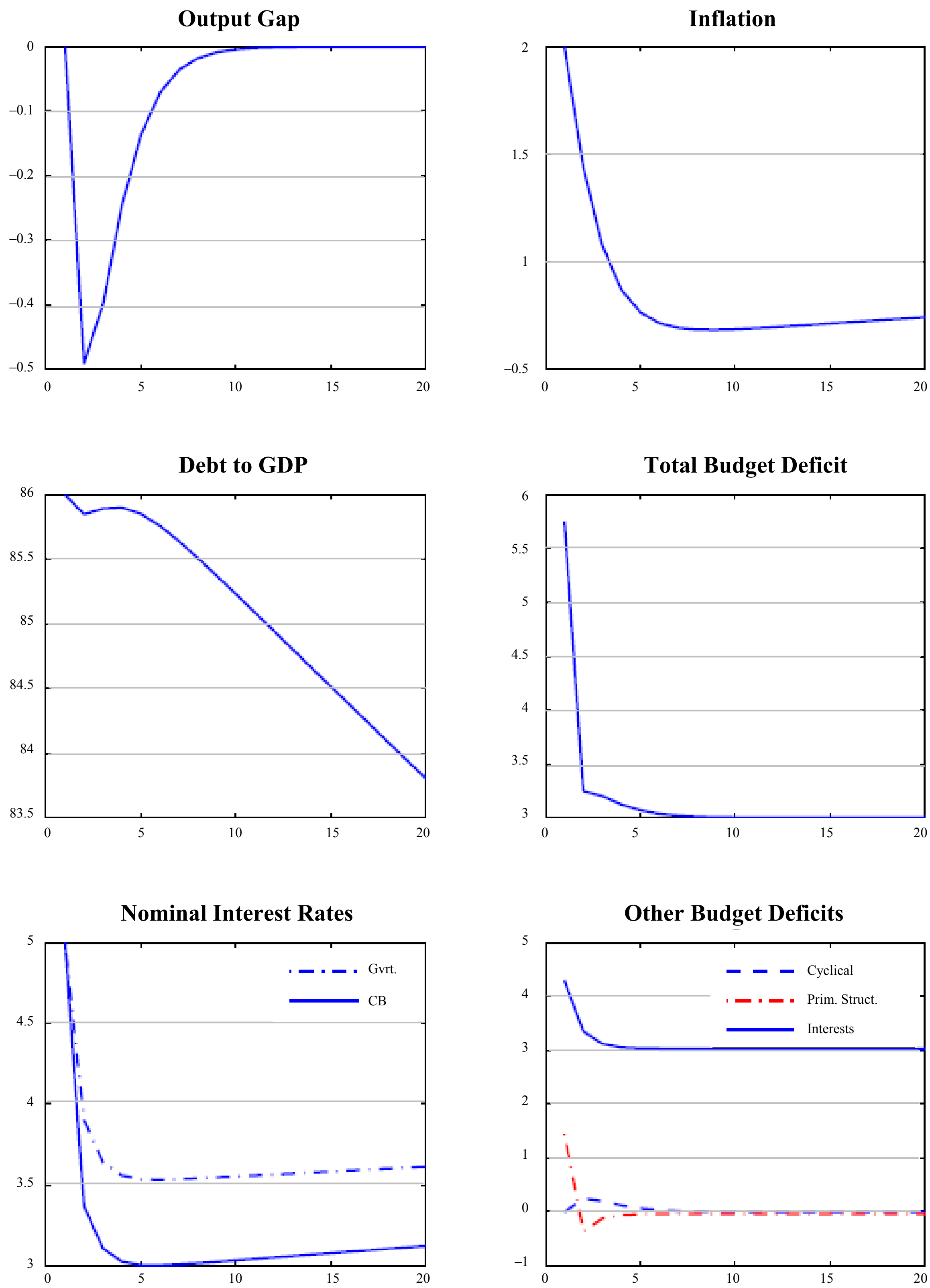


Other Budget Deficits



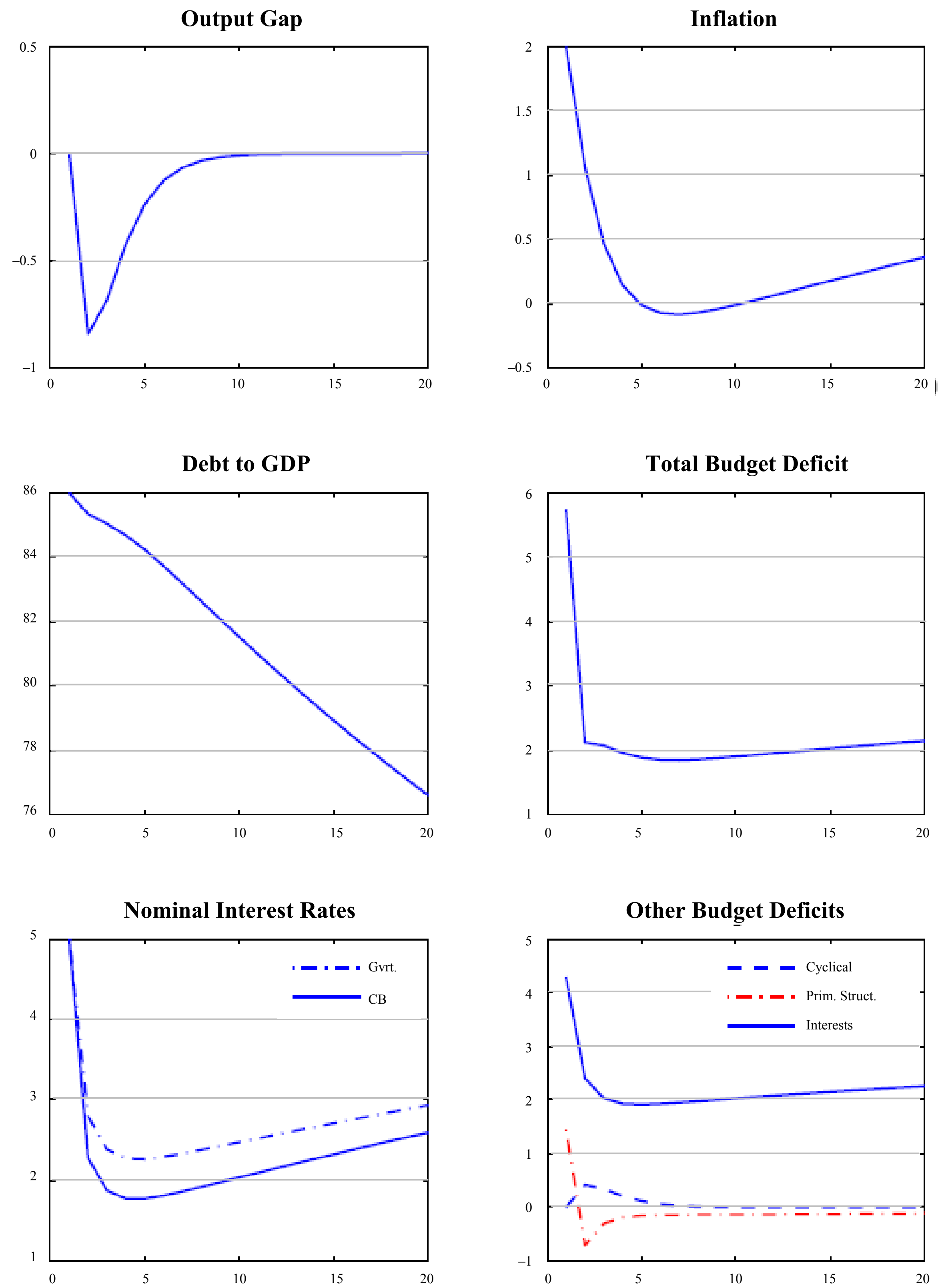
FRANCE

Investment Rule



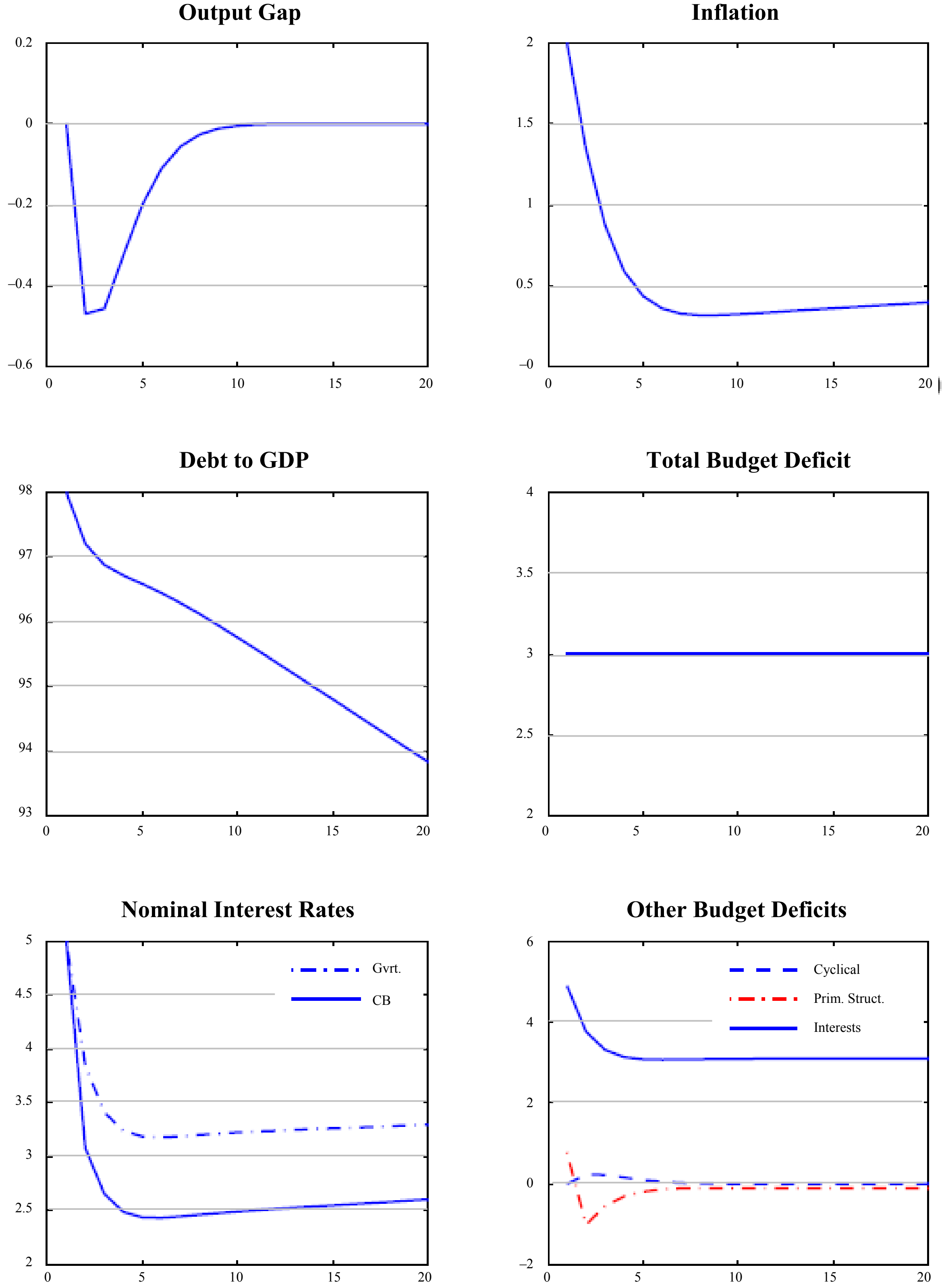
FRANCE

Fiscal Compact



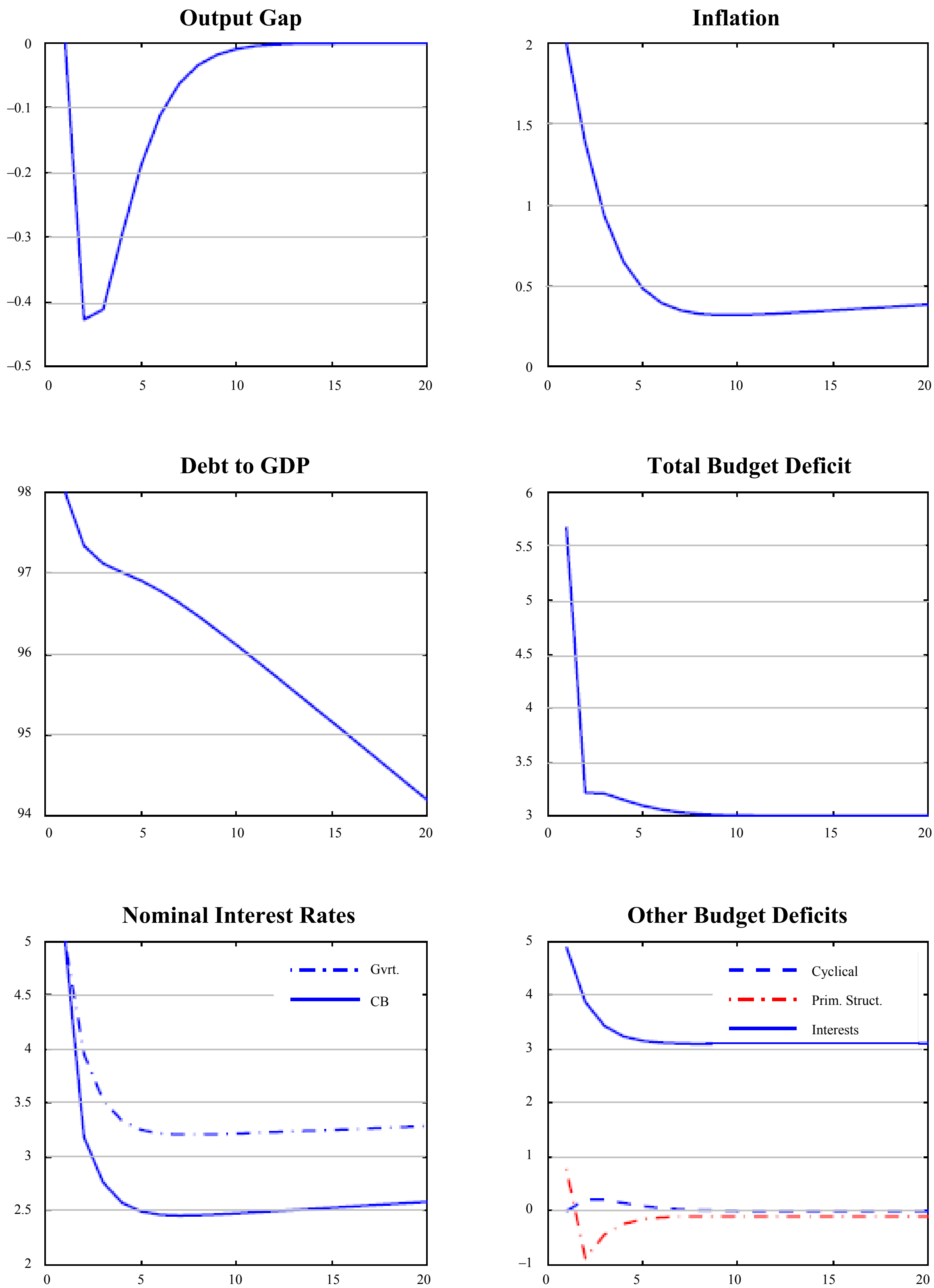
BELGIUM

Status Quo



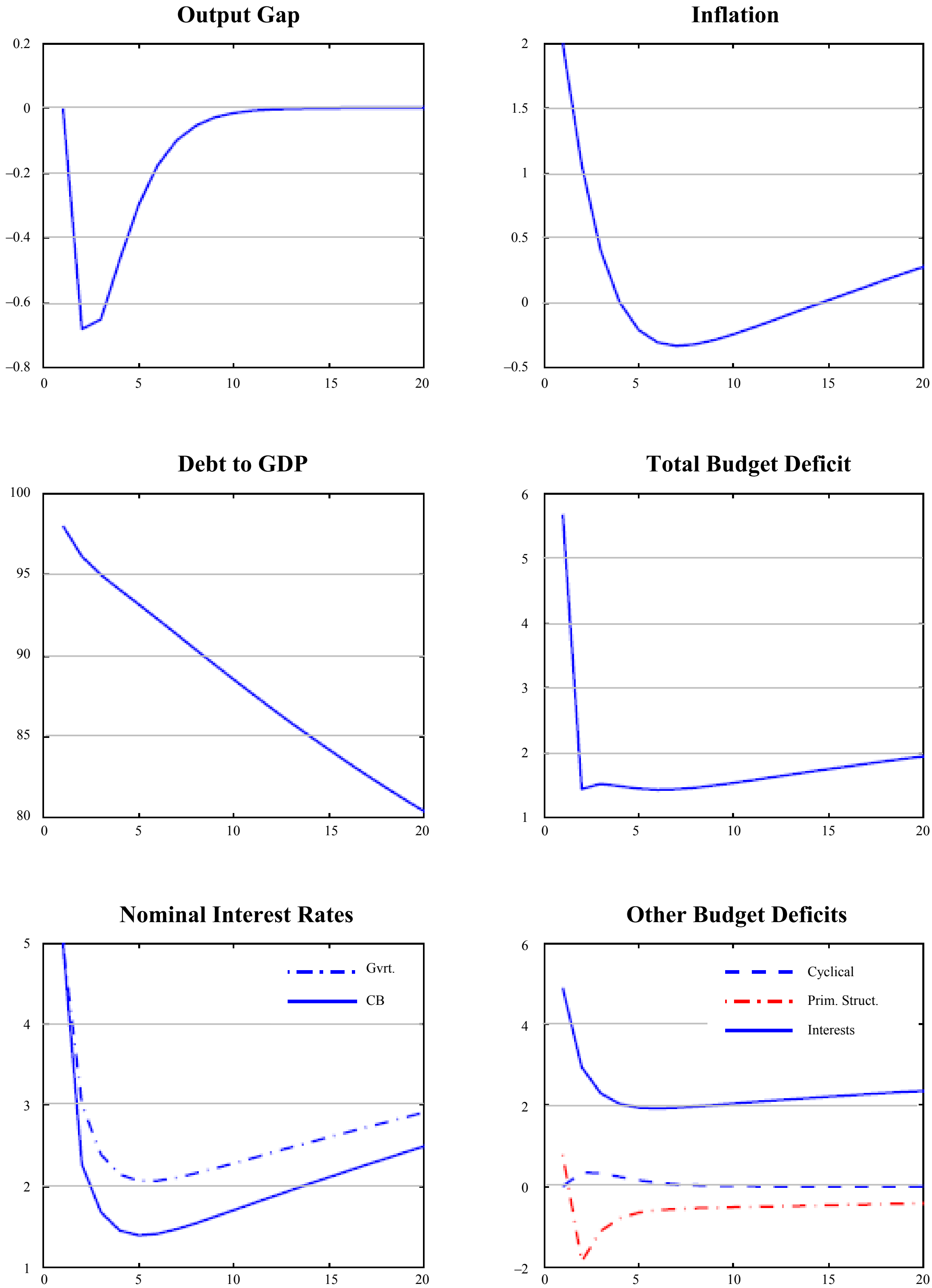
BELGIUM

Investment Rule



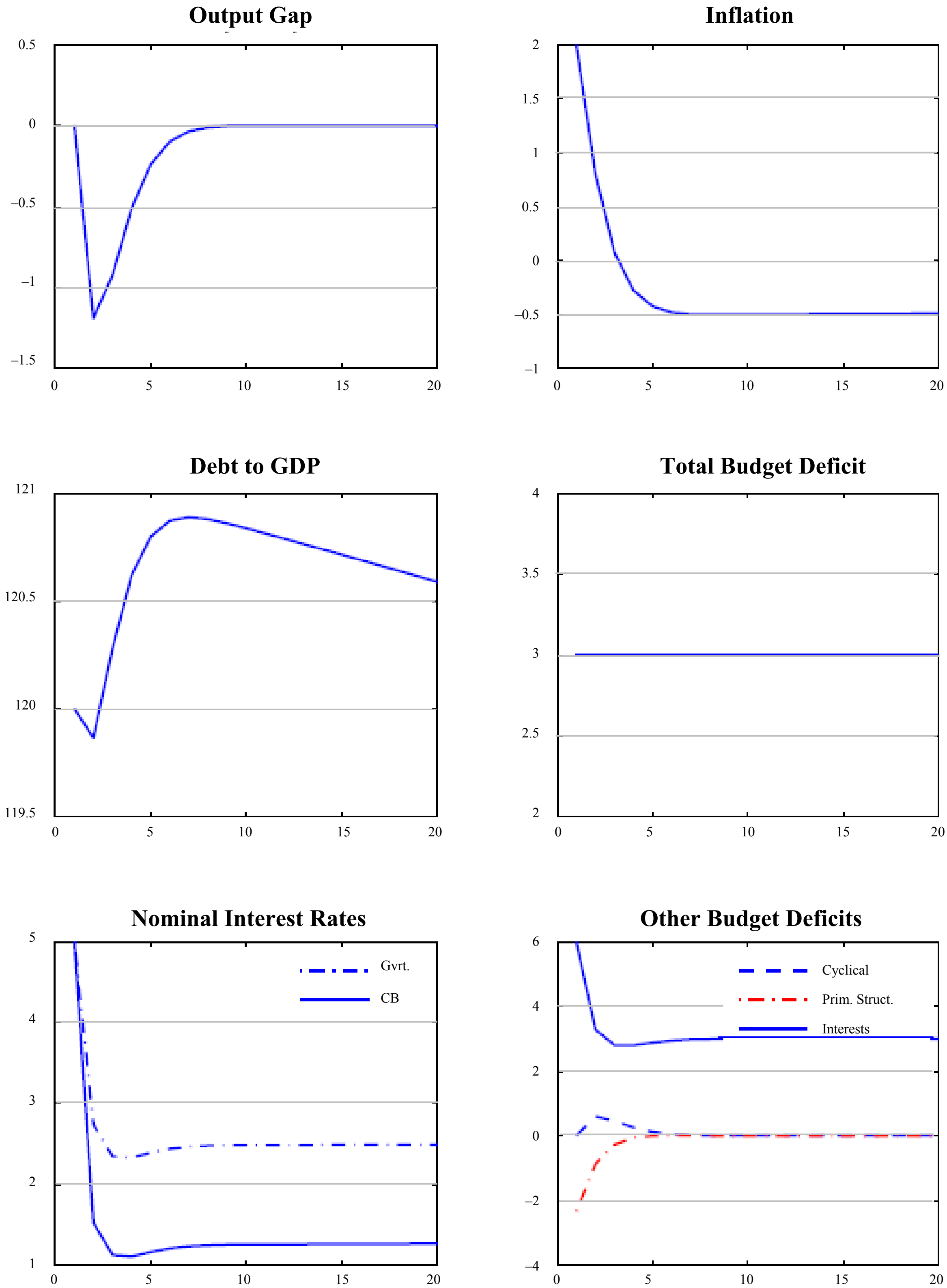
BELGIUM

Fiscal Compact



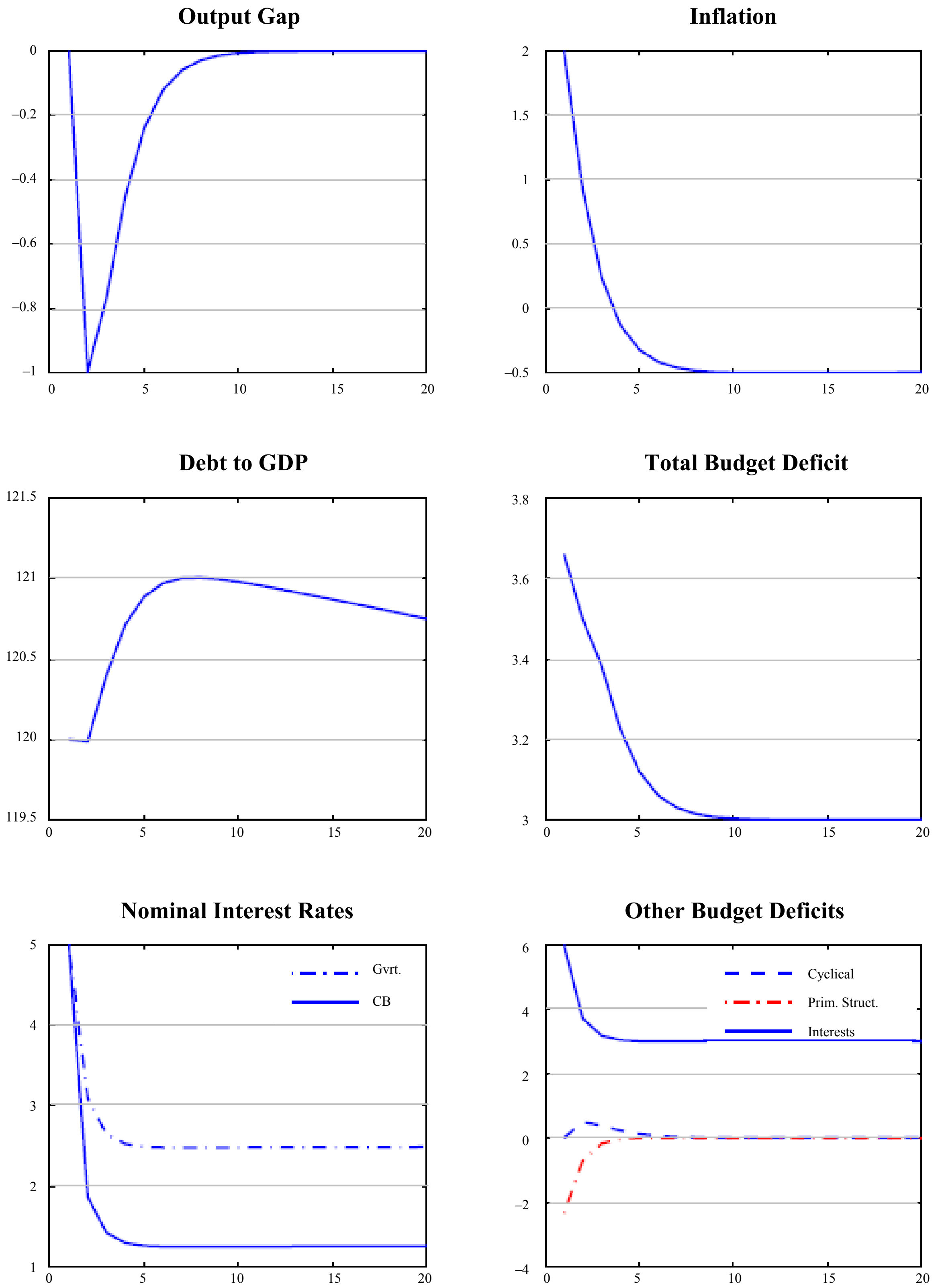
ITALY

Status Quo



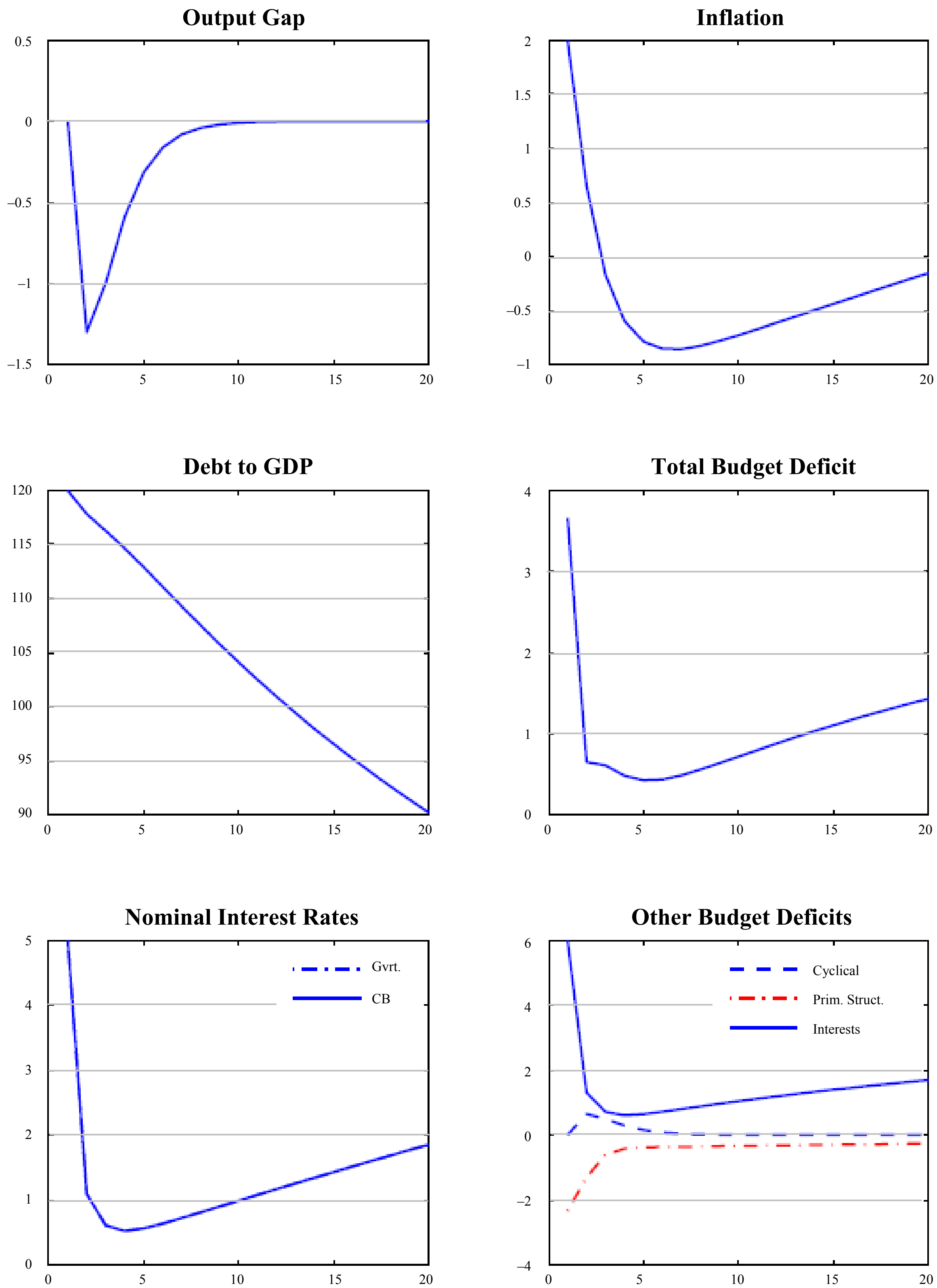
ITALY

Investment Rule



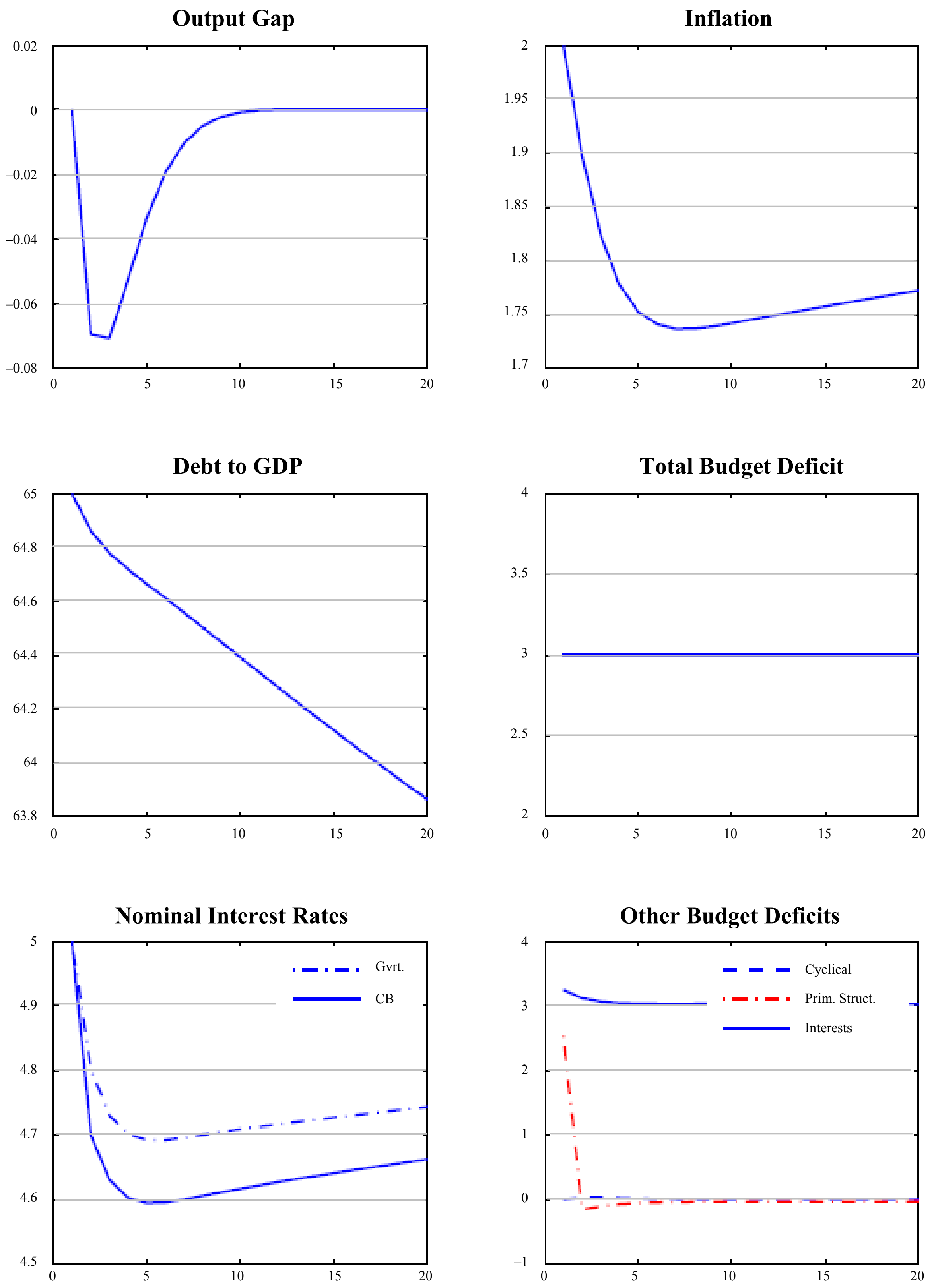
ITALY

Fiscal Compact, with $\gamma = 0$



NETHERLANDS

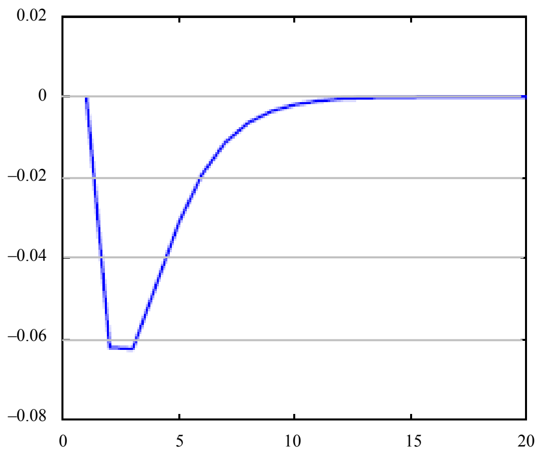
Status Quo



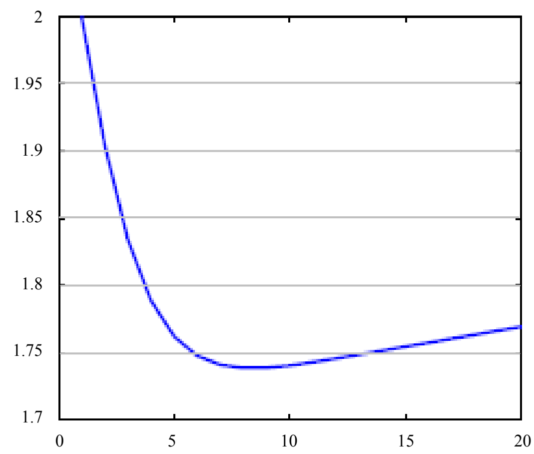
NETHERLANDS

Investment Rule

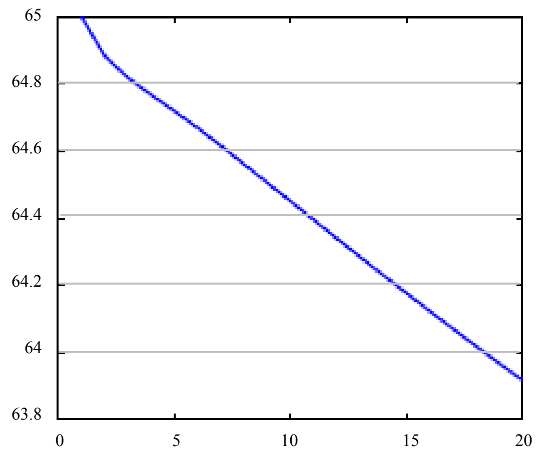
Output Gap



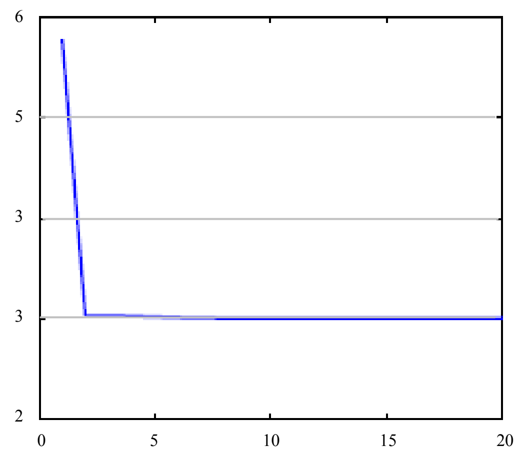
Inflation



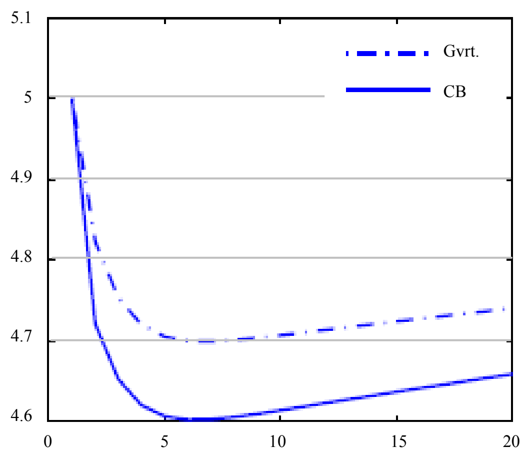
Debt to GDP



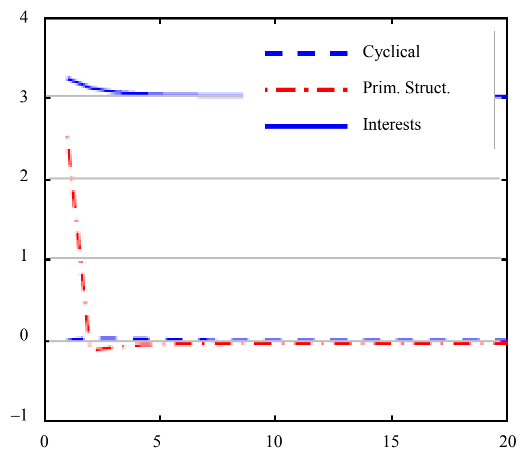
Total Budget Deficit



Nominal Interest Rates

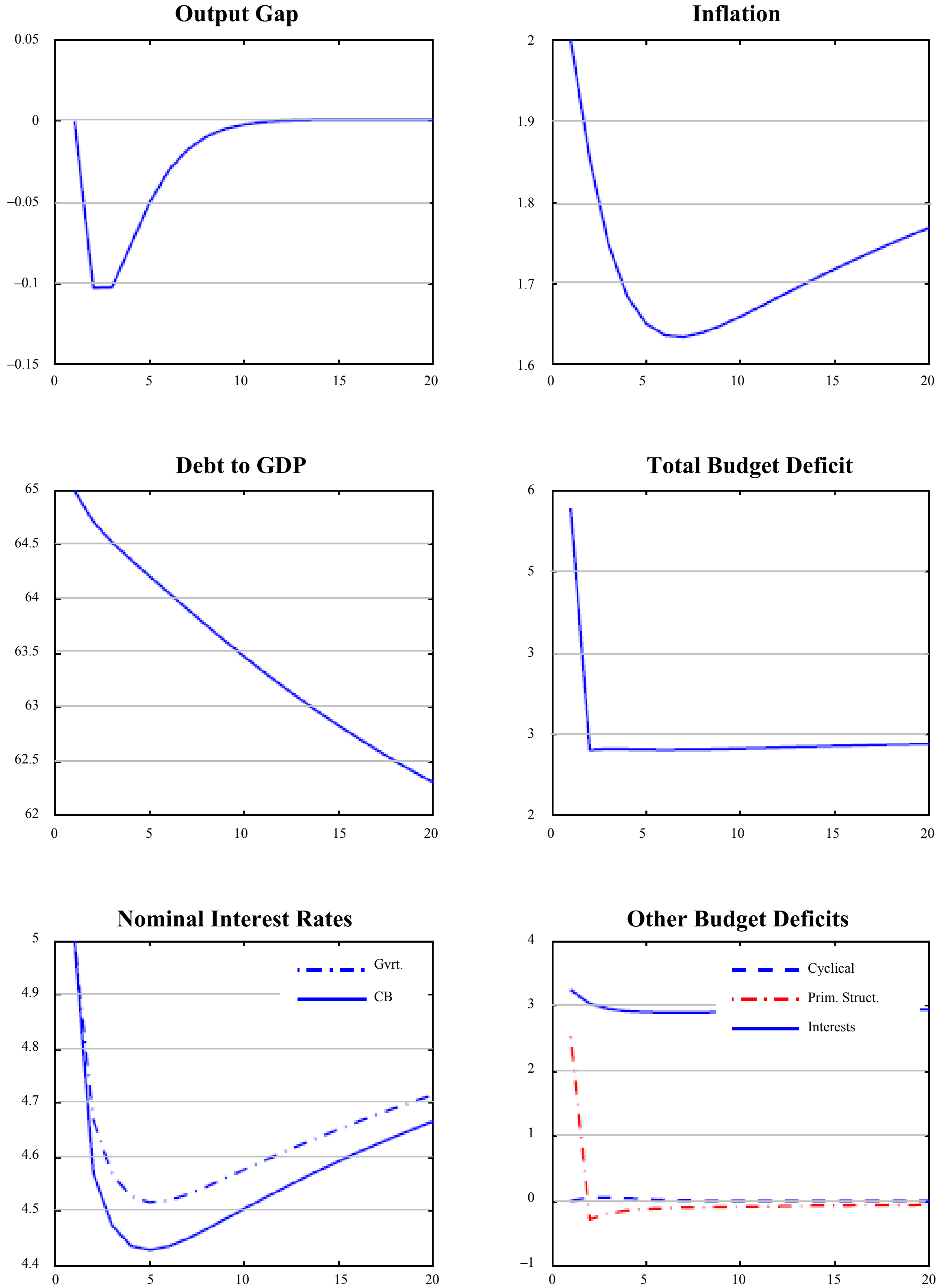


Other Budget Deficits



NETHERLANDS

Fiscal Compact



REFERENCES

Adolfson, M., S. Laseen, J. Linde and M. Villani (2007), "Bayesian Estimation of an Open Economy DSGE Model with Incomplete Pass-through", *Journal of International Economics*, No. 72, pp. 481-511.

Barro, R.J. (1986), "U.S. Deficits Since World War I", *Scandinavian Journal of Economics*, Vol. 88, No. 1, pp. 195-22.

Buiter, W.H. (2001), "Notes on 'A Code for Fiscal Stability'", *Oxford Economic Papers*, Vol. 53, No. 1, pp. 1-19.

Buti, M., D. Franco and H. Ongena (1998), "Fiscal Discipline and Flexibility in EMU: The Implementation of the Stability and Growth Pact", *Oxford Review of Economic Policy*, Vol. 14, No. 3, pp. 81-97.

Buti, M. and G. Giudice (2002), "Maastricht's Fiscal Rules at Ten: An Assessment", *Journal of Common Market Studies*, Vol. 40, No. 5, pp. 823-48.

Christoffel, K., K. Kuester and T. Linzert (2009), "The Role of Labor Markets for Euro Area Monetary Policy", *European Economic Review*, No. 53, pp. 908-36.

Clarida, R., J. Galí, and M. Gertler (1999), "The Science of Monetary Policy: A New Keynesian Perspective", *Journal of Economic Literature*, Vol. 37, No. 4, pp. 1661-707.

Coenen, G., C.J. Erceg, C. Freedman, D. Furceri, M. Kumhof, R. Lalonde, D. Laxton, J. Linde, A. Mourougane, D. Muir, S. Mursula, C. de Resende, J. Roberts, W. Roeger, S. Snudden, M. Trabandt and J. in't Veld (2012), "Effects of Fiscal Stimulus in Structural Models", *American Economic Journal: Macroeconomics*, No. 4, pp. 22-68.

Coenen, G., P. McAdam and R. Straub (2008), "Tax Reform and Labour-market Performance in the Euro Area: A Simulation-based Analysis Using the New Area-Wide Model", *Journal of Economic Dynamics and Control*, Vol. 32, No. 8, pp. 2543-83.

Cogan, J.F., T. Cwik, J.B. Taylor and V. Wieland (2010), "New Keynesian versus Old Keynesian Government Spending Multipliers", *Journal of Economic Dynamics and Control*, Vol. 34, pp. 281-95.

Creel, J., P. Hubert and F. Saraceno (2012), "Assessing the 'New Stability Pact': A Counterfactual Experiment on Alternative Proposals and Their Effects on Growth", *Journal of Economic Integration,* forthcoming.

Creel, J. and F. Saraceno (2010), "European Fiscal Rules after the Crisis", *Journal of Innovation Economics*, Vol. 6,No. 2, pp. 95-122.

Cwik, T. and V. Wieland (2011), "Keynesian Government Spending Multipliers and Spillovers in the Euro Area", *Economic Policy*, No. 26, pp. 493-549.

Dieppe, A., K. Kuester, P. McAdam (2005), "Optimal Monetary Policy Rules for the Euro Area: An Analysis Using the Area Wide Model", *Journal of Common Market Studies*, Vol. 43, No. 3, pp. 507–537.

Eichengreen, B. and C. Wyplosz (1998), "The Stability Pact: More Than a Minor Nuisance?", *Economic Policy*, No. 26, pp. 65-104.

Fuhrer, J.C. and G.D. Rudebusch (2004), "Estimating the Euler Equation for Output", *Journal of Monetary Economics*, Vol. 51, No. 6, pp. 1133-53.

Galí, J., M. Gertler and D. López-Salido (2005), "Robustness of the Estimates of the Hybrid New Keynesian Phillips Curve", *Journal of Monetary Economics*, Vol. 52, No. 6, pp. 1107-18.

Gelain, P. (2010), "The External Finance Premium in the Euro Area: A Dynamic Stochastic General Equilibrium Analysis", *North American Journal of Economics and Finance*, No. 21, p. 49-71.

Girouard, N. and C. André (2005), "Measuring Cyclically-adjusted Budget Balances for OECD Countries", OECD, Economics Department, Working Paper, No. 434.

Goodhart, C. and B. Hofmann (2005), "The Phillips Curve, the IS Curve and Monetary Transmission: Evidence for the US and the Euro Area", CESifo, Economic Studies, Vol. 51, No. 4, pp. 757-75.

Hauptmeier, S., A.J. Sánchez-Fuentes and L. Schuknecht (2011), "Towards Expenditure Rules and Fiscal Sanity in the Euro Area", *Journal of Policy Modeling*, Vol. 33, No. 4, pp. 597-617.

International Monetary Fund (2012), "Coping with High Debt and Sluggish Growth", *World Economic Outlook*, October.

Juillard, M. (1996), "Dynare: A Program for the Resolution and Simulation of Dynamic Models with Forward Variables Through the Use of a Relaxation Algorithm", CEPREMAP, Working Paper, No. 9602.

Kopits, G. and S. Symansky (1998), "Fiscal Policy Rules", International Monetary Fund, Occasional Paper, No. 162.

Monperrus-Veroni, P. and F. Saraceno (2005), "Reform of the Stability and Growth Pact: Reducing or Increasing the Nuisance?", OFCE, Working Paper, No. 2005-01.

Paloviita, M. (2008), "Comparing Alternative Phillips Curve Specifications: European Results with Survey-based Expectations", *Applied Economics*, Vol. 40, No. 17, pp. 2259-70.

Ratto, M., W. Roeger, J. in't Veld (2009), "QUEST III: An Estimated Open-economy DSGE Model of the Euro Area with Fiscal and Monetary Policy", *Economic Modelling*, Vol. 26, No. 1, pp. 222-33.

Rudd, J. and K. Whelan (2006), "Can Rational Expectations Sticky-price Models Explain Inflation Dynamics?", *American Economic Review*, Vol. 96, No. 1, pp. 303-20, March.

Schuknecht, L., P. Moutot, P. Rother and J. Stark (2011), "The Stability and Growth Pact: Crisis and Reform", European Central Bank, Occasional Paper, No. 129, September.

Smets, F. and R. Wouters (2003), "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area", *Journal of the European Economic Association*, Vol. 1, No. 5, pp. 1123-75.

Taylor, J.B. (1993), "Discretion Versus Policy Rules in Practice", *Carnegie-Rochester Conference Series on Public Policy*, Vol. 39, No. 1, pp. 195-214, December.

TSCG (2012), *Treaty on Stability, Coordination and Governance*, European Union.

Whelan, K. (2012), "What is the One-twentieth Rule and Does It Imply Additional Austerity?", blog entry, May 6th, available at: http://karlwhelan.com/blog/?p=434

Woodford, M. (2001), "Fiscal Requirements for Price Stability", *Journal of Money, Credit and Banking*, Vol. 33, No. 3, pp. 669-728.

FISCAL POLICY, STRUCTURAL REFORMS AND EXTERNAL IMBALANCES: A QUANTITATIVE EVALUATION FOR SPAIN

Angel Gavilán,[*] Pablo Hernández de Cos,[*] Juan F. Jimeno[*] and Juan A. Rojas[*]

This paper builds a large overlapping generations model of a small open economy featuring imperfect competition in the labor and product markets to understand i) which were the main determinants of the large expansionary phase experienced in Spain from the mid-1990s until the arrival of the global financial crisis in 2007-2008, ii) what role fiscal policy and structural reforms could have played to avoid the build-up of large external imbalance over this period, and iii) how these policies could affect the recovery of economic activity in Spain after the crisis. Our results indicate that falling interest rates and demographic changes were the main drivers of the Spanish expansionary phase and that, over this period, a tighter fiscal policy or structural reforms designed to foster competition in the labor and product markets could have not avoided the build-up of a large external imbalance. As for the macroeconomic behavior of the Spanish economy after the crisis our model highlights the trade-off faced by tighter fiscal policies: they may reduce the output losses induced by the crisis in the medium-term but at the expense of a mild output loss in the years immediately after the crisis. Instead, structural reforms do not face this trade-off and they may contribute to reduce output losses in the short- and medium-term, while inducing a positive long-run effect on the level of output.

1 Introduction

From the mid-1990s to 2008, the Spanish economy enjoyed a phase of sustained economic growth in which real convergence with the core EMU member countries advanced notably. This expansionary phase was mostly driven by two factors. First, by a significant expansion of credit, that was induced by the fall in interest rates that followed Spain's adhesion to the EMU and, more broadly, by a pervasive relaxation in the conditions of access to credit. And second, by the large immigration inflows into Spain over the period that substantially modified the demographic structure of the Spanish population.[1]

Yet significant imbalances built up in the process. On the one hand, the Spanish economy became increasingly more dependent of external financing over the period. The fall in interest rates and the overall expansion of credit led to an investment boom, much of which materialized in the housing sector, that increased the share of investment in GDP from around 22 in 1995 to 29 per cent in 2008. Thus, despite a move toward fiscal consolidation by the public sector, the Spanish current account deficit, that was close to zero in 1998, increased nearly monotonically over the period, reaching almost 10 per cent of GDP by 2008. On the other hand, price-competitiveness of the Spanish economy also deteriorated significantly, due to very low productivity growth and to the existence of important distortions in the domestic labor and product markets.

---

[*] Banco de España.

We thank António Afonso, Pablo Burriel, Miguel Cardoso, Rafael Domenech, Paul Hiebert, Miguel León-Ledesma, Carlos Martnez-Mongay, Gernot Müller, Anton Nakov and seminar participants at the 14th Public Finance Workshop (Perugia), the XXIII Symposium of Moneda y Crédito, European Central Bank, National Bank of Slovakia (Conference on "The Euro Area and the Financial Crisis"), Banco de España (Conference on "Interactions between Monetary and Fiscal Policies") and European Commission (Workshop on "External Imbalances and Public Finances in the EU") for helpful comments. We also benefited from useful suggestions by an editor and one anonymous referee.

[1] For a recent account of the evolution of the Spanish economy during this period, see Estrada, Jimeno and Malo de Molina (2009).

When the global financial crisis struck and the very favorable international credit conditions suddenly disappeared, the Spanish economy began an inevitable adjustment process, with a substantial reduction in consumption and investment by 2008q4, when housing investment plummeted. This adjustment, that has helped to correct the excessive indebtedness of the private sector, has led however to a large decrease in economic activity, with GDP growth in 2009 at around -3.6 per cent. At the same time, the work of automatic stabilizers and the expansionary fiscal programs put in place by the government to mitigate the effects of the crisis, have led to a very rapid deterioration of public accounts, that have moved from a surplus of around 2 per cent of GDP in 2007 to deficits of around 4 and 11 per cent in 2008 and 2009, respectively. All together, the Spanish economy has very quickly reduced its need for external financing as its current account deficit has decreased from around 10 per cent of GDP in 2008 to around 6 per cent in 2009, being now mostly driven by fiscal deficits rather than by private indebtedness as in the expansionary phase.

With this evidence in mind, several questions arise: i) to what extent are the fall in interest rates and the profound demographic changes witnessed in the Spanish economy over the last decade responsible for the expansionary phase and the build-up of imbalances?, ii) could have fiscal policy contributed more to avoid the build-up of these imbalances?, iii) how would structural reforms increasing competition in the product and labor markets have diminished the saving-investment gap and the loss of price-competitiveness of that period?, and iv) looking ahead, once the economy has been hit by the global financial shock in 2008, how would alternative fiscal policies and reforms in the labor and product markets may affect the expected macroeconomic evolution of the Spanish economy?

In order to address these questions, this paper constructs and calibrates a small open economy model for Spain. The model economy is composed by households, firms and a government. To properly incorporate the intense demographic changes that the Spanish economy experienced over the last decade, and those expected to happen in the future, this paper considers a large scale overlapping generations model.[2] In each period, households take consumption, labor and savings decisions to maximize their lifetime utility. There are four types of firms in the economy, that produce a final consumption good, intermediate goods, labor services and capital services. As the Spanish economy is characterized by rigid labor and product markets, the model incorporates distortions in these markets via monopoly power of intermediate goods and labor services producers. This approach, relatively standard in the new Keynesian literature, is less common in the large scale OLG literature, that typically considers perfect competition in all markets. The government in the model consumes, gives lump-sum transfers, runs a social security system, levies taxes (on consumption and on labor and capital income) and issues debt. The description of the social security system in the model is particularly rich. This is very relevant since, undoubtedly, one needs to take into account the pressures on the social security system generated by the aging of the Spanish baby-boom generation in the near future in order to properly analyze the role played by fiscal policy in the recent and future macroeconomic developments of the Spanish economy.

The model is calibrated to match the main macroeconomic features of the Spanish economy in 1998 and then its performance of over the period 1998-2008 is analyzed under different scenarios concerning interest rates, demographic developments, fiscal policies and labor and product market distortions. Our results indicate that, in line with Izquierdo, Jimeno and Rojas (2010), interest rates and demographic changes are the main responsible for the investment boom and the build-up of a sizable external imbalance (measured as the ratio of net foreign assets to

[2] In this sense, the model is an extended version of the general equilibrium model with overlapping generations used in Izquierdo, Jimeno and Rojas (2010) to evaluate the impact of immigration on the Spanish economy, and in Jimeno, Rojas and Puente (2008) and Rojas (2005) to analyze the consequences of population ageing in Spain.

GDP) witnessed in the Spanish economy during the expansionary phase. In this context, we find a very limited role for fiscal policy in reducing the external imbalance accumulated in Spain over the period 1998-2008. In particular, our results show that a temporary reduction of government expenditure over the expansionary phase would have reduced the size of the Spanish external imbalance by 2008 only very slightly. A more permanent tightening of fiscal policy could have even increased this imbalance. With respect to the effects of structural reforms in product and labor markets pursuing an increase in competition in these markets, we find that, although they would have not helped in reducing the external imbalance of the Spanish economy over the period 1998-2008, they would have led to a short- and long-run expansion of output, employment and investment, and to a substantial improvement in competitiveness and in public accounts. It is precisely due to these positive effects on the economy that these structural reforms may naturally induce in the short-run an increase in the external indebtedness of the economy, as forward-looking households anticipate lower taxes and a more efficient economy in the future and try to smooth their consumption.

As for the macroeconomic behavior of the Spanish economy beyond 2008, our model suggests that, even without the arrival of the crisis, in the short-run the Spanish economy would have entered into a phase of lower GDP growth where the external imbalance of the economy would have been reduced but where public accounts would have deteriorated. The arrival of the global financial crisis has accentuated the aforementioned dynamics. Furthermore, the model highlights the trade-off faced by tighter fiscal policies in the post-crisis scenario: they may reduce the output losses induced by the crisis in the medium-term but at the expense of a more intense output loss in the years immediately after the crisis. In contrast, structural reforms do not face this trade-off and may contribute to reduce output losses in the short- and medium-term, while inducing a positive long-run effect on the level of output.

The rest of the paper is organized as follows. Section 2 lays down the main facts with respect to the macroeconomic evolution of the Spanish economy during the expansionary phase. Then, Section 3 describes the model and Section 4 its calibration. Departing from the model economy calibrated to 1998, Section 5 quantifies the role played by demographic developments and interest rates in shaping the Spanish macroeconomic evolution over the period of analysis and performs counterfactual exercises regarding alternative fiscal policies and labor and product markets reforms. Then, Section 6 introduces into our model economy the global financial crisis that hit the economy in 2008 and shows the predictions of the model beyond that date. Finally, Section 7 concludes.

2 The expansion: driving factors and imbalances

The expansionary phase that the Spanish economy enjoyed from the mid-1990s to 2008 was characterized, among other things, by a process of fiscal consolidation in the public sector and by the build-up of a sizable external imbalance, with large and increasing current account deficits over the period that significantly deteriorated the international investment position of the country. This expansionary process was mostly fuelled by two factors: the fall in interest rates and the expansion of credit, and the large immigration inflows into Spain over the period. This section lays down the evolution of these variables for the period 1995-2008.[3]

Interest rates – In terms of the evolution of ex-post real long-term and short-term interest rates in Spain, despite a slight increase after 2005, the fall in these rates during the period was truly

[3] In this section, data come from the OECD Economic Outlook, except that of the Spanish current account balance and international investment position (Banco de España). Population data is from Instituto Nacional de Estadstica (INE).

remarkable: between 1995 and 2005 long-term (short-term) rates fell by around 7 (6) percentage points.[4] Nominal convergence in the run-up to EMU, lax monetary policy since the early 2000s, anchoring of inflation expectations, and a positive inflation differential in Spain are behind that large decrease. As already mentioned above, this cheaper access to credit, joint with a relaxation in credit standards, that allowed for a wider access to credit, were one of the main push factors behind the Spanish economic expansion during the 1995-2008 period.

I*mmigration inflows* – Immigration inflows were another important factor behind the last expansionary process in the Spanish economy. In Spain, traditionally an out-migration country, these inflows reached a significant scale in the years immediately before the creation of EMU and, since then, they have intensively transformed the Spanish population. Thus, foreign population residing in Spain has increased from 0.35 millions (1 percent of total population) in 1995 to 5.22 millions (11 per cent of total population) in 2008. In addition, these inflows have modified the age distribution in the Spanish population reducing its dependency ratio since, as usual, the age distribution of the immigrants that have entered into Spain has been younger than that of natives.

Fiscal consolidation – Up to 2007, fiscal consolidation in Spain was achieved both through a reduction in expenditures and through an increase in revenues. Thus, public deficit, which was around 6.5 per cent of GDP in 1995, gradually disappeared, to reach a surplus of almost 2 per cent of GDP in 2007. In 2008, however, with the arrival of the global financial crisis, government disbursements increased again, revenues fell and public deficit reached 4 per cent of GDP. Overall, the process of fiscal consolidation over this period contributed to a considerable reduction of public debt, that decreased from 63.3 per cent of GDP in 1995 to 39.7 per cent in 2008. In this sense, Spain significantly over-performed other EMU countries on this account.

External imbalance – The Spanish current account balance as a percentage of GDP fell almost monotonically during the 1995-2008 period and led to a very intense deterioration in the share of net foreign assets in GDP, that decreased from around –22 per cent in 1995 to around –80 per cent in 2008. The increase in current account deficits over this period, despite the process of consolidation of public accounts, clearly points to the rise in private indebtedness as the main origin of this external imbalance. In a cross-country comparison, it is evident that the Spanish increasing dependence on external financing over this period is truly remarkable, only comparable to that of Portugal and Greece and more intense than that of the U.S. It also contrasts with the situation of other countries in the EMU. Thus, while Germany and Finland exhibited sizeable current account surpluses, France and Italy showed a considerable less intense deterioration in their current account balance than that of Spain.

3 The model

This section describes the model used to perform the quantitative experiments reported on the following sections on the macroeconomic effects of interest rates, demographic changes, fiscal policy and product and labor market reforms. In essence, it is a model for a small open economy within a monetary union that combines, on the household side, the large scale overlapping generations structure of Auerbach and Kotlikoff (1987) and, on the supply side, the now standard framework in the new Keynesian literature with firms producing final and intermediate goods, labor and capital services in the presence of monopolistic competition in the intermediate goods

[4] It is somehow controversial, however, to what extent this fall truly resembles a reduction in the cost of financing. For some (see, for instance, Blanco and Restoy (2007) and Gimeno and Marques (2008)) the reduction in inflation uncertainty explains a great deal of the decline in real interest rates, so that the actual real cost of financing might have decreased significantly less than that indicated by ex-post real rates.

and labor markets.[5] The model economy is completed by a government that has a wide range of fiscal instruments at its disposal.

3.1 Households

Demographics – The economy has overlapping generations of agents who live a maximum of I periods. The agents differ in their age, $i \in \{1,2,...,I\}$, and in their place of birth, $n \in \{1,2,...,I\}$, where n=1 identifies a *native* and n=n_0≥2 denotes an *immigrant* who first entered the economy with age n_0.[6] We use $N_{i,n,t}$ to denote the total number of agents of type (i, n) in the economy at period t and $\mu_{i,n,t}$ to denote the share of these agents over the total population at that period. The former evolves over time in the following fashion:

$$N_{1,1,t} = \sum_{i}\sum_{n} N_{i,n,t-1} f_{i,t-1} \qquad (Births) \qquad (1)$$

$$N_{i,1,t} = N_{i-1,1,t-1} s_{i-1,t-1}, \forall i \geq 2 \qquad (Natives) \qquad (2)$$

$$N_{i,n,t} = N_{i-1,n,t-1} s_{i-1,t-1} + NI_{i,n,t}, \forall i \geq 2, \forall n \geq 2 \qquad (Immigrants) \qquad (3)$$

where $s_{i-1,t-1}$ denotes the conditional probability of surviving from age $i-1$ to age i at period $t-1$, $f_{i,t-1}$ is the probability of an agent of age i of having an offspring at that period, and $NI_{i,n,t}$ is equal to 0 when $i \neq n$ and to the number of immigrants of age i exogenously entering the economy at the beginning of period t when $i = n$. We assume that the survival and fertility probabilities are common to natives and immigrants, since there is no independent data readily available for these two population groups, and we consider the offspring of immigrants as natives.

Decision problem – At an exogenous age I_A agents start taking decisions. At that time they have no assets, besides transfers emanating from accidental bequests. In each period, agents take consumption and labor decisions in order to maximize lifetime utility. At period t an agent of type (v,n) solves the following problem:

$$Max \sum_{i=v}^{I} \beta^{i-v} \psi_{v,t}^{i} U(c_{i,n,t+i-v}, h_{i,n,t+i-v}) \qquad (4)$$

subject to:

$$\begin{aligned}(1+\tau_t^c)c_{i,n,t} + a_{i+1,n,t+1} \leq (1-\tau_t^l - \tau_t^{ss})w_{i,t}h_{i,n,t} + (1+r_t(1-\tau_t^a))(a_{i,n,t}+b_t) + \\ + ss_{i,n,t} + tr_t + div_t \quad \forall t\end{aligned} \qquad (5)$$

In the expression above, c is consumption, a denotes beginning of period assets and h is time spent at work. Agents are endowed with one unit of time per period. Between ages I_A and $I_R - 1$ this unit of time must be allocated between labor and leisure. Afterwards, agents are forced to retire. Only then, they receive social security benefits, ss, and devote their entire time endowment

---

5 Unlike the new Keynesian literature we do not consider price rigidities.

6 We need to keep track of the age at which immigrants entered the economy because we assume that they arrive with no assets (as in, for instance, Storesletten (2000, 2003) and Razin and Sadka (1999)). Thus, conditional on age, two immigrants arriving to the economy at different ages take different consumption and labor decisions because they do not have the same wealth level.

to leisure. In each period, regardless of their type, agents receive lump transfers, *tr*, accidental asset bequests, *b*, and dividends from the different firms operating in the economy, div. Regarding prices and taxes, we normalize the price of the final good consumed by households to one, w is the age-dependent wage (in units of *c*) agents receive for their working time, r is the net real interest rate paid on savings, τ_t^l and τ_t^{ss} are labor income taxes (the latter being the social security tax), and τ_t^c and τ_t^a denote proportional taxes on consumption and capital income, respectively. Finally, β is the discount parameter and $\psi_{v,t}^i$ is the unconditional probability of reaching age i for an individual that has age v at period *t*. Thus, $\psi_{v,t}^i = \prod_{k=v+1}^{i} s_{k-1,t+k-v-1}$ with $\psi_{v,t}^v = 1$.

3.2 Firms

Final good firm – In each period, a final consumption good, Y_t, is produced within the small open economy by a perfectly competitive firm. The firm does so by combining a continuum of domestic intermediate goods, $y_{H,j,t}$, $j \in (0,1)$, and a continuum of foreign intermediate goods, $y_{F,z,t}$, $z \in (0,1)$, using the following technology:

$$Y_t = \left[(\alpha_C)^{\frac{1}{\eta_C}} Y_{H,t}^{\frac{\eta_C - 1}{\eta_C}} + (1-\alpha_C)^{\frac{1}{\eta_c}} Y_{F,t}^{\frac{\eta_C - 1}{\eta_C}} \right]^{\frac{\eta_C}{\eta_C - 1}} \tag{6}$$

where $Y_{H,t}$ and $Y_{F,t}$ are composites of the continuum of domestic and of foreign intermediate goods, respectively, and follow the constant elasticity of substitution functions:

$$Y_{H,t} = \left[\int_0^1 y_{H,j,t}^{\frac{1}{\lambda_t^i}} dj \right]^{\lambda_t^i}, 1 \le \lambda_t^i < \infty \tag{7}$$

$$Y_{F,t} = \left[\int_0^1 y_{F,z,t}^{\frac{1}{\lambda_t^{i^*}}} dz \right]^{\lambda_t^{i^*}}, 1 \le \lambda_t^{i^*} < \infty \tag{8}$$

In the expressions above, $(1-\alpha_C)$ is the share of imports in consumption, η_c is the elasticity of substitution between the domestic and foreign composite goods, and λ_t^i ($\lambda_t^{i^*}$) denotes the time-varying substitutability of domestic (foreign) intermediate goods in the production of $Y_{H,t}$ ($Y_{F,t}$). Let $p_{H,j,t}$ ($p_{F,z,t}$) denote the price (in units of c) of the domestic (foreign) intermediate good *j*(*z*) in period *t*. Profit maximization by the final good firm implies the following demands for the composite goods $Y_{H,t}$ and $Y_{F,t}$, and for each intermediate good $y_{H,j,t}$ and $y_{F,z,t}$:

$$Y_{H,t} = \alpha_C p_{H,t}^{-\eta_C} Y_t, \qquad y_{H,j,t} = \left(\frac{p_{H,j,t}}{p_{H,t}} \right)^{\frac{\lambda_t^i}{1-\lambda_t^i}} Y_{H,t} \tag{9}$$

$$Y_{F,t} = (1-\alpha_C) p_{F,t}^{-\eta_C} Y_t, \quad y_{F,z,t} = \left(\frac{p_{F,z,t}}{p_{F,t}} \right)^{\frac{\lambda_t^{i^*}}{1-\lambda_t^{i^*}}} Y_{F,t} \tag{10}$$

where prices are related in the following fashion:

$$1 = \left[\alpha_C p_{H,t}^{1-\eta_C} + (1-\alpha_C) p_{F,t}^{1-\eta_C} \right]^{\frac{1}{1-\eta_C}} \tag{11}$$

$$p_{H,t} = \left[\int_0^1 p_{H,j,t}^{\frac{1}{1-\lambda_t^i}} dj \right]^{1-\lambda_t^i} \tag{12}$$

$$p_{F,t} = \left[\int_0^1 p_{F,z,t}^{\frac{1}{1-\lambda_t^{i^*}}} dz \right]^{1-\lambda_t^{i^*}} \tag{13}$$

Intermediate good firms – Each domestic intermediate good, $y_{H,j,t}$, $j \in (0,1)$, is produced within the small open economy by a monopolist who rents capital, K, and labor, L, in the market and uses the technology:

$$y_{H,j,t} = K_{H,j,t}^{\xi} \left(A_t L_{H,j,t} \right)^{1-\xi} \tag{14}$$

where $0 < \xi < 1$ and A_t denotes economy-wide labor augmenting technological change. As in Christiano *et al.* (2005), we rule out entry and exit into the production of intermediate goods. Profit maximization by these monopolists implies that, in each period, they set the price for their differentiated goods with a markup over their marginal costs. Namely:

$$p_{H,j,t} = \lambda_t^i MC_{H,j,t} \tag{15}$$

where the marginal cost, $MC_{H,j,t}$, depends on the rental price of the labor and capital inputs, W_t and $p_{K,t}$, respectively, according to:

$$MC_{H,j,t} = \left(\frac{\xi}{1-\xi} \right)^{-\xi} \left(\frac{W_t}{p_{K,t}} \right)^{-\xi} \frac{W_t}{A_t (1-\xi)} \tag{16}$$

At the end of each period, the profits of these monopolists, $\pi_{H,j,t}$, $j \in (0,1)$ are distributed to the households in the form of dividends.

Firms producing labor services – In each period, a representative competitive firm buys labor hours of households of different ages and transforms them into an aggregate labor input, L_t, which then sells to the domestic intermediate producers, using the following technology:

$$L_t = \left[\sum_{i=I_A}^{I_R-1} \left(\frac{1}{I_R - I_A} \right)^{\frac{\lambda_t^l - 1}{\lambda_t^l}} \left(e_i L_{i,t} \right)^{\frac{1}{\lambda_t^l}} \right]^{\lambda_t^l}, 1 \leq \lambda_t^l < \infty \quad (17)$$

where $L_{i,t}$ is the total number of labor hours supplied by age-i households, e_i is an age-specific index which transforms those raw labor hours into efficient units of labor, and λ_t^l measures the time-varying substitutability of labor hours of households of different ages in the production of the aggregate labor input. Profit maximization by this representative firm in the labor market implies that its demand for labor hours of age-i households is equal to:

$$L_{i,t} = \left(\frac{1}{I_R - I_A} \right) \left(\frac{\widetilde{w}_{i,t}}{W_t} \right)^{\frac{\lambda_t^l}{1-\lambda_t^l}} \left(e_i \right)^{\frac{1}{1-\lambda_t^l}} L_t \quad (18)$$

where $\widetilde{w}_{i,t}$ denotes the price that this firm pays for one hour of labor of an age-i household and W_t is the unit price of the aggregate labor input. These are related via:

$$W_t = \left[\sum_{i=I_A}^{I_R-1} \left(\frac{1}{I_R - I_A} \right) \left(\frac{\widetilde{w}_{i,t}}{e_i} \right)^{\frac{1}{1-\lambda_t^l}} \right]^{1-\lambda_t^l} \quad (19)$$

In Erceg, Henderson and Levin (2000), among others, each household is considered to be a monopoly supplier of a differentiated labor service implying that they can set their own wage. In this paper, due to the overlapping generations nature of our model, we follow a slightly different route to incorporate this friction in the labor market. Namely, we consider that, for each age $i \in I_A, I_R - 1]$, there is a monopoly who buys labor hours directly to the households of age i at price $w_{i,t}$, and sells them to the representative firm producing the aggregate labor input at price $\widetilde{w}_{i,t}$. As usual, these monopoly suppliers set their price with a markup over their marginal cost which, in this case, implies that $\widetilde{w}_{i,t} = \lambda_t^l w_{i,t}$, $i \in I_A, I_R - 1]$. At the end of each period, these firms distribute their profits, $\pi_{L,i,t}$, $i \in I_A, I_R - 1]$, to the households in the form of dividends.

The monopoly power of these firms comes from the fact that, as considered in (17), labor hours of households of different ages are imperfect substitutes in the production of the aggregate labor input. In this set up, as opposed to Erceg, Henderson and Levin (2000), households do not have any monopoly power because their labor hours are perfect substitutes in the production of the aggregate labor input with those of all the other households in the economy with the same age. Nevertheless, for the purposes of this paper, the relevant issue is that there exists a distortion in the labor market that leads to a misalignment between prices and marginal costs and not whether the monopoly power is held by the households or by these intermediate labor producers.

Investment firm – In this small open economy, all capital is owned by a representative firm which rents it to the domestic intermediate producers at a unit price $p_{K,t}$ and takes investment decisions. Investment is assumed to be given by a CES aggregate of domestic and imported goods. Namely:

$$I_t = \left[(\alpha_I)^{\frac{1}{\eta_I}} I_{H,t}^{\frac{\eta_I - 1}{\eta_I}} + (1-\alpha_I)^{\frac{1}{\eta_I}} I_{F,t}^{\frac{\eta_I - 1}{\eta_I}} \right]^{\frac{\eta_I}{\eta_I - 1}} \tag{20}$$

where $(1-\alpha_I)$ is the share of imports in investment, I_H and I_F are the same composites of the continuum of domestic and of foreign intermediate goods as in (7) and (8), respectively, and η_I is the elasticity of substitution between these composite goods in investment. Thus, the unit price of this investment aggregate is given by:

$$p_{I,t} = \left[\alpha_I p_{H,t}^{1-\eta_I} + (1-\alpha_I) p_{F,t}^{1-\eta_I} \right]^{\frac{1}{1-\eta_I}} \tag{21}$$

and the demands of the domestic and foreign composites of the continuum of domestic and foreign intermediate goods, respectively, are given by:

$$I_{H,t} = \alpha_I \left(\frac{p_{H,t}}{p_{I,t}} \right)^{-\eta_I} I_t \tag{22}$$

$$I_{F,t} = (1-\alpha_I) \left(\frac{p_{F,t}}{p_{I,t}} \right)^{-\eta_I} I_t \tag{23}$$

We follow Christiano *et al.* (2005) and assume that this firm's investment decisions are conditioned by the existence of quadratic investment adjustment costs. As argued in Lucca (2007), these adjustment costs are equivalent, up to a first order linearization, to a time-to-build representation of the investment process. Furthermore, along the lines of Garrett and Priestley (2000), among others, we also consider that this firm faces costs of changing the amount of dividends it distributes to households at the end of each period. Thus, in each period t this representative firm chooses an investment sequence to maximize, given prices, its discounted flow of future dividends, net of the dividends adjustment costs:

$$\max_{\{I_s\}_{s=t,\,t+1,..}} \sum_{s=t}^{\infty} \left(\prod_{j=t+1}^{s} \frac{1}{1+r_j} \right) \left[d_s - \frac{\varsigma}{2} \left(\frac{d_s}{d_{s-1}} - 1 \right)^2 \right] \tag{24}$$

subject to:

$$d_s = (1-\tau_s^k) p_{K,s} K_s - p_{I,s} I_s \left(1 + S\left(\frac{I_s}{I_{s-1}} \right) \right) \tag{25}$$

$$K_{s+1} = (1-\delta) K_s + I_s \tag{26}$$

where $\varsigma > 0$ gives a measure of the dividends adjustment costs, τ^k is a proportional tax rate on this firm's capital rents, and, as usual, the investment adjustment cost function, $S(\cdot)$, satisfies that $S(1) = S'(1) = 0$ and $S''(1) \equiv \chi > 0$.

3.3 Government

The government of this small open economy consumes, gives lump-sum transfers, runs a social security system, levies taxes and issues debt. In each period, the government devotes an exogenously given amount of resources to consume, G_t, and to give lump-sum transfers to the households, $TR_t = tr_t\left(\sum_n \sum_{i=I_A}^{I} N_{i,n,t}\right)$. It is assumed that the government consumes the same final consumption good as households.[7] The government also spends resources in social security benefits $SS_t = \sum_n \sum_{i=I_R}^{I} N_{i,n,t}\ ss_{i,n,t}$. For each retired worker these benefits are assumed to represent a fraction ϖ of its average labor earnings in the last I_{SS} periods before retirement. In order to finance these expenditures, the government may issue debt, D_{t+1}, or levy proportional taxes on households' consumption (τ_t^c), labor income (τ_t^l and τ_t^{ss}) and capital income (τ_t^a), and on the investment firm's capital rents (τ_t^k). Thus, the government's budget constraint in period t is:

$$G_t + TR_t + SS_t + (1+r_t)D_t = D_{t+1} + \tau_t^c C_t + \left(\tau_t^l + \tau_t^{ss}\right)\sum_n \sum_{i=I_A}^{I_R-1} N_{i,n,t} w_{i,t}\ h_{i,n,t} + \\ + \tau_t^a r_t A_t + \tau_t^k p_{K,t} K_t \quad (27)$$

where $C_t = \sum_n \sum_{i=I_A}^{I} N_{i,n,t} c_{i,n,t}$ and $A_t = \sum_n \sum_{i=I_A}^{I} N_{i,n,t}(a_{i,n,t} + b_t)$ denote aggregate households' consumption and financial assets, respectively, D_t is the stock of public debt outstanding at the beginning of period t, and r_t is the exogenous interest rate in the small open economy.

As usual in models like this, a fiscal rule is needed so as to avoid explosive dynamics of public debt. We follow Kilponen *et al.* (2006) and, for the simulation exercises described in Section 5, we consider that the proportional labor income tax rate τ^l adjusts in each period to accommodate deviations of this rate and of the debt to GDP ratio from corresponding target levels according to the following rule:[8]

$$\tau_t^l = \tau_{t-1}^l - \kappa_1(\tau_{t-1}^l - \overline{\tau}^l) + \kappa_2\left(\frac{D_t}{GDP_t} - \overline{\frac{D}{GDP}}\right) \quad (28)$$

where $\kappa_1 > 0$ and $\kappa_2 > 0$ measure the sensitivity of τ^l to deviations of τ^l and $\frac{D}{GDP}$, respectively, from their targets.

---

7 Aggregate household consumption (C_t) involves consuming domestic ($C_{H,t}$) and foreign goods ($C_{F,t}$). Given equation (6), which defines the final consumption good, it is possible to derive that $C_{H,t} = \alpha_C\left(p_{H,t}\right)^{-\eta_C} C_t$ and $C_{F,t} = (1-\alpha_C)\left(p_{F,t}\right)^{-\eta_C} C_t$. The same applies to G_t, which can be divided into $G_{H,t}$ and $G_{F,t}$.

8 As it will be clear in Section 4, for the calibration of the model we do not use this fiscal rule. We simply fix the ratio of public debt to GDP to a target level and obtain a labor income tax rate τ^l endogenously so as to balance the government budget.

In our simulation exercises below, we use (28) looking for a compromise between long-term and short-term dynamics in the following sense. On the one hand, from a long-term perspective, it seems reasonable to consider, as in (28), that excessive debt scenarios need to be corrected via tax rate adjustments. For instance, there is a general consensus that the pressure on public accounts that will be induced in the near future by the aging of the Spanish baby-boom generation would require significant tax reforms. On the other hand, in the short run, in line with the small changes in tax rates typically observed in the data, it seems more reasonable to assume that (28) does not apply. Thus, in order to accommodate these long- and short-term perspectives, in the simulation exercises presented in Section 5 we assume that (28) only operates beyond 2008, so that τ^l stays constant at its value in the calibration exercise prior to that date. Beyond 2008, when the rule is at work, we consider that $\overline{\frac{D}{GDP}} = 0.60$, in line with the Stability and Growth Pact in the euro area, that $\bar{\tau}^l$ is equal to its value in the calibration exercise and, following Kilponen *et al.* (2006), that $\kappa_1 = 0.3$ and $\kappa_2 = 0.1$.

3.4 *Foreign economy*

As already mentioned above, in this small open economy a fraction $(1-\alpha_I)$ of aggregate investment and a fraction $(1-\alpha_C)$ of aggregate private and public consumption correspond to imports of foreign goods. By a symmetric argument, a fraction of the domestic production of intermediate goods is exported abroad in each period to become part of foreign aggregate consumption (private and public) and investment. In this sense, in terms of the composite good defined in (7), which aggregates all domestic intermediate goods, we assume that, in each period t, domestic exports abroad equal X_t, defined as:

$$X_t = \left(\frac{p_{H,t}}{p_{F,t}}\right)^{-\eta_X} Y_t^* \tag{29}$$

where η_X is the elasticity of substitution in the world economy between the domestic and foreign composite goods, and Y^* is a measure of the total demand in that economy, which is completely exogenous to the domestic economy. Equation (29) closes the model. See the Appendix for a formal definition its equilibrium.

4 Calibration

An initial goal of this paper is to evaluate, in the context of the model described in the previous section, the role played by interest rates, demographic developments, fiscal policy and market distortions in explaining the macroeconomic performance of the Spanish economy over its last expansionary phase. To carry out this quantitative exercise we first need to set the values of the parameters, the initial conditions and the exogenous sequences of the model. This section describes our calibration strategy and Table 1 summarizes our parameter choices.

Target year – We choose 1998 as our calibration target year, so as to focus on the post-Euro performance of the Spanish economy. In this sense note that, although the expansion of economic activity in Spain began some years before the creation of the EMU, this process and the build-up of imbalances in the economy clearly accelerated after 1998.

Initial distributions – A complete characterization of the model requires initial distributions of financial assets and of social security entitlements across households to be specified. Rather than setting those initial distributions arbitrarily in 1998, we set them for the year 1950, which then becomes the first year in our computations. By starting our numerical analysis in 1950, we are able to obtain (initial) distributions of financial assets and of social security entitlements across households in 1998 which are optimally derived from the model and, given that 1950 is far from 1998, do not depend on the initial distributions assumed in 1950.[9]

Demographics – A period in the model corresponds to 1 year. Agents start taking economic decisions at age 16, they are forced to retire at age 65, and die with probability 1 at age 100. We take the age structure of the population in 1950 from the UN World Population Prospects. For the period 1951-98 we propagate that population according to equations (1) and (2) using age-specific fertility and survival probabilities consistent with the evolution of average fertility and life-expectancy in the data. Thus, we consider the 1950-98 calibration period as a non-immigration period (only *native* households are active). In this sense note that it was after 1998, when immigrants represented less than 3 per cent of the Spanish population, when the largest immigration inflows into Spain took place.[10] Beyond 1998, we propagate the population under the assumption that households expect constant (at their 1998 levels) fertility and survival probabilities and do not anticipate the immigration inflows happening after 1998. These flows will be described and incorporated into the analysis in Section 5.

Preferences – We assume a standard CRRA specification of the per period utility function:

$$U(c,h) = \frac{\left(c^{\theta}(1-h)^{(1-\theta)}\right)^{1-\sigma}}{1-\sigma} \tag{30}$$

where σ and θ determine households' risk aversion and the relative importance of consumption over leisure, respectively. We set $\sigma = 3$, which falls within the standard range of this parameter in the literature, and choose θ so that in 1998 households in the model economy spend on average one-third of their time endowment at work.

Technology, foreign economy and discount factor – From households' point of view, in order to generate an empirically plausible age profile of asset holdings, it is necessary to account for the fact that earnings grow with experience. In this sense, the standard practice in the literature is to endow agents with an age-specific profile of productivity which in our model is represented by e_i. We have obtained this profile by computing average age-specific hourly wages from the Structural Earnings Survey (SES) in Spain in 2002.

The depreciation rate of capital δ, the discount parameter β and the exogenous world interest rate and output, r and Y^* respectively, are chosen simultaneously to reproduce the following targets in 1998: i) a ratio of investment to GDP of 23.5 per cent, ii) a ratio of international investment position to GDP of –31.7 per cent, iii) a ratio of net exports to GDP of –0.22 per cent and iv) a ratio of exports to GDP of 26.7 per cent. The values generated by these targets are δ=9.95 per cent, β=0.9964, $r = 3$ per cent and $Y^* = 0.0197$.[11]

---

[9] An additional reason to start our analysis in 1950 is that the demographic information provided by the UN World Population Prospects also starts in that year.

[10] Furthermore, if we were to consider the presence of immigrants in Spain before 1999, we would not have information available about the years in which those immigrants first entered in the country.

[11] In 1998, the calibrated model exhibits a ratio of current account to GDP of –1.04 per cent and a capital-output ratio of 2.25. The values of these non-targeted variables are reasonably close to those observed in the data (–0.11 per cent and 2.42, respectively).

We set $\alpha_C = 0.60$, $\alpha_I = 0.40$, $\lambda^l = 1.20$ and $\lambda^i = 1.10$. This implies a 60 per cent (40 per cent) share of domestic goods in consumption (investment) and a 20 per cent (10 per cent) markup in the labor (product) market. These values fall within the typical range for these parameters considered in the literature for the Spanish economy. Following Domenech and Taguas (1995) we also set $\xi = 0.375$. Finally, regarding the elasticities of substitution between domestic and foreign goods, we follow the work of Adolfson *et al.* (2007) and consider that $\eta_C = 5$ and $\eta_I = 2.5$. Furthermore, we assume that $\eta_X = \eta_C$.

Government – In 1998, government consumption, government transfers and public debt represented 17.3 per cent, 5 per cent and 64.1 per cent of the Spanish GDP, respectively. In line with this evidence, in the calibration exercise we choose sequences of government consumption, government transfers and public debt such that, in each period, the model economy replicates those ratios. As for taxes, the ratios of consumption taxes to private consumption, of social security contributions to labor income and of capital taxes to GDP observed in Spain in 1998 were equal to 18 per cent, 25.7 per cent and 5.2 per cent, respectively. Consequently, in the calibration, we consider a constant proportional consumption tax rate, τ_t^c, equal to 18 per cent, a constant proportional social security tax rate, τ_t^{SS}, equal to 25.7 per cent, and set constant capital income tax rates $\tau_t^a = \tau_t^k = 14.8\%$ such that the model economy replicates the latter ratio.[12]

Regarding to the social security system, the Spanish Regimen General de la Seguridad Social considers the last 15 years of contributions prior to retirement to compute the pension. Thus, we choose $I_{SS} = 15$ in our numerical exercises. As for the pension replacement rate, ϖ, we set this parameter such that in 1998 our model economy matches the ratio of social security expenses to GDP observed in the Spanish economy in that period (9.5 per cent). Finally, in each period of the calibration exercise we determine τ_t^l endogenously so that the government budget constraint (27) is met. The value of the labor income tax that satisfies this restriction is 10.3 per cent in 1998 (its counterpart in the data is 12.1 per cent).

5 Findings: the expansion

In Section 4, the model economy described in Section 3 was calibrated to replicate the main features of the Spanish economy in 1998. In this section we analyze the quantitative performance of this model economy beyond 1998 with a two-fold purpose. First, we want to evaluate to what extent the large decline in interest rates and the intense demographic changes observed in the Spanish economy after 1998 may explain the evolution of the main macro-aggregates in this country over the period 1998-2008. And second, we aim to quantify how this evolution could have changed under different scenarios concerning fiscal policy and labor and product markets distortions.

5.1 The role of demographic changes

After 1998 the Spanish economy has experienced a profound demographic change. This has been induced, not only by the large immigration inflows into the economy, but also by a change in the survival and fertility probabilities. To evaluate the macroeconomic impact of these developments

[12] Data on tax revenues are available at http://www.meh.es. National accounts data is available at http://www.ine.es.

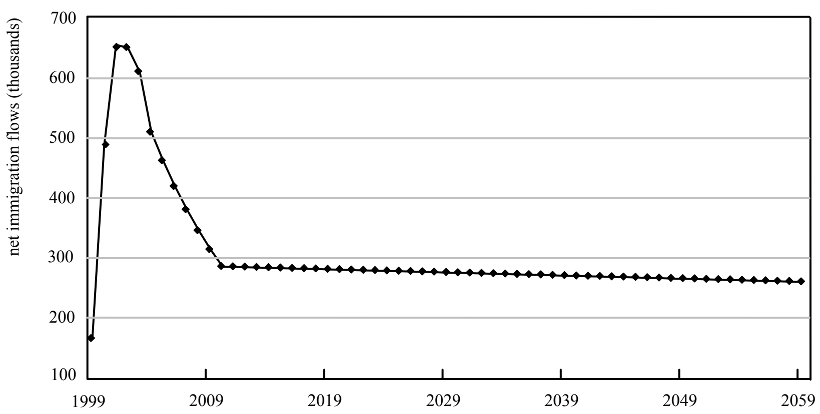
Table 1

Calibrated Parameters

Parameter	Value	Target
Demographics		
I_A	16	Assumed
I_R	65	Assumed
I	100	Assumed
$s_{i,t}$ and $f_{i,t}$	UN Population Prospects	Data
Preferences		
σ	3	Assumed
θ	0.5	Average labor hours = (1/3)
Tech. and foreign ec.		
e	Structural Earnings Survey	Data
δ	0.0995	Investment ratio=23.5%
β	0.9964	IIP over GDP=-31.7%
r	0.03	Net Exports over GDP=–0.22%
Y^*	0.0197	Exports over GDP=26.7
α_C	0.60	Assumed
α_I	0.40	Assumed
λ^l	1.20	Assumed
λ^{i}	1.10	Assumed
ξ	0.375	Domenech and Taguas (1995)
η_C	5	Adolfson *et al.* (2007)
η_I	2.5	Adolfson *et al.* (2007)
η_X	5	Assumed
χ	2	Christiano *et al.* (2005)
ς	15	Assumed
Government		
G/GDP	0.173	Data
TR/GDP	0.05	Data
D/GDP	0.641	Data
τ^c	0.18	Data
τ^{ss}	0.257	Data
τ^a=τ^k	0.148	Capital taxes/GDP=5.2%
I_{SS}	15	Data
ϖ	0.56	S.S. expeditures/GDP=9.5%

Figure 1

Net Immigration Flows in Spain, 1999-2059



in the context of our model, we depart from our economy calibrated to 1998 and assume that demographic variables evolve as follows:

Immigration inflows – We assume that, beginning in 1999, immigration flows behave according to Scenario 1 of the long-term demographic projections of the Instituto Nacional de Estadstica (Figure 1).[13] The age distribution of these immigration inflows is assumed to be constant over time and equal to that of new immigrants entering in Spain in 1999 according to the Estadistica de Variaciones Residenciales.[14] For computational reasons, we also assume that immigrants can only enter the country with ages between 16 and 44. In 1999, the immigrants who entered in Spain with these ages accounted for 60 per cent of the total.

Survival and fertility probabilities – Rather than assuming that these probabilities stay constant beyond 1998, as in the calibration exercise, we assume that they change over time in order to match the average fertility and life-expectancy data in the UN World Population Prospects for Spain. These projections end in 2050. Afterwards, we assume that the survival and fertility probabilities stay constant at their 2050 levels.

Findings – The response of our calibrated model economy to the demographic developments described above over the period 1998-2008 is summarized in Table 2. A first implication of these demographic changes is a fall in the dependency ratio of the economy. This can be seen in Figure 2, which compares, for the period 1998-2008, the dependency ratio in the calibration exercise (Baseline) and the associated to the demographic changes witnessed in Spain after 1998. In the model, this expansion of working-age population leads to a rise in aggregate employment, in aggregate investment and, consequently, in GDP. In this sense, according to the model, the observed demographic changes in Spain would have been responsible, on its own, for 60 per cent of the observed expansion in aggregate investment (Table 2).

The impact of these demographic changes is also strong in terms of public accounts. In particular, in the model the share of public debt in GDP gets significantly reduced from 64.1 per cent in 1998 to 53.8 per cent in 2008 (Table 2). This corresponds to 42 per cent of the improvement observed in this variable in the data and it has to do mostly with the increase in tax revenues associated to the expansion of economic activity.[15] Also note that the fall in the dependency ratio further contributes to a reduction in public debt via improving the balance of the social security system.

---

[13] Projections end in 2059. Afterwards, we assume that net immigration inflows stay constant at the 2059 level.

[14] These data refer to 10-years age groups. We make it annual by fitting a second order polynomial to the available age distribution.

[15] In this exercise we are keeping the share of government spending in GDP constant at a 17.3 per cent, as in the calibration exercise.

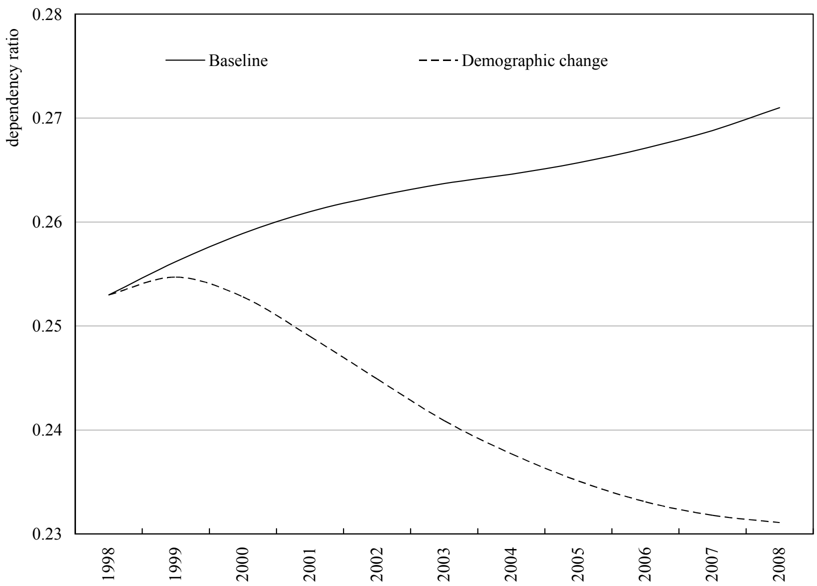
Table 2

Role of Demographic Changes

	Data		Model	
	1998	**2008**	**1998**	**2008**
Investment/GDP	23.5%	29.3%	23.4%	26.9%
Public Debt/GDP	64.1%	39.5%	64.1%	53.8%
Foreign Assets/GDP	–31.7%	–80.6%	–31.7%	–44.9%

Figure 2

The Dependency Ratio in the Short Run



As for impact of demographic developments on the external imbalance of the economy (measured as the ratio of net foreign assets to GDP) the aforementioned increase in investment in the model, together with minor changes in aggregate savings, impacts negatively on the current account and leads to a deterioration in the economy's international investment position. Thus, according to the model, 27 per cent of the deterioration in the ratio of net foreign assets to GDP observed in Spain over the period 1998-2008 could be explained by the demographic changes hitting the economy. However, a better accounting of these observed dynamics requires incorporating additional elements into the model economy. We do this next.

5.2 *The role of interest rates*

As illustrated in Figure 4, real interest rates in the Spanish economy fell significantly during its last expansionary phase. To evaluate the macroeconomic impact of this interest rate evolution, in this section we depart from our model economy calibrated to 1998 and consider, together with the demographic developments described above, two alternative interest rate scenarios.

Scenarios – To construct these scenarios we take the smoothed evolution of the *ex post* real short-term interest rate (measured as the one year Euribor) in Spain over the period 1998-2008 and consider two possibilities. In the first scenario we assume that the fall in interest rates observed over the period 1998-2008 is transitory, so that by 2010 the interest rate gets back to its 1998's

Figure 3

The Dependency Ratio in the Long Run

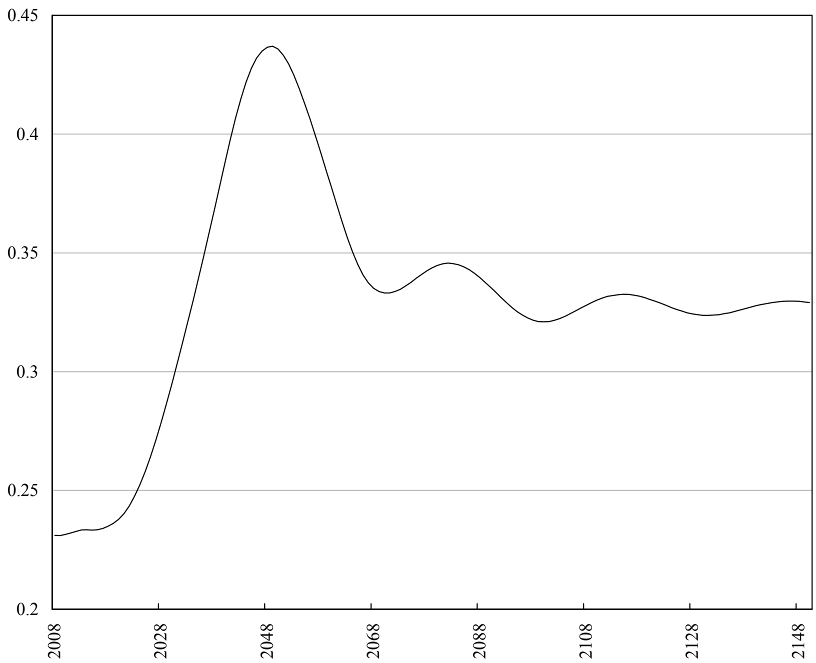
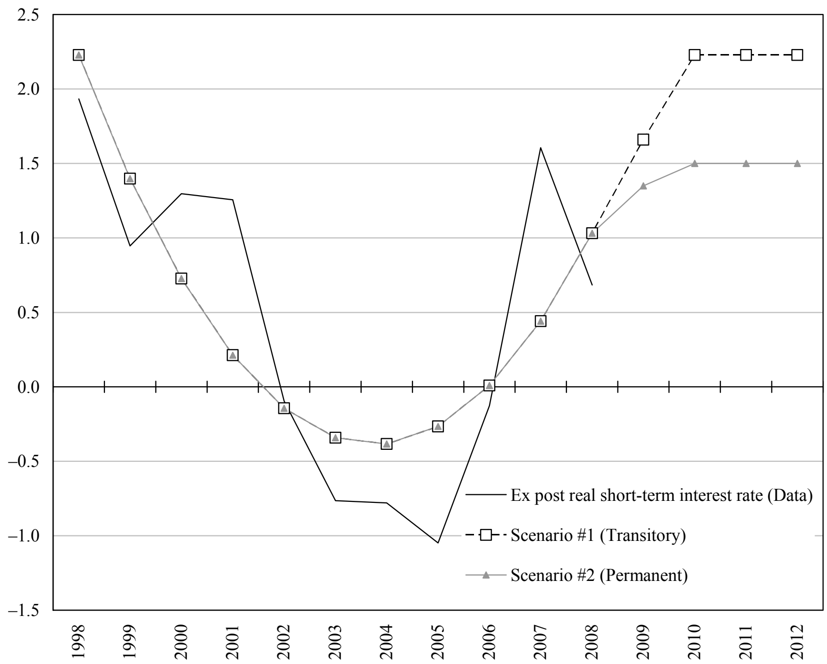


Figure 4

The Interest Rate Shock
(percent)



level, staying constant afterwards. In the second scenario, instead, we consider that the fall in interest rates is permanent. Namely, rather than returning to their 1998's levels, interest rates increase slightly between 2008-10 and stay constant at a 1.5 per cent level afterwards. These scenarios are depicted in Figure 4.

Findings – Departing from our initial state in 1998, we now incorporate into our model economy both the demographic changes described in Section 5.1 and the interest rate scenarios in Figure 4. In addition to the macroeconomic effects discussed above induced by demographic changes, the fall in interest rates (in both scenarios) has two main effects in the economy. Very intuitively, it contributes to a further expansion in aggregate investment and, via a reduction in the debt burden, to a more intense improvement in public debt (Table 3). It turns out that the former effect dominates so that the fall in interest rates leads to a further deterioration of the economy's international investment position. Naturally, all these effects are larger when the fall in interest rates is permanent (Scenario No. 2) rather than transitory (Scenario No. 1).

Table 3

Role of Interest Rates and Demographic Changes

	Data		Model (year 2008)	
	1998	**2008**	**Scenario No. 1**	**Scenario No. 2**
Investment/GDP	23.5%	29.3%	27.4%	30.1%
Public Debt/GDP	64.1%	39.5%	41.3%	40.1%
Foreign Assets/GDP	–31.7%	–80.6%	–63.9%	–77.3%

Not surprisingly, by adding the observed interest rate developments, the model delivers a better description of the evolution of the ratios of public debt and of net foreign assets to GDP in the Spanish economy over the period 1998-2008. Thus, according to the model, the developments in interest rates and demographic variables observed in the Spanish economy over this period would have been responsible for much of the observed improvement in public accounts in Spain (93 per cent in Scenario No. 1 and 97 per cent in Scenario No. 2) and of the deterioration of its external imbalance (66 per cent in Scenario No. 1 and 93 per cent in Scenario No. 2).

5.3 *The role of fiscal policy*

The previous section showed that much of the investment boom, the consolidation of public accounts and the increase in external indebtedness observed in the Spanish economy over the period 1998-2008 can be rationalized, in the context of our model economy, as the natural reaction of the economy to the observed developments in interest rates and demographic variables. This section analyzes to what extent this macroeconomic behavior would have changed if a different fiscal policy would have been in place. In particular, we study whether a more restrictive fiscal policy, involving a reduction in government consumption, could have attenuated the dramatic deterioration of the Spanish external position over the 1998-2008 period.

Thus, rather than assuming, as in the simulation exercises described above, that government consumption represents a constant fraction of GDP (17.3 per cent) in each period, we now consider two alternative fiscal policy scenarios. In these scenarios government consumption stays constant, in per capita terms, at its 1998's level for 10 (Scenario No. 1) and 20 (Scenario No. 2) years. Beyond 2008 in Scenario No. 1, 2018 in Scenario No. 2, government consumption represents again a 17.3 per cent of GDP in each period. Given that, as mentioned above, GDP increases in the model over the period of analysis in response to interest rate and demographic developments, these fiscal policy scenarios imply, in practice, a temporary reduction in the share of government expenditure to GDP, being this more permanent in Scenario No. 2. Namely, in Scenario No. 1 (Scenario No. 2) this share decreases smoothly from 17.3 per cent in 1998 to 15.8 per cent (14.8 per cent) in 2008 (2018).

Table 4 shows the results of these counterfactual exercises using as a benchmark the exercise described in Section 5.2 that incorporates demographic changes and a transitory fall in interest rates (Scenario No. 1). Intuitively, the model predicts that in both fiscal scenarios less government consumption over the period 1999-2008 would have led to a more intense improvement in public accounts by 2008 than that in the benchmark case. This fiscal tightening, however, would have

Table 4

Role of Fiscal Policy

	Model (Year 2008)		
	Benchmark	**Fiscal Scenario No. 1**	**Fiscal Scenario No. 2**
Investment/GDP	27.4%	27.5%	28.0%
Public Debt/GDP	41.3%	36.7%	36.9%
Foreign Assets/GDP	–63.9%	–63.0%	–75.9%

helped very little in attenuating the build-up of the economy's external imbalance over this period. In particular, the transitory tightening of government consumption in Scenario No. 1 would have only reduced the size of this imbalance by 2008 by 1 percentage point. The more permanent tightening of fiscal policy in Scenario No. 2 would have even increased that imbalance.

The intuition behind this little effectiveness of fiscal policy in addressing the economy's external imbalance lies on the forward-looking behavior of households in the model. Certainly, a reduction in government consumption leads to an improvement in public accounts and this, by itself, attenuates the economy's need for external financing. However, to the extent that households anticipate that the reduction in the share of public debt to GDP is going to imply a reduction in labor income taxes in the future (once the fiscal rule operates), they immediately modify their labor and consumption profiles so that current private borrowing increases. This increase therefore counteracts the fall in public financing needs and, depending on the temporal dimension of the fiscal tightening, it may even imply a more intense deterioration in the economy's external imbalance.

5.4 *The role of labor and product market distortions*

The Spanish economy is characterized by the existence of important distortions in the labor and product markets, which hinder productivity growth, a proper allocation of resources and, more broadly, damage the external competitiveness of the economy. In this section we explore how structural reforms on these markets could have affected the macroeconomic performance of the Spanish economy over the period 1998-2008. In particular, we study the reaction of our model economy to a 2 percentage points decrease in the labor and product markets markups. Christopoulou and Vermeulen (2008) report that the markup in the U.S. manufacturing sector was, on average over the period 1993-2004, 6 percentage points greater than in the Euro Area. In a similar vein, Andres, Ortega and Valles (2008) argue that a 5 percentage points differential in the product market markup is a conservative estimate of the importance of markup differences across European markets. In this sense, our simulated reduction in markups would entail closing around one third of these differences in the competitive environment. A number of papers in the literature have conducted quantitative exercises similar to ours. For instance, Gomes *et al.* (2009) show the macroeconomic implications of a decline in German markups in the product and labor markets of 5, 10 and 15 percentage points Moreover, Kilponen and Ripatti (2006) show the effects for the Finnish economy of a reduction of 5 percentage points in the labor market markup and of 2 percentage points in the product market markup.

As in the previous section, we use the exercise described in Section 5.2 that incorporates demographic changes and a transitory fall in interest rates (Scenario No. 1) as a benchmark.

Departing from that benchmark, we first consider a labor market reform setting λ^l= 1.18 (rather than 1.20) and then a product market reform setting λ^i = 1.08 (rather than 1.10). Not surprisingly, reducing the inefficiencies in these markets leads, compared to the benchmark, to an expansion of economic activity, with increases in aggregate investment and employment, and to an improvement in external competitiveness. According to the model, the positive effects of the same 2 percentage points reduction in the markup are larger if the reform is carried out in the product market rather than in the labor market. Namely, with a product market reform, GDP, employment and the terms of trade would have been, by 2008, 1.5 per cent higher, 0.7 per cent higher and 0.3 per cent lower, respectively, than with a labor market reform. On the long run, these differences persist: GDP, employment and the terms of trade would have been 1.6 per cent higher, 0.6 per cent higher and 0.3 per cent lower, respectively, with a 2 percentage points decrease in the product market markup than with the same decrease in the labor market markup.

The impact that these structural reforms would have had on Spanish publics accounts and on the economy's external imbalance over the period 1998-2008 is summarized in Table 5. Due to the aforementioned expansion of economic activity, the consolidation of public accounts over this period would have been more intense with the reforms. The external imbalance of the economy, however, would have been higher by 2008 if the reforms had been carried out. The reason for this result is that, as in the case of fiscal policy, households anticipate lower taxes and a more efficient economy in the future. Thus, in order to smooth consumption, they increase current private borrowing what, together with the increase in aggregate investment, dominate the improvement in public accounts and then lead to a more intense external indebtedness.

Table 5

Role of Labor and Product Market Distortions

	Model (Year 2008)		
	Benchmark	**Labor Market Reform**	**Product Market Reform**
Investment/GDP	27.4%	27.6%	28.4%
Public Debt/GDP	41.3%	38.0%	36.6%
Foreign Assets/GDP	–63.9%	–65.0%	–68.4%

Fiscal policy vs. structural reforms – According to the model, an structural reform in the product market (like the one considered in this section) could achieve a short-run reduction (over the 1998-2008 period) in the ratio of public debt to GDP similar to that achieved with the fiscal tightening exercises presented in Section 5.3. This, together with fact that the long-term positive effects of structural reforms on GDP, employment, investment and competitiveness are absent with alternative fiscal policy experiments, strongly point to structural reforms as a powerful instrument to pursue improvements in the economy's public accounts, not only for the period 1998-2008 but for the future. In this sense, as mentioned above, the fact that these reforms may lead to a deterioration in the economy's external imbalance in the short run should not be worrisome. It naturally comes from i) increased investment (once inefficiencies have been reduced) and ii) households' smoothing behavior (as the economy will be wealthier in the future with less distortions).

6 Beyond the financial crisis

The previous section showed the performance of the model over the period 1998-2008 under alternative scenarios. This section, instead, studies its implications beyond 2008. To do that in a meaningful way, we first introduce several shocks into the model economy in 2008 that aim to capture the arrival of the global financial crisis to the Spanish economy. Once these shocks have been incorporated, we address two questions: i) how the macroeconomic evolution predicted for the Spanish economy beyond 2008 has changed due to the global financial crisis and ii) to what extent that predicted evolution may be altered by fiscal policy and structural reforms.

6.1 The global financial crisis

In order to incorporate into our model economy the arrival of the global financial crisis in 2008 we take, as a benchmark, the exercise described in Section 5.2 including demographic changes and a transitory fall in interest rates (Scenario No. 1), and assume that the economy is hit by the following shocks in 2008:

(S1) Between 2008 and 2009 capital depreciates an additional 10 per cent and the depreciation rate goes back smoothly to its initial calibrated level ($\delta = 0.0995$) in 5 years.

(S2) Beyond 2008, during 20 years, TFP growth is 1 percentage point smaller than in the benchmark.

(S3) In 2009, the age-specific index, e_i, which transforms households' raw labor hours into efficient units of labor, decreases by 10 per cent and it goes back to its initial level after 5 years.

(S4) In 2009, the share of government consumption plus government transfers in GDP increases by 5 percentage points with respect to the benchmark and it goes back smoothly to its calibrated level ($\frac{G+TR}{GDP} = 0.223$) in 10 years.[16]

Certainly, the global financial crisis has hit the Spanish economy in many different dimensions, some of which can not be incorporated into the framework developed in this paper. Consequently, the aim of this exercise is not to perform a full account of the quantitative implications of the crisis, but to broadly incorporate its main consequences. In this sense, shock (S1) tries to capture the view that much of the investment made in Spain during the last expansionary phase was not as productive as initially thought so that its value will need to adjust gradually. The second shock (S2) we consider is a fall in TFP growth. This could be justified on two grounds: as a revision of households' expectations about future growth prospects (maybe too optimistic before the crisis) and/or as a way of capturing the real effects of the financial turmoil. The recent crisis has also led to a very rapid and intense increase in unemployment, that rose from around 9 per cent in 2007 to around 19 per cent in 2009, whose future reduction is expected to be very gradual. Shock (S3) incorporates into our model economy the effects of this fall in the workforce. Finally, with the arrival of the crisis public accounts in Spain have experienced a quick and intense deterioration. This has been mostly due to the work of automatic stabilizers and to the expansionary programs put in place by the government to mitigate the effects of the crisis. In our model, this is captured by shock (S4). It should be noted that, although the magnitude of these shocks has not been chosen to replicate the observed deterioration of macroeconomic aggregates

[16] In each year, the increase in $\frac{G+TR}{GDP}$ with respect to the benchmark is split evenly between government consumption and transfers.

over the crisis period, they are able to generate a contraction of real GDP between 2008 and 2009 of –1.7 per cent, which is in line with the HP-filtered growth rate observed in the data (–1.66 per cent). Furthermore, in terms of the dynamics of public accounts, the model generates an increase in public deficit of 6 percentage points (7 percentage points in the data).

6.2 *The effects of the crisis*

Departing from our benchmark in 2008, Figure 5, 6 and 7 show the short-run behavior of our model economy beyond that date with and without the arrival in 2008 of the global financial crisis, represented by the shocks described in Section 6.1.

Two main conclusions can be extracted from this figure. First, even without the arrival of the crisis, in the short-run the Spanish economy would have entered into a phase of lower, even negative, GDP growth where the external imbalance of the economy would have been reduced but where public accounts would have deteriorated. In this non-crisis scenario, that already incorporates an increase in interest rates

Figure 5

GDP During the Crisis

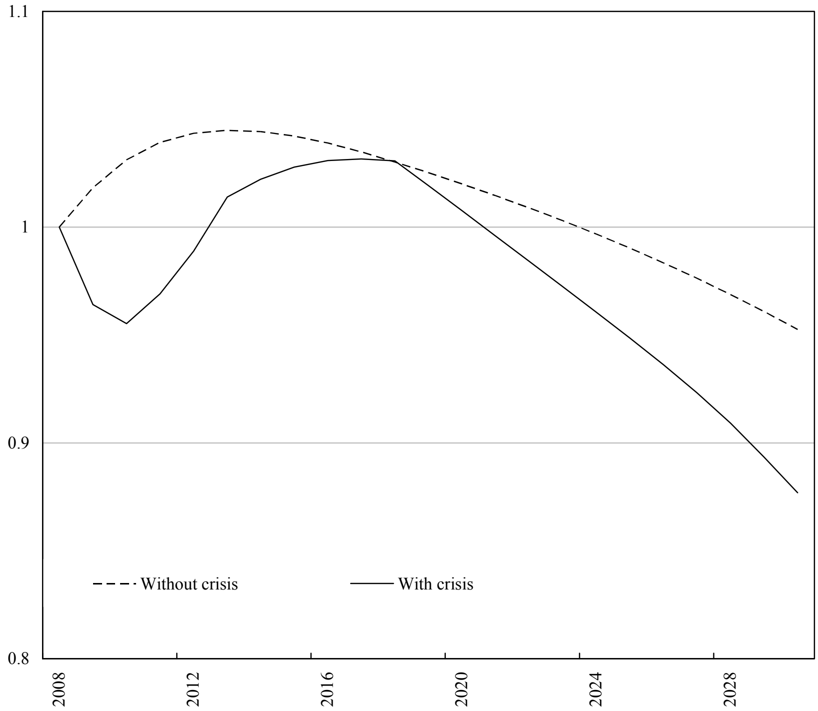


Figure 6

Public Debt During the Crisis

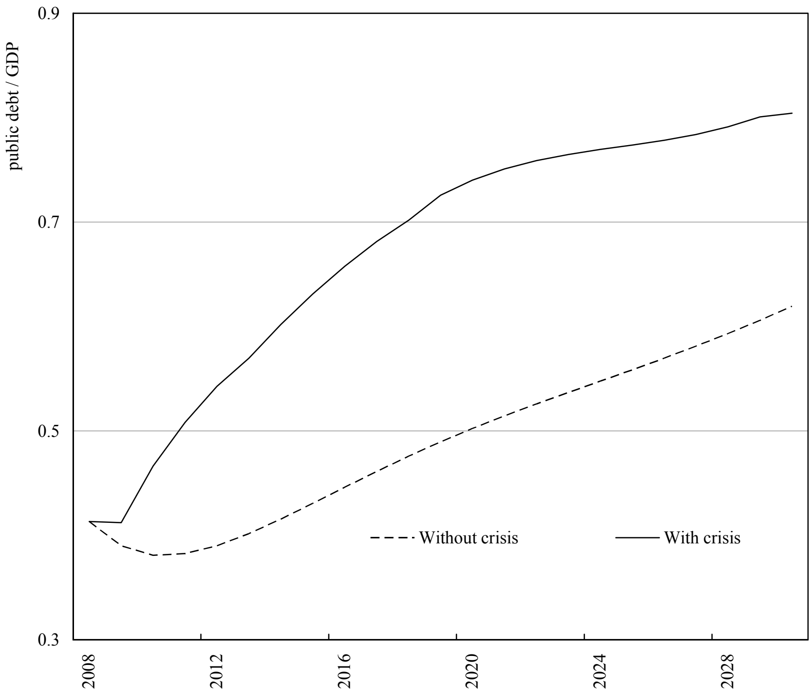
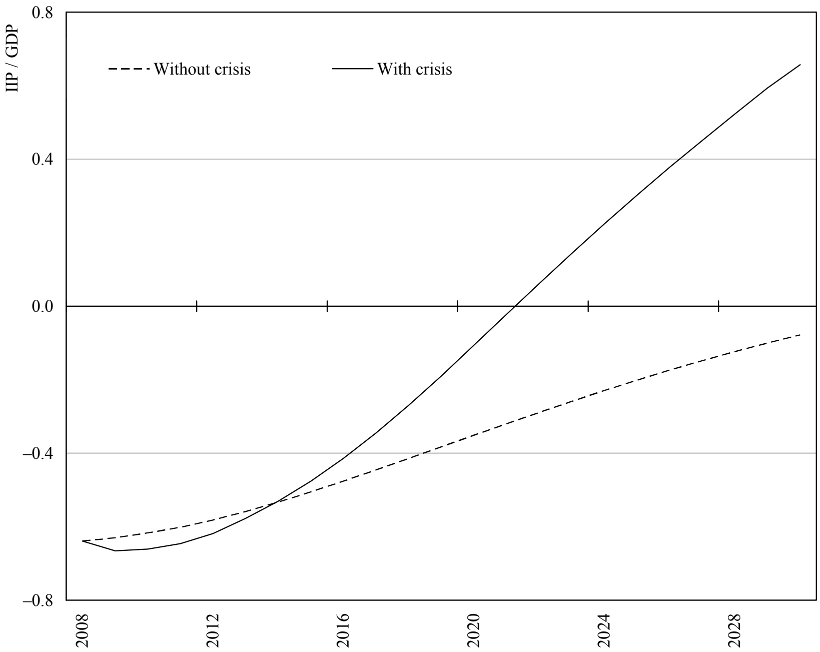


Figure 7

International Investment Position During the Crisis



beyond 2008, these dynamics are driven to a large extent by demographic changes. Namely, by the behavior of the dependency ratio in Spain that, after decreasing over the period 1998-2008 due to immigration inflows, increases again in 2009 and especially over the period 2020-50 with the aging of the Spanish baby boom generation. In the short-run, this reduction of the working-age population has an obvious negative effect on GDP and investment (which helps to correct the external imbalance of the economy) and, due to social security expenditures, on public accounts.

And second, in the short-run, the arrival of the global financial crisis accentuates the aforementioned dynamics that would have taken place without the crisis. Namely, the deterioration of GDP growth and public accounts and the improvement of the economy's external imbalance. To the extent that the crisis constitutes a negative wealth shock for the economy, households' consumption reduces substantially with respect to the non-crisis scenario, what explains the more intense improvement in the economy's the external imbalance. In addition, the shock increasing government consumption and transfers clearly leads to a deeper deterioration in public accounts. Finally, the negative shocks to the capital stock and to the age-specific index, e_i, cause an immediate fall in GDP after the crisis, which only recovers gradually. In this sense, the model predicts that by 2018 the economy is able to get back to the GDP level that would have had without the arrival of the crisis. At that moment, however, GDP starts deviating again from the non-crisis scenario. The reason is that in 2018 the fiscal rule begins to operate in the crisis scenario and, given the increase in the debt to GDP ratio over the period 2009-18, it leads to a substantial increase in labor taxes. Intuitively, this tax increase improves the dynamics of public accounts beyond 2018 but cause a fall in aggregate labor what has an adverse effect on GDP. Note, however, that the global financial crisis does not have any long-term effect on the economy and, in particular, on GDP. This is a natural consequence of the way in which we have modelled the crisis in Section 6.1, exclusively through temporary shocks.

6.3 *The role of structural reforms and fiscal policy*

Product market reform – Figures 8 and 9 show the short-run behavior of our model economy beyond 2008 when, besides the shocks coming from the global financial crisis, in 2009 the economy experiences a permanent fall in the markup in the intermediate goods market of

2 percentage points (λ^i falls from 1.10 to 1.08). Very intuitively, with respect to a non-reform scenario (*With crisis*), in the short-run this structural reform mitigates the deterioration in output and in public accounts caused by the global financial crisis without significantly affecting the economy's external imbalance. The effect of the reform on GDP is particularly strong. Namely, with the reform the economy, not only is able to get back to the GDP level that it would have had without the arrival of the crisis before compared to a non-reform scenario (2014 vs. 2018), but it may even enjoy for some time (over the period 2014-25) of greater output levels compared to a non-crisis scenario. In addition to these short-term effects, as mentioned in Section 5.4, this structural reform also has important long-term effects. In particular, as shown in Figure 14, with a reduction in the distortions in the product market output is higher and public debt is lower in the long-run.

Fiscal policy – In the crisis scenario (*With crisis*) considered so far the arrival of the global financial crisis was accompanied by a shock to the share of government consumption plus government transfers in

Figure 8

The Role of Product Market Reform During the Crisis: GDP Dynamics

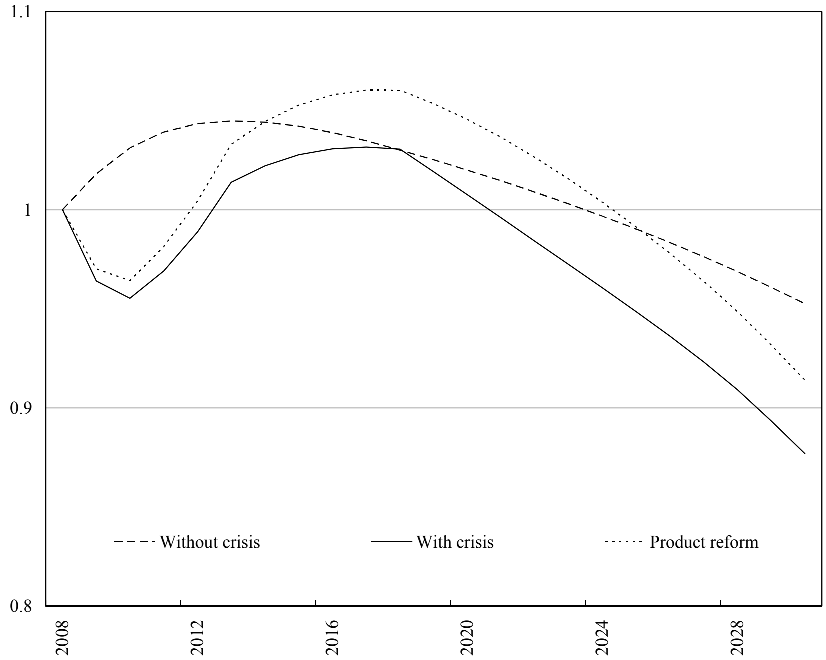
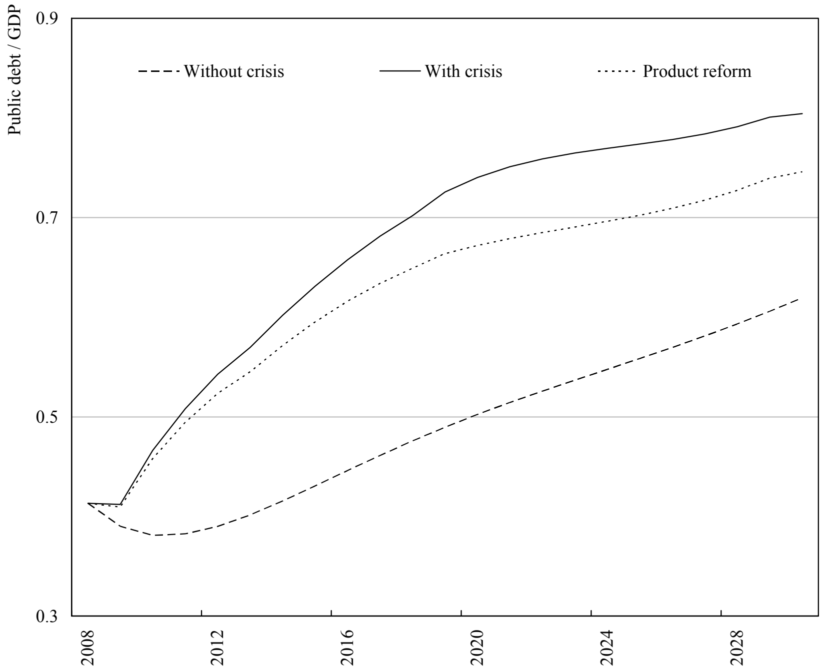


Figure 9

The Role of Product Market Reform During the Crisis: Public Debt Dynamics



GDP. In particular we have assumed that in 2009 this share increases by 5 percentage points with respect to the non-crisis scenario and it goes back smoothly to its calibrated level ($\frac{G+TR}{GDP}=0.223$) in 10 years. In order to evaluate how fiscal policy can modify the macroeconomic evolution of the Spanish economy beyond 2008, we now consider two alternative fiscal scenarios: Scenario No. 1, in which the initial size of the shock is smaller ($\frac{G+TR}{GDP}$ increases by 2.5 percentage points with respect to the benchmark in 2009) and it takes 10 years to get back to the benchmark, and Scenario No. 2, in which the initial size of the shock is the same but it reverts to the benchmark in 5 years rather than 10. In this context, a tighter fiscal policy leads to a less intense deterioration in public accounts and, for the reasons detailed in Section 5.3, it has very little effects on the economy's external imbalance. In terms of output, however, a tighter fiscal policy causes a slightly more intense deterioration of GDP in the years immediately after the arrival of the crisis, but more importantly, it substantially mitigates the output loss predicted beyond 2018 compared to the *With crisis* scenario. With a less expansive fiscal policy the debt to GDP ratio increases less over the period 2009-18 so that the increase in taxes needed afterwards to correct this imbalance is smaller, what has a less adverse effect on aggregate labor and output. Therefore, there is trade-off: while a more expansive fiscal policy helps mitigating the immediate output losses caused by the crisis, it has a negative effect on output recovery in medium-term.

Fiscal policy vs. structural reforms – Summing up, as pointed out in Section 5, the model predicts a very little role for fiscal policy and structural reforms to modify the behavior of the economy's external imbalance in the short-run. In terms of limiting the deterioration of public accounts, however, both tighter fiscal policies and structural reforms may help substantially. More importantly, in terms on output, the model predicts a different reaction of the economy to these alternative policies and structural reforms seem to outperform tighter fiscal policies. Thus, while structural reforms help in reducing output losses in the short- and medium-term and have a positive long-run effect on the level of output, tighter fiscal policies, which do not exhibit long-term effects, face a trade-off: they may reduce the output losses caused by the crisis in the medium-term but at the expense of a more intense output loss in the years immediately after the arrival of the crisis.

7 Concluding remarks

The emergence of a huge current account deficit was one of the main characteristics marking developments in the Spanish economy during the period of robust economic growth prior to the current crisis. This paper tries to disentangle the main drivers behind this upswing. To this end, we calibrate a small open-economy model for Spain that replicates relatively well the main features of the Spanish economy during the last decade. According to this model two main factors arise as particularly relevant in explaining these developments. First, the decline in interest rates derived from Spain's participation in the European Monetary Union; and further, the far-reaching demographic change brought about by huge immigration flows.

Apart from the role of these two factors, which have already been emphasized by the existing literature, our paper investigates the role played by economic policies in the build-up of the Spanish external imbalance. First, considerable attention has been given in the related literature to the potential role fiscal policy might play in the reduction of this imbalance. In this paper, the role of fiscal policy is analyzed by means of two counterfactual scenarios that try to measure what the external imbalance would have been if significantly tighter fiscal policies had been applied during the last decade. This restrictive fiscal policy is simulated through lower public expenditure growth than that observed in the data. Our results show that the role that a tightening of fiscal policy could have played in the reduction of the Spanish external imbalance would have been very limited and would have depended on the temporal dimension of this tightening. A transitory change in fiscal policy would have reduced the economy's external imbalance only very slightly, by affecting public savings without significantly distorting private ones. Instead, a permanent fiscal tightening would have had a negative effect on the economy's net foreign assets as it would have distorted optimal decisions by forward looking agents and reduced private savings.

These results need to be put in a new perspective under the current economic conditions, where the Spanish budget deficit has increased dramatically over the past two years, causing a significant increase in public debt, and where financing conditions for private agents are considerable tighter. Under these conditions, the situation is closer to the classical "twin deficits" scenario in which the current account imbalance is in close relation with public deficit. In this scenario, fiscal consolidation is needed to correct the external imbalance.

Second, we investigate the role played by labor and product market reforms in the correction of this imbalance. This is relevant insofar as the Spanish economy experienced a progressive increase in its prices and costs relative to those of its main competitors during the economic boom, which may have had an effect on net exports, and there is evidence that this rise in relative prices and wages is related to labor market rigidities and insufficient competition in some markets. Our results show that, if structural reforms in labor and product markets had been adopted in the Spanish economy over the period 1998-2008, the expansion of economic activity, investment and employment would have been more intense than the one observed over that period. The external competitiveness of the economy would have also improved relative to a non-reform scenario and the improvement in public accounts would have been larger. These reforms, however, would have implied a further deterioration of the Spanish external imbalance over the 1998-2008 period. Increased investment, once market distortions had been reduced, and reduced private savings, as households try to smooth their consumption anticipating lower taxes and a more efficient economy in the future, would be responsible for this further deterioration. Nevertheless, it is worth mentioning that, despite this short-run effect on the economy's external imbalance, according the model structural reforms, besides improving GDP, employment, investment and competitiveness in the long-run, constitute a very effective policy instrument to achieve fiscal consolidation.

The framework set out in this paper has also been used to analyze the different policy options faced by the Spanish economy after the crisis. Several results are worth mentioning. First, even without the arrival of the crisis, in the short-run the Spanish economy would have entered into a phase of lower GDP growth where the external imbalance of the economy would have been reduced but where public accounts would have deteriorated. The arrival of the global financial crisis has accentuated the aforementioned dynamics. Second, the model highlights the trade-off faced by tighter fiscal policies in the post-crisis scenario: they may reduce the output losses induced by the crisis in the medium-term but at the expense of a more intense output loss in the years immediately after the crisis. In contrast, structural reforms do not face this trade-off and may contribute to reduce output losses in the short- and medium-term, while inducing a positive long-run effect on the level of output. In light of the potential benefits of those policies aimed at improving competition, it would be interesting to further explore their interactions with total factor productivity growth.

APPENDIX

Definition of equilibrium

The equilibrium of the model is a list of sequences of:

- prices $\{r_t, p_{H,t}, p_{H,j,t}, p_{F,t}, p_{F,z,t}, p_{K,t}, p_{I,t}, W_t, \widetilde{w}_{i,t}, w_{i,t}\}$,
- taxes $\{\tau_t^c, \tau_t^l, \tau_t^{ss}, \tau_t^a, \tau_t^k\}$,
- transfers $\{b_t, ss_{i,n,t}, tr_t, SS_t, TR_t\}$,
- and quantities $\{N_{i,n,t},\ c_{i,n,t},\ C_t,\ C_{H,t},\ C_{F,t},\ a_{i,n,t},\ A_t,\ h_{i,n,t},\ Y_t,\ Y_{H,t},\ y_{H,j,t},\ Y_{F,t},\ L_t,\ L_{i,t},\ L_{H,j,t},\ K_t,\ K_{H,j,t},\ I_t,\ I_{H,t},\ I_{F,t},\ G_t,\ G_{H,t},\ G_{F,t},\ D_t,\ X_t,\ Y_t^*\}$,

such that, at each point in time t:

- the age structure of the population follows the law of motions (1)-(3),
- agents maximize lifetime utility (4) subject to the period by period budget constraints (5),
- all firms maximize profits,
- accidental bequests are given by:

$$b_t = \frac{\sum_n \sum_{i=I_A}^{I} \mu_{i-1,n,t-1} a_{i,n,t} (1 - s_{i-1,t-1})}{(1 + np_{t-1}) \sum_n \sum_{i=I_A}^{I} \mu_{i,n,t}} \tag{31}$$

where np_{t-1} is the population growth rate between periods $t-1$ and t,

- dividends received by households are equal to:

$$div_t = \frac{\int_0^1 \pi_{H,j,t} dj + \sum_{i=I_A}^{I_R - 1} \pi_{L,i,t} + d_t}{\sum_n \sum_{i=I_A}^{I} N_{i,n,t}} \tag{32}$$

- the budget constraint of the government (27) is satisfied,
- labor markets clear:

$$L_{i,t} = \sum_n \mu_{i,n,t} h_{i,n,t} \tag{33}$$

$$L_t = \left[\sum_{i=I_A}^{I_R - 1} \left(\frac{1}{I_R - I_A} \right)^{\frac{\lambda_t^l - 1}{\lambda_t^l}} \left(e_i L_{i,t} \right)^{\frac{1}{\lambda_t^l}} \right]^{\lambda_t^l} \tag{34}$$

$$\int_0^1 L_{H,j,t} dj = L_t \tag{35}$$

- the market for physical capital clears:

$$K_t = \int_0^1 K_{H,j,t} dj \tag{36}$$

- the market for the composite of domestic intermediate goods clears:

$$Y_{H,t} = C_{H,t} + I_{H,t}(1+S(\cdot)) + G_{H,t} + X_t \tag{37}$$

- and the aggregate budget constraint of the economy holds:

$$C_t + p_{I,t} I_t (1+S(\cdot)) + G_t + DAC_t + FA_{t+1} - (1+r_t) FA_t = p_{H,t} Y_{H,t} \tag{38}$$

where $FA_{t+1} = A_{t+1} - D_{t+1}$ denotes the net foreign asset position of the country at the end of period t and DAC_t is the dividends adjustment cost.

REFERENCES

Adolfson, M., S. Laseen, J. Linde and M. Villani (2007), "Evaluating an Estimated New Keynesian Small Open Economy Model", Sveriges Riksbank, Working Paper, No. 203.

Andres, J., E. Ortega and J. Valles (2008), "Competition and Inflation Differentials in EMU", *Journal of Economic Dynamics and Control*, Vol. 32, No. 3, pp. 848-74.

Arce, O., J.M. Campa and A. Gavilan (2008), "Asymmetric Collateral Requirements and Output Composition", Banco de España, Working Paper, No. 0837.

Auerbach, A.J. and L.J. Kotlikoff (1987), *Dynamic Fiscal Policy*.

Banco de España (2006), *Informe Anual*.

Blanco, R. and F. Restoy (2007), "Have Real Interest Rates Really Fallen That Much in Spain?", Banco de España, Working Paper, No. 0704.

Christiano, L.J., M. Eichebaum and C.L. Evans (2005), "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy", *Journal of Political Economy*, Vol. 113, No. 1, pp. 1-45.

Christopoulou, R. and P. Vermeulen (2008), "Markups in the Euro Area and the US over the Period 1981-2004 – A Comparison of 50 Sectors", European Central Bank, Working Paper, No. 856.

Collado, M.D., I. Iturbe-Ormaetxe and G. Valera (2004), "Quantifying the Impact of Immigration on the Spanish Welfare State", *International Tax and Public Finance*, Vol. 11, No. 3, pp. 335-53.

Domenech, R. and D. Taguas (1995). "Potential Output Estimations in the Spanish Economy", paper presented at the conference on *Medium Term Economic Assessment*, Helsinki.

Erceg, C., D. Henderson and A. Levin (2000), "Optimal Monetary Policy with Staggered Wage and Price Contracts", *Journal of Monetary Economics*, Vol. 46, No. 2, pp. 281-313.

Estrada, Á., J.F. Jimeno and J.L. Malo de Molina (2009), "The Spanish Economy in EMU: The First Ten Years", Banco de España, Occasional Paper, No. 0901.

Garrett, I. and R. Priestley (2000), "Dividend Behavior and Dividend Signaling", *Journal of Financial and Quantitative Analysis*, Vol. 35, No. 2, pp. 173-89.

Gimeno, R. and J.M. Marques (2008), "Uncertainty and the Price of Risk in a Nominal Convergence Process", Banco de España, Working Paper, No. 0802.

Gómes, S., P. Jaquinot, M. Mohr and M. Pisani (2009), "Macroeconomic Implications of Greater Competition in the Euro Area", mimeo.

Izquierdo, M., J.F. Jimeno and J. Rojas (2010), "On the Aggregate Effects of Immigration in Spain", *SERIEs: Journal of the Spanish Economic Association*, Vol. 1, No. 4, pp. 409-32.

Jimeno J.F., J. Rojas and S. Puente (2008), "Modelling the Impact of Demographic Aging on Social Security Expenditures", *Economic Modelling*, Vol. 25, No. 2, pp. 201-24.

Kilponen, J., H. Kinnunen and A. Ripatti (2006), "Population Ageing in a Small Open Economy: Some Policy Experiments with a Tractable General Equilibrium Model", Bank of Finland, Research Discussion Paper, No. 28/2006.

Kilponen, J. and A. Ripatti (2006), "Labour and Product Market Competition in a Small Open Economy – Simulation Results using a DGE Model of the Finnish Economy", Bank of Finland, Research Discussion Paper, No. 5/2006.

Krueger, D. and A. Ludwig (2007), "On the Consequences of Demographic Change for Rates of Returns to Capital, and the Distribution of Wealth and Welfare", *Journal of Monetary Economics*, Vol. 54, No. 1, pp. 49-87.

Lucca, D.O. (2007), "Resuscitating Time-to-Build", Board of Governors of the Federal Reserve System, mimeo.

Razin, A. and E. Sadka (1999), "Migration and Pension with International Capital Mobility", *Journal of Public Economics*, Vol. 74, pp. 141-50.

Rojas, J.A. (2005), "Life-cycle Earnings, Cohort Size Effects and Social Security: A Quantitative Exploration", *Journal of Public Economics*, Vol. 89, pp. 465-85.

Storesletten, K. (2000), "Sustaining Fiscal Policy Through Immigration", *Journal of Political Economy*, Vol. 108, No. 2, pp. 300-23.

—————(2003), "Fiscal Implications of Immigration: A Net Present Value Calculation", *Scandinavian Journal of Economics*, Vol. 105, No. 3, pp. 487-506.

EQUITY ASPECTS OF VAT IN EMERGING EUROPEAN ECONOMIES: THE CASE STUDY OF SERBIA

Nikola Altiparmakov[*] and Milojko Arsić[**]

Studies of VAT incidence in developed European economies reveal a regressive distribution in any particular year, but mildly progressive lifetime incidence. Micro-simulation analysis of Serbian expenditure survey data yields similar conclusions. However it is important to clearly recognize two distinctive features of emerging European economies when analyzing the VAT incidence. Firstly, we show that significant presence of own-source small farming production in many emerging European countries, including Serbia, presents an additional progressivity-enhancing buffer compared to VAT incidence in developed economies. Secondly, the high level of shadow economy and evasion of direct income taxes in many emerging European countries suggests that household expenditures are a more meaningful indicator of the living standard and ability to pay taxes than the registered income. Overall, we conclude that common beliefs of regressive VAT taxation, often encountered in the general public, are vastly overstated and poorly founded in economic reality of emerging European countries.

Introduction

Tax systems around the world are continuously changing in response to economic, political and administrative developments. Rapid globalization during the last couple of decades introduced unprecedented international mobility of capital, goods and services, and (to a certain extent) labor, consequently causing a world-wide trend of reducing custom duties, corporate income taxes and tax wedges on labor. Significant reductions in corporate and personal tax rates on capital and labor incomes have been especially stark in emerging European countries, which experienced a fierce (income) tax competition during the last decade in order to attract foreign investors – the so called "race to bottom" phenomenon.

Faced with reduced revenues from other sources, EU countries are increasingly relying on consumption taxation. European Commission (2009) notes that reliance on consumption taxes, and VAT in particular, has been continually increasing in EU member states in the 2000-07 period. Policy importance of consumption taxation is also highlighted by the renewed attention to the optimal tax-mix issues, due to strong theoretical and empirical evidence that consumption taxes are less disruptive to economic growth than direct income taxes (Johansson *et al.*, 2008). Some European countries have already implemented efficiency-driven tax reforms which shift the burden from income to consumption taxation – Germany in 2007 and Hungary in 2008 being the most obvious examples, with France and Croatia being the most recent ones. Similar efficiency-driven tax reforms are being analyzed in other European countries, both developed (Belgium, Netherlands) and emerging ones (Serbia, Czech Republic).

Implementing aforementioned reforms which shift the burden from income to consumption taxation is challenging in practice due to political considerations and common (mis)belief in the general public that VAT is a regressive tax that causes adverse distributional effects by creating disproportionate tax burden on the poor households. Public perception of regressive consumption taxation has been reinforced by the early empirical tax incidence analysis, including the classical work of Pechman (1985). However, more recent research has unambiguously shown that much of

[*] Fiscal Council, Republic of Serbia.

[**] University of Belgrade.

the estimated extremely regressive incidence of consumption taxes against annual income originates from measurement errors inherent in expenditure surveys. Furthermore, the theoretical basis for assessing the VAT incidence against annual income instead of annual expenditures or lifetime income is rather weak (Caspersen and Metcalf, 1994; Creedy, 1998). Recent empirical estimates in EU member states, based on the lifetime tax incidence approach, reveal slightly progressive VAT incidence (DeCoster *et al.*, 2010).

We will use micro-level data for Serbia to investigate equity aspects of value added taxation in a typical emerging European country. Compared to developed European countries, many emerging European countries, especially Poland, Romania and Serbia, feature a significant presence of own-source small farming production and associated in-kind consumption. As we will show, this feature tangibly enhances the progressivity of VAT systems in these countries. Furthermore, significant presence of shadow economy and evasion of direct income taxes in many emerging European countries suggests that household expenditures are a more meaningful indicator of the living standard and ability to pay taxes than the registered income. We conclude that common beliefs of regressive VAT taxation, often encountered in the general public, are vastly overstated and poorly founded in economic reality of emerging European countries.

This paper is organized as follows: Section 1 presents basic results from the existing literature, highlighting the difference between annual and lifetime tax incidence analysis and noting the inherent presence of income measurement errors in expenditure surveys. Section 2 describes features of the existing Serbian VAT system and explains the estimation methodology used in our analysis. Section 3 presents empirical estimates of annual and lifetime VAT incidence in Serbia. Section 4 quantifies the poor redistributive performance of the reduced VAT rate mechanism and highlights the fact that government transfer policies are the optimal tool for achieving social redistribution goals. Section 5 simulates three alternative approaches to increasing the VAT burden and compares their distributional effects. Concluding remarks are presented in Section 6.

1 Theoretical background and literature survey

Consumption taxes, and VAT in particular, are often deemed to be inherently regressive by the general public. Throughout the years, this point of view has been shared by a tangible number of economic practitioners and tax experts. The argument most often quoted in the general public in support of the regressive consumption taxation hypothesis is the observation/belief that poor individuals spend most or all of their incomes, while rich individuals are able to save significant shares of their income. Thus, one is led to believe that consumption taxation is inherently regressive – since it burdens poor individuals more heavily than it does the rich ones.

1.1 Empirical evidence

Early empirical evidence, such as the seminal study of Pechman(1985), seemed to be fully supportive of the general public beliefs. Namely, using annual income and expenditure data from a survey of US households, Pechman shows the sales tax incidence to be distinctively regressive, representing a higher share of income for poor households that for the rich ones. Similar results, based on annual data from expenditure surveys, have been obtained in many countries throughout the years. The most recent analysis with respect to EU member states is Decoster et al (2010), which also confirms the belief that VAT incidence is regressive, when measured against annual income.

When conducting empirical analysis of VAT incidence, *ex ante* one would expect to obtain results such that the effective VAT rate faced by any household lies in the range from 0 per cent to

Table 1

Effective VAT Rates in Selected EU Member States Across Annual Income Deciles

Decile	Belgium	Hungary	UK	Greece	Ireland
Poorest	26.7%	28.2%	16.1%	33.2%	46.4%
2	13.4%	20.2%	11.2%	22.1%	16.6%
3	13.0%	18.1%	10.3%	19.6%	13.6%
4	12.4%	17.1%	9.4%	18.5%	11.6%
5	12.0%	16.0%	8.8%	18.5%	12.2%
6	11.2%	15.6%	8.2%	16.7%	11.4%
7	11.0%	15.2%	8.2%	15.3%	10.3%
8	10.3%	14.7%	7.5%	15.1%	9.5%
9	10.1%	14.3%	7.1%	13.4%	8.5%
Richest	8.8%	12.5%	5.8%	11.6%	6.3%
Legal VAT rate	21%	25%	17.50%	19%	21%

Source: Rearranged from Decoster *et al.* (2010) by the authors.

the (standard) legal VAT rate. Namely, very rich households might be able to save most or virtually all of their annual income – thus facing an effective VAT rate of 0 per cent in a given year. On the other extreme, a poor household that is forced to spend all of its annual income on goods and services taxed at the standard VAT rate would face the maximum possible effective VAT rate, which equals the legally prescribed standard VAT rate in a given country. However, existing empirical studies most often present VAT burden as a percentage of (disposable) income across different deciles, ie on the tax-inclusive basis. Since VAT is legally charged on the tax-exclusive, we believe it is more meaningful to present VAT incidence results on the tax-exclusive basis. The two approaches of representing research results are completely equivalent, but representing results on the tax-exclusive basis clearly shows the effective VAT rates and allows us to easily validate our *ex ante* research expectations – that estimated effective VAT rates should lie in the range from 0 per cent to the legal VAT rate.[1]

Empirical evidence from five EU member states in Table 1 seems to confirm the entrenched belief in the general public that VAT is regressive – effective VAT rates faced by the poor households in lower income deciles are significantly higher than the effective tax rates faced by the rich households in higher income decile. However, it should be stressed that in four out of five countries analyzed – the estimated effective VAT faced by the poorest households in the lowest decile are significantly higher than the legally prescribed standard VAT in respective countries. These results are opposite to our *ex ante* research expectations! Extraordinarily high estimates of effective VAT rates in the lowest income deciles are encountered in most empirical studies based on annual income and expenditure data. For example, O'Donoghue *et al.* (2004) investigate VAT

---

[1] If VAT burden is estimated to equal x per cent of disposable household income, than effective VAT rate on the tax-exclusive basis is easily calculated as $y = x / (1–x)$.

Table 2

Saving Rates and Estimated VAT Rates Excluding the Dissaving Effect

Decile	Belgium		Hungary		Greece	
	Saving	Corrected VAT rate	Saving	Corrected VAT rate	Saving	Corrected VAT rate
Poorest	–63.4%	16.4%	–50.4%	18.8%	–117.3%	15.3%
2	–17.5%	11.4%	–14.3%	17.7%	–62.8%	13.6%
3	–8.1%	12.0%	–3.9%	17.4%	–36.3%	14.4%
4	–2.1%	12.1%	1.6%	17.1%	–3.2%	17.9%
5	3.8%	12.0%	6.4%	16.0%	–26.2%	14.6%
6	9.3%	11.2%	10.1%	15.6%	–14.3%	14.6%
7	13.3%	11.0%	12.1%	15.2%	–8.5%	14.1%
8	18.0%	10.3%	14.4%	14.7%	–5.0%	14.4%
9	22.7%	10.1%	17.6%	14.3%	1.6%	13.4%
Richest	33.3%	8.8%	27.1%	12.5%	15.8%	11.6%

Source: Rearranged from Decoster *et al.* (2010) and expanded by the authors.

incidence in twelve EU member states and their estimates of effective VAT rate in the lowest income deciles are tangibly higher than the legally prescribed VAT rates in ten out of the twelve countries being analyzed.[2]

Unrealistically high effective VAT rate estimates in most countries are caused by the extremely high dissaving rates estimated from expenditure survey data. Decoster et al (2010) describe the dissaving rates in the lowest deciles in Table 2 as "unbelievably high" and resulting from under-reporting bias of income data – a phenomenon reported by most authors dealing with data from expenditure surveys, including Sebelhaus and Groen (2000) and Meyer and Sullivan (2003). If we correct for unrealistically high dissaving rates by assuming that incomes in the lowest deciles have to equal the recorded expenditures, we can observe that estimated VAT incidence in Table 2 is still regressive, but significantly less than the estimates in Table 1 imply.[3]

Overall, we can conclude that the abundance of empirical evidence from many countries over the years, which shows extremely regressive VAT incidence over annual income deciles – only partially supports the general public's beliefs regarding the inherently regressive VAT taxation. Namely, while general public beliefs are driven by regressive incidence of saving, the empirical results are mostly driven by the regressive incidence of dissaving, stemming from measurement errors and under-reporting of income bias inherent in expenditure surveys.

---

[2] Interestingly enough, while Decoster *et al.* (2010) estimate effective VAT rate for the lowest decile to be lower than standard VAT rate in UK, O'Donoghue *et al.* (2004) estimate the opposite situation – effective VAT rate for the lowest decile was higher than standard VAT rate in this study. This points to the instability of empirical results that underlie common belief that VAT is a regressive tax.

[3] VAT estimates corrected for dissaving in Table 2 are obtained by dividing estimates in Table 1 *by (1 – the dissaving rate rate)*, for deciles which exhibiting the dissaving behavior.

1.2 Theoretical considerations

Correcting for income measurement errors significantly reduces the regressive character of VAT incidence estimates, but they remain invariably regressive when compared against annual income. However, one can rightfully question whether annual income represents the relevant measure of well-being against which VAT incidence should be assessed? In their seminal paper, Caspersen and Metcalf (1994) explain that low-annual-income households may include four very different kinds of individuals: those with volatile annual income who merely had a bad year, those that are young and just beginning a high-income career, those that are old and have just finished a high-income career, and those who are truly long-term poor. The identification of households that are truly poor requires that we look at the longer time horizon – moving from annual income framework to the entire lifetime income framework.

Permanent income theory (Friedman, 1957) suggests that annual income is not the relevant measure of well-being when assessing the VAT incidence, since households engage in consumption smoothing over their lifetime, saving temporary incomes in "good" years and dissaving accumulated funds in "bad" years. Similarly, the *lifecycle hypothesis* (Modigliani, 1986) implies that a typical individual moves from one income group to another during his lifetime, dissaving in youth and old-age and saving in the most productive middle-ages.[4] As a result, a typical individual is expected to face high VAT burden against annual income in some years, but low VAT burden in others. Overall, if we exclude the effects of inheritances and bequests, the average VAT rate an individual faces throughout his lifetime is exactly equal to the legally prescribed VAT rate – since the individual is assumed to spend all his lifetime earnings, although with certain temporal reallocations.[5] Thus, based on this theoretical consideration, one would expect consumption taxation (at uniform rate) to have proportional lifetime incidence. Furthermore, since most European VAT systems feature reduced VAT rates applicable to basic necessities, we could even expect to observe a degree of progressivity when analyzing lifetime incidence. This indeed is the case, as most research undertaken in the last couple of decades suggests somewhat progressive lifetime VAT incidence estimates, for example Caspersen and Metcalf (1994) for the United States and Decoster *et al.* (2010) for selected EU member states.

Empirical studies mostly support the lifecycle hypothesis and permanent income theory. Two approaches are most often relied upon when estimating lifetime VAT incidence – either using panel data to estimate the lifetime income (Fullerton and Rogers, 1991, Caspersen and Metcalf, 1994) or using current household consumption as a proxy for appropriate lifetime income (Caspersen and Metcalf, 1994; Decoster *et al.*, 2010; Slintakova and Klazar, 2010). Due to lack of appropriate panel data in most countries other than the United States, most of the research on VAT incidence in Europe thus far has relied on approximating the lifetime income with some form of non-durable household expenditures from expenditure surveys.[6] This is the approach we will follow in this study.

2 Data and methodological background

Serbian VAT system broadly follows the "EU model" and requirements laid-out in the

---

4 Lifetime income represents the present value of all incomes earned throughout individual's life plus any inheritance (s)he might receive.

5 Caspersen and Metcalf (1994) explain that empirical evidence suggests bequests to be U-shaped with respect to lifetime income. Thus, ignoring inheritances and bequests has the effect of overestimating progressivity for the richest households and underestimating progressivity for the majority of poor and medium-income households.

6 Current (non-durable) expenditures share many characteristics with the permanent or lifetime income, being rather stable from year to year, unlike current income which is very volatile over the years.

European Commission Sixth Directive. Standard VAT rate equals 18 per cent while the reduced rate is 8 per cent. Financial, postal, health and education services are tax-exempt, without the right to deduct the input-VAT (VAT Law Article 25). No domestic turnover of goods and services is zero-rated and basically only exports and services related to international trade and travel are zero-rated with the right to deduct the input-VAT (VAT Law Article 24). The list of goods subject to the reduced VAT rate (VAT Law Article 23) is somewhat long compared to best international practices, and includes food, medicines, utilities, textbooks, newspapers, hotel accommodation, fertilizers, firewood, natural gas, computer equipment, newly built apartments, utility services and tickets for cultural events (detailed list is given in the Appendix).

In order to estimate VAT incidence in Serbia, we will use the annual 2009 data from the Serbian Statistics Office Household Budget Survey (HBS). HBS is conducted in line with practices suggested by EuroStat, and the data from this source can be considered comparable to data obtained from expenditure surveys in other European countries. HBS 2009 data contains detailed expenditure information on 4592 representative households. HBS information is detailed enough so that particular households' expenditures can be identified as being subject to standard VAT rate, reduced VAT rate, being exempt from VAT or representing natural in-kind consumption due to own-source farming production. However, there are two limitations to be noted when using HBS data to estimate the VAT incidence:

1) Sale of food (fruit, vegetables, meat) on greenmarkets is legally exempt from VAT in Serbia. However, HBS data (and most expenditure surveys in general) does not allow differentiating between purchases of food on greenmarkets (VAT exempt) and purchases of food in grocery stores (subject to reduced VAT rate). Since poor households purchase more food from greenmarkets than rich households, using HBS data as the basis of VAT incidence overstates the actual VAT burden for the poor households.[7]
2) HBS surveys in general cover household consumption and don't include purchases of newly built apartments, which formally represent investment spending. However, since purchases of newly built apartments are subject to VAT, HBS data understates VAT incidence of rich households – which save for many years in order to afford one-time high cost of buying an apartment.

Aforementioned limitations of the HBS data cause a systematic regressivity-bias in VAT incidence estimates, by overstating the actual regressivity or understating the actual progressivity. These limitations should be kept in mind when discussing empirical VAT incidence estimates.

In what follows, we will assume full forward-shifting of VAT to consumer prices and will use HBS data to conduct a micro-simulation static analysis of VAT incidence (which assumes no behavioral responses). Although rather restrictive, these assumptions and this modeling framework are most often utilized when assessing incidence of consumption taxes in practice.[8]

We will analyze VAT incidence of the existing system and perspective reforms based on the estimated average effective VAT rate by income and expenditure groups (deciles). Average effective VAT rate in income group i is calculated as the ratio of total estimated VAT burden and total income for the income group i. In particular, average effective VAT rate in income group i (i = 1, 2 … 10) will be calculated as the weighted average of tax rates t_j (j = 1, 2, 3, 4) which

7 Jenkins *et al.* (2006) argue that "in developing countries the commodities on which poor households spend most of their income, even if they are included in the legal tax base, are administratively impractical to tax." Based on the detailed information on expenditure patterns and types of establishments from which items are purchased, they estimate effective (annual) VAT burden in the Dominican Republic to be progressive. Similar detailed information is not available for Serbia, but one should keep in mind this source of tangible regressivity-bias in our estimates.

8 Warren (2008) states some of the drawbacks of this approach and recommends use of input-output tables as the preferred approach to modeling incidence of consumption taxes. However, he notes that only a handful of most developed OECD countries are currently able to provide comprehensive information required to properly calibrate this type of models.

correspond to four different legal VAT rates present in the Serbian system – standard rate, reduced rate, exempt services and natural in-kind consumption due to own-source farming production. Weighting will be done according to the structure of consumption in each income group. Thus, we will use the formula:

$$\bar{t}_i = \sum_{j=1}^{4} t_j C_j \qquad (1)$$

where t_j stands for four different tax rates possible under the Serbian VAT system and C_j represents the share of consumption in income group i subject to the tax rate t_j. Alternatively, average effective tax rates will also be estimated for different groups according to the expenditure ranking.

Average effective VAT rates by income and expenditure groups informally indicate whether the system is progressive or regressive. However, we will compliment these statistics with formal global progressiveness indices. Gini coefficient is the most often quoted index with respect to the (in)equality of income and expenditure distributions:

$$G = \frac{\sum_{i=1}^{n}\sum_{r=1}^{n} |y_i - y_r|}{2n^2 \bar{y}} \qquad (2)$$

where $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$, n is number of individual households in the sample, y_i is the income of the individual household i, and y_r is the income of individual household r. We will be measuring the difference between inequality of income distribution before and after taxation by the difference of respective Gini coefficients (also known as the Reynolds-Smolensky index).

The Gini coefficient is focused on the middle portion of the distribution, mostly ignoring the developments in the tails of the distribution. Thus, it is usefully to complement the Gini statistics with General Entropy statistics that can assign higher weight to any particular portion of a distribution (Litcfield, 1999):

$$GE(\alpha) = \frac{1}{\alpha^2 - \alpha}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i}{\bar{y}}\right)^{\alpha} - 1\right] \qquad (3)$$

We will calculate General Entropy index with the alpha parameter 0.2 to investigate the behavior in the lower distribution tail and General Entropy index with alpha parameter 2.0 to investigate developments in the upper tail of income and expenditure distributions.[9] We will be investigating how Gini and General Entropy statistics change after imposing the VAT burden on before-tax distributions. Lower values of these statistics associated with after-tax distributions indicate progressive VAT incidence, while higher values of Gini and General Entropy statistics indicate regressive VAT incidence.

3 Estimates of VAT incidence in Serbia

This section estimates annual and lifetime VAT incidence in Serbia and compares the results with the existing literature on VAT incidence in other countries. Due to the lack of any relevant

[9] Both Gini and General Entropy statistics with a value of 0 indicate perfectly equal distribution of income, while increasing values of these statistics indicate increasingly unequal income distribution.

Table 3

Cross-referencing Annual and Lifetime Income Rankings
(percent)

		Deciles by Expenditure Ranking									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Deciles by Income Ranking	1	**41.8**	16.1	10.5	8.9	7.4	4.6	2.2	3.5	3.1	2.0
	2	23.1	**21.8**	13.7	11.1	8.9	7.4	5.9	3.7	2.8	1.5
	3	12.6	16.6	**18.3**	14.6	9.4	7.8	7.6	5.7	4.4	3.1
	4	10.5	16.6	13.5	**15.3**	10.5	6.5	10.5	7.8	4.6	4.4
	5	5.0	12.2	15.3	11.3	**12.9**	14.4	10.0	9.8	6.3	2.8
	6	2.6	5.4	10.9	15.0	15.5	**14.4**	12.0	10.5	8.9	4.8
	7	2.0	5.0	5.9	9.4	15.0	13.5	**14.4**	13.7	11.8	9.4
	8	0.7	2.8	5.4	7.6	11.5	15.3	14.6	**15.5**	15.7	10.9
	9	0.7	2.8	5.0	4.1	6.3	8.9	13.3	17.0	**21.4**	20.5
	10	1.1	0.7	1.5	2.6	2.6	7.2	9.6	12.9	21.1	**41.2**

panel data on household income and consumption behavior over time, VAT incidence analysis for Serbia will be based on annual income and expenditure data from the HBS. Annual income data will be used as the relevant measure of living standard in assessing the annual VAT incidence. We will use annual expenditure data as a proxy for permanent income in assessing the lifetime VAT incidence. Since Serbian HBS basically ignores household investment expenditures, we will use total recorded expenditures as a proxy for current (non-durable) household expenditures.[10] Annual income and expenditure rankings of households are cross-referenced in Table 3 in order to investigate how different are the results from these two alternative indicators of the living standard and ability-to-pay taxes.

We can observe that only about 42 per cent of households in the lowest decile by income ranking are also classified in the lowest decile according to expenditure ranking. On average, 21.7 per cent of households (along the main diagonal) are given the same ranking according to income and expenditure criteria. Information in Table 3 closely resembles original calculations by Caspersen and Metcalf (1994), although the data for Serbia in 2009 seems to be more dispersed than the original data for United States in 1988. Both cross-reference tables imply a significant difference between income and expenditure rankings, ie between annual and lifetime VAT incidence.

We have used HBS data to divide total expenditures of each representative household into four categories – expenditures subject to the standard VAT rate, those subject to reduced VAT rate, expenditures that are VAT exempt and expenditures due to natural in-kind consumption of

[10] In order to check the robustness of our results, we have also performed simulations which exclude vehicle purchases from total expenditures, in line with the Caspersen and Metcalf (1994) modeling approach. The results obtained were virtually identical and thus we have opted for the simplest approach of using total household expenditures – which facilities simple cross-country comparison of our results.

own-source farming production. Households were then sorted according to the two alternative indicators of the living standard: registered income and registered expenditures. In both cases, OECD equivalence scale was used to account for different sizes of households.[11] Summary results, by income and expenditure deciles, are presented in Tables 4a and 4b.

We can observe that existing VAT system in Serbia creates a tangibly progressive structure of the tax burden, especially when we consider expenditure ranking of households. Namely, reduced rate goods represent a higher share of expenditures in case of poor households than for the rich households. The progressive structure is somewhat diminished by regressive incidence of VAT exempt services. Besides the well known issue of exempting financial services, Serbian VAT system (like most other European systems) exempts health and education services. Due to the existence of public health and education system, mostly rich households can afford additional expenditures on these services – thus yielding a distinctively regressive incidence.[12]

It is important to note that natural in-kind consumption due to own-source farming production significantly increases the effective progressive layout of the Serbian VAT system by providing a tangible VAT-exempt buffer to poor households. This is a distinctive characteristic of VAT incidence in Serbia compared to other developed European countries where agricultural production and small scale own-source farming is only marginally present. Namely, agricultural production accounts for only 1.3 per cent of GDP in the European Union, while it accounts for 13 per cent of GDP in Serbia. Furthermore, agricultural employment accounts for 20 per cent of total employment in Serbia – compared to about 5 per cent in the European Union. As mentioned, own-source small farming production is also tangibly present in other emerging European countries, foremost Poland and Romania, where agricultural employment also accounts for 20 per cent of the overall employment.

In order to estimate average effective VAT burden for each household, we apply appropriate tax rates to each expenditure category. Effective 0 per cent VAT rate is applied to VAT exempt and natural in-kind consumption categories.[13] The results of annual and lifetime VAT incidence estimates are given in Tables 5a and 5b.

Despite broadly progressive structure of the Serbian VAT system in Tables 4a and 4b, we can see that annual VAT incidence in Table 5a is distinctively regressive, especially in the lower tail of income distribution. As in other countries, this is caused by the distinctively regressive estimated incidence of household (dis)saving. The estimated effective VAT rate in the lowest decile of 21.2 per cent is tangibly higher than the legally prescribed VAT rate of 18 per cent, indicating that under-reporting of income bias is present in the Serbian HBS data, as is the case in virtually all expenditure surveys analyzed in the existing literature.[14] After correcting for unrealistically high dissaving rates in the lowest deciles, which we know not to be realistic nor sustainable in economic reality, we can observe that regressive character of annual VAT incidence is significantly reduced.

---

[11] Since Serbia is significantly less developed than OECD countries, using OECD equivalence scale might not be a most suitable choice. For a detailed discussion on this topic, and alternative approaches to measuring income inequality in Serbia, see Jovičić and Milojević (2010).

[12] Possible legal improvements in this area are limited by the EU Sixth Directive which prescribes that VAT cannot be charged on publicly provided health and education services. However, taxing private provision of these services is not forbidden, which could, if introduced, eliminate a significant portion of these regressive effects.

[13] Although most VAT incidence studies assume 0 per cent rate for VAT exempt services, this is not a completely appropriate assumption since VAT exempt entities do bear a certain tax burden due to inability to reclaim VAT on business inputs. For this reason, a sensitivity analysis has been conducted by assuming different non-zero tax rates on VAT exempt services in Serbia. Main conclusions of the study remained valid even after assuming non-zero rate on exempt services.

[14] In fact, if we also recognize that 17.8 per cent of income in the first decile represents implicit in-kind income which can not possibly be saved, since it corresponds to in-kind consumption of own-source small farming production, we can conclude that effective VAT rate on monetary income equals 27 per cent in the first decile.

Table 4a

Expenditure Patterns by Annual Income Deciles
(percent)

Decile	Standard VAT Rate	Reduced VAT Rate	VAT Exempt	Natural Consumption
Poorest	46.4	40.8	2.4	10.3
2	48.7	39.0	2.6	9.7
3	49.5	40.3	2.1	8.1
4	50.0	38.9	2.1	9.0
5	51.4	39.5	2.0	7.1
6	51.7	39.8	2.0	6.5
7	54.5	39.1	2.2	4.3
8	52.9	39.7	3.5	4.0
9	55.5	36.9	3.8	3.9
Richest	58.1	34.6	3.4	3.9

Table 4b

Expenditure Patterns by Annual Expenditure Deciles

Decile	Standard VAT Rate	Reduced VAT Rate	VAT Exempt	Natural Consumption
Poorest	43.5	42.5	0.6	13.4
2	48.6	38.5	1.0	11.8
3	48.2	39.6	1.0	11.2
4	49.3	41.6	1.4	7.8
5	49.1	41.6	1.4	7.9
6	51.5	40.4	1.9	6.2
7	51.7	40.5	2.6	5.2
8	52.8	38.8	3.1	5.3
9	54.7	38.6	3.0	3.7
Richest	59.5	32.5	5.4	2.5

Table 5a

Annual VAT Incidence and VAT Incidence Excluding the Dissaving Effects

Decile	Effective VAT Rate	Saving	Corrected VAT Rate
Poorest	21.2%	–73.1%	12.3%
2	14.9%	–25.5%	11.9%
3	13.9%	–15.8%	12.0%
4	12.3%	–4.0%	11.8%
5	11.2%	6.7%	11.2%
6	11.0%	8.6%	11.0%
7	11.2%	9.9%	11.2%
8	10.0%	17.3%	10.0%
9	9.5%	22.4%	9.5%
Richest	7.9%	36.0%	7.9%
Global Progressiveness Indices			
	Before Tax	**After Tax**	**Difference**
Gini Coeff.	0.299338368	0.312989058	0.045602873
GE(0.2)	0.154817719	0.205870445	0.329760219
GE(2.0)	0.182480671	0.212404149	0.163981633

Table 5b

Lifetime VAT Incidence Statistics

Decile	Effective VAT Rate		
Poorest	10.8%		
2	11.5%		
3	11.5%		
4	11.8%		
5	11.8%		
6	12.2%		
7	12.2%		
8	12.3%		
9	12.6%		
Richest	13.0%		
Global Progressiveness Indices			
	Before Tax	**After Tax**	**Difference**
Gini Coeff.	0.27899635	0.27636354	–0.00263281
GE(0.2)	0.12826570	0.12590617	–0.00235953
GE(2.0)	0.16318280	0.16008532	–0.00309748

Information in Table 5b indicates a slightly progressive lifetime VAT incidence.[15] Global progressiveness indices indicate that lifetime VAT effects seem to be more progressive in the tails of the distribution than in the middle portion. Our estimate of lifetime VAT incidence in Serbia is in line with other empirical studies that indicate slightly progressive lifetime VAT incidence in other European countries (Decoster *et al.*, 2010). We can also notice that within the lifetime framework, estimated effective VAT rates are in the 10 to 13 per cent range, which is in line with our *ex-ante* expectation of the results that a VAT incidence study should yield.

We can notice from table 5b that the Gini coefficient associated with after-tax distribution of expenditures is lower than the Gini coefficient associated with the before-tax distribution of expenditures, which confirms progressive lifetime VAT incidence. It is instructive to decompose this total reduction in the Gini coefficient of 0.26328 per cent into two components – the effect of the structure of the VAT system itself and the effect due to significant presence of small scale own-source farming production in the Serbian economy. Our estimates show that the structure of the Serbian VAT system, which features taxation of necessities under the reduced VAT rate, is responsible for reducing the Gini coefficient by 0.11935 per cent. On the other hand, progressive incidence of own-source farming production reduces the Gini coefficient by 0.14375 per cent. Thus, we can conclude that more significant redistribution effects are achieved due to the presence of own-source farming production than due to the introduction of reduced VAT rate for taxing necessities.[16]

In order to reach the definite answer whether effective VAT incidence is regressive or mildly progressive, one needs to decide what is the most reliable indicator of the standard of living in Serbia – is it household annual income, or household expenditures? It is our belief that one should opt for household expenditures, both on theoretical and practical grounds. From theoretical point of view, it is widely recognized that household engage in consumption smoothing over the lifetime, implying that the lifetime VAT incidence, which can be decently approximated using annual household expenditures, represents a more suitable framework than the annual tax incidence analysis. Furthermore, Creedy (1998) explains that when analyzing distributional effects of consumption taxes in isolation, ignoring the remaining tax and benefit system – one should not mix apples and oranges and assess the incidence of consumption taxes against household incomes, but against their prescribed tax base – the household expenditures.

From practical point of view, it should be stressed that the transition process in emerging European countries has been accompanied with a large surge in shadow economy and tax evasion. In particular, Schneider(2005) approximates that shadow economy averages about 16.3 per cent of GDP in developed OECD countries, compared to 40.1 per cent of GDP in transitional European countries. He estimates the shadow economy in Serbia to be approximately 39 per cent of GDP in 2003.[17] Tax evasion is most pronounced in the case of direct income taxation, since undeveloped judiciary and tax administration capacities in many emerging European economies limit the possibilities for effective identification and effective prosecution of income tax evasion cases. On the other hand, the VAT system is the best available tool for combating tax evasion – evasion has to be organized throughout the entire production and distribution chain, since identification of unregistered invoices at any stage of production implies the collection of VAT corresponding to the

---

15 Lifetime VAT incidence estimates based on expenditure data in Table 5b basically eliminate the effects of dissaving in lower deciles and saving in upper deciles. Corrected annual VAT rates presented in Table 5a had been obtained by basically correcting only for the dissaving effect in the lowest deciles. Remembering regressivity-bias from Section 2, we can conclude that actual lifetime progressivity of VAT incidence is somewhat higher than observed in Table 5b.

16 It should be noticed that progressive effects of the reduced VAT rate are diminished by the regressive incidence of VAT-exempt services.

17 Replacing sales tax with VAT in 2005 reduced the extent of tax evasions and shadow economy in Serbia. However, tax compliance in Serbia seems to have deteriorated as the result of 2008-09 economic crisis.

entire value added since the beginning of the production process.[18] Thus, we believe that in Serbia, and many other emerging European countries, household expenditures represent a more meaningful and more reliable indicator of the standard of living and ability-to-pay taxes.[19] In the remainder of this paper, we will focus on analyzing the (lifetime) VAT incidence against household expenditures.

4 Targeting of reduced rate VAT subsidies

Due to social considerations, VAT systems in basically all European countries feature reduced rates for certain basic necessity goods. The idea behind reduced rates is to try to introduce redistributive social elements into the VAT structure. By subjecting basic necessities, such as food or medicine, to a reduced tax rate – the VAT system basically subsidizes the consumption of these goods by the difference between the standard and the reduced VAT rate. Since these necessities represent a higher share of expenditures for poor households than for rich hoseholds, it is hoped that poor households would capture most of the economic benefit associated with the consumption of goods under the reduced VAT rate. Implicit tax subsidies associated with the reduced VAT rate totaled about one quarter of total VAT revenues actually collected.

Slightly progressive lifetime VAT incidence estimated by Decoster et al (2010) for EU member states is driven by the consumption of goods under the reduced VAT rate. As discussed, progressive VAT incidence in Serbia is further reinforced by the significant presence progressive own-source farming production and in-kind consumption. However, it is important to acknowledge that poor households spend more on necessities in relative terms (relative to their total income or total expenditure), but not in absolute terms. For example, we can see from the Appendix that consumption of bread, milk or medicine products is distinctively progressive across expenditure deciles. Milk and dairy products account for 7.3 per cent of monetary expenditures in the lowest decile and 3.5 per cent of expenditures in the highest decile. However, in absolute terms, monetary expenditures for milk and dairy products are about three times larger in the highest decile than in the lowest decile. Thus, households in the highest expenditure decile are receiving a tax subsidy that is in absolute terms three times larger than the amount of subsidy going to the lowest expenditure decile. This difference is even more pronounced in the case of reduced-rate goods whose consumption pattern is not very progressive, such as meat, hotel accommodation or utility services, see Table 6.

We may ask whether the reduced VAT rate is the most suitable approach for achieving redistribution objectives? It seems that it might be more efficient to have a uniform-rate VAT system and to use additional revenues (from eliminating reduced VAT rate subsidies) to fund government programs that are better targeted at poor households. Ebril *et al.* (2001) stresses this point of poor redistributive performance with reduced VAT rates and highlights that best practice strongly suggests that VAT systems should have a single uniform tax rate – leaving the redistribution role to other segments of the tax and benefit system.[20] However, this is the "first-best" result which assumes government transfer policies are properly targeted and appropriately funded. Unfortunately, this is not the case in many emerging European countries. Thus, in the case of "second-best" result, Bird and Gendron (2007) argue that reduced VAT rate on

---

18 Jimenez *et al.* (2010) based on empirical data from Latin America economies show that tax evasion in emerging economies is much higher in the area of direct income taxes than with indirect consumption taxes.

19 In fact, due to concerns regarding quality of income measurements from sample surveys, poverty levels in developing countries, such as Serbia, are mostly assessed based on expenditure data.

20 Atkinson and Stiglitz (1976) show that if utility function is weakly separable in leisure and consumption, preferences for goods do not depend on ability and progressive labor income tax is available, then differential commodity tax cannot improve social welfare, *i.e.*, uniform taxation of final goods is optimal in this case.

Table 6

Annual VAT Subsidies per (Equivalence Scales) Household Across Expenditure Deciles
(dinars)

Deciles	Bread and Baked Products	Milk and Dairy Products	Meat and Fish	Medicines	Utility Services
Poorest	1,237	1,125	1,547	765	325
2	1,375	1,538	2,292	792	641
3	1,355	1,789	2,933	906	832
4	1,496	2,010	3,546	995	1,156
5	1,634	2,264	4,259	1,082	1,188
6	1,637	2,410	4,645	1,272	1,462
7	1,738	2,553	5,290	1,447	1,631
8	1,787	2,736	5,883	1,506	1,759
9	1,802	3,117	7,299	1,835	2,114
Richest	2,097	3,650	9,614	2,350	2,501

certain most basic necessities might serve a useful redistributive purpose in developing countries where welfare transfer programs and progressive income taxation are not well developed.

5 Distributional effects of prospective increase of the VAT burden

Increasing the VAT burden in Serbia could create additional budget revenues that could finance a growth-enhancing (revenue-neutral) tax reform aimed at reducing the tax wedge on labor (Arsić et al., 2010). However, alternative approaches to increasing VAT burden will have different distributional effects. In this section we will use lifetime VAT incidence framework to analyze distributional effects of three alternative approaches that yield same budget revenues.

Existing VAT system in Serbia is structured so that about 40 per cent of taxable consumption is subject to the reduced VAT rate of 8 per cent while 60 per cent of taxable consumption is subject to the standard VAT rate of 18 per cent. This implies that average effective VAT rate in Serbia currently stands at 0.4 * 8% + 0.6 * 18% = 14%.[21] We will analyze three alternative scenarios that all increase average effective VAT rate to 18 per cent. In revenue terms, this should yield additional tax revenues of about 2.5 per cent of GDP per year.

- Scenario 1: Transferring all goods from the current reduced rate of 8 per cent to the standard VAT rate of 18 per cent and basically introducing a uniform-rate VAT system.

[21] 14 per cent average VAT rate is not with respect to total household expenditures, but with respect to the taxable portion of expenditures, ie excluding tax-exempt services and in-kind consumption.

Table 7

Effective VAT Rates Relative Changes Under Alternative Reform Scenarios

Decile	Current System	Relative Changes to Current System		
		Scenario 1	Scenario 2	Scenario 3
Poorest	10.8%	4.6%	3.5%	2.9%
2	11.5%	4.1%	3.6%	3.1%
3	11.5%	4.3%	3.6%	3.3%
4	11.8%	4.5%	3.7%	3.6%
5	11.8%	4.5%	3.7%	3.6%
6	12.2%	4.4%	3.8%	3.6%
7	12.2%	4.4%	3.8%	3.7%
8	12.3%	4.2%	3.8%	3.7%
9	12.6%	4.2%	3.8%	3.8%
Richest	13.0%	3.5%	3.8%	3.6%
Gini Coeff.	0.27636	0.00162	–0.00029	–0.00079
GE(0.2)	0.12591	0.00160	–0.00022	–0.00067
GE(2.0)	0.16009	0.00362	0.00005	0.00004

- Scenario 2: Maintaining the existing VAT structure and increasing the tax rate by 4 per cent – reduced rate from 8 to 12 per cent and standard rate from 18 to 22 per cent.
- Scenario 3: Increasing the VAT rate by 2 per cent (reduced rate to 10 per cent and standard rate to 20 per cent) and transferring certain goods from the reduced rate to the standard rate, so that about 20 per cent of taxable consumption remains subject to the reduced rate, while 80 per cent of taxable consumption becomes subject to the standard rate.[22]

Table 7 presents information on the absolute incidence of the existing VAT system and relative incidence increments for each scenario. Looking at the VAT incidence according to expenditure deciles, we can notice that Scenario 1 causes slightly regressive effects, while Scenarios 2 and 3 cause slightly progressive effects. This conclusion is supported by Gini and General Entropy statistics, which indicate higher income inequality in Scenario 1 and lower income inequality in Scenarios 2 and 3.

[22] There are many different ways in which Scenario 3 could be designed, depending on which goods are chosen to be transferred from the reduced-rate to standard-rate status. In reality, this process would be driven by social and political preferences, subject to the constraint that only 20 per cent of taxable consumption should be left at the reduced VAT rate. For the purpose of this study, in order to achieve best distributional effects, we have decided to transfer to the standard VAT rate those goods with the least progressive (or even regressive in some cases) consumption patterns. Thus, we have implemented Scenario 3 by transferring fruit, meat, fish, computer equipment, hotel accommodation, firewood, natural gas, utility services and tickets for cultural events from the reduced-rate to standard-rate status.

Ebril *et al.* (2001) recommendation of a single uniform-rate VAT system might not be most suitable for the current Serbian environment characterized with low coverage of major welfare transfer programs (such as material family support or child allowance) and basically proportional system of income taxes. Thus, authors believe that Scenario 1 should be accompanied with progressivity-enhancing reforms in other segments of the tax and benefit system – such as expanding welfare transfer programs or increasing the progressivity of income taxation. Scenarios 2 and 3 do not cause regressive effects and could be implemented as standalone measures. Best distributional effects are achieved with Scenario 3. However, implementing this scenario in practice might be politically challenging since it includes significant tax increase on such basic items as meat or fruit.

6 Concluding remarks

The belief that consumption taxes, and VAT in particular, are inherently regressive is entrenched with a significant number of individuals in the general public. This belief, which seemed to had been supported with strong empirical evidence, presents a significant political challenge to implementing growth-enhancing tax reforms which shift tax burden from income to consumption. However, in their classical public finance textbook, Rosen and Gayer (2007) state that the final verdict on the incidence of consumption taxes and VAT is still undecided, despite seemingly strong empirical evidence from annual incidence studies.

We have shown that existing results from annual incidence studies only partially confirm the common belief regarding consumption taxation, since most of the estimated regressive VAT incidence stems from measurement errors and not from regressive savings incidence. From theoretical point of view, the annual VAT incidence approach suffers from many drawbacks and it is thus more meaningful to analyze lifetime tax incidence. When analyzing the VAT incidence in isolation, disregarding the other components of the tax and benefit system, it is particularly inappropriate to mix apples and oranges and compare VAT incidence against annual income and not against household expenditures – which is the prescribed tax-base for VAT assessment (Creedy, 1998).

Overall, it is authors' conclusion that claims regarding inequitable and regressive VAT taxation are vastly overstated and poorly founded in theoretical and empirical evidence. Similarly to the demise of common acceptance of the simple Keynesian consumption function few decades ago, the authors believe that contemporary evidence points to the demise of common beliefs regarding regressive consumption taxation. The case for regressive VAT claims is particularly weak in emerging European economies, due to large scale evasion of direct income taxes and significant presence of own-source farming production which enhances the progressive layout of the VAT burden in these countries.

APPENDIX

Following goods and services are currently subject to the reduced VAT rate of 8 per cent: food (bread and other baked products, milk and other dairy products, flour, sugar, eggs, edible oils and fats, honey, fruit, vegetables, meat, fish), medicines, fertilizers, textbooks, newspapers, computer equipment, hotel and motel accommodation, natural gas, firewood, utility services (including water), tickets for cultural events and newly built apartments.

This Appendix shows VAT incidence for reduced-rate goods, except for newly built apartments whose sales are not recorder in the Serbian HBS. VAT incidence is given across expenditure deciles, as a percentage of monetary expenditures. Data on natural consumption of own-source farming production have been purposely excluded, to highlight the fact that natural consumption of food would not be affected if certain food items are transferred from the reduced-rate to the standard-rate status.

Table 8

Incidence of Reduced Rate Goods

(percent of monetary expenditures, by expenditure deciles)

Deciles	Bread and Baked Products	Milk and Dairy Products	Flower, Sugar, Eggs, Oil, Honey	Fruit	Vegetables	Meat & Fish	Medicines
1	7.9	7.3	5.4	1.8	3.9	10.1	4.8
2	6.0	6.8	4.6	2.0	3.6	10.2	3.4
3	5.0	6.7	4.4	2.1	3.5	11.0	3.2
4	4.6	6.3	4.4	2.2	3.3	11.1	3.0
5	4.4	6.2	4.0	2.0	3.2	11.8	2.9
6	3.9	5.8	3.7	2.0	3.1	11.3	2.9
7	3.6	5.4	3.4	1.9	3.1	11.3	2.9
8	3.2	5.0	3.4	2.0	3.1	10.9	2.7
9	2.7	4.7	3.0	1.9	2.7	11.0	2.6
10	2.0	3.5	2.0	1.6	2.2	9.3	2.2

Deciles	Textbooks and Newspapers	Utility Services	Fertilizers	Natural Gas and Firewood	Hotel Accommodation	Tickets for Cultural Events	Computer Equipment
1	0.8	2.0	3.2	1.5	0.3	0.0	0.2
2	1.0	2.7	1.1	1.7	0.4	0.0	0.2
3	1.1	2.9	1.5	2.7	0.5	0.0	0.1
4	1.0	3.3	1.6	3.4	0.6	0.0	0.2
5	1.3	3.0	2.0	3.3	0.7	0.0	0.2
6	1.2	3.3	1.7	2.8	1.1	0.0	0.2
7	1.1	3.2	1.6	4.0	0.7	0.0	0.3
8	1.2	3.0	1.1	4.3	0.8	0.1	0.2
9	1.0	2.9	1.3	3.6	2.1	0.1	0.4
10	1.0	2.2	0.7	3.0	3.0	0.3	0.4

REFERENCES

Atkinson, A. and J. Stiglitz (1976), "The Design of Tax Structure: Direct versus Indirect Taxation", *Journal of Public Economics*, No. 6.

Arsić, M., N. Altiparmakov, S. Ranđelović, A. Bućić, D. Vasiljević and T. Levitas (2010), *Tax Policy in Serbia – Looking Forward*, Foundation for Advancement of Economics and USAID, Serbia.

Bird, R. and P.P. Gendron (2007), *The VAT in Developing and Transitional Countries*, Cambridge: Cambridge University Press.

Caspersen, E. and G. Metcalf (1994), "Is Value Added Tax Regressive? Annual Versus Lifetime Incidence Measures", *National Tax Journal*, Vol. 47, No. 4.

Creedy, J. (1998), "Are Consumption Taxes Regressive?", *The Australian Economic Review*, Vol. 31, No 2.

Decoster, A., J. Loughrey, C. O'Donoghue and D. Verwerft (2010), "How Regressive are Indirect Taxes? A Microsimulation Analysis of Five European Countries", *Journal of Policy Analysis and Management*, Vol. 29, No. 2.

Ebrill, L., M. Keen, J.P. Bodin and V. Summers (2001), *The Modern VAT*, International Monetary Fund, Washington (D.C.).

European Commission, Directorate-General for Economic and Financial Affairs (2008), *Public Finances in EMU – 2008*.

European Commission, Taxation and Customs Union (2009), "Taxation trends in the European Union".

Friedman, M. (1957), *A Theory of the Consumption Function*, Princeton University Press, Princeton (NJ).

Fullerton, D. and D. Rogers (1991), "Lifetime Versus Annual Perspectives on Tax Incidence", *National Tax Journal*, Vol. 44, No 3.

Jenkins, G., H. Jenkins and C.Y. Kuo (2006), "Is the Value Added Tax Naturally Progressive?", Queen's University, Working Paper, No. 1059, Kingston (ON, Canada).

Jiménez, J.P., J.C.G. Gómez Sabaini and A. Podestà (2010), "Tax Gap and Equity in Latin America and Carribean", *Fiscal Studies*, No. 16.

Johansson, A., C. Heady, J. Arnold, B. Brys and L. Vartia (2008), "Taxation and Economic Growth", OECD, Economics Department, Working Paper, No. 620.

Jovičić, M. and T. Milojević (2010), "Poverty and Inequality changes in Serbia as the Result of Global Instability", 11th bi-Annual conference of the EACES *Comparing Responses to Global Instability*, Tartu, Estonia, ISBN 978-9985-4-0613-7.

Kasselman, J. and R. Cheung (2004), "*Tax Incidence, Progressivity, and Inequality in Canada*", Simon Fraser University at Harbour Centre, Centre for Public Policy Research.

Keen, M. and S. Smith (2006), "VAT Fraud and Evasion: What Do We Know and What Can Be Done?", *National Tax Journal*, Vol. 59, No. 4.

Litchfield, J. (1999), "Inequality: Method and Tools", available on The World Bank website at: www.worldbank.org/poverty/inequal/index.htm

Meyer, B. and J. Sullivan (2003), "Measuring the Well-Being of the Poor Using Income and Consumption", *Journal of Human Resources*, No. 38.

Modigliani, F. (1986), "Life Cycle, Individual Thrift, and the Wealth of Nations", *American Economic Review*, No. 76.

Metcalf, G. (1994), "Lifecycle vs. Annual Perspectives on the Incidence of a Value Addes Tax", NBER, Working Paper, No. 4619.

Norregaard, J. and T. Khan (2007), "Tax Policy: Recent Trends and Coming Challenges", IMF, Working Paper, WP/07/274.

O'Donoghue, C., M. Baldini and D. Mantovani (2004), "Modeling the Redistributive Impact of Indirect Taxes in Europe: An Application of EUROMOD", EUROMOD, Working Paper, EM7/01, University of Cambridge (UK).

Pechman, J. (1985), *Who Bears the Tax Burden?*, Brookings Institution, Washington (D.C.).

Poterba, J. (1989), "Lifetime Incidence and the Distributional Burden of Excise Taxes", *American Economic Review*, Vol. 79.

Rosen, H. and T. Geyer (2007), *Public Finance*, eight edition, McGraw-Hill Publishing.

Sabelhaus, J. and J. Groen (2000), "Can Permanent Income Theory Explain Cross-sectional Consumption Patterns?", *Review of Economics and Statistics*, Vol. 82.

Slintáková, B. and S. Klazar (2010), "Impact of Harmonisation on Distribution of VAT in the Czech Republic", Prague Economic Papers, Vol. 19, No 2.

Warren, N. (2008), "A Review of Studies on the Distributional Impact of Consumption Taxes in OECD Countries", OECD Social, Employment and Migration, Working Paper, No. 64.

Value Added Tax Law, Official Gazette of Republika Srbija, No. 84/04, 86/04, 61/05 and 61/07.

COMMENTS ON SESSION 4
POLICIES TO PROMOTE SUSTAINABLE GROWTH

DOES FISCAL POLICY MATTER?

Werner Ebert and Sarah Ciaglia**

In the context of the current EMU debate on austerity and stimulus, the papers by Bouthevillain and Dufrénot and Gavilán *et al.* address important questions. As fiscal policy is the only policy area which instruments affect growth very differently the question on how heterogeneous growth patterns in the euro area can be shaped by fiscal policy measures compared to structural reforms is topical.[1] Coming from a finance ministry, our perspective is necessarily more practical than academic. Hence, we focus on the question: Does fiscal policy matter? That includes a discussion of the possible use of the results of these papers for the discussion on shaping the institutional context of the EU and the euro area.

1 Common motivation: need for disaggregation

When addressing the "fiscal policy and growth" issue,[2] historical experience with fiscal policy measures shows that a "one-size-fits-all" approach does not work well, particularly in a common currency area. Although aggregate models undisputedly have their merits, concerning these policy issues it is wise to disaggregate and to be more country-specific in order to derive practical policy conclusions. Therefore, both papers follow a quite sensible approach of explicitly taking heterogeneity into account: Bouthevillain and Dufrénot do so by disaggregating public expenditures and revenues and by selecting different growth periods, Gavilán *et al.* by following a country-specific long-term approach including open economy and external imbalances variables. The first paper concentrates on fiscal policy and growth while the second one focuses on structural policies with a specific view on macroeconomic imbalances and growth.

2 Models and main findings

Bouthevillain and Dufrénot raise the following question: does a common fiscal policy (taxation and expenditure measures) become growth enhancing or reducing in a similar way across countries? They run a double quantile fixed effects regression on the effects of fiscal variables on growth. Using the period between 2000 and 2010, they look at real and per capita GDP as that allows differentiating between fiscal policies' effects that are different by country and time. For the analysis of growth effects of social expenditure vs. "economic" spending, taxation vs. social security contributions and direct vs. indirect taxation, they make use of COFOG data. Concerning practical economic policy, their basic assumption is that the "recipes" for generating growth by

* Contact: Werner.Ebert@bmf.bund.de and Sarah.Ciaglia@bmf.bund.de
The authors work at the German Federal Ministry of Finance, Berlin.
The views expressed here represent their personal opinion, and not necessarily the view of the Ministry.

1 One could refer to the recent research on growth in the EMU by the EU COM.

2 Refer to the work by the EU COM (Pench *et al.*), the EU Economic Policy Committee with its Working Group on the Quality of Public Finances, the OECD (Heady on tax issues) and also by the ECB (Afonso and his team on efficiency and effectiveness of public spending). We are delighted that new literature on this topic is being provided at this conference in Perugia (e.g., by WIFO Austria, Afonso).

fiscal policies are very different in high and low growth countries. In their specific model they come to mixed and partly counterintuitive results:

1) welfare expenditure can foster growth in general while human capital expenditure can foster growth only in low-growth countries and can be even harmful for growth in high-growth countries;
2) the effect of a social VAT (replacing social security contributions) on growth is mixed, in low-growth countries positive, in high-growth countries neutral;
3) replacing direct by indirect taxation has a significant impact in high growth countries, not so in low-growth countries.

Gavilán *et al.* use an overlapping generations model of a small open economy characterized by imperfect competition. They focus on three periods, one between the mid-late nineties until the beginning of the crisis in Europe in 2008, the crisis years and a simulated post-crisis period. The basic question they try to answer is: How did external imbalances (in Spain and the euro area) evolve? As the main drivers of the performance of growth and external imbalances they identify demography and immigration causing changes in the work age population and the interest rate channel (permanent nominal convergence). The counterfactual question is what potential role fiscal policy could play to avoid imbalances and to foster growth. They conclude that a continuous tight fiscal policy does not reduce imbalances because of the forward looking behavior of households. Even structural reforms via reducing markups are not considered to help reducing imbalances. However, beyond the crisis the scenario changes as negative wealth shocks on consumption materialize, external imbalances decrease while public deficits increase. During the scenario, GDP falls first and recovers gradually. In that scenario structural reforms and frontloading of fiscal consolidation help mitigating the short-term drop of output and avoiding the medium-term output loss.

3 Possible extensions

Concerning Gavilán *et al.*, while the overlapping generations model nicely captures the effect of demography on external imbalances in principle it is indeed a surprising result that fiscal policy would have no correcting impact on imbalances in the “normal times” period before the (post) crisis scenario. Maybe more differentiation is needed and the impact on imbalances needs to be considered in more detail. In fact, we are confronted with the issue of reversed causality between demography and imbalances. There is an economic intuition that the built-up of the house price bubble triggered immigration, particularly of low-skilled labor. Extending the causal chain, one could expand the model by incorporating the other side of the coin, which is capital flows. And one could check where capital inflows came from. One hypothesis could be that capital flows have been starting after the German reunification (big open economy). Capital released due to the German consolidation process in the Nineties and the beginning of the 2000s complemented or may even have triggered the imbalances in south European countries. Recent OECD analyses support that hypothesis since they show a strong correlation between taxes on housing and the house price bubble in Spain possibly affecting the external current account balance. Therefore, somehow fiscal policy influenced imbalances also before the crisis and the question is if such an effect can be integrated in the model. The sound rational behavior assumption of private borrowing replacing public borrowing might be challenged by introducing myopic behavior of households as an alternative assumption. Bouthevillain and Dufrénot differentiate between high growth countries showing Keynesian behavior and low growth countries showing Ricardian behavior. This could be introduced in the model by Gavilán *et al.*, too.

Concerning the paper by Bouthevillain and Dufrénot, the model could differentiate further regarding the conflict between output and efficiency, basically asking: Do public expenditures, e.g. education, health, and R&D, improve efficiency? The authors indeed point to possible inefficiencies in high growth countries. Therefore, the question is if the analysed countries lie on the 'efficiency frontier' and if there is a systematic link between effectiveness of public spending and its impact on growth. That in principle is an invitation to combine the work by Gavilán *et al.* and also Afonso with country samples regarding their growth level. Nevertheless, it is very difficult to separate productive and unproductive expenditures (see Brender's intervention in this session).

Concerning the data used, we would encourage the authors to disaggregate the dataset further. The current dataset only differentiates between 10 categories displaying functions of government expenditures, whereas the Eurostat dataset knows around 70 subcategories which can be assigned to productive or unproductive spending (COFOG 2 digit structure). This data structure makes it possible to better assess those subcategories that gather growth enhancing policies and this could render the model's results more specific. With respect to the structure of public expenditures one could refer to a German case study by FiFo Köln which tries to assess the effects on growth of different types of expenditures in Germany using the disaggregated COFOG l/2 digit level. Additionally, one should differentiate between several growth indicators and what they should measure: either short-term growth (GDP or GDP per capita) or long-term sustainable growth. The latter one is difficult to assess. A well-known indicator to describe medium-term growth is the potential GDP. Nonetheless, there are more ways to describe sustainable growth as for example environmental accounting or accounting considering ageing and demography. Using "growth potential" could help to take supply side effects into account. These seem to be neglected in the presented models since they are incorporated only tentatively. This may cause the model's result that fiscal policy does not affect imbalances, and, hence, this result might be misleading. Furthermore, one could control for fiscal institutions and measure the effect of changes in debt rules for example. This would be especially interesting with regard to the current developments in the euro area regarding the enhanced Stability and Growth Pact (SGP) and the Fiscal Compact.

4 Lessons for the EU governance

From the point of view of a ministry, it is especially interesting to ask for the "practical" relevance of these papers. Do they provide useful information to improve policies? The reformed SGP 3.0 that now focuses on fiscal sustainability has a very limited view on growth. On the other hand, the new macroeconomic surveillance process (Macroeconomic Imbalance Procedure, MIP) focuses on internal and external imbalances and hence looks at growth, although only indirectly. Also, the strategy 'Europe 2020' as a follow-up to the Lisbon Strategy, referred to in the paper by Bouthevillain and Dufrénot (Guideline 3 of the Integrated Guidelines), is diluted and has a very imprecise focus on 'sustainable' growth. All three processes are quite isolated although the EU Commission tried to gather them under an integrated framework, the European Semester.

What can we learn from the papers for strengthening the governance in the euro area? The approach by Bouthevillain and Dufrénot calls for a renewed agenda on the quality of public finances which should be integrated in 'Europe 2020' and the SGP. The approach of Gavilán *et al.* could help to analyse the links between the MIP and fiscal policy observation under the SGP. While currently no "one-size-fits-all" approach for EU member states' fiscal policies is possible or desired, one could think of alternative measures:

- the medium-term objectives (MTO) could be country-specific differentiating with respect to the country's business cycle, growth rate or effectiveness of public finances;

- the SGP thresholds could be made country-specific, modified with respect to the country's sustainability of public finances and MIP variables;
- "Europe 2020" should be redefined with regard to structural policies enhancing potential growth and be linked to the Euro Plus Pact.

In general, the institutional link between different fiscal policy measures and growth is weak and the impact of structural reforms on fiscal sustainability is widely neglected in the current framework. Therefore, both papers are highly relevant for the current debate on the EU and euro area governance architecture. We encourage the EU COM and the member states to have a close look at these different channels of fiscal policies and to make use of the general ideas of the two papers.

COMMENTS ON SESSION 4
POLICIES TO PROMOTE SUSTAINABLE GROWTH

David A. Heald[*]

I have three papers to discuss, and I will not pretend that there are common themes. The first is the paper by Balázs Égert about debt thresholds, then the paper by Jérôme Creel, Paul Hubert and Francesco Saraceno about the effects of alternative fiscal rules, and finally the paper by Ernesto Rezk, María de los Ángeles Mignon and Agustín Ramello De la Vega about human capital growth, with particular reference to Argentina. I am not an econometrician; my interest is on the public policy side. There are people here who know far more about the econometric methodology and are better able to argue about that than I am. So, other than asking one or two questions, I am not going to talk about that aspect.

Turning to the Égert paper from OECD, I think that is very interesting because there does seem to be something of a policy demand for evidence that there are thresholds. I sense that, given how high debt ratios are now, there is a policy demand for saying that they are now too high: above 90%, that is damaging. The Égert paper convinced me that the Reinhart and Rogoff (2010) results are not robust. I am not quite sure whether the argument is that there are no thresholds or whether the endogenous thresholds around 20% and 50% ought to be taken seriously.

The paper makes clear its counter-intuitive result that, beyond 90 per cent, the effects on growth become less negative or neutral; that is obviously extremely worrying in terms of making intuitive sense. And the question I would throw out to the author, and other people working in the field, is to what extent that is a result of the particular data or of the particular econometric techniques that are used. As a user of this kind of research, when the results are counter-intuitive, one needs to understand what exactly is driving those results.

It is not heavily emphasised in the paper but one of the points I noticed was the fact that the Reinhart and Rogoff data are not publicly available, and there has had to be a reconstruction which makes this paper not an exact replication of their work. The Égert data analysis looks at two time series; a long series from 1790 to 2010 that looks at central government debt; and a shorter time series from 1960 to 2010 that looks at general government debt rather than central government debt. The results are not substantially different. But that provoked a number of reflections on my part, very much about whether one ought to be thinking about central government debt or general government debt or public sector debt.

I am a Professor of Accountancy and one of the things that I have very much noticed is arbitrage techniques using accounting rules, particularly in the context of Public-Private Partnerships (Heald and Georgiou, 2010 and 2011a) and also exploiting the difference between general government and public sector (Heald, 2012). As fiscal austerity bites and with fiscal consolidation generally, one should start watching for arbitrage mechanisms. These might damage value for money; they are also going to contaminate the macroeconomic data.

The second point is that net debt misses lots of things, pension liabilities having been mentioned several times in this Workshop.

The third point I would make is that, in public debate, there is remarkable neglect of the assets side of the public sector balance sheet. In accruals-based government financial reports, the focus is on the net assets figure or in national accounts the net worth figure. Now I recognise that the data often are not very good but, when one is thinking about what kind of policy response there

[*] University of Aberdeen.

should be to particular levels of government debt, I would want to have some idea what the assets side of the government balance sheet looks like. You could be a country with very good infrastructure assets and a relatively high level of debt, but might well be in a better position than a country with very poor infrastructure assets and a lower level of debt. As an accountant, I want to think about both sides of the balance sheet. The United Kingdom has now produced Whole of Government Accounts and that is one of the issues that academically I have been working on (Heald and Georgiou, 2011b).

So my questions about the Égert paper are:

a) to what extent do the different results from Reinhart and Rogoff (2010) depend on the data used or on the econometric techniques used?
b) particularly with regard to the data, are the results driven by certain countries and certain years or certain periods within those very long time series?
c) is the policy conclusion that there is not an obvious threshold? I will come back to the debt issue during the discussion of the Creel *et al.* paper.

Francesco Saraceno presented the paper extremely well, so I am not going to talk about what he and his colleagues did (*Creel et al.*, 2012); other than make the point that this is modelling, from outside the official institutions, of the new European Union fiscal rules. The paper chooses four formulations of possible fiscal rules.

This raises in my mind two questions. First of all, do the modelling assumptions, which are described as New Keynesian, determine the results? To what extent do the judgements made behind the model building prejudge the results that are going to be achieved? Secondly, given the fact that this is unofficial modelling, it makes me interested in what official modelling has been done within the European Commission or elsewhere. The Creel *et al.* results favour the old UK-style golden rule (where investment is outside the golden rule) rather than the new European fiscal compact. Would alternative modelling, which can be defended on technical economic grounds, generate different results? I worry that sometimes economic and political judgements can become obscured by modelling complexities. Would differences in results – between the Creel *et al.* model on a New Keynesian basis and an alternative official model that may lie behind the European Commission and Council decisions – be driven just by different assumptions on how the economy works? A far broader range of people can become involved in discussions about how the economy works than in the econometric arguments.

Table 3 shows what Creel *et al.* (2012) call the investment rule leads to higher debt ratios than the new golden rule. A fundamental aspect of the present debate is uncertainty about the reaction of financial markets to much bigger debt ratios, which is what following their favoured rule would involve. Leaving aside the doubts created by the previous paper I was discussing, let us assume for the time being that higher public net debt does damage growth. The issue that would influence my policy view, probably more than anything else, is the question about how big a cushion economies need in the context of another collapse of the financial sector. The United Kingdom tends to talk about net debt; that was 35.8 per cent (31 March 2007) but has reached 66.2 per cent (31 March 2012), even when "excluding financial interventions" (139.9 per cent with them) (Office for National Statistics, 2012, Table PSF1). So there has been a remarkable increase in UK net debt. My policy instinct is to want as big a cushion as I can possibly have, if you do not have confidence that the financial sector has really changed.

Moving on to Rezk *et al.* (2012) paper, which has two parts. The printer did not work properly but Ernesto very kindly annotated my copy for me, so I could work through it again. The expositional part, on the theory, I found very helpful. Some of the things I had read before; some I just did not know. The way it developed successive models, about how human capital might influence growth, was very helpful and I found it very informative.

The second part of the paper is the empirical results and I think there is evidence of haste in the way that these are reported. I found Ernesto's presentation more helpful than the actual paper, in terms of interpretation and of giving me some idea of which of the various results he placed more confidence in. The paper brought home to me the issue of finding good data for human capital. In no sense am I going to criticise the data that have been used, but years of schooling does seem a fairly limited measure. It also emphasises the importance of good social statistics, as well as good economic statistics. If one thinks that there is a complete separation between the social data and the economic data, you can concentrate on securing good economic data. But, as soon as you start arguing that human capital development is important in a growth context, it is very important to emphasise good social statistics and making sure that national statistics and social statistics do not themselves become a casualty of fiscal consolidation.

A final point on issues that came up in the previous papers in this Session of the Workshop. I become very worried about this discussion of "productive" and "non-productive" public expenditure. Hence I very much like the way that the Banque de France paper (Bouthevillain and Dufrénot, 2012), which I am not commenting on, went to the COFOG data. In my own country, when people talk about productive public expenditure, some of what they want to call productive does not strike me as particularly productive at all; some of it is just industrial subsidies. So I think that one has to be careful about the language of productive and non-productive. The more that you stress human capital development, as Ernesto does, the effect that public spending has on human capital may well come from diverse types of spending. This is an important area for research and that research is going to depend on securing better data, as well as protecting existing sources of data.

REFERENCES

Bouthevillain, C. and G. Dufrénot (2012), "Fiscal Policies Enhancing Growth in the Euro Area Countries: Can We Apply Common Remedies to Different Countries?', in this volume.

Creel, J., P. Hubert and F. Saraceno (2012), "An Assessment of Stability and Growth Pact Reform Proposals in a Small-scale Macro Framework", in this volume.

Égert, B. (2012), "Public Finances and Economic Growth in the Long Run: Evidence from Non-linear Bayesian Model Averaging", in this volume.

Heald, D.A. (2012), "Why Is Transparency About Public Expenditure so Elusive?", *International Review of Administrative Sciences*, Vol. 78, Issue 1, pp. 30-49.

Heald, D.A. and G. Georgiou (2011a), "The Substance of Accounting for Public-private Partnerships", *Financial Accountability & Management*, Vol. 27, Issue 2, pp. 217-47.

——————(2011b), ''The Macro-fiscal Role of the U.K. Whole of Government Account", *Abacus*, Vol. 47, Issue 4, pp. 446-76.

Office for National Statistics (2012) *Statistical Bulletin: Public Sector Finances, September 2012*, Newport, Office for National Statistics.

Reinhart, C.M. and K.S. Rogoff (2010), "Growth in a Time of Debt", *American Economic Review*, Vol. 101, Issue 5, pp. 1676-706.

Rezk, E., M. De los Ángeles Mignon and A. Ramello De la Vega (2012), "Human Capital Formation in Argentina: Contribution to Gross Domestic Product", in this volume.

COMMENTS ON SESSION 4
POLICIES TO PROMOTE SUSTAINABLE GROWTH

Sergey Vlasov[*]

The two papers on which I would like to comment, prepared by OECD and Serbian colleagues, cover very different topics. So, let me take these papers up in turn.

1 Comments on "Fiscal Consolidation Needs and Implications for Growth" by Jørgen Elmeskov and Douglas Sutherland

The paper prepared by Elmeskov and Sutherland studies the instruments of public debt levels adjustment with the implications for growth in OECD area. It is a very accurate, well-built paper, mainly descriptive, but supported with the corresponding calculations made by authors themselves or their OECD colleagues.

The motivation of the paper is the dramatic increase in the public debt in the OECD area that has trended upwards since early 1970s and passed annual GPD in 2011 both because of the pro-cyclical fiscal policy during the expansion before the crisis and severe crisis consequences. High public debt, in its turn, has adverse effects on economic growth by raising the costs of capital – reducing productivity and leading to a level shift in potential output in the short run and reducing investment in research and development in the long run – and by crowding out effect – the real interest shock. So, there is a clear necessity to reduce debt overhangs, including creating some room to react to possible future shocks.

Now let me turn to the main findings as well as to make some comments. Talking about the size of adjustment in OECD area the authors present the calculations of what has already been done or is under way as well as what should be done in the long run. What is interesting to note is that Greece, Portugal, Spain and Ireland, the countries associated with possible debt crisis in Euro area, not just make the largest cumulative fiscal tightening between the deficit trough and 2012 (as we can see in Figure 4) but also have the modest adjustment needs on a period till 2050 under conditions of bringing down gross financial liabilities to 50 per cent of GDP (with the only exception for Ireland – see Figure 5). First, how big is the risk that in 2012 the reporting figures would not correspond to those planned? For example, the situation in Greece – I mean considerable public protest against fiscal consolidation measures adopted by the government – allows to suppose that the expected effect will not be fully achieved (directly or indirectly). Second, how do the authors' estimations correlate with the low sovereign ratings given to these countries by international rating agencies and, for instance, the speech of the head of the World Bank, R. Zoellick, who told relatively not long ago that Greece will inevitably default and it is just a matter of time? Possibly there is a preliminary estimation on 2012 to support authors' calculations?

The debt overhang can be worked off in two ways. The first is by primary balance tightening and the second is by using the real growth and real interest rate effects. Let me start from the latter one by looking at first at the Russian experience. Following the Asian financial crisis and the deterioration of external conditions the Russian government had to declare a sovereign default and to allow the depreciation of the Russian ruble. The federal government debt increased dramatically to 137.4 per cent of GDP by the end of 1998. But only about half a decade later GDP growth effect,

[*] Bank of Russia.

Figure 1

Components of the Russian Federal Government Debt Dynamics in 1996-2000 (left axis) and 2001-2011 (right axis)
(percent of GDP)

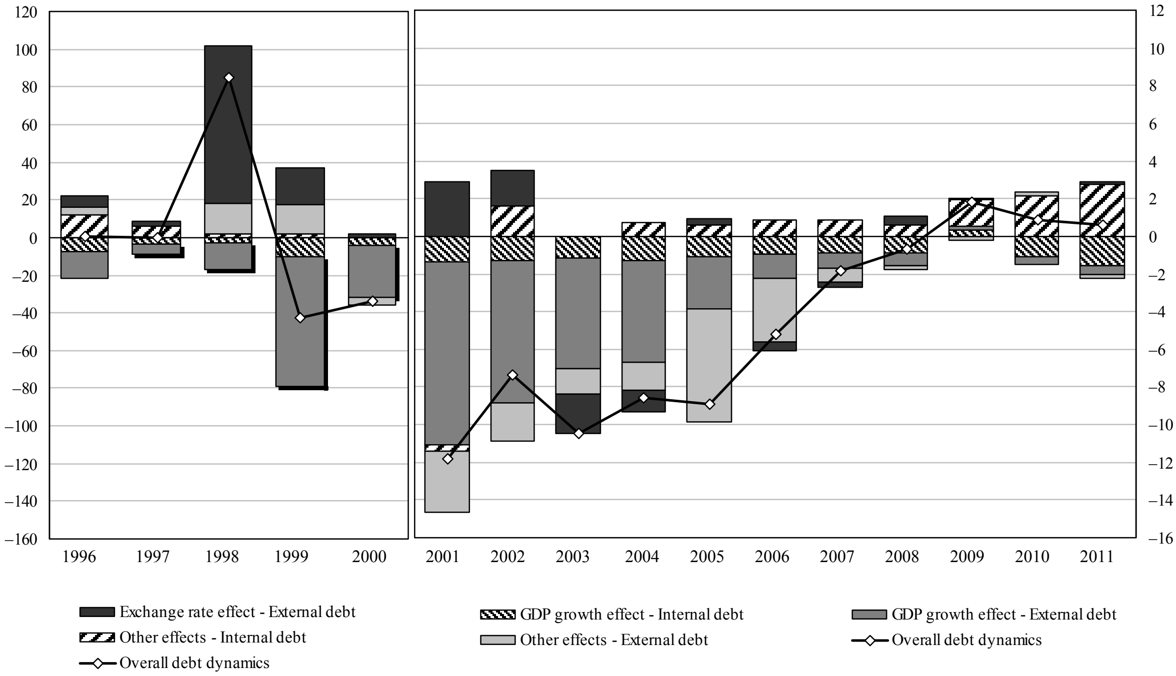
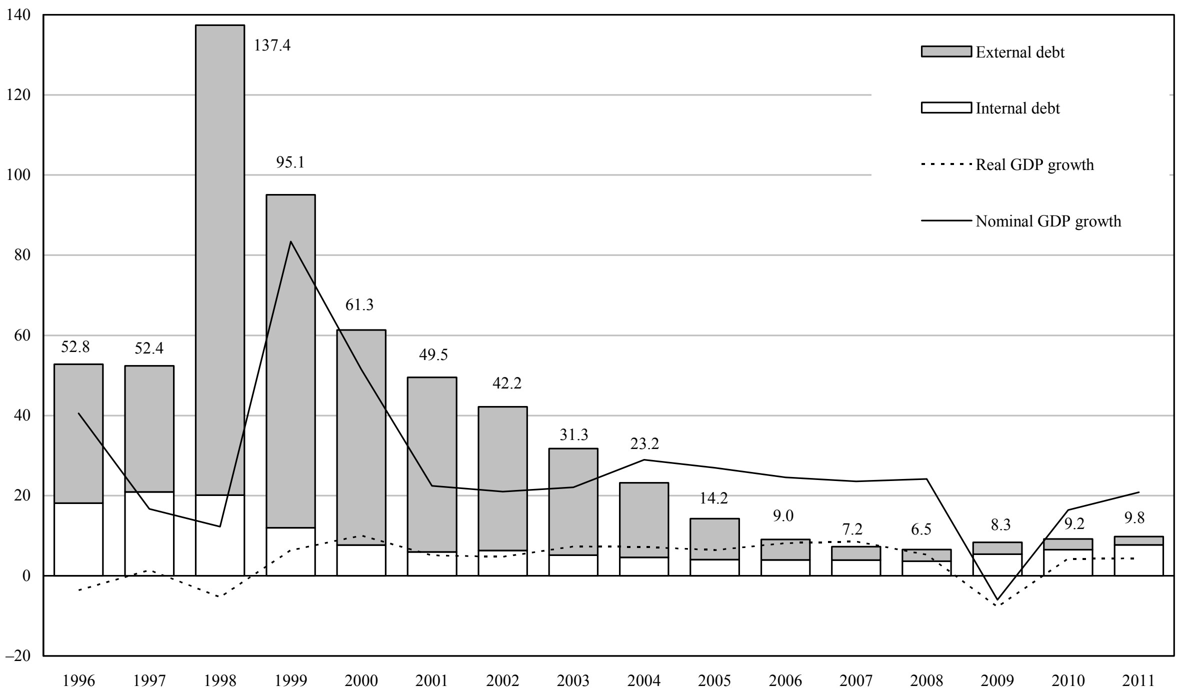


Figure 2

Dynamics and Structure of the Russian Federal Government Debt
(percent of GDP and nominal and real GDP growth rates in 1996-2011)



supported by favourable external conditions and prudent fiscal policy, allowed to reduce the public debt level from one of the highest in the world to one of the lowest ones (see Figures 1 and 2).

Probably this is not the case for the OECD countries and authors' calculations fully support this idea (see Table 2 for the corresponding calculations of real growth effect). So, the authors analyze a wide range of possible instruments of fiscal consolidation and quantify their contribution to primary balance tightening for each country (see Table 3). While implementation of the most of them seems reasonable there are some risks and limitations that have to be taken into account. Also several general comments could be made.

First, I have some doubts about the possibility to use OECD average as a target value for a set of instruments, e.g. social transfers, subsidies, tuition fees, as countries' peculiarities seem worthwhile to be taken into account. For instance, high values of social indicators in the Scandinavian region is a distinctive feature of the policy in these countries, partly concerned with their tax system.

Second, the level of discontent among the population as a result of possible employees' layoffs, social spending cuts and even increase in so-called "sin" taxes should be taken into account.

Third, it seems worthwhile to pay more attention in the paper to the pension reform: how it should be carried out and the corresponding calculations. This issue is very complicated and should become an important contribution to the fiscal tightening. For example, the Russian government, in contrast to the most of the countries, has an intension to avoid retirement age rise and find the required financial resources by reviewing the pension system components. This decision reflects possible political costs as the share of pensioners in the electorate is more than a third at present and presumably will rise gradually in the long run.

Fourth, it should be interesting to see how substantially can the primary balance be improved through the of automatic stabilizers operation. Although the model that is used to calculate fiscal gaps (see Appendix) implies the use of automatic stabilizers while the economy moves back to potential output, the paper does not present the corresponding calculations of their size for OECD and/or individual countries.

Finally, the paper contains the discussion of the choice of instruments based on their possible effect on long-term growth, e.g. that the differences in multipliers make tax increases more attractive than the government spending cuts. But it would be a nice contribution to the paper if the authors could estimate (at least make rough calculations of) the effect of fiscal consolidation (measures) on GDP growth rates for OECD and/or individual countries.

2 Comments on "Equity Aspects of VAT in Emerging European Economies: The Case Study of Serbia" by Nikola Altiparmakov and Milojko Arsić

The paper by Altiparmakov and Arsić is rather specific and aimed at proving VAT progressiveness in Serbia by investigating the VAT system in very details.

The motivation of this paper is the existence of the world-wide trend, aimed specifically at boosting the economic growth, of shifting the tax burden from income to consumption taxation, in particular VAT. At the same time there is common belief in the general public that VAT is a regressive tax that has an adverse distributional effects by creating disproportionate tax burden on the poor households. So, using the micro-level data the authors investigate equity aspects of VAT in Serbia as a typical emerging European country to prove its progressiveness.

Let me turn to the main findings of the paper. In order to determine whether the VAT is progressive or regressive the authors make calculations of the effective VAT rate for ten deciles of the population: from the poorest to the richest. Following the common practice the authors first investigate VAT burden by household annual income. It yielded unreasonable results indicating that the poorest group's of population effective VAT rate lies outside the range from 0 per cent the standard VAT rate (see Table 5a). These results show VAT regressiveness and are explained by the significant presence of the shadow economy and the evasion of direct income taxes in emerging market economies and by relying on annual income framework instead of lifetime income framework.

On the contrary, investigating VAT burden by looking at household annual expenditure leads to reasonable results and allows disclosing VAT progressiveness (see Table 5b). In about 60 per cent of the effect of progressivity is explained by taxation under reduced rate, while the remaining by progressive incidence of own-farming production, which is widely used in Serbia.

Finally, the authors investigate the prospective increase in VAT burden in Serbia by examining three possible scenarios. They conclude that introducing the uniform rate by eliminating the reduced rate contributes to regressivity and should be chosen only if accompanied with progressivity-enhancing reforms (see Table 7 for results).

There are three comments that I would like to make.

First, as a representative of another typical emerging European country which is very close to Serbia in many aspects, including those close to the topic of the paper (significant presence of the shadow economy and the evasion of direct income taxes, widely used own-farming production and VAT system as a whole), and basing on the Russian experience, let me express some doubts about the common (mis)belief in general public in Serbia that VAT is a regressive tax. I believe that scientists or politicians care about this issue, in particular the latter ones can use it in their political activity. At the same time I guess common people resist the prospective increase in VAT burden not because they believe that this tax burdens poor individuals more but because they resist the increase in tax burden at all.

Second, it would be a good contribution to the paper if authors would present their proposals how to modify VAT system in Serbia basing on such aspects as offering special consumption incentives, boosting economic growth, improving fiscal sustainability etc. They may include not only the suggestions of eliminating reduced rate or exemption from VAT for certain goods but any tax relief as well. For instance, in accordance with the Russian VAT system goods for children are a subject to the reduced VAT rate. It allows to promote their consumption and, therefore, to increase fertility rate.

Finally, is there any effective VAT rate for achieving the most of redistributive objectives, e.g., as the share of the standard VAT rate, especially for the poorest households? May be some literature on this issue exists?