Session 2

EVALUATING THE EFFICIENCY AND EFFECTS OF PUBLIC SPENDING

THE ROLE OF THE "QUALITY OF PUBLIC FINANCES" IN THE CONTEXT OF THE REFORMED STABILITY AND GROWTH PACT

Florian Höppner and Christian Kastrop[*]

1. The concept of "quality"

The concept of the "Quality of Public Finance" has gained increased importance in recent years within the framework of budgetary discussions at the national as well as at the EU level. The March 2003 Ecofin Council concluded that "greater attention should be paid, within the overall constraints of the Stability and Growth Pact, to the quality of public finances with a view to raising the growth potential of the EU economies".

The concept of quality pays tribute to the fact that a long term strategy for the consolidation of public finances has to take into account not only *quantitative* aspects of consolidation but also the issue of *qualitative* or *structural* consolidation. The growth-enhancing restructuring and the efficiency-improving design and management of public expenditure (and revenues) can quite clearly be described as a major policy challenge with evident macro- and microeconomic implications, an aspect often neglected in the practice of fiscal policy. It should be asked not only "how much money is spent", but just as well "how is the money spent". In this context, even traditionally accepted indicators of "good expenditures" such as public investment should be reviewed.

At EU level the issue has gained further political importance in view of the Lisbon goals and the "Broad Economic Policy Guidelines" (BEPGs) which emphasise already that "governments can contribute to achieve the Lisbon goals by spending money as efficiently as possible, by redirecting public expenditure towards growth-enhancing cost-effective investment and human capital and knowledge subject to overall budgetary constraints, and by seeking a higher leverage of public support on private investment". Before going into detail in the context of the reform of economic governance in the EU, clarification is needed on what "quality of public finances" really means.

Following recent work by the Commission, with respect to the present focus of the EU fiscal framework on macroeconomic aspects three dimensions of budgeting can be identified:

- ensure fiscal control and fiscal discipline;
- to provide a degree of stabilisation of the economy;

---

[*] Fiscal Policy and Economic Affairs Directorate, Federal Ministry of Finance, Wilhelmstraße 97, 10117 Berlin, Germany. E-mail: Florian.Hoeppner@bmf.bund.de and Christian.Kastrop@bmf.bund.de

The authors present their personal opinions and not necessarily the opinion of the German Federal Ministry of Finance.

- to promote allocative and technical efficiency when using public resources.

The issue of quality of public finances refers in particular to the third dimension because of increasing recognition that fiscal policy should contribute more systematically to the Lisbon objectives and that the refocused Lisbon Agenda (to long-term growth and employment) is to be reflected in the budget. Key questions that have to be answered in this context are: Is the allocation of resources in line with the strategic objectives of the government (currently in particular long-term growth)? What is the role of structural reforms in trying to achieve sustainable public finances by enhancing long-term growth? In implementing policies, are public resources used in the most efficient and effective way (value for money)? Which institutional arrangements would help to redirect public expenditures systematically with a view to long-term growth?

The questions of allocative and administrative efficiency draw attention to the composition of public expenditure as well as the structure of the tax system. On the expenditure side, a key issue is that of identifying potentially "productive" or "growth-enhancing" expenditure (covered by categories like R&D, education and infrastructure investment). Additionally an integrated overall cost/benefit and input/output (efficiency and effectiveness) assessment is needed to judge whether the benefits of a particular type of expenditure outweigh the costs, of course answering the question of "market failure" before the brackets. Detailed and country-specific assessments are therefore needed to guide the composition of national public expenditure. An issue relevant to all EU Member States is to closely monitor the expenditure dynamics to prevent productive expenditure from being crowded out by increasing ageing-related expenditure categories or interest payments. On the revenue side, it is important to set up tax structures which strengthen the growth potential by promoting employment creation and investment. In addition, an analysis of tax expenditures (e.g. tax exemptions) as substitutes for direct expenditures may produce significant insights.

These are several important facets of the complex issue of "quality of public finances", while we do not aim at defining the concept once and for all. We rather intend to indicate several questions that are at the core of actual fiscal policy making and that at the same time are underrepresented in the current framework of economic governance in the EU. In the following we aim to discuss the issue of quality in more detail with particular reference to the Stability and Growth Pact.

2. The reform of the Stability and Growth Pact

During the last year an intense and controversial discussion about the reform of the Stability and Growth Pact took place within the EU. A first agreement on fundamental principles of a "reformed" Pact had been reached at the special meeting of the Ecofin Council on 20th of March 2005, while meanwhile also the legal implementation of this agreement in the Council Regulations has been successfully completed. However, the general discussion about the reform of the Stability and

Growth Pact has shown that there is no clear common understanding on how fiscal coordination in the EU should function in detail. Therefore, although the reform of the Pact has been accomplished "on the paper", it remains an open question how exactly the new Pact will be implemented in the future – a consensus on a new "philosophy" of the application of the Pact still needs to be found. In the following we want to analyse in some more detail the role of the "Quality of Public Finances" in this debate.

Before the question of *how* in detail fiscal policy coordination should be organised within the European Monetary Union, the question of why we need a "Stability Pact" at all – *i.e.* what is the core that has to be "regulated" by the instrument of coordination of fiscal policies – needs to be looked at. With respect to this question a widely accepted consensus exists: the key is to achieve and maintain the stability of the common currency (inflation and exchange rate), for fiscal policies this means the securing of sound budget policies to support the common monetary policy. Hence, to avoid in particular negative spillover effects through different channels like e.g. interest rates or inflation, a common currency area needs some sort of "boundaries" for fiscal policy. The Stability and Growth Pact in connection with Article 104 of the Treaty is one such possible boundary, with the well known reference values for the deficit and the debt level, and in particular the clarification and operationalisation of the Excessive Deficit Procedure (EDP) of the Treaty.

We are now able to look back on some years of experience with the Stability and Growth Pact. In our opinion one can identify several problems that result from the initial construction of the Pact and that were also central issues in the debate of its application that eventually led to the reform initiative by the Commission:

- the short-run cyclical development had not been taken into account adequately when assessing the fiscal policies of the member states and when formulating recommendations as part of the EDP. Where monetary policy can consider only the monetary union as a whole, asymmetric shocks can be offset only by letting the automatic fiscal stabilizers work. But this is impossible if the Pact is interpreted too mechanistically;
- structural reforms on the one hand are seen as positive in the overall economic policy approach of the EU-coordination mechanism, however, potential short-run effects of the implementation of structural reforms had not been taken into account systematically in the Stability and Growth Pact. There is widespread agreement that after the abandonment of the real exchange rate as a means to adjust for differences in Member States' relative competitiveness, greater supply-side flexibility is needed. However, the necessary structural reforms may be politically unfeasible if there is no scope for fiscal policy to offset potential adverse effects on output in the short run;
- a too narrow focus on the 3 per cent deficit reference value led in some cases to pro-cyclical fiscal policies and the use of one-off measures, as a deficit had to be corrected in the year after its diagnosis.

The last issue is a major point, as the experience with the application of the Pact during the last years has revealed in particular that the narrow focus on the

3 per cent deficit limit does not adequately take into account the actual complexity of fiscal policy. The philosophy of relying mechanically on quantitative controls as the exclusive decision parameter of the Pact has proved to be very problematic. One could observe that the concept which has so far underlain the application of the Pact is an illusion – namely that the key instrument to prevent and cure excessive deficits is to set detailed consolidation targets down to a fraction of a percentage point.

When decisions under the Stability and Growth Pact are based solely on quantitative indicators rather than primarily on an analysis of the underlying fiscal policy, the danger of economically false recommendations or targets is quite high. As the Pact has been applied, in some cases countries with deficits exceeding 3 per cent have been urged to follow a pro-cyclical fiscal policy in order to achieve a short-term deficit reduction even at the risk of prolonging a period of weak growth – contrary to the Lisbon objectives – and thus making it more difficult to achieve the medium to long term consolidation objectives. In some circumstances countries have equally been impeded from initiating and applying necessary structural reforms which would have strengthened growth, while being obliged to push through one-off measures in order to demonstrate progress in consolidation in the short term. Thus a very mechanical interpretation of the Pact in our view not only undermined the economic rationale of the instrument but would also hamper the credibility of the Pact within the markets or the general public even if accepting a certain credibility trade-off between more sophisticated economic reasoning and a very mechanistic, simplistic but transparent Pact.

What instead is needed in our opinion is to broaden the perspective on the Stability and Growth Pact and in particular taking aspects related to the quality of public finances more seriously. The application of the Pact in this concept has to focus on the "right" policies, *not primarily* on quantitative indicators.

In the context of the reformed Pact, quite often the accusation is made that politicians are mainly interested in a pure deducting, e.g. of expenditure categories or the short term costs of structural reforms, from the official budget figure. For this reason, any discussion in the direction of an "overall assessment" of fiscal policies of the type we support risks being blocked without an unprejudiced look at the underlying argument. In our view, life is more complex: a simple deducting as well as a mechanistic approach to the Pact looking only at 3.0 per cent are both extremes that hamper a proper functioning of the Pact.

Again: of course there is a trade-off between the need for a simple and clear fiscal rule and the degree of discretion that has to be exercised in an individual assessment of the respective country in an EDP. Nevertheless we argue that in the application of the "old" Stability and Growth Pact this balance had been biased too much towards an overly mechanistic approach. It is without any doubt to some extent easier and more transparent to have just one single indicator and judge policy mainly by numbers, but we think that this approach can not ensure an adequate assessment as well as the right recommendations in every case.

3. The role of "quality" in the future application of the Stability and Growth Pact

In the reformulated Stability and Growth Pact the quality aspect is now mentioned explicitly. With respect to the "corrective arm" of the Pact (Council regulation (EC) 1467/97), article 2 states: "The commission, when preparing a report under Article 104 (3) of the treaty shall take into account all relevant factors as indicated in the article. The report shall appropriately reflect developments in the medium term economic position … and developments in the medium term budgetary position (in particular, fiscal consolidation efforts in "good times", debt sustainability, public investment and the overall quality of public finances)".

In line with that we argue that in an individual and overall assessment of fiscal policies in the context of the Pact more qualitative criteria should be used, without questioning the nominal anchor function of the quantitative reference values of the Treaty as the guiding principle of fiscal policy coordination. This would also take account of the fact that as a general rule there exist no blanket, universally applicable "patent remedies" for all cases, although in particular textbook economics often argues along this line.

Now, while the concept of quality certainly is a theoretically appealing concept, how can it be made operational especially in the context of the Stability and Growth Pact? So far there exist no clear cut and established answers, nevertheless some first thoughts shall be offered in the following.

Firstly, an analysis of the quality of the public-sector budgets must take account of the general structure of expenditure and revenue. Expenditure for past obligations should be reduced as much as possible in public budgets (here is a direct connection to the aspect of structural reforms mentioned before), while expenditure on "future-oriented" and growth enhancing items should be strengthened. It should be noted that the conventional concept of public investment is not a suitable measure of the quality of the budget in this respect. Rather, an overall analysis of expenditure and revenue should take into account the effects on growth of these components of the budget. High-quality expenditure in this context may be expenditure on education or Research and Development. This analysis should be made against the background of the type of expenditure and investment which is needed in a particular country – as regards the quality of public-sector budgets, as in other fields, there are no universally valid answers or *a priori* quality items.

Secondly, an assessment of the fiscal policy of a country should take into account if necessary structural reforms have been or are in the course of being implemented. The reformed Stability and Growth Pact now explicitly makes an allowance for possible short-term deficit-raising effects of reforms, e.g. dampened growth due to uncertainty and reticence of investors and consumers in the short term, or a transitional phase of higher expenditures. An important type of measures which should be mentioned here are also tax reforms. Tax reforms may be needed to strengthen growth and improve the competitiveness of a country, an example would be the German tax reform implemented in the year 2000, which was also positively

assessed by the Commission. Altogether a better coordination between the Lisbon Agenda and the Stability and Growth Pact is needed to avoid inconsistencies between the different coordination instruments.

Of course it has to be stressed again that the quality issue is only one aspect in the overall concept of the Stability and Growth Pact – but one that in our opinion had clearly been underrepresented in the "old" Pact, while its recent reform made some improvements in this direction. However, to be able to proceed even further along the lines sketched out very briefly, the following questions have to be addressed.

How can we analyse and assess growth-relevant public expenditures (e.g. on investment, R&D, education) when assessing public budgets in the procedures defined by the preventive as well as the corrective arm of the Stability and Growth Pact? How can a reform process aiming at a quality oriented consolidation be considered in such an assessment? Moreover, within the Lisbon process, how can we take qualitative aspects of expenditure policy into account as "high quality" public expenses help realise the Lisbon targets, while recognising that consolidation policy must not concentrate on quantitative issues alone? In a further perspective how could we succeed in combining the Stability and Growth Pact and the Lisbon agenda to a coherent strategy?

A pragmatic way forward would be to find a non-mechanistic assessment methodology that explicitly links quality-oriented reforms, budget composition and its restructuring, evaluation of effectiveness, efficiency and growth performance with the formal budget constraints of the Stability and Growth Pact.

Last but not least, we should also look closely at the underlying variable for all of these: the institutional arrangement. This means looking at all relevant fiscal rules, budgetary procedures and – especially in federal states – at the intra-federal structures like equalization systems or National Stability Pacts. The Ecofin conclusions on the Stability and Growth Pact reform from March 2005 state: "National budgetary rules should be complementary to the member states commitments under the Stability and Growth Pact. Conversely, at EU level, incentives should be given and disincentives removed for national rules to support the objectives of the Stability and Growth Pact … The implementation of existing national rules could be discussed in stability and growth programmes".

The Ecofin Council and at a more technical level the Economic Policy Committee (EPC) of the EU would be well suited to do a lot of preparatory work along these lines (a first starting point is the work of the newly established "Working Group on the Quality of Public Finances" of the EPC). From that framework general guidelines should be derived that should serve as policy recommendations in this field. The EPC could propose options for the institutional setting of such an assessment and a package of recommendations reflecting the analysis of the relationship between the quality of public finances and the Stability and Growth Pact.

EFFICIENCY OF PUBLIC SPENDING IN DEVELOPING COUNTRIES: AN EFFICIENCY FRONTIER APPROACH

Santiago Herrera and Gaobo Pang[*]

1. Introduction

Governments of developing countries typically spend resources equivalent to between 15 and 30 per cent of GDP. Hence, small changes in the efficiency of public spending could have a significant impact on GDP and on the attainment of the government's objectives whichever these are. The first challenge faced by stakeholders is measuring and scoring efficiency. This paper attempts such quantification. Additionally it verifies statistically some empirical regularities that describe the cross-country variation in the estimated efficiency scores.

The paper has four chapters following this Introduction. The first one presents the methodology that defines efficiency as the distance from the observed input/output combinations to an efficient frontier. This frontier, defined as the maximum attainable output for a given input level, is estimated using the Free Disposable Hull (FDH) and Data Envelopment Analysis (DEA) techniques. The exercise focuses on health and education expenditure because they absorb the largest share of most countries' budgets, and because of lack of data availability for international comparisons in other types of expenditures.

The second chapter estimates the efficiency frontiers for nine education output indicators and four health output indicators, based on a sample of 140 countries and data for 1996-2002. Both input efficiency (excess input consumption to achieve a level of output) and output efficiency (output shortfall for a given level of inputs) are scored. The chapter presents both the single input/single output and the multiple inputs/multiple outputs frameworks. In addition, this chapter explores how expenditure efficiency has changed over time.

The third chapter seeks to identify empirical regularities that explain cross-country variation in the efficiency scores. Using a Tobit panel approach, this chapter shows that higher expenditure levels are generally associated with lower efficiency scores. Similarly, countries in which the wage bill is a larger share of the total budget tend to have lower efficiency scores. Three other variables that explain

---

[*] World Bank.

The paper is a draft of work in progress that carries the names of the authors and should be cited accordingly. The findings, interpretations and conclusions are the authors' responsibility and do not necessarily represent the view of the World Bank or its Executive Directors. The authors thank Antonio Estache, Jorge Garcia, Ulrich Lachler, Vikram Nehru, participants of the Economic Policy and Debt seminar, and participants of the Banca d'Italia Workshop on Public Finance for helpful comments on earlier drafts of the paper. The dataset, specific-country results, and programs are available in the World Bank PRMED (Economic Policy and Debt Department) website http://intranet.worldbank.org/WBSITE/INTRANET/SECTORS/INTECONOMICPOLICY/INTQFA/0,,menuPK:369265~pagePK:151716~piPK:176772~theSitePK:369258,00.html

the cross country variation in efficiency scores are the degree of urbanization (positively correlated with efficiency, the prevalence of the HIV/AIDS epidemic (negatively associated with efficiency scores), and inequality in income distribution (higher inequality associated with lower efficiency).

The fourth and last chapter summarizes the conclusions.

2. Measuring efficiency: methodologies and overview of the literature

The object of this chapter is to briefly describe the specific empirical methods applied in this paper to measure efficiency and to survey the literature more directly related to the analysis of public expenditure efficiency. Empirical and theoretical measures of efficiency are based on ratios of observed output levels to the maximum that could have been obtained given the inputs utilized. This maximum constitutes the efficient frontier which will be the benchmark for measuring the relative efficiency of the observations. There are multiple techniques to estimate this frontier, surveyed recently by Murillo-Zamorano (2004), and the methods have been recently applied to examine the efficiency of public spending in several counties. These are the topics of the next two sections.

2.1 Methods for measuring efficiency

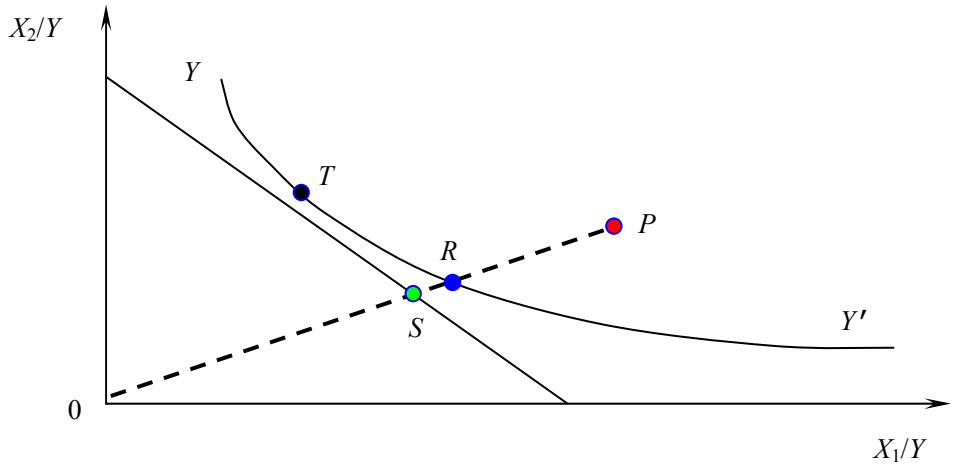
The origin of the modern discussion of efficiency measurement dates back to Farell (1957), who identified two different ways in which productive agents could be inefficient: one, they could use more inputs than technically required to obtain a given level of output, or two, they could use a sub-optimal input combination given the input prices and their marginal productivities. The first type of inefficiency is termed technical inefficiency while the second one is known as allocative inefficiency.

These two types of inefficiency can be represented graphically by means of the unit isoquant curve in Figure 1. The set of minimum inputs required for a unit of output lies on the isoquant curve *YY*'. An agent's input-output combination defined by bundle *P* produces one unit of output using input quantities X_1 and X_2. Since the same output can be achieved by consuming less of both inputs along the radial back to bundle *R*, the segment *RP* represents the inefficiency in resource utilization. The technical efficiency (*TE*), input-oriented, is therefore defined as $TE = OR/OP$. Furthermore, the producer could achieve additional cost reduction by choosing a different input combination. The least cost combination of inputs that produces one unit of output is given by point *T*, where the marginal rate of technical substitution is equal to the input price ratio. To achieve this cost level implicit in the optimal combination of inputs, input use needs to be contracted to bundle *S*. The input allocative efficiency (*AE*) is defined as $AE = OS/OR$.

The focus of this paper is measuring technical efficiency, given the lack of comparable input prices across the countries. This concept of efficiency is narrower

Figure 1

Technical and Allocative Inefficiency



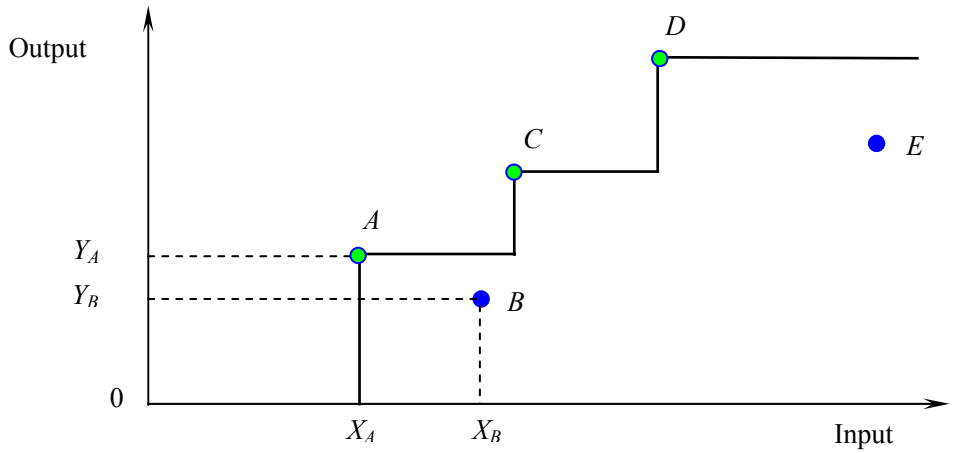
than the one implicit in social welfare analysis. That is, countries may be producing the wrong output very efficiently (at low cost). We abstract from this consideration (discussed by Tanzi, 2004), focusing on the narrow concept of efficiency.

Numerous techniques have been developed over the past decades to tackle the empirical problem of estimating the unknown and unobservable efficient frontier (in this case the isoquant *YY*"). These may be classified using several taxonomies. The two most widely used catalog methods into parametric or non-parametric, and into stochastic or deterministic. The parametric approach assumes a specific functional form for the relationship between the inputs and the outputs as well as for the inefficiency term incorporated in the deviation of the observed values from the frontier. The non-parametric approach calculates the frontier directly from the data without imposing specific functional restrictions. The first approach is based on econometric methods, while the second one uses mathematical programming techniques. The deterministic approach considers all deviations from the frontier explained by inefficiency, while the stochastic focus considers those deviations a combination of inefficiency and random shocks outside the control of the decision maker.

This paper uses non-parametric methods to avoid assuming specific functional forms for the relationship between inputs and outputs or for the inefficiency terms. A companion paper will explore the parametric approach, along the lines proposed by Greene (2003). The remainder of the section briefly describes the two methods: the Free Disposable Hull (FDH) and the Data Envelopment Analysis (DEA).

Figure 2

Free Disposal Hull (FDH) Production Possibility Frontier

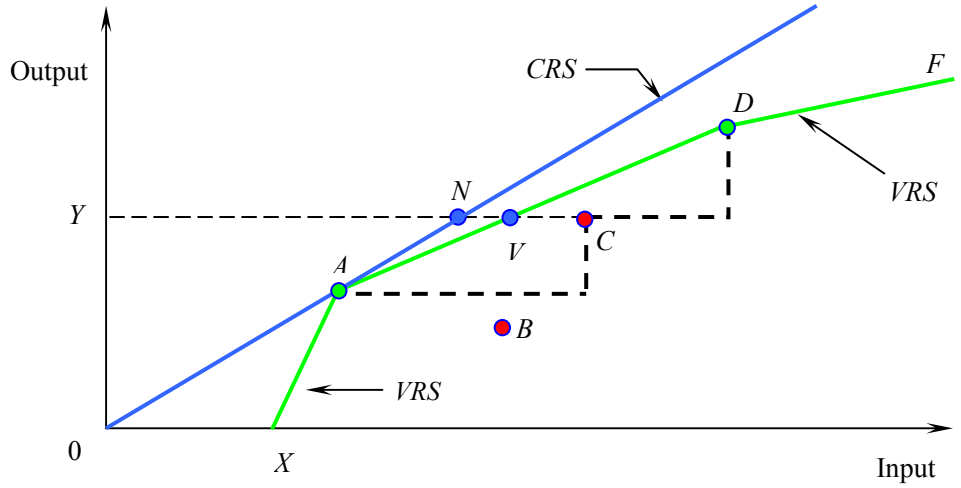


The FDH method imposes the least amount of restrictions on the data, as it only assumes free disposability of resources. Figure 2 illustrates the single input/single output case of FDH production possibility frontier. Countries A and B use input X_A and X_B to produce outputs Y_A and Y_B, respectively. The *input efficiency* score for country B is defined as the quotient X_A/X_B. The *output efficiency* score is given by the quotient Y_B/Y_A. A score of one implies that the country is on the frontier. An input efficiency score of 0.75 indicates that this particular country uses inputs in excess of the most efficient producer to achieve the same output level. An output efficiency score of 0.75 indicates that the inefficient producer attains 75 per cent of the output obtained by the most efficient producer with the same input intake. Multiple input and output efficiency tests can be defined in an analogous way.

The second approach, Data Envelopment Analysis (DEA), assumes that linear combinations of the observed input-output bundles are feasible. Hence it assumes convexity of the production set to construct an envelope around the observed combinations. Figure 3 illustrates the single input/single output DEA production possibility frontier. In contrast to the vertical step-ups of FDH frontier, DEA frontier is a piecewise linear locus connecting all the efficient decision-making units (DMU). The feasibility assumption, displayed by the piecewise linearity, implies that the efficiency of C, for instance, is not only ranked against the real performers A and D, called the peers of C in the literature, but also evaluated with a virtual decision maker, V, which employs a weighted collection of A and D inputs to yield a virtual output. DMU C, which would have been considered to be efficient by FDH, is now lying below the variable returns to scale (*VRS*, further defined below) efficiency frontier, *XADF*, by DEA ranking. This example shows that FDH tends to assign

Figure 3

DEA Production Possibility Frontier



efficiency to more DMUs than DEA does. The input-oriented technical efficiency of *C* is now defined by *TE* = *YV*/*YC*.

If constant returns to scale (*CRS*) characterize the production set, the frontier may be represented by a ray extending from the origin through the efficient DMU (ray *0A*). By this standard, only A would be rated efficient. The important feature of the *XADF* frontier is that this frontier reflects variable returns to scale. The segment *XA* reflects locally increasing returns to scale (*IRS*), that is, an increase in the inputs results in a greater than proportionate increase in output. Segments *AD* and *DF* reflect decreasing returns to scale. It is worth noticing that constant returns to scale technical efficiency (*CRSTE*) is equal to the product of variable returns to scale technical efficiency (*VRSTE*) and scale efficiency (*SE*). Accordingly, *DMU D* is technically efficient but scale inefficient, while *DMU C* is neither technically efficient nor scale efficient. The scale efficiency of *C* is calculated as *YN*/*YV*. For more detailed exploration of returns to scale, readers are referred to Charnes, Cooper, and Rhodes (1978) and Banker, Charnes and Cooper (1984), among others.

The limitations of the non-parametric method derive mostly from the sensitivity of the results to sampling variability, to the quality of the data and to the presence of outliers. This has led recent literature to explore the relationship between statistical analysis and non-parametric methods (Simar and Wilson, 2000). Some solutions have been advanced. For instance, confidence intervals for the efficiency scores can be estimated using asymptotic theory in the single input case (for input efficiency estimators) or single output (in the output efficiency) case, given these are shown to be maximum likelihood estimators (Banker, 1993 and Goskpoff, 1996).

For multiple input/output cases the distribution of the efficiency estimators is unknown or quite complicated and analysts recommend constructing the empirical distribution of the scores by means of bootstrapping methods (Simar and Wilson, 2000). Other solutions to the outlier or noisy data consist in constructing a frontier that does not envelop all the data point, building an expected minimum input function or expected maximum output functions (Cazals, Florens and Simar, 2002, and Wheelock and Wilson, 2003). Another limitation of the method, at least in the context in which we will apply it, is the inadequate treatment of dynamics, given the lag between input consumption (public expenditure) and output production (health and education outcomes).

2.2 Overview of precursor papers

There is abundant literature measuring productive efficiency of diverse types of decision making units. For instance, there are papers measuring efficiency of museums (Bishop and Brand, 2003), container terminals (Cullinane and Song, 2003), electric generation plants (Cherchye and Post, 2001), banks (Wheelock and Wilson, 2003), schools (Worthington, 2001) and hospitals (Bergess and Wilson, 1998), among others. Few papers, however, analyze aggregate public sector spending efficiency using cross-country data. These are the direct precursors of this paper and are the focus of this section's survey.

Gupta and Verhoeven (2001) employ the input-oriented FDH approach to assess the efficiency of government spending on education and health in thirty-seven African countries in 1984-1995. Using several output indicators for health and education, they construct efficiency frontiers for each of the indicators and for each of the time periods they considered. That is, they used a single input/single output for each time period. They find that, on average, African countries are inefficient in providing education and health services relative to both Asian and the Western Hemisphere countries. They also report, however, an increase in the productivity of spending through time, as they document outward shifts in the efficiency frontier. Finally the authors report a negative relationship between the input efficiency scores and the level of public spending, which leads them to conclude that higher educational attainment and health output requires efficiency improvement more than increased budgetary allocations.

Evans and Tandon (2000) adopt a parametric approach to measure efficiency of national health systems for the World Health Organization, by estimating a fixed effects panel of 191 countries for the period 1993-97. Health output was measured by the disability-adjusted life expectancy (DALE) index, while health expenditures (public and private aggregated) and the average years of schooling of the adult population were considered as inputs. The output efficiency score is defined as the ratio of actual performance above the potential maximum. The authors also introduce the square of the inputs (average years of schooling and expenditure), arguing it's a second-order Taylor-series approximation to an unknown functional form. The fact that the quadratic terms are significant may be an indication of the

importance of non-linearity, but may also reflect neglected dynamics or heterogeneity in the sample (Haque, Pesaran and Sharma, 1999), given that both developed and developing nations were included. An interesting contribution of the paper is a construction of a confidence interval for the efficiency estimates through a Monte Carlo procedure. These authors document a positive relationship between their efficiency scores and the level of spending. The more efficient health systems are those of Oman, Chile and Costa Rica. The more inefficient countries are all African: Zimabawe, Zambia, Namibia, Botswana, Malawi and Lesotho.

Jarasuriya and Woodon (2002) also adopt a parametric approach to estimate efficiency of health and education provision in a sample of developing countries. The authors estimate the efficiency frontier by econometric methods. These authors consider separately an educational attainment indicator (net primary enrolment) and a health output indicator (life expectancy) and estimate a functional linear relationship between these output indicators and three inputs: per capita GDP, spending per capita, and the adult literacy rate. Using a panel of 76 countries for the period 1990 to 1998, they found no relationship between expenditure and the educational or health output variables when they include the per capita GDP. This led the authors to conclude that spending more is not a guarantee to obtain better education or health results. The authors do not point at the correlation between the two variables as a possible cause of this problem, which we discuss in the next section. The countries with the lowest efficiency in health indicators are all African (Malawi, Zambia, Mozambique, Ethiopia) as well as in education attainment (Ethiopia, Niger, Burkina Faso).

The authors go further by attempting to explain the cross-country variation in efficiency and find that the degree of urbanization and the quality of bureaucracy are the most relevant variables. To capture possible non-linearity, the authors introduce these variables squared. This stage of their work poses several problems. First, it is possible that the (non-linear) quadratic terms reflect heterogeneity across countries and dynamics across time. As shown by Haque, Pesaran and Sharma (1999), this would produce inconsistent estimates. Second, the authors do not adjust for the fact that the dependent variable (efficiency scores) is censored, given that it can adopt only values between zero and one. And third, the authors do not consider the serial correlation of the efficiency scores (Simar and Wilson, 2004).

Greene (2003a) combines the previous two papers in the sense he concentrated on health efficiency only using the WHO panel data and explained inefficiency scores variation across the sample of counties. Greene's stochastic frontier estimation is much more general and flexible, as it allows for time variation of the coefficients and heterogeneity in the countries' sensitivity to the explanatory variables. The author first estimates a health production function using expenditure (public and private together) and education as inputs, and then explains inefficiency with a set of explanatory variables of which the only significant ones are the income inequality measure, GDP per capita and a dummy variable for tropical location.

Afonso, Schuknecht and Tanzi (2003) examine the efficiency of public spending using a non-parametric approach. First, they construct composite indicators

of public sector performance for 23 OECD countries, using variables that capture quality of administrative functions, educational and health attainment, and the quality of infrastructure. Taking the performance indicator as the output, and total public spending as the input, they perform single input/single output FDH to rank the expenditure efficiency of the sample. Their results show that countries with small public sectors exhibit the highest overall performance.

Afonso and St. Aubyn (2004) address the efficiency of expenditure in education and health for a sample of OECD countries applying both DEA and FDH. This paper presents detailed results by comparing input-oriented and output-oriented efficiency measurements. The small overlap of the samples limits the comparability of these results with those presented in the next section. An apparently strange result, reported in earlier drafts of the paper, was the inclusion of Mexico as one of the benchmark countries (on the efficiency frontier). The result is strange given that the sample is the OECD countries, and it counterintuitive. This is the result of Mexico having very low spending and low education attainment results, hence it can be considered as the "origin" of the efficiency frontier. The next chapter discusses this topic and reports similar counterintuitive results but for other countries.

3. Empirical results

3.1 Input and output indicators: description, assumptions and limitations

Cross-country comparisons assume some homogeneity across the world in the production technology of health and education. There are two particular aspects in which the homogeneity assumption is important. First, the comparison assumes that there is a small number of factors of production that are the same across countries. Any omission of an important factor will yield as a result a high efficiency ranking of the country that uses more of the omitted input. Second, the comparison requires that the quality of the inputs is more or less the same, with the efficiency scores biased in favor of countries where the quality is of higher grade.

Factor heterogeneity will not be a problem, as long as it is evenly distributed across countries. It will be problematic if there are differences between countries in the average quality of a factor (Farrell, 1957). The exercise that we present suffers from this limitation, given that the main input in both production technologies is used more intensively in richer countries (with higher per capita GDP). The main input is public spending per capita on education and health measured in constant 1995 US dollars in PPP terms. A clear positive association between this variable and per capita GDP can be verified (Figures 4 and 5).

This positive association between expenditure and the level of economic development (as measured by per capita GDP) may be explained by several reasons. One of them could be the Balassa-Samuelson effect, according to which price levels

in wealthier countries tend to be higher than in poorer countries.[1] This applies to both final goods and factor prices. Thus price of the same service (health or education, for instance) will be higher in the country with higher GDP. Similarly, wages in the relatively richer counties are higher, given the higher marginal productivity of labor, which will tend to increase costs, especially in labor-intensive activities as health and education.

Figure 4 can be interpreted as evidence of the validity of Wagner's hypothesis at the cross-country level. This hypothesis, postulates that there is a tendency for governments to increase their activities as economic activity increases. Since 1890 Wagner postulated that economic development implied rising complexities that required more governmental activity, or that the elasticity of demand for publicly provided services, in particular education was greater than one. This hypothesis has been tested econometrically (Chang, 2002) in time series and cross-country settings, showing that this is nothing particular of the series used for the present study.

Previous studies that measured the efficiency of public spending recognized the positive association and suggested alternative solutions. One possibility is to split the sample by groups of countries (Gupta and Verhoeven, 2001). We follow this approach by excluding the industrialized nations from the sample, and by presenting most of the results clustered regionally (Africa (AFR), East Asia and Pacific (EAP), Latin America and Caribbean (LAC), Middle East and North Africa (MNA) and South Asia (SAS)). A second alternative incorporates directly the per capita GDP as a factor of production, jointly with expenditure and other inputs (Jarasuriya and Woodon, 2002). The problem with this approach is that it combines variables derived from a production function approach, and hence with clear interpretation, with others (GDP per capita) that are difficult to interpret from any viewpoint. When the two types of variables are combined, their effects cannot be disentangled.

A third option consists in using as an input the orthogonal component of public expenditure to GDP.[2] We scored the efficiency using as input both the original expenditure variable and the orthogonalized variable. The goodness-of-fit of each model was gauged based on the frequency distribution of the inefficiency measures, as suggested by Farrell (1957) and Varian (1990). Comparing the efficiency distributions (Figure 5) it is clear that the orthogonalized expenditure version produces distributions that are not skewed towards extreme inefficient outcomes. On this basis, the paper considered the orthogonal component of expenditure on health and education.

---

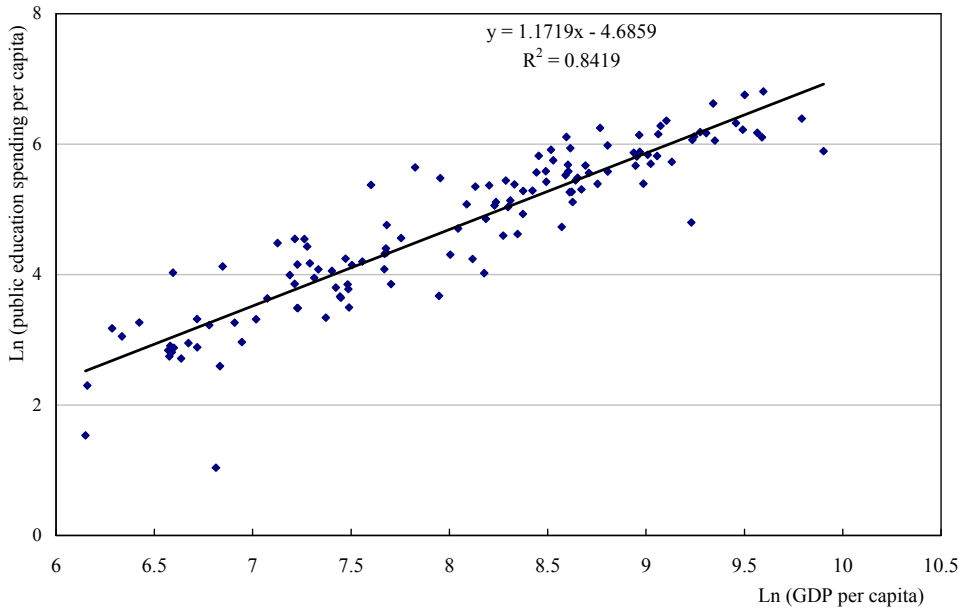
[1] The Balassa-Samuelson effect refers to the fact that price levels are higher in richer countries than in poorer countries. It can be shown that relative wages and relative prices are a function of the marginal productivity of labor in the traded goods. Given higher capital abundance in the richer countries, the productivity of labor tends to be higher in these countries, and hence will be wages and prices.

[2] The orthogonalized expenditure variable is the residual of the linear regression between pubic expenditure and GDP per capita. Since residuals may take positive and negative values, the variable was right-shifted to avoid negative values to facilitate graphical presentation of the frontiers.

Figure 4

Public Expenditure and GDP (Both per capita and in Logs)

Education Spending vs. GDP per capita



Health Spending vs. GDP per capita

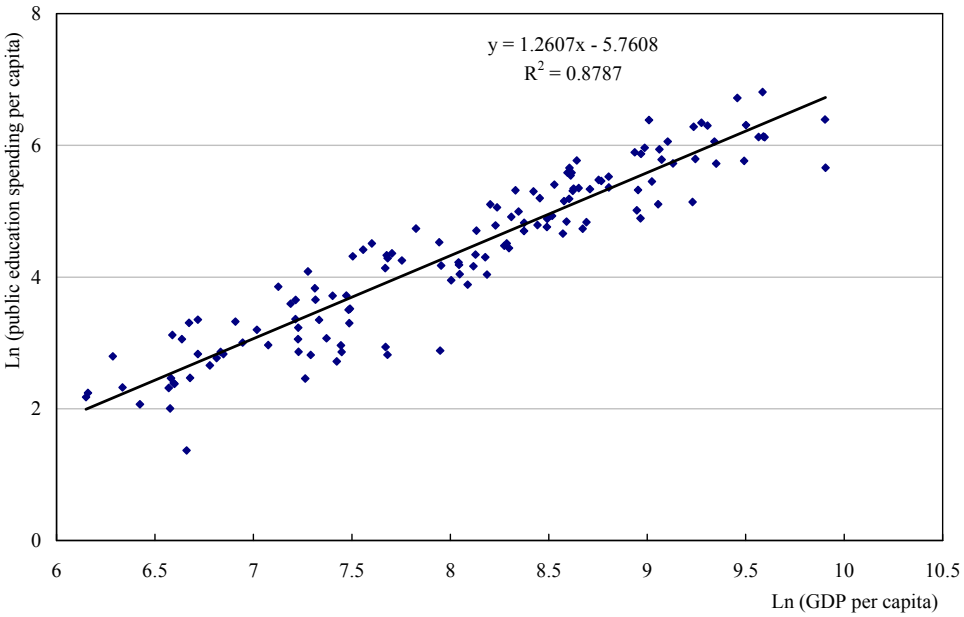
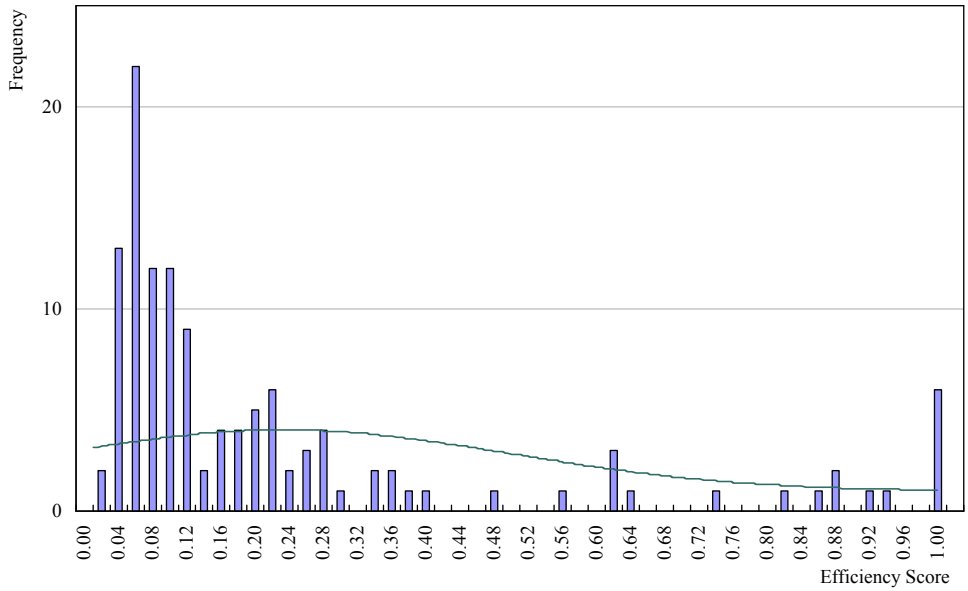


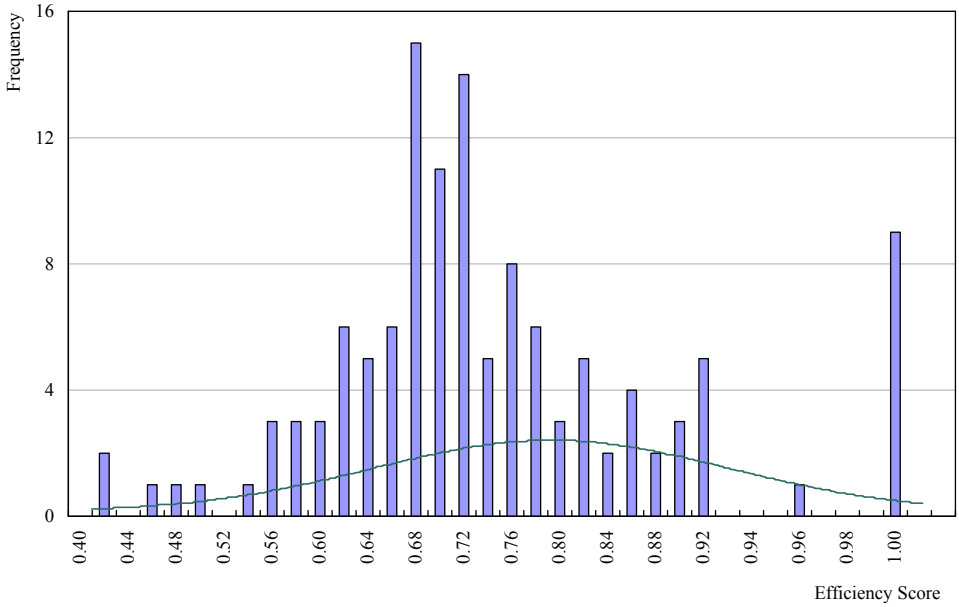
Figure 5

Density of Efficiency Scores – Gross Primary School Enrolment

Unorthogonalized Public Expenditure



Orthogonalized Public Expenditure



This paper uses nine indicators of education output and four indicators of health output.[3] The education indicators are: primary school enrolment (gross and net), secondary school enrolment (gross and net), literacy of youth, average years of school, first level complete, second level complete, and learning scores. Though the ideal educational output indicator are comparable learning scores, international assessments are based on samples mostly composed of developed nations, limiting the applicability to the present paper. However, Crouch and Fasih (2004) recently combined several international assessments to obtain a larger sample of comparable results.[4] Unfortunately they only do it for one period. The correlation between the learning scores and other output variables is high (.81 with net secondary school enrolment and .76 with average years of school), as shown in Figure 6.[5] The health output indicators are: life expectancy at birth, immunization (DPT[6] and measles), and the disability-adjusted life expectancy (DALE).

The cross-country comparisons with this set of indicators assume some form of data homogeneity, which might be problematic given the diversity of counties in the sample considered. Even for a more homogeneous group of countries, such as the OECD, there is call for caution when comparing expenditure levels in member countries (Jounard *et al.*, 2003). There is very little to do to overcome this limitation, except subdivide the sample into different groups. Probably a regional aggregation can be useful, but even at that level there may be extreme heterogeneity.

Other four limitations of the analysis arising from the particular data sources are: first, the level of aggregation. The paper uses aggregate public spending on health and education, while using disaggregate measures of output, such as. primary enrolment or secondary enrolment. Ideally, the input should be use separately public spending in primary and secondary education. Similarly, health care spending could be disaggregated into primary care level care and secondary level. The data can be disaggregated even further, by analyzing efficiency at the school or hospital levels. Second, there are omitted factors of production. This is especially true in education, as the paper did not consider private spending due to data constraints for developing nations. If this factor were used more intensively in a particular group of countries, then the efficiency scores (reported in the next section) would be biased favoring efficiency in that group.

---

[3] The data sources are: the World Bank World Development Indicators (WDI), Barro-Lee database, Crouch and Fasih (2004), and the World Health Organization (Mathers *et al.,* 2000).

[4] Crouch and Fasih (2004) consider several international tests of learning achievement in math, science and literacy applied at different levels of the school system. The tests are the following: TIMSS (Third International Mathematics and Science Survey), PIRLS (Progress in International Literacy Study), PISA (Program for International Student Assessment), Reading Literacy Study, LLECE (Laboratorio Latinoamericano de Evaluacion de la Calidad de la Educacion), SACMEQ (Southern Africa Consortium for Monitoring of Education Quality), MLA (Monitoring Learning Achievement). Since the tests have different samples, they converted all test scores through iterative comparisons to a single numeraire.

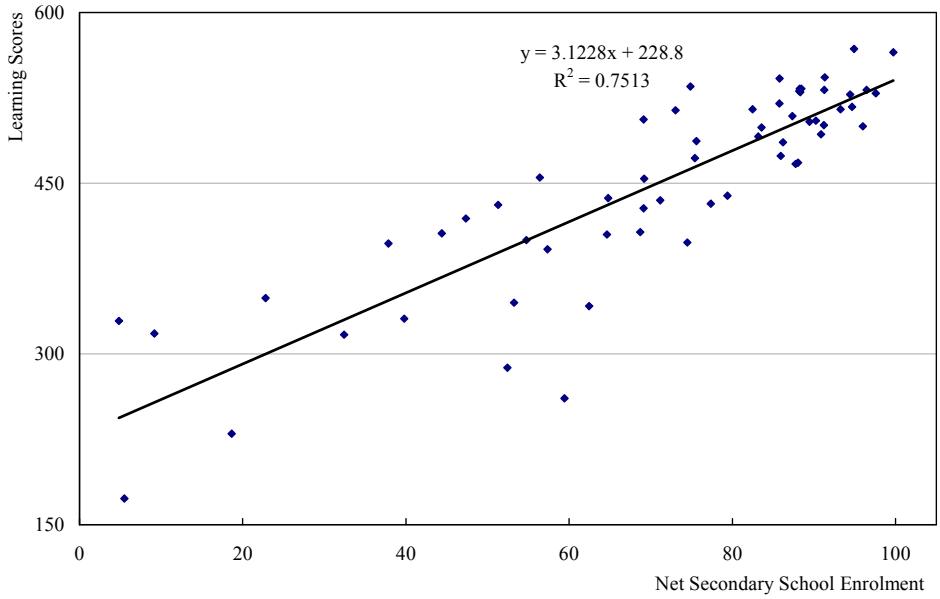
[5] The correlation coefficients and Figure 6 exclude developed nations for the Crouch and Fasih (2004) sample.

[6] DPT is Diphtheria-Pertussis and Tetanus.

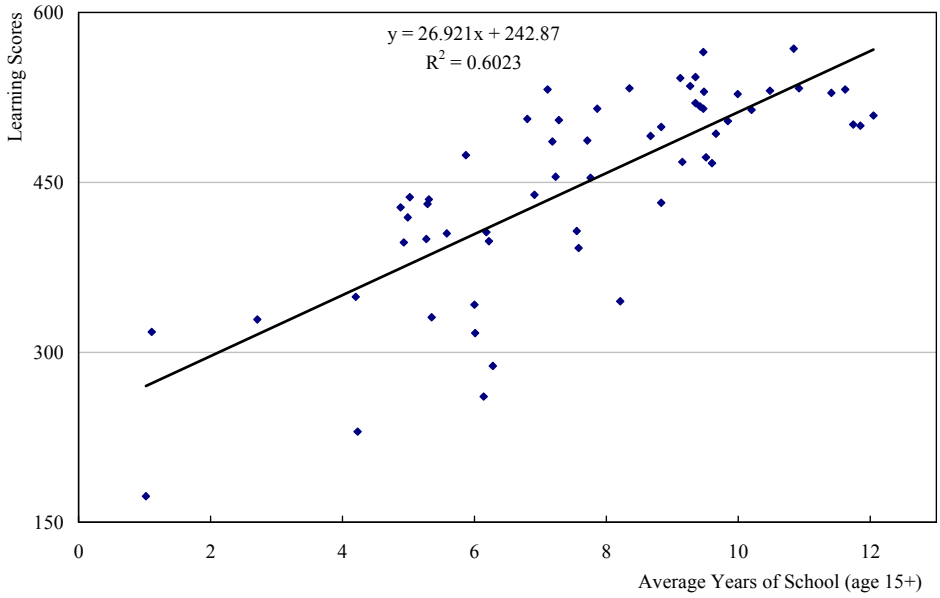
Figure 6

Correlation between Learnings Scores and Other Education Indicators

Correlation between Learning Scores and Net Secondary Enrolment



Correlation between Learning Scores and Average Years of School



Source for the figures on this page: World Bank WDI and Crouch and Fasih (2004).

The third limitation arising from the data is the combination of monetary and non-monetary factors of production. The paper uses together with public expenditure, other non-monetary factors of production such as the ratio of teachers to students, in the case of education, or literacy of adults in the case of health and education. Other factors of production that could have been used were the physical number of teaching hours (in education) or the number of doctors or in-patient beds, as Afonso and St. Aubyn did for the OECD countries. However, inexistent data for a large number of developing countries constrained the options. Fourth, data availability constrained a better differentiation between outputs and outcomes. For instance, most of the indicators of education, such as completion and enrolment rates do not measure how much learning is taking place in a particular country. In education, this paper advances by considering the learning scores as one of the indicators. In health, other outcomes such as the number of sick-day leaves or the number of missed-school days because of health-related causes could be better reflections of outcomes.

3.2 Single input/output results

3.2.1 FDH and DEA analysis: education

Figures 7a-c show both FDH and DEA estimation of the efficiency frontier for three of the nine output indicators: gross primary school enrolment, first level complete and learning scores.

Figure 7d illustrates the efficiency frontier for the learning scores if the developed countries are included in the sample, demonstrating the sensitivity of the results to the sample definition. This fact is particularly acute in the case of learning scores which capture the quality of education dimension that no other indicator captures. While in the sample of developing countries Chile, Hungary and the Czech Republic are on the frontier; once the developed nations are included they appear as inefficient.[7]

Several results may be highlighted:

a) In general, the rankings are robust to the output indicator selected. This can be can be verified by the Spearman rank-correlation coefficient: all are positive, significant and high. The range oscillates from a minimum of .53 to a maximum of .94, with the mean of .70. This result implies that countries appearing as efficient (or inefficient) according to one indicator, are ranked similarly when other output indicator is used.

b) Despite the orthogonalization by GDP, the relatively rich countries tend to be in the less efficient group, *i.e.* countries with higher per capita GDP spend more than other countries in attaining similar education outcomes. Higher spending

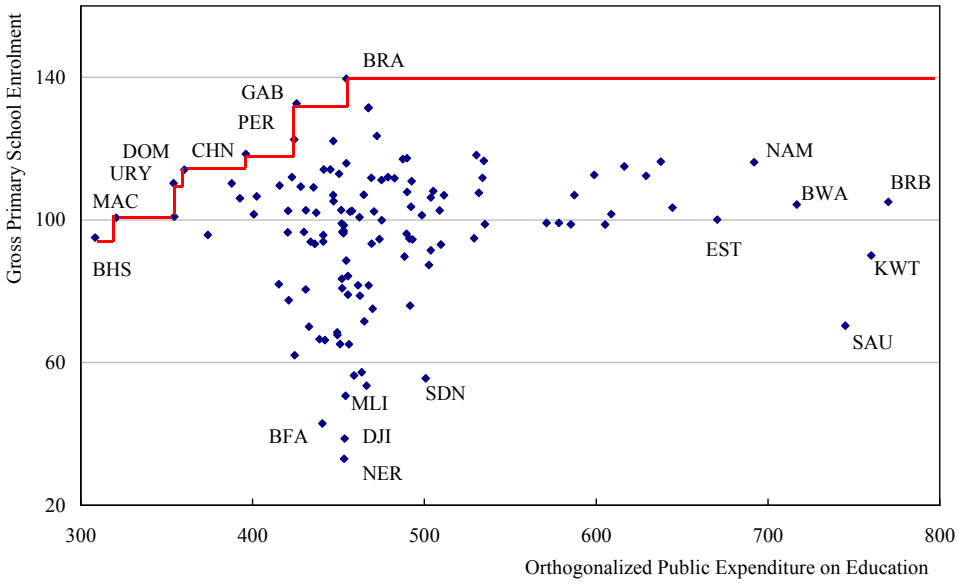
---

[7] The frontier depicted in Figure 7d excludes Japan, Korea, Ireland and Belgium to facilitate comparisons with the frontier without developed nations.

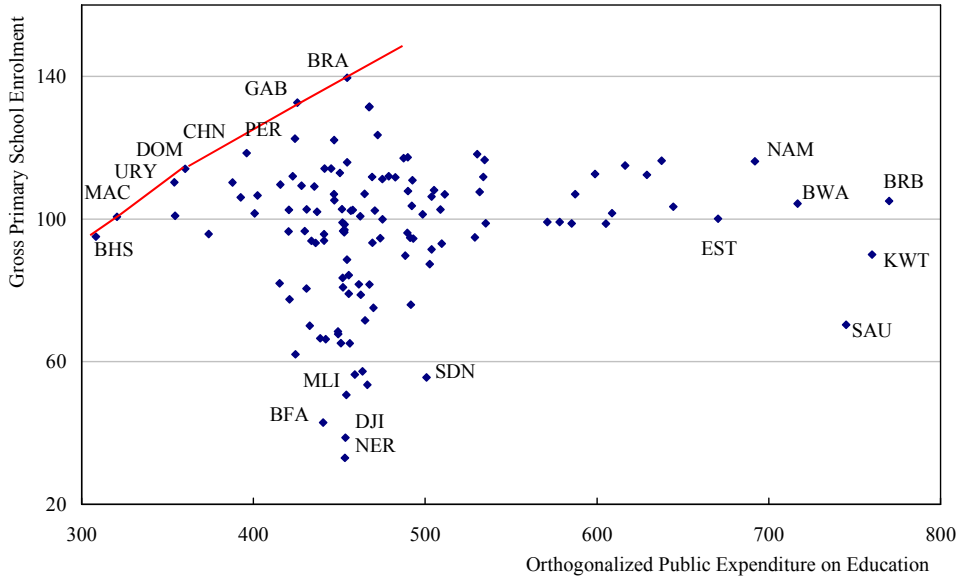
Figure 7

Education Efficiency Frontier: Single Input and Single Output

a.1: Gross Primary School Enrolment vs. Education Expenditure (Free Disposable Hull, FDH)



a.2: Gross Primary School Enrolment vs. Education Expenditure (Data Envelopment Analysis, DEA)

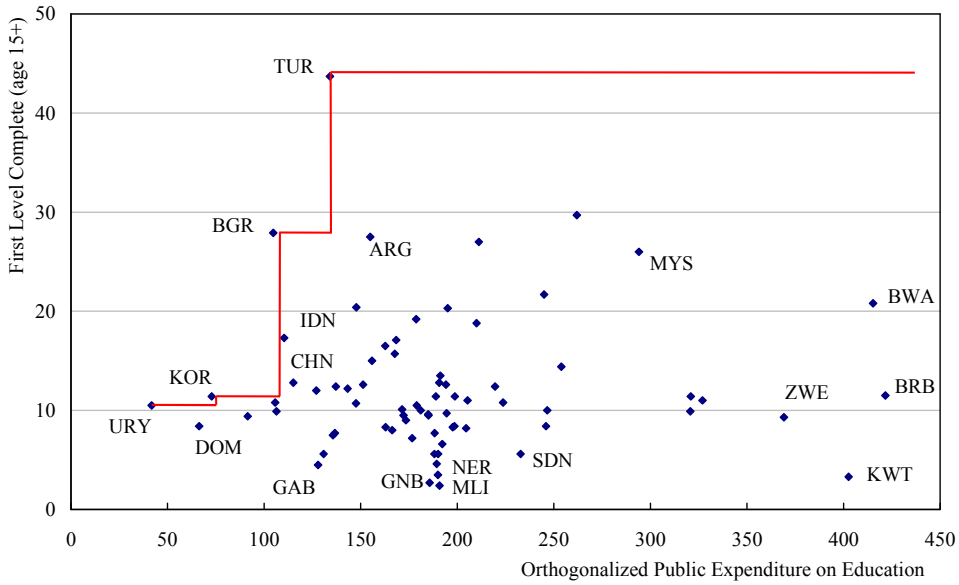


Source for the figures on this page: World Bank WDI.

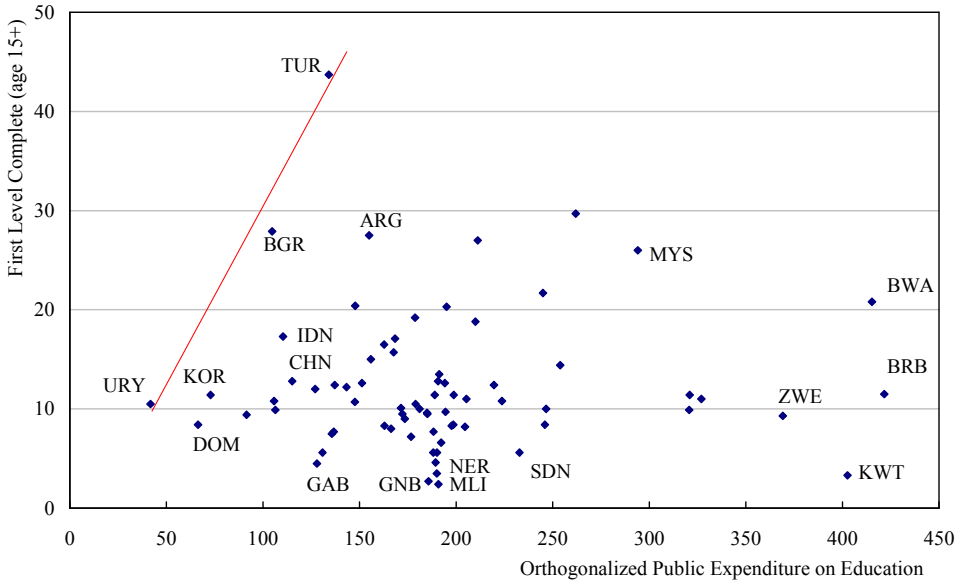
Figure 7 (continued)

Education Efficiency Frontier: Single Input and Single Output

b.1: First Level Complete vs. Education Expenditure (Free Disposable Hull, FDH)



b.2: First Level Complete vs. Education Expenditure (Data Envelopment Analysis, DEA)

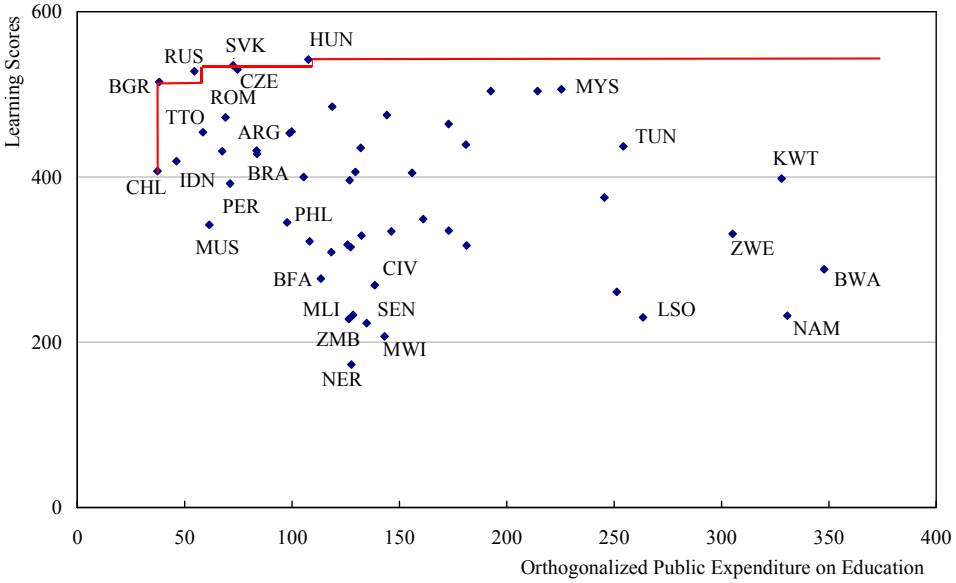


Source for the figures on this page: World Bank WDI, Barro-Lee database.

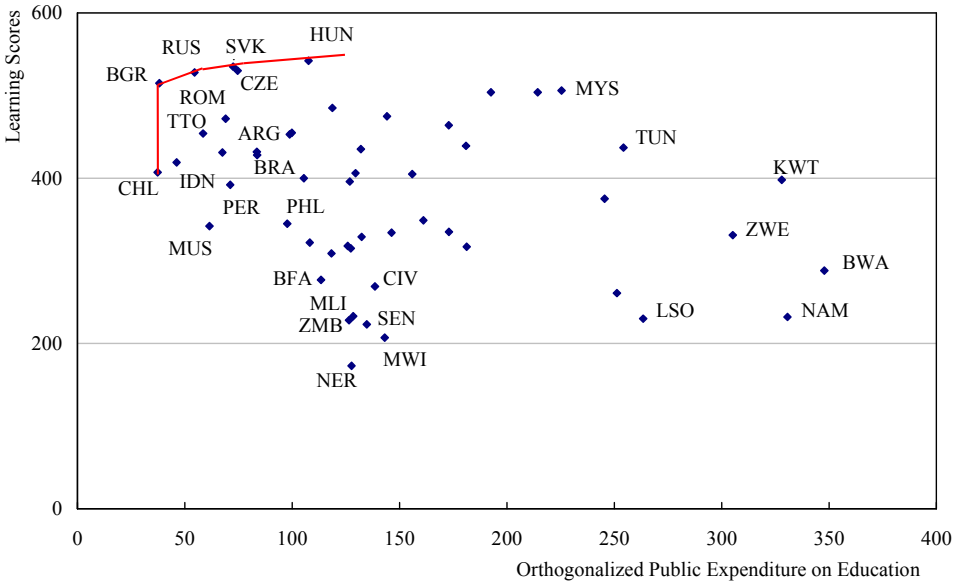
Figure 7 (continued)

Education Efficiency Frontier: Single Input and Single Output

c.1: Learning Scores vs. Education Expenditure (Free Disposable Hull, FDH)



c.2: Learning Scores vs. Education Expenditure (Data Envelopment Analysis, DEA)

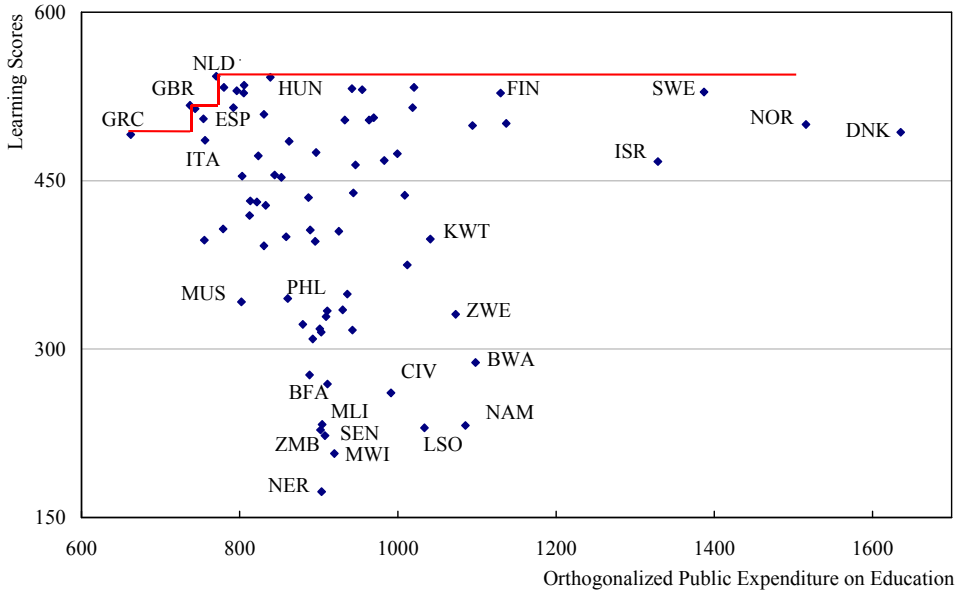


Source for the figures on this page: World Bank WDI and Crouch and Fasih (2004).

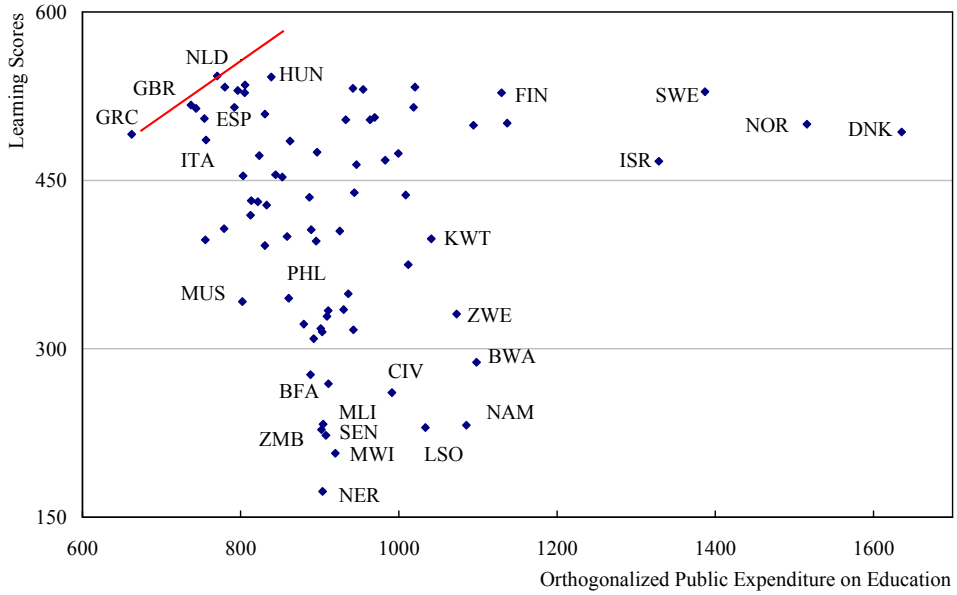
Figure 7 (continued)

Education Efficiency Frontier: Single Input and Single Output

d.1: Learning Scores vs. Education Expenditure (Free Disposable Hull, FDH)



d.2: Learning Scores vs. Education Expenditure (Data Envelopment Analysis, DEA)



Source for the figures on this page: World Bank WDI and Crouch and Fasih (2004).

may reflect the higher cost of tertiary education. This is one factor that may help explain the stand-out of Estonia, Latvia, and Poland. Oil-rich countries, such as Kuwait and Saudi Arabia, tend to be in the group of relatively more inefficient producers.

c) Another group of relatively inefficient producers are those with "average" expenditure levels but extremely low education attainment. Among those are mostly African counties (Angola, Niger, Burkina Faso, Sudan, and Ethiopia), some Middle Eastern countries (Djibouti, and Yemen) and South Asia (Bangladesh and Pakistan).

d) Output efficiency rankings also vary with the selected output indicators. The Spearman correlation coefficient of the output efficiency scores shows that these are robust to the selected indicator, though the mean of the correlation coefficients is lower (.52) and the range is somewhat higher (.30 to .95) than those registered in the input efficiency rankings.

e) In an attempt to identify clusters of more efficient countries and more inefficient countries, the top (and bottom) 10 per cent of the efficiency ranking were selected for each of the indicators. If a country appeared in the efficient (inefficient) tail in three or more of the indicators, it was included in Table 1.

f) This clustering exercise reveals (Table 1) a group of African countries as the most inefficient. Two oil-rich countries are included in this group as well. Among the more efficient group of countries we consistently find Uruguay, Korea, Bahamas, and Bahrain. Explaining why these particular sets of countries appear in each cluster requires more in-depth analysis. The last section of this paper attempts to associate efficiency results with some explanatory variables.

g) To grasp the order of magnitudes of the deviations from the efficiency frontier, we computed an average for all indicators for the inefficient countries. The input efficiency estimations indicate that the most inefficient decile could reach the same educational attainment levels by spending approximately 50 per cent less. The output efficiency estimators indicate that, on average, with the expenditure level this group could reach educational attainment levels four times as high.

h) It is critical to note that even if a country appears as efficient, there might still be a significant discrepancy between the observed output level and the desired or target output level. For instance, Bahamas, Bahrain, Dominican Republic and Guatemala appear as efficient countries on the efficiency frontier or very close to it (Figure 7 a.1). However, these countries are still far away from where Gabon or Brazil are, and could consider desirable to achieve those target enrolment rates. Both Guatemala and Dominican Republic spend 2 per cent of GDP on education but have (net) secondary enrolment rates below 40 per cent. And net primary enrolment is about 80 per cent. It would be difficult to argue that that is a desirable outcome, though it is an efficient one. Similarly, though Chile appears as efficient with learning scores of about 400, the country could still achieve higher learning scores of over 500 points at the cost of additional public spending. The important thing is that the country moves along the efficiency frontier to the higher target output level. Countries can even improve efficiency

Table 1

Education Attainment – Single Input/Single Output

	Input-Efficient	**Output-Efficient**
More efficient	Uruguay, Korea, Dominican Republic, Indonesia, Guatemala, China, Bahamas, Bahrain, El Salvador	Uruguay, Korea, Bahrain,
Least efficient	Botswana, South Africa, Kuwait, Tunisia, Lesotho, Barbados, Saudi Arabia, Zimbawe, Namibia, Malaysia, St. Lucia, Jamaica, St. Vincent, Latvia	Niger, Mali, Tanzania, Burkina Faso, Guinea-Bissau, Ethiopia, Guinea, Burundi, Sudan, Sierra Leone, Chad

by exploiting scale economies if they are operating in the increasing returns to scale zone of the production possibility frontier (output levels smaller than that of point A, Figure 3).

i) The regional aggregation of the efficiency scores by each individual output indicator shows that scores are lower when they are input oriented (Table 2) than when they are output oriented (Table 3).[8] This is especially true for ECA. In general, we observe higher efficiency scores when primary enrolment is considered as the output indicator. Scores are lower for secondary enrolment, especially when output-oriented measures are considered. Africa and MNA have similar levels of input-inefficiency: in most cases, both regions use public spending in excess of 35 per cent than the benchmark cases. EAP, ECA, LAC and SAS spend in excess between 20-30 per cent of the benchmark level. The output efficiency scores are lower in Africa.

3.2.2 FDH and DEA analysis: health

This section considers the case of one input (public expenditure on health per capita in PPP terms) and four alternative output indicators: life expectancy at birth, DPT immunization, measles immunization, and the disability-adjusted life expectancy (DALE) index which takes into account both mortality and illness. The efficiency frontiers for each indicator are computed using both the FDH and DEA methodologies. Figures 8 a-d show the efficiency frontier for one indicator.

---

[8] The regional aggregation is for illustrative purposes only and was computed as the simple average of the individual country scores obtained for the whole sample. The scores were not computed by constructing separate efficiency frontiers for each region. Hence, they do not reflect the heterogeneity in the individual country scores and possibly do not reflect adequately variations across regions.

Table 2

Educational Attainment: Input Efficiency Scores by Regions across the World Single Input/Single Output

	AFR	EAP	ECA	LAC	MNA	SAS
Gross Primary Enrolment	.69	.74	.67	.74	.65	.75
Net Primary Enrolment	.68	.78	.72	.77	.68	.71
Gross Secondary Enrolment	.65	.69	.67	.69	.63	.70
Net Secondary Enrolment	.64	.71	.71	.69	.64	.72
Average Years of School	.21	.36	.37	.32	.18	.25
First Level Complete	.21	.43	.48	.36	.20	.26
Second Level Complete	.22	.37	.33	.32	.19	.27
Literacy of Youth	.66	.73	.86	.72	.63	.72

Table 3

Educational Attainment: Output Efficiency Scores by Regions across the World Single Input/Single Output

	AFR	EAP	ECA	LAC	MNA	SAS
Gross Primary Enrolment	.62	.79	.72	.82	.67	.72
Net Primary Enrolment	.64	.93	.90	.93	.79	.78
Gross Secondary Enrolment	.23	.50	.70	.61	.54	.39
Net secondary Enrolment	.26	.58	.84	.66	.60	.44
Average Years of School	.32	.63	.79	.60	.53	.38
First Level Complete	.19	.49	.50	.36	.22	.20
Second Level Complete	.09	.37	.38	.24	.26	.22
Literacy of Youth	.72	.95	.99	.94	.88	.66

Several results may be highlighted:

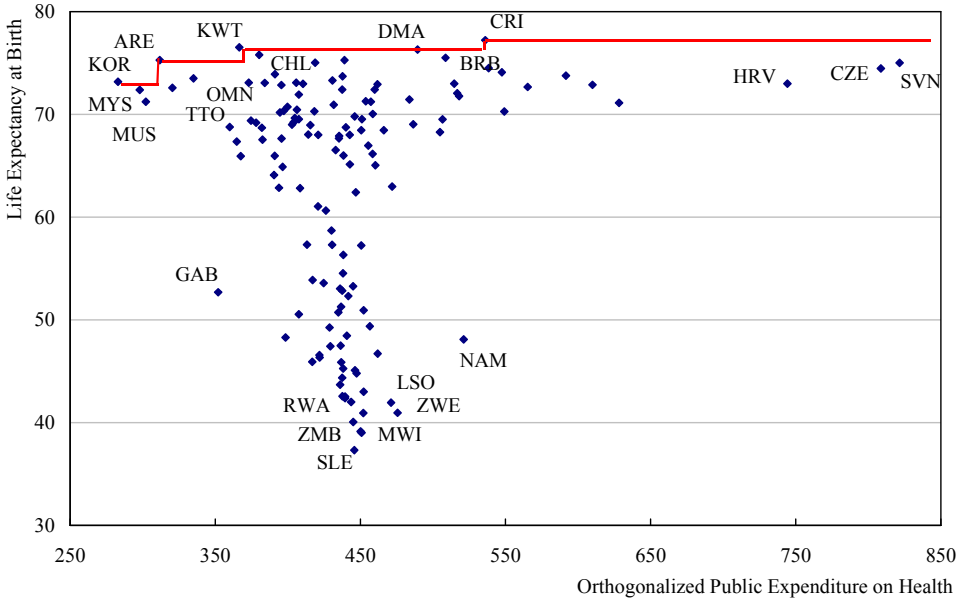
a) The input efficiency scores obtained for each of the output indicators are highly correlated. The Spearman rank-order correlation coefficient oscillates between .66 and .94, with a mean of 0.81. This indicates that the efficiency ranking is very similar regardless of the output indicator being used.

b) Despite the orthogonalization by GDP the relatively rich countries tend to be in the less efficient group. The group of inefficient producers tend to concentrate in

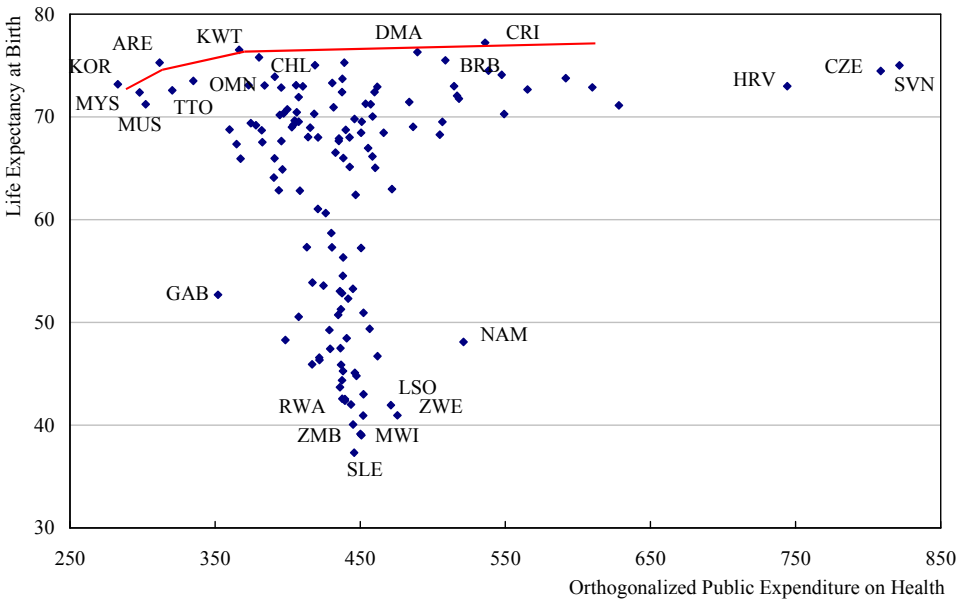
Figure 8

Health Efficiency Frontier – Single Input/Single Output

a.1: Life Expectancy vs. Health Expenditure (Free Disposable Hull, FDH)



a.2: Life Expectancy vs. Health Expenditure (Data Envelopment Analysis, DEA)

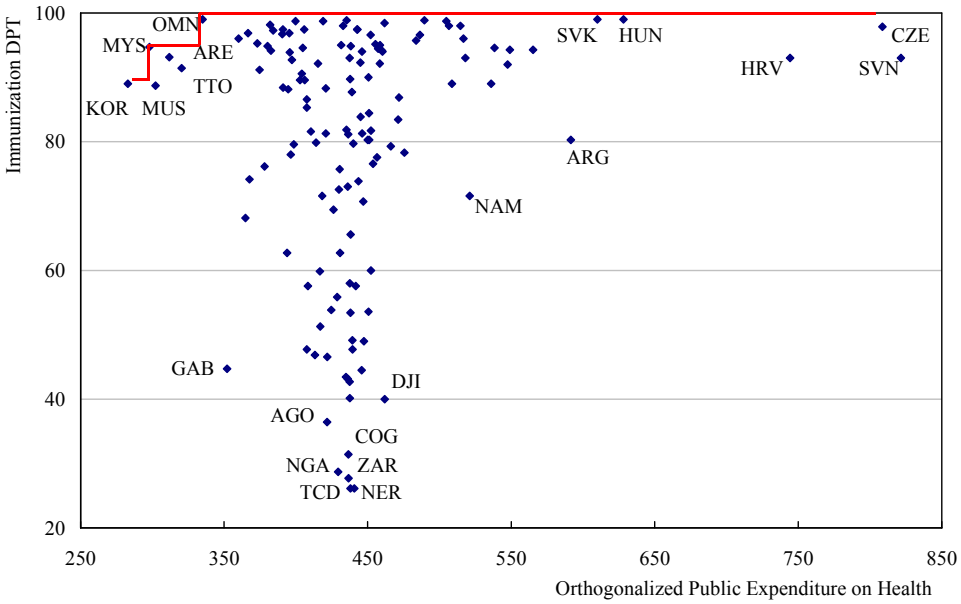


Source for the figures on this page: World Bank WDI.

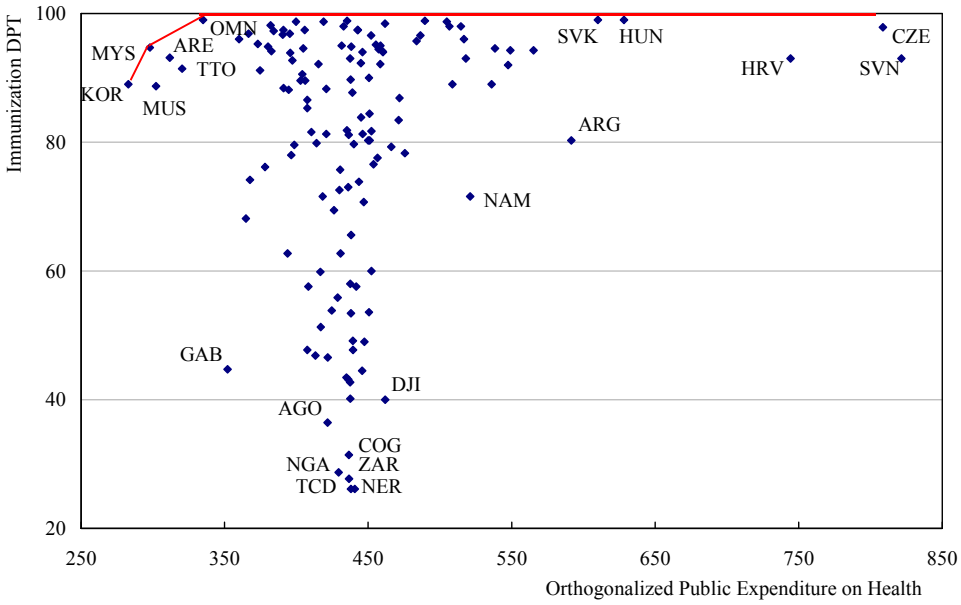
Figure 8 (continued)

Health Efficiency Frontier – Single Input/Single Output

b.1: Immunization DPT vs. Health Expenditure (Free Disposable Hull, FDH)



b.2: Immunization DPT vs. Health Expenditure (Data Envelopment Analysis, DEA)

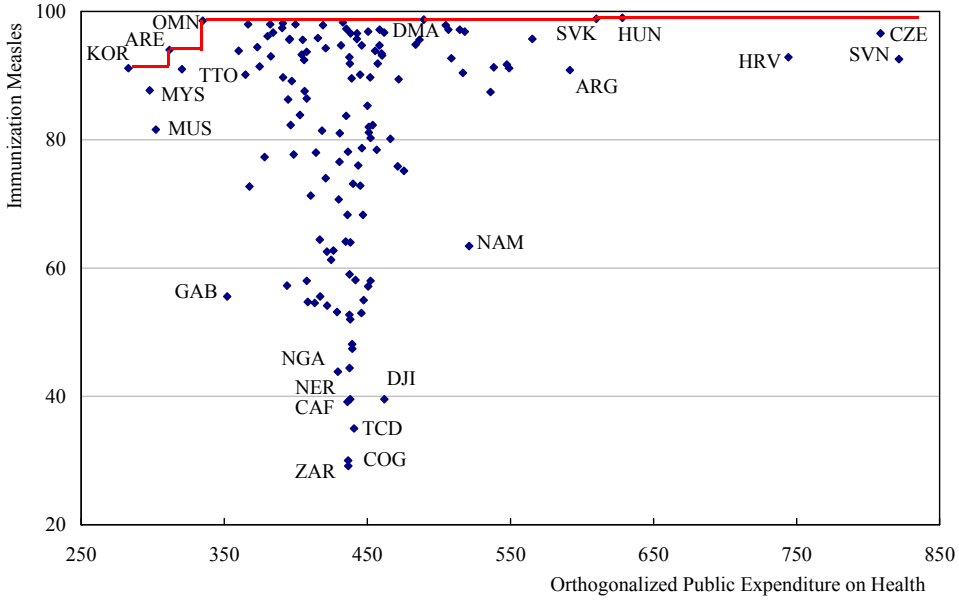


Source for the figures on this page: World Bank WDI.

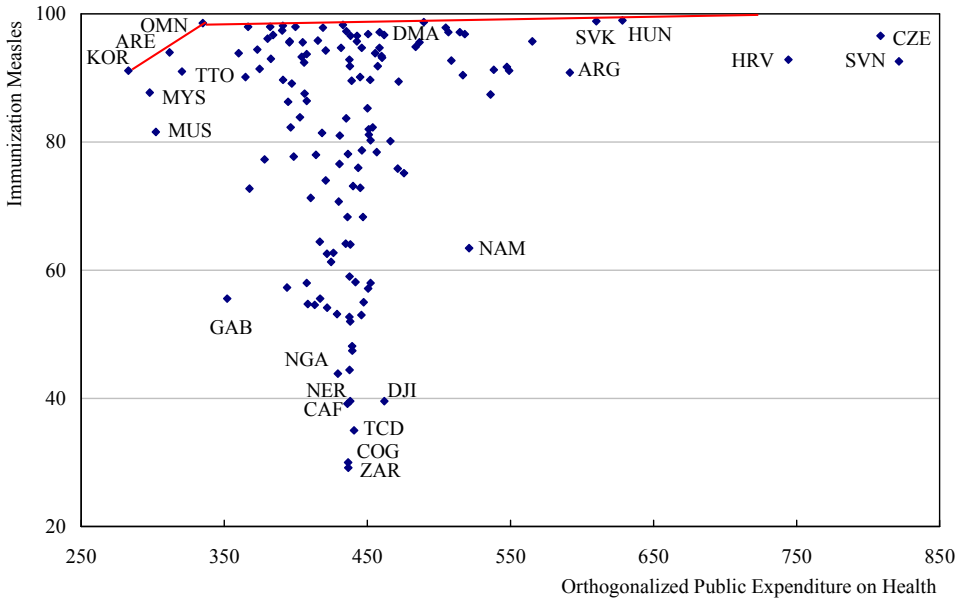
Figure 8 (continued)

Health Efficiency Frontier – Single Input/Single Output

c.1: Immunization Measles vs. Health Expenditure (Free Disposable Hull, FDH)



c.2: Immunization Measles vs. Health Expenditure (Data Envelopment Analysis, DEA)

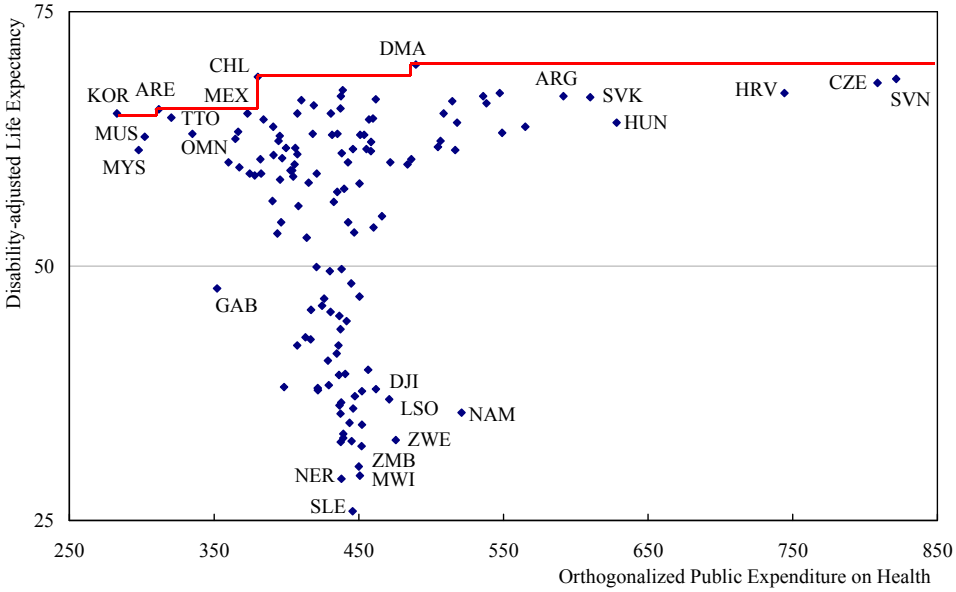


Source for the figures on this page: World Bank WDI.

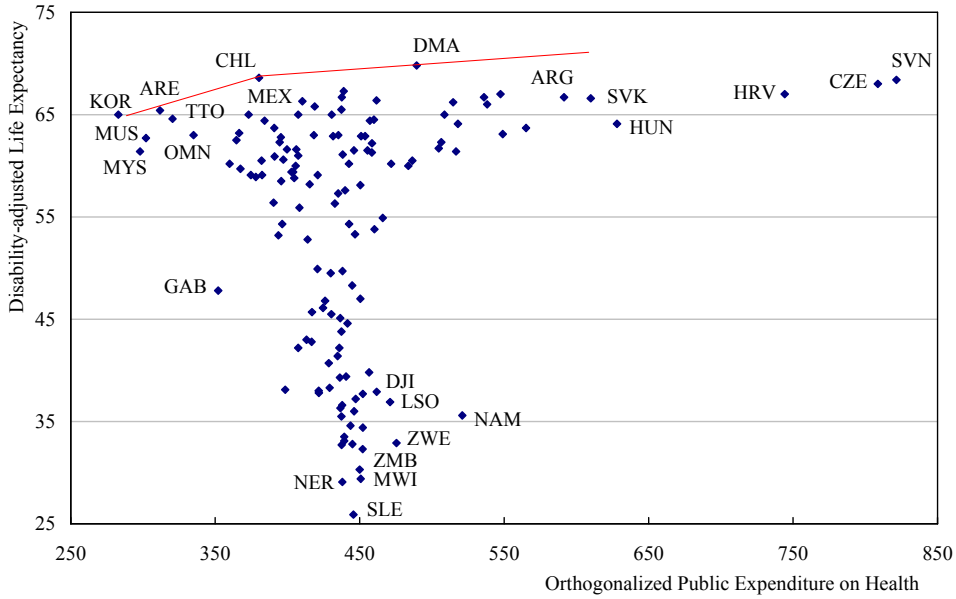
Figure 8 (continued)

Health Efficiency Frontier – Single Input/Single Output

d.1: DALE vs. Health Expenditure (Free Disposable Hull, FDH)



d.2: DALE vs. Health Expenditure (Data Envelopment Analysis, DEA)



Source for the figures on this page: World Bank WDI and Mathers *et al.* (2000).

two groups of countries: one group of relatively rich countries like the Czech Republic, Croatia, Slovenia, and Hungary that have big expenditure levels and not extremely high output (input inefficiency) and other group of countries that spend relatively little but their output indicators could be substantially larger, like Sierra Leone, Namibia, Zimbawe, and Lesotho.

c) To capture this difference, it is convenient to examine the output efficiency scoring. The rankings between input and output orientations are highly correlated.

d) With the four output indicators deciles, more efficient and least efficient countries are listed in Table 4. The group of least efficient countries could, on average, increase output significantly for a given expenditure level. For instance, the decile of most inefficient countries could almost double the disability-adjusted life expectancy (DALE) index to achieve the same efficiency as the benchmark. Similarly the DPT immunization would have to triple to achieve the same efficiency level than the benchmark developing countries.

e) The regional aggregation of the efficiency scores, by each individual output indicator shows that input efficiency scores (Table 5) are lower than output efficiency scores (Table 6). This is especially true in ECA, LAC and MNA, and to a lesser extent in EAP and SAS. In Africa, both scores are strikingly similar, indicating that, on average, the region spend about 35 per cent in excess of the benchmark cases to achieve the same output level. Alternatively, the output level is 35 per cent below comparable efficient countries that use the same input (expenditure) level.

3.3 Multiple inputs and multiple outputs

Both education and health attainment are not solely determined by public spending. Other inputs, such as private spending also affect the output indicators. For health, the World Bank WDI database reports a comparable statistic across countries. Unfortunately, a comprehensive database of this variable does not exist for education: for the education production technology we have multiple indicators of educational attainment, and three inputs (public spending, teachers per pupil, and adult literacy rate). In health, besides public spending, two other inputs were included: private spending and the education level of adults. The analysis was limited to include up to three outputs. Too many output indicators will complicate the analysis, biasing efficiency scores towards one, increasing the variance of the estimators, and reducing their speed of convergence to the true efficiency estimators (Simar and Wilson, 2000; Groskopff, 1996).

In education, the selected input-output combinations produce rankings that are somewhat similar: the average rank correlation coefficient is .53 The frequency distribution of the efficiency estimators is similar in all the models, and as the model shifts from a basic two-input two-output model to a more complex three-input/three-output model, the frequency distribution shifts to the right, that is, more concentrated around more efficient results.

Table 4

Health Attainment – Single Input/Single Output

	Input-Efficient	**Output-Efficient**
More efficient	Korea, Malaysia, Thailand, Trinidad and Tobago, Oman, United Arab Emirates, Mauritius, Kuwait, Chile	Korea, Dominica, Oman, United Arab Emirates, Anigua and Barbuda
Least efficient	Argentina, Estonia, Czech Republic, Slovenia, Macedonia, Croatia, Namibia, Tunisia, Latvia, Hungary, Barbados	Sierra Leone, Ethipia, Burkina Fasso, Central African Republic, Mali

Table 5

Health Attainment: Input Efficiency Scores by Regions across the World Single Input/Single Output

	AFR	**EAP**	**ECA**	**LAC**	**MNA**	**SAS**
Life Expectancy at Birth	.65	.72	.58	.69	.73	.69
Immunization DPT	.66	.73	.63	.68	.76	.71
Immunization Measles	.65	.73	.67	.69	.76	.71
DALE	.65	.72	.60	.70	.71	.69

Table 6

Health Attainment: Output Efficiency Scores by Regions across the World Single Input/Single Output

	AFR	**EAP**	**ECA**	**LAC**	**MNA**	**SAS**
Life Expectancy at Birth	.63	.87	.91	.92	.90	.83
Immunization DPT	.62	.83	.95	.87	.90	.75
Immunization Measles	.63	.83	.95	.91	.90	.71
DALE	.56	.83	.90	.90	.86	.79

Table 7

Educational Attainment – Multiple Inputs/Multiple Outputs

	Input-Efficient	**Output-Efficient**
More efficient	Bangladesh, Bahrain, Dominican Republic, Argentina, Estonia	Argentina, Bangladesh, Chile, Brazil, Bahrain, Dominican Republic, Congo
Least efficient	Zimbabwe, Lesotho, Botswana, Costa Rica, Swaziland, Saudi Arabia, Malaysia	Sudan, Ghana, Tanzania, Ethiopia, Kenya, Niger

The multi-input output model results (Table 7) in general confirm the results of Table 1. Some new countries that appear as efficient are Bangladesh, Congo and Argentina. In the case of Bangladesh and Congo, this is the result of considering literacy of adults as a factor of production, that in these countries is low, and hence, appearing as very efficient. Congo has also extremely low ratio of teachers per student, the other factor of production, reinforcing the bias towards the efficient score. Within the least efficient countries, the models point at Zimbabwe, Lesotho, Botswana, Malaysia, and Saudi Arabia as the single-input models. In addition, Costa Rica and Swaziland appear as input-inefficient.

The regional aggregation for input and output efficiency scores using the multiple input-output framework show (Tables 8 and 9) that as the model becomes more complex (adding inputs or outputs), scores tend to show more efficient regions. The input efficiency regional aggregation allows several interesting comparisons across the regions on the impact of an additional input on the efficiency scores. For instance, the first two rows of Table 8 allow examination of the impact of adding literacy of adults as an additional input. The biggest impact is in the MNA region, followed by ECA and LAC, while in the others the increase in efficiency scores is more marginal.[9] Output efficiency scores change substantially in MNA and Africa.

Rows 4 and 5 of Table 8 allow comparing the impact of adding the variable teachers per pupil as an additional input. In Africa the change is dramatic, while in ECA and MNA there is no significant change. Further analysis is required to explain this differential response to the inclusion of this input.

In health there are multiple combinations of inputs (public expenditure, private expenditure, and literacy of adults) and outputs (life expectancy at birth, immunization DPT, immunization measles, and disability-adjusted life expectancy (DALE)). The combinations we selected produce rankings that are more

[9] The statistical significance of these changes has yet to be determined. The tests developed by Banker , and used in previous sections do not apply to the multiple-output cases we are analyzing here (Simar and Wilson, 2000).

Table 8

Education Attainment: Input Efficiency Scores by Regions across the World Multiple Inputs/Multiple Outputs

	AFR	EAP	ECA	LAC	MNA	SAS
2 inputs (public expenditure, teachers per pupil) – 2 outputs (gross primary and secondary enrolment)	.88	.83	.72	.82	.73	.91
3 inputs (public expenditure, teachers per pupil, literacy of adult) – 2 outputs (gross primary and secondary enrolment)	.92	.89	.86	.89	.92	.96
3 inputs (public expenditure, teachers per pupil, literacy of adult) – 2 outputs (net primary and secondary enrolment)	.87	.94	.93	.93	.92	1.0
2 inputs (public expenditure, literacy of adult) – 3 outputs (first complete, second level complete, average years of school)	.78	.92	.95	.84	.80	.91
3 inputs (public expenditure, literacy of adult, teachers per pupil) – 3 outputs (first complete, second level complete, avg yrs of school)	.91	.97	.94	.89	.81	.95
3 inputs (public expenditure, teachers per pupil, literacy of adult) – 3 outputs (literacy of youth, first level complete, second level complete)	.91	.97	.94	.89	.80	.95

homogeneous. The rank correlation is in the range of .65 to .98. (Tables 10-12). In health, Bangladesh appears also as efficient, as well as Niger, this being the result of the low levels of literacy of adults that bias these countries to appear as efficient.

3.4 Efficiency change over time

To examine the evolution of input and output efficiency over time, we computed the efficiency scores in two different time periods: 1975-1980 and

Table 9

Education Attainment: Output Efficiency Scores by Regions across the World Multiple Inputs/Multiple Outputs

	AFR	EAP	ECA	LA	MNA	SAS
2 inputs (public expenditure, teachers per pupil) – 2 outputs (gross primary and secondary enrolment)	.68	.83	.80	.85	.71	.79
3 inputs (public expenditure, teachers per pupil, literacy of adult) – 2 outputs (gross primary and secondary enrolment)	.82	.88	.89	.89	.91	.90
3 inputs (public expenditure, teachers per pupil, literacy of adult) – 2 outputs (net primary and secondary enrolment)	.79	.97	.96	.96	.92	1.0
2 inputs (public expenditure, literacy of adult) – 3 outputs (first complete, second level complete, average years of school)	.64	.87	.94	.80	.79	.83
3 inputs (public expenditure, literacy of adult, teachers per pupil) – 3 outputs (first complete, second level complete, average years of school)	.86	.94	.93	.86	.80	.89
3 inputs (public expenditure, teachers per pupil, literacy of adult) – 3 outputs (literacy of youth, first level complete, second level complete)	.98	1.0	1.0	.98	.99	.99

Table 10

Health Attainment – Multiple Inputs/Multiple Outputs

	Input-Efficient	Output-Efficient
More efficient	Bangladesh, Malaysia, Costa Rica, Kuwait, Morocco, Oman, Mauritius, Niger	Bangladesh, Costa Rica, Kuwait, Malaysia, Morocco, Mauritius, Oman, Niger
Least efficient	Russia, Belarus, Namibia, Romania, Estonia, Croatia, Lituania, Hungary, Jordan	Namibia, Togo, Ethiopia, Mozambique, Cote d'Ivoire, Cameroon, Congo, Central African Republic, Nigeria, Uganda

Table 11

Health Attainment: Input Efficiency Scores by Regions across the World Multiple Inputs/Multiple Outputs

	AFR	EAP	ECA	LA	MNA	SAS
2 inputs (public expenditure, literacy of adult) – 2 outputs (life expectancy, immunization DPT.)	.85	.82	.72	.82	.91	.93
3 inputs (public expenditure, private spending, literacy of adult) – 2 outputs (life expectancy, immunization DPT.)	.86	.82	.74	.83	.91	.94
3 inputs (public expenditure, private spending, literacy of adult) – 2 outputs (life expectancy, immunization measles.)	.86	.82	.77	.83	.91	.94
3 inputs (public expenditure, private spending, literacy of adult) – 3 outputs (life expectancy, immunization DPT., DALE)	.86	.82	.80	.87	.93	.94

Table 12

Health Attainment: Output Efficiency Scores by Regions across the World Multiple Inputs/Multiple Outputs

	AFR	EAP	ECA	LA	MNA	SAS
2 inputs (public expenditure, literacy of adult) – 2 outputs (life expectancy, immunization DPT.)	.81	.91	.97	.93	.97	.96
3 inputs (public expenditure, private spending, literacy of adult) – 2 outputs (life expectancy, immunization DPT.)	.81	.91	.97	.94	.97	.96
3 inputs (public expenditure, private spending, literacy of adult) – 2 outputs (life expectancy, immunization measles.)	.80	.91	.96	.94	.98	.96
3 inputs (public expenditure, private spending, literacy of adult) – 3 outputs (life expectancy, immunization DPT., DALE)	.82	.91	.97	.95	.98	.97

1996-2002 for education study, and 1997-99 and 2000-02 for health study, the construction of which is driven by data availability.[10]

Comparison of different input-output bundles in different time periods has to be done carefully because the frontier can be shifting outward through time. In some cases the frontier displacement can be parallel (such as in the life expectancy case of Figure 9). In others, the frontier displacement can be very uneven (biased frontier shift in Figure 9) reflecting biased technological change.

The detailed comparison between observed input-output combinations in different time periods distinguishes whether variations in the levels of input utilization or output production levels are due to changes in efficiency or changes in technology. This testing is possible with observed levels of inputs and outputs, and are based on the concept of a Malmquist Index (Fare, Grosskpof, Norris and Zhang, 1994). This method has been used to study productivity change in the OECD economies, as well as productivity in agriculture across the world (Coelli and Rao, 2003; Nin, Arndt, and Preckel, 2003).

Results show that over the two decades output efficiency growth was faster in the most inefficient countries, showing that there is a "catching-up" phenomenon. However, when measuring input efficiency, the previous results do not hold: most regions increased expenditure levels without increasing output.[11]

4. Explaining inefficiency variation across countries

This chapter seeks to identify factors correlated with inefficiency scores variation across countries. This two-stage approach attempts to identify statistically significant regularities common to efficient or inefficient countries using the more basic statistical techniques. This exercise does not try to identify supply or demand factors that affect health and education outcomes, such as those described by Filmer (2003). The scope is limited to verifying statistical association between the efficiency scores and environmental variables.

4.1 Method, variables and data description

Given that the dependent variable, the efficiency scores, is continuous and distributed over a limited interval (between zero and one), it is appropriate to use a censored (Tobit) regression model to analyze the relationships with other variables. The panel consists of a large numebr of countries (varying from 70 to 140 depending on the output indicator) and only two time periods. The literature on panel estimation has shown that in panels with this configuration, that is, a large number of cross-section units (countries) and a relatively short time dimension (2 periods),

---

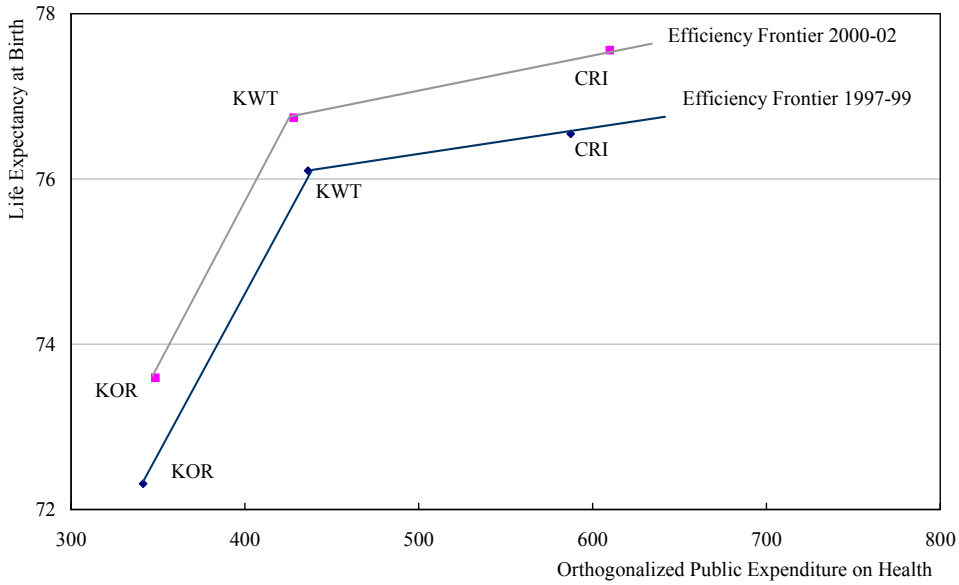
[10] Scores for individual countries are available at the PRMED website indicated in footnote 1.

[11] The results on country-by-country basis can be found at the PRMED website indicated in footnote 1.

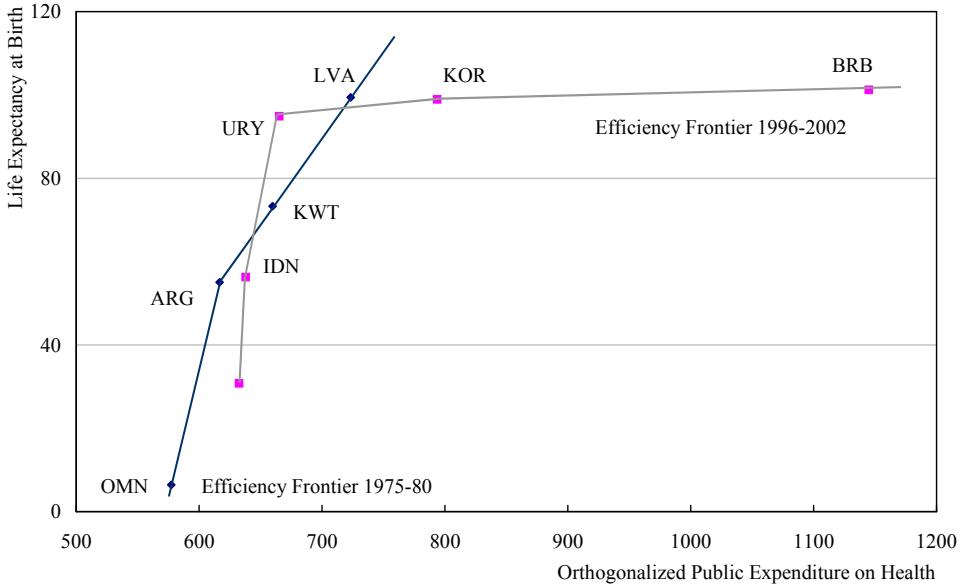
Figure 9

Efficiency Frontier Shift over Time

Parallel Shift



Biased Shift



Source for the figures on this page: World Bank WDI.

the fixed-effects estimators of the coefficients will be inconsistent (Maddala, 1987) and their variance will be biased downward (Greene, 2003b). Hence the random effects panel estimation method was preferred.

The dependent variable in the Tobit panel is the input efficiency score calculated by DEA method in the first stage. The input-oriented estimator reflects the consideration that input choices are more under the policymaker's control. The independent variables reflect environmental effects included in precursor papers, as well as suggested by others recently. We included the following independent variables:

a) *The size of government expenditure*. Most of the papers surveyed in the previous section explore the relationship between the size of the government (or expenditure as a percentage of GDP) and efficiency levels. The objective is to verify if additional pubic spending is associated with better education and health outcomes. While some papers have found a negative association between efficiency and expenditure levels (Gupta and Verhoeven, 2001, Jarasuriya and Woodon, 2003, and Afonso *et al.*, 2003), others have found a positive association (Evans *et al.*, 2003) and others have found no significant impact (Filmer and Pritchett, 1999).

b) *A government budget composition variable*. Given that both education and health are labor-intensive activities, the government's labor policies will determine the efficiency with which outputs are delivered. We choose a budget composition indicator to reflect this, in particular, the ratio of the wage bill to the total budget. A higher ratio is expected to be negatively correlated with efficiency.

c) *Per capita GDP*. We included the per capita GDP to control for the Balassa-Samuleson effect in comparing across countries. If richer countries tend to be more inefficient (given higher wages in these countries), a negative sign is expected. However, it must be recalled that to obtain the efficiency scores in the "fist stage" we constructed an auxiliary variable (the orthogonalized public expenditure). Hence the inclusion of this variable in the second stage is an attempt to control for any remaining Balassa-Samuleson effects.

d) *Urbanization*. The clustering of agents make it cheaper to provide services in urbanized areas rather than in rural. Higher degree of urbanization should reflect in higher efficiency, making positive as the expected sign of the coefficient on this variable.

e) *Prevalence of AIDS*. Based on WHO mappings of the disease, we included a dummy variable in the most severely affected countries to control for the role of this epidemic in the poor health outcomes. Evans *et al.* (2000) report that AIDS lowers the disability-adjusted life expectancy (DALE) by 15 years or more. Aids also affects education outcomes both directly and indirectly (Drake *et al.*, 2003), directly because school-age children are affected: UNAIDS estimates that almost 4 million children have been infected since the epidemic began, and two thirds have died. However, the indirect channel is relatively more important: AIDS leaves orphaned children that are more likely to drop out of school or repeat. All these factors reflect how AIDS affect the demand for education. But the supply is

also affected by the decreasing teacher labor force due to illness or death, or the need to care for family (Pigozzi, 2004). Prevalence of HIV/AIDS should be negatively associated with education and health outcomes. Consequently, efficiency scores should be negatively associated with the dummy variable.

f) Income distribution inequality. Ravallion (2003) argues that, besides the mean income, its distribution affects social indicators because their attainment is mostly determined by the income of the poor. Hence, we controlled for the distribution of income by including the Gini coefficient as an explanatory variable. Higher inequality is expected to be associated with lower educational and health attainments, making negative the expected sign of this variable.

g) Share of public sector in the provision of service. Services can be provided by both the public and private sectors, and efficiency indicators will differ across countries depending on the relative productivities of both sectors. Previous studies have included this variable to explain differences in outcomes (Le Grand, 1987; Berger and Messer, 2002) or efficiency scores (Greene, 2003a). The specific variable we included was the ratio of publicly financed service over the total spending (sum of private and public spending).

h) External aid. To the extent that countries do not have to incur the burden of taxation, they may not have the incentive to use resources in the most cost-effective way. Another channel through which aid-financing may affect efficiency is through the volatility and unpredictability of its flows. Given that this financing source is more volatile than other types of fiscal revenue (Bulir and Hamann, 2000), it is difficult to undertake medium-term planning within activities funded with aid resources. If this is the case, we would expect a negative association between aid-dependence and efficiency in those activities funded with aid, mostly health services. To our knowledge there are no previous attempts to establish a relationship between efficiency and the degree to which activities are financed by external aid. There is, however, recent evidence of a negative association between donor financing and some health outcomes (Bokhari, Gottret and Gai, 2005).

i) Institutional variables. Countries with better institutions, more transparency, and less corruption are expected to have higher efficiency scores. Similarly, countries that have suffered wars or state failures are expected to register lower efficiency scores. To capture these effects we included different indicators: the ICRG International Country Risk Indicators, the Worldwide Governance Research Indicators, in particular the Control of Corruption component (Kaufmann *et al.*, 2002). We also included a dummy variable if there had been some type of state failure, such as internal wars, from the State Failure Task force database.

The data on educational and health indicators are not available on a continuous annual basis for many countries. Thus, averages of the variables were computed over sub-periods both in the first stage calculation of efficiency score and in the second stage of regression analysis. Specifically, educational indicators are averaged over two periods (1975-80 and 1996-2002) and health indicators over two periods (1996-99 and 2000-02). This discrepancy in the sub-period construction is due exclusively to the lack of data for earlier years. The averages are treated as

separate observations. The advantages of this approach are threefold. First, the averages may serve as a better measure of the educational and health attainment, which can hardly be substantially improved on a yearly basis; second, the averaging maximizes the coverage of countries for each period, since one observation of a certain year is sufficient to help the country survive in the cross sectional comparison; Third, the time series thus constructed for each country, although short, facilitates the implementation of econometric techniques on panel data to explore the efficiency variations across countries and through time.

4.2 Results

The Tobit estimation on panel data is defined as follows:

$$VRSTE_{it} = f(WAGE_{it}, GOVEXP_{it}, PUBTOT_{it}, GDPPC_{it}, URBAN_{it}, AIDS_{it}, GINI_{it}, EXTAID_{it}, INST_{it}, CONS)$$

where:

$VRSTE_{it}$ = Variable returns to scale DEA efficiency score for single output and multiple output cases

$WAGE_{it}$ = Wages and salaries (percent of total public expenditure)

$GOVEXP_{it}$ = Total government expenditure (percent of GDP)

$PUBTOT_{it}$ = Share of expenditures publicly financed (public/total)

$GDPPC_{it}$ = GDP per capita in constant 1995 US dollars

$URBAN_{it}$ = Urban population (percent of total)

$AIDS_{it}$ = Dummy variable for HIV/AIDS

$GINI_{it}$ = Gini Coefficient

$EXTAID_{it}$ = External aid (percent of fiscal revenue)

$INST_{it}$ = Institutional indicators including ICRG country risk, World Governance Research Indicators (Corruption Control), or a dummy for state failures from the State Failure Task Force database

$CONS$ = Constant

Tables 13 and 14 report the results for the single input/single output case and the multiple input/multiple output case, respectively. The more interesting findings are:

a) We find that countries with larger expenditure levels also register the more inefficient scores. This result is robust to changes in the output indicator selected, to considering health or education, and to adopting either the single output or multiple output frameworks. The trade-off between size of expenditure and efficiency is quite robust.

b) Countries in which the wage bill represents a higher fraction of total expenditure tend to be more inefficient. This result does not hold for health in the multiple

output framework. This difference could be due partly to the relatively decreasing number of health care professionals in the world, especially in the poorer countries (Liese *et al*., 2003). Further investigation would be required to examine why this is not the case in education.

c) Countries in which public financing is a larger share of total expenditure on the service also register lower efficiency scores. This is probably due to differential productivity rates in the provision of services. Further research would be needed to explain why this is the case in health services. Recent case studies of water companies in Argentina show that private companies were more efficient than public ones and provided better service quality leading to lower child mortality rates (Galiani, Gertler and Schargrodsky, 2005). In education, there is some evidence that efficiency scores are lower in public schools (Alexander and Jaforullah, 2004), though the evidence regarding the impact of privatizing education on outcomes is mixed (World Bank, 2003).

d) Urbanization is positively associated with efficiency scores in both education and health. However, when life expectancy is included as an output, the relationship is non-significant (single output) or negative (multiple output). Possibly the urbanization variable is capturing other effects such as crime. There is ample literature studying the relationship between urbanization and crime (Glaeser and Sacerdote, 1999). Alternatively, as urbanization intensifies, communicable diseases are more difficult and costly to control, hence the negative association found between both variables in health.

e) The effect of the HIV/AIDS is clearly negative affecting health efficiency scores in the multiple-output models. However, its effect on education is less clear, as the expected negative sign is significant in few cases and has the opposite sign in equal number of cases. This confirms the difficulty of empirically verifying this relationship, reported in previous work (Wobst and Arndt, 2003).

f) Income distribution has the expected negative effect on the educational and health efficiency scores. The impact of inequality on health scores is less robust than in education, but confirms Greene's findings (2003). Other papers (Berger and Messer, 2002), have found a positive association between income inequality and health outcomes.

g) Results showed a negative relationship between some of the efficiency scores and the external aid dependency ratio. Only in one of the multiple output cases is the external aid associated with higher efficiency, but with borderline statistical significance. Though no causality relationship can be inferred from the exercise, this is one of the results that merit more detailed research. This result might be explained by the volatility of aid as a funding source that limits medium term planning and effective budgeting. Probably this is why the negative sign is more robust in health than in education, given that donor funding is mostly directed

Table 13

Explaining Cross-country Variation in Efficiency – Single Input/Single Output

Independent Variable	Gross Primary Enrolment	Net Primary Enrolment	Gross Secondary Enrolment	Net Secondary Enrolment	Literacy of Youth	Average Years of School	First Level Complete	Secondary Level Complete	Life Expect-ancy	Immun-ization DPT	Immun-ization Measles
WAGE	–.00117[c]	–.00357[a]	–.00172[b]	–.00680[a]	–.00189[b]	–.00570[a]	–.00470[b]	–.00546[a]	.00065	–.00052	–.00049
GOVEXP	–.00387[a]	–.00546[a]	–.00340[a]	–.00455[b]	–.00387[a]	–.00696[a]	–.00566[a]	–.00765[a]	–.00269[b]	–.00078	–.00227[c]
PUBTOT	-	-	-	-	-	-	-	-	–.00213[a]	–.00150[a]	–.00135[c]
GDPPC	–.00002[a]	–.00002[a]	–.00001[a]	.00002[b]	–.00002[a]	–1.5e–6	–.00001	–7.7e–6	7.6e–7	–.00001[a]	–.00001[a]
URBAN	.00167[a]	.00143[c]	.00168[a]	.00037	.00187[a]	.00532[a]	.00551[a]	.00555[a]	–.00018	.00099[b]	.00088
AIDS	–.04471[b]	–.08731[b]	–.02204	.01243	–.02974	.12717[c]	.1211[c]	.11041	–.05473	–.01108	–.02730
GINI	–.06688	.01507	–.19326[b]	–.42311	–.18484[c]	–.44658[b]	–.34402	–.45870[b]	.22118	.09510	.08692
EXTAID	–.00094	–.00196[b]	–.00021	–.00106	–.00054	.00089	–.00025	–.00006	–.00224[c]	–.00155	–.00324[b]
CONS	1.02996[a]	1.1282[a]	1.0472[a]	.84138[a]	1.0697[a]	.76791[a]	.70009[a]	.81705[a]	.79193[a]	.78734[a]	.84384[a]
# of Obs (# of Countrs)	79 (51)	44 (30)	79 (51)	34 (20)	72 (46)	71 (45)	71 (45)	71 (45)	118 (69)	118 (69)	118 (69)
Wald Chi2(6) (Prob > Chi2)	83.91 (.00)	66.09 (.00)	46.72 (.00)	55.31 (.00)	44.27 (.00)	64.13 (.00)	45.53 (.00)	61.94 (.00)	50.83 (.00)	123.97 (.00)	35.01 (.00)

Note: [a] 0.01 significance level, [b] 0.05 significance level, [c] 0.10 significance level, and insignificant otherwise.

Table 14

Explaining Cross-country Variation in Efficiency – Multiple Inputs/Multiple Outputs

Independent Variable	EDU2-2	EDU2-2n	EDU3-2	EDU3-2n	EDU3-3	EDU3-3bl	HEA2-2	HEA3-2	HEA3-2m	HEA3-3
WAGE	–.00212[b]	–.00767[a]	–.00219[b]	–.00425	–.001000	–.00340[c]	.00126[a]	.00205[a]	.00203[c]	.00203[c]
GOVEXP	–.00321[a]	–.00365	–.00203[c]	.00099	–.00123[c]	–.00316[c]	–.0012[c]	–.00273[a]	–.0009	–.00090
PUBTOT	-	-	-	-	-	-	–.00151[a]	–.00142[a]	–.00159[c]	–.00151[c]
GDPPC	–.00001[b]	–6.6e–7	–.00001[c]	–.00003	–4.2e–6	1.98e–6	–2.7e–6	4.2e–6[a]	–7.1e–7	–9.3e–7
URBAN	.00138[c]	–.00045	.00191[b]	.001997	.00127[a]	.00091	–.00095[a]	–.00148[a]	–.00106	–.00105
AIDS	–.03295	–.05843	–.00956	–.14763	.01797	.06022	–.04815[a]	–.033147[b]	–.07162	–.06999
GINI	–.06485	.43602	–.14717	.27058	–.17237[b]	–.15697	–.03997	–.07958[c]	–.01015	–.01387
EXTAID	.00010	–.00622	.00152	–.00274	–.00066	.00123	.00087	.00128[c]	–.00095	–.00106
CONS	1.0655[a]	1.0223	1.0642[a]	1.0124[a]	1.06570[a]	1.1218[a]	1.0098	1.0117[a]	.98891[a]	.98787[a]
# of Obs	76	34	69	32	69	63	97	98	98	98
(# of Countrs)	(49)	(20)	(44)	(19)	(44)	(40)	(55)	(56)	(56)	(56)
Wald Chi2(6)	24.48	11.69	20.84	7.44	18.72	9.18	185.21	229.98	19.25	18.62
(Prob > Chi2)	(.00)	(.11)	(.00)	(.38)	(.01)	(.24)	(.00)	(.00)	(.01)	(.02)

Notes:
[a] 0.01 significance level, [b] 0.05 significance level, [c] 0.10 significance level, and insignificant otherwise
EDU2-2: Inputs: orthogonalized public spending on education per capita, teachers per pupil
Outputs: gross primary and secondary enrolments
EDU2-2n: same inputs as EDU2-2, outputs: net primary and secondary enrolment
EDU3-2: literacy of adult is added to EDU2-2 as input
EDU3-2n: literacy of adult is added to EDU2-2n as input
EDU3-3: literacy of youth is added to EDU3-2 as output
EDU3-3bl: same inputs as in EDU3-2,
Outputs: average years of school, first level complete, and second level complete (Barro-Lee education indicators)
HEA2-2: Inputs: orthogonalized public spending on health per capita, literacy of adult
Outputs: life expectancy at birth, and immunization DPT
HEA3-2: orthogonalized private spending on health per capita is added to HEA2-2 as input
HEA3-2m: Immunization Measles is in place of DPT in HEA3-2 as output
HEA3-3: Immunization Measles is added to HEA3-2 as output.

towards the first. Recent research (Bokhari, Gottret and Gai, 2005) show a negative association between some health outcomes and the degree of donor funding, pointing in this same direction. This result also coincides with researchshowing that the quality of policies is not only unrelated to donor financing, but that highly indebted countries with "bad" policies received more net transfers as a share of GDP (Birdsall *et al.*, 2003).

h) None of the institutional variables proved to be statistically significant. We interpret this result as due to the data limitations, as some of the most crucial information, for instance the corruption index is only available since 1996 and the panel exercise was limited to a cross section. The state-failure dummy variable or the ICRG indicators did not prove to be significant either. Hence, these results are not reported in any of the tables.

To investigate the possibility of slope heterogeneity across countries, we followed the approach used in Haque, Pesaran, and Sharma (1999). Specifically, the slope coefficients in each country are assumed to be fixed over time, but varying across countries linearly with the individual sample mean of GDP per capita. The final results (Tables 15 and 16) only include the statistically significant interaction terms, in order to avoid co linearity arising from the correlation between original explanatory variables and the auxiliary variable capturing the interaction of these with the sample mean of GDP per capita. Hence the estimated model is:

$$VRSTE_{it} = f(WAGE_{it}, GOVEXP_{it}, GDPPC_{it}, URBAN_{it}, AIDS_{it}, GINI_{it}, WAGEG_{it}, GOVG_{it}, GINIG_{it} CONS)$$

where:

$VRSTE_{it}$ = Variable returns to scale DEA efficiency score for single output and multiple output cases

$WAGE_{it}$ = Wages and salaries (percent of total public expenditure)

$GOVEXP_{it}$ = Total government expenditure (percent of GDP)

$PUBTOT_{it}$ = Share of expenditures publicly financed (public/total)

$GDPPC_{it}$ = GDP per capita in constant 1995 US dollars

$URBAN_{it}$ = Urban population (percent of total)

$AIDS_{it}$ = Dummy variable for HIV/AIDS

$GINI_{it}$ = Gini Coefficient

$CONS$ = Constant

$WAGEG_{it} = WAGE_{it} * \overline{GDPPC_i}$

$GOVG_{it} = GOVEXP_{it} * \overline{GDPPC_i}$

$$GINIG_{it} = GINI_{it} * \overline{GDPPC_i}$$

$$\overline{GDPPC_i} = T^{-1} \sum_{t=1}^{T} GDPPC_{it}$$

Results show that the interaction terms are significant, especially for the health regression, implying that there is a heterogeneous response of efficiency scores to the different explanatory variables. This confirms Greene's (2003) results on the WHO data. One of the key results of this section is that the negative association between the size of government expenditure and efficiency is stronger in countries with higher per capita GDP. Similarly, this happens with the wage variable. Results are somewhat similar to those of the homogeneous slopes, though statistical significance of many of the coefficients is lower. This is the result of colinearity between the auxiliary variables and the original set of explanatory variables. This problem deserves further work in the future.

Interpretation of these results requires caution due to several limitations. First, education and health outcomes are explained by multiple supply and demand factors (Filmer, 2003) that are not included here. This is not the object of the present paper. The omission of one of these factors in the health or education production functions in the previous stage could explain some of the cross-country covariation of the efficiency results (Ravallion, 2003). The goodness-of-fit analysis of the first stage indicated that no important factor seemed to be omitted. Of course, there can always be additional factors that could be included but the curse of dimensionality[12] is particularly pressing in non-parametric statistical methods (even if the data were available).

The second limitation derives from the intuitive question why the set of explanatory variables used in the second stage were not included in the first stage. The answer lies in that most of these variables are environmental and outside the control of the decision-making unit. The inclusion of these environmental variables would have had little justification from the production function perspective. Additionally, by maintaining the production function as simple as possible the dimensionality curse is avoided.

Finally, the third limitation arises from the fact that if the variables used in the first stage to obtain the efficiency estimator are correlated with the second stage explanatory variables, the coefficients will be inconsistent and biased (Simar and Wilson, 2004; Grosskopf, 1996; Ravallion, 2003). To examine the extent of this potential problem we calculated correlation coefficients between the first-stage inputs and the second-stage explanatory variables. The largest correlation coefficients were between GDP per capita and the teachers per pupil ratio and the literacy of the adult. To examine the sensitivity of the results to the inclusion of GDP per capita, all the estimations were performed without this variable and none of the results changed.

---

[12] As the number of outputs increase, the number of observations must increase exponentially to maintain a given mean-square error of the estimator. See Simar and Wilson (2000).

Table 15

Explaining Cross-country Variation in Efficiency – Single Input/Single Output – Heterogeneous Slopes

Independent Variable	Gross Primary Enrolment	Net Primary Enrolment	Gross Secondary Enrolment	Net Secondary Enrolment	Literacy of Youth	Average Years of School	First Level Complete	Secondary Level Complete	Life Expectancy	Immun-ization DPT	Immun-ization Measles
WAGE	–.00006	.00076	–.00035	–.00228	–.00056	–.00200	–.00120	–.00419	–.00306[c]	–.00079	–.00241
GOVEXP	–.00363[a]	–.00255[c]	–.00377[a]	–.00727[c]	–.00552[a]	–.00595[c]	–.00453	–.00611[c]	.00337[b]	.00168[c]	.00221
PUBTOT	-	-	-	-	-	-	-	-	-	–.00162[a]	–.00097
GDPPC	–.00002[a]	–.00002[a]	–5.4e–6	.00003[a]	–.00002[c]	.00004[a]	.00003[c]	.00003[c]	.00002[b]	–.00002[a]	–.00001
URBAN	.00179[a]	.00132[b]	.00193[a]	.00139	.00212[a]	.00566[a]	.00601[a]	.00593[a]	–.00080	–.00117[a]	.00021
AIDS	–.03866[c]	–.06603b	–.03153	.01010	–.02177	.05491	.06656	.06464	–.02321	–.04147[b]	–.00826
GINI	–.14230	–.42098a	–.14976	–.29395	–.13107	–.09995	–.15463	–.24762	–.12865	–.38851[a]	–.42162[b]
WAGG	–4.4e–6[c]	–1.2e–6a	–4.6e–7[c]	–9.4e–7	–4.5e–7	–8.1e–7	–8.8e–7	–2.4e–7	8.9e–7[b]	6.95e–8	5.1e–7
GOVG	–8.6e–8	–5.2e–7c	4.3e–8	3.6e–7	4.0e–7	–4.3e–7	–4.4e–7	–5.3e–7	–1.4e–6[a]	–5.4e–7[a]	–9.4e–7[a]
GINIG	.00003	.00011[a]	–2.4e–6	–.00003	2.0e–6	–.00006	–.00005	–.00006	.00001	.00009[a]	.00006[c]
CONS	1.0156[a]	1.1036[a]	1.0098[a]	.74603[a]	1.0365[a]	.60371[a]	.53977[a]	.68648[a]	.82665[a]	1.0119[a]	.93820[a]
# of Obs (# of Countrs)	82 (52)	47 (31)	82 (52)	36 (21)	75 (47)	74 (46)	74 (46)	74 (46)	120 (70)	121 (71)	121 (71)
Wald Chi2(6) (Prob > Chi2)	87.32 (.00)	93.98 (.00)	62.74 (.00)	105.34 (.00)	58.40 (.00)	94.00 (.00)	69.32 (.00)	82.38 (.00)	74.33 (.00)	450.54 (.00)	52.71 (.00)

Note: [a] 0.01 significance level, [b] 0.05 significance level, [c] 0.10 significance level, and insignificant otherwise.

Table 16

Explaining Cross-country Variation in Efficiency – Multiple Inputs/Multiple Outputs – Heterogeneous Slopes

Independent Variable	EDU2-2	EDU2-2n	EDU3-2	EDU3-2n	EDU3-3	EDU3-3bl	HEA2-2	HEA3-2	HEA3-2m	HEA3-3
WAGE	.00051	–.00140	.00005	.00494	–.00018	–.00045	–.00063	–.00065	–.00093	–.00092
GOVEXP	–.00323[b]	.00501	–.00385[b]	.00520	–.00256[b]	–.00459	.00122[c]	.00063	–.00070	–.00064
PUBTOT	–	–	–	–	–	–	–.00180[a]	–.00145[b]	–.00149[c]	–.00141[c]
GDPPC	–8.6e–6	.00002	1.7e–6	.00003	–1.8e–6	–2.1e–6	–.00001[b]	–.00001	–.00003[b]	–.00003[b]
URBAN	.00137[b]	.00079	.00166[b]	.00096	.00134[a]	.00064	–.00246[a]	–.00167[c]	–.00160	–.00159
AIDS	–.04139	–.06211	–.04744	–.20362[a]	.00646	.04633	–.06289[a]	–.04001	–.07217	–.07025
GINI	–.14418	–.18676	.07096	–.02601	–.07474	–.20029	–.32844[a]	–.45695[b]	–.29885	–.30857
WAGG	–8.3e–7[b]	–1.2e–6	–6.4e–7[c]	–1.9e–6	–2.0e–7	–7.9e–7	7.8e–7[a]	7.2e–7	6.0e–7	6.0e–7
GOVG	–6.3e–8	–2.6e–6[c]	3.5e–7	–1.2e–6	3.0e–7	3.5e–7	–5.98e–7[a]	–4.9e–7	2.7e–8	1.4e–8
GINIG	.00003	.00012	–.00003	.00005	–.00002	.00003	.00005[a]	.00005[c]	.00006	.00006[c]
CONS	1.0515[a]	.89986[a]	1.0021[a]	.84756[a]	1.0464	1.1257[a]	1.1494[a]	1.1457[a]	1.1512[a]	1.1495[a]
# of Obs (# of Countrs)	79 (50)	36 (21)	72 (45)	34 (20)	72 (45)	66 (41)	101 (58)	101 (58)	101 (58)	101 (58)
Wald Chi2(6) (Prob > Chi2)	41.93 (.00)	18.57 (.03)	31.15 (.00)	18.71 (.22)	23.89 (.00)	13.22 (.15)	600.70 (.00)	37.22 (.00)	25.33 (.00)	24.74 (.01)

Notes:
[a] 0.01 significance level, [b] 0.05 significance level, [c] 0.10 significance level, and insignificant otherwise
EDU2-2: Inputs: orthogonalized public spending on education per capita, teachers per pupil
Outputs: gross primary and secondary enrolments
EDU2-2n: same inputs as EDU2-2, outputs: net primary and secondary enrolment
EDU3-2: literacy of adult is added to EDU2-2 as input
EDU3-2n: literacy of adult is added to EDU2-2n as input
EDU3-3: literacy of youth is added to EDU3-2 as output
EDU3-3bl: same inputs as in EDU3-2,
Outputs: average years of school, first level complete, and second level complete (Barro-Lee education indicators)
HEA2-2: Inputs: orthogonalized public spending on health per capita, literacy of adult
Outputs: life expectancy at birth, and immunization DPT
HEA3-2: orthogonalized private spending on health per capita is added to HEA2-2 as input
HEA3-2m: Immunization Measles is in place of DPT in HEA3-2 as output
HEA3-3: Immunization Measles is added to HEA3-2 as output.

5. Concluding remarks and directions for future work

The paper presented an application of non-parametric methods to analyze the efficiency of public spending. Based on a sample of more than 140 countries, the paper estimated efficiency scores for nine education output indicators and fourth health output indicators. Our results indicate that, in general, the least efficient countries could achieve substantially higher education and health output levels. Alternatively they could produce the same output level consuming approximately 50 per cent less of the inputs implicit in the efficiency frontier. It is crucial to identify what are the institutional or economic factors that cause some countries to be more efficient than others in the service delivery.

In terms of policy implications, it is crucial to differentiate between the technically efficient level and the optimal or desired spending level. Even if a country is identified as an "efficient" benchmark country, it may very well still need to expand its public spending levels to achieve a target level of educational or health attainment indicators. Such is the case of countries with low spending levels and low attainment indicators, close to the origin of the efficient frontier. The important thing is that countries expand their scale of operation along the efficient frontier.

The methods used in the paper can be interpreted as tools to identify extreme cases of efficient units and inefficient cases. Once the cases have been identified, more in-depth analysis is required to explain departures from the benchmark, as proposed and done by Sen (1981). Given that the methods are based on estimating the frontier directly from observed input-output combinations they are subject to sampling variability and are sensitive to the presence of outliers. Recent advances allow dealing with these problems such as in Wilson (2004). Additionally, it would be useful to contrast these results with those obtained with the use of parametric stochastic frontier estimation.

In a "second stage" the paper verified statistical association between the efficiency scores and environmental variables that are not under the control of the decision-making units. The panel Tobit regressions showed that the variables, which are negatively associated with efficiency scores, include the size of public expenditure, the share of the wage bill in the total public budget, the proportion of the service that is publicly financed, the prevalence of HIV/AIDS epidemic on health efficiency scores, income inequality on education efficiency scores, and external aid-financing on some of the efficiency scores. This last impact is probably due to the volatility of aid that impedes effective medium term planning and budgeting, and probably explains why the result is more robust in health than in education where most of the donor-funding is directed. This result points in the same direction of previous research showing that donor financing is unrelated to the quality of domestic policies and that, in the case of highly indebted counties, those with worse policies received more transfers. A positive association between urbanization and efficiency outcomes is also identified in education but some of the health efficiency scores are negatively associated. This last result probably is due to higher crime rates in the cities or the effect of communicable diseases that spread with agglomeration. These are topics for further research in case studies.

APPENDIX
DATA ENVELOPMENT ANALYSIS (DEA) MODEL

A measure of production efficiency, perhaps the simplest one, is defined as the ratio of output to input. It is, however, inadequate to deal with the existence of multiple inputs and outputs. The relative efficiency for all decision-making units (DMU), *j=1,…,n*, is then modified as the ratio of weighted outputs to weighted inputs, more precisely:

$$\textit{Relative efficiency} = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \tag{A.1}$$

where *x* and *y* are inputs and outputs, respectively, and *u* and *v* are the common weights assigned to outputs and inputs, respectively. A challenge of this measure immediately follows: it is difficult to justify the common weights given that DMUs may value inputs and outputs differently.

The seminal paper by Charnes, Cooper and Rhodes (1978) proposed the following ratio form to allow for difference in weights across DMUs, which establishes the foundation of data envelopment analysis (DEA).

$$\begin{aligned} &\textit{Max} \quad h_0 = \frac{\sum_{r=1}^{s} \mu_r y_{r0}}{\sum_{i=1}^{m} v_i x_{i0}} \\ &\textit{subject to}: \\ &\qquad \frac{\sum_{r=1}^{s} \mu_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1, \quad j = 1, \cdots, n \\ &\qquad \mu_r \ge \varepsilon, \quad r = 1, \cdots, s \\ &\qquad v_i \ge \varepsilon, \quad i = 1, \cdots, m \\ &\qquad \varepsilon > 0 \end{aligned} \tag{A.2}$$

In the model, there are *j = 1,…, n* observed DMUs which employ *i = 1,…, m* inputs to produce *r = 1,…, s* outputs. One DMU is singled out each time, designated as DMU_0, to be evaluated against the observed performance of all DMUs. The objective of model (A.2) is to find the most favorable weights, μ_r and v_i, for DMU_0 to maximize the relative efficiency. The constraints are that the same weights will make ratio for every DMU be less than or equal to unity. The optimal value of the ratio must be $0 \le h_0^* \le 1$ and DMU_0 is efficient if and only $h_0^* = 1$, otherwise it is considered as relatively inefficient. One problem with the ratio formulation is that there are an infinite number of solutions: if μ_r and v_i are solutions to (A.2), so are $\alpha\mu_r$ and αv_i, $\forall \alpha > 0$.

It is worth observing one important feature of model (A.2). In maximizing the objective function it is the relative magnitude of the numerator and the denominator that really matters and not their individual values. It is thus equivalent to setting the denominator to a constant, say 1, and maximizing the numerator. This transformation will not only lead to the uniqueness of solution but also convert the fractional formulation of model (A.2) into a linear programming problem in model (A.3).

$$
\begin{aligned}
&\textit{Max} \quad \sum_{r=1}^{s} \mu_r y_{r0} \\
&\textit{subject to}: \\
&\qquad \sum_{i=1}^{m} v_i x_{i0} = 1 \\
&\qquad \sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \quad j = 1, \cdots, n \\
&\qquad -\mu_r \le -\varepsilon, \quad r = 1, \cdots, s \\
&\qquad -v_i \le -\varepsilon, \quad i = 1, \cdots, m
\end{aligned}
\tag{A.3}
$$

Model (A.3) facilitates straightforward interpretation in terms economics. The objective is now to maximize the weighted output per unit weighted input under various conditions, the most critical one of which is that the virtual output does not exceed the virtual input for any DMU.

Since model (A.3) is a linear programming, we can convert the maximization problem into a minimization problem, e.g. a *dual* problem, by assigning a dual variable to each constraint in the *primal* (A.3). Specifically, dual variables $\theta, \lambda_j, s_r^+, s_i^-$ are assigned as follows.

Max	$\sum_{r=1}^{s} \mu_r y_{r0}$	Dual Variable
subject to:		
	$\sum_{i=1}^{m} v_i x_{i0} = 1$	θ (A.3')
	$\sum_{r=1}^{s} \mu_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, \quad j = 1, \cdots, n$	λ_j
	$-\mu_r \le -\varepsilon, \quad r = 1, \cdots, s$	s_r^+
	$-v_i \le -\varepsilon, \quad i = 1, \cdots, m$	s_i^-

A *dual* minimization problem is thus derived as model (A.4). It is clear that model (A.4) has *m+s* constraints while model (A.3) has *n+m+s+1* constraints. Since

n is usually considerably larger than $m+s$, the dual DEA significantly reduces the computational burden and is easier to solve than the primal.

$$\begin{aligned} &\textit{Min} \qquad \theta - \varepsilon\left[\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right] \\ &\textit{subject to}: \\ &\theta x_{i0} - \sum_{j=1}^{n} x_{ij}\lambda_j - s_i^- = 0 \\ &y_{r0} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ \\ &\lambda_j \geq 0,\ \ s_r^+ \geq 0,\ s_i^- \geq 0 \\ &i = 1,\cdots,m,\ \ r = 1,\cdots,s,\quad j = 1,\cdots,n \end{aligned} \tag{A.4}$$

More importantly, the duality theorem of linear programming states that the solution value to the objective function in (A.4) is exactly equal to that in (A.3). And, the dual variables, $(\lambda_1, \lambda_2, \cdots, \lambda_n)$, have the interpretation of Lagrange multipliers. That is, the value of a dual variable is equal to the shadow price of Lagrange Multiplier. It is also known that, from constrained optimization problem, $\lambda_j > 0$ normally when the constraint in (A.3') is binding and $\lambda_j = 0$ if not. Note that the binding constraint in (A.3) implies that the corresponding DMU is efficient. In another word, efficient units are identified by positive $\lambda's$ while inefficient units are given $\lambda's$ of zero. The DMU in question in model (A.4) is thus compared with the efficient DMUs only, named as comparison *peers* in the literature. The solution values of $\lambda's$ reflect the exact weights assigned to each peer in the evaluation of DMU_0.

Since only efficient DMUs exert effective constraints in model (A.4), as argued above, the input/output bundle, $(\sum_{j=1}^{n} x_{ij}\lambda_j, \sum_{j=}^{n} y_{rj}\lambda_j)$, is the most efficient combination for $i = 1,\cdots,m$ and $r = 1,\cdots,s$. To achieve an output level y_{r0}, which is as close as possible to $\sum_{j=}^{n} y_{rj}\lambda_j$, DMU_0 has to use an input bundle to meet the minimum requirement, $\sum_{j=1}^{n} x_{ij}\lambda_j$. This further implies that the solution θ^* is the lowest proportion of the current input bundle, x_{i0} used by DMU_0, that is actually required to meet the minimum input requirement and produce target output y_{r0}. The solution θ^* is defined as the efficiency score for DMU_0. For instance, $\theta^* = 0.60$ implies that 40 per cent of current input is a waste of resources.

Model (A.4) also offers the explanation why the data envelopment analysis is so named. The first constraint in (A.4) defines a lower limit of inputs and the second constraint an upper limit of outputs for DMU_0, and within the limits θ is minimized. The set of solutions to all DMUs forms an upper bound that envelops all observations.

REFERENCES

Afonso, A. and M. St. Aubyn (2004), "Non-parametric Approaches to Education and Health: Expenditure Efficiency in OECD Countries", Technical University of Lisbon, mimeo.

Afonso, A., L. Schuknecht and V. Tanzi (2003), "Public Sector Efficiency: An International Comparison", European Central Bank, Working Paper, No. 242.

Aigner, D.J., C.A.K. Lovell and P. Schmidt (1977), "Formulation and Estimation of Stochastic Frontier Production Function Models", *Journal of Econometrics*, Vol. 6, pp. 21-37.

Alexander, R. and M. Jaforullah (2004), "Explaining Efficiency Differences in New Zealand Secondary Schools", paper presented at the "International Conference on Policy Modeling", June, Paris.

Banker, R. (1993), "Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation", *Management Science,* Vol. 39, No. 10.

Berger, M. and J. Messer (2002), "Public Financing of Health Expenditures, Insurance, and Health Outcomes", *Applied Economics*, Vol. 34, pp. 2105-13.

Birdsall, N., S. Claessens and I. Diwan (2003), "Policy Selectivity Forgone: Debt and Donor Behavior in Africa", *The World Bank Economic Review*, Vol. 17, No. 3.

Bishop, P. and S. Brand (2003), "The Efficiency of Museums: A Stochastic Frontier Production Function Approach", *Applied Economics*, Vol. 35, pp. 1853-58.

Bokhari, F., P. Gottret and Y. Gai (2005), "Government Health Expenditures and Health Outcomes", The World Bank, mimeo.

Bulir, A. and A. Hamann (2000), "Aid Volatility: An Empirical Assessment", IMF, Staff Papers, Vol. 50, No. 1, pp. 64-89.

Burgess, J. and P. Wilson (1998), "Variation in Inefficiency in U.S. Hospitals", *Canadian Journal of Operational Research and Information Processing*, Vol. 6, pp. 84-102.

Bowlin, W.F. (1998), "Measuring Performance: An Introduction to Data Envelopment Analysis (DEA)", *Journal of Cost Analysis*, pp. 3-27.

Cazals, C., J. Florens and L. Simar (2002), "Non-parametric Frontier Estimation: A Robust Approach", *Journal of Econometrics*, Vol. 106, No. 1, pp. 1-25.

Chang, T. (2002), "An Econometric Test of Wagner's Law for Six Countries Based on Cointegration and Error-correction Modeling Techniques", *Applied Econometrics*, No. 34, pp. 1157-69.

Charnes, A., W. Cooper and E. Rhodes (1978), "Measuring Efficiency of Decision-making Units", *European Journal of Operational Research*, No. 3, pp. 429-44.

Cherchye, L. and T. Post (2001), "Methodological Advances in DEA: A Survey and an Application for the Dutch Electricity Sector", ERIM Report Series Research in Management, ERS-2001-53-F&A.

Crouch, L. and T. Fasih (2004), "Patterns in Educational Development: Implications for Further Efficiency Analysis", The World Bank, mimeo.

Cullinane, K. and D. Song (2003), "A Stochastic Frontier Model of the Productive Efficiency of Korean Container Terminals", *Applied Economics*, Vol. 35, pp. 251-67.

Drake, L., C. Maier, M. Jukes, A. Patrikios, D. Bundy, A. Gardner and C. Dolan (2003), "School-age Children: Their Nutrition and Health".

Evans, D., A. Tandon, C.J.L. Murray and J.A. Lauer (2000), "The Comparative Efficiency of National Health Systems in Producing Health: An Analysis of 191 Countries", World Health Organization, GPE Discussion Paper Series, No. 29.

Färe, R. and S. Grosskopf (1995), "Non-parametric Tests of Regularity, Farrell Efficiency, and Goodness of Fit", *Journal of Econometrics*, Vol. 69, No. 2, pp. 415-25.

Farrell, M. (1957), "The Measurement of Productive Efficiency", *Journal of the Royal Statistical Society*, Series A, Vol. 120, No. 3, pp. 253-90.

Filmer, D. (2003), "Determinants of Health and Education Outcomes", background note for *World Development Report 2004*, The World Bank, mimeo, May.

Filmer, D. and L. Pritchett (1999), "The Impact of Public Spending on Health: Does Money Matter?", *Social Science and Medicine*, No. 49, pp. 1309-23.

Galiani, S., P. Gertler and E. Schargrodsky (2005), "Water for Life: The Impact of the Privatization of Water Services on Child Mortality", *Journal of Political Economy*, Vol. 113, No. 1.

Glaeser, E. and B. Sacerdote (1999), "Why Is There More Crime in Cities?", *Journal of Political Economy*, No. 107, pp. 225-58.

Greene, W. (2003a), "Distinguishing Between Heterogeneity and Inefficiency: Stochastic Frontier Analysis of the World Health Organization's Panel Data on National Health Cares Systems", mimeo, New York University, www.stern.nyu.edu/~wgreene

———— (2003b), "Fixed Effects and Bias Due to the Incidental Parameters Problem in the Tobit Model", New York University, mimeo, www.stern.nyu.edu/~wgreene

Grosskopf, S. (1996), "Statistical Inference and Non-parametric Efficiency: A Selective Survey", *The Journal of Productivity Analysis*, Vol. 7, pp. 161-76.

Gupta, S. and M. Verhoeven (2001), "The Efficiency of Government Expenditure, Experiences from Africa", *Journal of Policy Modeling*, Vol. 23, pp. 433-67.

Haque, N., H. Pesaram and S. Sharma (1999), "Neglected Heterogeneity and Dynamics in Cross-country Savings Regressions", IMF, Working Paper.

Hanushek, E.A. (2002), "The Failure of Input-based Schooling Policies", National Bureau of Economic Research.

Hanushek, E.A. and J.A. Luque (2002), "Efficiency and Equity in Schools Around the World", National Bureau of Economic Research.

ICRG International Country Risk Indicators (2004), http://www.icrgonline.com

Jayasuriya, R. and Q. Wodon (2002), "Measuring and Explaining Country Efficiency in Improving Health and Education Indicators", mimeo, The World Bank, April.

———— (2003), "Explaining Country Efficiency in Improving Health and Education Indicators: The Role of Urbanization", background paper for the *World Development Report 2003*, The World Bank.

Jounard, I., M. Konsgrud, Y. Nam and R. Price (2004), "Enhancing the Cost Effectiveness of Public Spending: Experience in OECD Countries", OECD, Economics Department, Working Paper, No. 380, February.

Kaufmann, D., A. Kraay and P. Zoido-Lobaton (1999a), "Governance Matters", The World Bank, Policy Research, Working Paper, No. WPS2196.

———— (1999b), "Aggregating Governance Indicators", The World Bank, Policy Research, Working Paper, No. WPS2195.

———— (2002), "Governance Matters II – Updated Indicators for 2000-1", The World Bank, Policy Research, Working Paper, No. WPS2772.

Le Grand, J. (1987), "Inequalities in Health: Some International Comparisons", *European Economic Review*, No. 31, pp. 182-91.

Liese, B., N. Blanche and G. Dussault (2003), "The Human Resource Crisis in Health Services in Sub-Saharan Africa", background paper for the WDR 2004.

Maddala, G.S. (1987), "Limited Dependent Variable Models Using Panel Data", *The Journal of Human Resources*, Vol. XXII, No. 3.

Mathers, C.D., R. Sadana, J.A. Salomon, C.J.L. Murray and A.D. Lopez (2000), "Estimates of DALE for 191 Countries: Methods and Results", World Health Organization, GPE, Working Paper, No. 16.

Murillo-Zamorano, L. (2004), "Economic Efficency and Frontier Techniques", *Journal of Economic Surveys,* Vol. 18, No. 1, pp. 33-77.

Murray, C.J.L., J.A. Salomon and C. Mathers (1999), "A Critical Examination of Summary Measures of Population Health", World Health Organization, GPE, Discussion Paper, No. 2.

Nin, A., C. Arndt and P. Preckel (2003), "Is Agricultural Productivity in Developing Countries Really Shrinking? New Evidence Using a Modified Non-parametric Approach", *Journal of Development Economics*, No. 71, pp. 395-415.

Pigozzi, M. (2004), "Quality Education and HIV/AIDS", UNESCO, mimeo.

Ravallion, M. (2003), "On Measuring Aggregate Social Efficiency", The World Bank, Policy Research, Working Paper, No. 3166, November.

Sen, A. (1981), "Public Action and the Quality of Life in Developing Countries", *Oxford Bulletin of Economics and Statistics*, Vol. 43, No. 4, November.

Simar, L. and P. Wilson (2000), "Statistical Inference in Non-parametric Frontier Models: The State of the Art", *Journal of Productivity Analysis*, No. 13, pp. 49-78.

The State Failure Task Force (2004), "Political Instability (aka State Failure) Task Force Data", http://www.cidcm.umd.edu/inscr/stfail/sfdata.htm

Tanzi, V. (2004), "Measuring Efficiency in Public Expenditure", paper presented at the *Conference on Public Expenditure Evaluation and Growth*, The World Bank, October.

Varian, H. (1990), "Goodness-of-fit in Optimizing Models", *Journal of Econometrics,* No. 46, pp. 125-40.

Wheelock, D. and P. Wilson (2003), "Robust Non-parametric Estimation of Efficiency and Technical Change in U.S. Commercial Banking", Federal Reserve Bank of St. Louis, Working Paper, November.

Wilson, P. (2004), "A Preliminary Non-parametric Analysis of Public Education and Health Expenditures in Developing Countries", The World Bank, mimeo.

Wobst, P. and C. Arndt (2003), "HIV/AIDS and Primary School Performance in Tanzania", presented at the *25th International Conference of The International Association of Agricultural Economists*, August 16-22, Durban, South Africa.

World Bank (2004), *Making Services Work for the Poor. World Development Report 2004.*

Worthington, A. (2001), "An Empirical Survey of Frontier Efficiency Measurement Techniques in Education", *Education Economics*, Vol. 9, No. 3.

ASSESSING PUBLIC SECTOR EFFICIENCY: ISSUES AND METHODOLOGIES

Francisco Pedraja-Chaparro[*]*, Javier Salinas-Jiménez*[**] *and Peter C. Smith*[***]

Introduction

A more efficient public sector has become a universal target of central importance in economic policy. The increase in both the quantitative and qualitative relevance of the public sector within the economy (in terms of its size and functions), its contribution of budgetary discipline in the pursuit of macroeconomic stability and the difficulty of increasing public revenues are some of the reasons why attention is focusing on the public expenditure side of Public Finance.

A whole variety of initiatives ranging from privatisations to market simulations have been implemented in the pursuit of public sector efficiency. The introduction into the public sector of private sector management techniques such as decentralisation, management and performance measures, customer services, and so on comprise what is known as the New Public Management. These initiatives in the reform process have been implemented with varying levels of scope and intensity in western countries. The pioneers have been the UK and the USA, and the most demanding and comprehensive case has been New Zealand. The success of these initiatives to improve efficiency in the public sector depends crucially on the extent and confidence with which we are able to measure the performance of public services.

A central concern is to measure the relative efficiency of different public organisations providing the same public service. Two alternative approaches can be considered. The first is to develop a set of performance indicators, *i.e*. partial measures of some aspects in the behaviour of the organisation. The second is to try to develop a general index on the efficiency of the organisation. Although the first alternative has some virtues, its main flaws lie in its partial nature. As a consequence, contradictory results may arise, depending on the choice of indicator (Smith, 1990 and Smith and Goddard, 2003). The development of global efficiency scores seeks to overcome this weakness.

The traditional productivity literature characterises global measures of organisational efficiency as the distance of the unit under scrutiny from a frontier function, which is estimated using the best observed practice of the set of other similar units. There are two main methodologies for defining the frontier (Green and Coelli, 1998).

---

[*] Universidad de Extremadura (Spain).

[**] Instituto de Estudios Fiscales – Universidad de Extremadura (Spain).

[***] Centre for Health Economics – University of York (United Kingdom).

Parametric approaches that specify *a priori* a functional form with constant parameters to be estimated (*i.e*. Cobb-Douglas, translog, etc.). Efficiency is assessed in relation to this function and will be different depending on the chosen functional form. Stochastic-type frontiers are generally used where two components are identified in the residuals: inefficiency and all the other sources of error.

Non-parametric approaches that do not *a priori* specify a functional form, but nevertheless require certain assumptions about the structure of the production technology (e.g. free disposability, convexity, etc). This is done by solving a separate mathematical programming model for each observation. As in the parametric approach, the frontier changes, as will the estimated efficiency of each unit, in line with the assumptions made. This kind of approach generally is of a deterministic type, with all the distance from the frontier assumed to be caused by inefficiency.

Data Envelopment Analysis (DEA) is the dominant non-parametric technique in productivity analysis. Since it was introduced in 1978, it has undergone substantial theoretical development, and enjoyed a rapid growth in empirical applications in diverse fields, amongst which the public sector is very important.

The increased availability of data related to public sector performance and the ready availability of software enabling us to implement the DEA model raises the question of how to interpret efficiency scores obtained for different units for the purposes of improving their future performance. This paper therefore seeks to assess the usefulness of DEA in measuring the efficiency of a set of comparable units in the public sector. We must also note that our objective focuses exclusively on the extent to which DEA can yield an adequate measure of efficiency, and not on the implications of using such measures in a targets or reward scheme.

The paper is structured as follows. Some reflections are made on the special characteristics of public units' performance and the way in which these characteristics might influence the assessment of their efficiency. Secondly, the DEA model is briefly described. Finally, we offer some reflections on methodological issues that seem significant when assessing the efficiency of a set of public units through such non-parametric approaches.

1. Characteristics of "public supply" and efficiency

It is usual to start any discussion of public services by emphasising their unique characteristics, most especially the absence (or near absence) of any market in the conventional sense. When calculating the efficiency of a set of public units, we must therefore first briefly consider the special characteristics of public sector supply. This will shed light on some of the measurement and conceptual issues we are faced with.

A first restriction arises from the nature of the objective function for the public sector which is characterised by multiple criteria. In addition to efficiency,

public sector activities often try to achieve equity goals, and there often exists a trade off among these objectives. Diverse and conflicting objectives are components that must be borne in mind in any assessment of a public policy, to avoid hasty conclusions made when scrutinising only one of them. In addition to the existence of multiple objectives, the public sector differs from the private sector because of the diversity of principals (politicians, users, general public) that must be satisfied by agents (or 'bureaucrats', to use the terminology of the public choice literature). The multiplicity of tasks and principals causes serious problems when measuring public output.

Public output cannot be traded in the market, so it is difficult to define and to measure it. In practice, most of the time we resort to metrics in which aspects that are difficult to calculate (such as those related to the quality of the service) are not considered. The lack of a market prevents consumers from expressing a valuation of the services. Instead, the value of public output must often be inferred by observing public service activities. At best we can usually deal only with intermediate outputs (often with variables closer to measuring inputs than outputs) or mere proxies for final outputs (the services' effects on users). In addition to the measurement problems, there exist problems with attribution, in the sense that outputs may result from factors totally or partially out of the control of assessed units. In any case, the public sector is not a uniform body. Services and organisations of a different nature coexist, from the simplest ones to the most complex, where measurement and allocation problems appear at a different level.

The measurement and attribution issues lead to a monitoring and control problem. How should production be regulated? The regulatory problem is one of inferring optimal production in the absence of competitive pressure. In the absence of good regulation, public service units are likely to exhibit inefficient behaviour, in both an allocative and productive sense. However, the lack of competition can make the production technology uncertain and unknown. The question therefore is: does a comparison base exist that can furnish information on the technology?

Finally, there is in the public services an absence of the entry and exit options manifest in competitive markets. In particular, there is no guarantee that inefficient producers of public output are subject to the threat of bankruptcy that acts as a discipline in competitive markets.

The absence of competition, the monopolistic nature of public production and the absence of bankruptcy threat are some explanations for the difficulties in regulating public production. Moreover, or rather consequently, the schemes of internal incentives (positive or negative) found in most public services cannot guarantee efficient production. In the light of the preceding discussion, any measurement technique should be adapted to the following characteristics:

The lack of a market and the resulting difficulty of measuring the actual output make us use an intermediate output. Consequently, the technique must accommodate a measurement problem characterised by multiple output and input.

The technique should adjust to the characteristics of uncertainty surrounding public production technologies. Thus, it is advisable to use approaches which are flexible and do not require very strong assumptions on the production frontier.

The purpose of this paper is to assess the extent to which data envelopment analysis can assist regulators confronted by these difficulties, and seeking to assess the performance of public service institutions.

2. Data envelopment analysis (DEA)

DEA is a mathematical programming procedure developed by Charnes, Cooper and Rhodes (1978), based on the seminal work by Farrell (1957). The DEA model applies mathematical programming techniques to compare the efficiency of a set of units. DEA may be seen as an extension of the traditional output/input ratio analysis. The efficiency score of each unit can be represented as a ratio of the total weighted outputs to the total weighted inputs:

$$\text{Efficiency} = \frac{\text{Weighted sum of outputs}}{\text{Weighted sum of inputs}}$$

Mathematically, if we consider a set of n units consuming m inputs ($x_1 \ldots x_m$) and producing s outputs ($y_1 \ldots y_s$), the efficiency of a unit, say unit 0, can be measured as follows:

$$\text{Maximize } h_0 = \frac{\sum_{r=1}^{s} U_{ro} Y_{rj}}{\sum_{r=1}^{m} V_{io} X_{ij}} \leq 1$$

$$\text{subject to: } \frac{\sum_{r=1}^{s} U_{ro} Y_{ro}}{\sum_{r=1}^{m} V_{io} X_{io}} \qquad j = 0, 1, 2, \ldots, n$$

$$u_{ro} \geq 0; \qquad r = 1, 2, \ldots, s$$

$$v_{io} \geq 0; \qquad i = 1, 2, \ldots, m$$

By solving this problem, it is possible to calculate, for each one of the units, the set of inputs and outputs weights with which the unit may obtain the maximum efficiency score, with the restriction that using the same weights no other unit can achieve an efficiency score higher than one.

If, subject to this restriction, it is possible to find a set of weightings in which the efficiency ratio of the unit analysed equals one, that unit will be considered as

efficient. Otherwise, the unit will be evaluated as inefficient since, even if a more favourable set of weights is considered, one could find another unit obtaining a greater efficiency ratio. In assessing efficiency in this way, weightings assigned to inputs and outputs will vary by unit. As pointed out by Sexton (1986), since each unit uses differing combinations of inputs and outputs, a different set of weightings will in general be selected for each – that is, the set allowing the unit to obtain the greatest efficiency ratio. Therefore, the method assesses each unit in the most favourable light.

Presenting DEA as a weighted sum of outputs in relation to a weighted sum of inputs allows one to characterize the technique as an extension of the analysis of ratios and, as pointed out in the introduction, to approach efficiency from a global point of view.

The maximization problem specified above can be presented in a linear form, which is more convenient for solving, as follows:

$$\text{Maximize } \sum_{r=1}^{s} u_r y_{r0}$$

$$\text{subject to: } \sum_{i=1}^{m} v_i x_{i0} = 1$$

$$\sum_{r=1}^{s} \mu_r \gamma_{rj} - \sum_{i=l}^{m} v_i x_{ij} \leq 0 \qquad j = 1, 2, \ldots, n$$

$$u_{ro} \geq 0; \qquad r = 1, 2, ..., s$$
$$v_{io} \geq 0; \qquad i = 1, 2, ..., m$$

The dual problem is:

$$\text{Minimize } \theta_0$$

subject to:

$$\sum_{j=1}^{n} x_{ij} . \lambda_j - x_{i0} \theta_0 \leq 0 \qquad i = 1, 2, \ldots, m$$

$$\sum_{j=1}^{n} y_{rj} . \lambda_j \geq y_{r0} \qquad r = 1, 2, \ldots, s$$

with trivial restrictions $\theta_0, \lambda_j \geq 0 \qquad j = 1, 2, ..., n$

The interpretation of this programme is very simple, as with the fractional programme. The aim is to search, for each assessed unit, a linear combination of other units that produces at least the same quantity of output in each of the *s* dimensions considered by consuming a lesser proportion ($0<\theta<1$) of the *m* inputs. If this is not possible, the unit is efficient. On the other hand, if it is possible, the unit will be assessed as inefficient since there are other units in the sample (those from the reference group against which it is compared) performing better.

3. Assessment of efficiency in the public sector: some methodological reflections

Because *a priori* the production technology is unknown and therefore there is no single optimal approach towards assessing the efficiency of a set of productive units, we hesitate to advocate one particular analytic technique over another. In general, the characteristics of the sector analysed, together with the restrictions of information and the purpose of the analysis will determine the most suitable technique. Nevertheless, DEA seems suitable for the multidimensional character of public output, and its flexibility is particularly attractive given the lack of knowledge and the uncertainty involved in the public sector production process.

In this section we therefore discuss some of the most significant methodological issues in public service efficiency measurement that arise when using a non-parametric approach. In particular, we will analyse issues derived from:

1) the fact that the units may have different objectives;
2) from exogenous and environmental factors (out of the control of units) which may considerably influence their results; and
3) from limitations that arise from the deterministic and non-parametric nature of the DEA model.

3.1 Diversity of objectives and weights

One of the main strengths of DEA is that it enables us to analyse the global performance of the assessed units. In order to assess the efficiency of the units, we have to assume that they are homogenous, consuming the same inputs, producing the same outputs and trying to achieve similar objectives. However, in general, the objectives and outputs are not well defined and, even when they are, it is difficult to quantify them in practice. The analyst often has to turn to the only available data that approximately inform her on the performance of units. In this sense, let us think of the difficulties in areas such as education, health care or justice where final outputs might be the value added in schools, the health gains in health care or the protection granted by courts. The immense difficulties involved in directly or indirectly measuring these outcomes usually leads to the use of proxies such as pupils' results, the number of patients or the number of resolved cases respectively.

The case of decentralised services is particularly complex because we cannot guarantee uniformity in the objectives of the units under scrutiny. Indeed, some of the major gains in efficiency arising from decentralisation are said to derive from a better adjustment to differentiated preferences for these public services in the different jurisdictions, which contradicts the above identity of objectives assumed in DEA.

However, although the objectives are assumed to be universal in DEA, the weights attached to each input and output are calculated so that each assessed unit receives the most favourable possible treatment. This implies that there are no *a*

priori values or restrictions set on the different weights and that weights attached to the different inputs and outputs may differ when assessing the various units. Yet, although consistent with notions of decentralisation, total weight flexibility, considered as a significant advantage of DEA (Cooper *et al.*, 2000), may be criticised for various reasons:

- if the inputs and outputs included in the analysis are not equally important it is not sensible to claim that a unit is relatively efficient when the weights to the important inputs and outputs are zero. The total flexibility of the DEA model may lead to an unfounded emphasis on the efficient use of relatively unimportant factors, concealing inefficiencies in the most important activities carried out by the units;
- if we use the unbounded model, some inputs and outputs may be ignored in the analysis when assessing the relative efficiency of some units. As a result, the relative efficiency of a unit may not really reflect its performance in relation to inputs and outputs taken as a whole;
- the implicit assumption made when allowing weight flexibility in the unbounded DEA model is that the analyzed units may have individual objectives and particular circumstances that should be considered when assessing them. However, since the units compared using DEA are homogeneous units, it may be unacceptable to assume that the relative importance attached to the different inputs and outputs by each unit should differ greatly.

Therefore, there seems to be a strong case for considering the introduction of weight restrictions. The research has focused on technical aspects, setting the limits between which the weights can vary and evaluating how the introduction of weight restrictions may improve the results of the DEA model (See Dyson and Thanassoulis, 1988 and Pedraja-Chaparro *et al*., 1997).

3.2 Exogenous factors

When the efficiency of a set of public units is assessed, the conventional DEA model implies that there is no factor outside the control of the units in charge of providing the public service. This assumption is quite often very far from what actually happens in most public services.

We must distinguish between environmental and exogenous factors among those ones which are outside the control of productive units.

Environmental factors are not directly involved in the production process although they may provide useful information on how to explain different efficient behaviours. These are, among others, the kind of ownership, the degree of competition, geographical factors, etc.

We must particularly deal with exogenous factors that affect the production process but not entirely under the control of units. It would not make any sense to introduce an objective for some unit (*i.e.* a percentage decrease in inputs consumption for it to be efficient), if these inputs were beyond the control of the

unit. This type of factor is often found in a number of public services as, for instance, in educational services when pupils' skill and social economic background are more important determinants of results than the resources consumed by schools. DEA has shown a noteworthy adaptability to include these exogenous factors in efficiency assessment.

There are two broad approaches to including factors beyond the control of units. On the one hand, we have the one-stage approach, where the exogenous factors are included jointly with all the others inputs that can be controlled by the productive units. In this way, only one DEA analysis is run, in which all the inputs are included together. The principal model following this approach is that proposed by Banker and Morey in 1986. Its main advantage is that it enables us to introduce all relevant variables in a DEA single analysis, which simplifies to a large extent the calculation of efficiency indices. However this methodology has shortcomings. Units operating in the most disadvantaged circumstances will automatically be deemed efficient regardless of their performance (because there are no direct comparators). More generally, the increased number of variables introduced into the DEA model reduces its power to discriminate between units.

The second alternative is a multi-stage analysis. These models consist of several analytic stages. All of them have in common a first stage in which we only include those inputs that units can control. Afterwards, some adjustments on initial efficiency scores are made, avoiding biases that would lead to benefit unfairly units working in a relatively more favourable context. We can use different methodological options:

- *Two-stage models*. In these models, efficiency scores calculated in the first stage are included as dependent variables in a regression where explanatory variables are non-discretionary inputs. Although there are many ways to undertake this regression, the methodology proposed by Ray in his study of Connecticut schools in 1991 may be highlighted. Following this methodology corrected ordinary least squares are used to obtain consistent estimators of parameters. Their major appeal lies in that this correction guarantees that the units with the worst supply of non-controllable inputs enjoy the largest upwards adjustments.
- *Three-stage model* (Fried and Lovell, 1996). In this model, total slacks obtained in the first stage are included in a second DEA as controllable inputs, whereas outputs are non-controllable inputs. The aim of this second analysis is to identify the part of slacks that can be explained by exogenous factors and the part which reflects technical inefficiency. After separating both influences, the initial values of inputs and outputs are adjusted and then a third DEA model is run using these adjusted values.
- *Four-stage model* (Fried, Schimidt and Yaisawarng, 1999). This methodology can be considered as a mix between the three-stage model (since it also uses total slacks) and the two-stage approach (because slacks are included as dependent variables in a regression with non-controllable factors as explanatory variables). However, in this case a Tobit regression is used instead of ordinary least squares and only inputs slacks are included in regressions, one for each variable.

These models represent rational extensions to the basic DEA model. However, they too can be criticised. In our opinion, an important restriction comes from the possible bias in the results when there is a correlation between the inputs included in the fist stage and the independent variables considered in the second (see Chalos, 1997).

Moreover, two-stage and four-stage models have notable disadvantages, put forward by Simar and Wilson (2003). Specifically, there are problems related to the fact that DEA efficiency estimates are dependent in the statistical sense (they are computed using linear programming techniques) and, consequently, standard approaches to inference are invalid. They suggest employing bootstrap methods in order to overcome these problems.

Thus, in spite of the versatility and adaptability of DEA in handling non-discretionary inputs, there is no generally accepted methodology as to the appropriate way to introduce them when measuring efficiency. The analyst must often use judgement in the light of the characteristics of the specific area of application, data availability and a search for simplicity. This is especially relevant if very different results are obtained when applying the alternative approaches to the same sample. (See Cordero, Pedraja-Chaparro and Salinas-Jiménez, 2004).

3.3 Problems derived from the non-parametric and deterministic nature of DEA

One of the most serious shortcomings of DEA arises from the non-parametric and deterministic nature of the model. In this sense, the following issues are especially relevant:

i) the sensitivity of the results to model specification;

ii) the use of inappropriate data;

iii) the fact that efficiency estimates are point estimations; and

iv) the lack of adequate techniques for treating missing data.

a) Selection of variables and the specification of the model

The analyst faces two fundamental choices when assessing a set of units using DEA: on the one hand, the selection of variables that must be included in the efficiency analysis; and on the other, the type of returns to scale (constant or variable ones) that must be considered in the production function.

Given the deterministic and non-parametric nature of DEA, the choice of variables is a crucial decision that may considerably affect the results obtained in the analysis. As opposed to econometric models, where the analyst can use tests such as the t-test, one cannot apply any model selection test to DEA, and the researcher does not know if the results are robust or if they exclusively arise from the choice of variables used in the analysis.

In this sense, some studies have focused on the comparison of the results of a set of alternative models in order to prove the "validity" of the efficiency estimations. Thus, Gong and Sickles (1992) and Banker, Gadh and Gorr (1993), among others, compare efficiency indices obtained by DEA with those resulting from the application of alternative approaches. Nevertheless, this validation process of results has severe limitations. This is due to the fact that the results that are compared are derived from approaches based on very different assumptions related to the frontier production. Other authors have made a sensitivity analysis of results by calculating efficiency indices with several sets of variables and specifications. This is the path followed by Färe, Grosskopf and Weber (1989) when analysing educational centres and by Valdmanis (1992) when assessing a set of hospitals. However, if the results obtained were sensitive to the specification of the model it would not yet be clear what should be done, apart from relying on the analyst's common sense, bearing in mind the non-parametric nature of technique.

With the same aim, but on the theoretical level, Smith (1997) uses diverse models with simulated data in order to analyse the effects derived from model misspecification in DEA. The major conclusion is that errors derived from model misspecification are larger when the model is simple (*i.e.* with a small number of variables) and the sample is small; in such circumstances, it would be better to include non-relevant variables than leave major variables out of the model.

As regards to the type of returns to scale, it must be pointed out that in order to ensure the homogeneity of units studied in comparisons, the DEA model enables us to specify the type of returns to scale, including this assumption on the building of an efficient frontier. This aspect turns out to be crucial because when it is not considered we would mistake some inefficiency for scale problems. For instance, a wrong use of an assumption of variable returns to scale may favour units that operate at unusually large or small scale, making them incorrectly assessed as efficient. Thus, the first issue the analyst must face, in this sense, is to determine the type of returns to scale in order to estimate the production frontier. Previous studies on the sector and its characteristics may provide a first hint of the type of returns to scale that must be considered. Also, there are several alternative methods in the empirical literature that aim at contrasting the validity of the assumption on returns to scale. Among these alternatives are the following:

- to assume constant returns to scale and analyse the relation between efficiency indices and the size of the assessed units. For this, usually the Tobit model is used because the values of the dependent variable (efficiency indices) are included in the interval 0-1.
- to compare the similarity between results obtained under the assumption of constant returns to scale and that of variable returns, thus calculating in this way possible scale inefficiencies and in consequence finding the most convenient type of assumption for the analysis.

In those cases, where the production function is better known and the production process is simpler, one may use (as a complement) other types of parametric approaches in order to contrast the kind of returns to scale.

b) Inappropriate data

With regard to the use of inappropriate data, major problems arising in empirical papers result from measurement errors or outliers that may distort the efficiency analysis and from the relative small number of observations.

The problem of measurement errors may be abated, provided that they are occasional and are not repeated in the same units reiteratively, making the efficiency analysis multi-period; *i.e*. by repeating the efficiency analysis for different periods of time with one of the following tools:

- analysis of several periods of time and presentation of average results;
- calculation of average values for each variable in some period of time and subsequent efficiency assessment;
- use of the window analysis, a more sophisticated approach that involves considering observations of a same unit in different periods of time as if they were distinct units one from another.

In the public sector, a multi-period analysis is desirable not only because of the non-stochastic nature of the DEA model, but also because of the nature of the expenditure programmes surveyed. When analysing sectors such as education and health services, where the resources used may have medium and long-term effects, it would be advisable that the efficiency analysis were referred to relatively long periods of time.

Naturally, the data with greater impact on the efficiency analysis is that related to units on the frontier (efficient units) as they may affect the assessment of some inefficient units in the sample. For this reason, a range of more or less elaborated methods have been designed to detect which units, among those assessed as efficient, are especially influential. Among these methods we may highlight:

- superefficiency indices, built by Andersen and Petersen (1993) and Wilson (1995) which indicate to which extent an efficient unit is far from the frontier made up by the other units. A very high super efficiency index may be *prima facie* a cause for questioning the efficiency of this unit and for carrying out a particular scrutiny of its data.
- calculation of the number of times an efficient unit appears in reference groups of inefficient units; this device is a basic way to assess the influence of a unit on the others, paying special attention to measurement errors made by the most influential units.

As regard to the relatively small number of units in relation to the number of variables included in the efficiency analysis, there may be a very significant reliability problem referred to the results obtained through the DEA model. The number of free dimensions decreases as new variables are included and, consequently, it is more likely that each unit may be considered efficient because of the model flexibility. In spite of the significance of these aspects, only a small number of papers have dealt with issues related to the selection process of variables

and the analysis of degrees of freedom which must exist so that the results in the study may be considered reliable.

Unfortunately, in most empirical studies the choice of variables usually is too dependant on available data; and subsequent haste to apply the technique obviates basic theoretical aspects that should be considered. The criterion suggested by Banker *et al.* (1989) is usually used. They point out, as a rule of thumb, that the number of assessed units should be at least three times the number of variables included in the model. This rule, although it is an *ad hoc* criterion without any theoretical or empirical basis, has been used in many applied studies and considered as a valid criterion to ensure the reliability of results obtained, irrespective of the objectives of the efficiency analysis. Pedraja-Chaparro *et al*. (1999) show how the reliability of results obtained with the DEA model depends not only on the number of observations and variables but also on the distribution of actual efficiencies and on the correlation between inputs and outputs. Generally, Banker's rule, which considers only two of these four aspects, seems too generous. Nevertheless, it may be useful if the objective of the efficiency analysis is to obtain global information from one sector (average efficiency). But Banker's rule may lead to making mistakes if more detailed information is needed (such as efficiency rankings or individual efficiency indices of the different units). Some authors, such as Adler and Golany (2001), have decided to use principal components as outputs and inputs, making it possible to include information from a large set of variables but reducing problems related to the lack of degrees of freedom.

c) Point estimations

Another significant restriction of the DEA model, derived from its non-parametric nature is that the technique only enables us to obtain point efficiency indices of the units. So it is not possible to analyse if the differences between two estimates are statistically significant or make inferences. Wilson and Simar, Löthgren and Tambour (1997) and Simar and Wilson have proposed in recent years the use of "bootstrapping" techniques in order to overcome this restriction and build confidence intervals for the efficiency indices in order to make more accurate comparisons between the assessed units. In the same sense, Ferrier and Hirschberg (1997 and 1999) have proved that bootstrapping techniques may be applied to DEA efficiency scores. Various empirical papers dealing with efficiency measurement in the public sector have used this approach, among which the efficiency assessment in British primary care centres (Giuffrida, 1999) or the study on community care in England (Salinas-Jiménez, Pedraja-Chaparro and Smith, 2003) may be highlighted.

d) Missing data

Finally, due to the nonparametric nature of the model, missing data are a significant problem in DEA. In many empirical applications, blank entries for the data matrices are directly eliminated before the efficiency analysis, However, units

with missing data could be highly useful as reference or benchmark units, which span the efficient frontier and eliminating them may distort the efficiency evaluation. Only a few attempts to solve this problem can be found in the DEA literature. Among them, the paper by Kuosmanen (2002) uses dummy variables (zero for missing outputs; number large enough for inputs) and runs a DEA model with weight restrictions in such a way that the black entries are not considered. An alternative approach is proposed by Kao and Liu (2000) who use fuzzy sets to model the ranges for missing data.

4. Conclusions

The theoretical underpinnings DEA have developed to an extraordinary extent since the publication of the initial 1979 paper. However, there remain unresolved many conceptual and operational difficulties associated with the technique, as summarised in this paper. We nevertheless believe that DEA can furnish regulators with useful insights into the performance of public service organisations, if used with discretion, and viewed in conjunction with other techniques. It is especially useful for exploring complex datasets and identifying exceptionally good or poor performers. In its current state of development, it is not suitable for making definitive judgements on organisations or setting detailed targets.

REFERENCES

Adler, N. and B. Golany (2001), "Evaluation of Deregulated Airline Networks Using Data Envelopment Analysis Combined with Principal Components Analysis with an Application to Western Europe", *European Journal of Operational Research,* No. 132, pp. 260-73.

Andersen, P. and N.C. Petersen (1993), "A Procedure for Ranking Efficient Units in Data Envelopment Analysis", *Management Science*, Vol. 30, No. 10, pp. 1261-64.

Banker, R.D., A. Charnes, W.W. Cooper, J. Swarts and D.A. Thomas (1989), "An Introduction to Data Envelopment Analysis with Some of Their Models and its Uses", *Research in Governmental and Nonprofit Accounting*, No. 5, pp. 125-63.

Banker, R.D., V.M. Gadh and W.L. Gorr (1993), "A Monte Carlo Comparison of Two Production Frontier Estimation Methods: Corrected Ordinary Least Squares and Data Envelopment Analysis", *European Journal of Operational Research,* No. 67, pp. 332-43.

Banker, R.D. and R.C. Morey (1986), "Efficiency Analysis for Exogenously Fixed Inputs and Outputs", *Operations Research,* Vol. 34, No. 4.

Bosch, N., F. Pedraja-Chaparro and J. Suárez Pandiello (2000), "Measuring the Efficiency of Spanish Municipal Refuse Collection Service", *Local Government Studies,* Vol. 26, No. 3, pp. 71-90.

Coelli, T., D. Rao and G. Battese (1998), "An Introduction to Efficiency and Productivity Analysis", Boston, Kluwer Academic Publishers.

Cordero, J.M., F. Pedraja-Chaparro and J. Salinas-Jiménez (2004), "Efficiency Measurement in Education and Non-discretionary Inputs", presented at the *North American Productivity Workshop 2004*, Toronto, Canada, June.

Chalos, P. (1997), "An Examination of Budgetary Inefficiency in Education Using Data Envelopment Analysis", *Financial and Accountability & Management*, No. 13, pp. 55-69.

Charnes, A., W.W. Cooper and E. Rhodes (1978), "Measuring the Efficiency of Decision Making Units", *European Journal of Operational Research,* Vol. 2, pp. 429-44.

Cooper, W.W., L.M. Seiford and K. Tone (2000), "Data Envelopment Analysis: A Comprehensive Text With Models, Applications and References", Boston, Kluwer.

Cullis, J.G. and P.R. Jones (1987), "Microeconomics and the Public Economy: A Defence of Leviathan", Blackwell.

Dixit, A. (2002), "Incentives and Organizations in the Public Sector: An Interpretive Review", *Journal of Human Resources,* Vol. 37, No. 4, pp. 696-727.

Downs, A. (1957), "An Economic Theory of Democracy", New York, Harper and Row.

Dyson, R. and E. Thanassoulis (1988), "Reducing Weight Flexibility in Data Envelopment Analysis", *Journal of the Operational Research Society*, Vol. 39, No. 6, pp. 563-76.

Färe, R., S. Grosskopf and W.L. Weber (1989), "Measuring School District Performance", *Public Finance Quartely,* Vol. 17, No. 4, pp. 409-28.

Farrell, M.J. (1957), "The Measurement of Productive Efficiency", *Journal of the Royal Statistical Society* (A), Vol. 3, pp. 253-90.

Ferrier, G.D. and J.G. Hirschberg (1997), "Bootstrapping Confidence Intervals for Linear Programming Efficiency Scores: With an Illustration Using Italian Banking Data", *Journal of Productivity Analysis,* Vol. 8, pp. 19-33.

Ferrier, G.D. and J.G. Hirschberg (1999), "Can We Bootstrap DEA Scores?", *Journal of Productivity Analysis,* Vol. 11, pp. 81-92.

Frantz, R. (1988), "X-Efficiency: Theory, Evidence and Applications", Boston, Kluwer.

Fried, H.O. and C.A.K. Lovell (1996), "Searching for the Zeds", Georgia Productivity Workshop.

Fried, H.O., S. Schmidt and S. Yaisawarng (1999), "Incorporating the Operating Environment into a Non-parametric Measure of Technical Efficiency", *Journal of Productivity Analysis*, No. 12, pp. 249-67.

Fried, H.O., C.A.K. Lovell, S.S. Schmidt and S. Yaisawarng (2002), "Accounting for Environmental Effects and Statistical Noise in Data Envelopment Analysis", *Journal of Productivity Analysis*, Vol. 17, pp. 157-74.

Green, W.H. (1993), "The Econometric Approach to Efficiency Analysis", in H.O. Fried, C.A.K. Lovell and S.S. Schmidt (eds.), *The Measurement of Productive Efficiency: Techniques and Applications*, New York, Oxford University Press.

Gong, B.H. and R.C. Sickles (1992), "Finite Sample Evidence on the Performance of Stochastic Frontiers and Data Envelopment Analysis Using Panel Data", *Journal of Econometrics,* Vol. 51, pp. 259-84.

Kao, C. and S.T. Liu (2000), "Data Envelopment Analysis with Missing Data: An Application to University Libraries in Taiwan", *Journal of the Operational Research Society,* Vol. 51, No. 8, pp. 897-905.

Koopmans, T.C. (1957), Three Essays on the State of Economic Science, New York, McGraw Hill.

Kuosmanen, H. (2002), "Modelling Blank Data Entries in Data Envelopment Analysis", Working Paper, Department of Social Sciences, Wageningen University, The Netherlands.

Legrand, J. (1992), "The Theory of Government Failure", *British Journal of Political Science,* Vol. 21, pp. 423-42.

Löthgren, M. and M. Tambour (1997), "Bootstrapping the Data Envelopment Analysis Malmquist Productivity Index", *Applied Economics,* Vol. 31, No. 4, April.

Maniadakis, N., B. Hollingsworth and E. Thanassoulis (1999), "The Impact of the Internal Market on Hospital Efficiency, Productivity and Service Quality", *Health Care Management Science*, Vol. 2, No. 2, pp. 75-85.

Mueller, D.C. (1989), *Public Choice II*, Cambrige, Cambrige University Press.

Musgrave, R.A. (1981), "Leviathan Cometh: Or Does He?", in H.F. Ladd and T.N. Tideman (eds.), *Tax and Expenditures Limitations*, Washington (D.C.), Urban Institute Press.

Niskanen, W.A. (1971), *Bureaucracy and Representative Government*, Chicago and New York, Aldine Atherton.

Pedraja-Chaparro, F., J. Salinas-Jiménez and P. Smith (1997), "On the Role of Weight Restrictions in Data Envelopment Analysis", *Journal of Productivity Analysis,* Vol. 8, No. 2, pp. 215-30.

—————— (1999), "On the Quality of the Data Envelopment Analysis", *Journal of the Operational Research Society.*

Pestieau, P. and H. Tulkens (1990), "Assessing the Performance of Public Sector Activities: Some Recent Evidence from the Productive Efficiency Viewpoint", Discussion Paper, No. 9060, November, Center for Operations Research & Econometrics, Université Catholique de Louvain.

Ray, S.C. (1991), "Resource Use Efficiency in Public Schools: A Study of Connecticut Data", *Management Science*, Vol. 37, No. 12, pp. 1620-28.

Salinas-Jiménez, J., F. Pedraja-Chaparro and P. Smith (2003), "Evaluating the Introduction of a Quasi-market in Community Care: Assessment of a Malmquist Index Approach", *Socio Economic Planning Science* (2003), Vol. 37, No. 1, pp. 1-13.

Sexton, T.R. (1986), "The Methodology of Data Envelopment Analysis", in R. Silkman (ed.), *Measuring Efficiency: An Assessment of Data Envelopment Analysis*, San Francisco, Jossey Bass.

Smith, P. and M. Goddard (2003), "Los Indicadores de Gestión en el Sector Público: Fortalezas y Debilidades", Papeles de Economía Española, No. 95, pp. 35-47.

Simar, L. and P.W. Wilson (1998), "Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models", *Management Science*, Vol. 44, No.1, pp 49-61.

———————— (2003), "Estimation and Inference in Two-stage, Semiparametric Models of Production Processes", Discussion Paper, No. 0307, Institut de Statistique, Université Catholique de Louvain.

Stiglitz, J.E. (1989), *The Economic Role of the State*, Basil Blackwell, (v. c. Instituto de Estudios Fiscales, 1993).

Smith, P. (1997), "Model Misspecification in Data Envelopment Analysis", *Annals of Operations Research*, Vol. 73, pp. 233-52.

Tulkens, H. (1993), "On FDH Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Courts and Urban Transit", *Journal of Productivity Analysis,* Vol. 4, No. 1-2, pp. 183-210.

Valdmanis, V. (1992), "Sensitivity Analysis for DEA Models – An Empirical Example Using Public Vs. NFP Hospitals", *Journal of Public Economics,* Vol. 48, pp. 185-205.

Wilson, P.W. (1989), *Bureaucracy*, New York, Basic Books.

———————— (1995), "Detecting Influential Observations in Data Envelopment Analysis", *Journal of Productivity Analysis*, No. 6, pp. 27-45.

Wilson, P.W. and L. Simar (1995), "Bootstrap Estimation for Non-parametric Efficiency Estimates", Discussion Paper, No. 9517, Institut de Statistique, Université Catholique de Louvain.

Wolf, C. (1979), "A Theory of Non-market Failure", *Journal of Law and Economics,* Vol. 22, No. 1, pp. 107-39.

———————— (1987), "Market and Non-Market Failures: Comparison and Assessment", *Journal of Public Policy,.* Vol. 7, No.1, pp. 43-70.

———————— (1988), *Markets or Governments: Choosing Between Imperfect Alternatives*, Cambridge (Mass.), M.I.T. Press.

ASSESSING EDUCATION AND HEALTH EFFICIENCY IN OECD COUNTRIES USING ALTERNATIVE INPUT MEASURES

António Afonso and Miguel St. Aubyn[*]

1. Introduction and motivation

Economics is often defined as the science that studies the allocation of resources to alternative uses, so that some concept of satisfaction is maximised. Moreover, economists make a useful distinction between resources that are the "ingredients", or inputs, that allow the production of some other goods or services, sometimes final, the outputs. The relationship between inputs and outputs is usually rationalised in terms of a production function, and, in many theoretical and empirical work, it is assumed that this relationship or function holds for all production units involved.

In this paper, we do not assume that all units produce the maximum quantities allowed by input resources available to them. On the contrary, we are especially interested in allowing for, and providing estimates of, possible inefficiencies in production. These inefficiencies are, in fact, distances to a production frontier. They are a measure of what is lost when inputs are not put into the best of possible uses, so that output is lower than the one that could be attained.

Following Afonso and St. Aubyn (2005a), we assess efficiency in providing education and health services across OECD countries. These are two sectors where public expenditure is of great importance, and also determinant for economic growth and welfare.[1] Therefore, if there are important inefficiencies in one country, this may well mean that either education or health provision could be improved significantly without more pressure on the public purse, or else that resources could be freed to alternative uses, be they public or private.

Results presented in Afonso and St. Aubyn (2005a) are based in physically measured inputs. For example, education inputs are the number of teachers per 100 students and hours per year students spend at school. Here, we compare results using physically measured inputs to results attained when inputs are measured in financial

---

[*] António Afonso: ISEG/UTL – Technical University of Lisbon; CISEP – Research Centre on the Portuguese Economy, R. Miguel Lupi 20, 1249-078 Lisbon, Portugal, E-mail: aafonso@iseg.utl.pt. European Central Bank, Kaiserstraße 29, D-60311 Frankfurt am Main, Germany, E-mail: antonio.afonso@ecb.int

Miguel St. Aubyn: ISEG/UTL - Technical University of Lisbon; UECE – Research Unit on Complexity in Economics, R. Miguel Lupi 20, 1249-078 Lisbon, Portugal, E-mail: mstaubyn@iseg.utl.pt

We are grateful to participants at the Workshop in Public Finances of the Bank of Italy held in Perugia, 31 March-2 April 2005, for useful comments. The opinions expressed herein are those of the authors and do not necessarily reflect those of the author's employers.

[1] According to Afonso and St. Aubyn (2005a), Table 1, public expenditure on education and health averaged 88.4 and 72.2 per cent of total expenditure in those activities, respectively, in 2000, in OECD countries. Also, verage total expenditutre on education and on health was equal to 5.4 an 8.0 of GDP.

terms, *i.e.,* expenditure on education or health in monetary units. As it will be shown later, estimated inefficiencies are not the same when these two different approaches are followed. In the conclusion to this paper, we develop some possible explanations to this, and defend that physically measured inefficiencies are probably more meaningful.

Not many other authors have previously studied public expenditure inefficiency in an international and aggregate framework. Fakin and Crombrugghe (1997) and Afonso, Schuknecht and Tanzi (2003) have done so for public expenditure in the OECD, Clements (2002) for education spending in Europe, Gupta and Verhoeven (2001) for education and health in Africa and St. Aubyn (2002, 2003) for health and education expenditure in the OECD. Although these studies use methods similar to ours (Free Disposable Hull or Data Envelopment Analysis, to be described later), inputs are always measured in monetary terms only.

Note that our purpose is to measure inefficiency across countries, and not to explain it. For an attempt to do the latter, we refer the interested reader to Afonso and St. Aubyn (2005b).

The rest of the paper is organised as follows. Section 2 presents our methodology. Section 3 explains how we measure the education and health inputs and outputs. Section 4 presents the empirical results and section 5 contains our concluding remarks.

2. Methodology

We apply two different non-parametric methods – Free Disposable Hull (FDH) analysis and Data Envelopment Analysis (DEA). These methods have been developed and applied to Decision Making Units (DMUs) that convert inputs into outputs.[2] These units may include public organisations, such as hospitals, schools, universities, local authorities or regional governments.[3]

2.1. Free Disposable Hull (FDH) analysis

We apply a so-called FDH analysis, which is a non-parametric technique first proposed by Deprins, Simar, and Tulkens (1984). Suppose that under efficient conditions, the education or health status of a population *i*, measured by an

---

[2] See Coelli, Rao and Battese (1998), Sengupta (2000) and Simar and Wilson (2003) for an introduction to this literature.

[3] De Borger and Kerstens (1996) provide an analysis of Belgian local governments, Coelli (1996) assess the efficiency of Australian universities, Afonso and Fernandes (2005) study the efficiency of local municipalities in the Lisbon region, Afonso and Scaglioni (2005) analyse the efficiency of Italian regions, while Afonso and Santos (2005) investigate the performance of public tertiary education in Portugal.

indicator y_i, the output, depends solely on education or health expenditure per habitant, x_i, the input:

$$y_i = F(x_i) \tag{1}$$

If $y_i < F(x_i)$, it is said that country *i* exhibits inefficiency. For the observed expense level, the actual output is smaller than the best attainable one. FDH is one of the different methods of estimating function *F*, the efficiency frontier.

In a simple example, four different countries display the following values for indicator *y* and expense level *x*:

Table 1

Fictitious Values for Countries A, B, C and D

	Indicator	**Expenditure**
Country A	65	800
Country B	66	950
Country C	75	1,000
Country D	70	1,300

Expenditure is lower in country A (800), and the output level is also the lowest (65). Country D exhibits the highest expenditure (1300), but it is country C that attains the best level of output (75).

Country D may be considered inefficient, in the sense that it performs worse than country C. The latter achieves a better status with less expense. On the other hand, countries A, B or C do not show as inefficient using the same criterion.

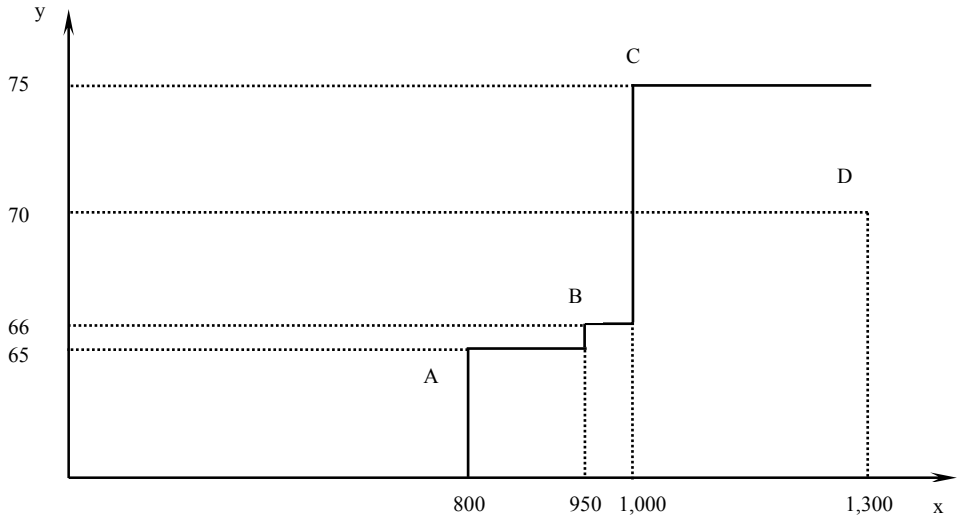
In FDH analysis, countries A, B and C are supposed to be located on the efficiency frontier.[4] This frontier takes the following form in this example:

---

[4] Of course, it could still be the case that there are inefficiencies in those countries, in the sense that they could improve outcomes without increasing resources used. The point here is that there are no other countries in the sample that provided evidence this is so. As in a court, a country is presumed efficient till inefficiency evidence is provided.

Figure 1

FDH Frontier



$$y = F(x) = \begin{cases} 65, & 800 \le x < 950 \\ 66, & 950 \le x < 1000 \\ 75, & 1000 \le x \le 1300 \end{cases} \tag{2}$$

This function is represented in Figure 1.

It is possible to measure country D inefficiency, or its efficiency scores, in two different ways:

i) inefficiency may be measured as the vertical distance between point D and the efficiency frontier. Here, one is evaluating the difference between the output level that could have been achieved if all expense was applied in an efficient way, and the actual level of output. In this example, the efficiency loss equals 5 – country D should, at least, achieve the same indicator level as country C, under efficient conditions.

ii) if one computes the horizontal distance to the frontier, the efficiency loss is now 300, in units of expense. It can be said that efficiency losses in country D are about 24 per cent (=300/1300) of total expense. To attain an indicator level of 70, it is necessary to spend no more than 1000, as shown by country C.

FDH analysis is also applicable in the multiple input-output cases. We sketch here how this is done, supposing the case of *k* inputs, *m* outputs and *n* countries.[5]

---

[5] The interested reader may refer to Gupta and Verhoeven (2001) and to Simar and Wilson (2003).

For country i, we select all countries that are more efficient – the ones that produce more of each output with less of each input. If no more efficient country is found, country i is considered as an efficient one, and we assign unit input and output efficiency scores to it. If country i is not efficient, its input efficiency score is equal to:

$$\underset{n=n_1,\ldots,n_l}{MIN}\ \underset{j=1,\ldots,k}{MAX}\frac{x_j(n)}{x_j(i)}$$

where $n_1,\ldots,n_l$ are the l countries that are more efficient than country i.

The output efficiency score is calculated in a similar way and is equal to:

$$\underset{n=n_1,\ldots,n_l}{MIN}\ \underset{j=1,\ldots,m}{MAX}\frac{y_j(i)}{y_j(n)}$$

Following the input and output scores calculation, countries can be ranked accordingly. Efficient countries are the same in both the input and output perspective, but the ranking and the efficiency scores of inefficient countries is not necessarily similar from both points of view.

2.2 Data Envelopment Analysis (DEA)

Data Envelopment Analysis, originating from Farrell (1957) seminal work and popularised by Charnes, Cooper and Rhodes (1978), assumes the existence of a convex production frontier, constructed using linear programming methods.[6]

Similarly to FDH, DEA allows the calculation of technical efficiency measures that can be either input or output oriented. The two measures provide the same results under constant returns to scale but give different values under variable returns to scale. Nevertheless, and since the computation uses linear programming, not subject to statistical problems such as simultaneous equation bias and specification errors, both output and input-oriented models will identify the same set of DMUs.

The analytical description of the linear programming problem to be solved, in the variable returns to scale hypothesis, is sketched below. Suppose there are k inputs and m outputs for n DMUs. For the i-th DMU, y_i is the column vector of the outputs and x_i is the column vector of the inputs. We can also define X as the $(k\times n)$ input matrix and Y as the $(m\times n)$ output matrix. The DEA model is then specified with the following mathematical programming problem, for a given i-th DMU:[7]

---

[6] Coelli *et al.* (1998), and Thanassoulis (2001) introduce the reader to the DEA.

[7] We simply present here the equivalent envelopment form, derived by Charnes *et al.* (1978), using the duality property of the multiplier form of the original programming model. See Coelli *et al.* (1998) for more details.

$$\begin{aligned} &MIN_{\theta,\lambda}\ \theta \\ &\text{s.to}\ \ -y_i + Y\lambda \geq 0 \\ &\qquad \theta x_i - X\lambda \geq 0 \\ &\qquad n1'\lambda = 1 \\ &\qquad \lambda \geq 0 \end{aligned} \tag{3}$$

In problem (3), θ is a scalar (that satisfies $\theta \leq 1$), more specifically it is the efficiency score that measures technical efficiency of unit (x_i, y_i). It measures the distance between a decision unit and the efficiency frontier, defined as a linear combination of best practice observations. With $\theta<1$, the decision unit is inside the frontier (*i.e.* it is inefficient), while $\theta=1$ implies that the decision unit is on the frontier (*i.e.* it is efficient).

The vector λ is a ($n \times 1$) vector of constants, which measures the weights used to compute the location of an inefficient DMU if it were to become efficient. The inefficient DMU would be projected on the production frontier as a linear combination, using those weights, of the peers of the inefficient DMU. The peers are other DMUs that are more efficient and therefore are used as references for the inefficient DMU.

$n1$ is a n-dimensional vector of ones. The restriction $n1'\lambda = 1$ imposes convexity of the frontier, accounting for variable returns to scale. Dropping this restriction would amount to admit that returns to scale were constant.

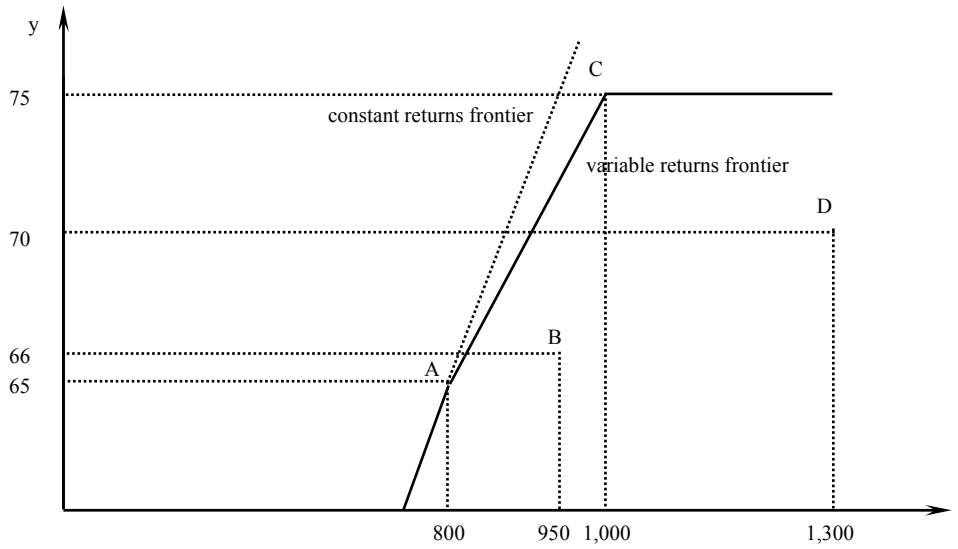
Notice that problem (3) has to be solved for each of the n DMUs in order to obtain the n efficiency scores.

Figure 2 illustrates DEA frontiers with the very same invented data of Table 1. The variable returns to scale frontier unites the origin to point A, and then point A to point C. If one compares this frontier to the FDH frontier in Figure 1, one notices that country B is now deemed inefficient. This is the result of the convexity restriction imposed when applying DEA. In fact, DEA is more stringent than FDH – a country that is efficient under FDH is not always efficient under DEA, but a country efficient under DEA will be efficient under FDH. In more general terms, input or output efficiency scores will be smaller with DEA.

The constant returns to scale frontier is represented in Figure 2 as a dotted line. In this one input – one output framework, this frontier is a straight line that passes through the origin and country A, where the output/input ratio is higher. Under this hypothesis, only one country is considered as efficient. In the empirical analysis that follows, *a priori* conceptions about the shape of the frontier were kept to a minimum and the constant returns to scale hypothesis is never imposed.

Figure 2

DEA Frontiers



3. Input and output measurement

3.1 Education indicators

As in Afonso and St. Aubyn (2005a), our main source of education data is OECD (2002a).[8] Concerning education achievement we selected two frontier models: one model where the input is a financial variable and another version where we use only quantity explanatory variables as inputs. In both specifications, the output is measured by the performance of 15-year-olds on the PISA reading, mathematics and science literacy scales in 2000 (simple average of the three scores for each country).

In the first specification, inputs are measured by the annual expenditure on educational institutions per student in equivalent US dollars converted using Purchasing Power Parities, in secondary education, based on full-time equivalents, 1999.

In the second specification, we use two quantitative input measures:

---

[8] The data and the sources used in this paper are presented in the Appendix. Note that total expenditure (public and private) was considered.

- the total intended instruction time in public institutions in hours per year for the 12 to 14-year-olds, 2000,
- the number of teachers per student in public and private institutions for secondary education, calculations based on full-time equivalents, 2000.

Since with these non-parametric approaches, higher performance is directly linked with higher input levels, we constructed the variable "Teachers Per Student," *TPS*:

$$TPS = \left(\frac{Students}{Teachers} \right)^{-1} \times 100 \tag{4}$$

using the original information for the students-to-teachers ratio (see Appendix). Naturally, one would expect education performance to increase with the number of teachers per student.

3.2 Health indicators

Following Afonso and St. Aubyn (2005a), we took two usual measures of health attainment, infant mortality and life expectancy, from OECD (2000b) and have calculated an "Infant Survival Rate", *ISR*:

$$ISR = \frac{1000 - IMR}{IMR} \tag{5}$$

which is interpretable as the ratio of children that survived the first year to the number of children that died; and increases with a better health status.

We have chosen to measure health spending in per capita terms and in purchasing power parities, therefore allowing for the fact that poorer countries spend less in real and per capita terms, even if their health spending is hypothetically comparable to richer nations when measured as a percentage of GDP.[9]

Therefore, our first frontier model for health has two outputs:

- the infant survival rate, and
- life expectancy,

the input being per capita health expenditure in purchasing power parities.

In a second formulation, and following the same reasoning that was made for education, we compared physically measured inputs to outcomes. In our second frontier model for health outputs are the same as before. Quantitative inputs are the number of doctors, of nurses and of in-patient beds per thousand habitants.

---

[9] As with education, total expenditure (public and private) was considered. See the Appendix for data details.

4. Empirical results

4.1 Education – financial input results

Concerning the education performance for the secondary level in the OECD countries, we present in Table 2 the results of the FDH analysis using a single output, the PISA rankings for 2000, and a single input, annual expenditure per student in 1999.

From the results it is possible to conclude that five countries are located on the possibility production frontier: Hungary, Japan, Korea, Mexico, and Poland. Overall, average input efficiency is around 0.61 implying that on average countries in our sample might be able to achieve the same level of performance using only 61 per cent of the per capita expenditure they were using. In other words, there seems to be a "waste" of input resources of around 39 per cent on average.

The scope for input efficiency improvement is quite large since for some countries (Italy, Portugal) the input efficiency score is roughly half of the average score. For instance, countries such as Italy and Germany, where expenditure per student is above average, deliver a performance in secondary attainment below the average of the PISA index.

Some important differences have to be mentioned when looking at the set of efficient countries in terms of education performance. Japan and Korea are located in the efficient frontier because they do indeed perform quite well in the PISA survey, getting respectively the first and the second position in the overall education performance index ranking. However, in terms of annual spending per student, Japan ranks above the average (6039 *versus* 5595 US dollars) and Korea (3419 US dollars) is clearly below average.[10]

On the other hand, countries like Mexico, Poland and Hungary are deemed efficient in the FDH analysis because they are quite below average in terms of spending per student. Given the expenditure allocated to education by these countries, their performance in the PISA index is not comparable to any other country with similar or inferior outcome and with less expenditure per student. Moreover, one has to note that Mexico, Poland and Hungary all have PISA outcomes below the country sample average.[11]

In Table 3 we present the DEA variable-returns-to-scale technical efficiency results using the same one-input and one-output framework. We report for each country its peers, *i.e.* the countries that give the efficient production for each decision unit.[12]

---

[10] See Appendix for details on the data.

[11] Notice that, by construction, the country that spends less is always on the frontier, even if its results are poor. Afonso and St. Aubyn (2005a) report results which exclude these countries from the sample.

[12] Additionally, and as a measure of comparison, we also present the constant returns to scale results. All the DEA computations of this paper were performed with the computer software DEAP 2.1 provided by Coelli *et al.* (1998).

Table 2

FDH Education Efficiency Scores: 1 Input *(annual expenditure on secondary education per student in 1999)* **and 1 Output** *(PISA 2000 survey Indicator)*

Country	Input efficiency		Output efficiency		Dominating producer *
	Score	Rank	Score	Rank	
Australia	0.499	14	0.975	9	Korea/Japan
Austria	0.402	20	0.946	12	Korea/Japan
Belgium	0.531	13	0.935	14	Korea/Japan
Canada	0.572	11	0.983	7	Korea/Korea
Czech Republic	0.991	6	0.924	17	Korea/Korea
Denmark	0.448	17	0.916	20	Korea/Japan
Finland	0.583	9	0.998	6	Korea/Korea
France	0.478	16	0.934	15	Korea/Japan
Germany	0.359	21	0.897	22	Hungary/Japan
Greece	0.545	12	0.943	13	Poland/Hungary
Hungary	**1.000**	**1**	**1.000**	**1**	
Ireland	0.780	7	0.950	10	Korea/Korea
Italy	0.243	24	0.872	23	Poland/Japan
Japan	**1.000**	**1**	**1.000**	**1**	
Korea	**1.000**	**1**	**1.000**	**1**	
Mexico	**1.000**	**1**	**1.000**	**1**	
Norway	0.448	18	0.923	18	Korea/Japan
Poland	**1.000**	**1**	**1.000**	**1**	
Portugal	0.306	23	0.842	24	Poland/Korea
Spain	0.487	15	0.899	21	Hungary/Korea
Sweden	0.578	10	0.947	11	Korea/Korea
Switzerland	0.350	22	0.933	16	Korea/Japan
United Kingdom	0.610	8	0.976	8	Korea/Korea
United States	0.419	19	0.918	19	Korea/Japan
Average	**0.610**		**0.966**		

* In terms of input efficiency/in terms of output efficiency.

Table 3

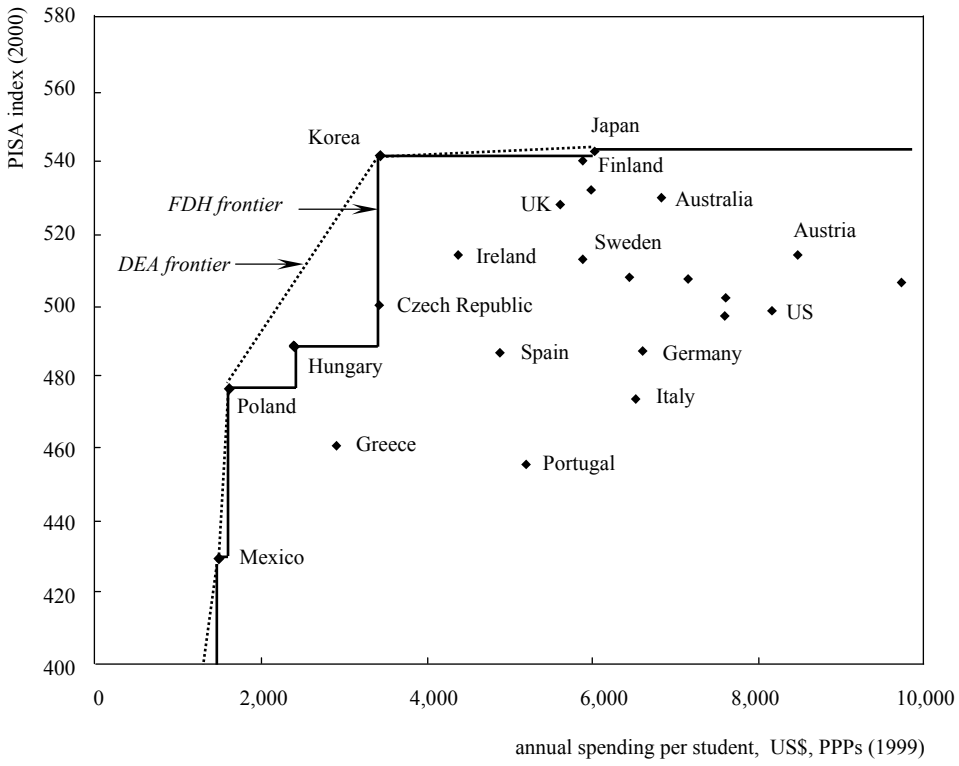
DEA Results for Education Efficiency in OECD Countries
1 Input *(annual expenditure on secondary education per student in 1999)*
and 1 Output *(PISA 2000 survey indicator)*

Country	Input oriented		Output oriented		Peers Input/output	CRS TE
	VRS TE	Rank	VRS TE	Rank		
Australia	0.453	12	0.976	7	Korea, Poland/Japan	0.257
Austria	0.311	17	0.947	11	Korea, Poland/Japan	0.201
Belgium	0.384	14	0.936	13	Korea, Poland/Japan	0.262
Canada	0.528	11	0.98	6	Korea, Poland/Japan, Korea	0.295
Czech Republic	0.650	6	0.924	16	Korea, Poland/Japan, Korea	0.481
Denmark	0.283	20	0.915	19	Korea, Poland/Japan	0.216
Finland	0.578	8	0.995	5	Korea, Poland/Japan, Korea	0.306
France	0.342	16	0.934	14	Korea, Poland/Japan	0.235
Germany	0.283	21	0.897	21	Korea, Poland/Japan	0.245
Greece	0.533	10	0.879	22	Mexico, Poland/Korea, Poland	0.526
Hungary	0.802	5	0.968	9	Korea, Poland/Korea, Poland	0.684
Ireland	0.603	7	0.949	10	Korea, Poland/Japan	0.389
Italy	0.242	24	0.871	23	Mexico, Poland/Japan	0.241
Japan	**1.000**	**1**	**1.000**	**1**	Japan/Japan	0.298
Korea	**1.000**	**1**	**1.000**	**1**	Korea/Korea	0.525
Mexico	**1.000**	**1**	**1.000**	**1**	Mexico/Mexico	0.962
Norway	0.298	18	0.923	17	Korea, Poland/Japan	0.218
Poland	**1.000**	**1**	**1.000**	**1**	Poland/Poland	1.000
Portugal	0.297	19	0.841	24	Mexico, Poland/Japan, Korea	0.292
Spain	0.384	15	0.898	20	Korea, Poland/Japan, Korea	0.332
Sweden	0.443	13	0.945	12	Korea, Poland/Japan, Korea	0.288
Switzerland	0.248	23	0.932	15	Korea, Poland/Japan	0.172
United Kingdom	0.543	9	0.973	8	Korea, Poland/Japan, Korea	0.312
United States	0.271	22	0.919	18	Korea, Poland/Japan	0.203
Average	**0.520**		**0.942**			**0.373**

Notes: CRS TE = constant returns to scale technical efficiency,
VRS TE = variable returns to scale technical efficiency.

Figure 3

Production Possibility Frontier, 24 OECD Countries, 2000



It seems interesting to point out that in terms of variable returns to scale, the set of efficient countries that comes out from the DEA approach, Japan, Korea, Mexico and Poland, are basically the same countries that were on the production possibility frontier built previously with the FDH results. In the DEA analysis only Hungary is no longer efficient.

Using the results obtained from both DEA and FDH analysis, we constructed the production possibility frontiers for this set of OECD countries (see Figure 3), concerning spending per student and the PISA report outcomes. The graphical portray of the production possibility frontiers helps locating the countries in terms of distance from those frontiers. The dotted line represents the DEA frontier, while the full line stands for the FDH one. It is visually apparent how Hungary is dropped from the efficiency frontier when convexity is imposed. Notice that, while some countries are positioned rather away from the frontier, such as the already mentioned cases of Portugal, Germany and Italy, other countries are relatively close to it, such as the Czech Republic, Finland, Australia or the UK.

4.2 Education – results with quantitatively measured inputs

We broadened our education efficiency analysis by looking at quantity measures of inputs used to reach the recorded outcome of education secondary performance. This implied an alternative specification, still using the PISA index as the output but now with two input measures instead of one. These new input measures are the following quantity variables: number of hours per year spent in school and the number of teachers per student (see details in the Appendix).

The results of the FDH analysis for this 2 inputs and 1 output alternative are reported in Table 4. We can observe that three of the countries that are now labelled

Table 4

FDH Education Efficiency Scores: 2 Inputs *(hours per year in school, 2000, teachers per 100 students, 2000)* **and 1 Output** *(PISA 2000 survey indicator)*

Country	Input efficiency		Output efficiency		Dominating producers *
	Score	Rank	Score	Rank	
Australia	0.850	13	0.975	7	Korea/Japan
Belgium	0.689	18	0.935	9	Sweden/Japan
Czech Republic	0.931	7	0.926	11	Sweden/Finland
Denmark	0.912	10	0.916	12	Sweden/Japan
Finland	**1.000**	**1**	**1.000**	**1**	
France	0.832	14	0.934	10	Korea/Japan
Germany	0.961	6	0.897	15	Korea/Japan
Greece	0.758	16	0.848	17	Sweden/Japan
Hungary	0.801	15	0.899	14	Sweden/Japan
Italy	0.730	17	0.872	16	Sweden/Japan
Japan	**1.000**	**1**	**1.000**	**1**	
Korea	**1.000**	**1**	**1.000**	**1**	
Mexico	**1.000**	**1**	**1.000**	**1**	
New Zealand	0.914	9	0.982	6	Korea/Korea
Portugal	0.879	11	0.844	18	Sweden/Finland
Spain	0.876	12	0.901	13	Sweden/Finland
Sweden	**1.000**	**1**	**1.000**	**1**	
United Kingdom	0.922	8	0.973	8	Korea/Japan
Average	**0.892**		**0.939**		

* In terms of input efficiency/in terms of output efficiency.

as efficient, Japan, Korea, and Mexico, are precisely the same as before, when we used a financial measure as the sole input variable. However, now Hungary is no longer efficient, while Poland, another efficient country in the financial input setup was dropped from the sample due to the unavailability of data concerning the number of hours per year spent in school.

Mexico is still deemed efficient essentially due to the fact that it has the highest students-to-teachers ratio in the country sample. On the other hand Hungary has now worse efficiency rankings and is dominated by Sweden and by Japan, that have a lower number of hours per year spent in school and a higher students-to-teachers ratio. Furthermore, both Japan and Sweden had a better performance outcome than Hungary in the PISA education index. Additionally, Sweden and Finland now come up as efficient since they have a students per teacher ratio not very different from the average, they are below average in terms of hours per year spent in school, and are above average concerning the PISA index ranking.

Therefore, this supplementary set of results, using quantity measures as inputs instead of a financial measure, seems to better balance the relative importance of the inputs used by each country. Indeed, it seems natural that in more developed countries like Sweden and Finland the cost of resources is higher than in less developed countries like Hungary and Mexico. Both Sweden and Finland were being somehow penalised when only a financial input was being used but this bias can be corrected using quantity measures as inputs.

Additionally, this set of results also reveals a higher average input efficiency score than before, placing the average "wasted" resources at a lower threshold of around 11 per cent. Concerning the average output efficiency score the results are nevertheless similar either using a financial input measure or two quantity input measures.

In Table 5 we report similar DEA variable-returns-to-scale technical efficiency results for 2 inputs and 1 output case.

With these quantity inputs one notices that three countries are still labelled efficient as before (DEA with 1 input and 1 output) assuming variable returns to scale: Japan, Korea, and Mexico. However, now two new countries appear as well as efficient, Sweden and Finland, in line with the results we obtained with the FDH analysis. Again Poland was dropped from the sample due to data unavailability and Hungary is once more no longer located on the frontier.

4.3 Health – financial input results

Results using input measured in monetary terms are a tentative answer to the following questions: do countries that spend more on health attain a better health status for their population? Or else are there a number of countries that spend comparatively more on health without an improved result?

Table 5

DEA Results for Education Efficiency in OECD Countries, 2 Inputs *(hours per year in school and teachers per 100 students)* **and 1 Output** *(PISA survey indicator)*

Country	Input oriented		Output oriented		Peers Input/output	CRS TE
	VRS TE	Rank	VRS TE	Rank		
Australia	0.788	14	0.976	7	Sweden, Finland, Korea/Japan	0.783
Belgium	0.689	18	0.936	9	Sweden, Korea/Japan	0.683
Czech Republic	0.880	6	0.921	11	Sweden, Korea/Japan, Finland	0.849
Denmark	0.857	12	0.915	12	Sweden, Korea/Japan	0.823
Finland	**1.000**	**1**	**1.000**	**1**	Finland/Finland	0.981
France	0.762	15	0.934	10	Sweden, Korea/Japan	0.736
Germany	0.891	6	0.897	15	Sweden, Korea/Japan	0.823
Greece	0.715	17	0.847	17	Sweden, Korea/Japan	0.636
Hungary	0.801	13	0.899	13	Sweden/Japan	0.762
Italy	0.728	16	0.871	16	Sweden, Korea/Japan	0.671
Japan	**1.000**	**1**	**1.000**	**1**	Japan/Japan	0.942
Korea	**1.000**	**1**	**1.000**	**1**	Korea/Korea	1.000
Mexico	**1.000**	**1**	**1.000**	**1**	Mexico/Mexico	1.000
New Zealand	0.878	9	0.979	6	Sweden, Korea/Japan, Finland	0.874
Portugal	0.880	8	0.842	18	Sweden/Japan, Finland	0.782
Spain	0.877	10	0.899	14	Sweden/Japan, Finland	0.832
Sweden	**1.000**	**1**	**1.000**	**1**	Sweden/Sweden	1.000
United Kingdom	0.859	11	0.972	8	Sweden, Finland, Korea/Japan	0.859
Average	**0.867**		**0.938**			**0.835**

Notes: CRS TE = constant returns to scale technical efficiency.
VRS TE = variable returns to scale technical efficiency.

Table 6 displays FDH results when a financial input, total per capita expenditure, is considered. In 30 considered countries, 11 were estimated to be on the efficiency frontier – the Czech Republic, Finland, Greece, Iceland, Japan, Korea, Mexico, Poland, Portugal, Spain and Turkey.

Note again that, by construction, the country that spends less is always on the frontier, even if its results are poor. This is why Mexico and Turkey are considered here as efficient, as both spend clearly below average and have results also clearly below average.

Another group of countries located in the frontier is the "less than average spenders" that attains "average to good results." Here, we can include the Czech Republic, Greece, Korea, Portugal and Spain. Finally, Finland, Iceland and Japan belong to a third group – those that have very good results without spending that much.

If we analyse the inefficient group of countries, the ones not in the FDH frontier, a number of countries display strong spending inefficiency. The United States have an input efficiency score of 0.313 with Greece as a reference, meaning that Greece spends less than a third of what the US spends, having better results. From this point of view, the US wastes more than two thirds of its spending. Similarly, Spain, an efficient country, spends slightly more than half (56.5 per cent) of German expenditure, being better off. Germany therefore is estimated to waste 43.5 per cent of its spending.

Results for this 1 input – 2 output model using DEA are summarised is Table 7.

In general terms, DEA results are not very different from FDH ones, the efficient group of countries being a subset of those previously efficient under FDH analysis. Specifically, Finland, Greece, Portugal and Spain are now inefficient, and the Czech Republic, Iceland, Japan, Korea, Mexico, Poland and Turkey define the frontier. The most striking difference is for Portugal – under DEA, this country is now near the end of the ranking, either in terms of input or output scores. Indeed, Portugal is dominated by the Czech Republic, Korea, and Japan, the first two countries having lower per capita spending in health and similar life expectancy.

4.4 Health – results with quantitatively measured inputs

When using quantitatively measured inputs, we are simply comparing resources available to the health sector (doctors, nurses, beds) with outcomes, without controlling for the cost of those resources. It is therefore possible that a country is efficient under this framework, but not in a model where spending is the input.

Half among the 26 countries analysed with this second formulation for health was estimated as efficient under FDH analysis (see Table 8). These are Canada, Denmark, France, Japan, Korea, Mexico, Norway, Portugal, Spain, Sweden, Turkey,

Table 6

FDH Health Efficiency Scores: 1 Input *(*per capita *total health expenditure)* **and 2 Outputs** *(infant survival rate and life expectancy)*

Country	Input efficiency		Output efficiency		Dominating producers *
	Score	Rank	Score	Rank	
Australia	0.843	18	0.981	16	Japan
Austria	0.882	15	0.969	22	Japan
Belgium	0.689	24	0.964	27	Spain/Japan
Canada	0.759	22	0.981	17	Japan
Czech Republic	**1.000**	**1**	**1.000**	**1**	
Denmark	0.682	25	0.952	29	Finland/Japan
Finland	**1.000**	**1**	**1.000**	**1**	
France	0.823	20	0.979	18	Japan
Germany	0.565	29	0.965	26	Spain/Japan
Greece	**1.000**	**1**	**1.000**	**1**	
Hungary	0.839	19	0.936	30	Korea
Iceland	**1.000**	**1**	**1.000**	**1**	
Ireland	0.878	17	0.972	21	Spain
Italy	0.780	21	0.975	19	Spain/Japan
Japan	**1.000**	**1**	**1.000**	**1**	
Korea	**1.000**	**1**	**1.000**	**1**	
Luxembourg	0.586	28	0.969	23	Spain/Japan
Mexico	**1.000**	**1**	**1.000**	**1**	
Netherlands	0.678	26	0.968	24	Spain/Japan
New Zealand	0.954	14	0.995	13	Spain
Norway	0.717	23	0.974	20	Japan
Poland	**1.000**	**1**	**1.000**	**1**	
Portugal	**1.000**	**1**	**1.000**	**1**	
Slovak Republic	0.983	13	0.967	25	Korea
Spain	**1.000**	**1**	**1.000**	**1**	
Sweden	0.993	12	1.000	12	Japan
Switzerland	0.588	27	0.990	14	Japan
Turkey	**1.000**	**1**	**1.000**	**1**	
United Kingdom	0.881	16	0.983	15	Spain
United States	0.313	30	0.953	28	Greece/Japan
Average	**0.848**		**0.982**		

* In terms input efficiency/in terms of output efficiency.

Table 7

DEA Results for Health Efficiency in OECD Countries, 1 Input *(*per capita *total expenditure in health)* **and 2 Outputs** *(infant survival rate and life expectancy)*

Country	Input oriented		Output oriented		Peers Input/output	CRS TE
	VRS TE	Rank	VRS TE	Rank		
Australia	0.670	17	0.981	13	Japan, Mexico/Japan	0.385
Austria	0.634	19	0.969	20	Czech Republic, Japan, Korea/Japan	0.502
Belgium	0.556	25	0.964	25	Czech Republic, Japan, Korea/Japan	0.447
Canada	0.604	21	0.981	14	Japan, Mexico/Japan	0.369
Czech Republic	**1.000**	**1**	**1.000**	**1**	Czech Republic/Czech Republic	1.000
Denmark	0.526	26	0.952	29	Czech Republic, Japan, Korea/Japan	0.462
Finland	0.906	10	0.981	15	Czech Republic, Iceland, Japan/ Czech Republic, Iceland, Japan	0.768
France	0.641	18	0.979	16	Korea, Japan, Mexico/Japan	0.479
Germany	0.490	29	0.965	24	Czech Republic, Japan, Korea/Japan	0.395
Greece	0.892	12	0.992	9	Japan, Mexico/Japan, Mexico	0.564
Hungary	0.757	14	0.928	30	Czech Republic, Poland/ Japan, Korea, Mexico	0.751
Iceland	**1.000**	**1**	**1.000**	**1**	Iceland/Iceland	0.823
Ireland	0.591	22	0.958	27	Czech Republic, Japan, Korea/ Japan, Mexico	0.515
Italy	0.711	15	0.975	17	Japan, Mexico/Japan	0.490
Japan	**1.000**	**1**	**1.000**	**1**	Japan/Japan	0.737
Korea	**1.000**	**1**	**1.000**	**1**	Korea/Korea	0.973
Luxembourg	0.511	28	0.969	21	Czech Republic, Japan, Korea/Japan	0.402
Mexico	**1.000**	**1**	**1.000**	**1**	Mexico/Mexico	0.839
Netherlands	0.559	24	0.968	22	Japan, Korea, Mexico/Japan	0.419
New Zealand	0.837	13	0.987	12	Japan, Mexico/Japan, Mexico	0.571
Norway	0.580	23	0.974	18	Czech Republic, Japan, Korea/Japan	0.460
Poland	**1.000**	**1**	**1.000**	**1**	Poland/Poland	1.000
Portugal	0.628	20	0.959	26	Czech Republic, Japan, Korea/ Japan, Mexico	0.593
Slovak Republic	0.895	11	0.966	23	Czech Republic, Poland/ Japan, Korea, Mexico	0.895
Spain	0.955	8	0.996	8	Japan, Korea, Mexico/Japan, Mexico	0.700
Sweden	0.948	9	0.988	11	Czech Republic, Iceland, Japan/ Czech Republic, Iceland, Japan	0.732
Switzerland	0.523	27	0.990	10	Japan, Mexico/Japan	0.323
Turkey	**1.000**	**1**	**1.000**	**1**	Turkey/Turkey	1.000
United Kingdom	0.672	16	0.972	19	Japan, Korea, Mexico/Japan, Mexico	0.509
United States	0.206	30	0.953	28	Japan, Korea, Mexico/Japan	0.157
Average	**0.743**		**0.978**			**0.609**

Notes: CRS TE = constant returns to scale technical efficiency.
VRS TE = variable returns to scale technical efficiency.

Table 8

FDH Health Efficiency Scores: 3 Inputs *(doctors, nurses and beds)* **and 2 Outputs** *(infant mortality and life expectancy)*

Country	Input efficiency		Output efficiency		Dominating producers *
	Score	Rank	Score	Rank	
Australia	0.926	18	1.000	14	Canada
Austria	0.967	16	0.981	19	Sweden
Canada	**1.000**	**1**	**1.000**	**1**	
Czech Republic	1.000	15	0.949	21	France
Denmark	**1.000**	**1**	**1.000**	**1**	
Finland	0.935	17	0.974	20	Sweden
France	**1.000**	**1**	**1.000**	**1**	
Germany	0.935	24	0.949	24	Sweden
Greece	0.923	19	0.992	16	Spain
Hungary	0.663	26	0.913	26	Korea/Spain
Ireland	0.902	25	0.946	25	Canada
Italy	0.837	22	0.997	15	Spain
Japan	**1.000**	**1**	**1.000**	**1**	
Korea	**1.000**	**1**	**1.000**	**1**	
Luxembourg	1.000	14	0.991	18	Spain
Mexico	**1.000**	**1**	**1.000**	**1**	
Netherlands	0.935	23	0.974	22	Sweden
New Zealand	0.913	20	0.991	17	Canada
Norway	**1.000**	**1**	**1.000**	**1**	
Poland	0.902	21	0.946	23	United Kingdom
Portugal	**1.000**	**1**	**1.000**	**1**	
Spain	**1.000**	**1**	**1.000**	**1**	
Sweden	**1.000**	**1**	**1.000**	**1**	
Turkey	**1.000**	**1**	**1.000**	**1**	
United Kingdom	**1.000**	**1**	**1.000**	**1**	
United States	**1.000**	**1**	**1.000**	**1**	
Average	**0.959**		**0.987**		

* In terms input efficiency/in terms of output efficiency.

the United Kingdom and the United States. Again one can distinguish different reasons for being considered efficient. Some countries have few resources allocated to health with corresponding low results (Mexico, Turkey); a second group attains better than average results with lower than average resources (e.g. the United Kingdom); finally, there is a third group of countries which are very good performers (e.g. Japan and Sweden).

Again, under DEA, the efficient group is smaller than under FDH. DEA results are summarised in Table 9, and there are 8 countries in the frontier: Canada, Japan, Korea, Mexico, Spain, Sweden, Turkey and the United Kingdom. All these countries were already considered efficient under FDH, but half of the "FDH-efficient" nations are not efficient now (Denmark, France, Norway, Portugal, Spain, Sweden, Turkey, and the United States). It is interesting to note that a group of ex-communist countries and European Union 2004 newcomers (the Czech Republic, Hungary, and Poland) are among the less efficient in providing health, when resources are physically measured.

5. Conclusion

Table 10 summarises our results, in terms of the countries that we found out as being efficient.

In general terms, similarly to Afonso and St. Aubyn (2005a), results suggest that efficiency in spending in education and health, two sectors where public provision is predominant, is an important issue. In the education sector, the average input inefficiency varies between 0.520 (1 input, 1 output, DEA) and 0.892 (2 inputs, 1 output, FDH), depending on the model and method, and on health, it varies between 0.743 (1 input, 2 outputs, DEA) and 0.959 (3 inputs, 2 outputs, FDH). Less efficient countries can therefore attain better results using the very same resources.

However, measuring efficiency when one considers the financial resources allocated to a sector is different from assessing efficiency from the measurement of resources in physical terms. The case of Sweden clearly illustrates this point. This is a country that only arises as efficient, in both education and health sectors, when inputs are physically measured. In our interpretation, this may well result from the fact that resources are comparatively expensive in Sweden. An opposite example is provided by the twin cases of the Czech Republic and Poland in what concerns health and by Hungary and Poland in the education sector. They are not efficient in physical terms. Probably because resources considered (doctors, nurses, hospital beds, teachers) are comparatively cheaper, they become efficient in financial terms.

Some countries always appear as efficient, either in health or in education – Mexico, Japan and Korea. Mexico is the country that spends fewer resources in these sectors and also gets the worse results. It appears as efficient for this sole reason. Japan is the best performer in health and education as far as outputs are concerned, and does not spend too many resources. Korea is a very good education

Table 9

DEA Results for Health Efficiency in OECD Countries, 3 Inputs *(doctors, nurses and beds)* **and 2 Outputs** *(infant mortality and life expectancy)*

Country	Input oriented		Output oriented		Peers Input/output	CRS TE
	VRS TE	Rank	VRS TE	Rank		
Australia	0.832	11	0.990	12	Canada, Japan, Spain, United Kingdom/ Canada, Japan, Spain, Sweden	0.691
Austria	0.703	21	0.976	15	Japan, Korea, Sweden/Japan, Sweden	0.703
Canada	**1.000**	**1**	**1.000**	**1**	Canada	0.978
Czech Republic	0.681	22	0.936	24	Japan, Korea, Sweden/Japan, Sweden	0.675
Denmark	0.808	14	0.965	21	Korea, Mexico, Spain, Sweden/ Japan, Spain, Sweden	0.802
Finland	0.806	15	0.970	19	Japan, Korea, Sweden/Japan, Sweden	0.802
France	0.835	10	0.991	10	Japan, Korea, Spain, Sweden, United Kingdom/ Japan, Spain, Sweden	0.768
Germany	0.604	24	0.972	18	Japan, Korea, Sweden/Japan, Sweden	0.604
Greece	0.820	13	0.991	11	Korea, Mexico, Spain/Japan, Spain, Sweden	0.695
Hungary	0.480	26	0.892	26	Korea, Mexico, Turkey, United Kingdom/ Japan, Spain	0.460
Ireland	0.716	19	0.958	23	Japan, Korea, Sweden/Canada, Japan, Sweden	0.715
Italy	0.798	16	0.995	9	Mexico, Spain, Sweden/Japan, Spain, Sweden	0.743
Japan	**1.000**	**1**	**1.000**	**1**	Japan	1.000
Korea	**1.000**	**1**	**1.000**	**1**	Korea	1.000
Luxembourg	0.707	20	0.979	14	Japan, Korea, Spain, Sweden, United Kingdom/ Japan, Spain, Sweden	0.683
Mexico	**1.000**	**1**	**1.000**	**1**	Mexico	1.000
Netherlands	0.579	25	0.973	17	Canada, Japan, Korea, United Kingdom/ Japan, Sweden	0.577
New Zealand	0.830	12	0.986	13	Canada, Japan, Korea, United Kingdom/ Canada, Japan, Sweden	0.802
Norway	0.726	17	0.976	16	Japan, Korea, Sweden/Japan, Sweden	0.725
Poland	0.679	23	0.934	25	Mexico, Turkey, United Kingdom/ Canada, Japan, Spain, United Kingdom	0.675
Portugal	0.844	9	0.961	22	Korea, Mexico, Spain, Sweden/ Mexico, Spain, Sweden	0.836
Spain	**1.000**	**1**	**1.000**	**1**	Spain	1.000
Sweden	**1.000**	**1**	**1.000**	**1**	Sweden	1.000
Turkey	**1.000**	**1**	**1.000**	**1**	Turkey	1.000
United Kingdom	**1.000**	**1**	**1.000**	**1**	United Kingdom	1.000
United States	0.725	18	0.968	20	Mexico, Sweden, United Kingdom/ Canada, Mexico, Sweden	0.724
Average	**0.814**		**0.977**			**0.795**

Notes: CRS TE = constant returns to scale technical efficiency.
VRS TE = variable returns to scale technical efficiency.

Table 10

OECD Efficient Countries in Education and in Health Sectors: Two Non-parametric Approaches and Different Input and Output Measures

Sector	Inputs, Outputs	N-p M	Countries
Education	• Spending per student (in) • PISA (out)	FDH	Japan, Korea, Mexico, Poland, Hungary
		DEA	Japan, Korea, Mexico, Poland
	• Hours per year in school (in) • Teachers per 100 students (in) • PISA (out)	FDH	Japan, Korea, Mexico, Sweden, Finland
		DEA	Japan, Korea, Mexico, Sweden, Finland
Health	• Per capita health spending (in) • Life expectancy (out) • Infant mortality (out)	FDH	Czech Republic, Finland, Greece, Iceland, Japan, Korea, Mexico, Poland, Portugal, Spain, Turkey
		DEA	Czech Republic, Iceland, Japan, Korea, Mexico, Poland, Turkey
	• Doctors (in) • Nurses (in) • Hospital beds (in) • Life expectancy (out) • Infant mortality (out)	FDH	Canada, Denmark, France, Japan, Korea, Mexico, Norway, Portugal, Spain, Sweden, Turkey, UK, US
		DEA	Canada, Japan, Korea, Mexico, Spain, Sweden, Turkey, UK

N-p M = Non-parametric Method.

performer, and it spends very little on health with surprisingly good results in comparative terms.

Assessing efficiency across countries opens the way to a related line of research – one would like not only to measure inefficiency, but also to explain international differences. In Afonso and St. Aubyn (2005b), we find a statistically significant influence of GDP per head and of educational attainment by the adult population in explaining cross-country variation of output scores.[13] Measuring and explaining inefficiency, and quantifying the systemic and the environment contributions to it, is something that, we believe, is of great relevance in economic policy terms.

[13] The importance of these variables in explaing student achievement was already reported by Barro and Lee (2001), with different methods, countries, data and time period.

APPENDIX
DATA AND SOURCES

Table 11

Education Indicators

Country	PISA (2000) (1)	Spending per student (2)	Hours per year in school (3)	Students per teacher (4)
Australia	530	6850	1019	12.6
Austria	514	8504	1148	
Belgium	508	6444	1075	9.7
Canada	532	5981		18.8
Czech Republic	500	3449	867	13.1
Denmark	497	7626	890	12.8
Finland	540	5863	808	13.8
France	507	7152	1042	12.5
Germany	487	6603	903	15.2
Greece	460	2904	1064	10.7
Hungary	488	2368	925	11.2
Iceland	506		809	
Ireland	514	4383	891	
Italy	473	6518	1020	10.3
Japan	543	6039	875	15.2
Korea	541	3419	867	21.2
Luxembourg	436			9.2
Mexico	429	1480	1167	31.7
Netherlands			1067	17.1

Table 11 (continued)

Education Indicators

Country	PISA (2000) (1)	Spending per student (2)	Hours per year in school (3)	Students per teacher (4)
New Zealand	531		948	16.3
Norway	501	7628	827	
Poland	477	1583		15.5
Portugal	456	5181	842	9.0
Slovak Republic				13.2
Spain	487	4864	845	11.9
Sweden	513	5911	741	14.1
Switzerland	506	9756		
Turkey			796	14.0
United Kingdom	528	5608	940	14.8
United States	499	8157		15.2
Mean	500	5595	932	14.4
Median	506	5946	897	13.8
Minimum	429	1480	741	9.0
Maximum	543	9756	1167	31.7
Standard deviation	30	2186	117	4.6
Observations	27	24	24	25

(1) Average of performance of 15-year-olds on the PISA reading, mathematics and science literacy scales, 2000. Source: OECD (2001).
(2) Annual expenditure on educational institutions per student in equivalent US dollars converted using PPPs, secondary education, based on full-time equivalents, 1999. Source: OECD (2002a).
(3) Total intended instruction time in public institutions in hours per year for 12 to 14-year-olds, 2000. Source: OECD (2002a).
(4) Ratio of students to teaching staff in public and private institutions, secondary education, calculations based on full-time equivalents, 2000. Source: OECD (2002a).

Table 12

Health Indicators

Country	Life expectancy (1)	Infant mortality (2)	Per capita spending in health (3)	Doctors (4)	Nurses (5)	Hospital beds (6)
Australia	79.0	5.7	2058	2.5	8.1	7.9
Austria	78.0	4.4	1968	3.0	9.0	8.8
Belgium	77.6	4.9	2008	3.8		7.3
Canada	79.0	5.3	2285	2.1	7.5	3.9
Czech Republic	74.8	4.6	944	3.0	8.2	8.7
Denmark	76.6	4.2	2241	3.4	7.3	4.5
Finland	77.4	3.7	1529	3.1	14.4	7.6
France	78.8	4.3	2109	3	6	8.4
Germany	77.7	4.5	2451	3.5	9.5	9.2
Greece	78.1	6.2	1307	4.4	3.9	4.9
Hungary	70.7	8.4	751	3.2	5.0	8.3
Iceland	79.5	2.4	2204	3.4	14.2	
Ireland	76.5	5.5	1576	2.3	8.7	9.7
Italy	78.5	5.1	1774	5.9	4.5	4.9
Japan	80.5	3.4	1735	1.9	7.8	16.4
Korea	75.5	7.7	630	1.3	1.4	5.5
Luxembourg	78.0	4.6	2361	3.1	7.1	8
Mexico	75.0	25.9	431	1.7	1.2	1.1
Netherlands	77.9	5.2	2040	3.1	12.7	11.1
New Zealand	78.3	5.4	1450	2.3	9.6	6.2
Norway	78.4	3.9	2421	2.8	10.1	14.4

Table 12 (continued)

Health Indicators

Country	Life expectancy (1)	Infant mortality (2)	Per capita spending in health (3)	Doctors (4)	Nurses (5)	Hospital beds (6)
Poland	73.2	8.9	543	2.3	5.1	5.1
Portugal	75.6	5.6	1345	3.2	3.8	4
Slovak Republic	73.0	8.3	641		7.3	8.1
Spain	78.7	4.5	1384	3.1	3.6	4.1
Sweden	79.5	3.4	1748	2.9	8.4	3.7
Switzerland	79.7	4.6	2952	3.4		18.3
Turkey	68.4	40.3	303	1.2	1.1	2.6
United Kingdom	77.4	5.8	1527	1.8	4.6	4.1
United States	76.7	7.1	4178	2.8	8.3	3.6
Mean	76.9	7.1	1696.5	2.9	7.1	7.3
Median	77.8	5.2	1741.5	3.0	7.4	7.3
Minimum	68.4	2.4	303.0	1.2	1.1	1.1
Maximum	80.5	40.3	4178.0	5.9	14.4	18.3
Standard deviation	2.7	7.5	827.6	0.9	3.5	4.0
Observations	30	30	30	29	28	29

(1) Years of life expectancy. Total population at birth. 1999. Greece: 1998. Italy: 1997. Source: OECD (2002b).
(2) Deaths per 1000 live births. 1999. Korea: 1997. New Zealand: 1998. Source: OECD (2002b).
(3) Total expenditure on health per capita, purchasing power parities, US dollars. 1998. Source: OECD (2002b).
(4) Practising physicians, density per 1000 population. 1999. Australia, France and Japan: 1998. Source: OECD (2002b).
(5) Practising nurses, density per 1000 population. 1999. Australia, France: 1997. Japan: 1998. Slovakia: 2000. Source: OECD (2002b).
(6) Total in patient care beds per 1000 population. 1999. Denmark, Ireland, Luxembourg, New Zealand, Portugal: 1998. Belgium: 1997. Source: OECD (2002b).

REFERENCES

Afonso, A., L. Schuknecht and V. Tanzi (2003), "Public Sector Efficiency: An International Comparison," ECB, Working Paper, No. 242, forthcoming in *Public Choice*.

Afonso, A. and S. Fernandes (2005), "Local Government Spending Efficiency: DEA Evidence for the Lisbon Region," forthcoming in *Regional Studies*.

Afonso, A. and M. Santos (2005), "Students and Teachers: A DEA Approach to the Relative Efficiency of Portuguese Public Universities," ISEG/UTL–Technical University of Lisbon, Department of Economics, Working Paper, No. 7/2005/DE/CISEP.

Afonso, A. and C. Scaglioni (2005), "Public Services Efficiency Provision in Italian Regions: A Non-Parametric Analysis", Department of Economics, ISEG-UTL, Working Paper, No. 02/2005/DE/CISEP, forthcoming in F. Columbus (ed.), *State and Local Economics: Issues and Developments*, Nova Science Publishers, New York.

Afonso, A. and M. St. Aubyn (2005a), "Non-parametric Approaches to Public Education and Health Expenditure Efficiency in OECD Countries", forthcoming in *Journal of Applied Economics*.

———— (2005b), "Cross-country Efficiency of Secondary Education Provision: A Semi-parametric Analysis with Non-discretionary Inputs", European Central Bank, Working Paper, No. 494, June. ISEG/UTL–Technical University of Lisbon, Department of Economics, Working Paper, No. 5/2005/DE/CISEP/UECE.

Barro, R. and J.W. Lee (2001), "Schooling Quality in a Cross-section of Countries", *Economica*¸ No. 68, pp. 465-88.

Charnes, A., 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. (1996), "Assessing the Performance of Australian Universities Using Data Envelopment Analysis," University of New England, mimeo.

Coelli, T., P. Rao and G. Battese (1998), *An Introduction to Efficiency and Productivity Analysis,* Boston, Kluwer.

De Borger, B. and K. Kerstens (1996), "Cost Efficiency of Belgian Local Governments: A Comparative Analysis of FDH, DEA, and Econometric Approaches," *Regional Science and Urban Economics*, Vol. 26, No. 2, pp. 145-70.

Deprins, D., L. Simar and H. Tulkens (1984), "Measuring Labor-efficiency in Post Offices", in M. Marchand, P. Pestieau and H. Tulkens (eds.), *The Performance of Public Enterprises: Concepts and Measurement*, Amsterdam, North-Holland.

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", Policy Research, Working Paper, No. 1803, Washington (D.C.), The World Bank.

Farrell, M. (1957), "The Measurement of Productive Efficiency," *Journal of the Royal Statistical Society, Series A*, No. 120, Part 3, pp. 253-90.

Gupta, S. and M. Verhoeven (2001), "The Efficiency of Government Expenditure – Experiences from Africa", *Journal of Policy Modelling*, No. 23, pp. 433-67.

OECD (2001), *Knowledge and Skills for Life – First Results from Pisa 2000*, Paris, OECD.

—————— (2002a). *Education at a Glance – OECD Indicators 2002*, Paris, OECD.

—————— (2002b), *OECD Health Data 2002*, Paris, OECD.

Sengupta, J. (2000), *Dynamic and Stochastic Efficiency Analysis – Economics of Data Envelopment Analysis*, Singapore, World Scientific.

Simar, L. and P. Wilson (2003), *Efficiency Analysis: The Statistical Approach,* lecture notes, July.

St. Aubyn, M. (2002), "Evaluating Efficiency in the Portuguese Health and Education Sectors," paper presented to the conference *Desenvolvimento Económico Português no Espaço Europeu: Determinantes e Políticas*, Banco de Portugal, May, Lisbon, Fundação Calouste Gulbenkian.

—————— (2003), "Evaluating Efficiency in the Portuguese Education Sector", *Economia*, No. 26, pp. 25-51.

Thanassoulis, E. (2001), *Introduction to the Theory and Application of Data Envelopment Analysis*, Kluwer Academic Publishers.

PUBLIC EXPENDITURE AND OPTIMAL GOVERNMENT SIZE IN AN ENDOGENOUS GROWTH MODEL: AN ANALYSIS OF THE ARGENTINE CASE

Ernesto Rezk[*]

1. Introduction

In spite of the valuable contributions the Solow Swan Model[1] rendered to the modern theory of Economic Growth the approach, based on a neoclassical production function with diminishing returns to labour and capital and combined with the assumption of a constant saving rate, yielded the uncomfortable prediction that per capita growth would eventually cease unless exogenous technological progress took place.

By acknowledging this deficiency in the model, many theorists enriched the theory of Economic Growth in diverse ways; Cass (1965) and Koopmans (1965), for instance, resorted to Ramsey's contribution[2] to the analysis of consumer optimization in order to provide an endogenous determination of the saving rate. Let it however be said that this improvement of the neoclassical growth model did not solve the problem of dependence of the long run growth rate on exogenous technical advances.

In aiming at sorting out the shortcomings of exogenous growth models, new lines of research, represented by the works of Romer (1986) and Lucas (1988), developed into what is known as endogenous growth models, allowing for a broader capital definition also including human capital and whose main feature was that the long run growth rate could be constant and positive as diminishing capital marginal product did not take place.[3]

In following the latter line of analysis, it results interesting to consider the inclusion of government in endogenous growth models in order to address the questions of what the optimal government size and the tax rate maximizing per capita consumption, capital and income growth rates should be and what implications they will bear upon the analysis, should one allow for distorting taxes to be used.

---

[*] Institute of Economics and Finance – Faculty of Economic Sciences – National University of Cordoba, Argentina. E-mail: ernerezk@eco.unc.edu.ar

The author is highly indebted to the members of his Research Team: Maria Victoria Sarjanovich, María Cecilia Avramovich and Martín Basso for their hard work and true commitment to the project; he is also grateful to Blanca Moreno Dodson, Peter Wierts and Robert Woods for their fruitful comments, observations and cites concerning objectives, modeling and analysis of results, during the Banca D'Italia's Seventh Workshop on Public Finance. They do not share however any responsibility for errors or omissions that may remain in the paper.

[1] Based on Solow (1956) and Swan (1956).

[2] Ramsey (1928).

[3] An instance of this are the so-called AK models of growth.

In this connection the paper aims at identifying for Argentina, by using an AK endogenous growth model and resorting to taxes likely to alter incentives upon savings and investment, the government size that makes maximum the per capita growth rate. Furthermore, and whatever magnitude the estimation of government size may render, the empirical exercise carried out seeks to demonstrate that an intertemporal fiscal balance is possible if a more efficiency-oriented and better administered tax system is aimed at, free from distorting taxes and with respect to which existing evasion levels are curtailed.

Notwithstanding the fact that the government size here equals the productive public spending share on GDP, the point may be differently regarded as the literature embodies at least two variants[4] for public expenditure: in the first place, the standard Samuelsonian approach to public goods in which consumption is neither rival nor excludable; in the second case, public spending refers to government activities entering private production functions as inputs subject to congestion as many firms coincide in the use of facilities.

No public spending, either current or capital outlays, is completely free from the congestion problem and, therefore, growth perspectives will tend to worsen when the former's provision falls short real of demand needs for all kinds of public services for a sustained period of time, not to mention the negative impact upon private production of externality-creating public investment shortage.[5]

In this respect, preliminary statistical analyses realized with the Argentina public spending, as of the Nineties (Table 1), showed that the public spending's and public investment's growth rate lagged in general well behind that of product for what – and to the extent that this is not reverted – public facilities scarcity may at some moment hinder the process of outuput growth. On these grounds, the congestion model of productive government services, due to Barro and Sala-i-Martín (1992), is used here as the conceptual framework for the evaluation of the optimal government size.

In extending the empirical support for the congestion model chosen, it should be noticed from figures above that the GDP's annual growth rate not only outweighed that of public spending 7 times out of 10, but also that its overall figure for the period considered reached 58.9 per cent compared with 35.6 per cent in total public spending and the modest 27.5 per cent exhibited by public investment; this gap between growth rates helps to explain why the public expenditure's proportion of GDP fell from 15.2 in 1993 to 13 per cent in 2003.

The optimal public spending share (as a proportion of GDP) definitionally equals, via the government budget constraint, the average tax rate and, for that,

---

[4] Barro (1990) also refers to the case in which public spending enters the private production function as another input (free public services to producers) whose use will be both rival and excludable.

[5] The point is worth emphasizing here that public provision of services and investment is not to be confused with production, as the latter can be either public or private (*i.e.* privatization of construction and maintenance of a part of the road network in Argentina).

Table 1

Argentine GDP and Total Public Spending, 1993-2003

Year	GDP (1)	Current Public Spending (1, 2)	Public Investment (1)	Total Public Spending (1)	Total Public Spending/ GDP (%)	GDP's Annual Growth Rate (%)	Public Spending's Annual Growth Rate (%)
1993	236.5	32.0	4.0	36.0	15.2	-	-
1994	257.4	33.9	4.6	38.5	15.0	8.84	6.94
1995	258.0	34.4	3.9	38.3	14.8	0.23	–0.52
1996	272.2	34.0	3.1	37.1	13.6	5.50	–3.13
1997	292.9	35.3	4.4	39.7	13.6	7.60	7.00
1998	298.9	37.4	4.6	42.0	14.1	2.05	5.79
1999	283.6	38.9	4.5	43.4	15.3	–5.12	3.33
2000	284.2	39.2	3.0	42.2	14.8	2.12	–2.76
2001	268.7	38.0	2.8	40.8	15.2	–5.45	–3.32
2002	312.6	38.2	2.3	40.5	13.0	16.34	–0.74
2003	375.9	43.7	5.1	48.8	13.0	20.25	17.28
Δ	**58.9%**	**36.5%**	**27.5**	**35.5%**	**-**	**-**	**-**

(1) Billions of current Argentine pesos, rate of exchange with the U.S. dollar: 1 dollar = 2.93 pesos.
(2) Only Wages, Goods and Services included. Interests, Social Security Payments and Transfers not included.

Source: Own estimates based on information from the National Institute of Statistics and Censuses (INDEC) of Argentina.

the model's empirical results will permit also to compare the optimal and actual average tax rates in Argentina and to suggest policy changes in the existing tax regime, either feasible in terms of tax yield capability (emphasis in efficiency) or convenient in terms of changes in income distribution (emphasis in welfare).

A no minor point is however worth clarifying concerning the scope of the paper: although the point is acknowledged that not only quantity but also quality of public spending bears a hold on long run economic growth, no qualitative assessment is carried out in the paper assuming – as said above – a uniform quality of provided services and facilities[6] by the government.

---

[6] The author is particularly grateful to Blanca Moreno Dodson who pointed out the convenience of focusing also in efficiency aspects of public spending. Let it in this connection be said that the no consideration of the quality dimension of public spending was here decided on simplicity grounds, in view of the objectives of the paper.

As for the structure of the paper: Section 2 includes a description of the model used whereas Section 3 and 4 are respectively devoted to the empirical exercise of determining the optimal government size and of suggesting tax changes in the light of achieved results and its comparison with the structure and revenue yield of the present Argentine Tax System; finally, Section 5 concludes.

2. An endogenous model of economic growth with public spending subject to congestion[7]

As Barro (1990) pointed out, the inclusion of public spending within an AK model amounts to enhancing the level of technology implied by A and will in consequence affect the long run per capita growth. The spending activities (subject to congestion) carried out by the government, and included in the model developed below, will therefore be considered to cause an effect on coefficient A regardless of their current or capital outlays' nature.

According to Barro and Sala-i-Martín (1992), the expression (1) below stands for the per capita production function for the *i*th producer:

$$y_i = Ak_i\, f(G / Y) \tag{1}$$

in which:

y_i = per capita product
k_i = per capita capital
G = productive public spending subject to congestion
$Y(\Sigma y_i)$ = aggregate product

As is easily noticed in (1), the functional expression $f(G / Y)$ implies that, given k_i, an increase in public spending relative to aggregate product will enhance y_i and in turn Y; conversely and due to congestion ($\Delta Y > \Delta G$), an increase in product relative to G will dwindle y_i.

By making the functional expression $f(G / Y)$ equal to $(G / Y)^{1-\alpha}$ and having:

$$f' = (1-\alpha)(G / Y)^{-\alpha} > 0 \qquad \text{and} \qquad f'' = -\alpha(1-\alpha)(G / Y)^{-\alpha-2} < 0$$

expression (1) above turns into:

$$y_i = Ak_i (G / Y)^{1-\alpha} \tag{1'}$$

where $0 < \alpha < 1$.

The demonstration that production function (1') exhibits constant returns to scale asks for all firms to have similar technology, for what α will be the same for each of them and for the economy as a whole.[8]

---

[7] This section includes a synthesis of the model used.

[8] The author is aware that criticisms can be raised in respect of the simplifying assumption that sector *i*'s factor shares also apply to the aggregate production function but, allowing that disparities may exist in reality regarding factors' intensity of use among sectors, results are still sound given the macroeconomic nature of the paper.

By dividing (G / Y) by the population, this quotient can be expressed in per capita terms, as in (2) below:

$$(G / Y) = [(G / N) / (Y / N)] = g / y \tag{2}$$

and since $\Sigma y_i = N y_i = Y$, the ensuing expression (3) will also hold:

$$G / Y = g / y_i \tag{3}$$

By substituting in (1'), and rearranging, (4') will be used to show constant returns to scale in the function:

$$y_i = A\, k_i\, (g / y_i)^{1-\alpha} \tag{4}$$

$$y_i = A\, k_i y_i^{-(1-\alpha)}\, g^{1-\alpha} \tag{4'}$$

rearranging as follows:

$$y_i^{2-\alpha} = A\, k_i\, g^{1-\alpha}$$

and solving, (4') will turn out into (5) below:

$$y_i = A\, k_i^{1/(2-\alpha)}\, g^{(1-\alpha)/(2-\alpha)} \tag{5}$$

It can be shown, from (5), that:

$$1 / (2 - \alpha) + (1 - \alpha) / (2 - \alpha) = 1 \tag{6}$$

and this in turn stands for constant returns to scale in the production function.

Infinite-lived households, on their part, maximize the following utility function:

$$U(0) = \int_0^{\infty} e^{-(\rho - n)\, t} [\, c^{1-\theta} - 1 / (1 - \theta)]\, \mathrm{d}t \tag{7}$$

subject to the budget constraint (8) stating that private consumption plus gross investment equal net of taxes per capita income:

$$\mathrm{d}k / \mathrm{d}t = (1 - \tau)\, \mathrm{A}\, k_i\, (G / Y)^{1-\alpha} - c - (\delta + n)\, k \tag{8}$$

where ρ, δ and n respectively stand for the temporal rate of preference, the depreciation rate and the population growth rate; θ in turn indicates the degree of concavity of the utility function while τ is the rate of a proportional tax on the aggregates of domestic gross product whose revenue yield is used by the government to run a balanced budget,[9] according to the ensuing budget constraint:

$$G = \tau Y \tag{9}$$

The expression (9), which depicts the government size in terms of public spending, may also be viewed as the average tax rate imposed upon the economy, according to (9') below:

$$\tau = G / Y \tag{9'}$$

Once the maximization process is performed, and all substitutions completed, the model renders per capita consumption and growth rates as follows:

$$\gamma = 1 / \theta [(1 - \tau)\, A\, (G / Y)^{1-\alpha} - (\delta + \rho)] \tag{10}$$

or, in terms of the tax rate τ:

---

[9] It would be more appropriate to state that the government could temporarily incur in surpluses or deficits, but the budget should in the long run be balanced.

$$\gamma = 1 / \theta[(1 - \tau) A \tau^{1-\alpha} - (\delta + \rho)] \quad (10')$$

Several points are worth emphasizing concerning the expression (10') above: in the first place, in so far as the government takes resources from the private sector, taxation reduces the per capita growth rate[10] but, at the same time it helps enhancing the latter through the corresponding provision of public facilities and services. Furthermore, by being a function of constans, the per capita growth rate is itself a constant and no dynamic transition will take place towards zero growth in the steady state; in other words, the growth rate will be positive and constant in the long run.

How does the growth rate achieved in (10') relate with the optimal government size? By taking derivatives in (10') with respect to τ, setting the derivative to zero and rearranging terms the expression (11) is achieved:

$$(1 - \tau^*) = f(\tau^*) / f'(\tau^*) \quad (11)$$

where:

τ^* is the tax rate that maximizes γ; $f(\tau) = \tau^{1-\alpha}$ and $f'(\tau) = (1 - \alpha)\ \tau^{-\alpha}$

After conveniently rearranging it, the expression (11) becomes:

$$\tau^* = (1 - \alpha) / (2 - \alpha) \quad (12)$$

The expression (12) shows that τ^*'s value will depend, under the assumption of a Cobb-Douglas production function that exhibits constant returns to scale, on the public spending share in product. Under the mentioned assumption, payments to factors according to their marginal product will exhaust the produced income, as indicated below:

$$PY = rK + \gamma G \quad (13)$$

Dividing both members by PY, the ensuing expression is obtained:

$$1 = 1 / (2 - \alpha) + (1 - \alpha) / (2 - \alpha) \quad (14)$$

where:

$$1 / (2 - \alpha) = rK / PY \qquad \text{and} \qquad (1 - \alpha) / (2 - \alpha) = \gamma G / PY$$

and, finally:

$$\alpha = 2 - PY / rK \quad (15)$$

3. The application of the model to the Argentine case

3.1 The Argentine fiscal scenario

In spite of Argentina being a three-tier federation embodying one national government, twenty four provincial governments and over 1,100 municipalities, all of which are constitutionally endowed with ample faculties to raising taxes and carrying out expenditure programmes, the existing interjurisdictional fiscal

[10] Barro and Sala-i-Martín (1995) refers to this as the negative effect of taxation on the after tax marginal product of capital.

arrangements (the so called revenue sharing system) whereby provinces delegate to the national level the collection of main taxes (that is VAT and the Corporation and Individual Income Tax) places in the national level's hands the responsibility of collecting 77-78 per cent of all tax revenues (as shown by Table 15 in the Statistical Appendix), while the subnational governments account for a rather modest 22-23 per cent. All in all, figures also show that – for the benchmark year 2003 – the real overall average tax rate[11] (including all government layers) amounted to 33.75 points of GDP (Table 14).

The nature of the Argentine Tax Regime, and the structure of tax revenues, as depicted by Tables 13 and 15 in the Statistical Appendix are well deserving some comments. Following the introduction of VAT in 1974, tax revenues in Argentina were practically made up with a handful of taxes, namely VAT, Social Security Contributions, Corporate and Personal Income Tax and Fuel Taxes; the fiscal status-quo was firstly disturbed when – as of 1994 – the new Pension Scheme came into being and a part of Social Security Contributions (the employees'dues) went thereafter to Private Pension Funds.

The second great change in the structure of tax revenues took place in 2001 when the national government, in the middle of a political and economic turmoil and in view of the serious budgetary restraint caused by the impossibility of acceding to new loans from international organisms or of placing new debt in financial markets, embarked itself in a so called "zero deficit budgetary policy" for what new taxes had to be resorted to.

In terms of the Tax System, the main consequences of the zero deficit policy were the reintroduction of Export Tariffs, which had been done away by the Government at the beginning of the Convertibility period (in 1991), and the Tax on Financial Transactions, both strongly resisted by economic agents on the grounds that the distorting impact upon exports' competitiveness and the wrong incentives they would give economic agents to move to the shadow economy seriously challenged the convenience and economic efficiency of their use.

The fiscal consequences of these tax changes are clearly depicted by Table 14's figures, showing a mounting tax pressure in 2002, and by Figure 1 overleaf in which Property and Foreign Trade Taxes Revenue's shares are seen to markedly increase since 2001-02.

In comparison, provinces' fiscal performance (Figure 2) makes only noticeable a slight improvement in the case of Taxes on Goods and Services explained by some boost in consumption accompanied by a nominal revenue rise following devaluation in 2002.

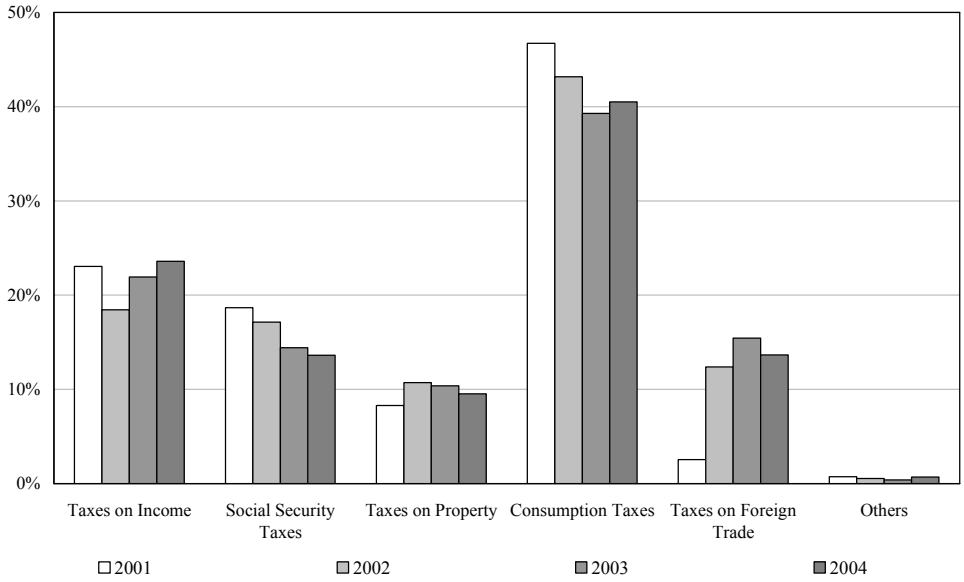
Whatever decisive against inefficiency the preceding arguments may be, Tables 13, 14 and 16 highlight the importance export tariffs and financial

---

[11] The real overall average tax rate results from the quotient between Overall Revenues and GDP.

Figure 1

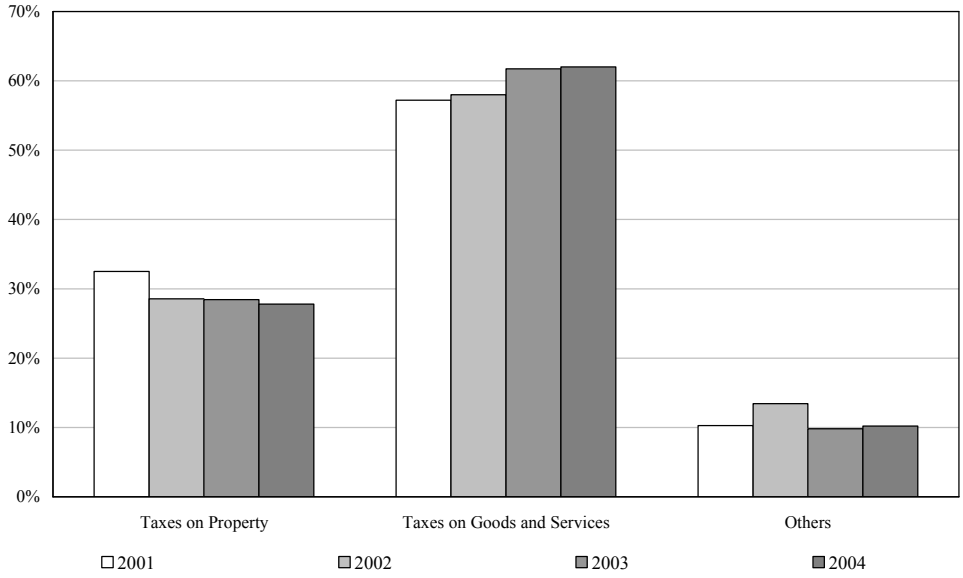
Argentina: National Taxes' Revenue Percentage Share



Source: Table 16 in the Statistical Appendix.

Figure 2

Argentina: Provincial Taxes' Revenue Percentage Share



Source: Table 16 in the Statistical Appendix.

transactions taxes have reached, in terms of GDP and as a percentage of the national tax revenue (2.32-2.03 points and 8.72-7.78 per cent, respectively, in 2004), for what any substitution would only be feasible if the lost yield caused by their replacement could immediately be made up with revenues coming from other sources[12] and these precisely are the foundations of the performed simulation exercise, whose results are found in Section 4.

On the expenditure side, and due to decentralization processes set in motion at the end of the Eighties and furthered during the Nineties, subnational governments (provinces and municipalities) were responsible in 2003 for practically 50 per cent of consolidated current and capital spending; their share was overwhelmingly high in the provision of certain public goods and services – especially in the fields of Education, Public Health and Housing – in which they accounted for almost 100 per cent of incurred expenses and in Welfare and Economic Services where the subnational share can by no means considered a minor one. Again, if overall figures are taken for 2003 (Tables 2 and 3), total public expenditure reached 27.62 points of GDP in 2003 and this figure, compared to the 29.01 points of current and capital revenue, rendered a fiscal surplus of almost 1.40 per cent of GDP.[13] In turn the total primary surplus, let alone payments of interest on domestic and foreign debt, reached 3.77 points of GDP in the same year.

3.2 The calculation of the optimal government size for Argentina

Section 2 showed that the budget constraint could be rearranged in order to have the average tax rate τ to stand for the government size [expression (9')] and that its magnitude, obtained by solving equation (12), would in turn guarantee that the requirement of a maximum economic growth rate was met.

The expression (12) also stated that τ^*'s value depended on $(1 - \alpha)$ standing for the public spending share in product. Under the quoted assumption of a constant returns to scale production function $1 / (2 - \alpha)$ and $(1 - \alpha) / (2 - \alpha)$ will respectively equal to rK / PY and $\gamma G / PY$.[14]

The empirical application of the model called in the first place for the choice of benchmark values for γG and PY to be made; in this connection, and in the light of relatively normal macroeconomic conditions in 2003, following the country's abnormal situation of default of its sovereign debt and the exit of

---

[12] The argument will be more easily understood if one takes into account that these two taxes' yields are crucial in the strategy followed by the Government of building the surplus required to meet the post-default incoming financial burden. Some estimates are given below by the author.

[13] Nevertheless, this surplus can not by any means be considered sustainable in the long run as it is somehow hiding the fact that no payments (interest and capital) are so far being made with respect to the defaulted public debt.

[14] It must be borne in mind that, by having *K*, *G* and *Y* multiplied by their prices, both these quotients are expressed in monetary terms.

Table 2

Argentina: Revenues, Expenditures and Financial Results by Government Level, 2003
(millions of current Argentine pesos)

ITEMS	NATIONAL LEVEL	PROVINCIAL LEVEL	MUNICIPAL LEVEL	TOTAL
Current Revenue	**65,080**	**35,356**	**7,690**	**108,126**
Tax Revenue	44,511	30,299	3,596	78,406
Social Security Contributions	10,470	-	-	10,470
Non-tax Revenue	3,344	4,350	4,097	11,790
Accrued Interest	4,471	257	-	4,727
Others	2,285	450	-	2,735
Current Expenditure	**53,110**	**36,577**	**7,380**	**97,067**
Consumption and Operating Surplus	12,404	24,351	6,940	43,695
Interest Payments	7,095	1,808	45	8,948
Social Security Benefits	18,868	-	-	18,868
Current Transfers	14,413	10,418	395	25,226
Other Current Expenses	331	-	-	331
Current Savings	**12,861**	**–1,221**	**310**	**11,950**
Capital Revenue	**206**	**691**	**34**	**931**
Capital Outlays	**1,267**	**4,410**	**1,080**	**6,756**
Transfers from Upper Levels	**15,706**	**6,606**	**835**	**23,147**
Transfers to Lower Levels	**22,276**	**872**	**-**	**23,147**
Total Primary Surplus	**11,423**	**2,603**	**145**	**23,147**
Total Primary Surplus(1)	**11,216**	**1,911**	**111**	**13,239**
Financial Result	**4,341**	**794**	**99**	**5,234**

(1) Exclusive of Capital Revenue.

Source: Ministry of Economics, National Direction of Fiscal Research and Analysis. Internet site: www.mecon.gov.ar/hacienda

Table 3

Argentina: Revenues, Expenditures and Financial Results by Government Level, 2003
(percent of GDP)

ITEMS	NATIONAL LEVEL	PROVINCIAL LEVEL	MUNICIPAL LEVEL	TOTAL
Current Revenue	**17.31**	**9.41**	**2.05**	**28.76**
Tax Revenue	11.84	8.06	0.96	20.86
Social Security Contributions	2.79	-	-	2.79
Non-tax Revenue	0.89	1.16	1.09	3.14
Accrued Interest	1.19	0.07	-	1.26
Others	0.61	0.12	-	0.73
Current Expenditure	**14.13**	**9.73**	**1.96**	**25.82**
Consumption and Operating Surplus	3.30	6.48	1.85	11.62
Interest Payments	1.89	0.48	0.01	2.38
Social Security Benefits	5.02	-	-	5.02
Current Transfers	3.83	2.77	0.11	6.71
Other Current Expenses	0.09	-	-	0.09
Current Savings	**3.42**	**–0.32**	**0.08**	**3.18**
Capital Revenue	**0.05**	**0.18**	**0.01**	**0.25**
Capital Outlays	**0.34**	**1.17**	**0.29**	**1.80**
Transfers from Upper Levels	**4.18**	**1.76**	**0.22**	**6.16**
Transfers to Lower Levels	**5.93**	**0.23**	**-**	**6.16**
Total Primary Surplus	**3.04**	**0.69**	**0.04**	**3.77**
Total Primary Surplus(1)	**2.98**	**0.51**	**0.03**	**3.52**
Financial Result	**1.15**	**0.21**	**0.03**	**1.39**

(1) Exclusive of Capital Revenue.

Source: Own estimates based on figures in Table 2.

convertibility in 2002, it was advisable to resort to 2003 data for calculating the model's optimal value for τ.

The choice of public spending series that would adjust best to a theoretical model of economic growth in which public facilities' congestion existed was addressed to by observing the performance of public spending as of the Nineties, as depicted by Table 1 and taking also into consideration the evidence given by Table 4.

Notwithstanding the fact that figures in Table 1 permit somehow to infer that congestion in public services and facilities is by all means a likely outcome, if proper attention is paid to the fact that the Argentine overall public expenditure (excluded interests and social security benefits) fell – in the period under analysis – from more than 15 to 13 points of its GDP;[15] the main bottleneck makes itself evident in capital outlays (embodying externality creating public investment), whose participation fell from an already low average figure of 1.8 points in the mid-Nineties to less than 1.0 point of GDP in the most recent years.[16] Therefore, and in the light of the mentioned empirical evidence, it appears reasonable to resort to data on public fixed capital stock on the understanding that they will better reflect the congestion hypothesis assumed in the theoretical model.

By furthering the empirical analysis, the evidence given by Table 4 strengthens even more the case for the use of public fixed capital stock (excluding private construction) in the determination of the optimal government size in Argentina. As may be seen, the 6.31 per cent rise in public construction during the period fell well short of overall capital stock and private construction, which exhibited rises of almost 23 and 30 per cent respectively; all the same, during the difficult 1999-2003 period for the Argentine economy, overall capital stock and private construction still managed to have an increase of 2.10 and 4.86 per cent whereas public construction practically stagnated and machinery and equipment fell by 6.21 per cent.

The preceding verification suffices to say that G in expression (1) above could be well represented by "Public Construction" as, in line with the theoretical foundations of the growth model resorted to, it embodies most of the fields in which users could more easily congest public facilities. Nevertheless, a closer analysis of Table 4 also avails the inclusion of "Domestic Transport Means and Materials" and "Machinery and Equipment" on grounds that these items also comprise diverse

---

[15] Proper attention means here that there are no grounds to believe that the the reduction in public spending – relative to GDP – was somehow matched by an enhanced productivity or quality of services rendered to the public.

[16] Although the thread of the argument still holds it must be said that, following the widespread privatization process that took place in the Nineties, private owned public utilities firms are now largely responsible for investment in communication, energy, transport and water distribution.

Table 4

Argentina: Aggregate Fixed Capital Stock, 1993-2003
(millions of Argentine pesos of 1993)

Year	Aggregate Capital Stock	Machinery and Equipment	Domestic Transport Means and Materials	Imported Transport Means and Materials	Private Constr-uction	Public Constr-uction
1993	543,164	103,648	18,234	7,621	279,367	116,514
1994	564,398	107,043	19,405	8,982	292,050	118,153
1995	580,001	108,105	20,118	10,067	301,848	119,163
1996	593,887	110,770	20,670	11,402	311,996	119,518
1997	615,345	115,737	21,295	13,179	323,615	121,055
1998	636,592	120,484	21,976	15,402	336,040	122,509
1999	652,937	122,817	22,069	16,800	345,894	123,922
2000	663,113	124,325	22,249	18,046	352,843	124,027
2001	668,841	122,441	22,379	18,868	358,850	124,100
2002	661,870	115,564	22,174	19,147	359,787	123,324
2003(1)	666,660	115,186	22,319	19,712	362,696	123,870
1993-2003	Δ 22.74%	Δ 11.13%	Δ 22.40%	Δ 158.65%	Δ 29.82%	Δ 6.31%
1999-2003	Δ 2.10%	Δ –6.21%	Δ 1.13%	Δ 17.33%	Δ 4.86%	Δ –0.04%

(1) Provisional data.

Source: DNCN-INDEC: PROJECT BID-UNPRE STUDY 1.EE.88 (2004), *The National Wealth in Argentina*. National Director: Lic. Fernando Cerro. Coordinator: Ariel Coremberg. August.

items subject to congestion investment in public services.[17] Let it be mentioned, in passing, that 1999-2003 figures show that public investment building-up did not keep in this case pace either with that of overall fixed capital stock or with the increase of GDP for what its performance will aid to better reflecting the theoretical concept underlying (G / Y) in expression (1).

[17] This still holds in the case of several public facilities whose services have been privatized in the Nineties, such as railways or underground trains, with the firms' express compromise of building up investment on account of the conceding government level.

In computing α, according to expression (15), figures (in constant prices) for the values of production (*PY*) and the aggregate fixed capital *rK* (excluding private construction) were estimated for the benchmark year 2003 according to the ensuing procedure: the value of production was obtained by multiplying the 2003 GDP by the coefficient relating the value of industrial production and the product in the 1997 Input-Output Matrix, that is:

$$1.517 \text{ x } 256{,}023.0 \text{ millions} = 388{,}387 \text{ millions}$$

The figure for *rK* resulted from adding machinery and equipment, transport means and materials and public construction;[18] that is:

$$115{,}186 + 0.78^{19} \text{ x } 42{,}031 + 123{,}870 = 271{,}840$$

By estimating next expression (15):

$$\alpha^{20} = 2 - (388{,}387 / 271{,}840) = 0.571$$

the value of τ^* can finally be achieved:

$$\tau^* = (1 - 0.571) / (2 - 0.571) = 0.30$$

Thus, this figure indicates the optimal government size, in terms of the long-run maximum economic growth rate determined by expression (10') above.

4. Optimal growth and fiscal sustainability

The immediate first conclusion, when comparing the arithmetical solution for equation (12) for the benchmark year (0.30) with the effective public spending share in the same year (27.62 per cent of GDP, Table 3 above, when the 1.39 per cent surplus is not considered) is self explaining: the actual government size in Argentina falls short of the optimal size required for long run economic growth, according to the model which explicitly accounts for the possibility of congestion in the use of public goods and facilities. In other words, the investment effort will have to be deepened in Argentina should the government expect to remove the negative impact of congestion upon long run economic growth.

Second, even though Table 3 showed that the three government levels runned altogether an overall surplus of 1.39 points of GDP in the benchmark year, the question may be raised of whether the Public Sector in Argentina is in a position of enlarging this fiscal surplus while at the same time doing away with distortionary taxes on exports and financial transactions.

---

[18] In order to keep coherence with the condition stated by expression (13) the used figure for Public Construction reflects the monetary value of public capital stock (stock in physical terms by its price).

[19] The rationale followed here was that as much as 75 to 80 per cent of Transport Means somehow serve a productive end, either in secondary or tertiary sectors and can therefore be considered part of fixed capital stock.

[20] All figures in million of Argentine pesos of 1993.

Having posed this challenge, the rest of this section is devoted to showing that there is in fact room in Argentina for a more efficient tax regime and yet producing revenue yields consistent with the requirements of the optimal government size, according to the endogenous model of economic growth developed in Section 2, and of long run fiscal sustainability that respects the necessary provision of public goods and services and meets the country's new financial commitments towards domestic and foreign creditors.[21]

The exercise rests on the assumption that the pressure already mounting over economic authorities will sooner or later lead to gradual reductions of export tariffs whereas, and by the same token, the tax on financial transactions could either disappear or be maintained with the possibility of using it as a tax credit for the Income Tax of Individuals and Firms.[22] Last but not least, suggestions for making the Tax Regime more efficient (by not curtailing through taxes individuals' and firms' right incentives) do not rule out the possibility of having also a more equitable Tax Sistem in terms of income distribution; this, not dealt with in this preliminary version of the paper, may be achieved by reducing the flat rate in VAT which – as all indirect taxation – hits more heavily to consumers placed in the lower income deciles. It goes without saying that the exercise's main appeal resides in showing that an equal yield scenario will be possible once all changes take place.

Simply put, the proposal deals on the one hand with a proven possibility of enhancing revenue yields of the three taxes that make up almost 50 per cent of overall tax revenues (see Table 15 in the Statistical Appendix), that is, Value Added Tax, Individuals' and Firms' Income Tax, and Employers' Contributions on the Payroll and, on the other, with the possibility of replacing the revenue yield of Financial Transaction Taxes and Export Tariffs, whose share in overall revenue reached 13-14 per cent according to 2003-04 figures. Such a fiscal re-engineering could only be possible by effectively curtailing tax evasion[23] which is reckoned[24] to be greater than 30 per cent, in the case of VAT, superior to 43 per cent in Individuals' Income Tax and not less than 38 per cent in Employers' Social Security Contribution, the latter based on recent reports on the amount of informal or not declared labour.

Although data on fiscal evasion are not so straightforwardly known in the Corporate Income Tax, it may be inferred that it is lower in large firms, whose accounting records permit their tax liability's better assessment and greater in

---

21 On the basis of the government's recent proposal to bondholders that closed on 25 February 2005.

22 This solution is favoured by many specialists on grounds that will help to check traditionally high evasion levels particularly in the Individual Income Tax.

23 By referring to evasion reduction as the mechanism upon which the proposal is founded, the point is here worthmentioning that the economic authorities in Argentina have also set in motion policies and devoted resources conducive to evasion curtailing.

24 Data from different Reports on Fiscal Evasion confirm in general figures mentioned. In this case, the percentage of evasion in Income and Value Tax was taken from the paper by Avramovich (2004).

middle sized or smaller companies whose annual balance sheets may not reflect the actual situation *vis-à-vis* their tax dues.[25]

Avramovich's estimation of evasion in Value Added Tax, for year 2003 and based on the methodology developed by the Federal Administration of Public Revenues (AFIP) of Argentina, is summarized in the ensuing table:

Table 5

Argentina: Evasion in Value Added Tax, 2003
(thousands of current Argentine pesos and percentage)

Presumed Real Tax Base	142,824,808
Declared Tax Base	99,232,591.2
Effective Tax Rate	21.11%
Potential Tax Yield	30,150,317
Actual Tax Yield	20,948,000
Evasion	9,202,317
Percentage of Evasion in VAT	30.52%

Source: Avramovich (2004).

In assuming that evasion in VAT could be checked by one fifth, by far much more modest a target that the one set by the Argentine economic and fiscal authorities, figures in Table 5 would now turn into the ones shown in Table 6.

In considering next how tax revenues from the Individuals' Income Tax would have behaved should evasion had been one fifth smaller in 2003 the following two features, emphasized by Avramovich in her paper and supporting figures in Table 7, are worth mentioning:

- the variety of personal deductions (medical expenses, pension payments, family allowances and specific deductions for the employed) and a relatively high threshold for non taxable minimum income reduce significantly the number of taxpayers;
- 97 per cent of the revenue is collected from taxpayers in population decile 10 and the remaining 3 per cent from those in the population decile 9.

---

[25] The size of the shadow economy could well be a proxy for inferring the evasion level in this tax. In this connection, Schneider and Klinglmair (2004) deemed that the shadow economy in Argentina reached 25.4 points of GDP in year 2000.

Table 6

Argentina: Value Added Tax Yield
under the Hypothesis that Evasion Is Reduced by One Fifth in Year 2003
(thousands of current Argentine pesos and percentage)

Presumed Real Tax Base	142,824,808
Declared Tax Base	107,946,992
Effective Tax Rate	21.11%
Potential Tax Yield	30,150,317
Actual Tax Yield	22,787,610
Evasion	7,362,707
Percentage of Evasion in VAT	24.42%
Additional Tax Yield	**1,839,610**

Source: Own estimates based on figures from Table 5.

Table 7

Argentina: Evasion in Personal Income Tax, 2003
(thousands of current Argentine pesos and percentage)

Presumed Real Tax Base	87,794,966.7
Effective Marginal Tax Rate	10%
Potential Tax Yield	8,779,496.7
Declared Tax Base	54,933,333.3
Effective Marginal Tax Rate	9%
Actual Tax Yield	4,944,000
Evasion	3,835,496.7
Percentage of Evasion in PIT	43.69%

Source: Avramovich (2004).

Table 8 shows the new values for revenue from the Individuals' Income Tax obtained by adopting a similar hypothesis of one fifth evasion reduction.

Although figures on evasion are rather scanty with respect to the Corporate Tax, contrariwise to other taxes, it is not adventurous to assume that possibilities of a revenues' better performance in the tax will certainly depend on the success in achieving a sizeble shrink of the informal economy in Argentina, given the straightforward relationship between the firms' sales and their tax base.

Table 8

Argentina: Personal Income Tax Yield under the Hypothesis that Evasion Is Reduced by One Fifth in Year 2003
(thousands of current Argentine pesos and percentage)

Presumed Real Tax Base	87,794,966.7
Effective Marginal Tax Rate	10%
Potential Tax Yield	8,779,496.7
Declared Tax Base	63,448,855.6
Effective Marginal Tax Rate	9%
Actual Tax Yield	5,710,397
Evasion	3,069,099.7
Percentage of Evasion in PIT	34.96%
Additional Tax Yield	**766,397**

Source: Own estimates based on figures from Table 7.

It is also true that in upholding the same hypothesis of one fifth reduction, in this case with respect to the shadow economy, will hardly result in a tax yield increase of similar proportions as firms now entering the formal circuit will not be the largest ones already making up – and assumedly with relatively low evasion levels – most of the Corporate Tax Revenue. Therefore, the assumption of a successful one fifth reduction of the shadow economy, from 25.4 to 20.32 points of the GDP, will be taken here to be conducive to only 15 per cent increase in the 2003 tax yield, as shown by Table 9 below.

Table 9

Argentina: Corporate Tax Yield under the Hypothesis that the Shadow Economy Is Reduced One Third in Year 2003
(thousands of current Argentine pesos and percentage)

Actual Tax Yield	8,559,000
Shadow Economy	25.4%
Corrected Shadow Economy	20.32%
Yield's Correction Coefficient	1.15%
Impact on CIT Yield	9,842,850
Additional Tax Yield	**1,283,850**

Source: Own estimates based on Schneider and Klinglmair (2004) and figures from Table 13.

Table 10 includes official statistical information on labour markets and the performance of the Tax Administration with relation to Social Security Taxes.

Table 10

Argentina: Labour Markets and Social Security Taxes in Year 2003

Total Employees and Workers	7,303,226
Declared Employees and Workers	4,528,000
Undeclared Employees and Workers	2,775,226
Average Monthly Wage(1)(2)	867
Total Annual Earnings of Declared(3)	47,109,312
Tax Rate	16%
Actual Yield of Employers' Contributions(3)	7,539,000
Percentage of undeclared labour	38%

(1) Declared labour only.
(2) In current Argentine pesos.
(3) Thousands of current Argentine pesos.

Source: Own estimates based on figures from the Ministry of Economy. Internet site: www.mecon.gov.ar

By adopting also the assumption that Undeclared Labour could be reduced by one fifth, in line with what has so far been done, Table 11 shows the figures that will result for Employers' Contributions in 2003.

The results shown by these tables were intended to show, for the benchmark year 2003, that there was ground to assert that evasion checking could be an alternative to revenues from economically unwanted taxes. Nevertheless, a static exercise falls short of yielding conclusive evidence as long run fiscal sustainability – more akin to dynamic scenarios – is what really matters in relation to economic growth. In this connection, Table 12 depicts results obtained when spending requirements for the optimal government size and needed efficiency enhancing changes in the Tax Regime, in order to render the latter less distorting, are matched within a period extending till 2008 with the government's enhanced financial situation brought about by improvements in its tax administration. In line with the need to assess dynamic fiscal sustainability, the simulation exercise was carried out on the following assumptions: as of 2005, the inflation rate exhibits decreasing annual figures of 10, 8, 6 and 4 per cent respectively, whereas the occurrence of positive economic growth is also assumed with the GDP experiencing a constant growth rate of 4 per cent per year; this permits in turn to achieve the corresponding additional revenue yields in value added tax, individuals' income tax, corporate tax and social security taxes as percentages of product once the reduction in evasion is accounted for.

Table 11

Argentina: Labour Markets and Social Security Taxes under the Hypothesis that Undeclared Labour Is Reduced One Fifth in Year 2003

Total Employees and Workers	7,303,226
Declared Employees and Workers	5,083,045
Undeclared Employees and Workers	2,220,181
Average Monthly Wage(1)(2)	867
Total Annual Earnings of Declared(3)	52,884,000
Tax Rate	16%
Actual Yield of Employers' Contributions(3)	8,461,440
Percentage of undeclared labour	25.33%
Additional Tax Yield(3)	**922,440**

(1) Declared labour only.
(2) In current Argentine pesos.
(3) Thousands of current Argentine pesos.

Source: Own estimates based on figures from the Ministry of Economy. Internet site: www.mecon.gov.ar

As the simulation mainly rests on the idea that – for the period under analysis – there will be an impact on revenues due to a once and for all successful evasion curtailing of 20 per cent in the four main national taxes, Table 12's upper part shows the corresponding additional revenue yields in value added tax, individuals' income tax, corporate tax and social security taxes, resulting from computing the reduction in evasion and once the product's benchmark figure was corrected by growth and inflation in order to correctly estimate improvements in the tax yield.

Second, and in line with the declared objective of improving the Tax Regime profile, by gradually doing away with distortionary taxation, Table 12 reflects the revenue's replacement of Financial Transactions Tax and Export Tariffs subject to the condition that the fiscal balance is not altered. The rationale resorted to here is that Export Tariffs are at present and on economic grounds the more damaging fiscal instrument since, to the negative impact upon the competitiveness of exporting sectors, it has to be added the inflationary risk derived from a rate of exchange conditioned by fiscal needs;[26] the proposal's core consists of a cumulative annual export tariff reduction reaching not less than 12.5 per cent of its present level.[27] As for Financial Transactions Taxes, the also proposed 12.5 per cent cumulative

[26] As the fiscal yield of export tariffs is based on two components: the rate of exchange and the international price of commodities, the latter's falls induces the government to intervine to keep a high exchange rate.

[27] The proposal considers both the cases of an annual 12.5 per cent linear reduction in all export tariffs or case by case reduction which final overall impact reaches 12.5 per cent.

Table 12

Argentina: Optimal Growth and Fiscal Sustainability as of 2005
(millions of current Argentine pesos)

ITEMS	2005	2006	2007	2008
Improvements due to a more effective tax administration	**11,764**	**13,213**	**14,565**	**15,754**
Additional Value Added Tax Yield	6,921	7,775	8,570	9,270
Additional Individuals' Income Tax Yield	1,224	1,375	1,515	1,639
Additional Corporate Tax Yield	2,942	3,305	3,643	3,940
Additional Social Security Taxes Yield	676	759	836	905
Overall budget surplus (1.39% of GDP)	**7,113**	**7,990**	**8,807**	**9,526**
Reductions proposed in tax revenues	**–2,642**	**–5,936**	**–9,814**	**–14,153**
Reduction in Financial Transaction Tax	–1,397	–3,138	–5,188	–7,482
Reduction in Export Tariffs	–1,245	–2,798	–4,626	–6,671
Additional Public Capital Outlays in line with requirements of Optimal Government Size	**–5,066**	**–5,690**	**–6,273**	**–6,785**
Financial Commitments to Public Debt Creditors(1)	**–2,805**	**–2,805**	**–2,805**	**–2,805**
Expected Fiscal Outcome	**8,364**	**6,772**	**4,480**	**1,537**

(1) Only interest payments have been considered.

Source: Own estimates based on figures from and from the Government's recent and accepted proposal for the debt in default.

reduction could either mean a change in the existing tax rate or its taxpayer's use as a tax credit applicable to Individuals' Income and Corporate Tax.[28]

Third, Table 12 also shows required additional public spending, as determined by the solution to the endogenous model of economic growth developed in Section 2. In reason of the alternative chosen for public spending and

[28] As mentioned above, the second possibility is favoured on grounds that it will help to reduce evasion without increasing fiscal pressure.

acknowledging that congestion mainly affects existing infrastructure stock, it goes without saying that is not envisaged in the simulation exercise a current spending increase but the formation of new public fixed capital stock.

Fourth, the case is also considered in Table 12 of the additional budgetary burden that new financial responsibilities towards domestic and foreign bondholders of the defaulted debt, following the recent response to the government's offer,[29] will impose to the public sector. In this case, the table includes only figures for interest payments (as capital amortization will be due only as of 2024) and acknowledges the financial surplus for the overall Public Sector in Argentine, which amounted in 2003 to 1.39 points of GDP.

Let it however an important conclusion, suggested by figures in Table 12 above, be stressed: notwithstanding the fact that the expected fiscal outcome shows fiscal surpluses all throughout the period considered, the latter shrink as the cumulative reduction in Transaction Tax and Export Tariffs takes place for what, and unless the growth rate increases or further evasion checking helps reinforcing tax revenues, a complete elimination of the former two taxes is not envisaged in the very short run.

5. Concluding remarks

The paper highlighted the relationship between public spending and the rate of economic growth, in the frame of a model of endogenous growth in which public services and facilities are subject to congestion.

A natural empirical extension consisted in comparing the optimal government size, as derived from the mentioned model, and the actual government size based on overall budgetary commitments of the three government levels in Argentina, including revenue items as well as expenditure items. Figures showing that the actual government size was slightly smaller than the optimal one hide however the fact that most public spending is devoted to non capacity creating outlays or to finance public services whose congestion level is much more difficult to assess whereas public investment (mainly public construction) in facilities like roads, transport and the like, which can more easily be congested by users, practically stagnated in the last five years.

In the light of the achieved results and of the evidence furnished by public spending figures in Argentina, a dynamic simulation exercise was intended whereby the gap between optimal and actual government size could be closed by resorting to the application of measures that meet, from the fiscal viewpoint, the long run requirements of positive economic growth.

It appears necessary, in the first place, and given the real risk of hindrance on growth likely to be imposed by public facilities' scarcity in the very short run, that

---

[29] At the closing date, on 25 February 2005, the proposal gathered an acceptance level of 76.06 per cent.

any expansion of expenditure be carried out at the expense of current spending share in total public spending.

Second, and from the revenue side, the exercise proved that additional financial needs, as well as revenues required to partially do away with damaging taxation as Financial Transactions Taxes and Export Tariffs, would not alter the fiscal balance provided that the extremely high evasion levels in main taxes (Value Added Tax, Income Tax and Social Security Contributions) could be reduced to more reasonable standards. As a matter of fact, the hypothesis of one fifth reduction in evasion sufficed, in the simulation carried out for the period 2005-08, to match the needed extra fiscal revenues.

Nevertheless, the simulation exercise gave clear evidence that a complete elimination of both distorting taxes would require further efforts in evasion curtailing, new tax instruments or higher growth rates, should the equal yield principle be met.

It is also worth mentioning that the exercise's results allowed also for the margin necessary in order that the additional financial burden, arising from the prospective settlement of the defaulted public debt, be met.

Last but not least, the paper's conclusions also pointed out that the results of the exercise carried out could only be conducive to long run dynamic fiscal sustainability if – and only if – the model's prediction of a constant and positive rate of growth of GDP is finally validated by reality.

STATISTICAL APPENDIX

Table 13

Argentina: Tax Revenues from All Government Levels
(millions of current Argentine pesos)

Items	2001	2002	2003	2004[1]
I. National Taxes				
Taxes on Income Benefits and Capital Gains of Individuals and Firms	**10,719**	**9,514**	**16,170**	**23,560**
Personal Income Tax	3,634	3,493	4,944	6,120
Corporate Tax	5,683	4,343	8,559	15,082
Taxes on Firm Assets	10	11	7	4
Taxes on Minimum Presumed Income	550	535	1,363	1,224
Taxes on Benefits Abroad	774	1,083	1,247	1,088
Others	68	49	50	43
Social Security Taxes	**8,683**	**8,841**	**10,628**	**13,601**
Employees' Contributions	2,164	1,894	2,373	2,768
Employers' Contributions	5,505	6,184	7,539	9,767
Self Employed Individuals	1,013	763	716	1,065
Taxes on Properties	**3,848**	**5,527**	**7,646**	**9,515**
Taxes on Financial Transactions	3,021	4,944	5,966	7,771
Taxes on Individuals' Assets	769	524	1,603	1,661
Others	57	60	77	83
Consumption Taxes	**21,725**	**22,285**	**28,976**	**40,461**
Value Added Tax	15,351	15,242	20,948	30,977
Taxes on Goods and Services	5,620	6,773	7,819	9,248
Fuel and Gas Taxes	3,420	4,484	4,973	5,380
Others	2,200	2,289	2,846	3,868
Others	**754**	**270**	**209**	**236**
Taxes on Foreign Trade and International Transactions	**1,185**	**6,398**	**11,394**	**13,642**
Import Duties	1,575	1,308	2,289	3,250
Export Tariffs (net of refunds)	(480)	3,800	7,845	8,708
Others	90	69	(106)	120
Others	**340**	**279**	**292**	**693**
TOTAL NATIONAL REVENUE	**46,501**	**51,622**	**73,740**	**99,908**
II. Provincial Taxation				
Taxes on Property	**3,178**	**3,028**	**4,079**	**4,881**
Taxes on Goods and Services Transaction	**5,593**	**6,145**	**8,848**	**10,890**
Others	**1,005**	**1,424**	**1,405**	**1,794**
TOTAL PROVINCIAL REVENUE	**9,775**	**10,596**	**14,332**	**17,565**
III. Municipal Taxes				
Taxes on Property, Business and Services	**5,274**	**5,696**	**7,690**	**9,382**
TOTAL MUNICIPAL REVENUE	**5,274**	**5,696**	**7,690**	**9,382**
TOTAL REVENUE	**61,550**	**67,914**	**95,762**	**126,854**

(1) Provisional figures.

Source: Ministry of Economy, National Direction of Fiscal Research and Analysis, internet site: www.mecon.gov.ar/hacienda

Table 14

Argentina: Tax Revenues from All Government Levels
(percentage of GDP)

Items	2001	2002	2003	2004(1)
I. National Taxes				
Taxes on Income Benefits and Capital Gains of Individuals and Firms	**3.77**	**3.54**	**5.17**	**6.27**
Personal Income Tax	1.28	1.30	1.58	1.63
Corporate Tax	2.00	1.62	2.74	4.01
Taxes on Firm Assets	0.00	0.00	0.00	0.00
Taxes on Minimum Presumed Income	0.19	0.20	0.44	0.33
Taxes on Benefits Abroad	0.27	0.40	0.40	0.29
Others	0.02	0.02	0.02	0.01
Social Security Taxes	**3.06**	**3.29**	**3.40**	**3.62**
Employees' Contributions	0.76	0.70	0.76	0.74
Employers' Contributions	1.94	2.30	2.41	2.60
Self Employed Individuals	0.36	0.28	0.23	0.28
Taxes on Properties	**1.35**	**2.06**	**2.45**	**2.53**
Taxes on Financial Transactions	1.06	1.84	1.91	2.07
Taxes on Individuals' Assets	0.27	0.19	0.51	0.44
Others	0.02	0.02	0.02	0.02
Consumption Taxes	**7.64**	**8.29**	**9.27**	**10.76**
Value Added Tax	5.40	5.67	6.70	8.24
Taxes on Goods and Services	1.98	2.52	2.50	2.46
Fuel and Gas Taxes	1.20	1.67	1.59	1.43
Others	0.77	0.85	0.91	1.03
Others	0.27	0.10	0.07	0.06
Taxes on Foreign Trade and International Transactions	**0.42**	**2.38**	**3.65**	**3.63**
Import Duties	0.55	0.49	0.73	0.86
Export Tariffs (net of refunds)	–0.17	1.41	2.51	2.32
Others	0.03	0.03	–0.03	0.03
Others	**0.12**	**0.10**	**0.09**	**0.18**
TOTAL NATIONAL REVENUE	**1.36**	**19.21**	**23.59**	**26.58**
II. Provincial Taxation				
Taxes on Property	**1.12**	**1.13**	**1.31**	**1.30**
Taxes on Goods and Services Transaction	**1.97**	**2.29**	**2.83**	**2.90**
Others	**0.35**	**0.53**	**0.45**	**0.48**
TOTAL PROVINCIAL REVENUE	**3.44**	**3.94**	**4.59**	**4.67**
III. Municipal Taxes				
Taxes on Property, Business and Services	**1.86**	**2.12**	**2.46**	**2.50**
TOTAL MUNICIPAL REVENUE	**1.86**	**2.12**	**2.46**	**2.50**
TOTAL REVENUE	**21.66**	**25.28**	**30.64**	**33.75**

(1) Provisional figures.

Source: Own estimates based on official figures for the GDP and of revenue data in Table 13.

Table 15

Argentina: Tax Revenues from All Government Levels

(yield percentage share in overall tax revenues)

Items	2001	2002	2003	2004(1)
I. National Taxes				
Taxes on Income Benefits and Capital Gains of Individuals and Firms	**17.42**	**14.01**	**16.89**	**18.57**
Personal Income Tax	5.90	5.14	5.16	4.82
Corporate Tax	9.23	6.39	8.94	11.89
Taxes on Firm Assets	0.02	0.02	0.01	0.00
Taxes on Minimum Presumed Income	0.89	0.79	1.42	0.96
Taxes on Benefits Abroad	1.26	1.60	1.30	0.86
Others	0.11	0.07	0.05	0.03
Social Security Taxes	**14.11**	**13.02**	**11.10**	**10.72**
Employees' Contributions	3.52	2.79	2.48	2.18
Employers' Contributions	8.94	9.11	7.87	7.70
Self Employed Individuals	1.65	1.12	0.75	0.84
Taxes on Properties	**6.25**	**8.14**	**7.98**	**7.50**
Taxes on Financial Transactions	4.91	7.28	6.23	6.13
Taxes on Individuals' Assets	1.25	0.77	1.67	1.31
Others	0.09	0.09	0.08	0.07
Consumption Taxes	**35.30**	**32.81**	**30.26**	**31.90**
Value Added Tax	24.94	22.44	21.87	24.42
Taxes on Goods and Services	9.13	9.97	8.17	7.29
Fuel and Gas Taxes	5.56	6.60	5.19	4.24
Others	3.57	3.37	2.97	3.05
Others	**1.22**	**0.40**	**0.22**	**0.19**
Taxes on Foreign Trade and International Transactions	**1.93**	**9.42**	**11.90**	**10.75**
Import Duties	2.56	1.93	2.39	2.56
Export Tariffs (net of refunds)	–0.78	5.60	8.19	6.86
Others	0.15	0.10	–0.11	0.09
Others	**0.55**	**0.41**	**0.31**	**0.55**
TOTAL NATIONAL REVENUE	**75.55**	**76.01**	**77.00**	**78.76**
II. Provincial Taxation				
Taxes on Property	**5.16**	**4.46**	**4.26**	**3.85**
Taxes on Goods and Services Transactions	**9.09**	**9.05**	**9.24**	**8.58**
Others	**1.63**	**2.10**	**1.47**	**1.41**
TOTAL PROVINCIAL REVENUE	**15.88**	**15.60**	**14.97**	**13.85**
III. Municipal Taxes				
Taxes on Property, Business and Services	**8.57**	**8.39**	**8.03**	**7.40**
TOTAL MUNICIPAL REVENUE	**8.57**	**8.39**	**8.03**	**7.40**
TOTAL REVENUE	**100.00**	**100.00**	**100.00**	**100.00**

(1) Provisional figures.

Source: Own estimates based on revenue figures in Table 13.

Table 16

Argentina: Tax Revenues from All Government Levels

(yield percentage share in tax revenues by government level)

Items	2001	2002	2003	2004(1)
I. National Taxes				
Taxes on Income Benefits and Capital Gains of Individuals and Firms	**23.05**	**18.43**	**21.93**	**23.58**
Personal Income Tax	7.82	6.77	6.71	6.13
Corporate Tax	12.22	8.41	11.61	15.10
Taxes on Firm Assets	0.02	0.02	0.01	0.00
Taxes on Minimum Presumed Income	1.18	1.04	1.85	1.22
Taxes on Benefits Abroad	1.66	2.10	1.69	1.09
Others	0.15	0.10	0.07	0.04
Social Security Taxes	**18.67**	**17.13**	**14.41**	**13.61**
Employees' Contributions	4.65	3.67	3.22	2.77
Employers' Contributions	11.84	11.98	10.22	9.78
Self Employed Individuals	2.18	1.48	0.97	1.07
Taxes on Properties	**8.27**	**10.71**	**10.37**	**9.52**
Taxes on Financial Transactions	6.50	9.58	8.09	7.78
Taxes on Individuals' Assets	1.65	1.01	2.17	1.66
Others	0.12	0.12	0.10	0.08
Consumption Taxes	**46.72**	**43.17**	**39.29**	**40.50**
Value Added Taxes	33.01	29.53	28.41	31.01
Taxes on Goods and Services	12.09	13.12	10.60	9.26
Fuel and Gas Taxes	7.35	8.69	6.74	5.38
Others	4.73	4.43	3.86	3.87
Others	**1.62**	**0.52**	**0.28**	**0.24**
Taxes on Foreign Trade and International Transactions	**2.55**	**12.39**	**15.45**	**13.66**
Import Duties	3.39	2.53	3.10	3.25
Export Tariffs (net of refunds)	–1.03	7.36	10.64	8.72
Others	0.19	0.13	–0.14	0.12
Others	**0.73**	**0.54**	**0.40**	**0.69**
TOTAL NATIONAL REVENUE	**100.00**	**100.00**	**100.00**	**100.00**
II. Provincial Taxation				
Taxes on Property	**32.51**	**28.58**	**28.46**	**27.79**
Taxes on Goods and Services Transaction	**57.21**	**57.99**	**61.73**	**62.00**
Others	**10.28**	**13.44**	**9.81**	**10.21**
TOTAL PROVINCIAL REVENUE	**100.00**	**100.00**	**100.00**	**100.00**
III. Municipal Taxes				
Taxes on Property, Business and Services	**100.00**	**100.00**	**100.00**	**100.00**
TOTAL MUNICIPAL REVENUE	**100.00**	**100.00**	**100.00**	**100.00**

(1) Provisional figures.

Source: Own estimates based on revenue figures in Table 13.

REFERENCES

Avramovich, M.C. (2004), "Recaudación Tributaria: El Desafío de Acabar con la Evasión Fiscal de un País", *Annals of the 37° International Meeting on Public Finance,* Córdoba, Argentina.

Barro, R.J. (1990), "Government Spending in a Simple Model of Economic Growth", *Journal of Political Economy,* No. 98, October, part II, pp. S103-S125.

Barro, R.J. and X. Sala-i-Martín (1992), "Public Finance in Models of Economic Growth", *Review of Economic Studies*, Vol. 59, No. 4, October, pp. 645-61.

—————— (1995), *Economic Growth,* McGraw Hill Inc., USA, p. 154-55.

Cass, D. (1965), "Optimum Growth in a Aggregative Model of Capital Accumulation", *Review of Economic Studies*, No. 32, July, pp. 223-40.

Koopmans, T. (1965), "On the Concept of Optimal Economic Growth", in *The Econometric Approach to Development Planning,* Amsterdam, North Holland.

Lucas, R.E.Jr. (1988), "On the Mechanics of Development Planning", *Journal of Monetary Economics*, Vol. 22, No. 1, July, pp. 3-42.

Ramsey, F. (1928), "A Mathematical Theory of Saving", *Economic Journal*, Vol. 38, No. 152, December, pp. 543-59.

Romer, P.M. (1986), "Increasing Returns and Long Run Growth", *Journal of Political Economy*, Vol. 94, No. 5, October, pp. 1002-37.

Schneider, F. and R. Klinglmair (2004), "Shadow Economies around the World: What Do We Know?", CESifo, Working Paper, No. 1167, April, Austria.

Solow, R. (1956), "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics,* No. 70, February, pp. 65-94.

Swan, T. (1956), "Economic Growth and Capital Accumulation", *Economic Record*, No. 32, November, pp. 334-61.

FEDERAL TAXES AND TRANSFERS ACROSS CANADA: IMPACT ON FAMILIES

Marie-Anne Deussing[*]

1. Introduction

In Canada, there is a continual and long-standing debate over the regional impact of federal spending and taxation. In order to shed light on this issue, federal fiscal balances are often used to characterize the provincial distribution of federal revenues and expenditures. What is typically overlooked, however, is the provincial distribution of federal taxes and spending across family income groups. The goal of this paper is to address this issue by analysing the distribution of federal taxes and transfers across provinces and across income groups, while taking into account the role of federal intergovernmental transfers (i.e., indirect transfers). Surprisingly, very few studies have analysed both the provincial differences in federal net transfers (*i.e.,* direct and indirect transfers received minus taxes paid) and their distribution across income groups. The most recent study "Where the Money Goes: The Distribution of Taxes and Benefits in Canada" by Finn Poschmann, dates back to 1998 and analysed the provincial distribution of federal taxes and transfers (including intergovernmental transfers) for 1997 across family income groups. His study concluded that the federal government collects taxes from low-income Canadians in high-income provinces in part to fund transfers to higher-income residents of poorer provinces.

This paper looks once again at the provincial distribution of federal taxes and transfers across provinces and across family income groups, this time using 2000 data. However, while both studies use Statistics Canada's Social Policy Simulation Database (SPSD/M) to derive distributional estimates for both federal taxes and transfers (direct and indirect), each is unique in its treatment of intergovernmental transfers. More specifically, in Poschmann's study, Equalization is treated as a tax point transfer. However, while it is often argued that Equalization funds lower provincial taxes rather than services, this approach makes some strong assumptions, especially with respect to the way Equalization tax points are distributed across income groups. As such, this paper takes a more neutral approach by treating Equalization as a block transfer. This approach has not only the advantage of requiring fewer and weaker assumptions, it is also more factual as Equalization is a federal cash payment to less prosperous provincial governments. A key result of this paper is that in contrast to Poschmann's study, smaller variations are found in federal net transfers among provinces for high-income groups, reflecting the

---

[*] Ministry of Finance/Ministère des Finances – Canada.

The views expressed in this paper are those of the author and do not reflect those of the Department of Finance. The author would like to thank Chris Matier, Finn Poschmann, Mélanie Raymond, Michel Poitevin, Émile Allie and Isabelle Amano for their very helpful discussions and comments.

sensitivity of the results to the treatment of intergovernmental transfers. As such, this study fails to support Poschmann's conclusion that lower-income Canadians in high-income provinces are funding transfers to higher-income residents of low-income provinces to a large extent.

The structure of this paper is as follows. Section 2 first provides some background and then discusses the methodology and the choice of income concept. Section 3 examines the distributional profile of federal taxes and Section 4 analyses the distributional profile of federal transfers (direct and indirect) across income groups and provinces. Section 5 presents the distribution of net federal transfers on families and Section 6 provides a summary and conclusions.

2. Background and methodology

2.1 Federal fiscal balances

Federal fiscal balances are often used to illustrate the provincial distribution of federal revenues and (current) expenditures. This balance represents the difference between federal expenditures made and federal revenues raised in each province. A province characterized by a positive (negative) balance is one that receives more (less) in federal expenditures than it contributes to federal revenues and is referred to as a "net recipient" ("net contributor").

Table 1 presents federal fiscal balances for the provinces in 2000.[1] As in previous years, Ontario, Alberta and British Columbia were net contributors, with Ontario contributing the most ($26.4 billion) and British Columbia contributing the least ($2.7 billion).

Per capita fiscal balances (Table 2) ranged from $5,145 in Prince Edward Island to –$2,389 in Alberta. Quebec's per capita fiscal balance stood at $411, the smallest of all net recipient provinces.

While federal fiscal balances provide a useful measure of the distribution of federal taxes and transfers across provinces, they convey very little information as to how federal taxes and revenues are distributed across family income groups within each province. As such, this paper attempts to provide a thorough analysis of the federal fiscal flows between household income groups within and across provinces in 2000.

---

[1] The federal fiscal balances are measured using the Provincial Economic Accounts annual estimates. In the Provincial Economic Accounts, provincial distribution of federal revenues is based on residence of the person or establishment making the payment, while federal expenditures are allocated according to where consumption of resources occurs. Furthermore, debt charges are allocated across provinces by population to better reflect the consumption of resources funded by the accumulated public debt. These data should in no way be interpreted as the benefit or cost of Confederation, since the data cannot capture the true economic impact of the federal government.

Table 1

Federal Fiscal Balance by Province, 2000

(millions of Canadian dollars)

	Program Spending	Debt Charges	Revenues	Fiscal Balance
Newfoundland	3,962	787	2,072	**2,677**
P.E.I.	1,171	202	663	**710**
Nova Scotia	7,221	1,378	4,451	**4,148**
New Brunswick	5,019	1,106	3,366	**2,759**
Quebec*	33,742	10,807	41,518	**3,030**
Ontario	40,360	17,117	83,911	**–26,434**
Manitoba	6,736	1,679	5,590	**2,825**
Saskatchewan	5,221	1,497	4,774	**1,944**
Alberta	9,345	4,407	20,942	**–7,190**
British Columbia	14,003	5,945	22,680	**–2,732**
Canada	129,108	45,070	190,893	**–16,716**

Note: These figures include an upward adjustment to both federal expenditures and revenues for the Quebec (CHST) Abatement.

Source: Statistics Canada, *Provincial Economic Accounts, 2001 Annual Estimates*, cat. 13-213-PPB, November 2002.

Table 2

Per Capita Federal Fiscal Balance by Province, 2000

(Canadian dollars per capita)

	Program Spending	Debt Charges	Revenues	Fiscal Balance
Newfoundland	7,378	1,466	3,858	**4,985**
P.E.I.	8,486	1,464	4,804	**5,145**
Nova Scotia	7,674	1,464	4,730	**4,408**
New Brunswick	6,648	1,465	4,458	**3,654**
Quebec*	4,573	1,465	5,627	**411**
Ontario	3,454	1,465	7,181	**–2,262**
Manitoba	5,878	1,465	4,878	**2,465**
Saskatchewan	5,109	1,465	4,671	**1,902**
Alberta	3,106	1,465	6,960	**–2,389**
British Columbia	3,450	1,465	5,588	**–673**
Canada	**4,196**	**1,465**	**6,204**	**–543**

Note: These figures include an upward adjustment to both federal expenditures and revenues for the Quebec Abatement.

2.2 The database

In this analysis, SPSD/M, release 9.0, is used to compute the distribution of federal taxes and transfers across provinces and (census) family income groups for 2000. The SPSD/M is a static microsimulation model that combines individual administrative data from 1997 T1 personal income tax returns and employment insurance claimant histories with 1997 survey[2] data on family incomes and on expenditure patterns. As such, it estimates taxes and transfers at the individual and household level and aggregates the results to arrive at provincial and national estimates.

2.3 The income concept

The income concept used for this distributional analysis is based on post-tax, post-transfer income, which is defined in SPSD/M as the sum of market income (income from employment, self-employment, investment and other private sources) and transfer income to persons (cash transfers from federal and provincial governments) less all taxes (personal income tax, indirect taxes, benefit repayments and employment insurance premiums).[3] Although other income concepts exist,[4] a post-tax, post-transfer income concept is considered to be the most appropriate for two reasons. First, it is reasonable to assume that families consider transfers to be part of their income given that they have full discretionary control over how they will spend the money they receive through them. Second, it is a comprehensive measure of income, and as such, the incidence rates that use this income base provide a clearer picture of how the government "takes" on the tax side, after it has "given" on the spending side.[5]

However, although federal taxes and transfers are expressed relative to the post-tax, post-transfer income concept, it is important to note that family total income (market income and direct transfers) is the measure used to differentiate the income groups in the distributional tables presented in the sections below. This treatment follows Poschmann (1998) and allows the reader to readily recognize his or her own status *vis-à-vis* provincial and national averages.

---

2 The survey includes the Survey of Consumer Finance (SCF), the Survey of Household Spending (SHS) and the Survey of Income and Labour Dynamics (SLID).

3 Following Poschmann (1998), some adjustments are made to the SPSD/M post-tax, post-transfer income concept. First, the employer share of EI premiums is attributed directly to households and is therefore included as taxes. As well, both CPP/QPP contributions and the resultant pensions are removed to better accord with the National Accounts definitions. The income concept also includes intergovernmental transfers, although the benefits that arise from direct government spending on goods and services are not included.

4 For a more detailed description of existing income bases, see Vermaeten, F., W.I. Gillespie and A. Vermaeten, "Tax Incidence in Canada", *Canadian Tax Journal*, Vol. 42, No. 2, 1994, pp. 353-54.

5 *Ibid,* pp. 354.

3. Federal taxation

This analysis begins by looking at the distribution of the federal tax burden across income groups and provinces. Following the previous study, the analysis generally assumes that the tax burden is borne by those paying the tax. Furthermore, keeping with the focus on the individual, the incidence of the corporate income tax is not estimated here.

Federal taxes included in this analysis can be classified into two categories. The first category, federal direct taxes, accounts for 57 per cent of federal revenues and includes federal income tax on personal and unincorporated business income as well as employee/employer Employment Insurance (EI) contributions. The second category, federal indirect taxes, makes up about 20 per cent of federal revenues and includes custom import duties, excise duties, excise taxes, other energy taxes, and the Goods and Services Tax (GST). Corporate income taxes and federal own investment income make up the remaining share of federal revenue, but these are not included in the analysis.

3.1 Personal Income Tax

SPSD/M draws on Revenue Canada's sample of T1 personal income tax returns and as such, models the personal income tax in considerable detail. Furthermore, to account for the Quebec (CHST) Abatement, which reflects 13.5 percentage points of Basic Federal Tax for Alternative Payments for Standing Programs, an upward adjustment is applied to the personal income tax collected in Quebec.

Table 3 shows the distribution of the federal personal income tax (PIT) burden across income groups and provinces. As a proportion of (post-tax, post-transfer) income, federal PIT rises quite steadily moving from lower to higher (total) income groups. Indeed, on average, for families in the lowest income group, PIT amounts to 2.7 per cent of their post-tax, post-transfer income, while for families in the highest income group, these taxes amount to 27.2 per cent on average.

Across provinces, there is very little variation in PIT rates for given income groups. Indeed, for families in the $20,000 to $30,000 income group, PIT rates range from 6.1 per cent for Newfoundland to 8.4 per cent for Alberta. This can be attributed in part to the fact that families in given income groups are subject to the same federal income tax rates, regardless of where they reside. However, variations in family characteristics and in other components of taxes and transfers can affect the PIT income shares.

Overall, given the progressive nature of PIT, residents of Ontario and Alberta, which both have higher-than-average per capita income, pay proportionately more federal income taxes. In fact, the average personal income tax rate of 18.7 per cent in Ontario is about 7 percentage points higher than that estimated for Newfoundland.

Table 3

Personal Income Taxes as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	1.3	6.1	8.2	11.9	14.1	15.9	18.8	25.3	**11.9**
PE	2.3	7.1	8.0	11.3	13.6	13.9	17.9	23.4	**12.3**
NS	2.2	6.5	11.1	14.0	15.7	15.9	18.3	24.1	**13.7**
NB	2.0	6.4	10.1	12.6	14.4	16.6	18.2	27.7	**14.0**
QC	2.4	7.5	10.7	13.5	15.6	17.9	19.8	27.9	**15.8**
ON	2.8	7.5	10.4	14.2	15.7	17.4	19.5	27.6	**18.7**
MN	2.5	7.6	10.6	13.1	14.6	16.4	17.6	24.1	**15.1**
SK	2.5	7.3	10.7	13.3	15.6	16.9	18.4	28.5	**16.0**
AB	3.3	8.4	11.1	14.5	16.7	18.0	19.3	26.2	**18.6**
BC	3.1	8.1	10.4	14.3	16.3	17.7	19.4	26.9	**17.3**
ALL	**2.7**	**7.5**	**10.5**	**13.9**	**15.8**	**17.5**	**19.4**	**27.2**	**17.3**

3.2 Employment insurance contributions

SPSD/M models employee EI contributions by drawing on an administrative database of employment insurance claim histories. However, since SPSD/M does not account for the employer portion of EI contributions, this analysis also attributes the employer share of EI contributions directly to the employees based on the assumption that the employer contribution to payroll taxes is ultimately borne by employees.

In 2000, the employee contribution rate was $2.40 for every $100 of insurable earnings, and the employer contribution rate was 1.4 times the employee rate or $3.36 per $100 of insurable earnings. Furthermore, the maximum annual EI contribution was $936 for an employee and $1,310 for an employer.

Table 4 provides the distribution profile of EI contributions. This distribution is progressive over the lower-income range, although it ceases being progressive and becomes regressive over the higher-income range because of the upper limit on EI contributions and because the higher-income households receive a substantial proportion of their income from self-employment[6] and investment income, which is not considered insurable earnings.

---

[6] This excludes self-employed fisherman, who contribute to the EI program.

Table 4

EI Contributions as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	0.5	2.4	3.5	4.8	5.2	5.8	5.5	4.5	**3.8**
PE	1.1	2.8	3.9	4.9	5.5	5.4	4.9	4.8	**4.2**
NS	0.9	2.8	4.5	5.0	5.4	5.6	5.2	4.0	**4.2**
NB	0.9	2.7	4.4	5.0	5.4	5.6	5.6	3.6	**4.1**
QC	0.9	2.9	4.7	5.7	5.6	5.8	5.7	4.0	**4.4**
ON	1.1	2.7	4.1	5.1	5.2	5.3	5.2	3.4	**4.1**
MN	1.0	2.7	4.0	5.1	5.2	5.5	5.4	3.9	**4.2**
SK	1.2	2.8	3.9	5.2	5.1	5.4	5.6	3.0	**4.0**
AB	2.3	3.6	4.4	5.5	5.5	5.1	5.2	3.2	**4.2**
BC	1.3	3.3	4.0	5.1	5.1	5.1	5.3	3.3	**4.0**
ALL	**1.1**	**2.9**	**4.3**	**5.3**	**5.3**	**5.4**	**5.3**	**3.5**	**4.2**

However, in contrast to the provincial average PIT rates, the average EI contribution rates across provinces are very similar, ranging from 3.8 per cent for Newfoundland to 4.4 for Quebec. Given the upper limit on annual EI contributions, one would expect higher-income provinces to have lower average effective EI contribution rates. This (all else equal) would likely be the case if employment rates were uniform across provinces. However, employment rates vary considerably across provinces, and as such, this affects the amount of EI contributions collected. Higher-income provinces have higher employment rates, resulting in a greater number of EI contributors, which helps to attenuate variations across provinces in terms of average EI contributions relative to post-tax, post-transfer incomes.

3.3 Indirect taxes

This study uses the federal commodity tax variable provided by SPSD/M to model the distribution of indirect taxes. This variable includes federal custom import duties, excise duties, excise taxes, other energy taxes, and the federal GST.

Table 5 presents the distributional profile of indirect taxes. In general, a regressive distribution is observed across income groups. Indeed, on average, for families in the lowest income group, indirect taxes amount to 7.7 per cent of their

Table 5

Indirect Taxes as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	5.7	6.3	5.9	6.1	6.3	5.9	5.3	5.0	**5.8**
PE	8.0	7.9	7.4	8.1	8.0	5.8	4.9	5.1	**6.9**
NS	6.7	6.9	6.8	6.3	6.2	6.1	5.7	4.6	**6.1**
NB	6.5	6.6	7.0	6.4	6.7	5.8	5.6	4.9	**6.1**
QC	7.3	7.0	6.5	6.4	6.4	6.1	5.6	4.7	**6.1**
ON	8.1	7.0	6.8	6.4	6.0	5.7	5.5	4.6	**5.7**
MN	6.3	6.3	6.2	5.9	6.0	5.5	5.2	4.6	**5.5**
SK	7.3	7.1	6.7	5.9	6.2	5.5	5.4	4.5	**5.9**
AB	9.1	8.1	7.6	7.2	6.9	6.2	5.9	5.0	**6.3**
BC	8.4	7.4	6.5	6.5	6.2	5.8	5.6	4.7	**6.0**
ALL	**7.7**	**7.1**	**6.7**	**6.5**	**6.3**	**5.8**	**5.5**	**4.7**	**5.9**

post-tax, post-transfer income, while for families in the highest income group, these taxes amount to 4.7 per cent of their income on average.

When looking at the distribution of indirect taxes across provinces, there are slight variations that exist for given income groups. Indeed, for families with incomes of less than $20,000, indirect taxes amount to 9.1 per cent of post-tax, post-transfer income for families in Alberta, while they equal 5.7 per cent of post-tax, post-transfer income for those in Newfoundland. Nevertheless, the dispersion of indirect tax rates narrows in the higher-income groups resulting in average rates ranging from 5.5 per cent in Manitoba to 6.9 per cent in Prince Edward Island.

3.4 Total federal taxes

Table 6 shows the distribution of the total federal tax burden across income groups and provinces. For the lowest income group, the average tax rate hovers around 11.5 per cent of post-tax, post-transfer income. With increasing incomes, the influence of the progressive personal income tax takes over, with average federal taxes in the neighbourhood of 35.4 per cent observed in the highest income group.

However, as a share of post-tax, post-transfer income, federal taxes across provinces for given income groups are relatively uniform. Indeed, for families with incomes between $50,000 and $60,000, federal taxes vary slightly from 25.5 per cent for Newfoundland to 29.1 per cent for Alberta.

Table 6

Federal Taxes as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	7.4	14.8	17.5	22.8	25.5	27.7	29.6	34.8	**21.5**
PE	11.4	17.8	19.4	24.3	27.0	25.2	27.7	33.3	**23.3**
NS	9.8	16.3	22.3	25.3	27.2	27.6	29.1	32.7	**23.9**
NB	9.5	15.6	21.4	23.9	26.5	28.0	29.4	36.1	**24.2**
QC	10.6	17.4	21.8	25.6	27.5	29.8	31.1	36.6	**26.3**
ON	12.0	17.2	21.3	25.7	27.0	28.4	30.1	35.6	**28.5**
MN	9.7	16.6	20.9	24.1	25.8	27.4	28.2	32.6	**24.8**
SK	11.0	17.3	21.3	24.3	26.8	27.8	29.4	35.9	**25.8**
AB	14.6	20.1	23.1	27.2	29.1	29.4	30.4	34.5	**29.1**
BC	12.8	18.8	20.9	25.9	27.6	28.5	30.3	34.9	**27.3**
ALL	11.5	17.5	21.5	25.6	27.3	28.7	30.2	35.4	**27.4**

Overall, the distribution of average federal taxes across provinces follows provincial income patterns. Total federal tax rates range from 27.3 per cent to 29.1 per cent for those provinces with higher-than-average income, while in the remaining provinces, the rate varies between 21.5 per cent and 26.3 per cent of post-tax, post-transfer income. This again primarily reflects the progressivity of the federal tax system.

These results are in line with those obtained by the Poschmann (1998) study. Indeed, Poschmann observed average total federal tax rates that ranged from 10 per cent for lower-income groups to 38.5 per cent for higher-income groups. Furthermore, he observed very little variation across provinces for given income groups, although average federal tax rates across provinces followed provincial income patterns, with British Columbia's average tax rate one-quarter higher than Newfoundland's rate.

4. Federal transfers

This section examines the distribution of federal transfers across provinces. Federal transfers are classified into two categories. The first, federal *direct* transfers to persons, accounts for 42 per cent of federal program spending and includes

elderly benefits, Employment Insurance (EI) benefits, Child Tax Benefit (CTB) transfers and GST credits. The second, federal intergovernmental or *indirect* transfers to persons, makes up 26 per cent of federal program spending and includes Equalization entitlements and the Canada Health and Social Transfer (CHST).

4.1 Federal direct transfers

4.1.1 Elderly benefits

Elderly benefits account for 44 per cent of federal direct transfers and can be divided into three categories. The first, Old Age Security (OAS) is a pension available to all residents of Canada 65 years of age and older who meet the residence requirements. OAS pensions are taxed under the personal income tax, and individuals with an annual income in excess of $57,879 must repay part or the entire maximum OAS pension amount. Furthermore, the full OAS pension is eliminated when a pensioner's net income is $94,148 or above. The full pension, which is provided to those who have lived in Canada for at least 40 years after age 18, was $419.92 per month in January 2000.

The second, Guaranteed Income Supplement (GIS) operates like a negative income tax program. In January 2000, single individuals with no income other than the OAS pension received $499.05 per month, and a married couple each received $325.06 a month. Benefits are reduced by 50 cents for each dollar of income (other than OAS pension) that the individual or couple receives.

The third, the Spouse's Allowance (SPA) is an income-tested benefit that is paid to the spouse of an OAS pensioner, or to a widow or widower. The recipient must be 60 to 64 years of age and have lived in Canada for at least ten years after the age of 18. The maximum SPA was $839.84 in 2000, and the benefit is reduced by 75 cents for each dollar of non-OAS income received by the recipient or couple.

Table 7 below shows the distribution of OAS/GIS/SPA benefits across income groups and provinces. As a proportion of (post-tax, post-transfer) income, OAS/GIS/SPA benefits decrease substantially moving from lower to higher (total) income groups. On average, for families in the lowest income group, elderly benefits amount to 32.5 per cent of their post-tax, post-transfer income, while for families in the highest income group, these benefits make up 0.2 per cent of their income. This decrease is observed primarily because seniors are found disproportionately in lower-income families and because these benefits are reduced if seniors receive non-OAS income.

Across provinces, for families in the less than $20,000 income group, there is significant variation as reflected in elderly benefit rates that range from 28 per cent for Alberta to 41.4 per cent for Saskatchewan. This variation can be attributed in part to the different demographic profiles of each province, with Alberta characterized by a younger population relative to all other provinces.

Table 7

OAS/GIS/SPA as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	32.8	8.6	2.9	2.0	0.7	0.5	0.0	0.0	**7.6**
PE	37.5	10.8	3.9	3.5	1.0	1.0	0.6	0.1	**7.3**
NS	31.2	9.4	4.6	3.2	1.6	0.8	0.4	0.1	**6.7**
NB	32.6	9.6	4.0	3.0	1.2	1.0	0.4	0.1	**6.8**
QC	34.0	9.4	3.4	1.7	2.3	1.2	0.6	0.2	**6.3**
ON	32.5	11.9	6.2	3.5	1.9	1.1	0.9	0.3	**4.2**
MN	32.5	10.1	5.9	3.9	1.5	0.9	0.6	0.1	**5.8**
SK	41.4	12.4	5.7	2.8	2.3	1.1	0.6	0.5	**7.1**
AB	28.0	12.0	6.4	2.6	1.1	1.1	0.6	0.2	**3.6**
BC	30.1	10.5	5.0	2.9	1.9	1.0	0.6	0.2	**4.9**
ALL	**32.5**	**10.8**	**5.1**	**2.8**	**1.8**	**1.1**	**0.7**	**0.2**	**5.0**

Overall, average elderly benefit rates follow provincial income patterns as these benefits target lower-income seniors. Elderly benefit rates range from 3.6 per cent to 4.9 per cent for higher-income provinces, while in the remaining provinces, the rate varies between 5.8 per cent and 7.6 per cent.

4.1.2 Employment insurance benefits

EI benefits make up about 18 per cent of federal direct transfers and are based on an individual's hours worked in a year, earnings and previous regional unemployment rates. In particular, for 2000, the minimum required number of hours for eligibility ranged from 700 hours over the last 52 weeks if the regional unemployment rate was 6.0 per cent or less, to 420 hours if the regional unemployment rate exceeded 13.1 per cent, with longer benefit periods the higher the unemployment rate.

Table 8 shows the distributional profile of EI benefits. The distribution of EI benefits is regressive up to the $30,000 to $40,000 income group, at which point the distribution of EI benefits becomes progressive. On average, families with incomes below $20,000 receive EI benefits equalling 2.0 per cent of their post-tax, post-transfer income, while families with incomes between $30,000 and $40,000 receive benefits amounting to 4.4 per cent of their post-tax, post-transfer income. As

Table 8

Employment Insurance Benefits as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	3.4	14.0	16.9	14.5	12.1	8.3	5.3	2.6	**9.3**
PE	5.5	11.3	14.1	10.0	10.3	5.2	2.6	1.7	**7.4**
NS	2.5	5.3	6.3	5.2	4.4	3.5	3.4	1.3	**3.8**
NB	3.4	9.9	9.5	8.4	6.7	3.7	3.6	0.7	**5.2**
QC	2.3	6.3	5.6	4.8	3.7	3.4	2.1	0.9	**3.2**
ON	1.4	2.1	2.5	2.2	2.0	1.5	1.1	0.5	**1.3**
MN	1.6	3.0	3.0	3.2	2.1	1.7	1.5	0.8	**1.8**
SK	1.4	3.0	3.4	2.2	2.4	1.3	1.6	0.7	**1.8**
AB	2.0	4.0	3.8	2.7	1.4	1.0	1.6	0.5	**1.5**
BC	2.2	4.3	4.7	3.0	2.4	2.8	1.9	0.8	**2.3**
ALL	**2.0**	**4.3**	**4.4**	**3.6**	**2.8**	**2.2**	**1.7**	**0.6**	**2.2**

a share of post-tax, post-transfer income, EI benefits then decline significantly for families in income groups $40,000 to $50,000 and above.

Moreover, for a given income group, the Atlantic provinces tend to have much higher EI benefit rates compared to other provinces. Looking across provinces, this is reflected in the average rates that range from 3.8 to 9.3 per cent for the Atlantic provinces compared to the average rates in other provinces that range from 1.5 per cent to 3.2 per cent of post-tax, post-transfer income. This can be attributed to the fact that the Atlantic provinces have higher unemployment rates, and as such, 1) more people are collecting benefits, 2) more families are eligible to receive EI benefits since it is easier to qualify, and 3) the benefit periods are longer.

4.1.3 Net employment insurance program analysis

Given that almost everyone who has employment income must make EI contributions, it may be of interest to look at the overall net program impact.[7] Table 9 sets out the results.

---

[7] The net EI program is defined as EI benefits less EI employee and employer contributions.

Table 9

EI Benefits Less Contributions
as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	2.9	11.6	13.4	9.7	7.0	2.5	–0.2	–1.9	**5.5**
PE	4.5	8.5	10.2	5.1	4.9	–0.2	–2.3	–3.1	**3.2**
NS	1.5	2.4	1.8	0.2	–1.0	–2.2	–1.8	–2.6	**–0.4**
NB	2.5	7.2	5.1	3.4	1.3	–1.9	–2.0	–2.9	**1.1**
QC	1.4	3.4	0.9	–0.9	–1.9	–2.4	–3.7	–3.1	**–1.2**
ON	0.3	–0.6	–1.6	–3.0	–3.2	–3.8	–4.0	–2.9	**–2.8**
MN	0.6	0.3	–1.0	–1.9	–3.1	–3.8	–3.9	–3.1	**–2.3**
SK	0.2	0.2	–0.5	–3.0	–2.7	–4.1	–4.1	–2.3	**–2.2**
AB	–0.2	0.5	–0.6	–2.8	–4.1	–4.1	–3.6	–2.7	**–2.7**
BC	0.9	1.0	0.7	–2.1	–2.7	–2.3	–3.4	–2.6	**–1.7**
ALL	**0.9**	**1.4**	**0.2**	**–1.7**	**–2.5**	**–3.2**	**–3.7**	**–2.9**	**–2.0**

For 2000, contributions exceed benefits paid, and therefore, on a net basis, the average rate is negative: an average of –2.0 per cent of post-tax, post-transfer family income. Furthermore, on a net basis, the EI program is generally progressive, with the exception of families with incomes of less than $20,000 and more than $100,000.

However, and more importantly, some considerable discrepancies exist across provinces for given income groups. Indeed, families in Alberta with incomes less than $20,000 are net contributors to the program, while families in Newfoundland with incomes between $60,000 and $75,000 are net recipients, receiving more from the EI program than they are paying into it. Furthermore, families in Ontario are net contributors to the EI program at all income groups except for those with incomes less than $20,000 while families in Newfoundland are net recipients at all income groups except for those with incomes above $75,000. Again, these discrepancies can be attributed to the difference in unemployment rates across provinces since these influence not only the number of hours of insurable employment required to be eligible to receive EI benefits, they also influence the length of the benefit period. For example, in Ontario where the unemployment rate is between 7 and 8 per cent, it would take 630 to 664 hours of insurable employment to qualify for 17 weeks of benefits. In contrast, in a high unemployment region in Newfoundland, where the

unemployment rate is over 16 per cent, it would take 420 hours of insurable employment to qualify for 32 weeks of benefits.

4.1.5 The Canada child tax benefit

Canada Child Tax Benefit (CCTB) transfers are responsible for about 12 per cent of federal direct transfers and can be broken up into two main elements. The first element is a basic benefit available to 80 per cent of families with children. The annual basic benefit in 2000 was $1,104 per child under age 18 for the first and second child in a family, and $1,181 for the third and each additional child. The basic benefit was taxed back (on combined net income of parents over $32,960) at 5 per cent where there were two or more children and 2.5 per cent if there was only one child.

The second element of the CCTB is the National Child Benefit Supplement (NCBS), which targets low-income families, and as such, for 2000, the maximum was paid only if family net income was less than $21,214. When family net income exceeds the NCBS threshold, the benefit is reduced by a percentage amount that depends on the number of children in a family. On average, the annual NCBS in 2000 was $977 per child under age 18 for the first child in a family, $771 for the second child and $694 for the third.

Table 10 sets out the distribution of CCTB transfers across provinces and income groups. There is an increase of 1.6 percent in average CCTB transfer rates as we move from families with incomes less than $20,000 to families with incomes between $20,000 and $30,000, mainly because young single mothers are found disproportionately in the less than $20,000 income groups while two parent families tend to have incomes of more than $20,000. This would tend to distort the CCTB transfer rates for the lowest income group. However, the distribution of CCTB is progressive for income groups of $20,000 and above, with families with income between $20,000 and $30,000 receiving 3.7 per cent of their post-tax, post-transfer income in the form of CCTB transfers while families in the higher income groups receive 0.4 per cent or less.

Across provinces, there is very little variation in CCTB rates as reflected in the average rates that range from 1.1 per cent for Alberta to 2.0 per cent for Newfoundland and Prince Edward Island. This is as expected since uniform CCTB transfer rates are applied across provinces for each income group and families belonging to a given income group are subject to the same claw back rules, regardless of where they live.

4.1.5 Refundable GST credit

The GST credit is a tax-free quarterly payment that helps individuals and families with low and modest incomes offset all or part of the GST and as such, it helps to compensate for the regressive nature of the GST. In 2000, the GST quarterly credit was $205 for each eligible adult and $107 per child under the age

Table 10

Canada Child Tax Benefits
as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	2.5	4.5	3.7	2.1	1.7	1.0	0.4	0.1	**2.0**
PE	1.0	3.7	4.9	3.0	2.3	1.3	0.5	0.1	**2.0**
NS	1.9	4.0	2.8	2.0	1.8	1.3	0.4	0.0	**1.6**
NB	2.5	3.8	3.7	2.0	1.7	0.8	0.4	0.0	**1.7**
QC	1.6	3.6	3.3	2.3	1.8	1.2	0.5	0.0	**1.5**
ON	2.0	3.9	3.3	2.1	1.7	1.0	0.4	0.1	**1.1**
MN	2.1	3.2	3.6	2.5	1.9	1.2	0.5	0.1	**1.5**
SK	2.3	4.3	4.2	3.2	2.2	1.3	0.4	0.1	**1.9**
AB	2.2	2.9	2.8	2.2	1.7	1.2	0.6	0.1	**1.1**
BC	2.9	3.8	3.0	1.8	1.8	1.1	0.4	0.0	**1.4**
ALL	**2.1**	**3.7**	**3.3**	**2.2**	**1.8**	**1.1**	**0.4**	**0.0**	**1.3**

of 19, and the credit was phased out for households with income above a threshold level of $32,500 if the household was comprised of a single person and $38,700 if it was comprised of a married/common law couple with two children.

Table 11 sets out the distribution profile for the refundable GST credit. Since the GST credit targets lower-income families, it is not surprising that the transfer rates are highest for those families in the less than $20,000 income group and lowest for those with incomes above $100,000.

There is also very little variation in GST credit rates across provinces for given income groups as reflected in average effective rates ranging from 0.5 per cent for Alberta to 1.0 per cent for Newfoundland.

Indirect taxes net of the GST credit

Given that the refundable GST credit is meant to offset the regressive nature of the GST, it may be of interest to look at the distribution of indirect taxes net of the GST credit. Table 12 sets out the results.

As a proportion of post-tax, post-transfer income, indirect taxes net of the GST credit increase as we move from the lower than $20,000 income group to the

Table 11

GST Credit as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	2.5	1.9	1.2	0.5	0.3	0.2	0.2	0.2	**1.0**
PE	2.5	1.8	1.2	0.6	0.3	0.3	0.1	0.2	**0.9**
NS	2.6	1.7	0.8	0.4	0.2	0.2	0.2	0.1	**0.8**
NB	2.6	1.7	1.0	0.4	0.2	0.2	0.2	0.1	**0.8**
QC	2.8	1.8	1.0	0.5	0.3	0.2	0.2	0.1	**0.8**
ON	3.1	1.8	1.0	0.4	0.2	0.2	0.2	0.1	**0.5**
MN	2.5	1.6	0.9	0.4	0.2	0.2	0.2	0.1	**0.6**
SK	3.0	1.9	1.1	0.4	0.2	0.2	0.1	0.1	**0.8**
AB	3.1	1.7	0.9	0.3	0.2	0.1	0.1	0.1	**0.5**
BC	3.3	1.8	0.9	0.4	0.2	0.1	0.1	0.1	**0.7**
ALL	**2.9**	**1.8**	**1.0**	**0.4**	**0.2**	**0.2**	**0.2**	**0.1**	**0.6**

Table 12

Indirect Taxes net of GST credits
as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	3.2	4.4	4.6	5.7	6.0	5.8	5.1	4.8	**4.8**
PE	5.5	6.1	6.2	7.5	7.6	5.5	4.7	4.9	**6.0**
NS	4.1	5.2	5.9	5.9	5.9	5.9	5.4	4.5	**5.3**
NB	3.9	4.8	5.9	5.9	6.5	5.5	5.4	4.7	**5.3**
QC	4.5	5.2	5.4	5.9	6.0	5.9	5.4	4.6	**5.3**
ON	5.0	5.2	5.9	6.0	5.8	5.5	5.3	4.5	**5.2**
MN	3.7	4.7	5.3	5.5	5.8	5.3	4.9	4.4	**4.9**
SK	4.2	5.2	5.6	5.4	5.9	5.4	5.2	4.4	**5.1**
AB	5.9	6.4	6.7	6.9	6.7	6.1	5.7	4.9	**5.8**
BC	5.2	5.6	5.6	6.1	6.0	5.6	5.5	4.6	**5.3**
ALL	**4.8**	**5.3**	**5.7**	**6.0**	**6.0**	**5.7**	**5.4**	**4.6**	**5.3**

$40,000 to $50,000 income group, suggesting that the refundable GST credit is in fact successful in eliminating the regressivity of the GST. However, with the GST credit substantially reduced for families with higher incomes, the regressivity of indirect taxes is still present among higher income groups, with rates ranging from 6.0 percent for families with incomes between $50,000 and $60,000 to 4.6 per cent for those with incomes above $100,000.

When looking at the distribution of indirect taxes net of the GST credit across provinces, it can again be said that the GST credit reduces the regressivity of the GST. Indeed, without the GST credit, the distribution of indirect taxes across provinces yields slight variations for given income groups (see paragraph 3.3). With the GST credit, the dispersion of indirect tax rates narrows somewhat. For example, for families with incomes of less than $20,000, prior to the inclusion of the GST credit, rates range from 9.1 per cent for Alberta to 5.7 per cent for Newfoundland. With the GST credit, rates now range from 5.9 per cent for Alberta to 3.2 per cent for Newfoundland, reflecting a 0.7 per cent decrease in variability.

4.1.6 Total federal direct transfers to persons

Table 13 provides the distributional profile of federal direct transfers. This distribution is progressive in the sense that the contribution relative to income is more important for lower-income groups. Indeed, direct transfers amount to 39.6 per cent of post-tax, post-transfer income for those in the lowest income group while higher-income groups receive transfers totalling 1 per cent of their post-tax, post-transfer income.

There is a considerable drop of almost 50 per cent in average direct transfer rates as we move from families with incomes less than $20,000 to families with incomes between $20,000 and $30,000. This is mainly because a high proportion of the elderly, who receive the bulk of federal direct transfers through OAS/GIS/SPA, fall into the less than $20,000 income group. Indeed, families with incomes less than $20,000 receive elderly benefits (OAS/GIS/SPA) equalling 32.5 per cent of their post-tax, post-transfer income, while those with incomes between $20,000 and $30,000 receive benefits amounting to 10.8 per cent (67 per cent less) of their post-tax, post-transfer income (see Table 7).

Furthermore, with the exception of the lower income groups, there is significant variation in federal direct transfer rates across provinces for a given income group. This variation is reflected in the average transfer rates that range from 6.7 per cent for Alberta to 19.9 per cent for Newfoundland, mainly because of the influence of the EI program. Indeed, families residing in Ontario with (total) incomes between $30,000 and $40,000 receive EI benefits which amount to 2.5 per cent of their (post-tax, post-transfer) income while families in Newfoundland in the same income group receive proportionately more (16.9 per cent) (see Table 8).

Overall, federal direct transfers follow provincial income patterns, with higher-income provinces experiencing rates between 6.7 per cent and 9.2 per cent

Table 13

Federal Direct Transfers
as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	41.2	29.0	24.7	19.0	14.8	10.0	5.9	2.9	**19.9**
PE	46.5	27.6	24.0	17.0	13.9	7.9	3.8	2.1	**17.6**
NS	38.2	20.4	14.5	10.8	7.9	5.7	4.4	1.6	**12.9**
NB	41.2	25.0	18.3	13.9	9.8	5.7	4.7	0.9	**14.5**
QC	40.8	21.1	13.3	9.3	8.1	6.0	3.4	1.3	**11.8**
ON	38.9	19.7	12.9	8.1	5.8	3.8	2.6	0.9	**7.2**
MN	38.8	17.8	13.4	9.9	5.8	4.0	2.8	1.1	**9.7**
SK	48.2	21.6	14.3	8.6	7.0	3.8	2.7	1.3	**11.5**
AB	35.3	20.7	13.9	7.7	4.4	3.4	2.8	0.9	**6.7**
BC	38.4	20.4	13.7	8.1	6.2	5.1	3.0	1.1	**9.2**
ALL	**39.6**	**20.6**	**13.8**	**9.0**	**6.7**	**4.6**	**3.0**	**1.0**	**9.1**

while lower-income provinces observed rates between 9.7 per cent and 19.9 per cent.

Poschmann observed broadly similar results with average federal direct transfer rates ranging from 36.9 per cent for families in the less than $20,000 income group to 1.5 per cent for those in the highest income group. Furthermore, across provinces, Poschmann observed substantial variations, due mainly to the effect of the differing prevalence of low-income families, with average transfer rates ranging from 9.3 per cent for Alberta to 26.8 per cent for Newfoundland.

4.2 Indirect transfers

In updating the Poschmann (1998) study, we also treat social services that are provided by provincial governments and funded by federal intergovernmental transfers as in-kind transfers; however, a different approach is taken in allocating these transfers to families. In the previous study, CHST cash transfers were allocated to health, post-secondary education and social services according to each province's historical share of CAP and EPF. In fiscal year 1995-96, prior to the consolidation of EPF and CAP, 43.1 per cent of the combined value of EPF and CAP cash transfers was allocated to health, while 14.5 per cent and 42.4 per cent was allocated to post-secondary education and social services, respectively.

Table 14

Share of Provincial Program Spending Related to Health, Education and Social Services by Province, 2000
(millions of Canadian dollars)

		Share of Social Spending		
	CHST	**Health**	**Education**	**Social Services**
NF	300	47%	31%	22%
PEI	75	49%	37%	15%
NS	513	52%	31%	17%
NB	408	54%	35%	11%
QC	4163	52%	33%	15%
ON	5105	56%	26%	19%
MN	619	53%	30%	16%
SK	534	55%	29%	15%
AB	1380	47%	39%	14%
BC	2356	48%	35%	17%
CA	**15453**	**52%**	**31%**	**16%**

Source: Department of Finance, Fiscal Policy Division.

This study takes a different approach in allocating CHST cash transfers. First, although CHST was intended to cover only certain provincial expenditures related to health, education, and social services, this study allocates CHST cash to all areas related to these services. Given that the CHST is a block transfer, it is reasonable to assume that provincial governments have considerable flexibility to allocate federal CHST cash transfers according to their spending needs. Moreover, in recent analyses of federal support for health care, federal and provincial governments adopt essentially the same approach.[8]

Table 14 below shows provincial spending related to health, education and social services across provinces as a share of provincial social program spending. Overall, health care spending commands a greater share of provincial social spending (52 per cent) while the smallest share (16 per cent) of provincial social spending is allocated to social services.

Table 15

Provincial Expenditures on Health, Education and Social Services
As a Percentage of Total Program Spending, 2000
(millions of Canadian dollars)

		Share of Program Spending			
	Equalization Entitlements[9]	**Health**	**Education**	**Social Services**	**Other Program Spending**
NF	1,138	34%	22%	16%	28%
PEI	273	30%	22%	9%	38%
NS	1,413	42%	25%	13%	19%
NB	1,255	40%	25%	8%	27%
QC	5,293	38%	24%	11%	27%
MN	1,291	41%	23%	13%	23%
SK	198	39%	21%	11%	28%
CA	**10,861**	**39%**	**24%**	**11%**	**26%**

Source: Department of Finance, Fiscal Policy Division.

Furthermore, this paper also takes a different approach in the allocation of Equalization entitlements. Contrary to Poschmann, who viewed Equalization as a tax point transfer, this study treats Equalization as another block transfer. Moreover, given that Equalization cash transfers have never been tied to any particular provincial expenditure, they are allocated to health, education and social services according to each program's share of total provincial program spending.

Table 15 describes provincial spending related to health, education and social services as a percentage of total program spending. Again, the bulk of provincial program spending is allocated to health, while the smallest share of total program spending is used to fund social services.

4.2.1 Federal transfers for health

In the previous study, Poschmann allocated 43 per cent of the CHST cash transfer to health for all provinces. This 43 per cent was the previous share of the combined value of EPF and Canada Assistance Plan that was set aside for health. It was then distributed across income groups according to the number of hospital

[8] *Federal Support for Health Care: The Facts.* Department of Finance, July 2002. The report can be accessed at http://www.fin.gc.ca/acces/fedprove.html

[9] Includes CHST Associated Equalization.

patient-days "consumed" by individuals by age, by sex and by province. The reasoning behind this allocation was that the number of hospital patient-days, by age and by sex represented the risk of an individual requiring health services, and as such, the dollar value of this risk was the portion covered by the federal cash transfer.

This study takes a more current approach and uses aggregate provincial spending patterns to allocate CHST cash and Equalization entitlements. As Table 14 shows, the share of CHST allocated to health ranges from 47 per cent of social spending for Alberta to 56 per cent for Ontario, while the share of the Equalization entitlement allocated to health (see Table 15) ranges from 30 per cent of total program spending for Prince Edward Island to 42 per cent for Nova Scotia.

For each province, the cash amounts of CHST and Equalization related to health care are then assigned to individuals based on their age and sex, using detailed estimates of public sector health expenditures from the Canadian Institute of Health Information (CIHI).

Figure 1 provides a look at public health expenditures for each province across three age categories: 1 to 44, 45 to 64, and 65 and above.[10] Interestingly, the share of provincial spending that is assigned to each age category varies across provinces. This can be attributed to 1) the unique demographic profile of each province and 2) the different spending choices made by each provincial government in the area of health care.

Figure 1 shows that all provinces, with the exception of Alberta, devote the largest share of their health spending to those aged 65 and above, with the spending shares ranging from 36 per cent to 43 per cent. Alberta spends the largest share (43 per cent) of its total provincial health care budget on individuals younger than 45 years of age and compared to all other provinces, it spends the least (36 per cent) on those 65 and above. This largely follows from the fact that relative to all other provinces, the population in Alberta is, on average, younger. Nova Scotia, on the other hand, spends the most (48 per cent) on its older generation relative to all other provinces, and the least (30 per cent) on those younger than 45 years of age, mainly because of the provincial government's health care spending choices.

Table 16 provides the distributional profile of health benefits derived from CHST cash transfers and Equalization entitlements by income and by province.

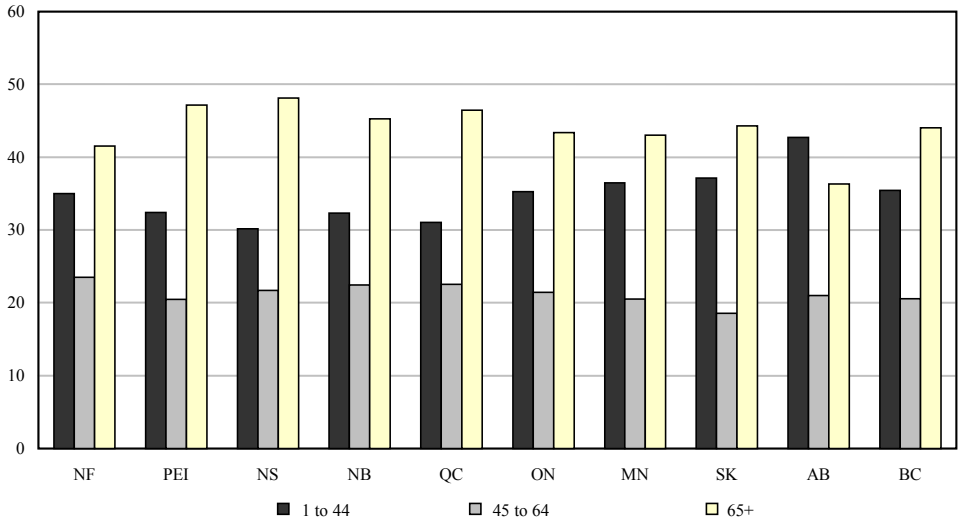
Across income levels, the distribution is progressive in the sense that benefits from the federal indirect health transfer relative to income are larger for lower-income groups. Indeed, the average health benefits derived from federal transfers ranges from 10.3 per cent for lower-income families to 0.7 per cent for higher-income families. However, there is a considerable drop (over 50 per cent) in average health benefit rates as we move from families with incomes less than $20,000 to families with incomes between $20,000 and $30,000. This is mainly

---

[10] Eight CIHI age groups, however, are used to allocate health spending to individuals.

Figure 1

Provincial Health Spending by Age
as a Proportion of Total Provincial Health Spending, 2000
(percent)



Source: Canadian Institute of Health Information, *National Health Expenditure Database*, *National Health Expenditure Trends, 1975-2002*.

because seniors are the most intensive consumers of health care services and are found disproportionately in families with incomes below $20,000.

Across provinces, considerable variation exists for given income groups, in part because Ontario, Alberta and British Columbia did not receive Equalization entitlements in 2000. However, considerable variation exists even among the provinces receiving Equalization transfers. For example, families in Saskatchewan with incomes less than $20,000 receive health care benefits amounting to 9.8 per cent of their post-tax, post-transfer income while families in Newfoundland belonging to the same income group receive 20.8 per cent. However, this variation is somewhat reduced in the higher-income groups. Overall, Ontario, Alberta and British Columbia have average effective benefit rates of 1.4 per cent, 1.3 per cent and 1.8 per cent, respectively, while in the remaining provinces, the range varies between 2.8 per cent and 7.9 per cent.

4.2.2 Federal transfers for education

In a similar way, Poschmann allocated 14.5 per cent (the historical share of EPF and CAP) of each province's CHST to post-secondary education. The

Table 16

CHST and Equalization Entitlements Allocated to Health as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	20.8	10.3	6.4	5.2	4.6	4.1	3.1	2.4	**7.9**
PE	19.2	9.8	5.9	5.0	4.4	3.4	2.8	2.3	**6.6**
NS	19.5	9.1	6.4	5.5	4.2	3.9	3.1	2.0	**6.8**
NB	20.2	9.6	6.8	5.8	4.6	3.8	3.2	1.9	**7.0**
QC	12.9	6.0	3.7	3.0	3.0	2.4	2.0	1.3	**4.1**
ON	6.3	3.0	2.0	1.5	1.2	1.0	0.9	0.5	**1.4**
MN	17.2	7.6	6.1	4.8	3.4	3.0	2.6	1.7	**5.1**
SK	9.8	4.4	3.0	2.0	1.9	1.5	1.3	0.8	**2.8**
AB	5.4	2.9	1.9	1.4	1.1	0.9	0.8	0.5	**1.3**
BC	7.0	3.0	2.1	1.5	1.3	1.1	0.9	0.6	**1.8**
ALL	**10.3**	**4.6**	**3.0**	**2.4**	**2.0**	**1.6**	**1.3**	**0.7**	**2.6**

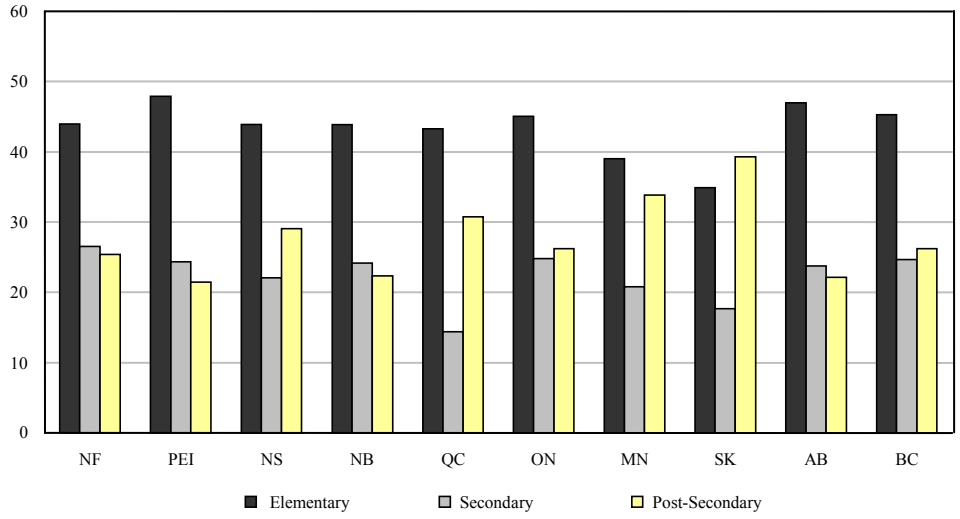
post-secondary education benefit was then allocated to students based on whether they were part-time or full-time enrolees. Each part-time student was given one-third the benefit of a full-time student. Subsequently, multiplying the number of students in each category by the share of CHST related to post-secondary education yielded the total notional post-secondary cash transfer.

In this study, however, given that CHST and Equalization are treated as block transfers, it is assumed that their share related to education funds education at all levels, including elementary and secondary levels. As such, the amount of CHST and Equalization transfers allocated to elementary, secondary and post-secondary levels is determined using the Financial Management System (FMS) publication. However, the FMS only classifies provincial education spending into two categories: 1) combined elementary and secondary education and 2) post-secondary education. Therefore, to determine the separate shares of provincial spending attributed to elementary and secondary education, enrolment rates are used since it is assumed that spending on elementary and secondary education is done on an equal per capita basis.

Figure 2 shows each province's spending pattern across different education levels. All provinces, except for Saskatchewan, spend a greater share of their education budget on elementary education, with Prince Edward Island spending the

Figure 2

Provincial Education Spending by Educational Level as a Proportion of Total Provincial Spending on Education, 2000



Source: Statistics Canada, *Public Sector Statistics, Financial Management System 2001-2002*, cat. 68-213-XIE, June 2002.

greatest share (48 per cent) relative to all other provinces. This can be attributed to the fact that more students attend elementary schools, given that these incorporate eight years of schooling as opposed to four provided by secondary and post-secondary institutions. On the other hand, relative to all other provinces, Saskatchewan spends the least (35 per cent) out of its education budget on elementary education, while it spends the most on post-secondary education. Given that Saskatchewan has the highest share of elementary students relative to all other provinces, a relatively lower spending share for elementary education reflects in part Saskatchewan's provincial government's budgeting decisions.

Moreover, the share of total spending on education that is allocated to the secondary level varies from 14 per cent for Quebec to 27 per cent for Newfoundland, with spending on secondary education exceeding spending on post-secondary education for Newfoundland, Prince Edward Island, New Brunswick and Alberta. This again in part reflects provincial government spending decisions.

Once the cash amounts of CHST and Equalization related to elementary, secondary and post-secondary education are established, they are assigned to individuals according to their age, highest level of education completed, and education status (*i.e.,* full-time or part-time). In particular, following Poschmann's

methodology, each part-time student receives one-third the education benefit of a full-time student.

Table 17 provides the distributional profile of education benefits across income groups and provinces. The distribution appears to be mainly proportional, although it becomes somewhat regressive through the highest income groups.

Across provinces receiving Equalization entitlements, there are relatively small variations within given income groups. For example, families in Quebec with incomes less than $20,000 have education benefits amounting to 1.8 per cent of their post-tax, post-transfer income while families in Newfoundland belonging to the same income group receive 5 per cent. Overall, average benefit rates range from 0.6 per cent to 1.3 per cent for Ontario, Alberta and British Columbia (non-Equalization receiving provinces), while the remaining provinces experience rates varying between 1.4 per cent and 5 per cent.

4.2.3 Federal transfers for social services

Although this study allocates the block transfers to social services according to provincial spending patterns rather than based on the federal portion of CAP

Table 17

CHST and Equalization Entitlements Allocated to Education as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	5.0	4.6	7.3	4.6	4.6	4.1	4.9	4.8	**5.0**
PE	3.2	4.0	6.3	5.6	4.9	7.3	3.6	2.5	**4.7**
NS	2.5	4.9	3.8	3.6	3.9	4.2	5.3	2.9	**3.9**
NB	4.0	4.0	5.5	4.1	4.2	3.9	5.0	2.2	**4.1**
QC	1.8	2.8	2.9	2.4	2.8	2.3	2.3	1.6	**2.3**
ON	0.8	0.8	0.9	0.8	0.8	0.7	0.7	0.4	**0.6**
MN	2.8	3.2	2.8	2.8	2.9	2.4	3.2	2.4	**2.8**
SK	2.1	1.5	1.2	1.2	1.5	1.2	1.2	1.2	**1.4**
AB	1.2	1.0	1.1	1.0	1.6	1.2	1.2	0.6	**1.0**
BC	2.0	1.6	1.4	1.4	1.5	1.3	1.1	0.8	**1.3**
ALL	**1.7**	**1.9**	**1.9**	**1.7**	**1.8**	**1.5**	**1.5**	**0.8**	**1.4**

Table 18

CHST and Equalization Entitlements allocated to Social Services as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	13.4	6.6	3.6	1.0	0.3	0.0	0.1	0.0	**3.8**
PE	8.8	3.5	2.6	0.4	0.5	0.0	0.4	0.0	**2.0**
NS	9.2	4.6	2.1	0.5	0.2	0.2	0.2	0.0	**2.2**
NB	7.0	2.4	1.1	0.2	0.2	0.0	0.0	0.0	**1.4**
QC	5.8	2.5	1.2	0.2	0.2	0.1	0.0	0.0	**1.2**
ON	3.3	1.8	0.9	0.3	0.2	0.1	0.0	0.0	**0.5**
MN	7.6	4.7	1.5	1.0	0.5	0.4	0.1	0.0	**1.6**
SK	3.3	2.4	1.0	0.5	0.2	0.0	0.0	0.0	**0.8**
AB	3.5	1.1	0.2	0.3	0.2	0.1	0.1	0.0	**0.4**
BC	4.4	1.2	0.7	0.2	0.1	0.0	0.1	0.0	**0.6**
ALL	**5.1**	**2.2**	**1.0**	**0.3**	**0.2**	**0.1**	**0.1**	**0.0**	**0.8**

money embodied in the CHST, it uses Poschmann's methodology to assign this amount across families. This study allocates the amount of CHST and Equalization attributed to social services according to the distribution of social assistance payments across income groups. This information is found in the SPSD/M database as it incorporates data from the Survey of Consumer Finances, and as such, records family welfare income.

As observed in Table 14, the share of CHST cash transfers allocated to social services ranges from 11 per cent for New Brunswick to 22 per cent for Newfoundland while the share of Equalization entitlements ranges from 8 per cent for New Brunswick to 16 per cent for Newfoundland (see Table 15).

Table 18 shows the distribution of social service benefits derived from CHST cash transfers and Equalization entitlements. Across income levels, the distribution of benefits related to social services is progressive, which is as expected given that social assistance payments target lower-income families. However, there appears to be some considerable variation across provinces for given income groups. For example, families in Newfoundland with incomes of less than $20,000 receive social service benefits amounting to 13.4 per cent of their post-tax, post-transfer income, while families in Saskatchewan and Ontario belonging to the same income group receive 3.3 per cent. Overall, average social service benefit rates follow provincial

income patterns, ranging from 3.8 per cent for Newfoundland to 0.4 per cent for Alberta. This follows largely from the fact that provinces with above average income do not receive Equalization payments and they also have a lower incidence of lower-income families.

4.24 Distribution of indirect transfers

Having examined at the distribution of health, education and social service benefits derived from CHST cash transfers and Equalization payments, it is now possible to assess the overall distribution of indirect transfers across income groups and provinces (see Table 19).

The distribution of the total indirect transfers financed by CHST cash transfers and Equalization entitlements is again quite progressive in that the transfer share is larger in lower-income families. Indeed, the average indirect federal transfer rate ranges from 17.2 per cent for lower-income families to 1.6 per cent for higher-income families. This result arises mainly because 1) the bulk of federal

Table 19

Federal Cash Transfers for Health, Education and Social Services as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	39.3	21.6	17.6	11.0	9.6	8.4	8.3	7.4	**16.9**
PE	31.3	17.5	15.1	11.3	10.1	11.1	7.1	4.9	**13.5**
NS	31.2	18.9	12.5	9.9	8.5	8.5	8.8	5.1	**13.0**
NB	31.5	16.4	13.8	10.5	9.4	8.1	8.8	4.3	**12.8**
QC	20.6	11.5	8.2	5.9	6.4	5.2	4.6	3.1	**7.8**
ON	10.4	5.6	3.9	2.7	2.1	1.8	1.6	0.9	**2.6**
MN	27.8	15.7	10.5	8.8	7.1	5.9	6.1	4.3	**9.6**
SK	15.3	8.3	5.3	3.9	3.7	2.9	2.7	2.1	**5.0**
AB	10.2	5.0	3.2	2.7	3.0	2.2	2.1	1.1	**2.7**
BC	13.4	5.9	4.2	3.1	3.0	2.5	2.1	1.5	**3.8**
ALL	**17.2**	**8.8**	**6.1**	**4.5**	**4.2**	**3.4**	**3.0**	**1.6**	**4.9**

indirect transfers is allocated to health and 2) seniors, who are the most intensive consumers of health care services, are found disproportionately in lower-income families.

When looking at the distribution of intergovernmental transfers across provinces for given income groups, considerable variations are observed. For families in the less than $20,000 income group residing in provinces receiving Equalization payments, rates range from 15.3 per cent for Saskatchewan to 39.3 per cent for Newfoundland, while those families residing in Ontario, Alberta and British Columbia observe indirect transfer rates varying between 10.2 per cent and 13.4 per cent. The dispersion narrows considerably in the higher income groups.

The pattern across provinces follows relative provincial income levels. However, the impact of including Equalization entitlements tends to inflate the size of federal indirect transfers (relative to income) for the Equalization receiving provinces. Ontario, Alberta and British Columbia (non-Equalization receiving provinces) have low average indirect transfer rates ranging from 2.6 per cent to 3.8 per cent. In all other provinces, this rate varies between 5.0 per cent and 16.9 per cent.

These results differ somewhat from those obtained by Poschmann (see Table 20). First, although the trend remains the same, Poschmann's benefit rates are somewhat smaller through the lower income groups and larger through the higher income groups. In general, he observed a progressive distribution across all income groups, with families in the lower-income groups receiving indirect transfers on average equalling 15.9 per cent of their post-tax, post-transfer income, while the highest-income families received indirect transfers equalling 3.3 per cent.

Second, Poschmann found smaller variations in indirect transfer rates across provinces for lower income groups as well as considerably greater variations in indirect transfer rates for higher income groups. For example, for families with incomes of less than $20,000 residing in provinces receiving Equalization payments, Poschmann found that indirect transfer rates ranged from 15.3 per cent for Saskatchewan to 27.1 per cent for Newfoundland. On the other hand, he found that families in the highest income groups residing in Equalization receiving provinces observed indirect transfer rates that ranged from 4.4 per cent for Saskatchewan to 35.7 per cent for Newfoundland.

It is also interesting to note that in Poschmann's study, for Equalization receiving provinces, the distribution of indirect transfers is progressive over the lower-income range, but ceases being progressive and becomes regressive over the higher income range. This can mainly be attributed to the fact that he treats Equalization as a tax transfer, which tends to benefit higher-income families more than lower-income families. However, the distribution of indirect transfers is progressive over all income groups for non-Equalization receiving provinces.

Across provinces, these results follow those obtained by Poschmann as the indirect transfer rates he observed also follow provincial income levels, with average federal indirect transfer rates ranging from 2.8 per cent to 3.1 per cent for

Table 20

Federal Cash Transfers for Health, Education and Social Services as a Percentage of Post-Tax, Post-Transfer Income, 1997 (Poschmann)
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	27.1	19.1	18.3	16.2	21.0	22.7	28.5	35.7	**23.0**
PE	21.6	15.3	13.1	13.6	13.5	13.8	14.5	16.9	**15.3**
NS	20.2	14.4	12.9	12.2	13.0	12.9	15.8	18.1	**15.0**
NB	20.8	13.5	12.0	12.3	11.6	12.7	15.3	18.3	**14.8**
QC	20.3	10.1	7.5	5.9	6.0	6.3	7.0	8.2	**9.1**
ON	11.1	6.9	3.5	2.7	1.7	1.5	1.3	0.9	**2.8**
MN	17.7	11.2	9.7	8.6	8.2	9.2	10.8	12.6	**11.0**
SK	15.3	7.7	5.3	4.0	4.1	4.0	4.5	4.4	**6.1**
AB	11.2	5.6	2.9	1.9	1.6	1.4	1.2	1.0	**2.9**
BC	15.9	5.4	3.5	2.2	2.0	1.3	1.4	0.8	**3.1**
ALL	**15.9**	**8.3**	**5.5**	**4.4**	**3.9**	**3.7**	**3.8**	**3.3**	**5.6**

non-Equalization receiving provinces, while all other provinces observed indirect transfer rates ranging from 6.1 per cent to 23 per cent.

4.2.5 Equal per capita *allocation across income groups*

A second approach that can be used to allocate federal indirect transfers to families is to distribute a province's CHST and Equalization transfers on an equal per capita basis. Table 20 sets out the results.

In this scenario, the distribution of federal indirect transfers remains progressive. However, the degree of progressivity is reduced somewhat, since, by distributing health, education and social service transfers equally across a province's population, these transfers are no longer attributed to particular types of families, and as such, the fact that lower-income families are predominant recipients of these transfers is ignored. Instead, a progressive distribution exists because transfers make up a larger portion of income for lower-income families.

However, as mentioned by Poschmann, the progressivity of our results may be somewhat overstated. First, although seniors are disproportionately found in lower-income groups based on their current income, they may in fact be drawing down savings made in earlier years. As such, this post-tax, post-transfer income base may not truly reflect their lifetime well being. Second, this study does not capture

the benefits of post-secondary education that accrue to higher-income individuals after they graduate. Indeed, although these individuals may no longer be students, they still enjoy a higher standing of living due to their extended years of education.

5. The distribution of net transfers

Having examined the allocation of federal taxes and transfers, we now have the components needed to calculate the federal net transfer for each family income group. The federal net transfer represents the difference between the amount a family receives from the federal government in terms of both direct and indirect (*i.e.,* social spending funded through federal intergovernmental transfers) transfers to persons and the amount paid in federal taxes. A positive (negative) net balance indicates that a family received more (less) in federal transfers than it paid in taxes. Table 21 sets out the results.

On average, families contribute to the federal government 13.3 per cent of their post-tax, post-transfer income, resulting in a net tax bill of about $4,773 (see Appendix). To a large extent, this net tax bill reflects the federal government's

Table 21

Federal Cash Transfers for Health, Education and Social Services as a Percentage of Post-Tax, Post-Transfer Income, Using an "Equal Per Capita" Imputation, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	27.0	21.4	20.0	16.2	14.8	13.3	11.2	8.6	**16.9**
PE	20.3	16.6	16.7	14.3	13.7	11.9	9.0	7.3	**13.5**
NS	21.9	17.0	14.5	13.0	12.2	11.7	9.5	6.5	**13.0**
NB	22.3	16.8	15.5	13.5	12.4	10.4	9.6	5.5	**12.8**
QC	13.6	10.3	9.2	8.2	7.7	7.1	6.1	3.8	**7.8**
ON	5.7	4.2	3.7	3.2	3.0	2.6	2.3	1.3	**2.6**
MN	17.3	12.6	12.1	11.0	9.7	8.9	8.2	5.4	**9.6**
SK	9.2	6.9	6.3	5.6	5.2	4.4	3.8	2.3	**5.0**
AB	6.3	4.1	3.7	3.4	3.1	2.8	2.5	1.4	**2.7**
BC	8.6	5.5	4.6	4.0	3.9	3.5	3.0	1.8	**3.7**
ALL	**10.8**	**7.6**	**6.7**	**5.9**	**5.4**	**4.7**	**3.9**	**2.1**	**4.9**

Table 22

Federal Net Transfers as a Percentage of Post-Tax, Post-Transfer Income, 2000
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	73.1	35.8	24.7	7.2	–1.1	–9.3	–15.4	–24.4	**15.2**
PE	66.3	27.3	19.7	4.1	–2.9	–6.1	–16.7	–26.3	**7.8**
NS	59.5	23.0	4.7	–4.6	–10.7	–13.4	–15.9	–26.0	**2.0**
NB	63.1	25.7	10.6	0.6	–7.3	–14.1	–15.8	–30.8	**3.2**
QC	50.7	15.2	–0.2	–10.4	–13.1	–18.6	–23.0	–32.1	**–6.7**
ON	37.3	8.2	–4.6	–15.0	–19.0	–22.7	–25.9	–33.8	**–18.7**
MN	56.8	16.9	3.1	–5.4	–12.9	–17.5	–19.3	–27.1	**–5.5**
SK	52.4	12.6	–1.7	–11.8	–16.0	–21.0	–24.0	–32.5	**–9.3**
AB	30.9	5.6	–5.9	–16.7	–21.7	–23.7	–25.5	–32.4	**–19.6**
BC	39.0	7.4	–3.0	–14.6	–18.3	–20.9	–25.3	–32.4	**–14.3**
ALL	**45.2**	**12.0**	**–1.6**	**–12.1**	**–16.5**	**–20.7**	**–24.3**	**–32.8**	**–13.3**

strong budgetary position. In 2000, the federal government posted a budgetary surplus of $16.7 billion, indicating that overall federal revenues exceeded federal expenditures.[11]

The distribution of net federal transfers overall is progressive, with families in the lowest income group receiving net transfers amounting to 45.2 per cent of their post-tax, post-transfer income on average while families in the highest income group contribute, on average, 32.8 per cent of their post-tax, post-transfer income to the federal government.

When looking at the distribution of net transfers across income groups, the first interesting inference that can be made is that for lower-income groups, there are considerable variations in the net federal transfers across provinces. For example, families with incomes of $30,000 to $40,000 in Quebec, Ontario, Saskatchewan, Alberta and British Columbia face net contribution rates ranging from 0.2 per cent to 5.9 per cent, while families in the Atlantic provinces and Manitoba are net recipients, with rates ranging from 3.1 per cent to 24.7 per cent. As well, families in Alberta with incomes between $50,000 and $60,000 have a net contribution rate of

[11] While the federal budgetary position in a given year tends to influence strongly the sign of the net balances, their relative positions can be used to make comparisons across time.

21.7 per cent, almost seventeen times more than families in Newfoundland belonging to the same income group. Also, census families in Alberta with incomes of less than $20,000 receive net transfers equalling 30.9 percent of their post-tax, post-transfer income while census families in Newfoundland with incomes between $20,000 and $30,000 receive benefits amounting to 35.8 per cent of their post-tax, post-transfer income. Much of these variations can be attributed to two factors. The first is the demographic profile of each province. For example, for a province characterized by an older population, such as Newfoundland, we would expect it to receive a larger share of federal direct transfers through OAS/GIS/SPA than Alberta, which has a population that is on average younger. Furthermore, we would expect health spending in Newfoundland to be higher than health spending in Alberta since seniors are the most intensive consumers of health care services. As such, since seniors are found disproportionately in lower-income families, and since OAS/GIS/SPA and health benefits make up the bulk of federal direct and indirect transfers, we would expect some variation to exist among these lower-income groups. The second factor that could explain the variations observed among the lower-income families is the inclusion of Equalization entitlements. We must not forget that including Equalization entitlements tends to inflate the size of federal indirect transfers (relative to income) for the Equalization receiving provinces and as

Table 23

Federal Net Transfers
as a Percentage of Post-tax, Post-transfer Income, 1997 (Poschmann)
(Census Family Total Income)

	≤ 20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ 100,001	All
NF	70.8	48.5	28.0	15.7	8.4	1.1	0.1	1.2	**25.4**
PE	61.2	36.8	19.8	7.1	–2.3	–8.9	–13.0	–12.8	**10.2**
NS	52.8	28.6	10.8	–1.1	–6.5	–10.4	–13.6	–18.0	**6.6**
NB	54.4	29.6	11.6	–1.9	–6.8	–11.3	–16.0	–14.8	**6.2**
QC	48.4	23.0	3.2	–7.6	–13.5	–17.7	–21.8	–27.0	**–2.2**
ON	36.2	18.8	–2.8	–12.3	–21.2	–24.9	–29.7	–36.9	**–16.1**
MN	43.8	22.4	2.1	–8.0	–12.1	–16.3	–20.8	–25.4	**–4.9**
SK	44.9	16.0	–1.5	–14.1	–18.2	–22.1	–25.6	–34.8	**–9.0**
AB	28.8	11.3	–9.0	–18.1	–21.9	–27.3	–31.4	–35.8	**–17.6**
BC	44.1	8.4	–5.2	–16.4	–21.4	–23.5	–29.9	–35.6	**–15.9**
ALL	**42.8**	**19.1**	**–0.5**	**–10.9**	**–17.9**	**–21.9**	**–27.0**	**–33.7**	**–10.7**

such, we would expect Ontario, Alberta and British Columbia (non-Equalization receiving provinces) to have lower indirect transfer rates than the remaining provinces.

However, more interestingly, in higher income groups, the dispersion of net transfer rates narrows considerably. Indeed, the families with incomes of $75,000 and above have broadly similar net federal transfer rates across provinces. For example, for families with incomes of $100,000 and above, net contribution rates range from 24.4 per cent for Newfoundland to 33.8 per cent for Ontario.

This latter result stands in contrast to Poschmann's findings (see Table 23); he found that considerable variations in federal net transfer rates existed even among higher-income families, with families in the $75,000 to $100,000 income group facing federal net transfer rates ranging from 0.1 per cent for Newfoundland to –31.4 per cent for Alberta. The discrepancy between this study and Poschmann's results is largely attributable to the difference in the treatment of indirect transfers, mainly Equalization entitlements, demonstrating that the results are sensitive to changes in the underlying assumptions. By treating Equalization as a block transfer rather than a tax transfer (a more appropriate treatment given that Equalization is a federal cash payment to the provinces) this study has reduced the regressivity in the distribution of indirect transfers that was experienced among the higher-income families of Equalization receiving provinces in Poschmann's study. As such, this analysis cannot lend strong support to Poschmann's conclusion that the federal government collects taxes from low-income Canadians in high-income provinces in part to fund transfers to higher-income residents in poorer provinces.

6. Conclusion

In order to reduce regional disparities and to ensure that provincial governments have sufficient revenues to provide reasonably comparable levels of public services at reasonably comparable levels of taxation, the federal government provides transfers to the provinces. However, given that spending priorities vary from province to province and that each province has unique demographic characteristics, the distribution of these transfers across families and provinces is difficult to discern precisely.

This paper revisits the question of how federal taxes and transfers are distributed across provinces and across income groups, a topic of an earlier study, "Where the Money Goes: The Distribution of Taxes and Benefits in Canada" by F. Poschmann. In doing so, this analysis has improved upon the treatment of intergovernmental transfers. Indeed, rather than using the historical share of EPF and CAP to allocate federal CHST cash transfers and rather than treating Equalization as a tax point transfer, federal CHST and Equalization cash transfers are treated as in-kind transfers to families that are allocated to health, education and social services according to provincial spending patterns, an approach currently adopted by federal and provincial governments to determine government support for health care.

The key results are as follows. First, the federal total tax incidence is found to be progressive for all provinces, with some variation across provinces for given income groups. This is mainly because of the influence of the progressive PIT. EI contributions are progressive up to an income of $50,000 and regressive thereafter and indirect taxes follow a regressive distribution.

Second, the relative size of federal direct transfers to persons varies significantly from province to province and across family income groups, a variation that is considerably greater than that observed for federal taxes. Furthermore, the impact of these transfers appears to be quite progressive, with average transfer rates declining sharply across income groups. Overall, these results are in line with those obtained in the previous study.

However, differences emerge once we analyse the distributional profile of indirect (intergovernmental) transfers across provinces. In this study, the distribution of total indirect transfers financed by CHST cash transfers and Equalization entitlements is progressive, with considerable variation across provinces for lower-income families and little variation across provinces for higher-income families. This stands in contrast to Poschmann's results. For Equalization receiving provinces, Poschmann observes a distribution of indirect transfers that is progressive across lower income groups but regressive across higher-income groups, with smaller variations across provinces for lower-income groups and significant variations among higher-income groups. This is further reflected in his net transfer rates that vary considerably across provinces for all given income groups. The results of this study, on the other hand, show little variation in federal net transfers among provinces for higher-income groups and as such, Poschmann's conclusion that low-income Canadians in high-income provinces are funding transfers to higher-income residents of low-income provinces cannot be strongly supported. However, in accord with the previous study, this study finds that some considerable variation in net transfers across provinces does exist among lower and middle-income groups.

APPENDIX

Table 24

Federal Net Transfers per Family, 2000

(Census Family Total Income, Post-tax, Post-transfer)
(Canadian dollars per Census Family)

	≤ $20,000	$20,001-30,000	$30,001-40,000	$40,001-50,000	$50,001-60,000	$60,001-75,000	$75,001-100,000	≥ $100,001	All
NF	11,095	8,216	7,637	2,528	–454	–4,539	–9,343	–20,121	**4,652**
PE	8,614	5,919	5,807	1,415	–1,187	–3,154	–10,368	–22,213	**2,438**
NS	7,911	5,255	1,335	–1,575	–4,332	–6,582	–9,914	–24,750	**619**
NB	8,244	5,854	3,063	197	–3,002	–6,997	–9,867	–32,153	**1,006**
QC	5,990	3,174	–67	–3,389	–5,124	–8,524	–13,190	–30,900	**–2,090**
ON	3,991	1,667	–1,241	–4,955	–7,604	–10,981	–15,682	–37,847	**–7,476**
MN	7,509	3,696	869	–1,857	–5,321	–8,455	–11,815	–25,950	**–1,956**
SK	5,699	2,532	–436	–3,857	–6,248	–9,949	–14,187	–33,646	**–2,946**
AB	3,215	1,125	–1,603	–5,564	–8,628	–11,509	–15,845	–35,891	**–7,874**
BC	3,968	1,475	–824	–4,812	–7,272	–9,982	–15,281	–34,038	**–4,839**
ALL	**5,089**	**2,472**	**–432**	**–4,014**	**–6,566**	**–9,896**	**–14,586**	**–35,091**	**–4,773**

REFERENCES

Canadian Institute of Health Information (2002), *National Health Expenditure Database, National Health Expenditure Trends, 1975-2002*.

Department of Finance Fiscal Policy Division (2002), *Federal and Provincial/Territorial Public Accounts, 1991-92 to 2001-2002*, November.

Poschmann, F. (1998), "Where the Money Goes: The Distribution of Taxes and Benefits in Canada", C.D. Howe Institute, Commentary No. 105, Ottawa, April.

Statistics Canada (2002), *Provincial Economic Accounts, Annual Estimates 1981-2001,* Catalogue No. 13-213-PPB, November.

Vermaeten, F., W. I. Gillepsie and A. Vermaeten (1994), "Tax Incidence in Canada", *Canadian Tax Journal*, Vol. 42, No. 2.

THE CROSS-COHORT DISTRIBUTION OF GOVERNMENT NON-RETIREMENT TRANSFERS AND ITS IMPACT ON WORKING AND EARNING

Jagadeesh Gokhale[*]

1. Introduction

During the last four decades, the composition of federal spending has shifted significantly from the provision of physical public goods – requiring government purchases for defense, public infrastructure etc. – to the provision of social insurance benefits that mainly involve tax transfer programs. Past expansion in federal retirement and health care programs has been studied extensively. Less attention has been devoted to growth in federal non-retirement transfers such as education and training subsidies, child-care benefits, unemployment support, and others. These transfers have more than doubled since the early Sixties from 1.9 per cent of GDP to 4.2 per cent today.

Federal non-retirement transfers could fulfill several roles. First, they constitute welfare payments meant to sustain the consumption of those affected by unforeseen economic shocks against which private insurance is not available or turns out to be inadequate *ex post*. This role of non-retirement transfers is often cited as the major motivation for undertaking them. Here, it is termed as a "defensive" role – to counter the loss of resources for current consumption from economic misfortunes.

Second, non-retirement transfers could fulfill an "offensive" role by improving the functioning of the economy and markets. This role, too, is a major motivation for some types of transfers. Their provision may help recipients and others to achieve better economic outcomes in the future such as higher employment and higher productivity and earnings.

Although no public transfers are motivated by a desire to worsen future economic outcomes, some non-retirement transfers may exert negative economic effects. Indeed, generous provision of welfare benefits could weaken economic performance by increasing dependency on public support, reducing labor force participation, and discouraging saving. Thus, a third possibility is that non-retirement transfers worsen future economic outcomes for recipients. This role of non-retirement transfers is termed "regressive."

The rationale for most government-provided transfers – unemployment insurance, child support, housing and energy assistance, health-care subsidies etc. is "defensive" – to support those suffering bad economic outcomes and misfortunes. In

---

[*] Senior Fellow, Cato Institute, Washington D.C. This paper was prepared for presentation at the Bank of Italy's conference on fiscal policy, March/April 2005. The author thanks William Niskanen and Doris Prammer for helpful comments and Arash Molavi Vassei and Roger D'Souza for excellent research assistance.

contrast federal education and training subsidies have an explicitly "offensive" rationale – to improve economic outcomes for their direct recipients and the operation of markets and the economy generally. However, transfers undertaken mainly for "defensive" purpose may also exert "offensive" effects. And those undertaken mainly for "offensive" reasons may not be very effective and may, indeed, exert "regressive" effects. The objective of this paper is to empirically examine whether federal transfers on account of unemployment insurance, child-care benefits, education and training subsidies and other non-retirement transfers exert mainly "offensive", "defensive" or "regressive" economic effects on recipients future economic outcomes.

To the author's knowledge, there have been no studies addressing the issues outlined above. The traditional literature has mostly tackled the question of whether government expenditures on public infrastructure and R&D outlays have productive effects. Most such studies examine whether that type of public spending affects the productivity of private firms.[1] Traditional public finance terminology does not attach the label of "productive" to public transfers because they are mainly intended to support recipients' consumption. Nevertheless, a useful distinction could be made between those non-retirement transfers that are fully consumed without any significant future repercussions versus those that exert positive (offensive) and negative (regressive) effects on recipients' future economic outcomes.

Whether or not public non-retirement transfers generate "offensive" effects is relevant for the ongoing debate on Social Security reform in the United States. One way to support the projected larger proportion of older individuals in the population is to save and invest ahead of their retirement. Amendments to U.S. Social Security laws in 1983 resulted in payroll tax surpluses that are expected to continue until the year 2017. However, those surpluses are invested exclusively in non-marketable U.S. Treasury securities and recent studies (Smetters, 2002; Nataraj and Shoven, 2004) suggest that not only does the federal government spend those surpluses but that their availability induces yet additional federal spending on non-Social Security federal operations. These studies imply that rather than saving those surpluses to help pay for the Social Security benefits of baby-boomers upon their retirement, the surpluses are being dissipated on other federal spending.

Because the share of federal spending on physical goods and services (investment in infrastructure, defense, R&D etc.) has been declining and the share of non-physical transfers has been increasing, one can deduce that any incremental resources available to the federal government – such as Social Security payroll tax surpluses – are directed toward the latter outlays.[2] That is, Social Security surpluses and the induced additional federal spending are used to finance the growth in federal transfers: Medicare and Medicaid benefits, unemployment support, education subsidies, child-care assistance and so on. Of these, retirement transfers would,

---

[1] For example, see Holtz-Eakin, D. (1992), *Public Sector Capital and the Productivity Puzzle*, National Bureau of Economic Research, Working Paper, No. 4122, July.

[2] The author gratefully acknowledges a discussion with Blair Comley on this point.

obviously, not improve markets and the economy directly, but non-retirement transfers could.

If non-retirement transfers, indeed, generate positive labor market outcomes for their recipients, the fact that Social Security surpluses are available for current government spending would not, in and of itself, imply that payroll tax surpluses are wastefully dissipated. On the contrary, such spending would constitute an effective "storage technology" – expanding economic output and the future tax base for meeting the government's future Social Security benefit obligations.

Federal non-retirement transfers and subsidy programs include education and job-training subsidies, unemployment insurance, and child support programs. These programs transfer resources to people during their working lifetimes, potentially assisting them to be more productive. In addition, they may have positive effects on non-recipients because higher education and lower unemployment among their direct recipients promotes better social and market environments.

It is well known that measuring the indirect effects of federal transfers on non-recipients is extremely difficult, if not impossible. Unfortunately, isolating the direct effects of particular federal expenditures on recipients' labor productivity is also very difficult because micro-survey datasets that track individuals and households over long periods of time do not contain adequate information on receipts from government non-retirement transfers. And micro-surveys that include detailed information on such federal transfers do not follow the same individuals over sufficiently long periods of time.[3]

Given these constraints, this paper's limited objective is to examine the distribution of federal non-retirement spending and to measure its direct impact on labor market outcomes at a *cohort* level of aggregation where cohorts are defined by single year of birth and gender of the family head. A dataset is constructed of cohort averages for earnings, labor-force participation, part-time/full-time status, demographic characteristics, and receipts of federal non-retirement transfers. The estimates are based on the Current Population Survey's March Supplement files spanning the years 1988 and 2001. This period is especially relevant for the Social Security reform debate because significant payroll tax surpluses began to accrue only by the late Eighties.

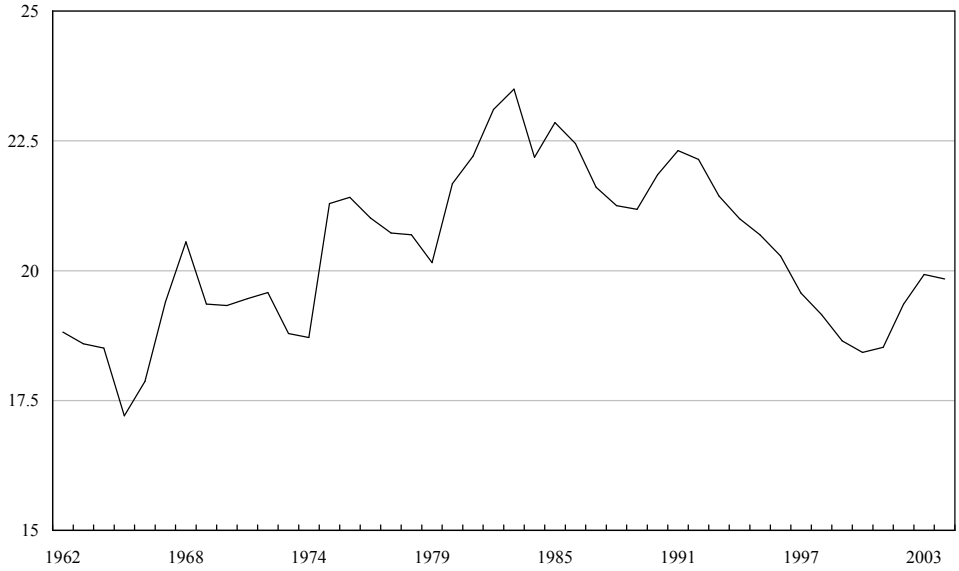
The first section below describes how the composition of federal expenditures has changed over the past four decades. It shows that along with retirement transfers, non-retirement transfers have also grown significantly as a share of GDP. Section 2 discusses the potential justifications for non-retirement transfers by classifying their objectives under "defensive," "offensive" and "regressive" categories. Section 3 describes the construction of the cohort dataset used to analyze the effects of

---

[3] The Survey of Income and Program Participation interviews households for a maximum of 32 months. Another potential data set is that of the University of Michigan's Panel Study of Income Dynamics which contains data on households between 1968 and 1999. However, that survey is known to over-sample relatively poorer households and does not contain detailed information on income on public transfers and subsidies.

Figure 1

Federal Outlays
(percent of GDP)



non-retirement transfers on cohort labor market outcomes. Section 4 describes the two-stage estimation method used to determine whether particular federal transfers exert any of the three effects described in Section 2. Section 6 provides details about the cohort data set used in the estimation; section 7 describes the findings; and Section 7 concludes the paper.

2. Changing composition of government spending

United States' federal spending as a percentage of U.S. GDP has remained stationary at about 20 per cent since the early Sixties (Figure 1). The constancy of the overall GDP share masks significant changes in its underlying components. Discretionary spending – that is directly determined by Congress via annual appropriations – has declined from 12.6 per cent of GDP in the early Sixties to about 6.5 per cent today (Figure 2). Almost all of that change can be explained by the decline in defense spending. This decline has accommodated a secular increase in mandatory spending – that is mainly determined by factors not within lawmakers' short-term control – such as demographics and business cycles (Figure 3).

As a share of GDP, federal mandatory spending increased from about 4.7 per cent during the early Sixties to 10.6 per cent by 1983, after which has remained in the vicinity of 10 per cent (Figure 3). This increase reflects the growing generosity

Figure 2

Components of Federal Spending – I
(percent of GDP)

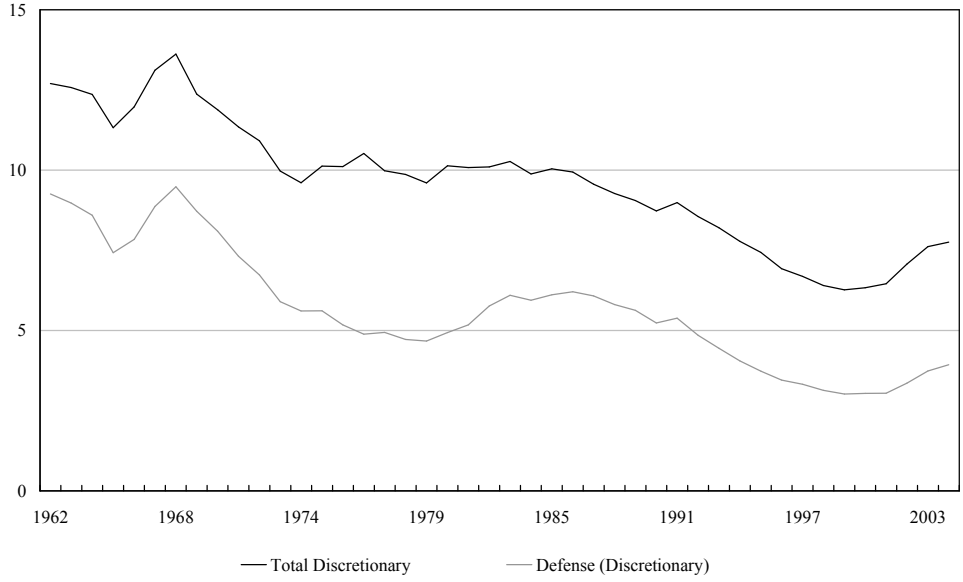
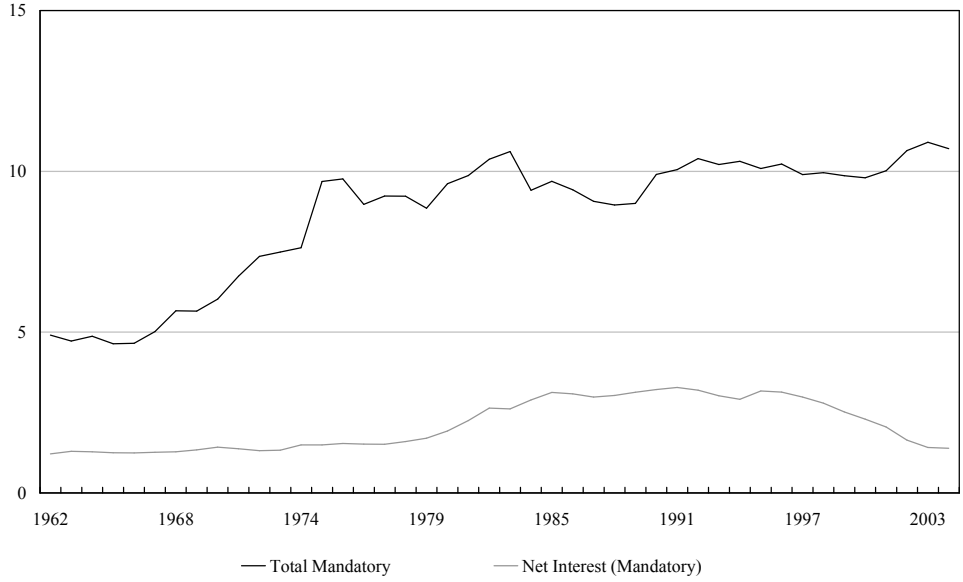


Figure 3

Components of Federal Spending – II
(percent of GDP)



of benefit payments – especially health care payments – and an increasingly longer-lived population requiring retirement and health care support.[4]

The fluctuations in federal payments to individuals over time closely match those in total federal mandatory outlays. Retirement transfers have increased from 3.2 per cent of GDP in 1962 to almost 8 per cent by 2004. Of the 4.7 percentage points increase in retirement transfers, 2.5 percentage points occurred due to growth in Medicare and Medicaid outlays.

Non-retirement payments to individuals have also increased: these transfers provide education subsidies, employment and training services, unemployment insurance, child-care assistance, community development, general welfare payments and so on.[5] The share of non-retirement transfers to individuals has more than doubled from 1.9 per cent of GDP in 1962 to 4.2 per cent today (see Table 6 in the Appendix).

The distribution of spending across retirement and non-retirement programs has shifted in favor of the former: In 1962, retirement transfers were 74 per cent larger than non-retirement transfers whereas they are larger by 91 per cent today. Again, most of this shift can be accounted for by rising federal outlays on retiree health care benefits.

Distributions of federal outlays per capita across male and female cohorts are shown in Figures 5 and 6. These include federal retirement benefits, non-retirement transfers, and non-transfer outlays. The Appendix provides details about the outlay categories and method used to estimate the distributions by age and gender. It is evident that total outlays on young cohorts have not increased significantly during the period considered, 1988 through 2001. Almost all of the increase over time in per capita outlays is concentrated among the oldest age groups. Again, increasing end-of-life health care expenses explain most of this increase.

Figures 7 and 8 contain distributions of federal non-retirement transfers – those outlays remaining after eliminating retirement benefits and non-transfer outlays from total federal outlays. The distribution of non-retirement transfers is considerably more volatile across male cohorts than across female ones. Middle-aged female cohorts receive significant amounts of non-retirement transfers on account of public assistance, child care, and other programs. Finally, for both genders, real non-retirement transfers appear to have increased substantially during the 1988-2001 period.

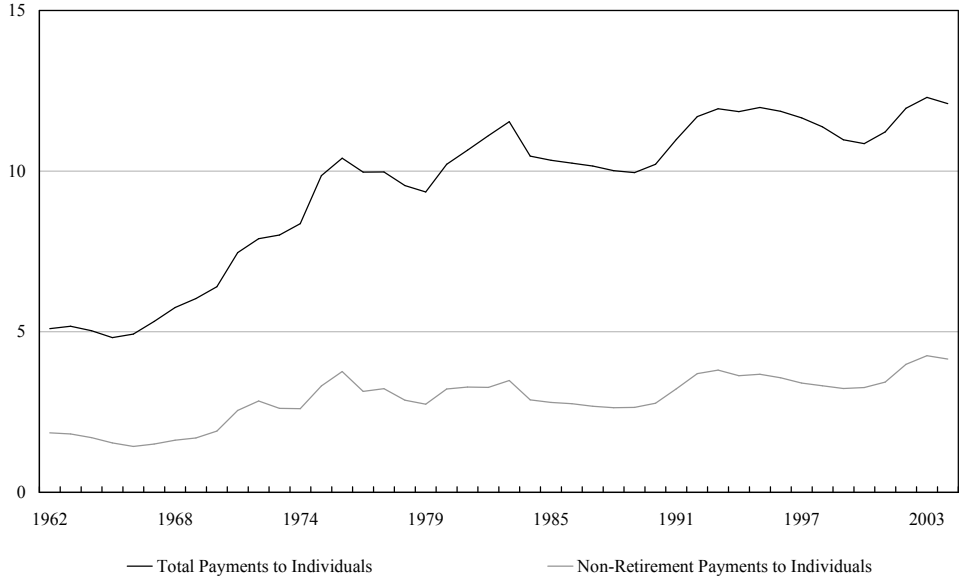
---

[4] Post-retirement life expectancy increased from 67.6 years in 1960 to 77.3 years in 2002 according to the life-tables of the U.S. National Center for Health Statistics.

[5] Non-retirement transfers include all programs except Social Security, Medicare, Federal Retirement (military and civil service), Supplemental Security Income (welfare payments for retirees with inadequate alternative income including Social Security benefits) and Veterans' pensions.

Figure 4

Transfer Payments to Individuals
(percent of GDP)



3. Federal non-retirement transfers: "defensive," "offensive," or "regressive?"

Federal non-retirement transfers have two potential (non mutually exclusive) motivations. One is simply to support the unfortunate and poor through welfare payments and the other is to improve the economy by subsidizing the acquisition of skills and provide services that increase workers' labor force participation. To what extent are government transfers successful in achieving the latter objective? And to what extent do they fail, that is produce undesirable economic outcomes?

This question appears to have drawn relatively little attention in the literature: Most studies focus on the impact of federal investment spending – infrastructure and R&D support – on private sector productivity. However, such federal spending constitutes only 2.5 per cent of GDP in 2004 whereas federal non-retirement payments to individuals (both, direct payments and through grants to State and Local governments) amounted to 4.2 per cent of GDP in the same year.

Diewert (2001) discusses the government's role as a provider of core and non-core services. The latter are defined as services that could be provided by the private sector but are provided by the government instead. Citing Bates (2001), Diewert's list of non-core services includes the provision of higher education, health services and insurance, pensions, income support to the poor, and unemployment insurance. Other services may also qualify – such as child care, nutrition, and housing assistance etc. The provision of non-core services by the government is

Figure 5

Federal Outlays Per Capita By Age, 1988-2001
(males)

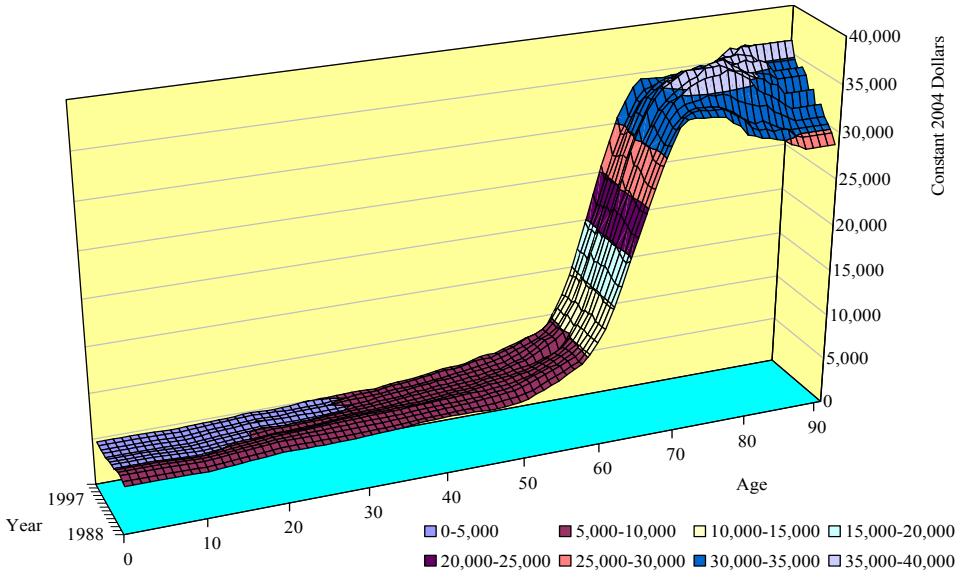


Figure 6

Federal Outlays Per Capita By Age, 1988-2001
(females)

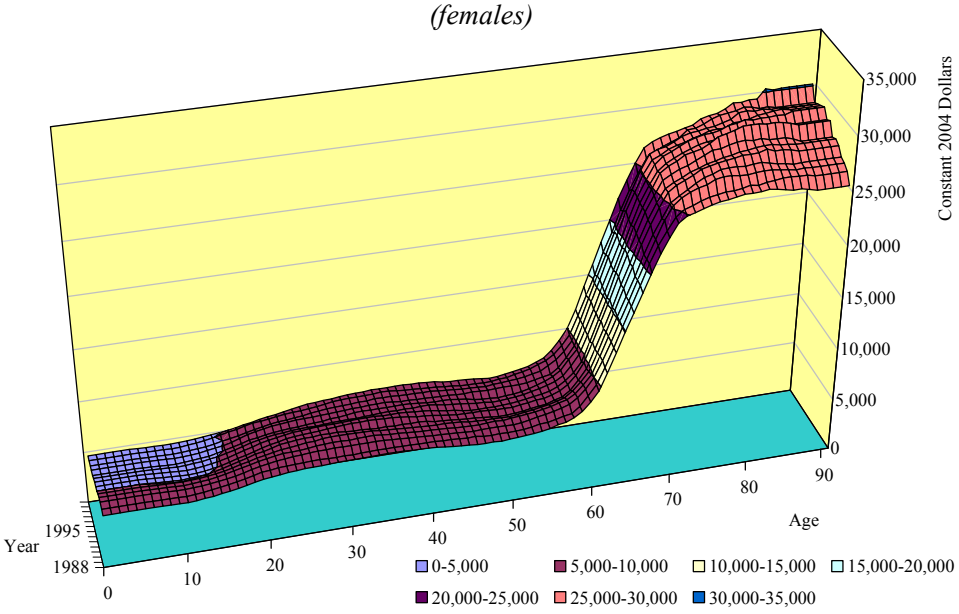


Figure 7

Non-retirement Transfers Per Capita By Age, 1988-2001
(males)

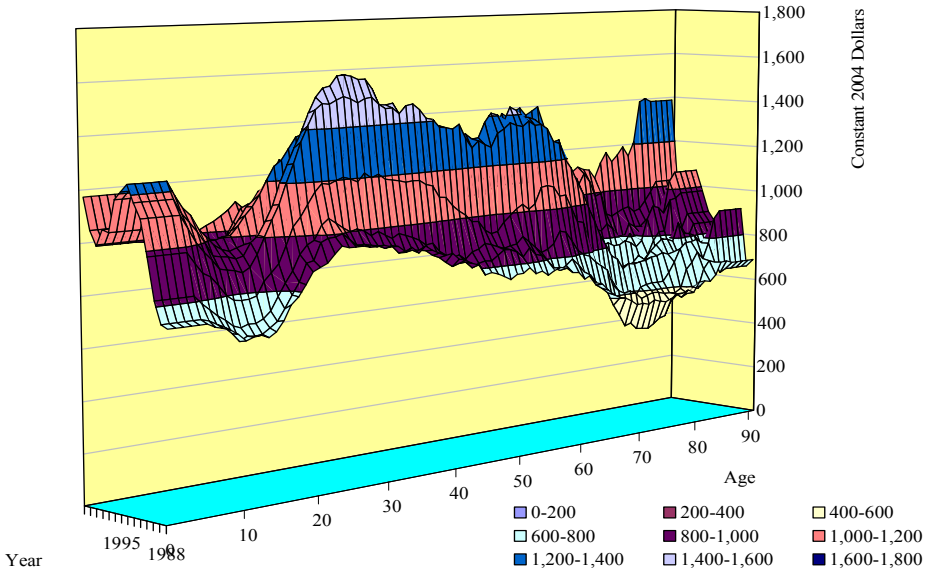
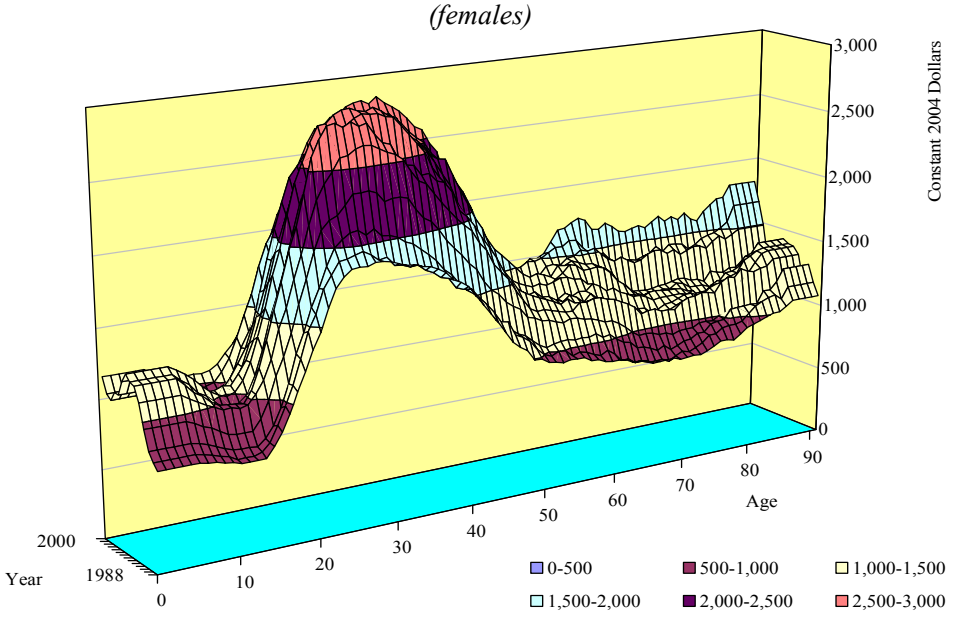


Figure 8

Non-retirement Transfers Per Capita By Age, 1988-2001
(females)



usually justified by citing a significant market failure in their provision. Such justifications usually allude to market failures in the provision of insurance against various kinds of economic risks – mainly reduced living standards from job losses caused by many potential factors – economic recessions, foreign competition, skill obsolescence, inability to work because of domestic constraints such as the lack of a home, child-care costs, lack of transportation, etc.

One could classify the provision of non-retirement transfers by the government under the labels "defensive," "offensive," and "regressive." Defensive transfers are those intended to insure against reduced consumption and living standards from market downturns – to protect the economically unfortunate by providing support during spells of unemployment, inadequate family support, child care for single parents, etc. In contrast, "offensive" non-retirement transfers are those intended to improve the operation of the economy and markets by providing services not adequately provided by the private sector – education and job-training subsidies, employment services, transportation subsidies, youth and adult education programs etc. Transfers of the offensive type should expand labor market opportunities and outcomes for their direct recipients and, if they generate significant positive externalities, for non-recipients as well.

There need not be a sharp dividing line between defensive versus offensive transfers. Indeed, some transfers may operate in both ways: For example, child-care subsidies obviously improve household living standards but may also enable household adults to spend more time at work – increasing their employment and earnings. Unemployment insurance benefits obviously support consumption after job losses but may also enable longer employment searches to improve the employer/employee matches – again improving labor market outcomes.

However, some non-retirement transfers may also exert negative economic effects. Increased dependency on welfare and health-care subsidies may reduce recipients' willingness to work and save. Some may have negative social effects – promoting divorce and out-of-wedlock births. These "regressive" effects may be so severe that they may negate the "defensive" rationale. The perception that this was increasingly the case prompted Congress to pass welfare reform legislation during the mid-Nineties that introduced time limits on eligibility to welfare benefits and required recipients to actively seek employment.

What are the observable effects of non-retirement transfers on future labor market outcomes of recipients? First, more individuals would be protected from the consequences of economic misfortunes – resulting in a lower incidence of poverty. Second, cohorts that receive larger non-retirement transfers could experience less frequent job losses, longer spells of employment, and more rapid earnings growth. Alternatively, recipients may experience the opposite – lower wages, longer spells of unemployment and smaller incomes because they are more dependent on non-retirement transfers.

This paper attempts to analyze whether different types of federal non-retirement transfers exert "offensive," or "regressive" economic effects on

recipients' future economic outcomes – in addition to the "defensive" effect that is assumed to exist by default. To do so, it exploits the cross-cohort variation in such transfers and cross-cohort variation in labor market outcomes – specifically, changes in the fraction of the cohort that does not work, works full-time, and works part-time; changes in average wage of those in the cohort working full- and part-time; and changes in total earnings of those in the cohort working full-time and part-time. The variation examined is in changes in cohort averages over time as computed from different years of CPS data that do not necessarily refer to the same individuals across years.

4. Construction of the cohort dataset from the CPS

The Bureau of Labor Statistic's Current Population Survey is widely used to draw inferences about demographic and economic trends in the United States. The survey's March Supplement files provide information on individual workers' earned and unearned incomes, including income from federal and state transfer programs.[6] Survey participants are asked whether household members received any of several types of public retirement and non-retirement transfers. CPS data on non-retirement transfers include income from veterans' survivors, education, and other benefits, Medicaid benefits, unemployment insurance receipts, family support benefits, TANF, general welfare payments, the earned income tax credit, student assistance, housing assistance, food stamps and other nutrition assistance and child-care assistance. The CPS collected individual level information using a consistent set of questions between 1988 and 2004 – the period selected for analysis.

Household surveys generally contain deficiencies that preclude analysis of the impact of federal transfers on labor market outcomes at the individual or household level. Unlike the Panel Survey of Income Dynamics (PSID) and other panel data surveys, the CPS does not follow the same households each year.[7] Unfortunately, micro-surveys that follow the same households over time do not contain sufficiently detailed information on federal non-retirement income transfers (PSID) or do not cover the entire population (Health and Retirement Survey [HRS] and National Longitudinal Surveys [NLS]) or do not interview each household for a sufficiently long period of time (Survey of Income and Program Participation [SIPP]). Hence, this paper constructs a cohort data set based on the CPS consisting of weighted cohort-specific averages of transfer receipts and cohort-specific demographic characteristics, where cohorts are distinguished by single year of birth and gender.

---

[6] This paper focuses exclusively on federal transfer programs although the authors acknowledge that state transfer programs are equally important and merit a similar analysis. However, implementing such an analysis is beyond the scope and capacity of the datasets available to the authors.

[7] Madrian and Lefgren (1999) have developed computer programs to match a subset of CPS households into a longitudinal data set. However, the matching procedure excludes the years 1994 and 1995 due to revisions in household identifies. This makes the longitudinal length of the individual-level time series of very limited duration.

Very young individuals' labor force experience is relatively much less stable compared to middle-aged individuals. And, those older than 55 face a variety of public and private retirement incentives, which could impart considerable variability in labor market behavior and experience unrelated to non-retirement transfer receipts. Hence, the cohort dataset constructed here includes only cohorts within the age range of 25 through 55.

The CPS asks households a consistent battery of questions on earnings and receipts from federal transfers between 1988 and 2004. The questions on transfer of various types do not distinguish between receipts from federal programs and those from state and local programs. Hence, these data reflect all transfers by function – whether for child care, unemployment benefits, education subsidies, etc. – and from all sources.

To accommodate 5 lagged terms of first differences in the regressions implemented, cohort-specific time series data spanning 10 years (1995 through 2004) are constructed. Hence, the oldest cohort in the dataset consists of those born in 1949 (aged 55 in 2004) and the youngest cohort is of those born in 1970 (aged 25 in 1995).

5. Estimation method

In attempting to estimate the "offensive" and "regressive" effects of government non-retirement transfers, it is necessary to address a simultaneity problem in determining the size of the transfers directed toward any cohort. Specifically, the transfers are determined by cohorts' economic conditions and, in turn, affect that condition. Hence, this paper devises a two-step estimation method by postulating, first, that government non-retirement transfers in period t produce a contemporaneous "defensive" impact but exert "offensive" or "regressive" effects only in period $t+1$ and later. A second assumption is that the existing policy on the size of transfers and economic activity are in equilibrium and only transfers that represent a deviation from that policy (the unexpected component of transfers) would exert positive or negative economic effects.

Under these assumptions, a two-step regression strategy can be employed to isolate the "offensive" or "regressive" impact of transfers on economic outcomes – as described below. Finally, regressions are implemented on annual *changes* in cohort-specific transfers and labor market outcomes to deal with the fact that both types of variables are non-stationary.

The cohort-specific annual change in the average of the specific federal non-retirement transfer to be examined, γ, in period t is first decomposed into two parts: the first component refers to the component of transfer growth undertaken in response to each cohort's "needs" – that is, it represents the "defensive" motivation. This component is estimated using the following regression specification:

$$\gamma i,t = \gamma(\phi i;\ Xi,t \ldots Xi,t-k;\ \omega t \ldots \omega t-k;\ \Gamma i,t \ldots \Gamma i,t-k;\ \gamma i,t-1,\ \gamma i,t-2) + ui,t \qquad (1)$$

Here, i indexes cohort birth years, t indexes calendar years, and k indexes annual lags. In equation (1), $\gamma_{i,t}$ is regressed on cohort fixed effects, ϕ_I; cohort demographic characteristics, $X_{i,t}$; cohort labor market characteristics, $\omega_{i,t}$; contemporaneous changes in other transfers to the cohort, $\Gamma_{i,t}$ (excluding γ); and a second-order autocorrelation component to take account of inertia in the change in federal non-retirement transfers directed at particular cohorts.

The resulting "explained" component of cohort-specific transfer growth, $\hat{\gamma}_{i,t}$, is interpreted as the "defensive" component. The residuals estimated from this regression, $\hat{u}_{i,t}$, are interpreted as the "policy change" or "unexpected" transfer component – the component unrelated to cohort-specific welfare needs or prior expectations regarding the change in $\gamma_{i,t}$. The residual $\hat{u}_{i,t}$, could potentially exert "offensive" or "regressive" cohort economic effects.

The matrix ω includes cohort-specific changes in labor market characteristics: the wage rate for full-time workers; wage rate for part-time workers; the fraction of the cohort not working; the fraction working full-time; the fraction working part-time; cohort earnings for those working full-time; earnings for those working part-time. The matrix X represents cohort demographic characteristics (in levels) such as birth year dummy variables (to capture cohort fixed effects); the cohort's age, age-squared, and gender; the fraction of cohort members who are married; the fraction non-white; the average number of children per family; average educational attainment of family head and spouse; the fraction of families with income below the official poverty limit; the cohort's average effective marginal income tax rate; and variables representing occupational composition according to wage growth across different occupations for those working full-time and part-time (described below).

When implementing the "first stage" regressions specified in equation (1), the changes in labor market variables, ω, are interacted with cohort age and gender variables. Note that current (year t) values of Γ, ω, and X are included as explanatory variables in the regression to estimate the "defensive" component of transfer growth, $\gamma_{i,t}$.

Occupational composition variables are useful as explanatory variables in determining the size of "defensive" government transfers in the "first stage" regression described above. They are also needed to distinguish between the cohort's wage growth and labor market participation variables arising from shifts in occupational composition and from "offensive" government transfers in the second stage regressions described below.

The occupational wage-growth variables for full-time and part-time workers are constructed in 3 steps: first, growth in within occupation average wages is calculated from the CPS for all occupations in each year compared to the previous year. Year-specific average wages and salaries are calculated for all occupations distinguished by 3-digit codes across all workers (regardless of cohort affiliation or age) in each year. Next, occupations are ranked according to the growth in average

wages and each CPS individual who participates in the labor market is assigned a number ρ ranging between 1 and 10 depending on the growth rate decile of the occupation he or she works in. Finally, the cohort average for ρ, $\overline{\rho}$, is taken to represent each CPS cohort's occupational composition. Thus, an increase in a cohort's $\overline{\rho}$ value from one year to the next would reflect a shift of its members from low-growth to high-growth occupations. Cohort averages, $\overline{\rho}$, are used to control for this source of cohort-specific changes in average wage rates when attempting to estimate the effects of past government transfers on workers' productivity and earnings.

The residuals $\hat{u}_{i,t}$ from estimating equation (1) are taken to represent the "offensive" policy components of government transfers. Cohorts with positive values of $\hat{u}_{i,t}$ received more transfers of type $\gamma_{i,t}$ than would be explained by their "defensive" need or expectations based on current economic conditions and past transfer growth. If the unanticipated change in transfers results in better future economic outcomes (higher wage growth, lower fraction non-working, higher fraction working full-time etc.), the transfer in question fulfills an "offensive" function. If government transfers, indeed, improve workers' labor market performance in future periods, lagged values $\hat{u}_{i,t-k}$ should enter significantly and with coefficients of the appropriate sign in "second stage" regressions of the following type:

$$\omega_{i,t} = \omega(\phi_i\ ;\ \gamma_{i,t}\ ;\ \Gamma_{i,t} \ldots \Gamma_{i,t-k}\ ;\ \hat{u}_{i,t-1} \ldots \hat{u}_{i,t-k}\ ;\ X_{i,t} \ldots X_{i,t-k}) + e_{i,t}, \qquad (2)$$

Up to 4 lagged innovations in transfers $\hat{u}_{i,t-k}$ are used in equation (2) because the "offensive" effects of extra transfers could arise after more than just 1 year. Demographic variables, $X_{i,t}$, cohort fixed effects, ϕ_i, and other contemporaneous transfers, $\gamma_{i,t}$ and $\Gamma_{i,t}$, are included as explanatory variables in the second stage regressions. Finally, current GDP growth, the current unemployment rate, the current inflation rate, and a 0-1 dummy for the recession year 2001, are also included as regressors in equation (2) to control for current macroeconomic conditions.[8]

6. Cohort CPS data

Figures 9 and 10 depict cohort-specific profiles of average real earnings by age derived from the CPS between 1993 and 2004.[9] Average earnings trajectories rise rapidly for younger cohorts compared to older cohorts and earnings growth is slower for females than for males. The cohort profiles exhibit much more volatility

---

[8] The macroeconomic variables are taken from the *2005 Economic Report of the President*.

[9] Annual earnings reported are deflated by the Bureau of Labor Statistics' Consumer Price Index for All Urban Consumers.

Figure 9

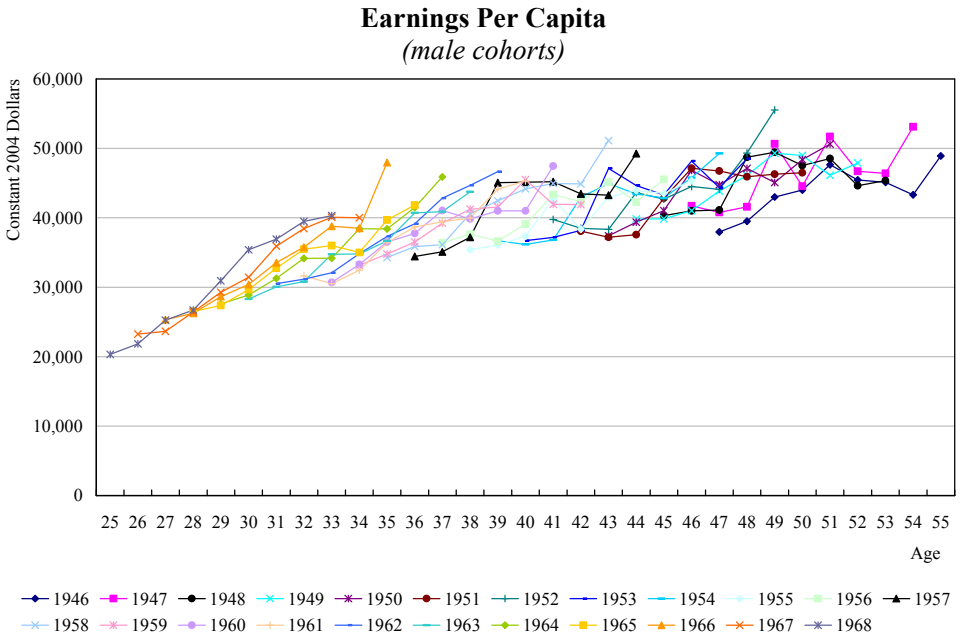
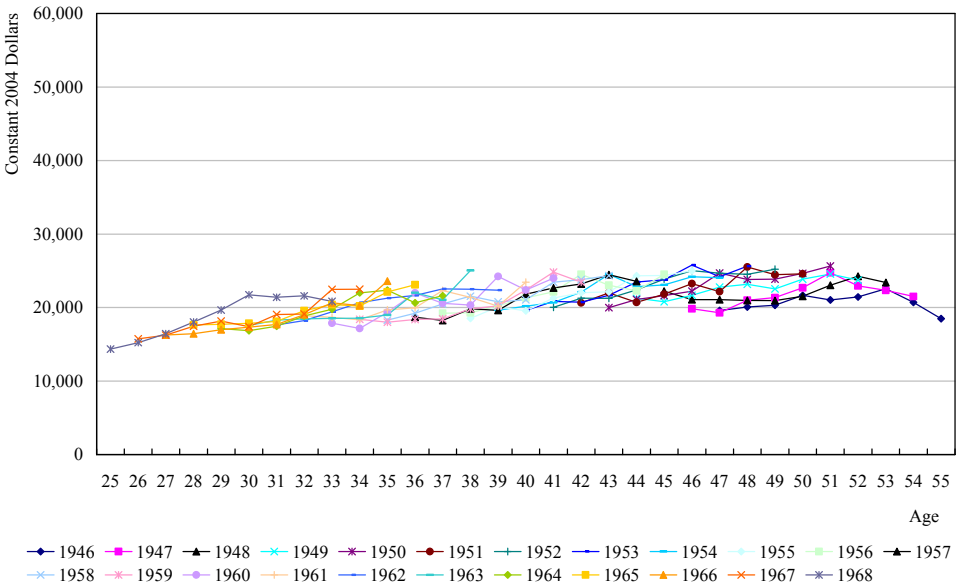


Figure 10



because unlike the data shown in Figures 7 and 8, these profiles are not smoothed (as described in the Appendix). Figures 11 and 12 contain average transfer receipts for the sample of cohort earnings. Transfer levels for females are approximately twice as large for younger females as for younger males. They decline with age for females and almost achieve parity with those awarded to males older than 50.

Figures 13 and 14 depict the fraction of males and females that are not working by cohort. For males, the fraction not working begins to creep up after age 40 and the cohort profiles exhibit considerable volatility, especially at older ages. A much larger percentage of females stay out of the labor force compared to males. However, the fraction of non-working females trends downward until their late Forties and begins to rise for women aged 50 and older.

Figures 15 and 16 report the percentages of males and females in each cohort that work part-time. Younger male cohorts experience steep declines in part-time labor force participation – presumably because many of them migrate to full-time jobs. About one quarter of middle-aged males appear to be part-timers. The percentage of female part-timers is twice that of males and within-cohort cross-year volatility in female part-time labor-force participation appears to be much smaller than that for males.

7. Findings

7.1 Demographic and labor market statistics

Table 1 shows averages across years and ages for all variables – dependent and explanatory – used in the regressions implemented in Tables 2 through 5. Averages of demographic variables appear to accord with well known facts.

Among labor market variables, real earnings growth along cohort trajectories has been very low. Males working full-time experienced zero growth in real wages and growth for females and males working part-time was just 1 per cent per year. The percent of the cohort not working was stable for males and declined by 1 per cent per year for females. For cohorts studies, the percent of those working full-time declined whereas the percent of those working part-time increased between 1995 and 2004.

Among transfers, growth in education and other government transfers per capita has been negative since 1995. In contrast, growth in child-care transfers was quite substantial, especially for females. That is not surprising as new programs for pregnant women and children were initiated during the late Nineties – notably as part of the Medicaid program. Growth in inflation adjusted unemployment transfers is modest for both males and females.

Figure 11

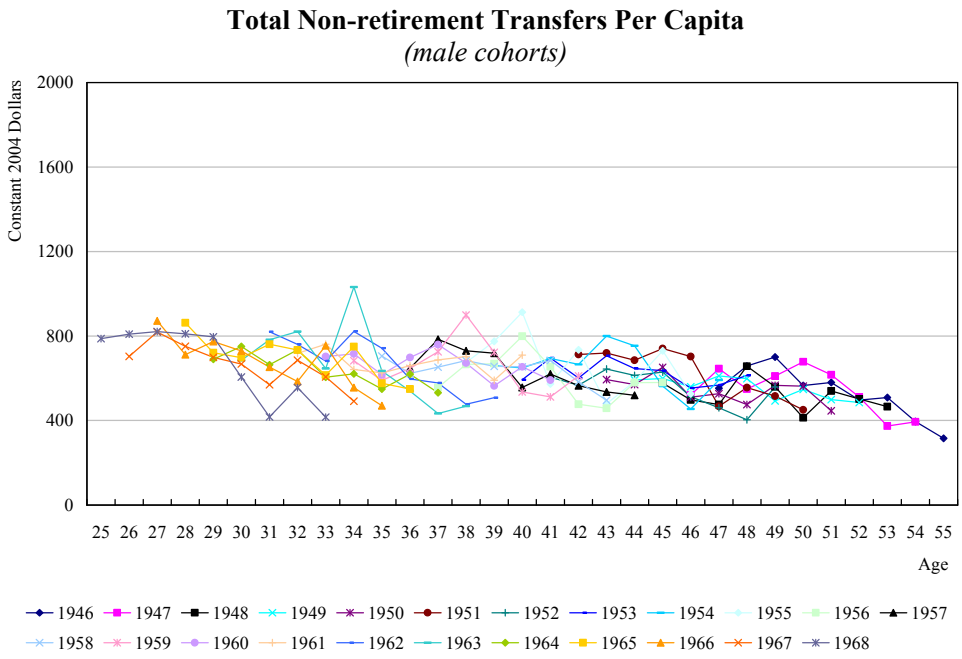


Figure 12

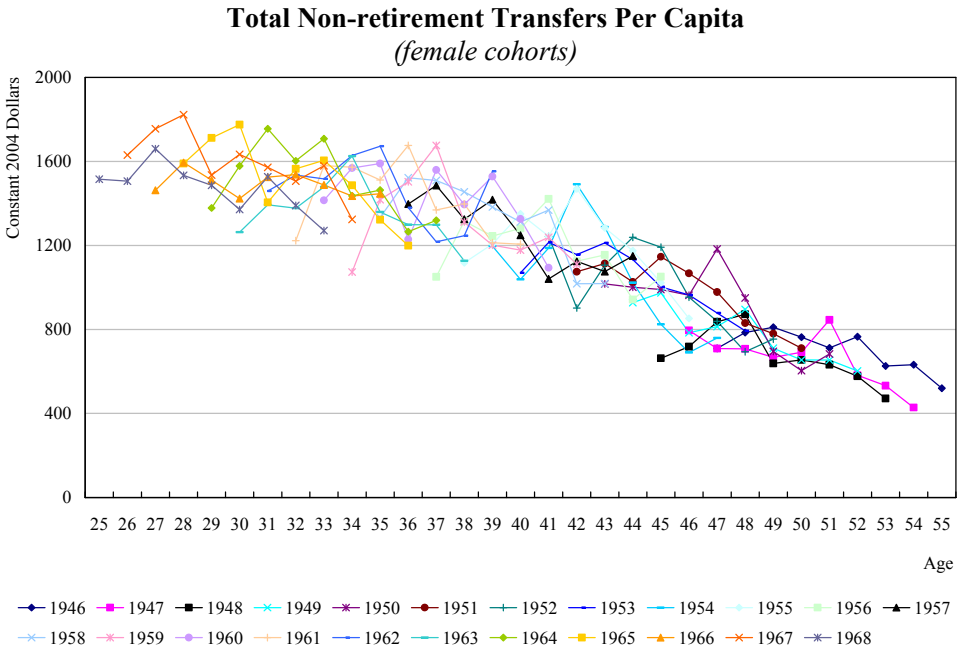


Figure 13

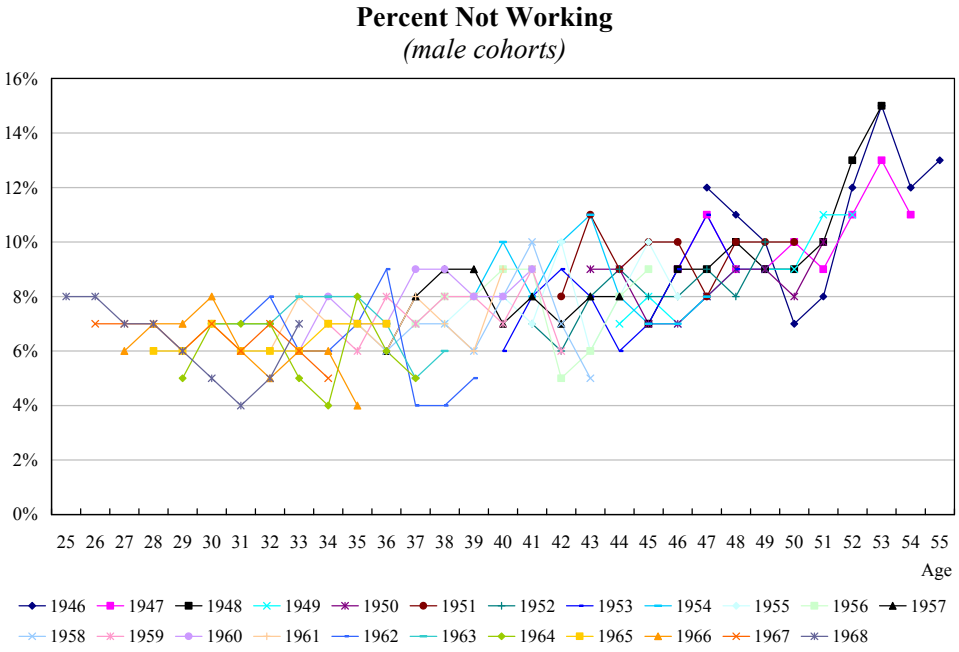


Figure 14

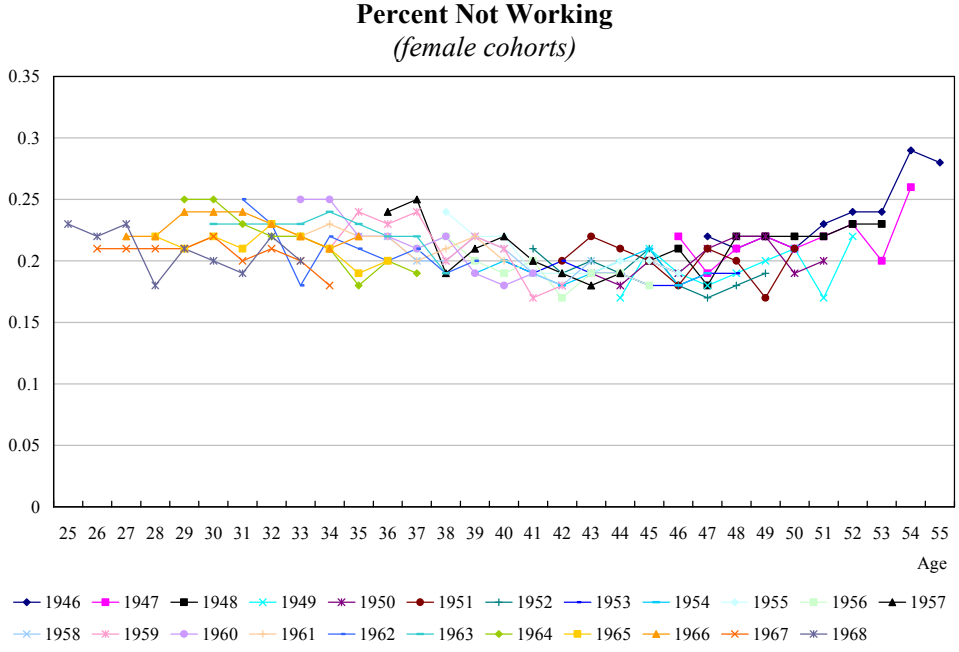


Figure 15

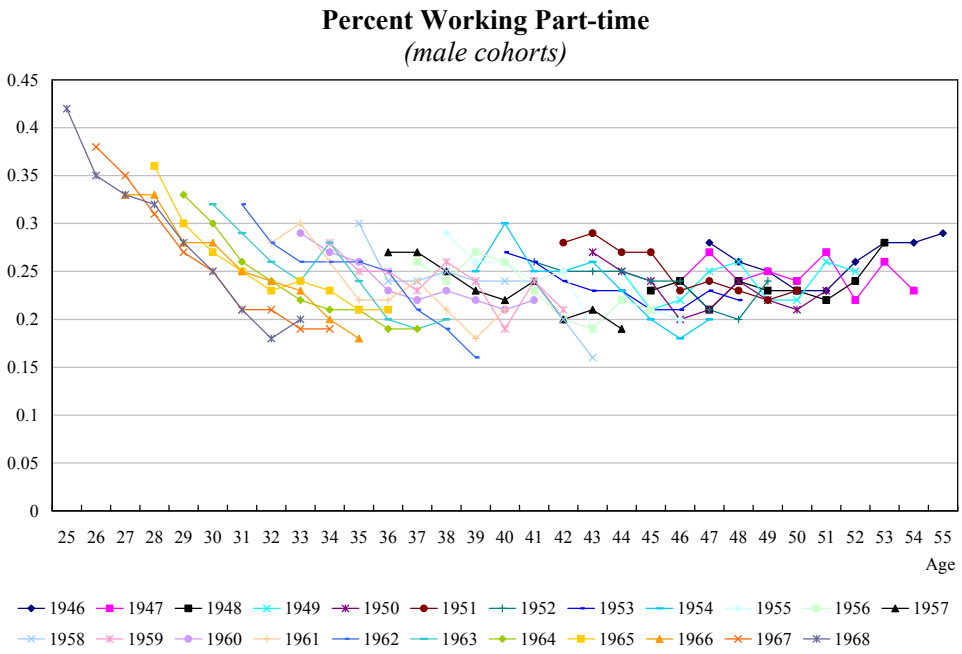


Figure 16

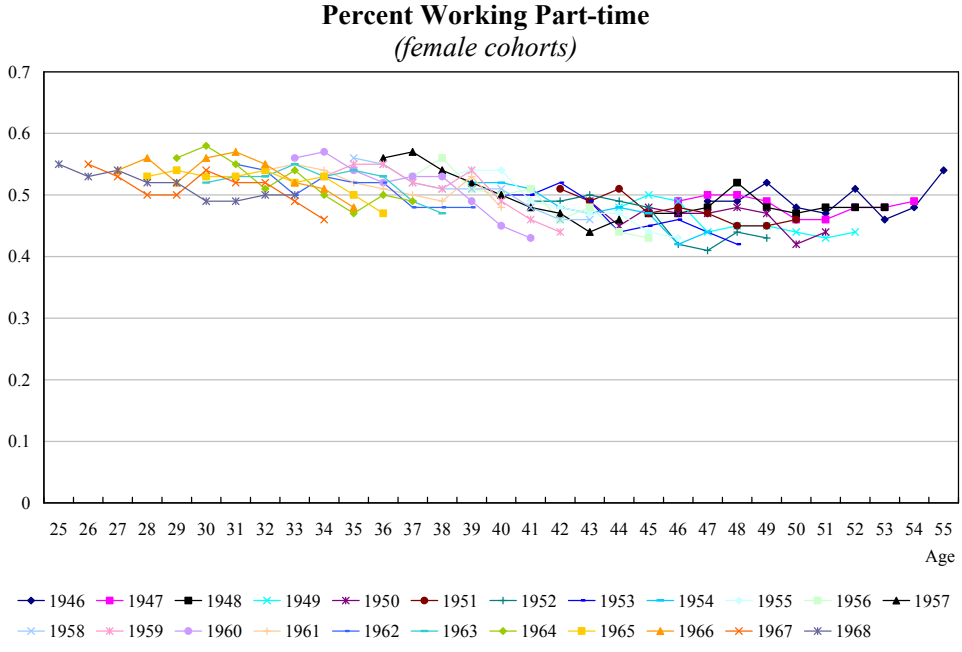


Table 1

Demographic Characteristics

Variable	Males		Females	
	Average Across All Ages and Years	Gross Growth Rate	Average Across All Ages and Years	Gross Growth Rate
Married *(percent)*	65.7	-	66.4	-
White *(percent)*	83.4	-	81.2	-
With Child(ren) *(percent)*	49.0	-	57.5	-
Completed High-School or More *(percent)*	87.5	-	89.0	-
With Earnings <= Poverty Limit *(percent)*	10.8	-	14.3	-
Average Marginal Tax Rate	13.0	-	6.8	-
Wages/week (full-time workers)*	-	1.00	-	1.01
Wages/week (part-time workers)*	-	1.01	-	1.01
Not Working** *(percent)*	-	1.00	-	0.99
Working Full Time** *(percent)*	-	0.99	-	0.99
Working Part Time** *(percent)*	-	1.02	-	1.02
Unemployment Insurance*	-	1.01	-	1.01
Child Care*	-	1.00	-	1.17
Education Benefits*	-	0.99	-	0.98
Other Government Transfers*	-	0.99	-	1.00

* In constant 2004 dollars.
** Growth rate refers to within-cohort growth in variable with advancing age.
Source: Authors' calculations.

7.2 Regression results

Tables 2 through 5 show the results from the second stage regression specified in equation (2).[10] The tables examine the impact of various types of federal non-retirement transfers on five variables of interest: cohorts' average weekly wage rate calculated only for its full-time workers (dwg_ft), the weekly wage rate for part-time workers (dwg_pt), the fraction of the non-working members in the cohort (dnw), the fraction of members working full-time (dft), and the fraction working part-time (dpt). Cohort fixed-effect coefficients are omitted from the tables.

Table 2 shows the impact of unemployment insurance transfers on cohort labor market variables. That is, the $\hat{u}_{\_k}$ (where k indicates the number of annual lags) refer to residuals from regressing cohort unemployment insurance transfers on the explanatory variables specified in equation (1). The Table shows that macroeconomic control variables (gdp growth, the unemployment rate and the

[10] Results from the first-stage regressions are available from the author upon request.

Table 2

Impact of Unemployment Insurance Transfers on Future Cohort Labor-force Outcomes

	dwg_ft	**dwg_pt**	**dnw**	**dft**	**dpt**
Intercept	0.820 ***	3.313 **	0.865	–2.977	4.650 **
Macro GDP growth	0.002	0.007	–0.009	–0.028 **	0.028 ***
Macro unemployment rate	0.005	–0.121	0.029	0.311 **	–0.281 **
Macro inflation rate	0.000	0.016 *	–0.001	–0.059 ***	0.049 ***
Change in other government transfers	–0.039	–0.707	–0.649	–1.516	1.596
Change in other government transfers x sex	–0.002	0.108	–0.081	0.480	–0.338
Change in other government transfers x age	0.001	0.010	–0.001	0.052	–0.038
Change in unemployment benefits	–0.055	0.138	–0.154	0.367	–0.180
Change in unemployment benefits x sex	–0.004	0.003	0.202 *	0.247 *	–0.307 ***
Change in unemployment benefits x age	0.001	–0.002	0.001	–0.008	0.006
Change in education benefits	0.009	–0.134	0.240	–0.250	0.008
Change in education benefits x sex	–0.005	0.008	–0.070 **	0.053	0.011
Change in education benefits x age	0.000	0.003	–0.004	0.005	0.000
Change in child–care benefits	0.013	0.048	0.037	0.035	–0.081
Change in child–care benefits x sex	–0.017	–0.034	0.000	–0.019	0.044
Change in child–care benefits x age	0.000	0.000	–0.001	0.000	0.001
Recession year dummy (2001)	0.005	0.187 *	–0.045	–0.548 **	0.482 ***
Age	0.003	–0.060 *	–0.007	0.109	–0.092 *
Age–Squared	0.000	0.000	0.000	0.000	0.000
Sex	0.022	–0.085	–0.079	–0.682	0.535
Fraction married	0.035	–0.013	0.049	0.013	–0.121
Fraction white	0.006	0.156	0.213	0.139	–0.161
Fraction with children	–0.037 **	–0.012	0.086	–0.083	0.023
Fraction with HS education or more	0.075	0.067	0.870 **	–0.483	–0.348
Fraction Poor	0.069	–0.052	–0.948 **	1.458 ***	–0.431
Average Marginal Income Tax Rate	0.001	0.000	0.001	–0.007	0.006
Occupational wage–growth: Full–time workers	0.005 ***	–0.001	–0.009	0.005	0.002
Occupational wage–growth: Part–time workers	0.002	0.007	0.004	0.004	–0.006
$\hat{u}$_1	**0.004**	**0.011**	**–0.049**	**–0.098**	**0.107**
$\hat{u}$_2	**0.000**	**0.031**	**–0.053**	**0.027**	**0.040**
$\hat{u}$_3	**0.000**	**0.048**	**–0.016**	**0.060**	**–0.025**
$\hat{u}$_4	**0.000**	**0.002**	**–0.080**	**0.048**	**0.048**
Number of observations	266	266	266	266	266
R–Squared	0.307	0.141	0.258	0.324	0.197
Root Mean Square Error	0.008 ***	0.036 **	0.064 *	0.073 *	0.058 *

* = Significant at the 10 percent confidence level; ** = Significant at the 5 percent confidence level; *** = Significant at the 1 percent confidence level.

Table 3

Impact of Education and Training Transfers on Future Cohort Labor-force Outcomes

	dwg_ft	dwg_pt	dnw	dft	dpt
Intercept	0.767 ***	3.235 **	1.438	–3.659	4.524 **
Macro GDP growth	0.002	0.006	–0.007	–0.025 *	0.024 **
Macro unemployment rate	0.009	–0.124 *	–0.008	0.381 ***	–0.293 **
Macro inflation rate	0.000	0.019 **	0.002	–0.064 ***	0.050 ***
Change in other government transfers	–0.042	–0.609	–0.723	–1.891	1.961 *
Change in other government transfers x sex	–0.001	0.084	–0.087	0.629	–0.442
Change in other government transfers x age	0.002	0.008	0.001	0.060 *	–0.047 *
Change in unemployment benefits	–0.057	0.166	–0.136	0.328	–0.159
Change in unemployment benefits x sex	–0.005	0.015	0.211 *	0.243 *	–0.310 ***
Change in unemployment benefits x age	0.001	–0.003	0.001	–0.006	0.004
Change in education benefits	0.008	–0.129	0.270 *	–0.271	–0.004
Change in education benefits x sex	–0.004	0.010	–0.075 ***	0.067 **	0.006
Change in education benefits x age	0.000	0.003	–0.005	0.005	0.000
Change in child–care benefits	0.013	0.052	0.022	0.036	–0.063
Change in child–care benefits x sex	–0.017	–0.033	0.010	–0.021	0.033
Change in child–care benefits x age	0.000	0.000	–0.001	0.000	0.001
Recession year dummy (2001)	–0.001	0.198 *	0.000	–0.625 ***	0.490 ***
Age	0.004	–0.058 *	–0.023	0.126 *	–0.088
Age–Squared	0.000	0.000	0.000	0.000	0.000
Sex	0.023	–0.073	–0.095	–0.834 **	0.664 *
Fraction married	0.043	–0.022	0.047	0.070	–0.174
Fraction white	0.009	0.161	0.197	0.122	–0.143
Fraction with children	–0.040 **	–0.007	0.101	–0.133	0.053
Fraction with HS education or more	0.071	0.084	0.853 **	–0.349	–0.421
Fraction Poor	0.067	–0.058	–0.931 **	1.403 ***	–0.407
Average Marginal Income Tax Rate	0.001	0.000	0.003	–0.009	0.006
Occupational wage–growth: Full–time workers	0.005 ***	–0.001	–0.009	0.009	–0.001
Occupational wage–growth: Part–time workers	0.002	0.005	0.004	0.004	–0.006
$\hat{u}$_1	**–0.002** *	**–0.001**	**0.008**	**–0.011**	**0.003**
$\hat{u}$_2	**0.000**	**0.005**	**0.003**	**–0.019**	**0.012**
$\hat{u}$_3	**0.001**	**–0.002**	**0.013**	**–0.008**	**–0.004**
$\hat{u}$_4	**0.001**	**–0.002**	**–0.007**	**0.021** **	**–0.010**
Number of observations	266	266	266	266	266
R–Squared	0.323	0.138	0.260	0.340	0.198
Root Mean Square Error	0.008 ***	0.036 **	0.063 *	0.072 *	0.058 *

* = Significant at the 10 percent confidence level; ** = Significant at the 5 percent confidence level; *** = Significant at the 1 percent confidence level.

Table 4

Impact of Child-care Transfers on Future Cohort Labor-force Outcomes

	dwg_ft	**dwg_pt**	**dnw**	**dft**	**dpt**
Intercept	0.848 ***	3.657 ***	2.869	–3.366	3.165
Macro GDP growth	0.002	0.004	–0.002	–0.029 **	0.024 **
Macro unemployment rate	0.007	–0.135 *	–0.112	0.315 **	–0.171
Macro inflation rate	0.000	0.019 *	0.019	–0.060 ***	0.035 **
Change in other government transfers	–0.077	–0.786	–0.502	–1.443	1.658
Change in other government transfers x sex	–0.007	0.048	–0.205	0.477	–0.230
Change in other government transfers x age	0.002	0.012	–0.005	0.052	–0.041
Change in unemployment benefits	–0.051	0.249	–0.132	0.385	–0.210
Change in unemployment benefits x sex	–0.004	0.037	0.229 *	0.221	–0.308 ***
Change in unemployment benefits x age	0.001	–0.005	0.001	–0.008	0.005
Change in education benefits	0.008	–0.177 **	0.280 *	–0.314 *	0.024
Change in education benefits x sex	–0.005	0.015	–0.071 ***	0.060 *	0.009
Change in education benefits x age	0.000	0.004 **	–0.005	0.007 *	–0.001
Change in child–care benefits	0.005	0.032	0.113	0.065	–0.137
Change in child–care benefits x sex	–0.015	–0.022	–0.002	–0.048	0.059
Change in child–care benefits x age	0.000	0.000	–0.002 *	–0.001	0.002
Recession year dummy (2001)	0.004	0.201 *	0.156	–0.555 **	0.330 *
Age	0.002	–0.072 **	–0.054	0.107	–0.048
Age–Squared	0.000	0.000	0.000	0.000	0.000
Sex	0.024	–0.082	0.004	–0.617	0.423
Fraction married	0.027	–0.042	0.153	–0.048	–0.160
Fraction white	0.014	0.195	–0.003	0.277	–0.152
Fraction with children	–0.037 *	0.000	0.102	–0.102	0.038
Fraction with HS education or more	0.074	0.079	0.704 *	–0.120	–0.524
Fraction Poor	0.084	–0.128	–0.946 **	1.237 ***	–0.255
Average Marginal Income Tax Rate	0.001	0.000	0.004	–0.010	0.006
Occupational wage–growth: Full–time workers	0.004 ***	–0.001	–0.008	0.008	0.000
Occupational wage–growth: Part–time workers	0.002	0.007	–0.001	0.005	–0.004
$\hat{u}$ **_1**	**0.000**	**–0.002**	**–0.001**	**–0.009**	**0.008**
$\hat{u}$ **_2**	**0.001**	**0.001**	**0.002**	**0.004**	**–0.004**
$\hat{u}$ **_3**	**0.000**	**0.000**	**0.001**	**0.012** **	**–0.009** **
$\hat{u}$ **_4**	**0.000**	**0.002** *	**0.004** **	**–0.003**	**–0.001**
Number of observations	255	255	255	255	255
R–Squared	0.301	0.163	0.300	0.337	0.216
Root Mean Square Error	0.008 ***	0.036 **	0.061 *	0.073 *	0.058 *

* = Significant at the 10 percent confidence level; ** = Significant at the 5 percent confidence level; *** = Significant at the 1 percent confidence level.

Table 5

Impact of Other Non-retirement Transfers on Future Cohort Labor-force Outcomes

	dwg_ft	dwg_pt	dnw	dft	dpt
Intercept	0.854 ***	3.271 **	1.352	–3.320	4.331 **
Macro GDP growth	0.002	0.007	–0.010	–0.022	0.024 **
Macro unemployment rate	0.001	–0.109	0.042	0.321 **	–0.282 **
Macro inflation rate	0.000	0.016 *	–0.003	–0.058 ***	0.048 ***
Change in other government transfers	–0.023	–0.725	–1.416	–1.238	1.967 *
Change in other government transfers x sex	–0.008	0.172	–0.058	0.610	–0.441
Change in other government transfers x age	0.001	0.008	0.014	0.045	–0.045 *
Change in unemployment benefits	–0.053	0.118	–0.202	0.407	–0.174
Change in unemployment benefits x sex	–0.003	0.002	0.168	0.281 **	–0.309 ***
Change in unemployment benefits x age	0.001	–0.002	0.003	–0.009	0.005
Change in education benefits	0.013	–0.159 *	0.255 *	–0.266	–0.001
Change in education benefits x sex	–0.006	0.014	–0.067 **	0.063 **	0.004
Change in education benefits x age	0.000	0.003 *	–0.005	0.005	0.000
Change in child–care benefits	0.011	0.066	0.071	0.028	–0.089
Change in child–care benefits x sex	–0.016	–0.041	–0.031	–0.004	0.048
Change in child–care benefits x age	0.000	–0.001	–0.001	–0.001	0.001
Recession year dummy (2001)	0.009	0.186 *	–0.073	–0.527 **	0.467 ***
Age	0.001	–0.052	–0.010	0.115 *	–0.089
Age–Squared	0.000	0.000	0.000	0.000	0.000
Sex	0.027	–0.144	–0.054	–0.868 **	0.657 *
Fraction married	0.039	–0.024	0.114	–0.020	–0.148
Fraction white	0.016	0.100	0.139	0.125	–0.110
Fraction with children	–0.034 *	–0.020	0.039	–0.069	0.052
Fraction with HS education or more	0.076	0.029	0.795 **	–0.418	–0.333
Fraction Poor	0.065	–0.072	–0.860 **	1.369 ***	–0.438
Average Marginal Income Tax Rate	0.001	0.000	0.003	–0.008	0.004
Occupational wage–growth: Full–time workers	0.005 ***	–0.001	–0.005	0.005	–0.001
Occupational wage–growth: Part–time workers	0.002	0.007	0.003	0.005	–0.006
$\hat{u}$_1	**0.025**	**–0.212**	**–0.524** **	**0.224**	**0.204**
$\hat{u}$_2	**0.017**	**–0.238** *	**0.122**	**–0.418**	**0.218**
$\hat{u}$_3	**0.004**	**0.097**	**0.176**	**0.019**	**–0.094**
$\hat{u}$_4	**0.036**	**–0.012**	**–0.133**	**0.173**	**–0.029**
Number of observations	266	266	266	266	266
R–Squared	0.313	0.160	0.276	0.331	0.195
Root Mean Square Error	0.008 ***	0.035 **	0.063 *	0.073 *	0.058 *

* = Significant at the 10 percent confidence level; ** = Significant at the 5 percent confidence level; *** = Significant at the 1 percent confidence level.

inflation rate and the recession-year dummy variable) enter significantly in explaining cohort labor market characteristics – especially the fractions working full-time and part-time. However, none of the $\hat{u}_{\_k}$ variables are significant. That implies that providing larger unemployment insurance transfers (that are unanticipated) do not result in higher future wages or employment. Thus, such transfers perform a purely "defensive" role.

Table 3 shows the impact of education and training transfers on cohort labor market characteristics. It shows $\hat{u}_{\_1}$ to be a small negative number significant at the 10 per cent confidence level in the regression for wage rate for full-time workers (dwg-ft). That is, providing additional education transfers results in a very small *decline* in the weekly wage after 1 year. In addition, $\hat{u}_{\_4}$ is positive and significant at the 5 per cent level in the regression for the fraction working full-time (dft). Thus, unanticipated increases in education transfers appear to have a small positive impact in the fraction of full-time workers after a lag of 4 years. However, most of the $\hat{u}_{\_k}$ coefficients are not significantly different from zero, suggesting, contrary to popular belief, that marginally higher education transfers would not significantly improve future labor market outcomes.

Table 4 examines the impact of child-care transfers on cohort labor market behavior. Extra child-care benefits appear to have no statistically and economically significant impact on wage rates. They exert a statistically, but not economically, significant positive impact on the fraction of non-working cohort members after a lag of 4 years. Child-care transfers appear to have small and offsetting effects on the fractions of cohort members working full-time and part-time after a long time lag of 3 years. Additional and unanticipated child-care transfers induce cohort members to shift from part-time to full-time work, suggesting the presence of a small "offensive" long-term employment effect.

Table 5 shows that other government non-retirement transfers (the sum of items such as veterans benefits, housing assistance, food stamps, earned income credit, health care assistance, etc.) reduces the fraction of cohort members that remain out of the labor force. The employment impact of government transfers is not surprising. Unexpectedly high government transfers are known to stimulate economic activity. However, as the results show, the impact is short-lived. Curiously, larger than expected government transfers also reduce the weekly wages of part-time workers: the coefficient on $\hat{u}_{\_2}$ is negative and significant at the 10 per cent level of confidence.

In summary, government transfers on unemployment insurance are predominantly "defensive." Those on education exert a marginally "offensive" role but the effects are small and offsetting. Other government transfers exert a significant "offensive" effect on employment, but a "regressive" effect on wage rates of part-time workers.

8. Conclusion

That retirement transfers have grown considerably is well known. Less appreciated is that non-retirement transfers have also increased as a share of GDP. Such transfers could be justified on "defensive" or "offensive" grounds. The former implies that transfers are made to protect and support the needy and those experiencing bad economic outcomes. The latter implies an effort to improve the functioning of the economy and markets. It is not possible to theoretically classify non-retirement transfers as being defensive or offensive in their effects. Indeed, some transfers may exert both types of effects and others may be predominantly "regressive".

This paper attempts to empirically estimate the impact of government non-retirement transfers at the cohort level. It develops a cohort dataset based on the Current Population Survey to study whether non-retirement transfers primarily fulfill a defensive or offensive role in the economy. This analysis bears on the current Social Security debate: if such transfers could be shown to fulfill a significant offensive role, it could be argued that the Social Security Trust Fund operates as an effective storage technology for saving and investing payroll tax surpluses despite the fact that such surpluses are invested in non-marketable Treasury securities and spent on federal non-Social Security outlays. Trends in federal spending during the last two decades suggest that much of the surpluses are being spent on transfers.

Estimates of the impact of federal non-retirement transfers on labor market outcomes at the cohort level suggest that they play a predominantly "defensive" role – that is finance recipients' consumption and maintain current labor market outcomes rather than improve upon them. However, this conclusion should be viewed with caution because, rather than individual or household level data, they are based on cohort averages of economic and demographic variables – data that are considerably noisy. It could also be the case that the results arise from the joint determination of transfers and labor market outcomes due to changes in other unobserved variables. Nonetheless, the results provide little support to the idea that non-retirement transfers improve the functioning of labor markets and the economy: observed "offensive" effects from some transfers on some labor market outcomes are small and appear to be neutralized by "regressive" effects on other economic outcomes.

APPENDIX
THE METHOD FOR DISTRIBUTING FEDERAL OUTLAYS BY AGE AND GENDER

The distribution of total federal outlays and its components is estimated by using data on aggregate transfers from the U.S. Budget (published by Office of Management and Budget), Area Population tables by age and gender provided by the Social Security Administration, and profiles of transfer receipts by age and gender developed from the Current Population Survey's March Supplement files for the years 1988 through 2001. Federal transfers to individuals are divided into 16 categories, of which 14 are distributed by age and gender using CPS profiles and 2, comprising of non-transfer discretionary outlays such as defense, administration, international affairs etc. and other health care outlays, are distributed equally across all individuals. Table 6 shows government spending on individuals and their share in GDP for 1962 and 2004.

Table 6

Federal Retirement and Non-retirement Transfers to Individuals: 1962 and 2004
(percent of GDP)

Federal Outlay Category	1962	2004
Federal Transfers to Individuals	**5.1**	**12.1**
Retirement Transfers	**3.2**	**7.9**
Social security and railroad retirement*	2.6	4.3
Federal employees retirement and insurance*	0.4	0.8
Medicare*	0.0	2.6
Supplemental Security Income	0.0	0.3
Veterans' Pensions and Disability	0.6	0.3
Non-retirement Transfers	**1.9**	**4.2**
Veterans' Survivors, Education, and Other Benefits	0.5	0.3
Medicaid	0.0	1.5
Unemployment Assistance	0.6	0.4
Family Support, TANF, and Other Public Assistance	0.4	0.3
Earned Income Tax Credit	0.0	0.3
Student Assistance	0.0	0.2
Housing Assistance	0.0	0.3
Food Stamps	0.0	0.3
Child Care Assistance	0.0	0.4
Other Health Services	0.2	0.3
Federal Non-transfer Outlays	**13.7**	**7.7**

Each item is distributed using relative age/sex profiles from the Current Population Survey. The age/gender profiles are generated using weighted mean receipt of transfer payments by age and sex in constant 2004 dollars. The Bureau of Labor Statistic's Consumer Price Index for All Urban Consumers (CPI-U) is used to convert nominal amounts into real ones. CPS sample weights for the March Supplement files are used (variable *wgt* divided by 100) to derive weighted averages. The resulting profiles display considerable volatility by age. A centered moving average is used across 9 age categories by gender to smooth the age/gender profiles. The profiles are then normalized to the per capita transfer of 40-year-old males (except for Federal Employee Retirement category, which is normalized to the value for 60-year-old males).

These relative profiles are used to distribute each of the 14 aggregates listed above by age and gender according to the following procedure. Let the $t^i_{a,t,x}$ represent the per capita transfer of type *i* to a person of age *a* and gender *x*, in period *t*. Let $T^i_{a,t}$ represent the corresponding aggregate federal transfer, and let $p_{a,t,z}$ represent the population of persons aged *a* in period *t*. Let $r^i_{a,t,x}$ represent the relative profile value defined earlier for transfer of type *i*. We first derive the per capita value of transfers received by *N*-year-old males, were *N* is the age to which the profiles are normalized. The procedure for doing so is:

$$t^i_{N,t,x} = \frac{T^i_{a,t}}{\sum_x \sum_a p^i_{a,t,x} r^i_{a,t,x}}$$

The transfer per capita for all other ages is simply the product $t^i_{N,t,x} r^i_{40,t,x}$. The sum of all per capita transfers and the per capita values of the equally distributed categories of federal outlays are shown in Figures 5 and 6. Figures 7 and 8 contain the distributions of non-retirement transfers by age, gender, and year.

REFERENCES

Bates, W. (2001), "How Much Government?: Do High Levels of Spending and Taxing Harm Economic Performance?", draft report commissioned by the New Zealand Business Roundtable, Wellington.

Diewert, E.W. (2001), "Productivity and the Role of Government", University of British Columbia, Department of Economics, Discussion Paper, No. 01-13, Vancouver, available at http://www.econ.ubc.ca/discpapers/dp0113.pdf

Dupor, B., L. Lochner, C. Taber and M.B. Wittekind (1996), "Some Effects of Taxes on Schooling and Training", *American Economic Review*, No. 86, pp. 340-46, May.

Gruber, J. and D. Wise (eds). (1999), *Social Security and Retirement Around the World*, Chicago, University of Chicago Press.

Haltiwanger, J. (2000), "Aggregate Growth: What Have We Learned from Microeconomic Evidence?", OECD, Economics Department, Working Paper, No. 267, Paris.

Harberger, A. (1998), "A Vision of the Growth Process", A*merican Economic Review*, No. 88, pp. 1-32.

Harris, R.G. (1999), "Making a Case for Tax Cuts", paper prepared for the *Business Council on National Issues Global Agenda Initiative*.

Harris, R.G. (2001), "Determinants of Canadian Productivity Growth: Issues and Prospects", forthcoming in A. Sharpe and S. Rao (eds.), *Productivity Issues in a Canadian Context*, Montreal, McGill-Queen's Press.

Kesselman, J.R. (1993a), "Squeezing Universities, Students, or Taxpayers? Issues in Designing a Canadian Income-Contingent Loan Program", in S.T. Easton (ed.), *Ending the Squeeze on Universities*, Montreal, Institute for Research on Public Policy.

———— (1993b), "Ends and Means in Recasting UI, Welfare, Training, and Job Programs", in E.B. Reynolds (ed.), *Income Security in Canada: Changing Needs, Changing Means*, Montreal, Institute for Research on Public Policy.

———— (1994), "Public Policies To Combat Child Poverty: Goals and Options", in K. Banting and K. Battle (eds.), *A New Social Vision for Canada? Perspectives on the Federal Discussion Paper on Social Policy Reform*, Kingston (Ontario), Queen's University, School of Policy Studies.

Kneller, R., M.F. Bleaney and N. Gemmell (1999), "Fiscal Policy and Growth: Evidence from OECD Countries", *Journal of Public Economics,* Vol. 74, No. 2, pp. 171-90.

Madrian, B., C. and L.J. Lefgren (1999), "A Note on Longitudinally Matching Current Population Survey (CPS) Respondents", National Bureau of Economic Research, Technical Working Paper, No. 247.

Nakamura, A.O. and W.E. Diewert (2000), "Insurance for the Unemployed: Canadian Reforms and their Relevance for the United States", in L.J. Bassi and S.A. Woodbury (eds.), *Long-Term Unemployment and Reemployment Policies*, Stamford (Connecticut), JAI Press.

Nakamura, A.O. and P. Lawrence (1994), "Education, Training and Prosperity", in T.J. Courchene (ed.), *Stabilization, Growth and Distribution: Linkages in the Knowledge Era*, The John Deutsch Institute for the Study of Economic Policy, Queen's University.

COMMENTS ON SESSION 2: EVALUATING THE EFFICIENCY AND EFFECTS OF PUBLIC SPENDING

Adi Brender[*]

The papers presented in this section cover several aspects of public expenditure (PE) efficiency from different angels. According to the division of labor between the discussants I shall focus on papers 1-3 (Höppner and Kastrop, Herrera and Pang, and Salinas-Jiménez, Pedraja-Chaparro and Smith). However, the third paper discusses the second one thoroughly, a discussion with which I concur, so I chose to focus on the theme emerging from the Höppner and Kastrop paper. Particularly, this paper is favorable to the decision to amend the EU fiscal rules in a way that will accommodate "quality spending". In this discussion I will try to highlight the requirements that the use of the concept of "quality spending" imposes on policy-makers, and argue that we may still be far from being able to use it broadly for policy purposes. Specifically, I will focus on the difficulties in evaluating the efficiency and effects of public spending – a key element in any meaningful definition of "quality spending".

The government's fundamental optimization problem can be presented as:

Max. social welfare (PE) s.t.: cost of revenue collection

where the cost of revenue collection is composed of:

a) forgone private consumption,

b) administration,

c) dead weight loss.

The optimal level of *PE* is achieved when:

$$MU\,(PE) = MC\,(revenue)$$

i.e. when the marginal utility from public expenditure is equal to the marginal cost of increasing revenue collection.

In recent years globalization has been pushing the marginal cost of collecting taxes up, due to intensifying international competition on the location of firms and individuals. Consequently, governments were pressed to adapt by increasing *MU*(*PE*), either by reducing *PE* or by improving its efficiency. Since the competition between governments is not just for low taxation but for the "value for money" that they deliver, the focus is shifting from cutting the overall level of expenditure to increasing its efficiency. Moreover, one should note that while increasing efficiency improves welfare it does not necessarily reduce the level of public expenditure. Such a reduction is likely to take place if the efficiency gains occur in activities that are far from the margin (necessities), because in these cases

[*] Research Department, Bank of Israel. E-mail: adib@boi.gov.il

only the income effect applies.[1] However, if the efficiency gains affect marginal activities (those that are less critical), both the substitution and income effects apply and the result may be an increase in public expenditure.

A decision to adapt the fiscal criteria to allow more room for "quality expenditure" requires that there would be a clear definition of what "quality expenditure" is and a system for the evaluation of various expenditures according to their quality. In the remainder of my comments I will depict what may be required for an evaluation of expenditures according to their quality, and what are the merits and risks of such an approach.

A key first step in a meaningful evaluation of *PE* is defining government objectives. These objectives should reflect the needs of the individual country and the tastes of its population. Clearly, a government should know what it wants to achieve to be able to evaluate whether a certain expenditure is useful or not. Based on these objectives, policy-makers should examine their programs and assign to them properly set (social) values**.**

If the government is able to set appropriate values for various programs and areas of activity, the next step is to adopt reliable methods for evaluating performance and measuring output. Such methods can contribute substantially to improving the quality of *PE* by identifying weaknesses and by facilitating the adoption of performance-based remuneration and/or budget allocations through the identification of success in reaching pre-specified targets (preferably as part of a long-term plan). More specifically, if the evaluation system properly measures the desired outputs, identifies their respective cost, and the outputs are properly related to outcomes, then the system indicates the "value for money" generated by *PE*, and may be useful in pointing to potential welfare-enhancing reallocations.

Unfortunately, the conditions mentioned above are quite demanding and require significant investment in the design, collection and processing of information. This type of evaluations is particularly useful where outputs – and outcomes – are easy to measure. However, this is the exception in the public sector; in most of its activities outputs are not easily identifiable and even more rarely are they measurable in a meaningful way.

Performance can be evaluated according to different measures, and one should clarify – in advance – which measure is to be used when defining the quality of expenditure:

- technical inefficiency: output is lower than the possible maximum – given the existing inputs (x-inefficiency);
- economic inefficiency: the input composition could be changed to produce more output at the current cost (e.g., by shifting to highly paid high quality staff);

---

[1] The income effect works to reduce public expenditure because the savings from the efficiency gains (the lower cost of producing the original level of services) will be reallocated between public and private consumption according to the population's tastes.

- technical ineffectiveness: the output is not useful in generating the desired outcome (e.g., hospitals may be much less effective than immunization in containing certain diseases);
- economic ineffectiveness: reallocating resources to activities that produce other desired outcomes may increase overall welfare at the same cost (e.g., building safer roads may be more effective than improving the health system in reducing mortality).

Beyond the question of which definition of efficiency or effectiveness underlies the definition of "quality spending" one should also decide which performance criteria are going to be set and on what targets will they be based, as these may vary considerably:

- outputs or outcomes. In most areas only outputs can be measured. However, outputs may not always provide sufficient information to evaluate effectiveness. For example, test results may not be a good enough indicator for the quality of education achieved by the school system;
- minimizing the cost of specific inputs by constraining their use or focusing on minimizing overall costs while allowing flexibility in the choice of inputs. The choice between these two options would depend on whether policy-makers are focusing on predictability and discipline, or on economic efficiency;
- monitoring outputs or inputs: Should entities' evaluations be based on their ability to produce a given range of outputs, using whatever inputs they wish (minimizing cost), or on their ability to produce outputs given a fixed set of inputs (maximizing output).

The choice between these criteria may be a difficult task even in a private firm. Clearly, identifying a meaningful criterion to be used for a cross-country pact, and being able to measure it, is a very complicated process.

Once the criterion/criteria for evaluating performance had been set, one needs to decide according to which dimensions will performance be evaluated. There are various dimensions that could be used such as past performance, plan, a predetermined standard or peer comparison. Policy-makers should clarify which concept, criterion and dimension are more relevant to each unit and each type of expenditure, otherwise evaluations may turn into a source of confusion. According to the concept used in the UK targets should be set in a SMART way (that is: Specific, Measurable, Achievable, Relevant and Timed) to be effective. These criteria may be a good guideline for evaluating the programs that qualify as "quality expenditure" once the other conditions mentioned above are also met.

Despite its substantial merits, the adoption of evaluation-based management and remuneration system may create some risks as well. First, units may focus on achieving the pre-set outputs while ignoring other, which are as important for achieving the desired outcome. Second, the government may find itself focusing on measurable, rather than important, criteria. Finally, lifting budgetary controls, while relying on evaluations, may hurt efficiency and budgetary discipline if the targets

are not properly set, and the evaluation/quality of information is not carefully monitored and verified.

Another risk to the efficiency and effectiveness of public expenditure, due to the focus on measurable quantitative outputs, is related to the quality of service and output. Since quality in public services is hard to measure there may be a tendency of units to curb quality in order to perform better on the quantitative scales. Theoretically, one could use in the evaluation process quality measures as well but, in practice, both internal and client quality valuations have substantial deficiencies.[2]

The usefulness of performance evaluations is also hampered by data and technical difficulties. To be able to improve the reallocation of public resources between activities the evaluation should provide information on the marginal output and efficiency of the various activities. However, in practice, only data on average efficiency are available, and these may not be indicative for efficient decisions. Additionally, a significant part of the public sector deals with prevention of undesired outcomes (defense, fire protection) and the measurement of efficiency in these areas is particularly difficult.

The Data Envelopment Analysis (DEA) discussed in three of the papers in this session may be a useful, though limited, technique for evaluating the efficiency of public expenditure. Notwithstanding the methodological issues raised by Salinas-Jiménez, Pedraja-Chaparro and Smith, this method deals with technical and economic efficiency and can be used for performance benchmarks in parallel units where output is measurable (e.g., local authorities, international). However, the use of this method suffers from the problem of output diversion which may be quite substantial.[3] Additionally, one should be careful in interpreting the results of DEA. For example, the findings by Herrera and Pang that the efficiency of *PE* is decreasing when the share of *PE* in GDP is rising do not necessarily imply that *PE* should be reduced; if social preferences are for more *PE*, then the increasing marginal cost is a price that society is willing to pay for public goods. Only if one can show that the marginal cost of *PE* exceeds the marginal value of that expenditure can these results be used as an indication for the need to reduce *PE*. Unfortunately, our ability to make such judgments is still limited.

---

[2] Although using client response may be more useful for generating decision-makers' (politicians') interest in quality performance.

[3] For example, teaching students for international tests rather than improving education systems.

COMMENTS ON SESSION 2: EVALUATING THE EFFICIENCY AND EFFECTS OF PUBLIC SPENDING

Blanca Moreno-Dodson[*]

The three papers that I will be discussing are quite different in nature. While the one by Rezk is essentially a macroeconomic paper attempting to define the optimal size of the Argentina government using an endogenous growth model, the two following ones deal with a rather microeconomic approach to expenditure efficiency. Afonso and St. Aubyn develop an application of non-parametric approaches to education and health expenditure efficiency in OECD Countries. Salinas-Jiménez, Pedraja-Chaparro and Smith analyze constructively critical issues and methodologies related to the application of non-parametric approaches, such as the ones presented in the previous paper, to assess public sector efficiency.

What do these three different papers have in common? Firstly, they focus on efficiency, and not effectiveness, of public spending. In other words they look to see if things are being done well, as opposed to asking whether those are the right things to do. Secondly, they use a conceptual input/output framework for analysis, implicitly treating the public sector like a private firm. Thirdly, they measure inefficiency by the "distance" from an efficiency frontier (or optimal state). Fourthly, they treat the congestion aspect of public goods (rival and excludable); as opposed to using the "pure" public good Samuelsonian approach to public goods. And finally, they focus on quantitative rather than qualitative aspects of public spending.

In the paper by Rezk on Argentina, public spending is considered as an input entering the production function. The model allows for defining the government size, measured by the ratio of public spending to GDP, which maximizes the per capita growth rate, considered as the final output. It introduces the revenue side of the public budget constraint, and allows for dynamic considerations. The model presents a consistency framework that could be easily applied to other countries as well.

The non-parametric approaches to expenditure efficiency, used by Afonso and St. Aubyn in the analysis of OECD countries, entail an undefined production function with assumptions about the input/output process. This is often the case in this kind of approaches, as indicated by Salinas-Jiménez, Pedraja-Chaparro and Smith more generally in their paper. The study presents different measures of output, related to health and education, without any revenue considerations, and using an analysis purely static.

Before turning into a broader discussion, I would like to make some specific comments on each one of the papers. The Argentina model establishes a clear link

[*] The World Bank.

between spending and revenues, within a long run fiscal sustainability framework and growth objectives. This framework, which leads very nicely to tax policy recommendations, is very appropriate for Argentina, especially at a time when the country is emerging from a severe financial and currency crisis, because it redirects the attention to the revenues that can be generated internally given the needs for public spending. The paper accomplishes its objectives well although it does not allow for a differentiation among different sectors/levels of public spending and does not enter into public spending effectiveness considerations.

The study by Afonso and St. Aubyn about OECD countries presents an interesting international cross-country comparison for developed countries, using homogenous and reliable data. By suggesting two alternative non-parametric methodologies, DEA ands FHD, which they apply masterfully, the authors are able to show the implications of imposing (or not imposing) an efficiency frontier to the data and the difference in the results obtained using both methodologies. Since they find harder to show efficiency under DEA, comparing to FDH, they seem to imply that assuming the convexity of the efficiency function leads to less optimistic results and more cautious implications for policy makers deciding about budget allocations.

Another interesting aspect of this paper can be found in the introduction of intermediary outcomes (called in the paper "quantity inputs"), which are by definition closer to final outputs. The choice of variables (inputs, intermediary outcomes, outputs) used to compare education and health indicators in order to assess welfare, seems to be very comprehensive and appropriate.

The limitations of their approach go well beyond the paper itself. First of all, the methodology attempts to measure, not to explain, efficiency. Therefore the question of which inputs/outputs may be critical remains unanswered. Second, it over-simplifies the problem of attribution since other sector inputs (including private) and outcomes may also influence the final impact on education/health (example of Mexico). Third, it does not introduce external factors, different in all countries, which may also affect education and health outcomes. Finally, the static analysis may be somehow limited since final impact on the welfare of the population, measured by health and education indicators, usually requires a longer term horizon to materialize and a dynamic analysis illustrating trade-offs among different sectors (for example, basic infrastructure and access to water also affect health outcomes) may be needed.

The third paper by Salinas-Jiménez, Pedraja-Chaparro and Smith on issues and methodologies highlights most of the limitations encountered by the non-parametric approaches to efficiency analysis, of which the previous paper is a very good example. The study sets up the stage on public sector considerations very nicely, making the reader aware of the differences from private sector decision units. It presents a very thorough review of the literature and methodologies, and explains in detail the limitations of using the DEA approach.

In my general remarks, I would like to focus rather on the last two papers and interrogate the audience about the following question: How suitable are the non

parametric cross country approaches to assess public spending efficiency in developing countries?

Developing countries often lack the kind of reliable and homogenous data needed for DE and FHD, and present a great diversity of exogenous factors that make international comparison difficult. Moreover, the trade-offs among sectors are of critical importance for the design of public expenditure programs, and effectiveness of spending is as important, if not more, as efficiency. Finally, the longer term horizon needed to find impact on the ground and final results often covers many more years than in developed countries, due to weak institutional capacities and sometimes political instability.

In our discussion, I would like to suggest that we think about public spending using a broader framework, which I call the "three tiers of performance measuring". In the efficiency tier (lowest level), one would find efficiency and cost-benefit ratios which measure how economically inputs (funds, resources, expertise...) are converted into outputs. Examples of outputs would be the number of schools constructed, the miles of roads built, and so on and so forth. The non-parametric approaches to efficiency presented two of the papers in this session are most useful to enlighten policy makers at this level.

In addition, at the effectiveness tier (second level), one would see intermediary outputs, using a rather multi-sectoral perspective. Performance would be measured by progress towards strategic goals, linking different sectors in the economy according to pre-established development priorities. In that respect, improvements in the health sector would be judged according to the priorities established, in connection with other sectors such as education or rural development, and not just in reference to a specific project or intervention. For example, instead of measuring the miles of roads built, the focus would be on how much access of the different groups of the population (by region, village, age, gender, occupation, etc.) has improved as a result of the overall strategy. Can nurses now make it to the hospital? Do children in remote areas use the road to attend school?

Finally, in the third and highest level tier, one would find the final growth and poverty reduction objectives, such as in the Millennium Development Goals (MDGs), including variables like GDP growth per capita, infant mortality and maternal mortality. These final goals can be achieved only as a result of improvement in the first two levels, efficiency and effectiveness. Not only things must be done right and efficiently, but actions need to be taken in the key priority areas for development, in an integrated manner. Trade offs among sectors are key to determine final results due to the numerous linkages among them.

The main message that I would like to convey to this audience is that in addition to searching for the best methodologies to assess public expenditure efficiency, we should also intensify our efforts to determine the factors that would trigger higher effectiveness and contribute to achieving ambitious final results, as reflected in the MDGs. Efficiency alone will not be enough to halve the population living in extreme poverty and improve the education and health standards of the poorest.

COMMENTS ON SESSION 2: EVALUATING THE EFFICIENCY AND EFFECTS OF PUBLIC SPENDING

Doris Prammer[*]

This contribution discusses the papers by Marie-Anne Deussing and Jagadeesh Gokhale, which give insights into the distributional effects of (net) transfers in Canada and the USA respectively. Even though both papers deal with the distribution of transfers, they approach the issue from different angles and focus on very different distributional effects. While Deussing concentrates on the provincial distribution of federal net transfers across family income groups and across provinces, Gokhale analyses the impact of general government non-retirement spending on working and earning.

"Federal Taxes and Transfers Across Canada: Impact on Families" by Marie-Anne Deussing

In contrast to many studies focusing on direct federal taxes and direct transfers to income groups, Deussing's paper provides a fuller picture. In addition to analysing the impact of direct federal taxes (such as personal income tax and employment insurance contributions) and the effect of indirect taxes on post-tax, post-transfer income, Deussing – in line with an earlier study by Finn Poschmann – also accounts for direct and indirect federal transfers. However, while the findings of Poschmann are based on shares in historical indirect transfer programmes, Deussing allocates indirect transfers stemming from the Canada Health and Social Transfer (CHST) to provinces and income groups according to the actual provincial spending pattern. Furthermore, this paper also takes a different approach with regard to the allocation of equalization entitlements. Here, equalization is treated as a block transfer instead of a tax point transfer.

The interesting key results on the distributional effects are that:

- the total federal tax is progressive for all provinces,
- the impact of direct federal transfers is progressive as well, with greater variation across provinces and across income groups than observed for federal taxes.

These two results are well in line with the 1998 study by Poschmann. However, his third main conclusion, namely that "*the federal government collects taxes from low-income Canadians in high-income provinces in part to fund transfers to higher-income residents of poorer provinces*" (Poschmann, p. 3), cannot be supported by the current study. The diverging conclusions basically reflect the different treatment of indirect transfers, which exert a considerable influence on the distribution of the overall net tax burden on income groups across provinces.

---

[*] National Bank of Austria.

Comments

Regarding the theoretical approach, I find Deussing's way of analysing the distributional effects of federal net transfers very attractive. Yet a short introduction into the Canadian public finance system – for example in an appendix – would have been helpful for readers who are not that familiar with the Canadian public finance system.

Furthermore, although the effective net transfer rates across income groups and across provinces provide a clear picture of the progressivity of the Canadian federal tax system , this is not the end of the story. The distributional power of a tax system is not only given by effective tax rates, but also by the number of people paying these tax rates. Consequently, some information on the distribution of income would be welcome. Hence, it would be interesting to see similar results calculated using family income percentiles.

As this study is closely related to Finn Poschmann's study – at least the part on direct taxes and direct transfer payments – the two papers could provide interesting comparative information. Comparing Poschmann's and Deussing's table on federal direct transfers as a percentage of post-tax, post-transfer income across income groups, I realised that while overall federal direct transfers decreased from 12.2 to 9.1 per cent of post-tax, post-transfer income, federal direct transfers to the lowest income group increased from 36.9 to 39.6 per cent. This gives rise to the question as to whether this change is the mere result of methodological differences – or is there any economic explanation to it? Have federal direct transfers been redesigned to decrease income inequality, or are the higher transfer payments just the result of an even more unequal primary income distribution?

In general, has the progressivity in the Canadian federal net transfer system been increasing or decreasing?

"The Cross-Cohort Distribution of Government Non-retirement Transfers and Its Impact on Working and Earning" by Jagadeesh Gokhale

Rather than analysing where government transfer spending goes like Deussing, Gokhale puts the focus on how these transfers affect the economy.

Gokhale classifies government non-retirement transfers according to their basic effects on recipients into *defensive* and *offensive* spending. While defensive transfers are intended to prevent the bad consequences of market failure (*i.e.* provide just the needy with welfare payments) offensive spending provides services that enhance worker productivity, thereby improving the overall economic and social environment.

Gokhale then investigates the direct impact of four types of federal transfers, namely unemployment insurance transfers, child care transfers, education subsidies and other government non-retirement transfers on labour market outcomes.

Applying regression analysis the author finds that unemployment benefits reduce earnings growth and accelerate the rate of non-work by males. Faster growth in educational spending seems to positively affect male earnings growth, while it appears to dampen female earnings growth. In contrast, child care transfers are associated with higher earnings growth and higher full-time job rates for women only.

Hence, overall these regressions provide only *"little support to the idea that non-retirement transfers improve the functioning of labor markets and the economy"* (Gokhale, p. 478).

Comments

Oversimplifying, one could draw the following policy conclusions from Gokhale's results: *"to improve the economy by subsidizing the acquisition of skills and provide services that increase workers' labor force participation"* (Gokhale, p. 459), the transfer spending of the federal government should be redirected. Unemployment benefits should be cut or eliminated altogether, as they negatively affect male earnings growth and accelerate the rate of non-work by males, while they do not exert any significant impact on female labour market variables. Furthermore, educational spending for women should be cut and invested either in educational spending for men or child care transfers for women. Both measures should enhance worker productivity according to Gokhale's results.

However, I would not dare to give such policy advice, since I am not ready to defend Gokhale's results on an economic ground.

First, reduced earnings growth need not necessarily reflect the level/rate of unemployment benefits; both variables might just as well indicate reactions to cyclical conditions – such as the overall level of unemployment. By decreasing trade union negotiation powers, both high levels of unemployment and long periods of unemployment – via decreasing human capital – would decrease earnings growth while unemployment benefits increase at the same time.

Second, I do not have any economic interpretation for the result that, while boosting male earnings, educational spending growth should negatively affect female earnings growth.

Also, the third result, namely that child care transfers are associated with higher earnings growth and higher full-time job rates for women, is not evident to me. In Austria, the newly-introduced lump-sum child care benefit has reduced the labour force (full-time) re-entry rate of women with children below the age of 2¼. According to a WIFO paper, the percentage of woman returning to the labour market before their children reach the age of 2¼ has dropped from 54 to 35 per cent. Obviously, the incentive effects of child care benefits do not only depend on the monetary transfer alone, but also on the "accompanying institutional setting" such as the possibility of maternal leave and its length or dismissal protection during parental leave. Hence, it would be interesting to learn to what extent Gokhale's

result would be altered with respect to the European child care benefit system in general and the Austrian in particular.

Even though the author himself admits that the dividing line between transfers serving defensive versus offensive objectives is not sharp (for example child care support may enable household adults to spend more time at work), I would refrain from such a classification on the basis of effects on the economy. Even the most defensive transfer might exert externalities that improve the overall economic and social environment as well as the functioning of the markets. In general, while findings are ambiguous on the size of the effect, there is evidence that "*redistributive policies that result in less income inequality could well promote growth*" (Tanzi and Zee, 1997, p. 198) (for a more recent literature survey, see Harris, 2002). Hence, to my mind, this classification is even more questionable than the separation into productive and unproductive general government expenditure made in other contributions to this book.

After these very general remarks, I have some more specific comments concerning the econometric part of the paper. First, the author is not very clear on the econometric model. Judging from the limited explanation available, I assume that the author used pooled panel regression. This specification foregoes the opportunity to control for unobserved time-fixed cohort specific effects. This is important, as cohort specific effects in the form of different preferences with regard to leisure or participation in child care might be prevalent, given that the difference between the youngest and the oldest cohort is as much as 20 years. Furthermore, as these preferences may be correlated with the variables of part-time work and unemployment they will lead to biased estimates of other parameters in the model. Hence, the fixed-effects panel data estimator may be more appropriate for this data set than the pooled OLS estimator.

Gokhale himself suspects that the regression results might be driven by changes in unobserved variables including macroeconomic shocks. Moreover, as both unemployment benefits and the level of employment (especially female labour force participation) are very susceptible to the cycle, I wonder if one can actually infer the causality that unemployment benefits cause lower earnings growth. Instead, I suggest controlling for the business cycle explicitly, by including some output gap variable.

Finally, the estimation of the effects of government non-retirement transfers on labour market outcomes does not take into account the life cycle position of the respective cohort. While I assume that child care transfers have a significant effect for women in their late 20s and early 30s, they will possibly not affect labour market decisions of older women. Interacting the child care benefit variable with an age dummy could control for this and separate the overall effect into distinct effects. This could provide additional insights into the effects of transfer payments on the labour market and earnings, thereby strengthening the assumption that transfer payment might be directly "offensive" for particular age groups.

REFERENCES

Harris, R.G. (2002), "Social Policy and Productivity Growth: What Are the Linkages?", Industry Canada Research Publications Program, Discussion Paper, No. 11.

Lutz, H. (2003), "Auswirkungen der Kindergeldregelung auf die Beschäftigung von Frauen mit Kleinkindern", *WIFO Monatsberichte*, 03/2003.

Poschmann, F. (1998), "Where the Money Goes: The Distribution of Taxes and Benefits in Canada", C.D. Howe Institute, Commentary No. 105, April, Ottawa.

Tanzi, V. and H.H. Zee (1997), "Fiscal Policy and Long-Run Growth", IMF, Staff Papers, Vol. 44, No. 2.