BANCA D'ITALIA

EUROSISTEMA

Temi di Discussione

(Working Papers)

Estimating DSGE models with unknown data persistence

by Gianluca Moretti and Giulio Nicoletti

March 2010

Number 750

BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working papers)

Estimating DSGE models with unknown data persistence

by Gianluca Moretti and Giulio Nicoletti

Number 750 - March 2010

The purpose of the Temi di discussione *series is to promote the circulation of working papers prepared within the Bank of Italy or presented in Bank seminars by outside economists with the aim of stimulating comments and suggestions.*

The views expressed in the articles are those of the authors and do not involve the responsibility of the Bank.

ESTIMATING DSGE MODELS WITH UNKNOWN DATA PERSISTENCE

by Gianluca Moretti* and Giulio Nicoletti*

Abstract

Recent empirical literature shows that key macro variables such as GDP and productivity display long memory dynamics. For DSGE models, we show that long memory data can substantially bias deep parameter estimates when a standard Kalman Filter-MLE procedure is adopted. We propose a 'Generalized' Kalman Filter to deal effectively with this problem: our method connects to and innovates upon data-filtering techniques already used in the DSGE literature. We show that our method produces more plausible estimates of the deep parameters as well as more accurate out-of-sample forecasts for macroeconomic data.

JEL Classification: C51, C53, E37.
Keywords: DSGE model estimation, Kalman Filter, long memory dynamics.

Contents

* Bank of Italy, Economics, Research and International Relations.

1 Introduction[1]

Over the last few years, DSGE models have become the workhorse of modern macroeconomic modeling and both academics and practitioners often use them to produce macroeconomic forecasts. The reduced form of these models, obtained once they are solved for expectations, is usually cast into a state space form, which describes the data as a (dynamic) linear combination of fundamental exogenous shocks and endogenous states. Under some general conditions this implies a (restricted) finite order VARMA representation for the data generating process (DGP).[2]

The appealing feature of DSGE models is that they are deeply grounded on economic theory. They link macroeconomic variables to the microeconomic behavior of agents, imposing meaningful restrictions on the data generating process, so called 'cross equation restrictions'. An important drawback of DSGE models is that their reduced form representation might be too stylized to possibly capture the persistent dynamics of the data, as highlighted by Cogley and Nason (1992) and, back in the early days, by Kydland and Prescott (1982). An important consequence is that when these models are taken to the data, the value of the estimated deep parameters fail to square well with available evidence either from national accounting (as for labor share, capital depreciation) or, more generally, from microeconometric evidence. This issue, as highlighted by Prescott and McGrattan (2007) (appendix), seriously weakens the use of 'microfounded models'.

Earlier attempts to reconcile the theoretical framework and the statistical properties of the data have mostly focused on modifying cross equation restrictions in order to add persistence to the simulated economy and have a better match with the data. This led to models which included a large number of 'frictions', endogenous state variables and exogenous state processes. While still at the forefront, Chari, Kehoe, and McGrattan (2008) criticized this line of research as 'lacking discipline', e.g. weak microfoundations and economic reasoning. A second line of research has focused on finding the "right" data transformation to resolve the mismatch between the data and the model implied dynamics. This has been done by either filtering the data before estimation, for example with either an HP or a Band Pass filter, or by explicitly including one or more unit roots in the model. The implications of these procedures are not always crystal clear: Chang, Doh, and Schorfheide (2007) show that we can not statistically discriminate *whether the 'true' DGP i*s better approximated by a model that includes unit roots in the exogenous states or by one that has more 'frictions'. At the same time, Canova (2009) and Canova and Ferroni (2009), show that different data filtering can influence the estimation results

[1] We thank Luc Bauwens, Michel Juillard and Raf Wouters for discussing with us a preliminary version of the paper. We are indebted to Stefano Neri, F. Mihoubi, participants at the 'Third Dynare Conference' and SNDE annual symposium for helpful comments at various stages of the paper. We are also grateful to Richard Baillie, Fabio Busetti, Michele Caivano, Sergio Santoro, Enrico Sette and Giordano Zevi for useful comments. The usual disclaimer applies here.

[2] The order of the VARMA depends upon the type of model at hand. In particular it is related to the order of the AR in the exogenous states as well as the number endogenous states which are treated as non observed variables, see Ravenna (2007) for more details.

in a non-trivial way.

In this paper we take a different route from most of previous literature. We propose a method that does not impose any strong assumptions on the statistical nature of model exogenous dynamics and makes the structure of a DSGE model consistent with the persistence of the data at hand. This is done by adding a non-parametric step to the standard Kalman filter to adapt the VAR(MA) order of the DSGE model to the data persistence without modifying cross equation restrictions. Our procedure relates to the Kalman filter in a similar way as the Generalized Least Squares refer to OLS, and therefore we named it 'Generalized' Kalman Filter ('GKF' heretofore). We show that our technique is equivalent to filtering the data in a way that optimally removes the persistence which is left unexplained by the model structure.

The theoretical justification for our approach is the following. The order of the VAR(MA) representation of the data implied by a DSGE model depends crucially on the autoregressive order of the exogenous state processes. Highly persistent data might require models with a large number of lags to approximate accurately their dynamics. However, on the economic side there is an important drawback in using 'ad hoc' high order exogenous processes. In fact, by formulating a specific type of process for exogenous states, the researcher also casts assumptions on how expectations about future variables are formed: if a researcher modifies the order of the exogenous state processes to get a better fit of the data, the economic interpretation associated with the model might change dramatically every time a new dataset is considered. Probably this is one of the main reasons why exogenous state processes tend to be cast in a few conventional ways, such as simple AR(1) or unit root process, with few exceptions in the literature. Moreover, using using high order AR processes would not be efficient from an econometric point of view, due to the dramatic increase in the number of parameters to be estimated, as we discuss in section 3.1

For these reasons our approach aims instead to make the structure of DSGE models consistent with the degree of persistence of the data at hand, allowing us to handle a very persistent data generating process, such as in the case of long memory process. As we show below, this implies that we do not violate the cross-equation restrictions, which are applied to a set of filtered, rather than to the raw data. The method is also parsimonious since we use a non-parametric approach which enables us to approximate the dynamics of persistent processes by estimating few parameters. In particular, our filter spans from the case of strictly stationary to the most extreme case of long memory processes. Those have the appealing feature of having non-negligible spectral mass at low frequencies, 'the typical spectral shape of economic time series' (Granger (1966)), but their approximation with short order ARMA processes might be very poor, in particular when standard statistical lag-selection criteria are employed (Granger and Joyeux (1980)).

Our results can be summarized as follows. First we show that a strong degree of persistence in the data can substantially bias the structural parameters estimates of a DSGE model when exogenous state dynamics is misspecified with respect to the true one. Following the thread of McGrattan (2006) and Ruge-Murcia (2007) we recur to

simulated data. We evaluate the ability of the Maximum Likelihood (ML) methods to estimate the DSGE structural parameters and find a relevant amount of bias in the estimates: the stronger the persistence of the data the larger the bias. We show in the simulation that our approach dramatically reduces the bias in estimated parameters, which turns out to be negligible even when data are highly persistent. Finally, we turn to real data, and in line with Ireland (2004), we estimate a standard RBC model on U.S. data using both the Standard Kalman filter and the GKF. We report different and more plausible parameter estimates with the GKF than with the standard filter. Further evidence of the ability of the generalized Kalman filter to better capture the dynamics of the data, comes from more accurate out of sample forecasts compared to the standard filter. Finally, as a further check, we show that our method outperforms the case where a unit root is imposed on the technology shock and the model is estimated in levels.

The plan of the work is the following. In section 2 we present references needed to apply long memory methods to DSGE models, appendix 9 contains further details on general long memory processes. In 3 we describe the Generalized Kalman filter and the related estimation technique. Since the method can be applied also outside the field of DSGE models, we make this section as self-contained as possible, in order to be used as a cookbook recipe even by non-DSGE researchers. In section 4 we run some simulation to evaluate the effects of dynamic misspecification of exogenous states for the estimation of unobserved endogenous states (e.g. the capital stock) and deep parameter estimates. In this section we explain how the cross equation restrictions of the DSGE model come into the picture and how in our approach measurements are filtered in way which is consistent with the model. In section 5 we estimate a RBC model using real data for the U.S. economy and we undertake a forecasting exercise as in Ireland (2004) in section 6. In section 7 we discuss how our analysis relates to the case of pure unit roots in technology shocks. Some conclusions follow.

2 Background and objectives

In this section we provide some background relating long memory processes and DSGE models, some more details on long memory processes are in appendix (9).

Long memory processes have been extensively studied in time series analysis and a good review of their properties can be found in Robinson (2003) and Baillie (1996). Their main characteristic is an autocorrelation function that decays slowly to zero: compared to ARMA processes, it decays hyperbolically rather than exponentially. A well known class of long memory processes, introduced by Granger (1980), is the Autoregressive Fractional Moving Average Process of order d (ARFIMA(p,d,m)), which is defined as

$$(1-L)^d \phi(L) y_t = B(L) e_t, \tag{1}$$

where e_t is a white noise process with variance σ^2, $\phi(L)$ is a lag polynomial of finite order p, $B(L)$ is a lag polynomial of order m and $d \in [0,1]$. The basic building block of fractionally integrated processes is called the 'fractional white noise': it is a special case,

defined by setting $p, m = 0$ in (1). Since d can take any value in the interval 0 and 1, ARFIMA processes fill the gap between a strictly stationary process ($I(0)$), when $d = 0$, and a unit root process, when $d = 1$. For $0 < d < 0.5$ the process is "stationary"with hyperbolic rather than exponential decay of the autocorrelation function. When $0.5 < d < 1$ the process is non stationary (the squared sums of its autocorrelations do not converge), but, differently from unit root processes, it is still mean reverting. When $d < 0$ similar properties hold, with the difference that the autocorrelation function oscillates around zero in a way which shows an 'overdifferencing' pattern of the time series.

Several empirical studies argue that many macroeconomic variables are better described by long memory rather than AR(I)MA models: Diebold and Rudebusch (1989) were the first to report evidence of long memory in many macroeconomic time series; their results were recently confirmed by Mayoral (2005) with an updated version of the dataset. Abadir, Caggiano, and Talmain (2006) show that long memory processes fit better the dynamics of the Nelson and Plosser dataset compared to ARIMA processes. Gadea and Mayoral (2006) provide robust empirical evidence supporting the hypothesis fractional integration for the inflation rates of the OECD countries while Altissimo, Mojon, and Zaffaroni (2009) reach the same conclusion for the euro area inflation rates. Evidence of long memory dynamics for hours worked has been found by Gil-Alana and Moreno (2006).

The main theoretical reason for long memory behaviour in macroeconomic data is the aggregation of heterogenous dynamic processes: fractional integrated process can be generated by *linearly* aggregating heterogenous ARMA models (Granger (1980)). More recently, Haubrich and Lo (2001) and Abadir and Talmain (2002) show that long memory processes can be produced in microfounded macroeconomic models with different sectors. In particular, Abadir and Talmain show that introducing heterogenous sectors in a monopolistic competition DSGE model leads to a *non-linear* long memory process for the aggregate output whose autocorrelation function reproduces the empirical shape found on the US GDP.

2.1 Background: DSGE and long memory

This section discusses our approach to incorporate potentially persistent data into the state space representation of a DSGE model. Consider a scalar state space model:

$$\theta_{t+1} = \Phi\theta_t + \varepsilon_{t+1} \quad (2)$$

$$y_t = H\theta_t, \quad (3)$$

where θ_t is an (unobserved) state variable, modeled by an AR(1) process; y_t is the measurement variable and the innovation ε_t is a normal i.i.d. process with standard deviation σ_ε. The dynamic properties of the data y_t closely depend on the specification of the transition equation for θ_t, in particular to its autoregressive order. For the case at hand, it is straightforward to show that y_t has an autoregressive representation of order one:

$$y_t(1 - \Phi L) = H\varepsilon_t.$$

The reduced form of DSGE models can be represented in a state space form; for instance, 2 - 3 is consistent with the case of DSGE models having only one unobserved exogenous state θ_t and one observable variable where the matrices $\Phi \equiv \Phi(\psi)$, $H \equiv H(\psi)$ are derived from the solution of the model as functions of some deep parameters ψ. In the more standard case (e.g. an RBC model) of one unobserved endogenous state variable (e.g. capital) and one AR(1) exogenous state (e.g. productivity), the measurement y_t (e.g. the GDP) would follow a (restricted) ARMA(2,1) process (see Ravenna (2007)).

A way to account for strong persistence in the y_t is by allowing the innovations ε_t to be a long memory process. Assume for example the following fractional noise representation:

$$(1-L)^d \varepsilon_t = e_t. \tag{4}$$

with $e_t \sim IID$, then by construction the exogenous variable θ_t is a fractional autoregressive process,

$$(1-\Phi L)(1-L)^d \theta_{t+1} = e_{t+1}, \tag{5}$$

and the measurements have the following representation:

$$y_t(1-\Phi L)(1-L)^d = He_t, \tag{6}$$

which is a ARFIMA(1,d,0) process.

As we show in the paper, this way of introducing long memory in a state space allows us to account for the persistence of the data at hand without any implication for the structural cross-equation restrictions of the DSGE model. Broadly viewed, our approach connects to well-established methods in DSGE modelling based on the idea of data-filtering. In fact, as discussed in details in section 4, the gist of our approach is to introduce a new variable $\tilde{y}_t$ defined as

$$\tilde{y}_t = \sum_{j=0}^{\infty} l_j y_{t-j},$$

where the weights $\{l_j\}_{j=0}^{\infty}$ are function of the persistence of the original data which is not captured by the model structure[3] and to estimate the model using this new variable:

$$\tilde{y}_t(1-\Phi L) = He_t. \tag{7}$$

7 embeds the same structure (Φ, H) of the original model but the potential 'excess' persistence of the data is removed.[4]

An alternative and more straightforward strategy would be to directly increase the order of the exogenous state process and use conventional state space techniques to estimate the enlarged set of parameters. However, adding (too many) ad hoc lags would

[3]The fractional noise representation used in this section is merely for expositional convenience and it does not play any role in our estimation strategy. Our method, and in particular the estimation of the weights l_j, can be applied in both the case of errors being fractionally integrated or following even more general types of long memory processes. For the case of fractionally integrated processes, the appendix (9) shows how the weights l_j relate to the parameter d.

[4]This is implemented with an iterative estimation procedure.

change the way agents form expectations about future variables, every time a new data set is considered. This would make the model exposed to some type of Lucas' critique. Some further discussion of important technical estimation issues of this alternative approach are relegated to a later section, (3.1).

3 Generalized Kalman Filter

In this section we introduce a modification of the Kalman Filter that can be used when the data persistence is unknown and possibly very strong. We start presenting the equations of the generalized filter and assume that the autocorrelation structure of the data is known. We remove this assumption in section 3.1 where we describe our estimation strategy.

Consider the following state-space model:

$$\theta_{t+1} = \Phi\theta_t + \varepsilon_{t+1} \tag{8}$$

$$\mathbf{y}_t = H\theta_t + v_t \tag{9}$$

$$E\left(\varepsilon_t\varepsilon'_{t+k}\right) = \left\{ \begin{array}{ll} Q & k=0 \\ 0 & otherwise \end{array} \right\} \tag{10}$$

$$E\left(v_t v'_{t+k}\right) = \left\{ \begin{array}{ll} R & k=0 \\ 0 & otherwise \end{array} \right\} \tag{11}$$

The $n-$dimensional state θ_t in the transition equation is modeled as an autoregressive process with innovation ε_t, Φ is its $n \times n$ transition matrix and $\mathbf{y}_t$ a $l-$dimensional vector of measurement variables, related to the states via an $n \times l$ matrix H. v_t is a vector of measurement errors, also of dimension l. Q is the variance covariance matrix of ε_t while R is an $l \times l$ diagonal matrix for the variance covariance matrix of the independent measurement errors.

Although we assume, for expositional simplicity, to have only one state variable in the θ_t vector, in the remaining part of this section we explain our methodology by referring to a generic state space model of type (8-11), our approach can be easily extended to more general cases. To provide a short reference here, the simple state space (8-11) refers to a DSGE where θ_t corresponds to exogenous states (e.g. total factor productivity), commonly modelled as autoregressive processes of order one. In section 4 we examine how to apply our method to DGSEs and to extend the state space in order to encompass the case of one endogenous state variable (e.g. capital).

In (8-11) the state variable θ_t is assumed to be not observed. When the transition and measurement equations are linear and the shocks ε_t and v_t are normally distributed, the Kalman filter provides the best estimate of the state, conditionally on the assumed autoregressive process of the transition equation. In the following discussion we assume that both linearity of the model and normality of the shocks are good approximations of the data generating process.[5]

[5]This is also consistent with what is typically done in the DSGE literature, where models are typically linearized around their steady state and normal shocks are considered.

Define:

$$\begin{aligned}\theta_{t|t-1} &= E\left(\theta_t|\Im_{t-1}\right),\\ P_{t|t-1} &= E\left[\left(\theta_t-\theta_{t|t-1}\right)\left(\theta_t-\theta_{t|t-1}\right)'\right],\end{aligned}$$

respectively the prediction of the state θ_t based on all information up to $t-1$ and its dispersion matrix, then the Kalman filter equations are given by:

$$\theta_{t|t-1} = \Phi\theta_{t-1|t-1}, \tag{12}$$
$$P_{t|t-1} = \Phi P_{t-1|t-1}\Phi' + Q, \tag{13}$$
$$\nu_t \equiv \mathbf{y_t} - H\theta_{t|t-1}, \tag{14}$$
$$F_t = E\left[\left(\mathbf{y_t}-H\theta_{t|t-1}\right)\left(\mathbf{y_t}-H\theta_{t|t-1}\right)'\right] = HP_{t|t-1}H' + R, \tag{15}$$
$$K_t = P_{t|t-1}H_t'\left(HP_{t|t-1}H' + R\right)^{-1}, \tag{16}$$
$$\theta_{t|t} = \theta_{t|t-1} + K_t\left(y_t - H\theta_{t|t-1}\right), \tag{17}$$
$$P_{t|t} = \left(I - K_tH\right)P_{t|t-1}, \tag{18}$$

Equation (12) and (13) are respectively the state prediction and its variance, given the information set at $t-1$. Equation (14) defines the one-step-ahead prediction error of the measurements and (15) its dispersion matrix. This group of equations is called the *projection step*, in which states are predicted forward in time, leaving the information set of the observer constant. The last block of equations gives the so called *information updating step* in which the projection of the states is revised due to the arrival of time t information. The Kalman gain, (16), is a weighting matrix that minimizes the variance of the state forecast error, given available information:

$$K_t \equiv \arg\min_{K_t} E\left[\left(\theta_t-\theta_{t|t-1}-K_t\left(\mathbf{y_t}-H\theta_{t|t-1}\right)\right)\left(\theta_t-\theta_{t|t-1}-K_t\left(\mathbf{y_t}-H\theta_{t|t-1}\right)\right)'\right]$$

and it is used to extract information from time t prediction errors ν_t. By construction, past predictions $\theta_{t|t-1}$ and the constructed forecast errors ν_t are assumed to be orthogonal, an assumption which holds true when the transition equation is not misspecified. Finally, $P_{t|t}$ in (18) is the variance covariance matrix of the state conditional on information at time t.

Equations (12) to (18) represent a system whose parameters can be estimated by maximum likelihood. In fact, if we assume that the innovations $\{e_t\}$ and $\{v_t\}$ are multivariate Gaussian, then the distribution of the $\mathbf{y_t}$ conditional on the state θ_t and the information set at time $t-1$, is given by

$$\begin{aligned}f_{\mathbf{y_t}|\theta_t,\Im_{t-1}}\left(\mathbf{y_t}|\theta_t,\Im t-1\right) = (2\pi)^{-\frac{1}{2}}\left|HP_{t|t-1}H' + R\right|^{-\frac{1}{2}}\\ \exp\left\{-\frac{1}{2}\left(\mathbf{y_t}-H\theta_{t|t-1}\right)\left(HP_{t|t-1}H' + R\right)^{-1}\left(\mathbf{y_t}-H\theta_{t|t-1}\right)'\right\}\end{aligned} \tag{19}$$

which can be maximized with respect to the unknown parameters.

When the dynamics of the state process are misspecified, (e.g. the order of the autoregressive process is too short to approximate the DGP), the state space model described above can no longer be taken as a good approximation of the data generating process: the stronger the persistence in the original DGP, the more severe this problem is likely to be. As a result, both the projection step and the updating step, as designed in the standard Kalman filter will fail to be optimal and the parameter estimates will be biased. We cope with this problem by setting up a state space model where we allow ε_t to be serially correlated,

$$\theta_{t+1} = \Phi\theta_t + \varepsilon_{t+1} \tag{20}$$

$$\mathbf{y_t} = H\theta_t + v_t \tag{21}$$

$$E\left(\varepsilon_t\varepsilon'_{t+k}\right) = \begin{cases} \sigma^2\rho_\varepsilon(k) & \forall k <= m \\ 0 & otherwise \end{cases} \tag{22}$$

$$\Omega^m = \sigma^2 \begin{bmatrix} 1 & \rho_\varepsilon(1) & \cdots & \rho_\varepsilon(m) \\ \rho_\varepsilon(1) & 1 & \cdots & \rho_\varepsilon(m-1) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_\varepsilon(m) & \rho_\varepsilon(m-1) & \cdots & 1 \end{bmatrix} \tag{23}$$

$$E\left(v_t v'_{t+k}\right) = \begin{cases} R & k = 0 \\ 0 & otherwise \end{cases} \tag{24}$$

where σ^2 is the variance of ε_t and Ω^m is the matrix containing the autocorrelation structure of ε_t, $\{\rho_\varepsilon(k)\}_{k=1}^m$, up to lag m.

This set up can be handled in such a way to reconcile it with the standard approach, by constructing a 'bridge variable' which is an AR(1) with i.i.d. innovations, as follows:

1. Consider the Cholesky decomposition of Ω^m, namely
$$\Omega^m = \sigma^2\Gamma\Gamma'$$
where Γ is a lower triangular matrix with elements $\{\gamma_{i,j}\}_{i,j=1}^{m+1}$.

2. Invert the matrix Γ and define $L \equiv \Gamma^{-1}$. Construct using the original states θ_t (all vectors are denoted with bold typeset in the remaining part of this section) a vector of transformed variables $\mathbf{z}_t$ defined as
$$\mathbf{z}_t = \begin{bmatrix} z_{t-m} \\ \vdots \\ z_{t-1} \\ z_t \end{bmatrix} = \begin{bmatrix} l_{0,1} & 0 & \cdots & 0 \\ l_{1,1} & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ l_{m,1} & l_{m,2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \theta_{t-m} \\ \vdots \\ \theta_{t-1} \\ \theta_t \end{bmatrix} \equiv L\theta_t \tag{25}$$

3. By construction, the transformed variable $z_t \equiv \theta_t + \sum_{j=0}^{m-1} l_{m,m-j}\theta_{t-j-1}$ is an AR(1) process, i.e.
$$z_{t+1} = \Phi z_t + e_{t+1}, \text{such that holds: } E(e_t \mid z_t, z_{t-1}, \ldots, z_{t-m}) = 0, \tag{26}$$

this step removes any autocorrelation in the residuals in equation (20), as in a GLS procedure; we refer to that as the 'GLS transformation'.

Coefficients of the L matrix, denoted as $\{l_{m,m-j}\}_{j=0}^{m-1}$, can be thought of as optimal weights such that $E\left(e_t\,|\mathbf{z_t}\right)=0$ is satisfied when regressing z_t on its lagged value. This is the moment condition that we implicitly impose in the next section when estimating the parameters of the model. It can be immediately seen that if the ε_t are serially uncorrelated, then the coefficient $\{l_{m,m-j}\}_{j=0}^{m-1}$ are all equal to zero [6] and θ_t is equal to z_t which leads us back to the standard case. We are now able to define what we call henceforth the Generalized State Space model.

Definition 1 *Let's consider the variables defined in the state space model (20-24), where v_t is a (vector) of independent white noise with (diagonal) variance matrix R. The generalized state-space model is then defined as*

$$\begin{bmatrix} z_{t-m+1} \\ \vdots \\ z_{t+1} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{(m)} & I_{(m)} \\ \mathbf{0}'_{(m)} & \Phi \end{bmatrix} \begin{bmatrix} z_{t-m} \\ \vdots \\ z_t \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{(m)} \\ \\ e_{t+1} \end{bmatrix} \tag{27}$$

or

$$\begin{aligned} \mathbf{z}_{t+1} &= \Psi \mathbf{z}_t + \mathbf{e}_{t+1} \\ \theta_t &= z_t + \sum_{j=0}^{m-1} \gamma_{m,m-j} z_{t-j-1} = \begin{bmatrix} \mathbf{0}_{m-1} & 1 \end{bmatrix}' \Gamma \mathbf{z_t} = D_m' \Gamma z_t \qquad (28) \\ \mathbf{y_t} &= H\theta_t + v_t, \qquad (29) \\ E\left(\mathbf{e}_t \mathbf{e}_t'\right) &= \left\{ \sigma^2 D_m D_m' = \tilde{Q} \right. \qquad (30) \\ E\left(v_t v_t'\right) &= \{R \qquad (31) \end{aligned}$$

where vectors are denoted with a bold and scalars with normal typeset: consistently with previous notation $\mathbf{y_t}$ is an $1 \times l$ vector, while $\mathbf{z}_t$ is an $m \times 1$ vector whose last element is z_t. $\mathbf{0}_{(m)}$ is a zero (column) vector of m-elements and $\gamma_{m,m-j}$ corresponds to the $m, m-j$ elements of the matrix Γ, D_m is a vector selecting the last row in the Γ matrix.

The main difference with the standard state-space models is given by equation (28) which maps the vector $\mathbf{z_t} \equiv [z_t, \ldots, z_{t-m}]$ into the scalar θ_t.[7] This can be considered as a 'bridge equation' that embodies all the information on the autocorrelation function of shocks ε_t and transforms them into the i.i.d. innovations e_t. Once again, if the ε_t are serially uncorrelated, then the generalized state space model reduces to the standard state space model.

We are now ready to write the Kalman filter equations for the state-space model defined in (27-31). Following previous notation and using the generalized state space

[6] since then $\Gamma\Gamma' = I$

[7] The filter by construction requires that we drop the first m observation in order to avoid any effect of the intimal condition at the beginning of the sample.

model, in particular equation (28), it is easy to construct prediction errors for the state θ_t as function of prediction errors of the GLS-transformed variable z_t and the weighting matrix Γ, as follows:

$$\theta_t - \theta_{t|t-1} = \left(z_t - z_{t|t-1}\right) + \sum_{j=0}^{m-1} \gamma_{m,m-j} \left(z_{t-j|t-j} - z_{t-j|t-j-1}\right) = D_m^{'} \Gamma \left(\mathbf{z_t} - \mathbf{z_{t|t-1}}\right), \tag{32}$$

where the vector of filtered states is defined as $\mathbf{z_{t|t-1}} \equiv [z_{t|t-1}, \ldots z_{t-m+1|t-m}]$ so that no smoothing process of the past states is involved. From the state prediction errors all moments which are relevant for the Kalman filter can easily be derived. For example the prediction error variance for θ is written as:

$$P^{\theta}_{t|t-1} = P^{z}_{t|t-1} + \sum_{j=0}^{m-1} \gamma^2_{m,m-j} P^{z}_{t-j-1|t-j-2} = D_m \Gamma^{'} P^{\mathbf{z}}_{t|t-1} \Gamma D^{\prime}_m, \tag{33}$$

which depends on matrix Γ and it is generally larger than the matrix $P_{t|t-1}$ defined in the standard Kalman filter (13).

The complete Kalman filtering equations can then be written as follows:

$$\begin{aligned}
\mathbf{z_{t|t-1}} &= \Psi \mathbf{z_{t-1|t-1}}, & (34)\\
P^{\mathbf{z}}_{t|t-1} &= \Psi P^{\mathbf{z}}_{t-1|t-1} \Psi^{'} + \tilde{Q}\,, & (35)\\
\theta_{t|t-1} &= D_m^{'} \Gamma \mathbf{z_{t|t-1}}, & (36)\\
P^{\theta}_{t|t-1} &= D_m \Gamma^{'} P^{\mathbf{z}}_{t|t-1} \Gamma D^{\prime}_m, & (37)\\
F_t &= E\left[\left(\mathbf{y_t} - H\theta_{t|t-1},\right)\left(\mathbf{y_t} - H\theta_{t|t-1}\right)^{\prime}\right] = H P^{\theta}_{t|t-1} H^{'} + R, \\
&= H D_m \Gamma^{'} P^{\mathbf{z}}_{t|t-1} \Gamma D^{\prime}_m H^{'} + R, & (38)\\
K^G_t &= P^{\mathbf{z}}_{t|t-1} H^{\prime}_t D^{'}_m \Gamma \left(H D_m \Gamma^{'} P^{\mathbf{z}}_{t|t-1} \Gamma D^{\prime}_m H^{'} + R\right)^{-1}, & (39)\\
\mathbf{z_{t|t}} &= \mathbf{z_{t|t-1}} + K_t \left(\mathbf{y_t} - H D_m^{'} \Gamma \mathbf{z_{t|t-1}}\right), & (40)\\
\theta_{t|t} - \theta_{t|t-1} &= = D_m^{'} \Gamma \left(\mathbf{z_{t|t}} - \mathbf{z_{t|t-1}}\right), & (41)\\
P^{\mathbf{z}}_{t|t} &= \left(\mathbf{I_m} - K_t H\right) P^{\mathbf{z}}_{t|t-1}. & (42)
\end{aligned}$$

The first two equations (34-35) correspond to the projection step applied to the transformed variable vector $\mathbf{z}$ rather than the original θ. The whole vector $\mathbf{z}$ is projected forward by using the matrix Ψ. Since the GLS transformation is a linear filter, coefficients in the Ψ matrix are the same as for the original untransformed state variable θ (e.g. the Φ), after taking into account lagged states as in our generalized state space model. The matrix $P^{\mathbf{z}}_{t|t-1}$ denotes the variance covariance matrix for the whole vector $\mathbf{z_{t|t-1}}$, while $P^{z}_{t|t-1}$ is its submatrix related only to $z_{t|t-1}$. If the $\{\varepsilon_t\}_{t=1}^{T}$ are uncorrelated then $\Gamma = I_m$ and $\theta_t = z_t$; we are back to the standard Kalman filter. The remaining main

difference with the standard filter lies in how the Kalman gain is defined. As equation (39) displays, our filter solves the following problem:

$$K_t^G \equiv \underset{K_t^G}{\arg\min}\, E\Bigg[\left(z_t - z_{t|t-1} - K_t^G\left(\mathbf{y}_t - HD_m^{'}\Gamma\mathbf{z_{t|t-1}}\right)\right) \left(z_t - z_{t|t-1} - K_t^G\left(\mathbf{y}_t - HD_m^{'}\Gamma\mathbf{z_{t|t-1}}\right)\right)\Bigg] \quad (43)$$

We use it to update the GLS transformed state, $z_{t|t-1}$, by using prediction errors defined on the original data vector $\mathbf{y}$. The Kalman gain K_t^G defined in (39) will be larger than its equivalent in the standard Kalman filter due to the larger $P_{t|t-1}^{\theta}$. This is because the generalized filter embodies all the information contained in the autocorrelation function of the state θ_t and consequently gives more weight (compared to the standard filter) to the observed variable when estimating the state variable at time t given the information at time $t-1$. On the other hand, the standard Kalman filter, by imposing a specified AR formulation to the state dynamics, would discard all the residual persistence and regard it as a noise component (e.g. measurement error) of the observed variable, as we discuss in section 4.

A problem of our generalized state space model is that in order to approximate persistent dynamics, a large number of lags m should be introduced in estimating the Ω^m matrix . A one-step estimation of such a parametric model by maximum likelihood is known to be likely to be troublesome. With this respect, the interested reader may want to check the summary of the maximum likelihood procedure (Sowell (1992)) given in Baillie (1996). In the next section we present an approach, borrowed from Abadir and Talmain (2005), which allows to estimate parsimoniously (e.g. few parameter estimates are needed) the autocorrelation structure of the residuals even assuming very long processes. This allows us to apply the Generalized State Space Model described in this section.

3.1 Estimation

In general the autocorrelation matrix Ω^m is unknown and it can be derived either from assumptions on the nature of the shock, for example if we assume it is a fractionally integrated process, we can estimate the parameter d, and infer from that the autocorrelation structure, or by approximating the autocorrelation function of the innovations by using a large number of parameters. The former way is not very convenient in a state space model,[8] while the latter is unfeasible in the case of highly persistent shocks since we would need a large number of parameters to be estimated by Maximum Likelihood.

Following Abadir and Talmain (2005), we take a different path and estimate $\rho_\varepsilon(k)$

[8]Furthermore since d is not a structural parameter, its estimation does not add any additional information from an economic point of view.

by fitting a functional form to the ACFs of the residual ε_t:

$$\rho_{\varepsilon}(k) \simeq \frac{1}{(1+a_1 k^{a_2})^{a_3}} \tag{44}$$

where a_1, a_2 and a_3 are parameters to be estimated. This functional form[9] was derived in Abadir and Talmain (2002) and corresponds to the decay rate of the ACF of a long memory process which includes the fractional integrated processes as a *special case.*

Using 44 and assuming that the innovations $\mathbf{e}_{t+1}$ and v_t are normally distributed, we can estimate the parameters of the generalized filter by maximizing the following likelihood function

$$\begin{aligned} f_{y_t|\theta_t,\Im_{t-1}}(\mathbf{y_t}|\theta_t,\Im_{t-1}) &= (2\pi)^{-\frac{1}{2}}\left|HP^{\theta}_{t|t-1}H^{'}+R\right|^{-\frac{1}{2}} \\ &\exp\Big\{-\frac{1}{2}\left(\mathbf{y_t}-H\theta_{t|t-1}\right)\left(HP^{\theta}_{t|t-1}H^{'}+R\right)^{-1}\left(\mathbf{y_t}-H\theta_{t|t-1}\right)^{'}\Big\} \\ &= (2\pi)^{-\frac{1}{2}}\left|HD\Gamma P^{\mathbf{z}}_{t|t-1}\Gamma^{'}D^{\prime}H^{'}+R\right|^{-\frac{1}{2}} \\ &\exp\Big\{-\frac{1}{2}\left(\mathbf{y_t}-HD^{'}\Gamma^{'}\mathbf{z_{t|t-1}}\right)\left(HD\Gamma P^{\mathbf{z}}_{t|t-1}\Gamma^{'}D^{\prime}H^{'}+R\right)^{-1} \\ &\left(\mathbf{y_t}-HD^{'}\Gamma^{'}\mathbf{z_{t|t-1}}\right)^{'}\Big\} \end{aligned} \tag{45}$$

where

$$\begin{aligned} \Gamma\Gamma^{\prime} &= \sigma^2\Omega^m; \\ \Omega^m &= \{\omega_{i,j}:\omega_{i,j}=\rho_{\varepsilon}(k),\ k=|i-j|,\ i,j=1,...,m\} \\ \rho_{\varepsilon}(k) &= \frac{1}{(1+a_1k^{a_2})^{a_3}} \end{aligned}$$

with respect to the parameters of the filter matrices and the parameters of the functional form, implicitly requiring the moment condition $E(e_t|z_t)=0$ to be satisfied.

This approach can be seen as a generalized method of moments estimation with a "GLS type" correction where Γ' is an optimal weighting matrix such that e_t and z_t are uncorrelated. In the next section, we show through simulation the ability of our approach to estimate the structural parameters of a simple DSGE model under the hypothesis that the DGP of the state is a fractional AR process.

[9]This functional form has been used in a number of papers. Abadir, Caggiano, and Talmain (2006) showed that it can capture well (and better than ARMA processes) the dynamic properties of the variables in the Nelson and Plosser dataset; Abadir and Talmain (2005) used it to construct a GLS approach to effectively deal with the uncovered interest rate puzzle; Moretti (2007) applied it in an Engle and Granger framework to test for cointegration between very persistent time series.

4 The artificial DSGE economy

In this section we evaluate the accuracy of our approach in estimating a DSGE model in which the Kalman filter reconstructs both the path of an exogenous state (e.g. the capital stock) and that of an endogenous state (e.g. the capital stock). We show that when the persistence of the data at hand is unknown and stronger than what assumed by the researcher, the performance of the standard Kalman filter in reconstructing of the unobserved endogenous states (e.g. the capital stock) deteriorates and parameter estimates can be very biased. Conversely, that our generalized filter is able to effectively cope with the issue at hand.

Our experiment is based on simulated data in line with the recent literature (see for an example Ruge-Murcia (2007)). The framework is as simple as possible, we choose the Ramsey model which is the core of many, if not all, DSGE/RBC models. We introduce the misspecification by simulating a Ramsey model where the technology is an autoregressive process of order one with an innovation that follows a fractional noise (the process is described in the appendix). A fractional noise allows us to easily control the degree of persistence by changing a single parameter, the order d of fractional integration; furthermore, it is also a good way to check that our method is able to approximate processes that, by nature, do not have a finite order state space representation.

A simple Ramsey model is sufficient to highlight one of the main points of the paper: unaccounted persistence in *exogenous* state processes distorts the path of the unobserved *endogenous* state variables and leads to substantial bias in the parameter estimates. The more persistent is the data generating process the larger is the root mean square error of the estimated capital stock. At the same time, the estimated deep parameters, which determine the reduced form parameters in the transition equation for capital accumulation, are more and more biased with respect to the true ones.

For each sample we first estimate the unobserved states and the deep parameters by the means of the usual Kalman filter, documenting the effects of dynamic misspecification. Then we apply our Generalized Kalman Filter to the same set of data and show that the estimation bias decreases significantly compared to the standard case. A further robustness check is conducted in subsection 7.1 where we investigate the extent to which long memory dynamics can be tackled by introducing a unit root in technology.[10]

Our set-up features a productivity shock (a_t) as unobserved exogenous state, capital (k_t) as unobserved endogenous state, consumption (c_t), output (y_t), the real interest rate (r_t) as measurements.[11] We simulate the Ramsey model to generate 1000 artificial data samples each one of 170 observations for consumption, output and the real interest rate. We repeat this procedure for each integer degree of fractional order d ranging from 0.1 to 0.9.

[10] In this case, in order to simulate the data, we use the same model used for real data estimation. We postpone the discussion accordingly. This also makes the presentation of the unit-root case more self contained.

[11] As there is a single structural shock it would also be possible to estimate the model using only one series, e.g. the GDP. Here we stick to the treatment of Ireland (2004), where one shock, technology, is used as the source of comovement of variables

In the following we provide first a short summary of the model we use and we explain how our method is applied to it. We show that our method is equivalent to building a filter on the data which removes the 'excess persistence', which is left unexplained by the model. Simulation results are then provided in subsection 4.1.

The model can be summarized as a standard (decentralized) market problem as follows. Households choose consumption C and investments I such as to maximize their objective function:

$$U_0 = E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \log \left(C_t \right) \right\}, \tag{46}$$

subject to a budget constraint:

$$C_t + I_t \leq W_t + R_t K_{t-1}, \tag{47}$$

where W_t is the real wage rate and R_t is the real interest rate on previous period capital stock K_{t-1}; since household face no leisure choice, we normalized labor to one.

Capital is set by the households with the standard law of motion, involving the capital depreciation parameter δ:

$$K_t = (1 - \delta) K_{t-1} + I_t,$$

where no growth trend in productivity is assumed. Firms use capital according to the production function:

$$Y_t = A_t K_{t-1}^{\alpha}, \tag{48}$$

where technology evolves as:

$$\log(A_t) = \rho \log(A_{t-1}) + \varepsilon_t, \tag{49}$$

where ε_t is considered as i.i.d. by the households with a constant variance $var(\varepsilon_t) = \sigma_{\varepsilon}^2$.

Using equation (47), the absence of profits, capital accumulation can be rewritten as:

$$K_t = (1 - \delta) K_{t-1} - C_t + W_t + R_t K_{t-1}, \tag{50}$$

The first order necessary condition associated with the maximization of the objective (46) subject to (50) is given by the standard Euler equation:

$$C_t^{-1} = E_t \left[\beta (1 + R_{t+1} - \delta) C_{t+1}^{-1} \right] \tag{51}$$

Firms rent capital K_t by paying a rental price R_t and maximize their profits by choosing capital such that the real interest rate equals the marginal productivity of capital minus the depreciation rate δ:

$$R_t + \delta = \alpha A_t K_{t-1}^{\alpha - 1} \tag{52}$$

Finally we have the goods market clearing condition

$$Y_t = C_t + I_t.$$

The competitive equilibrium for the economy is the sequence of prices $\{R_t, W_t\}_0^\infty$ and quantities $\{Y_t, K_t, C_t\}_0^\infty$ such that firms maximize profits, agents maximize utility and all markets clear. This amounts to estimating the system of equations given by the log-linear approximation of (48-52). When calibrating the model for the simulations deep parameters are set in a way which is consistent with the previous literature. The capital share α is set equal to 0.33; the preference term β is equal to 0.99 which corresponds to a real interest rate of 0.04 on annual basis; the depreciation rate δ is set equal to 0.025; the autoregressive term is set to 0.9. To gain on clarity σ_ε^2 is set equal to 1.[12] In order to remove any singularity in the system of equations we add two i.i.d. normal measurement errors with zero mean and standard deviation equal to 0.01 to consumption and the real interest rate. To keep the estimation process as clean as possible we do not introduce measurement errors as autocorrelated processes in order to ensure technology shock is the only dynamic factor in the model. The model is then log-linearized around its steady state [13] and solved under the assumption that innovations ε_t are i.i.d.

By denoting log-deviations from the steady state with lower case variables, we write the model solution in the following form:

$$a_{t+1} = \rho a_t + \varepsilon_{t+1}, \tag{53}$$

$$k_{t+1} = p_{kk} k_t + q_{ka} a_t, \tag{54}$$

$$\mathbf{y_t} = m_{yk} k_t + n_{ya} a_t, \tag{55}$$

For clarity we separate here exogenous and endogenous state variables, respectively

$$a_t \equiv \ln(A_t/A^{ss}), k_t \equiv \ln(K_t/K^{ss}),$$

and the vector $\mathbf{y_t}$ collects measurements:

$$\mathbf{y}_t = [\ln(C_t/C^{ss}),\, \ln(Y_t/Y^{ss}),\, \ln(R_t/R^{ss})].$$

The $\{p, q, m, n\}$ matrices describe the (linear) rational expectation solution of the model, the so called policy functions of the model. p and q are in this case scalars which denote how endogenous state variables respond respectively to lagged endogenous states and to exogenous states. Matrices m, n (3×1) denote how measurements react to the same set of variables.

It is straightforward to cast the model in a state space representation: define a vector of state variables, $\theta_\mathbf{t} \equiv [a_t, k_t]$, stack the equations (53-54) to form the transition equations of the state space. Add a set of measurement errors to (55) to specify the measurement equation. Finally, we can use the Kalman Filter innovations to estimate parameters.

When the unobserved innovations ε_t are serially correlated, equation (53) is not a good approximation of the exogenous state dynamics. Our Generalized State Space

[12] We checked that results are unchanged by setting more realistic values.

[13] The steady state values of the variables are denoted with ^{ss}

framework deals with this problem as follows. Append an equation to the state space model describing the residuals' autocorrelation:

$$\begin{aligned} a_{t+1} &= \rho a_t + \varepsilon_{t+1}, && (56)\\ k_{t+1} &= p_{kk} k_t + q_{ka} a_t, && (57)\\ \mathbf{y}_t &= m_{yk} k_t + n_{ya} a_t, && (58)\\ E[\varepsilon_t \varepsilon_{t+k}] &= \rho_\varepsilon(k),\ k = 1, \ldots, \infty && (59) \end{aligned}$$

where the autocorrelations ρ_ε, up to lag m, are estimated using the functional form in 44. Then, as in section 3, invert the cholesky of the variance matrix of ε_t and get a set of coefficients $\{l_{m,j}\}_{j=1}^{m-1}$. Define a new exogenous variable:

$$\tilde{a}_t = a_t - \sum_{j=1}^{m-1} l_{m,j} a_{t-j}, \tag{60}$$

where $\tilde{a}_t$ corresponds to the 'bridge variable' defined in section (3) and $\{l_{m,j}\}_{j=1}^{m-1}$ to the last row of the matrix L in the generalized state space model (25).[14] Once the bridge variable is described, one can form the generalized state space (27–31), use the formulas in (34–42) in section 3.1, to respectively build the filter and undertake the estimation. The economic interpretation of this approach is now given. First, notice that the 'GLS-transformed' variable $\tilde{a}_t$ is now by construction an AR(1) process. Then since the GLS transformation is a linear filter, it can be thought as applying to all variables through the exogenous shock $\tilde{a}_t$ resulting in the following model:

$$\begin{aligned} \tilde{a}_{t+1} &= \rho \tilde{a}_t + \eta_{t+1}, && (61)\\ \tilde{k}_{t+1} &= p_{kk} \tilde{k}_t + q_{ka} \tilde{a}_t, && (62)\\ \tilde{\mathbf{y}}_t &= m_{ka} \tilde{k}_t + n_{ya} \tilde{a}_t, && (63) \end{aligned}$$

Where $\eta_t \sim IID$ and variable $\tilde{x}$ is the GLS-filtered version x. The filtered capital stock $\tilde{k}_t$ and measurements $\tilde{\mathbf{y}}_t$ are defined as:[15]

$$\begin{aligned} \tilde{k}_t &= k_t - \sum_{j=1}^{m-1} l_{m,j} k_{t-j}, && (64)\\ \tilde{\mathbf{y}}_t &= \mathbf{y}_t - \sum_{j=1}^{m-1} l_{m,j} \mathbf{y}_{t-j}. && (65) \end{aligned}$$

Since the GLS-transformation is a linear filter, our approach is equivalent to applying the cross equation restrictions, defined in the matrices (p, q, m, n), to the set filtered

[14] Here $\tilde{a}_t$ and a_t play respectively the role of z_t and θ_t in the equation $z_t = \theta_t + \sum_{j=0}^{m-1} l_{m,m-j} \theta_{t-j-1}$.

[15] There is a little abuse of notation since $\mathbf{y_t}$ is a vector; in this case it is sufficient to stack the same filter in different columns, one for each variable of vector $\mathbf{y_t}$.

variables. In other words, our method can be seen as a way of filtering the data and applying model restrictions to filtered, rather than to raw, data. In this respect, the filter endogenously account for the part of the data not explained by the model and corrects variables consistently to that information.[16] The same type of logic, albeit within a different technical framework, is followed by Cogley and Sbordone (2008) where the cross equation restrictions (e.g. the New Phillips curve in their application) hold for an inflation gap obtained using a time varying trend inflation. Our framework is also related to the recent work of Canova and Ferroni (2009) where they show that different data filtering lead to different behaviour of the structural shocks of the model (e.g. in some cases there will be too much persistence left, in others some signs of overdifferencing). For these reasons, they propose the "naive" approach to combine in the measurement equation several filtered versions of the same variable (e.g. Beveridge Nelson filtered, HP and others) aggregated with estimated weights in order for residuals of the Kalman filter to be closer to i.i.d. Finally a way of dealing with persistence in DSGE models which shares some common points with ours can be found in Gorodnichenko and Ng (2009), who propose to treat the reduced form of the DSGE model with the filter defined by the quasi difference of exogenous states, e.g. $1 - \rho L$, where ρ is the persistence parameter of AR(1) process of the exogenous state and L is the lag operator.

4.1 Simulation results

We simulate 1000 artificial data for DSGE model described in the previous section. For each sample, we estimate the parameters and the unobserved state variables first with standard likelihood methods (see 19) and then with the generalized state space approach in sections 3-3.1. The functional form (44) is used to recover the autocorrelation matrix Ω^m of the residuals of technology, as explained in 3.1, then its Cholesky factorization to recover the weights $\{l_{m,j}\}_{j=1}^{m-1}$ contained in the matrix L defined above. We set the number of lags m equal to 30 since it seemed sufficient to successfully approximate fractional integration behaviour.

Table one reports simulation results for the standard filter. The second column shows the true model parameters, while in the following ones we report the sample means of each estimated parameter for different degree of persistence (e.g. an increasing d). The last two rows report our assessment, measured by the Root Mean Square Error, of how well the two filters are able to reconstruct the true unobserved state variables (the rows are denoted with $a_{t|t}$ and $k_{t|t}$, the RMSE refers to the error $a_t - a_{t|t}$ and $k_t - k_{t|t}$).[17] states example annual also reliable the Results show that, when the persistence of exogenous states is higher than what was assumed by the researcher, due to long memory, the reconstruction of the unobserved states will be poor. In the estimation process, deep parameters which determine the persistence of the transition

[16] As our procedure improves the fit of any model, a model evaluation criterion might be to assess how much of the raw data is explained by the model and how much is left to the filtering device. We leave this question to further research.

[17] For each simulation sample we estimate the structural parameters of the model

equation will tend to be biased in order for the reduced form of the model to try to match the persistence in the data.[18]

		Fractional Integration Parameter d								
	True	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
α	0.330	0.3478	0.3556	0.3645	0.3704	0.3702	0.3768	0.4109	0.4605	0.5483
δ	0.025	0.0266	0.0272	0.0279	0.0279	0.028	0.0285	0.0309	0.0343	0.0407
β	0.99	0.9875	0.988	0.9881	0.991	0.9901	0.9903	0.9902	0.9923	0.9954
ϕ	0.9	0.8903	0.8967	0.9004	0.9066	0.9067	0.9078	0.9126	0.9226	0.9388
σ_e	1	1.0404	1.0404	1.1358	1.2554	1.4744	1.703	1.8521	1.9403	1.9852
σ_{v_1}	0.01	0.1122	0.1384	0.0898	0.0698	0.0344	0.0258	0.0211	0.0211	0.0218
σ_{v_2}	0.01	0.1064	0.1314	0.1228	0.1289	0.1237	0.1359	0.2217	0.3609	0.6956
RMSE										
$a_t - a_{t\|t}$		1.0439	1.0967	1.1964	1.3862	1.7154	2.4849	5.6623	11.1167	19.6298
$k_t - k_{t\|t}$		0.3827	0.4714	0.5357	0.6491	0.8617	1.5441	5.3139	10.8163	19.0695

Table 1: Results for degree of fractional integration d: Standard Kalman filter

Our experiment delivers the following results. First, the bias in parameter estimates is relevant and it tends to increase as the persistence in the simulated data becomes stronger. For d larger than 0.5, the bias becomes quite substantial, especially for the capital share α, the standard deviation of the shock σ_e, e.g. the innovation on productivity, and the standard deviation of measurement errors. In particular, for large values of d the estimates of the capital share are similar to what found by Ireland (2004) on real data. The upward bias in the capital share is also related to the distortion in the capital dynamics; in fact, a value of α above its true value increases the return on capital and consequently it makes the capital more persistent in the reduced form. In our simulations this effect is partially compensated by an increase in the capital depreciation δ (which lowers the persistence of capital) but the resulting errors in the reconstruction of capital dynamics are still considerable. Further experiments showed that the increase in δ is mostly due to the inclusion of the real interest rate among the measurements; this variable is difficult to grasp in real data and it is almost never included in the dataset of a DSGE model.[19] Second, we find that the accuracy of the standard filter in reconstructing both the exogenous $a_{t|t}$ and the endogenous state $k_{t|t}$ is poor (RMSE shown in the last two rows of Table one) and it gets much worse as the parameter d increases. As both components are typically unobserved, this check is only possible by using simu-

[18]The actual pattern of deep parameter estimates is hard to tell in advance since it will depend on which measurements are used. If the researcher uses several measurements which are deeply correlated with the unobserved state one might expect a better estimate of the deep parameters in spite of the misspecification. We leave this point, to further research.

[19]When we estimate a similar model on U.S. data and we do not include a real rate measure among the observables, the estimate of the depreciation rate is much lower than what one should expect, see section 5.2 for discussion. We leave a more complete assessment of which variables should be selected in the measurement equation of a DSGE model to future research.

lated data, but it is important as, in general, the results of many empirical papers may depend on how unobserved states are reconstructed.

Our third result is that, as far as d increases, the variance of measurement errors also tends to increase: persistent dynamics is discarded as measurement error. One caveat holds for this result: a further investigation showed us that the pattern of measurement errors estimates depends by few very large estimates that bias upwards the mean of the simulation samples.[20] not depend on If we remove those samples, the estimates of all other parameters are almost unchanged, but we get an almost accurate estimate for σ_{v1} and a clearer pattern for σ_{v2} (e.g. its bias rises progressively as d increases): In Table 3

True $\sigma_{\mathbf{v2}}$	0.01
$d = 0.1$	0.0101
$d = 0.2$	0.0128
$d = 0.3$	0.0199
$d = 0.4$	0.0213
$d = 0.5$	0.0254
$d = 0.6$	0.0311
$d = 0.7$	0.1129
$d = 0.8$	0.2637
$d = 0.9$	0.5756

Table 2: Estimates for σ_{v2} after removing problematic samples

we report the results of the same exercise for the Generalized Filter.[21]

		Fractional Integration Parameter d								
	True	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
α	0.330	0.3302	0.3301	0.3301	0.3299	0.3299	0.3299	0.3309	0.3342	0.3411
δ	0.025	0.025	0.025	0.025	0.0249	0.025	0.02499	0.0251	0.0253	0.0258
β	0.99	0.99	0.99	0.99	0.99	0.99	0.9900	0.99	0.9901	0.9903
ϕ	0.9	0.90	0.90	0.90	0.90	0.9	0.899	0.9002	0.9007	0.902
σ_e	1	0.9964	1.0017	1.0106	1.01	1.0295	1.0437	1.0598	1.0836	1.1301
σ_{v_1}	0.01	0.0099	0.0099	0.0099	0.01	0.0099	0.0099	0.0099	0.0099	0.0099
σ_{v_2}	0.01	0.0099	0.0099	0.0099	0.01	0.0099	0.0100	0.0119	0.0249	0.0524
RMSE										
$a_t - a_{t\|t}$		1.0154	1.0172	1.0342	1.0654	1.1241	1.2155	1.3512	2.6863	5.3401
$k_t - k_{t\|t}$		0.2073	0.2542	0.3135	0.4014	0.5165	0.6899	0.9249	2.5774	5.1981

Table 3: Results for degree of fractional integration d: Generalized Kf

As it can be readily seen, the amount of bias provided by the generalized filter is almost negligible. This is true for all the degree of fractional integration, even when the

[20]The estimates were just large, they did not touch any of the boundaries imposed for the parameter estimation

[21]We choose a value of m, number of lags in productivity, equal to 30. This seems to be a reasonable compromise; while implementing a rather effective correction it does not exclude too many observations.

state variables become non-stationary ($d > 0.5$). The generalized filter also provides much more accurate prediction of the endogenous and exogenous state variable compared to the standard filter. In fact, the RMSE errors of the state predictions are up to 4 times smaller than those obtained with the standard filter.

This shows the ability of the modified filter to capture the dynamic properties of the data. In the last section we show how this accuracy in predicting the data dynamics also holds when undertaking an out-of-sample forecasting exercise.

5 Real data estimation

5.1 Model

In this section we take our model to the real data and repeat the Maximum Likelihood estimation for a RBC model as in Ireland (2004); the same type of model is also used by Ruge-Murcia (2007) in order to compare different estimation techniques for a different case of model misspecification. Households choose consumption and labor/leisure and save by investing in stocks of capital. Furthermore, they maximize the following utility function:

$$U_0 = E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \log(c_t) + \gamma(1 - n_t) \right\}, \tag{66}$$

subject to the budget constraint:

$$c_t + i_t \leq w_t n_t + r_t k_{t-1},$$

there is no population growth and the total amount of labor is normalized to one and leisure is given by $1 - n_t$. Capital accumulates with the standard law of motion:

$$\eta k_t = (1 - \delta) k_{t-1} + i_t, \tag{67}$$

expressed in efficiency units in order to take into account a log-linear trend in technology η. This gives rise to the (standard) first order conditions:

$$c_t^{-1} = E_t \left[\beta(\eta + r_{t+1} - \delta) \right] c_{t+1}^{-1}, \tag{68}$$

$$w_t = \gamma c_t \tag{69}$$

The production function is given by the Cobb-Douglas:

$$y_t = a_t (\eta^t n_t)^{1-\alpha} k_{t-1}^{\alpha}, \tag{70}$$

where technology evolves as:

$$\log(a_t) = (1 - \rho) \log(a^{ss}) + \rho \log(a_{t-1}) + \varepsilon_t. \tag{71}$$

In equilibrium the real interest rate equals the marginal productivity of capital minus the depreciation rate δ:

$$r_t + \delta = \alpha a_t n_t^{1-\alpha} k_{t-1}^{\alpha-1}, \tag{72}$$

and the real wage rate equals the marginal productivity of labor:

$$w_t = (1-\alpha) a_t n_t^{\alpha} k_{t-1}^{\alpha}. \tag{73}$$

The model is closed by the market clearing condition:

$$y_t = c_t + i_t. \tag{74}$$

The competitive equilibrium for the economy is the sequence of prices $\{r_t, w_t\}_0^\infty$ and quantities $\{y_t, k_t, c_t, n_t, i_t\}_0^\infty$ such that firms maximize profits, agents maximize utility and all markets clear. The model is linearized by using the Taylor expansion of the system of equations ((67)-(74) around the deterministic steady state of the model. Then we solve the model following Klein (2000) and rewrite the reduced form solution into the (generalized) state space representation.

As before, we define with $\mathbf{y_t}$ the vector of observable variables of the model: we use output, consumption and hours worked. $\theta_\mathbf{t}$ is the vector of unobservable states, both the capital stock and productivity. Following Ireland (2004), we introduce in the state space model three mutually independent autocorrelated measurement errors, $\eta_t = \left[\eta_t^y,\ \eta_t^c, \eta_t^h\right]'$ that are assumed to evolve as AR(1) processes

$$\begin{bmatrix} \eta_t^y \\ \eta_t^c \\ \eta_t^h \end{bmatrix} = \begin{bmatrix} \rho_{\eta^y} & 0 & 0 \\ 0 & \rho_{\eta^c} & 0 \\ 0 & 0 & \rho_{\eta^h} \end{bmatrix} \begin{bmatrix} \eta_{t-1}^y \\ \eta_{t-1}^c \\ \eta_{t-1}^h \end{bmatrix} + \begin{bmatrix} \zeta_t^y \\ \zeta_t^c \\ \zeta_t^h \end{bmatrix}$$

with

$$E\left[\zeta_t' \zeta_t\right] = \begin{bmatrix} \sigma_{\zeta^y}^2 & 0 & 0 \\ 0 & \sigma_{\zeta^c}^2 & 0 \\ 0 & 0 & \sigma_{\zeta^h}^2 \end{bmatrix}.$$

We have then

$$\mathbf{y}_t = [y_t, c_t, h_t]\ ; \theta_t = [k_t, a_t, \eta_t^y, \eta_t^c, \eta_t^h].$$

The following state space representation is then used to estimate the model:

$$\begin{aligned} \theta_{t+1} &= \Phi\theta_t + \varepsilon_{t+1} \\ \mathbf{y}_t &= H\theta_t, \end{aligned}$$

where

$$\Phi = \begin{bmatrix} p_{kk} & [q_{ka} & 0 & 0 & 0] \\ 0 & [& F & &] \end{bmatrix}; \; F = \begin{bmatrix} \rho & 0 & 0 & 0 \\ 0 & \rho_{\eta^y} & 0 & 0 \\ 0 & 0 & \rho_{\eta^c} & 0 \\ 0 & 0 & 0 & \rho_{\eta^h} \end{bmatrix}$$

$$H = \begin{bmatrix} m_{ck} & n_{ca} & 1 & 0 & 0 \\ m_{yk} & n_{ya} & 0 & 1 & 0 \\ m_{rk} & n_{ra} & 0 & 0 & 1 \end{bmatrix},$$

where the elements of the matrices Φ and H are obtained from the solution of the model, for example:

$$k_t = p_{kk} k_{t-1} + q_{ka} a_t,$$

$$c_t = m_{ck} k_{t-1} + n_{ca} a_t.$$

where the matrices (p, q, m, n) denote as before elements of the decision rules and in general x_{ij} denotes the effect of changes of variable j into the dynamics of variable i for $x \in (p, q, m, n)$.

As it is clear from the state space model, we focus only on one special treatment of measurement errors: three autocorrelated but mutually uncorrelated measurement errors. This is consistent with Ireland (2004) which shows that such specification, compared with one having correlated measurement errors, have the best out-of sample forecasting properties in spite of less plausible values of the deep parameters. As we show below, the trade off between forecasting accuracy and realistic parameter estimation vanishes once we allow for persistent dynamics in the data.

The measurements are hours worked, consumption and GDP, taken respectively from BLS data (Current Employment Statistics) and US NIPA national accounting: data run from 1948:1 to 2002:2.[22] We estimate the following set of structural parameters: $\psi \equiv [\alpha,\ \eta,\ \gamma,\ \delta,\ \rho,\ \rho_{\eta^y},\ \rho_{\eta^c},\ \rho_{\eta^h},\ \sigma_\varepsilon,\ \sigma_{\zeta^y},\ \sigma_{\zeta^c},\ \sigma_{\zeta^h}]$ plus the level of technology a^{ss} which enters the steady state expressions of the variables. The estimated parameters are the capital share, the log-linear trend of technology, the parameter which pins down the amount of hours worked in steady state, the depreciation rate of capital, the persistence parameter of the technology shock and the measurement errors together with their standard deviations. The discount preference term β is calibrated at the value of 0.99, as standard in the literature. Differently from Ireland, we also estimate the depreciation term δ it has important implication for the dynamics of the endogenous state k_t.

5.2 Parameter estimates

In this section we compare full-sample estimates of the deep parameters obtained with the Generalized and the Standard Kalman Filter. Table (4) reports the estimation results:

[22] We use the same dataset as in Ireland (2004) which is based on the 1996 chained data.

Parameters	α	δ	ρ_a	η	γ	σ_ε
Gen. Kalman Filter	0.2917	0.0246	0.9686	0.0055	0.0041	0.0048
Kalman Filter	0.5189	0.0031	0.9999	0.0057	0.0036	0.0092
Parameters	$\rho_{\eta_t^y}$	$\rho_{\eta_t^c}$	$\rho_{\eta_t^h}$	σ_{ζ^y}	σ_{ζ^c}	σ_{ζ^h}
Gen. Kalman Filter	0.998	-0.8041	0.9995	0.0039	0.0004	0.0062
Kalman Filter	0.9999	0.9994	0.9990	0.0036	0.0001	0.0060

Table 4: Generalized and standard Kalman filter parameter estimates for the U.S. economy.

Since the 'true' deep parameters are unknown, we can not directly asses the goodness of the estimates obtained with the standard and the generalized Kalman Filter. However, since some parameters have a clear structural interpretation it is possible to compare them with the evidence coming from either national accounts or previous microeconometric studies (see discussion in Prescott and McGrattan (2007). We start discussing the capital share α and the depreciation rate of capital δ: these estimates should be consistent with figures derived from national accounts data.[23] Table (4) shows that both the capital share and the depreciation rate estimated by the generalized filter are close to those obtained using national accounts data, while the standard filter falls short in providing good estimates. This latter finding is consistent with the results in Ireland (2004) and with the fact that in most of the literature those parameters are calibrated rather than estimated since it is believed that DSGE models fail in reproducing reasonable estimates for them. This failure can be attributed to the role of the capital share and the depreciation rate of capital in determining the persistence of the capital stock in the transition equation: larger values of the capital share and smaller depreciation rates imply higher returns on capital and lower depreciation in capital accumulation, this produces a more persistent capital dynamics. Since the capital persistence is transmitted to output and consumption through the other equations of the RBC model, such unrealistic parameter estimates might be seen as the consequence of the Kalman Filter trying to replicate the persistence of the data. As the second row of Table (4) shows, once we control for persistent data the RBC model is able to convey estimates which are in line with national accounts data.

With regards to the other parameters, estimation results differ substantially between the two filters.[24] There is a substantial difference in the values of both the persistence and variance parameters of the technology shock. The productivity trend η is also slightly larger than in the standard case. The measurement error on consumption is negatively autocorrelated, a result found also by Ireland (2004) when allowing mea-

[23]The capital share can be derived from the labor share, computed as the average share of value added which is paid to labor (around 0.3 for US data), while the depreciation rate should square with the average ratio between investments and the capital stock (around 0.025 assuming quarterly data)

[24] We found similar results when we calibrate $\delta = 0.025$.

surement errors to be cross-correlated.[25] An important remark concerns the high estimated value of the persistence of exogenous states (measurement errors and the technology shock) delivered by the Generalized Filter which might sound at odds with the fact that our device controls for persistence. As a robustness check, we re-estimated the model using i.i.d. measurement errors with both the Standard and the Generalized Filter: the persistence of the technology shock estimated by the Standard filter is $\rho = 0.98$ while the Generalized Filter delivers a much lower $\rho = 0.85$. This suggests that the high values of persistence shown in table 1 are mostly an artifact of the introduction of the autocorrelated measurement errors, rather than a problem of our method not being able to capture persistence. of parameter is filter.

To complete the section, we provide a warning against inferring too quickly (e.g. without looking carefully also at what innovation residuals have to say) that a unit root process is the correct description of exogenous states when a Standard Kalman Filter delivers an high value of persistence for the exogenous states, further discussion of the unit-root case is in section 7. In particular, consider the estimate delivered by the Standard Filter ($\rho = 0.9999$), shown in the second row of table 1, this value makes the process of technology as follows

$$a_t = 0.9999 a_{t-1} + \varepsilon_t,$$

which is undistinguishable from a unit root process, since $a_t - 0.9999 a_{t-1} \simeq \Delta a_t$. Now, if a unit root process describes correctly the dynamics of the technology, then we should expect the innovations $\{\varepsilon_{t=1}^T\}$ to be i.i.d.. However, if we estimate with a Local Whittle Estimator (Robinson(1995)) the fractional integration order d of the innovations $\{\varepsilon_{t=1}^T\}$, as reconstructed by the Kalman Filter, we get a significant negative value around -0.3, which hints to series being over-differenced.[26]

Reconstructed residuals $\{\varepsilon_{t=1}^T\}$	Estimated d
Exact LW estimator	-0.30

We can get to the same conclusion by looking at the dynamics of the log detrended GDP. If we assume that in our RBC model the output y_t has the same statistical nature as the technology a_t (this is consistent with Christiano and Vigfusson (2003)), we can gauge the order of fractional integration of a_t directly from that of (log-detrended) GDP. The Local Whittle Estimator for this series is equal to 0.7:

[25] Overall, it seems that the generalized filter produces estimates which are in line with estimates produced by models with less restrictions on the variance-covariance matrices of the exogenous shocks, such as the one with cross-correlated shocks in Ireland (2004).

[26] For the innovations from our generalized filter we could not reject the hypothesis that d is equal to zero.

Detrended Log-GDP	Estimated d
LW estimator	0.69
Exact LW estimator	0.71
Feasible ELW estimator	0.71
Feasible ELW estimator with detrending	0.71
2-step feasible ELW estimator	0.71
Feasible ELW estimator w/o detrending	0.71

Straightforward computations show that if $a_t \sim I(d = 0.7)$, meaning that $a_t(1 - L)^{0.7} = e_t$, where $e_t \ \ i.i.d.$, it follows that:

$$a_t - 0.9999a_{t-1} \simeq \Delta a_t = (1 - L)(1 - L)^{-0.7}e_t,$$

which, consistently with what already shown, tells us that the first difference of productivity is overdifferenced by an order 0.3:

$$\Delta a_t(1 - L)^{-0.3} = e_t, \text{that is: } \Delta a_t \sim I(d = -0.3).$$

Following Baillie (1996), long memory processes can be successfully employed to describe time series which appear as non stationary in their levels but they look overdifferenced when their first difference is considered.

6 Forecasting

In order to assess the ability of our approach to capture the dynamic properties of the data we compare the out of sample forecast of the Generalized filter with those of the standard Kalman filter. The rationale behind this exercise follows the results in Granger and Joyeux (1980) who showed that while a short order AR representation can adequately fit long memory dynamics in sample, the forecasts produced by such AR models will not be very accurate.

The exercise is implemented as follows. We estimate the RBC model described in the previous section with both the standard and our approach for a subsample of data, precisely from 1948:1 until 1987:4. We generate out-of-sample forecasts one through four quarters ahead for each variable and compare the root-mean-squared forecast errors from the modified model to those from the standard Kalman filter. We then extend the subsample by one period and repeat the estimation and forecasting. We continue this way until all the sample is covered in 2002.

Table 4 reports the RMSE together for the forecasts generated by the standard Kalman filter (KF) and those generated by the generalized filter (GKF). In order to asses whether any difference of the two RMSE is significant we also report the statistic proposed by Diebold and Mariano (1995).[27]

[27]The critical values for the Diebold and Mariano test have been obtained using a bootstrap procedure.

Steps ahead	1	2	3	4
Output				
RMSE: GKF	0.6492	1.2294	1.8137	2.3585
RMSE: KF	0.984	1.8935	2.881	3.7871
D-M test	-6.8582***	-5.9845***	-4.8295***	-3.7659***
Consumption				
RMSE: GKF	0.4588	0.7022	1.0155	1.3689
RMSE: KF	0.5443	0.961	1.4172	1.9213
D-M test	-3.6916***	-3.7701***	-3.0075***	-2.1948**
Hours				
RMSE: GKF	0.6352	1.1373	1.6295	2.0847
RMSE: KF	0.8135	1.4583	2.0558	2.6461
D-M test	-4.8812***	-3.4673***	-2.2781**	-1.8388*

Table 5: Root mean square error of the forecasts from one to four quarters ahead for the standard and Generalized Kalman filter. *** is significant at 1%, ** is significant at 5% and * is significant at 10%

The results indicate that forecasts from the modified filter significantly outperform those from the standard Kalman filter. In particular, for output, the RMSE of the generalized filter are up to 60% smaller than those of the standard filter. This shows the better performance of the Generalized Filter with respect to the normal one[28]. On one hand, the forecast from the generalized filter are more accurate than those from standard filter for all the steps ahead; on the other, the forecast accuracy of the modified filter improves, relatively to the standard filter, as the forecast horizon arises. This result indicates the presence of very persistent dynamics in the data that can not be fully captured by the autoregressive structure of the standard filter. On the other hand, it also shows that by accounting for such persistence it is possible to considerably improve the dynamic properties of the model.

7 Comparison with unit root case

Part of the literature specifies technology as a stochastic rather than a deterministic trend process. As a further check, we estimate a model where we assume that the level of technology follows a unit root process. We show that under this assumption we still get unrealistic parameter estimates as well as less accurate forecasts compared to those shown in section 6.

In the case of a unit root in technology and considering log-variables as before, the

[28] Similar, albeit less striking, results hold for the case when we calibrate the δ.

state space model should be written as:

$$\begin{aligned}
\hat{k} &= k_t - a_t && (75)\\
\hat{y}_t &= y_t - a_t && (76)\\
\hat{c}_t &= c_t - a_t && (77)\\
\hat{h}_t &= h_t && (78)\\
\Delta a_{t+1} &= \rho \Delta a_t + \varepsilon_{t+1}, && (79)\\
\hat{k}_{t+1} &= p_{kk}\hat{k}_t + q_{ka}\Delta a_t, && (80)\\
\hat{\mathbf{y}}_t &= m_{yk}\hat{k}_t + n_{ya}\Delta a_t, && (81)\\
a_t &= \mu + a_{t-1} + \varepsilon_{t,} && (82)
\end{aligned}$$

where the vector $\hat{\mathbf{y}}_t$ is now defined in detrended terms in order ensure stationarity: $\hat{\mathbf{y}}_t \equiv [\hat{y}_t, \hat{c}_t, \hat{h}_t]$. In the unit-root set-up hours worked are a strictly stationary variable ($I(0)$), driven by the growth rate of technology; in the trend stationarity approach they follow the level of technology which is allowed, by our approach (GKF), to be a long memory process.
The state space model (79-81) has now a non-stationary state variable. As common in the literature, we cope with this problem by implementing the Exact Kalman Filter as in Durbin and Koopman (2001).[29]

The model with three autocorrelated measurement errors does not seem to be the best specification under the unit root hypothesis as we encountered severe identification problems. For this reason, we resorted to the model with i.i.d. measurement errors whose parameter estimates are shown in Table [30]

Parameter estimates, reported below for the full sample, show a reverse pattern in comparison to what we obtained with the Standard Kalman Filter under the deterministic trend assumption. The capital share is too low and the depreciation rate too high compared to what we would expect from national accounts data. In the same line as we argued before, these can be interpreted as signs of having introduced 'too much persistence' in the data generating process of the model.

Concerning the out-of-sample forecasts, they seem fairly accurate for the cointegrated variables, e.g. consumption and output, even if much worse than in those from the deterministic trend model (from 2 to 4 percent is the RMSE of output, around 1 per cent for consumption). On the other hand, the root mean square errors for hours worked are ten times larger compared to the previous case. Consistently with our parameter estimates hours worked appear to be mainly 'driven' by measurement errors.

The inability of tracking hours worked with a unit root technology shock is not a novelty in the literature. Most studies, as (King, Plosser, Stock, and Watson (1991)),

[29] In particular we use their 'univariate version' where the updating step is repeated for each variable by adding one measurement at the time and the information set is expanded accordingly. This set-up is quite convenient in the case when the dataset comprise some variables that are cointegrated (output and consumption) and others that are not (hours)

[30] We also performed out-of sample forecasts with the generalized filter for the model with no autocorrelated measurement errors. Conclusions were not overturned.

Parameters	α	δ	ρ_a	η	γ	σ_ε
Gen. Kalman Filter	0.2917	0.0246	0.9686	0.0055	0.0041	0.0048
Unit Root Filter	0.1871	0.1073	0.7317	0.0050	0.0047	0.0050
Parameters	$\rho_{\eta_t^y}$	$\rho_{\eta_t^c}$	$\rho_{\eta_t^h}$	σ_{ζ^y}	σ_{ζ^c}	σ_{ζ^h}
Gen. Kalman Filter	0.998	-0.8041	0.9995	0.0039	0.0004	0.0062
Unit Root Filter	-	-	-	0.0197	0.0001	0.0540

Table 6: Full Sample parameter estimates: Generalized Kalman Filter against Unit Root.

Steps ahead	1	2	3	4
Output				
RMSE: Unit Root Model	2.4087	3.1181	3.6301	4.0673
Consumption				
RMSE: Unit Root Model	1.0172	1.0201	1.0406	1.3298
Hours				
RMSE: Unit Root Model	9.6450	10.0057	10.1787	10.2798

Table 7: Root mean square error of the forecasts from one to four quarters ahead for the unit root model.

based on models that specified technology as a permanent component, found a very limited explanatory power of the model on variables which are modelled as stationary, e.g. hours worked. The same result can be found in very recent studies (for example the New Area Wide Model) where the permanent component of technology has a small contribution in explaining non cointegrated variables variations, in particular for hours worked. This is probably related to the strong restriction that a unit root technology shock imposes on the data generating process: the level of technology represent a common trend for the cointegrated variables, while its growth rate movements in the stationary variables (such as hours worked). When technology is the only shock present in the model the unit root assumption implies a very tight relation between the variance of the cycle and that of the trend, which might results as being too restrictive. A further remark concerns the statistical nature of hours worked. Empirical studies show that hours worked have non-clear cut order of integration, e.g. they can be integrated of order one, zero or possibly be a long memory process, see Gil-Alana and Moreno (2006). The Generalized Kalman Filter procedure accounts for the possibility of long memory behaviour, since it filters hours in the same way as the other variables, allowing in this way for a 'common long range component' between GDP and hours worked.

From the results in this section, we can conclude that a long memory specification of technology, which lies between a strict stationarity and a pure unit root case, might be the preferred one. Nevertheless, we are aware that results can also be model-

dependent; as discussed by Chang, Doh, and Schorfheide (2007), the order of integration of the shocks of a DSGE strongly depends upon the structure of the model at hand. Further inquiry on the topic is left to future research.

7.1 Comparison with unit roots: simulated data

In this section we provide a further experiment using simulated data. We simulate data from the stationary model described by Ireland (2004), the same as in section 5 and we add a long memory component as described in section 4. We then use the simulated dataset to estimate a model with a unit root in technology such as (75-82). This complements the analysis of section 4. We do not use the same simple Ramsey model, as we got serious identification problems with it.[31] To run this experiment it is also crucial that the signal-to-noise ratio (e.g. the standard deviation of the technology shock relative to those of measurement errors) is set at values which are plausible with respect to those estimated on real data, when the signal-to-noise ratio is too high the optimizer simply fails (e.g. lack of identification) to estimate the unit-root model. To sum-up, the model used to simulate data is the one described by Ireland (2004), adding three i.i.d. measurement errors and the model used for estimation is its unit root version described by (75-82). The calibration of parameters to be estimated is as follows:[32]

$$\alpha = 0.33,\ \rho_e = 0.99,\ \delta = 0.025,\ \sigma_e = 0.002,\ \sigma_i = 0.001,$$

where α is the capital share in production ($1-\alpha$ is the labor share), δ is the depreciation rate, ρ_e, σ_e denote the persistence and the standard deviation of the growth rate of technology, while $\sigma_i, i \in \{y, c, h\}$ is the standard deviation of the three i.i.d. measurement errors, on output, consumption and hours worked. Results are means from 200 samples;[33] to save on space we only report the cases $d = 0, 0.1, 0.4, 0.6, 0.9$. The $d = 0$ case is relevant as in this case the estimated model is somehow misspecified with respect to the true Data Generating Process since we impose a unit root on stationary data.

The first worthy remark is that even when no long memory is introduced in the DGP, the unit root model produces substantial bias in the estimates.[34] This appears to be rather counterintuitive, as it occurs even if the persistence of the exogenous parameter is set at a very high value, close to a unit root; dramatic results, not shown to save space, appear also for a lower $\rho = 0.95$. After some trials we found that, as highlighted in the previous section, one serious problem seems to be the presence of a

[31] Results were cross-checked by using Dynare 4.0.2

[32] No substantially different results were found when the β was included in the estimated parameters. Also, in this simple example data are considered in deviation from their steady state, then we do not estimate a drift η nor the parameter which pins down the steady state level of hours worked γ.

[33] We eliminate those very few samples in which the optimizer got stuck at some bound; the sampling distribution looks quite symmetric and no important difference can be found from using median or means

[34] A potential critique is that the direction of bias shown here is not entirely consistent with what we have found in the previous section on real data. While this should deserve further enquiry in the future, we do not claim that a simple stationary RBC with a fractional noise in the innovations is the True Data Generating Process for US data, so we do not see this potential criticism as really critical

Parameter	d = 0	d =0.1	d=0.4	d=0.6	d=0.9
α	0.4021	0.4012	0.4220	0.4719	0.5941
ρ_e	0.0425	-0.0198	0.1527	0.5439	0.8444
δ	0.0589	0.0351	0.0359	0.0332	0.0295
σ_c	0.0002	0.0002	0.0004	0.0012	0.0094
σ_y	0.0029	0.0027	0.0022	0.0013	0.0011
σ_h	0.0027	0.0029	0.0060	0.0159	0.0419
σ_e	0.0016	0.0020	0.0035	0.0053	0.0090

Table 8: Deep parameter estimates : data generating process is stationary (plus long memory), estimating model embeds a unit root in technology

stationary time series among the measurements, hours worked. The true DGP delivers a relatively high amount of persistence in hours due to the presence of the level of technology, $a_t = 0.99a_{t-1} + \varepsilon_t$, in their reduced form equation; in the unit root model hours worked are instead a function of the first difference of technology, whose law of motion is given by $\Delta a_t = \rho_e \Delta a_{t-1} + \epsilon_t^u$. Given the true exogenous state process, a correct estimate should be $\rho_e = 0.99 - 1 = -0.01$ but such a low value would make persistence in hours worked very low in the unit-root model. When a long memory component is added to the innovations of the DGP this makes hours even more persistent and, as d goes high and it tends to exacerbate the problem. It is not surprising then that estimates of ρ_e are upward biased and so are those for the other parameters, especially the capital share, making capital to be more persistent. When removing hours worked from the set of measurements, estimates improve (not shown) but they still suffer from some upward bias. A plausible, though non exhaustive, explanation for this pattern is that the unit root model delivers deviations of cointegrated variables from the technology level (their trend) which are more limited and short-lived than the DGP. To see that, using the reduced form of the unit-root model we can write the deviation of the level of capital from its trend as:

$$k_t - a_t(1 - p_{kk}^u L) = q_{aa}^u \Delta a_t,$$

where p^u, q^u are the solution of the unit-root model and L is the lag operator. Under the true DGP Δa_t should almost be a white noise, as ρ_e is very low. Plugging the true values of the remaining deep parameters, we see that the resulting p^u, q^u give a more short lived variation of the term $k_t - a_t$ with respect to the DGP (this is also consistent with the findings of Rotenberg and Woodford, see below).

Overall, our results from this and the previous section suggest that even when the data generating process has an exogenous process which is very close a unit root ($\rho = 0.99$) a model which has unit-root in technology behaves fundamentally in a different way from a stationary model; furthermore it is impaired in reproducing the behaviour of the stationary time series. This is consistent with most of the finding in previous literature, coming from disparate sources: Rotemberg and Woodford (1996) find that

an RBC with unit roots cannot explain the variability of the predictable component of the business cycle. Other researchers such as Canova (2009) found that the variance decomposition of a shocks changes dramatically when either a unit root in technology is introduced or data are linearly detrended. This is related to our results, in the sense that when the unit-root is introduced it will be less able to capture cyclical fluctuations and it will make the other shocks look more important in the variance decomposition.

8 Conclusion

In this paper we have shown a simple method to bring general equilibrium dynamic models to persistent data. This can be done by allowing one (or more) common factor to have a long memory behaviour. By doing this one can improve the plausibility of estimates and the out-of-sample forecast performance with respect to both the pure stationary and the pure unit root case: we report an average (significant) reduction in the forecast error of about 30%.

Our method does not alter the cross equations restrictions implied by the DSGE model, but it rather applies them to observations which have been detrended. Contrary to other studies, the detrending procedure is not assumed a priori but it is directly derived from the data in a way which is consistent with the structure of the model, e.g. by exploiting information which the model have already deemed as not useful and put into innovation residuals. Our method also differs from other methods proposed for example by Canova (2009), where an hybrid model (DSGE plus filter) is estimated.

In simulation experiments we have shown that persistent data can have important effects in the way the Kalman filter reconstructs unobserved endogenous states. We also provide evidence that the introduction of a stochastic trend as a unit-root imposes is a constraint the data generating process which can come at some cost and should not be seen as a general way to cope with persistence. We leave the full development of this line of research to further work.

Our contribution is not limited to DSGE models but it can be used any time long memory data and latent component are present; moreover, while here we apply it to the case of a simple univariate unobserved shock, it can be readily applied to a multivariate framework. Simulation results are provided for the case of ARFIMA shocks, the method is also able to cope with general long memory processes, since it does not cast any strong assumption concerning the nature of the long memory dynamics, whether it is a fractionally integrated process or a different long memory process.

As already mentioned in the introduction, there is a widespread evidence about long memory behaviour of many macroeconomic time series, for example inflation; this would be a call for monetary and New Keynesian Models. Since these are mostly estimated by Bayesian methods we leave it to further research to extend our framework to that case.

Bibliography

ABADIR, K. M., G. CAGGIANO, AND G. TALMAIN (2006): "Nelson and Plosser revised: the ACF approach," University of Glasgow Working Papers 7.

ABADIR, K. M., AND G. TALMAIN (2002): "Aggregation, persistence and volatility in a macromodel," *Review of Economic Studies*, 69, 749–779.

——— (2005): "Distilling co-movements from persistence macro and financial series," ECB Working Papers 525.

ALTISSIMO, F., B. MOJON, AND P. ZAFFARONI (2009): "Fast micro and slow macro: can aggregation explain the persistence of inflation?," *Journal of Monetary Economics*, 56(2), 231–241.

BAILLIE, R. (1996): "Long Memory Processes and fractional integration in econometrics," *Journal of Econometrics*, 73, 5–59.

CANOVA, F. (2009): "Bridging Cyclical DSGE Models and the Raw Data," Unpublished.

CANOVA, F., AND F. FERRONI (2009): "Multiple filtering devices for the estimation of cyclical DSGE models," Economics Working Papers 1135, Universitat Pompeu Fabra.

CHANG, Y., T. DOH, AND F. SCHORFHEIDE (2007): "Non-stationary Hours in a DSGE Model," *Journal of Money, Credit and Banking*, 39(6), 1357–1373.

CHARI, V. V., P. KEHOE, AND E. MCGRATTAN (2008): "New Keynesian Models: Not Yet Useful for Policy Analysis," Staff Report 409 Minneapolis FED.

CHRISTIANO, L. J., AND R. J. VIGFUSSON (2003): "Maximum likelihood in the frequency domain: the importance of time-to-plan," *Journal of Monetary Economics*, 50(4), 789–815.

COGLEY, T., AND J. NASON (1992): "Output Dynamics in Real Business Cycle Models," *American Economic Review*, JUNE.

COGLEY, T., AND A. SBORDONE (2008): "Trend Inflation, Indexation and Inflation Persistence in the New Keynesian Phillips Curve," *American Economic Review*, 98(6), 2101–2126.

DIEBOLD, F. X., AND R. MARIANO (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13(3), 253–263.

DIEBOLD, F. X., AND G. D. RUDEBUSCH (1989): "Long Memory and Persistence in Aggregate Output," *Journal of Monetary Economics*, 24(2), 189–209.

DURBIN, J., AND S. J. KOOPMAN (2001): *Time series analysis by state space methods*. Oxford University Press.

GADEA, M., AND L. MAYORAL (2006): "The Persistence of Inflation in OECD Countries: A Fractionally Integrated Approach," *International Journal of Central Banking*, 2(1).

GIL-ALANA, L. A., AND A. MORENO (2006): "Technology Shocks and Hours Worked: A Fractional Integration Perspective," Working papers Universitad de Navarra 03/06.

GORODNICHENKO, Y., AND S. NG (2009): "Estimaton of DSGE Models When the Data are Persistent," Columbia University.

GRANGER, C. W. J. (1966): "The typical spectral shape of an economic variable," *Econometrica*, 34(1), 150–161.

——— (1980): "Long memory relationship and the aggregation of dynamic models," *Journal of Econometrics*, 14, 227–238.

——— (2001): "On Model Approximation for Long-Memory Processes: A Cautionary Result," *Annals of Economics and Finance*, 2, 97–100.

GRANGER, C. W. J., AND R. JOYEUX (1980): "An introduction to long-memory time series models and fractional differencing," *Journal of Time Series Analysis 1*, pp. 15–30.

HAUBRICH, J. G., AND A. W. LO (2001): "The sources and nature of long-term memory in aggregate output," *Economic Review, Federal Reserve Bank of Cleveland*, Q(II), 15–30.

IRELAND, P. (2004): "A method for taking models to the data," *Journal of Economic Dynamics and Control*, 28(6), 1205–1226.

KING, R. G., C. I. PLOSSER, J. STOCK, AND M. WATSON (1991): "Stochastic Trends and Economic Fluctuations," *American Economic Review*, 81(4), 819–40.

KLEIN, P. (2000): "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model," *Journal of Economic Dynamics and Control*, 24(10), 1405–1423.

KYDLAND, F. E., AND E. PRESCOTT (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50(6), 1345–1370.

MAYORAL, L. (2005): "Further evidence on the statistical properties of Real GNP," Economics Working Papers 955, Universitat Pompeu Fabra.

MCGRATTAN, E. (2006): "Measurement with Minimal Theory," Meeting Papers 338, Society for Economic Dynamics.

MORETTI, G. (2007): "Detecting long memory co-movements in macroeconomic time series," Bank of Italy working paper 642.

PRESCOTT, E. C., AND E. R. MCGRATTAN (2007): "Technical Appendix: Unmeasured Investment and the Puzzling U.S. Boom in the 1990s," Staff Report 395 Minneapolis FED.

RAVENNA, F. (2007): "Vector Autoregressions and Reduced Form Representations of DSGE models," *Journal of monetary economics*, 54(7).

ROBINSON, P. M. (2003): *Time Series with long memory*. Oxford University Press.

ROTEMBERG, J. J., AND M. WOODFORD (1996): "Real-Business-Cycle Models and the Forecastable Movements in Output, Hours and Consumption," *The American Economic Review*, 86(1), 71–89.

RUGE-MURCIA, F. J. (2007): "Methods to estimate dynamic stochastic general equilibrium models," *Journal of Economic Dynamics and Control*, 31(8), 2599–2636.

SOWELL, F. (1992): "Modeling long-run behavior with the fractional ARIMA model," *Journal of Monetary Economics*, 29, 277–302.

9 Appendix: Long memory processes

Long memory processes have been extensively studied in time series analysis and a good review of their properties can be found in Robinson (2003) and Baillie (1996). Long-memory processes are defined by autocorrelations decaying slowly to zero or spectral density displaying a pole at zero frequency. For instance, compared to ARMA processes, they are have a autocorrelation function that decays hyperbolically rather than exponentially. In this respect, long memory processes represent an alternative representation to the knife-edge distinction between I(0) and I(1) processes. A well known class of long memory processes, introduced by Granger (1980), is the Autoregressive Fractional Moving Average Process of order d (ARFIMA(p,d,m)), which is defined as

$$(1-L)^d \phi(L) y_t = B(L)\varepsilon_t, \tag{83}$$

where ε_t is a white noise process with variance σ^2, $\phi(L)$ is a lag polynomial of order p, $B(L)$ is a lag polynomial of order m and $d \in [0,1]$ indicates the fractional differencing parameter. As the parameter d can take any value in the interval between 0 and 1, ARFIMA processes fill the gap between a strictly stationary process, when $d = 0$, and a unit root process, when $d = 1$.

A special case of (83), which is the basic building block of fractionally integrated processes is the so called 'fractional white noise', defined by setting $p, m = 0$ in (83). For $0 < d < 0.5$ the process is "stationary" with hyperbolic rather than exponential decay of the autocorrelation function (ACF heretofore); when $0.5 < d < 1$ the process is non stationary, i.e. the squared sums of its autocorrelations do not converge. However, differently from unit root processes, it is still mean reverting. Finally, when d is negative the ACF presents an oscillatory behaviour around zero, which is the typical pattern we can observe in 'overdifferenced' time series.

Going more into the technical details, the term $(1-L)^d$ in 83 is called the fractional differencing polynomial: the term 'fractionally integrated processes' denotes the family

of processes generated by the application of a fractional differencing polynomial. For $d < 0.5$ the fractional differencing polynomial $(1-L)^d$ can be represented as in Granger (1980):

$$\begin{aligned}(1-L)^d &= \sum_{j=0}^{\infty} \frac{\Gamma(j-d)(-1)^j}{\Gamma(j+1)\Gamma(-d)} L^j \simeq \sum_{j=0}^{\infty} j^{-(d+1)} \frac{(-1)^j}{\Gamma(-d)} L^j \\ &\equiv \sum_{j=0}^{\infty} \pi_j L^j\end{aligned}$$

where $\Gamma(\cdot)$ is the Gamma function and the approximating term derives from a first order asymptotic expansion. Therefore, any stationary fractional noise y_t can be rewritten in an infinite order autoregressive representation (AR(∞)),

$$(1-L)^d y_t = \varepsilon_t \tag{84}$$

$$\sum_{j=0}^{\infty} \pi_j L^j y_t = \varepsilon_t \tag{85}$$

$$y_t = \sum_{j=1}^{\infty} \pi_j y_{t-j} + \varepsilon_t, \tag{86}$$

As Granger (2001) and Granger and Joyeux (1980) point out, even if it possible by using standard identification criteria, to fit an ARIMA model on long memory generated data, this would lead to a poor approximation of the true DGP and inaccurate out-of-sample forecasts. In particular, Granger and Joyeux (1980) show that short order autoregressive processes are not flexible enough to successfully approximate long memory behaviour and that statistical criteria have a very hard time in selecting processes with enough lags.

RECENTLY PUBLISHED "TEMI" (*)

N. 726 – *Low skilled immigration and the expansion of private schools*, by Davide Dottori and I-Ling Shen (October 2009).

N. 727 – *Sorting, reputation and entry in a market for experts*, by Enrico Sette (October 2009).

N. 728 – *Ricardian selection*, by Andrea Finicelli, Patrizio Pagano and Massimo Sbracia (October 2009).

N. 729 – *Trade-revealed TFP*, by Andrea Finicelli, Patrizio Pagano and Massimo Sbracia (October 2009).

N. 730 – *The riskiness of corporate bonds*, by Marco Taboga (October 2009).

N. 731 – *The interbank market after august 2007: what has changed and why?*, by Paolo Angelini, Andrea Nobili and Maria Cristina Picillo (October 2009).

N. 732 – *Tax morale and public spending inefficiency,* by Guglielmo Barone and Sauro Mocetti (November 2009).

N. 733 – *Dynamic macroeconomic effects of public capital: evidence from regional Italian data*, by Valter Di Giacinto, Giacinto Micucci and Pasqualino Montanaro (November 2009).

N. 734 – *Networks with decreasing returns to linking*, by Filippo Vergara Caffarelli (November 2009).

N. 735 – *Mutual guarantee institutions and small business finance*, by Francesco Columba, Leonardo Gambacorta and Paolo Emilio Mistrulli (November 2009).

N. 736 – *Sacrifice ratio or welfare gain ratio? Disinflation in a DSGE monetary model*, by Guido Ascari and Tiziano Ropele (January 2010).

N. 737 – *The pro-competitive effect of imports from China: an analysis of firm-level price data*, by Matteo Bugamelli, Silvia Fabiani and Enrico Sette (January 2010).

N. 738 – *External trade and monetary policy in a currency area*, by Martina Cecioni (January 2010).

N. 739 – *The use of survey weights in regression analysis*, by Ivan Faiella (January 2010).

N. 740 – *Credit and banking in a DSGE model of the euro area*, by Andrea Gerali, Stefano Neri, Luca Sessa and Federico Maria Signoretti (January 2010).

N. 741 – *Why do (or did?) banks securitize their loans? Evidence from Italy*, by Massimiliano Affinito and Edoardo Tagliaferri (January 2010).

N. 742 – *Outsourcing versus integration at home or abroad*, by Stefano Federico (February 2010).

N. 743 – *The effect of the Uruguay round on the intensive and extensive margins of trade*, by Ines Buono and Guy Lalanne (February 2010).

N. 744 – *Trade, technical progress and the environment: the role of a unilateral green tax on consumption*, by Daniela Marconi (February 2010).

N. 745 – *Too many lawyers? Litigation in Italian civil courts*, by Amanda Carmignani and Silvia Giacomelli (February 2010).

N. 746 – *On vector autoregressive modeling in space and time*, by Valter Di Giacinto (February 2010).

(*) Requests for copies should be sent to:
Banca d'Italia – Servizio Studi di struttura economica e finanziaria – Divisione Biblioteca e Archivio storico – Via Nazionale, 91 – 00184 Rome – (fax 0039 06 47922059). They are available on the Internet www.bancaditalia.it.

2007

S. SIVIERO and D. TERLIZZESE, *Macroeconomic forecasting: Debunking a few old wives' tales*, Journal of Business Cycle Measurement and Analysis , v. 3, 3, pp. 287-316, **TD No. 395 (February 2001).**

S. MAGRI, *Italian households' debt: The participation to the debt market and the size of the loan*, Empirical Economics, v. 33, 3, pp. 401-426, **TD No. 454 (October 2002).**

L. CASOLARO. and G. GOBBI, *Information technology and productivity changes in the banking industry*, Economic Notes, Vol. 36, 1, pp. 43-76, **TD No. 489 (March 2004).**

G. FERRERO, *Monetary policy, learning and the speed of convergence*, Journal of Economic Dynamics and Control, v. 31, 9, pp. 3006-3041, **TD No. 499 (June 2004).**

M. PAIELLA, *Does wealth affect consumption? Evidence for Italy*, Journal of Macroeconomics, Vol. 29, 1, pp. 189-205, **TD No. 510 (July 2004).**

F. LIPPI. and S. NERI, *Information variables for monetary policy in a small structural model of the euro area*, Journal of Monetary Economics, Vol. 54, 4, pp. 1256-1270, **TD No. 511 (July 2004).**

A. ANZUINI and A. LEVY, *Monetary policy shocks in the new EU members: A VAR approach*, Applied Economics, Vol. 39, 9, pp. 1147-1161, **TD No. 514 (July 2004).**

D. JR. MARCHETTI and F. Nucci, *Pricing behavior and the response of hours to productivity shocks*, Journal of Money Credit and Banking, v. 39, 7, pp. 1587-1611, **TD No. 524 (December 2004).**

R. BRONZINI, *FDI Inflows, agglomeration and host country firms' size: Evidence from Italy*, Regional Studies, Vol. 41, 7, pp. 963-978, **TD No. 526 (December 2004).**

L. MONTEFORTE, *Aggregation bias in macro models: Does it matter for the euro area?*, Economic Modelling, 24, pp. 236-261, **TD No. 534 (December 2004).**

A. NOBILI, *Assessing the predictive power of financial spreads in the euro area: does parameters instability matter?*, Empirical Economics, Vol. 31, 1, pp. 177-195, **TD No. 544 (February 2005).**

A. DALMAZZO and G. DE BLASIO, *Production and consumption externalities of human capital: An empirical study for Italy*, Journal of Population Economics, Vol. 20, 2, pp. 359-382, **TD No. 554 (June 2005).**

M. BUGAMELLI and R. TEDESCHI, *Le strategie di prezzo delle imprese esportatrici italiane*, Politica Economica, v. 23, 3, pp. 321-350, **TD No. 563 (November 2005).**

L. GAMBACORTA and S. IANNOTTI, *Are there asymmetries in the response of bank interest rates to monetary shocks?*, Applied Economics, v. 39, 19, pp. 2503-2517, **TD No. 566 (November 2005).**

P. ANGELINI and F. LIPPI, *Did prices really soar after the euro cash changeover? Evidence from ATM withdrawals*, International Journal of Central Banking, Vol. 3, 4, pp. 1-22, **TD No. 581 (March 2006).**

A. LOCARNO, *Imperfect knowledge, adaptive learning and the bias against activist monetary policies*, International Journal of Central Banking, v. 3, 3, pp. 47-85, **TD No. 590 (May 2006).**

F. LOTTI and J. MARCUCCI, *Revisiting the empirical evidence on firms' money demand*, Journal of Economics and Business, Vol. 59, 1, pp. 51-73, **TD No. 595 (May 2006).**

P. CIPOLLONE and A. ROSOLIA, *Social interactions in high school: Lessons from an earthquake*, American Economic Review, Vol. 97, 3, pp. 948-965, **TD No. 596 (September 2006).**

L. DEDOLA and S. NERI, *What does a technology shock do? A VAR analysis with model-based sign restrictions*, Journal of Monetary Economics, Vol. 54, 2, pp. 512-549, **TD No. 607 (December 2006).**

F. VERGARA CAFFARELLI, *Merge and compete: strategic incentives for vertical integration*, Rivista di politica economica, v. 97, 9-10, serie 3, pp. 203-243, **TD No. 608 (December 2006).**

A. BRANDOLINI, *Measurement of income distribution in supranational entities: The case of the European Union*, in S. P. Jenkins e J. Micklewright (eds.), Inequality and Poverty Re-examined, Oxford, Oxford University Press, **TD No. 623 (April 2007).**

M. PAIELLA, *The foregone gains of incomplete portfolios*, Review of Financial Studies, Vol. 20, 5, pp. 1623-1646, **TD No. 625 (April 2007).**

K. BEHRENS, A. R. LAMORGESE, G.I.P. OTTAVIANO and T. TABUCHI, *Changes in transport and non transport costs: local vs. global impacts in a spatial network*, Regional Science and Urban Economics, Vol. 37, 6, pp. 625-648, **TD No. 628 (April 2007).**

M. BUGAMELLI, *Prezzi delle esportazioni, qualità dei prodotti e caratteristiche di impresa: analisi su un campione di imprese italiane*, v. 34, 3, pp. 71-103, Economia e Politica Industriale, **TD No. 634 (June 2007).**

G. ASCARI and T. ROPELE, *Optimal monetary policy under low trend inflation*, Journal of Monetary Economics, v. 54, 8, pp. 2568-2583, **TD No. 647 (November 2007).**

R. GIORDANO, S. MOMIGLIANO, S. NERI and R. PEROTTI, *The Effects of Fiscal Policy in Italy: Evidence from a VAR Model*, European Journal of Political Economy, Vol. 23, 3, pp. 707-733, **TD No. 656 (January 2008).**

B. ROFFIA and A. ZAGHINI, *Excess money growth and inflation dynamics*, International Finance, v. 10, 3, pp. 241-280, **TD No. 657 (January 2008).**

G. BARBIERI, P. CIPOLLONE and P. SESTITO, *Labour market for teachers: demographic characteristics and allocative mechanisms*, Giornale degli economisti e annali di economia, v. 66, 3, pp. 335-373, **TD No. 672 (June 2008).**

E. BREDA, R. CAPPARIELLO and R. ZIZZA, *Vertical specialisation in Europe: evidence from the import content of exports*, Rivista di politica economica, v. 97, 3, pp. 189, **TD No. 682 (August 2008).**

2008

P. ANGELINI, *Liquidity and announcement effects in the euro area*, Giornale degli Economisti e Annali di Economia, v. 67, 1, pp. 1-20, **TD No. 451 (October 2002).**

P. ANGELINI, P. DEL GIOVANE, S. SIVIERO and D. TERLIZZESE, *Monetary policy in a monetary union: What role for regional information?*, International Journal of Central Banking, v. 4, 3, pp. 1-28, **TD No. 457 (December 2002).**

F. SCHIVARDI and R. TORRINI, *Identifying the effects of firing restrictions through size-contingent Differences in regulation*, Labour Economics, v. 15, 3, pp. 482-511, **TD No. 504 (June 2004).**

L. GUISO and M. PAIELLA,, *Risk aversion, wealth and background risk*, Journal of the European Economic Association, v. 6, 6, pp. 1109-1150, **TD No. 483 (September 2003).**

C. BIANCOTTI, G. D'ALESSIO and A. NERI, *Measurement errors in the Bank of Italy's survey of household income and wealth*, Review of Income and Wealth, v. 54, 3, pp. 466-493, **TD No. 520 (October 2004).**

S. MOMIGLIANO, J. HENRY and P. HERNÁNDEZ DE COS, *The impact of government budget on prices: Evidence from macroeconometric models*, Journal of Policy Modelling, v. 30, 1, pp. 123-143 **TD No. 523 (October 2004).**

L. GAMBACORTA, *How do banks set interest rates?*, European Economic Review, v. 52, 5, pp. 792-819, **TD No. 542 (February 2005).**

P. ANGELINI and A. GENERALE, *On the evolution of firm size distributions*, American Economic Review, v. 98, 1, pp. 426-438, **TD No. 549 (June 2005).**

R. FELICI and M. PAGNINI, *Distance, bank heterogeneity and entry in local banking markets*, The Journal of Industrial Economics, v. 56, 3, pp. 500-534, **No. 557 (June 2005).**

S. DI ADDARIO and E. PATACCHINI, *Wages and the city. Evidence from Italy*, Labour Economics, v.15, 5, pp. 1040-1061, **TD No. 570 (January 2006).**

M. PERICOLI and M. TABOGA, *Canonical term-structure models with observable factors and the dynamics of bond risk premia*, Journal of Money, Credit and Banking, v. 40, 7, pp. 1471-88, **TD No. 580 (February 2006).**

E. VIVIANO, *Entry regulations and labour market outcomes. Evidence from the Italian retail trade sector*, Labour Economics, v. 15, 6, pp. 1200-1222, **TD No. 594 (May 2006).**

S. FEDERICO and G. A. MINERVA, *Outward FDI and local employment growth in Italy*, Review of World Economics, Journal of Money, Credit and Banking, v. 144, 2, pp. 295-324, **TD No. 613 (February 2007).**

F. BUSETTI and A. HARVEY, *Testing for trend*, Econometric Theory, v. 24, 1, pp. 72-87, **TD No. 614 (February 2007).**

V. CESTARI, P. DEL GIOVANE and C. ROSSI-ARNAUD, *Memory for prices and the Euro cash changeover: an analysis for cinema prices in Italy*, In P. Del Giovane e R. Sabbatini (eds.), The Euro Inflation and Consumers' Perceptions. Lessons from Italy, Berlin-Heidelberg, Springer, **TD No. 619 (February 2007).**

B. H. HALL, F. LOTTI and J. MAIRESSE, *Employment, innovation and productivity: evidence from Italian manufacturing microdata*, Industrial and Corporate Change, v. 17, 4, pp. 813-839, **TD No. 622 (April 2007).**

J. SOUSA and A. ZAGHINI, *Monetary policy shocks in the Euro Area and global liquidity spillovers*, International Journal of Finance and Economics, v.13, 3, pp. 205-218, **TD No. 629 (June 2007).**

M. DEL GATTO, GIANMARCO I. P. OTTAVIANO and M. PAGNINI, *Openness to trade and industry cost dispersion: Evidence from a panel of Italian firms*, Journal of Regional Science, v. 48, 1, pp. 97-129, **TD No. 635 (June 2007).**

P. DEL GIOVANE, S. FABIANI and R. SABBATINI, *What's behind "inflation perceptions"? A survey-based analysis of Italian consumers*, in P. Del Giovane e R. Sabbatini (eds.), *The Euro Inflation and Consumers' Perceptions. Lessons from Italy*, Berlin-Heidelberg, Springer, **TD No. 655 (January 2008).**

R. BRONZINI, G. DE BLASIO, G. PELLEGRINI and A. SCOGNAMIGLIO, *La valutazione del credito d'imposta per gli investimenti*, Rivista di politica economica, v. 98, 4, pp. 79-112, **TD No. 661 (April 2008).**

B. BORTOLOTTI, and P. PINOTTI, *Delayed privatization*, Public Choice, v. 136, 3-4, pp. 331-351, **TD No. 663 (April 2008).**

R. BONCI and F. COLUMBA, *Monetary policy effects: New evidence from the Italian flow of funds*, Applied Economics , v. 40, 21, pp. 2803-2818, **TD No. 678 (June 2008).**

M. CUCCULELLI, and G. MICUCCI, *Family Succession and firm performance: evidence from Italian family firms*, Journal of Corporate Finance, v. 14, 1, pp. 17-31, **TD No. 680 (June 2008).**

A. SILVESTRINI and D. VEREDAS, *Temporal aggregation of univariate and multivariate time series models: a survey*, Journal of Economic Surveys, v. 22, 3, pp. 458-497, **TD No. 685 (August 2008).**

2009

F. PANETTA, F. SCHIVARDI and M. SHUM, *Do mergers improve information? Evidence from the loan market*, Journal of Money, Credit, and Banking, v. 41, 4, pp. 673-709, **TD No. 521 (October 2004).**

M. BUGAMELLI and F. PATERNÒ, *Do workers' remittances reduce the probability of current account reversals?*, World Development, v. 37, 12, pp. 1821-1838, **TD No. 573 (January 2006).**

P. PAGANO and M. PISANI, *Risk-adjusted forecasts of oil prices*, The B.E. Journal of Macroeconomics, v. 9, 1, Article 24, **TD No. 585 (March 2006).**

M. PERICOLI and M. SBRACIA, *The CAPM and the risk appetite index: theoretical differences, empirical similarities, and implementation problems*, International Finance, v. 12, 2, pp. 123-150, **TD No. 586 (March 2006).**

U. ALBERTAZZI and L. GAMBACORTA, *Bank profitability and the business cycle*, Journal of Financial Stability, v. 5, 4, pp. 393-409, **TD No. 601 (September 2006).**

S. MAGRI, *The financing of small innovative firms: the Italian case*, Economics of Innovation and New Technology, v. 18, 2, pp. 181-204, **TD No. 640 (September 2007).**

S. MAGRI, *The financing of small entrepreneurs in Italy*, Annals of Finance, v. 5, 3-4, pp. 397-419, **TD No. 640 (September 2007).**

V. DI GIACINTO and G. MICUCCI, *The producer service sector in Italy: long-term growth and its local determinants*, Spatial Economic Analysis, Vol. 4, No. 4, pp. 391-425, **TD No. 643 (September 2007).**

F. LORENZO, L. MONTEFORTE and L. SESSA, *The general equilibrium effects of fiscal policy: estimates for the euro area*, Journal of Public Economics, v. 93, 3-4, pp. 559-585, **TD No. 652 (November 2007).**

R. GOLINELLI and S. MOMIGLIANO, *The Cyclical Reaction of Fiscal Policies in the Euro Area. A Critical Survey of Empirical Research*, Fiscal Studies, v. 30, 1, pp. 39-72, **TD No. 654 (January 2008).**

P. DEL GIOVANE, S. FABIANI and R. SABBATINI, *What's behind "Inflation Perceptions"? A survey-based analysis of Italian consumers*, Giornale degli Economisti e Annali di Economia, v. 68, 1, pp. 25-52, **TD No. 655 (January 2008).**

F. MACCHERONI, M. MARINACCI, A. RUSTICHINI and M. TABOGA, *Portfolio selection with monotone mean-variance preferences*, Mathematical Finance, v. 19, 3, pp. 487-521, **TD No. 664 (April 2008).**

M. AFFINITO and M. PIAZZA, *What are borders made of? An analysis of barriers to European banking integration*, in P. Alessandrini, M. Fratianni and A. Zazzaro (eds.): The Changing Geography of Banking and Finance, Dordrecht Heidelberg London New York, Springer, **TD No. 666 (April 2008).**

G. FERRERO and A. NOBILI, *Futures contract rates as monetary policy forecasts*, International Journal of Central Banking, v. 5, 2, pp. 109-145, **TD No. 681 (June 2008).**

L. ARCIERO, C. BIANCOTTI, L. D'AURIZIO and C. IMPENNA, *Exploring agent-based methods for the analysis of payment systems: A crisis model for StarLogo TNG*, Journal of Artificial Societies and Social Simulation, v. 12, 1, **TD No. 686 (August 2008).**

A. CALZA and A. ZAGHINI, *Nonlinearities in the dynamics of the euro area demand for M1*, Macroeconomic Dynamics, v. 13, 1, pp. 1-19, **TD No. 690 (September 2008).**

L. FRANCESCO and A. SECCHI, *Technological change and the households' demand for currency*, Journal of Monetary Economics, v. 56, 2, pp. 222-230, **TD No. 697 (December 2008).**

G. ASCARI and T. ROPELE, *Trend inflation, taylor principle, and indeterminacy*, Journal of Money, Credit and Banking, v. 41, 8, pp. 1557-1584, **TD No. 708 (May 2007).**

S. COLAROSSI and A. ZAGHINI, *Gradualism, transparency and the improved operational framework: a look at overnight volatility transmission*, International Finance, v. 12, 2, pp. 151-170, **TD No. 710 (May 2009).**

M. BUGAMELLI, F. SCHIVARDI and R. ZIZZA, *The euro and firm restructuring*, in A. Alesina e F. Giavazzi (eds): Europe and the Euro, Chicago, University of Chicago Press, **TD No. 716 (June 2009).**

B. HALL, F. LOTTI and J. MAIRESSE, *Innovation and productivity in SMEs: empirical evidence for Italy*, Small Business Economics, v. 33, 1, pp. 13-33, **TD No. 718 (June 2009).**

2010

S. MAGRI, *Debt maturity choice of nonpublic Italian firms* , Journal of Money, Credit, and Banking, v.42, 2-3, pp. 443-463, **TD No. 574 (January 2006).**

FORTHCOMING

L. MONTEFORTE and S. SIVIERO, *The Economic Consequences of Euro Area Modelling Shortcuts*, Applied Economics, **TD No. 458 (December 2002).**

M. BUGAMELLI and A. ROSOLIA, *Produttività e concorrenza estera*, Rivista di politica economica, **TD No. 578 (February 2006).**

G. DE BLASIO and G. NUZZO, *Historical traditions of civicness and local economic development*, Journal of Regional Science, **TD No. 591 (May 2006).**

R. BRONZINI and P. PISELLI, *Determinants of long-run regional productivity with geographical spillovers: the role of R&D, human capital and public infrastructure*, Regional Science and Urban Economics, **TD No. 597 (September 2006).**

E. IOSSA and G. PALUMBO, *Over-optimism and lender liability in the consumer credit market*, Oxford Economic Papers, **TD No. 598 (September 2006).**

S. NERI and A. NOBILI, *The transmission of US monetary policy to the euro area*, International Finance, **TD No. 606 (December 2006).**

G. FERRERO, A. NOBILI and P. PASSIGLIA, *Assessing Excess Liquidity in the Euro Area: The Role of Sectoral Distribution of Money*, Applied Economics, **TD No. 627 (April 2007).**

A. CIARLONE, P. PISELLI and G. TREBESCHI, *Emerging Markets' Spreads and Global Financial Conditions*, Journal of International Financial Markets, Institutions & Money, **TD No. 637 (June 2007).**

Y. ALTUNBAS, L. GAMBACORTA and D. MARQUÉS, *Securitisation and the bank lending channel*, European Economic Review, **TD No. 653 (November 2007).**

M. IACOVIELLO and S. NERI, *Housing market spillovers: evidence from an estimated DSGE model*, American Economic Journal: Macroeconomics, **TD No. 659 (January 2008).**

F. BALASSONE, F. MAURA and S. ZOTTERI, *Cyclical asymmetry in fiscal variables in the EU*, Empirica, **TD No. 671 (June 2008).**

M. BUGAMELLI and F. PATERNÒ, *Output growth volatility and remittances*, Economica, **TD No. 673 (June 2008).**

F. D'AMURI, O. GIANMARCO I.P. and P. GIOVANNI, *The labor market impact of immigration on the western german labor market in the 1990s*, European Economic Review, **TD No. 687 (August 2008).**

A. ACCETTURO, *Agglomeration and growth: the effects of commuting costs*, Papers in Regional Science, **TD No. 688 (September 2008).**

L. FORNI, A. GERALI and M. PISANI, *Macroeconomic effects of greater competition in the service sector: the case of Italy*, Macroeconomic Dynamics, **TD No. 706 (March 2009).**

Y. ALTUNBAS, L. GAMBACORTA, and D. MARQUÉS-IBÁÑEZ, *Bank risk and monetary policy*, Journal of Financial Stability, **TD No. 712 (May 2009).**

V. DI GIACINTO, G. MICUCCI and P. MONTANARO, *Dynamic macroeconomic effects of public capital: evidence from regional Italian data*, Giornale degli economisti e annali di economia, **TD No. 733 (November 2009).**

F. COLUMBA, L. GAMBACORTA and P. E. MISTRULLI, *Mutual Guarantee institutions and small business finance*, Journal of Financial Stability, **TD No. 735 (November 2009).**

L. FORNI, A. GERALI and M. PISANI, *The macroeconomics of fiscal consolidations in euro area countries*, Journal of Economic Dynamics and Control, **TD No. 747 (March 2010).**