Euro area inflation persistence in an estimated nonlinear DSGE model*

Gianni Amisano† Università di Brescia

Oreste Tristani‡ European Central Bank

This draft: 18 August 2006

Abstract

We estimate the approximate nonlinear solution of a small DSGE model using Bayesian methods. Our results, based on euro area data, suggest that this approach delivers sharper inference compared to the estimation of the linearised solution. The nonlinear model can also account for richer economic dynamics. We document how the impulse responses of inflation to structural shocks vary depending on initial conditions. The differences are statistically significant if we compare the early eighties with more recent years. This time-varying pattern is captured by the nonlinear features of the model, not by the hypothesis of structural breaks in the inflation mean.

JEL classification:
Keywords: DSGE models, policy rules, inflation persistence, second order approximations, particle filter, Bayesian estimation.

---

*The opinions expressed are personal and should not be attributed to the European Central Bank. We wish to thank Gianni Lombardo for useful discussions.

†Address: Università di Brescia, Facoltà di Economia, Via S. Faustino 74/B, I - 25122 Brescia, Italy. E-mail: amisano@eco.unibs.it.

‡Address: European Central Bank, DG Research, Kaiserstrasse 29, D - 60311, Frankfurt, Germany. E-mail: oreste.tristani@ecb.int.

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models have become popular tools for monetary policy analysis. The central feature of these models, emphasised in the theoretical work of Yun (1996) and Woodford (2003), is the presence of nominal rigidities in the adjustment of goods prices. More recently, a number of additional frictions have been introduced in the basic sticky-price framework and the resulting models have been successfully taken to the data (see e.g. Christiano, Eichenbaum and Evans, 2005; Smets and Wouters, 2003).

In all cases, however, what is estimated is only the reduced form emerging from the solution of a linearised version of those models. This approach has obvious advantages in terms of simplicity and possibility of comparison to other well-known empirical tools, such as VARs. There are a number of reasons, however, to also be interested in exploring the implications of the many *nonlinear* features built in DSGE models. Amongst these, three such reasons can be mentioned here.

The first one is that they are more suited to characterise macroeconomic dynamics in presence of large deviations from the steady state. Even in relatively short samples of 20 or 25 years, such as the post-1979 sample typically explored in empirical monetary policy analyses after the work of Clarida, Galí and Gertler (1998, 2000), euro area inflation has reached a maximum and minimum of 12.0 and 0.6 percent, respectively, compared to an average of 4.2 percent. By construction, a linearised model is not suited to explain such large deviations and it might deliver distorted estimates, at least in principle, if forced to do so.

To provide a more concrete example, it is conceivable that inflation persistence should depend on its distance from the steady state. Small deviations could be characterised by a relatively small degree of persistence, but the persistence could become more pronounced in case of larger deviations, which would more easibly become entrenched in expectations. This economic feature could be captured by higher-order terms in a nonlinear solution, terms which, by construction, would start playing a nonnegligible role only when large deviations from steady state do take place. A linearised model, on the contrary, would be forced to account for all observed dynamics with linear terms. It could therefore deliver estimates of inflation persistence that are biased upwards for small deviations and downwards for large deviations from the steady state. We test this conjecture explicitly in our estimations.

The second reason to be interested in nonlinear models is empirical, since nonlinear models are likely to provide sharper estimates than their linearised counterparts. More specifically, the nonlinearities of a model can be seen as additional testable implications compared to those characterising the linearised version of the same model. A straighforward example can be made for the case of solutions obtained through second-order perturbation methods. These approximate solutions imply that the variance of exogenous shocks have an impact on the unconditional means of observable variables. This link amounts to a restriction on the size of those variances, which is ignored in linearised solutions.

A more general empirical advantage of the estimation of nonlinear models is highlighted by Fernandez-Villaverde, Rubio-Ramirez and Santos (2006). The paper shows that the approximation errors of a model are compounded over the sample when constructing the likelihood function. As a result, a model solved up to a second-order approximation is the minimum requirement to obtain a first order approximation to the true likelihood function of the nonlinear model.

The third reason to be interested in nonlinear models is related to the possibility of exploring the asset price implications of macroeconomic models. The notions of risk and risk aversion only assume concrete meaning in nonlinear models. Conversely, linear models deliver by construction equal returns on all assets regardles of their riskiness. The availability of nonlinear solutions is therefore a necessary condition to analyse satisfactorily the implications of any macroeconomic model for the term structure of interest rates and equity prices. These implications can be of interest from an asset pricing perspective, but also from a purely macroeconomic perspective. Exploring the asset price implications of macroeconomic models can in fact be a powerful test of those models, which highlights their weaknesses and can act as a catalist for further improvements.

The main objective of this paper is to explore the empirical implications of a small DSGE model of the euro area solved using second-order perturbation methods. While other authors have provided results based on second or higher order approximation methods using simulated data (e.g. An and Schorfeide, 2005; Fernandez-Villaverde and Rubio-Ramirez, 2004), ours is, to our knowledge, the first paper which takes the second-order approximation of a DSGE model directly to the data.

Our results show that the nonlinear version of DSGE models provides sharper inference and accounts for richer economic dynamics compared to linearised versions. Concerning the dynamic features of euro area inflation, we show, using two different model-specifications, that notable differences can be found between linear and nonlinear solutions.

More specifically, the persistence in the inflation response to shocks is shown to be dependent on initial conditions in a model which does not allow for the possibility of shifts in the mean of inflation. For example, a positive inflation target shock has much larger and more pronounced effects if it occurs in a high inflation environment, compared to an environment where price stability is maintained. No significant dependence on initial conditions is observed instead in a model allowing for smooth shifts in the mean of inflation. We compare formally the two model specifications in terms of their marginal likelihood and find that the model without shifts in the mean is overwhelmingly superior to its alternative.

The rest of the paper is organised as follows. Section 2 provides a broad description of the two models that we estimate in the empirical section. The main difference between those models concerns the behaviour of monetary policy. While always following a Taylor-type rule, the central bank is assumed to have a stationary stochastic inflation target in the first case and an integrated target in the second case. Section 3 discusses briefly the solution method. It is well-known that approximate nonlinear solutions can be computed using a variety of methods (see Aruoba, Fernandez-Villaverde and Rubio-Ramirez, 2006). We focus on second-order perturbation methods, because they are direct extensions of standard linearisations and because they are fast to implement. The estimation methodology is presented next, in Section 4, with particular emphasis on the construction of the likelihood function, which is performed using the particle filter. We also discuss briefly some of the choices available to the researcher in this context and the importance of a plausible specification of the priors for the variance of the shocks. Section 5 presents the estimation results. More specifically, we discuss the differences between estimates obtained through linear and nonlinear versions of each model, in terms of both posterior densities for the parameters and conditional and unconditional moments. Section 6 concludes.

2 The theoretical framework

One of the conclusions of the "Inflation Persistence Network" (IPN) coordinated by the European Central Bank is that the estimated persistence of aggregate euro area inflation changes depending on one key hypothesis (see e.g. Angeloni *et al.*, 2005), namely whether shifts in the mean of inflation are, or not, accounted for. Empirical estimates of inflation persistence are high in the first case, while they fall considerably in the second. For example, Bilke (2004) argues that a structural break in French CPI inflation occurred in the mid-eighties. Controlling for this break, both aggregate and sectoral inflation persistence are stable and low. Levin and Piger (2004) also find strong evidence for a break in the mean of inflation in the late 1980s or early 1990s for twelve industrial countries. Allowing for such break, the inflation measures generally exhibit relatively low inflation persistence. Similar results are obtained by Corvoisier and Mojon (2004) for most OECD countries. Dossche and Everaert (2004) find similar results when they allow for shifts in the inflation target in the form of a random walk.

By and large, the existence of shifts in the mean of inflation has been tested within statistical or reduced-form frameworks (see e.g. Levin and Piger, 2004; Corvoisier and Mojon, 2004). As a result, it could be argued that there are two difficulties with the interpretation of these results. First, it remains unclear whether the hypothesis of one or more shifts in the mean of inflation would be rejected within a richer model. Secondly, the reasons for a potential shift in the inflation mean are left unspecified, while it would be interesting to understand their determinants.

To shed further light on these issues, we explore the empirical plausibility of two variants of a simple DSGE model of inflation and output dynamics. The first one is a benchmark model which embodies the assumption of no permanent shifts in the average inflation rate. In this framework, we investigate whether the nonlinearities of the model can explain the persistently high inflation rates observed in the first part of our sample as reflecting the impact of higher order terms in the solution, terms which play instead a negligible role when inflation is low (and presumably close to the steady state).

We then compare these results to those obtained within a model where smooth shifts in the mean of inflation are allowed for through an integrated inflation target. We study the relevance of the nonlinear aspects of the model also in this case – even if one would expect them to be less relevant, because time variations in the mean should imply that deviations from the mean are normally small.

In the rest of this section, we present in more detail the main features of the microeconomic environment and the two policy rules which represent the main differences across the two specifications.

2.1 A simple DSGE model

The model is based on the framework developed by Woodford (2003) and extended in a number of directions by Christiano, Eichenbaum and Evans (2005).

Consumers maximize the discounted sum of the period utility

$$U(C_t, H_t, L_t) = \frac{(C_t - hC_{t-1})^{1-\gamma}}{1-\gamma} - \int_0^1 \chi L_t(i)^\phi \, di \tag{1}$$

where C is a consumption index satisfying

$$C = \left(\int_0^1 C(i)^{\frac{\theta-1}{\theta}} \right)^{\frac{\theta}{\theta-1}}, \tag{2}$$

$H_t = hC_{t-1}$ is the habit stock, $L(i)$ are hours of labour provided to firm i.

For consistency with Smets and Wouters (2003) and Christiano, Eichenbaum and Evans (2005), habit formation is modelled in difference form. However, habit is internal, so that households care about their own lagged consumption.

The households' budget constraint is given by

$$P_t C_t + B_t \leqslant \left(1 - \frac{\tau_t}{1+\tau_t}\right) \int_0^1 w_t(i) L_t(i) \, di + \int_0^1 \Xi_t(i) \, di + W_t \tag{3}$$

with the price level P_t defined as the minimal cost of buying one unit of C_t, hence equal to

$$P_t = \left(\int_0^1 p(i)^{1-\theta} \right)^{\frac{1}{1-\theta}}. \tag{4}$$

In the budget constraint, B_t denotes end of period holdings of a complete portfolio of state contingent assets. W_t denotes the beginning of period value of the assets, $w_t(i)$ is the nominal wage rate, $L_t(i)$ is the supply of labor to firm i and $\Xi_t(i)$ are the profits received from investment in firm i. Following Steinsson (2003), we also introduce a stochastic income tax, which will lead to a trade-off between inflation and the output gap. We write the tax rate as $\frac{\tau_t}{1+\tau_t}$ to ensure that the total tax is bounded between 0 and 1, given that

$$\log \tau_t = (1-\rho_\tau)\overline{\tau} + \rho_\tau \log \tau_{t-1} + v_t^\tau, \qquad v_t^\tau \sim N\left(0, \sigma_\tau^2\right). \tag{5}$$

The first order conditions w.r.t intertemporal aggregate consumption allocation and labour supply can be written as

$$\begin{aligned} \left(1 - \frac{\tau_t}{1+\tau_t}\right) \frac{w_t(i)}{P_t} &= \frac{\phi \chi L(i)^{\phi-1}}{\Lambda_t} \\ \Lambda_t &= (C_t - hC_{t-1})^{-\gamma} - \beta h \mathrm{E}_t \left[(C_{t+1} - hC_t)^{-\gamma} \right] \\ \frac{1}{I_t} &= \mathrm{E}_t \left[\beta \frac{P_t}{P_{t+1}} \frac{\Lambda_{t+1}}{\Lambda_t} \right]. \end{aligned} \tag{6}$$

where I_t is the gross nominal interest rate.

Turning to the firms problem, the production function is given by

$$Y_t(i) = A_t L(i)^{\alpha}, \qquad A_t = A_{t-1}^{\rho_a} e^{v_t^a} \tag{7}$$

where A_t is a technology shock and v_t^a is a normally distributed shock with constant variance σ_a^2.

We assume Calvo (1983) contracts, so that firms face a constant probability ζ of being unable to change their price at each time t. Firms will take this constraint into account when trying to maximize expected profits, namely

$$\max_{P_t^i} \mathrm{E}_t \sum_{s=t}^{\infty} \zeta^{s-t} \beta^s \frac{P_t}{P_{t+s}} \frac{\Lambda_{t+s}}{\Lambda_t} \left(P_s^i Y_s^i - TC_s^i \right), \tag{8}$$

where TC denotes total costs and, as in Smets and Wouters (2003), firms not changing prices optimally are assumed to modify them using a rule of thumb that indexes them partly to lagged inflation and partly to steady-state inflation $\overline{\Pi}$, namely $P_t^i \left(\overline{\Pi}\right)^{1-\iota} \left(\frac{P_{s-1}}{P_{t-1}}\right)^{\iota}$, where $0 \leq \iota \leq 1$. The exception is when we assume an integrated inflation target and steady state inflation is not defined. In that case, we set $\iota = 1$. We introduce indexation in the model for two reasons. First, aggregate inflation will be driven to some extent by lagged inflation, which is an empirically plausible hypothesis – though not immediately consistent with the microeconomic evidence. Second, firms not allowed to update their prices optimally for a long time will still find themselves with a price which is not too far from the optimum.

Under the assumption that firms are perfectly symmetric in all other respects than the ability to change prices, all firms that do get to change their price will set it at the same optimal level P_t^*. The first order conditions of the firms' problem can be written recursively as implying

$$\begin{aligned}
\left(\frac{P_t^*}{P_t}\right)^{1-\theta\left(1-\frac{\phi}{\alpha}\right)} &= \frac{\phi\chi\theta}{\alpha(\theta-1)}\frac{K_{2,t}}{K_{1,t}} \\
K_{2,t} &= \frac{A_t^{-\frac{\phi}{\alpha}}}{\left(1-\frac{\tau_t}{1+\tau_t}\right)\Lambda_t}Y_t^{\frac{\phi}{\alpha}} + \mathrm{E}_t\zeta\overline{\Pi}^{-\theta\frac{\phi}{\alpha}(1-\iota)}\beta\frac{\Lambda_{t+1}}{\Lambda_t}K_{2,t+1}\Pi_t^{-\theta\frac{\phi}{\alpha}\iota}\Pi_{t+1}^{\theta\frac{\phi}{\alpha}} \qquad (9) \\
K_{1,t} &= Y_t + \mathrm{E}_t\zeta\overline{\Pi}^{(1-\theta)(1-\iota)}\beta\frac{\Lambda_{t+1}}{\Lambda_t}K_{1,t+1}\Pi_t^{(1-\theta)\iota}\Pi_{t+1}^{\theta-1}
\end{aligned}$$

where Π_t is the inflation rate defined as $\Pi_t \equiv \frac{P_t}{P_{t-1}}$ and

$$\frac{P_t^*}{P_t} = \left(\frac{1-\zeta\left(\overline{\Pi}^{1-\iota}\frac{\Pi_{t-1}^{\iota}}{\Pi_t}\right)^{1-\theta}}{(1-\zeta)}\right)^{\frac{1}{1-\theta}} \qquad (10)$$

expresses the optimal price at time t as a function of aggregate variables.[1]

Note that we can use equation (10) in the system (9) to write aggregate inflation as an implicit function of expected future inflation

$$\Pi_t = \overline{\Pi}^{1-\iota}\Pi_{t-1}^{\iota}\left(\frac{1}{\zeta} - \frac{1-\zeta}{\zeta}\left(\frac{\phi\chi\theta}{\alpha(\theta-1)}\frac{K_{2,t}}{K_{1,t}}\right)^{\frac{1-\theta}{1-\theta\left(1-\frac{\phi}{\alpha}\right)}}\right)^{\frac{1}{\theta-1}} \qquad (11)$$

It is well known that a first order approximation of this equation yields the new-Keynesian Phillips curve, where inflation is positively related to expected future inflation. When inflation is high, the cost of not being able to freely adjust prices is higher: firms will therefore try to front-load on future price increases by increasing their mark-up over current costs. The second-order approximation of equation (11) is more elaborate, so that the relationship between current and future inflation is not immediately apparent (see Benigno and Woodford, 2005, for an example in the simpler case without habits nor inflation indexation). Nevertheless, equation (11) is suggestive of two features.

First, past inflation only enters linearly in the equation, since it never appears in the $K_{2,t}$ and $K_{1,t}$ terms. Even with indexation, the fact that past inflation is high does not *per se* matter in inducing a nonlinearity in inflation as a function of the state of the economy. Indexation does, of course, matter in changing expectations of future inflation.

The second known feature of the quadratic approximation of equation (11) is that this will be either concave or convex, regardless of whether inflation deviations from the long run mean are positive or negative. The effects of the second order terms in the solution

[1] Similar expressions are derived, amongst others, by Ascari (2004) and Benigno and Woodford (2005).

will therefore be asymmetric. If, *ceteris paribus*, inflation is a convex function of expected future inflation, firms will try to increase current prices more and more aggressively, the larger is the expected future deviation of inflation from steady state. They will however cut their prices less than they would in the linear case, in case of negative inflation deviations from the steady state.

2.2 Two Taylor rules

Equations (3), (6)-(8), (10)-(11) describe aggregate economic dynamics. We close the model with a Taylor rule with interest rate smoothing. A key decision that has to be taken in the specification of the rule concerns the inflation target. Since inflation displays a noticeable downward trend over the sample period, the assumption of a constant target is not very appealing. In empirical applications, it is therefore often assumed that the decline in inflation corresponds to a decline in the inflation target. This is also what we do here. However, this assumption is likely to have important implications in terms of the persistence of inflation. In order to explore this issue, we analyse two variants of the policy rule.

The first rule assumes that the inflation target follows a stationary AR(1) process. In this case, the idea is that the long run target of the central bank is actually constant, but that there are shifts in the horizon at which the central bank tries to get inflation back to that long run level. If the target is temporarily high when inflation is high, then the central bank is willing to tolerate a slow return to the long run target. If, instead, there are no changes in the long run target when inflation is high, inflation will be brought back on target more quickly.

In logarithmic terms (lower case letters), the first rule takes the form

$$i_t = (1-\rho_I)(\overline{\pi} - \ln\beta) + \psi_\pi(\pi_t - \pi_t^*) + \psi_y(y_t - y_t^n) + \rho_I i_{t-1} + v_t^i \tag{12}$$

$$\pi_t^* = (1-\rho_\pi)\overline{\pi} + \rho_\pi \pi_{t-1}^* + v_t^{\pi^*} \tag{13}$$

where i_t is the logarithm of the gross nominal interest rate, π_t^* is the inflation target, u_t^i is a policy shock and y_t^n is the logarithm of the level of natural output. The innovations v_t^i and $v_t^{\pi^*}$ are white noise with variances σ_i^2 and $\sigma_{\pi^*}^2$, respectively. In this model, considerable deviations from the mean of inflation can arise from short-term movements in the inflation target. The model solved using the first policy rule is dubbed M1.

The second policy rule is identical to the first, except for the property that the inflation target becomes integrated (and the steady state level of the interest rate is modified accordingly)

$$i_t = (1-\rho_I)\left((\pi_t^* - \ln\beta) + \psi_\pi(\pi_t - \pi_t^*) + \psi_y(y_t - y_t^n)\right) + \rho_I i_{t-1} + v_t^i \tag{14}$$

$$\pi_t^* = \pi_{t-1}^* + v_t^{\pi^*} \tag{15}$$

In this case, smooth changes of the mean occur over time as the central bank target is revised. The idea here is that the inflation target process captures true shifts in the objective of the central bank. Given the slow decline in inflation over our sample period, this should supposedly reflect a shift in public preferences in favour of lower and lower inflation levels. The integrated inflation target induces a non-stationary behaviour also in actual inflation and the nominal interest rate. These nominal variables are also co-integrated, so that the model can be written in stationary form in terms of the rate of growth of inflation, $\Delta\pi_t = \pi_t - \pi_{t-1}$, and the deflated inflation target and interest rate, defined as $\widetilde{\pi}_t^* = \pi_t^* - \pi_t$ and $\widetilde{i}_t = i_t - \pi_t$, respectively (see the appendix). This model is dubbed M2.

3 Second-order approximate solution

We solve the model using a second order approximation around the non-stochastic steady state. The model dynamics will then be described by two systems of equations: a quadratic law of motion for the predetermined variables of the model and a quadratic relationship linking each non-predetermined variable to the predetermined variables.

The solution is obtained numerically. A few methods have been proposed in the literature, including those in Schmitt-Grohé and Uribe (2002) and Kim *et al.* (2003). For our applications we select the implementation proposed by Klein (2005), that has the advantage of being relatively faster. Speed is particularly important for estimation, since the model needs to be solved at every evaluation of the likelihood. For this reason, we also rely on analytical derivatives to evaluate the second order terms of the approximation.

The solution can be characterised as follows. The vector $\widehat{x}_t$ of predetermined variables follows the quadratic law of motion

$$\widehat{x}_{t+1} = \frac{1}{2}k_x + P\widehat{x}_t + \frac{1}{2}\left(I_{n_x} \otimes \widehat{x}_t'\right) G\widehat{x}_t + v_{t+1} \tag{16}$$

where P, k_x and G are $n_x \times n_x$, $n_x \times 1$ and $n_x^2 \times n_x$ matrices, respectively. The vector of shocks has variance covariance matrix Σ. Non-predetermined variables, y_t, are linked to predetermined variables by the solution

$$\widehat{y}_t = \frac{1}{2}k_y + F\widehat{x}_t + \frac{1}{2}\left(I_{n_y} \otimes \widehat{x}_t'\right) E\widehat{x}_t \tag{17}$$

where k_y, F and E are $n_y \times 1$, $n_y \times n_x$ and $n_y n_x \times n_x$ matrices, respectively.

4 Estimation method

4.1 Non linear-non Gaussian state space models

The system (16)-(17) can be cast in the general form

$$\text{(measurement equation) } \mathbf{y}_t^o = G(\mathbf{x}_t, \mathbf{v}_t, \boldsymbol{\theta}) \tag{18}$$

$$\text{(state equation)} \mathbf{x}_t = H(\mathbf{x}_{t-1}, \mathbf{w}_t, \boldsymbol{\theta}) \tag{19}$$

where (omitting hats) $\mathbf{y}_t^o$ is the subset of observable elements of the vector $\mathbf{y}_t$, $\boldsymbol{\theta}$ is the parameter vector, $\mathbf{v}_t \equiv \left[v_t^a, v_t^{\pi^*}, v_t^{\tau}, v_t^i\right]'$ is the vector of structural shocks and $\mathbf{w}_t$ are measurement errors.

In order to be able to do inference on the unobservables (parameters and state vector) we need to solve a filtering problem, i.e. given $p(\mathbf{x}_t|\underline{\mathbf{y}}_t^o, \boldsymbol{\theta})$ obtain $p(\mathbf{x}_{t+1}|\underline{\mathbf{y}}_{t+1}^o, \boldsymbol{\theta}), t = 0, 1, ...T-1$, where

$$\underline{\mathbf{y}}_t^o = \left[\begin{array}{cccc} \mathbf{y}_1^{o'} & \mathbf{y}_2^{o'} & ... & \mathbf{y}_t^{o'} \end{array}\right]' \tag{20}$$

collects all the data evidence up to time t.

The filtering problem is conceptually very easy and consists of two steps:

- projection

$$p(\mathbf{x}_{t+1}|\underline{\mathbf{y}}_t^o, \boldsymbol{\theta}) = \int p(\mathbf{x}_{t+1}|\mathbf{x}_t, \boldsymbol{\theta}) p(\mathbf{x}_t|\underline{\mathbf{y}}_t^o, \boldsymbol{\theta}) d\mathbf{x}_t \tag{21}$$

- update

$$p(\mathbf{x}_{t+1}|\underline{\mathbf{y}}_{t+1}^o, \boldsymbol{\theta}) = \frac{p(\mathbf{x}_{t+1}|\underline{\mathbf{y}}_t^o, \boldsymbol{\theta}) p(\mathbf{y}_{t+1}^o|\mathbf{x}_{t+1}, \boldsymbol{\theta})}{p(\mathbf{y}_{t+1}^o|\underline{\mathbf{y}}_t^o, \boldsymbol{\theta})} \tag{22}$$

$$p(\mathbf{y}_{t+1}^o|\underline{\mathbf{y}}_t^o, \boldsymbol{\theta}) = \int p(\mathbf{x}_{t+1}|\underline{\mathbf{y}}_t^o, \boldsymbol{\theta}) p(\mathbf{y}_{t+1}^o|\mathbf{x}_{t+1}, \boldsymbol{\theta}) d\mathbf{x}_{t+1} \tag{23}$$

Equation (23) indicates that the filtering recursion yields the likelihood of each observation.

The integration steps which are inherent in the filtering recursion can be easily performed under two very special circumstances:

- when the support of the state variables is discrete (and finite): then the integrals are just summations;
- when the state and the measurement equations are both linear and the disturbances are Gaussian: in this latter case we can use the Kalman filter recursion.

In the more general context of this paper, where we have to deal with non-linearities, we use Sequential Monte Carlo (SMC) methods. The literature on these methods is vast and good and accessible introductions are Arulampalam *et al.* (2002), Doucet *et al.* (2001). Fernandez-Villaverde and Rubio-Ramirez (2004) and An and Schorfheide (2005) are the first studies in which these techniques are used for DSGE models.

SMC methods applied to models with latent variables allow to do filtering by simulation.

The intuition behind the simplest version of these methods, which is called the *particle filter* (PF) is to compute the likelihood $p(\mathbf{y}^o_{t+1}|\underline{\mathbf{y}}^o_t, \boldsymbol{\theta})$ by:

1. drawing a large number of realisations from the distribution of $\mathbf{x}_{t+1}$ conditioned on $\underline{\mathbf{y}}^o_t$;
2. assigning them a weight which is determined by their "distance " from (compatibility with) $\mathbf{y}^o_{t+1}$.

If we call $p(\mathbf{x}_{t+1}|\underline{\mathbf{y}}^o_t, \boldsymbol{\theta})$ the prior distribution (prior to observing $\mathbf{y}^o_{t+1}$) and $p(\mathbf{y}^o_{t+1}|\mathbf{x}_{t+1}, \boldsymbol{\theta})$ the "likelihood", the PF algorithm can be given a very simple Bayesian interpretation which immediately clarifies its limitations: it is as if we were doing posterior simulation drawing from the prior and then using the likelihood as weights. This is a very easy procedure to implement but hardly a computationally efficient one in the case the likelihood is much more concentrated than the prior.

The PF works as follows: imagine that at time t we have the availability of a large number N of draws to approximate $p(\mathbf{x}_t|\underline{\mathbf{y}}^o_t, \boldsymbol{\theta})$ (we have a so called *swarm of particles*):

$$\left(\mathbf{x}_t^{(i)}, w_t^{(i)}\right), i = 1, 2, ..., N \tag{24}$$

Each particle is endowed with a weight $w_t^{(i)}$ which might be the result of the fact that each $\mathbf{x}_t^{(i)}$ might have been drawn from a distribution $q(\mathbf{x}_t)$ which does not coincide with $p(\mathbf{x}_t|\underline{\mathbf{y}}_t^o, \boldsymbol{\theta})$[2]:

$$w_t^{(i)} = \frac{p(\mathbf{x}_t^{(i)}|\underline{\mathbf{y}}_t^o, \boldsymbol{\theta})}{q(\mathbf{x}_t^{(i)})} \tag{25}$$

Note that it is possible to use this sample to compute the expected value of any function f of $\mathbf{x}_t$ in two different ways:

- via direct Importance Sampling (IS):

$$E\left[f(\mathbf{x}_t^{(i)})|\underline{\mathbf{y}}_t^o, \boldsymbol{\theta}\right] \approx \frac{\sum_{i=1}^{N} w_t^{(i)} f(\mathbf{x}_t^{(i)})}{\sum_{i=1}^{N} w_t^{(i)}} \tag{26}$$

- Alternatively it is possible to resample the $\mathbf{x}_t^{(i)}$ drawing N times (with reimmission) from empirical distribution of the $\mathbf{x}_t^{(i)}$, with probabilities given by $w_t^{(i)}$. In this way we obtain a modified sample

$$\left(\mathbf{x}_t^{(j)}, 1\right), j = 1, 2, ..., N \tag{27}$$

  which can be directly used to form the desired estimate:

$$E\left[f(\mathbf{x}_t^{(i)})|\underline{\mathbf{y}}_t^o, \boldsymbol{\theta}\right] \approx \frac{\sum_{j=1}^{N} f(\mathbf{x}_t^{(j)})}{N} \tag{28}$$

The swarm of particles $\left(\mathbf{x}_t^{(i)}, w_t^{(i)}\right), i = 1, 2, ..., N$, can be used to perform filtering by simulation. The easiest way to do this is to construct the so called particle filter (PF) algorithm which consists of two simple steps. Assuming to have a swarm of particles with perfectly even weights ($w_t^{(j)} = 1, j = 1, 2, ..., N$), the simulation filtering consists in two steps:

- projection

$$p(\mathbf{x}_{t+1}|\underline{\mathbf{y}}_{t+1}^o, \boldsymbol{\theta}) \approx \frac{\sum_{j=1}^{N} p(\mathbf{x}_{t+1}|\mathbf{x}_t^{(j)})}{N} \tag{29}$$

[2] In this case we have the so-called Importance sampling. See Geweke 1989.

This step is empirically performed by taking each particle $\mathbf{x}_t^{(j)}$and drawing $\mathbf{x}_{t+1}^{(j)}$from the distribution

$$p(\mathbf{x}_{t+1}|\mathbf{x}_t^{(j)},\boldsymbol{\theta}) \tag{30}$$

This amounts to simulate from the state equation. In this way, the projection distribution is approximated by the sample

$$\left(\mathbf{x}_{t+1}^{(j)},1\right),j=1,2,...,N \tag{31}$$

- Update: we take into account that we have drawn from $p(\mathbf{x}_{t+1}|\underline{\mathbf{y}}_t^o,\boldsymbol{\theta})$ and not from $p(\mathbf{x}_{t+1}|\underline{\mathbf{y}}_{t+1}^o,\boldsymbol{\theta})$ by assigning weights proportional to $p(\mathbf{y}_{t+1}^o|\mathbf{x}_{t+1}^{(j)},\boldsymbol{\theta})$. So, the updated distribution $p(\mathbf{x}_{t+1}|\underline{\mathbf{y}}_{t+1}^o,\boldsymbol{\theta})$ is approximated by the sample

$$\left(\mathbf{x}_{t+1}^{(j)},w_{t+1}^{(j)}\right),j = 1,2,...,N, \tag{32}$$

$$w_{t+1}^{(j)} = p(\mathbf{y}_{t+1}^o|\mathbf{x}_{t+1}^{(j)},\boldsymbol{\theta}) \tag{33}$$

This sample can be resampled using the weights $w_{t+1}^{(j)}$ as probabilities.

This resampling step, performed at the end of each cycle of the filter (after updating) is performed in order to generate a sample with even weights (by definition these weights are all equal after resampling). It is certainly possible to avoid this resampling step at each observation t but the consequence is that while the filter progresses towards the end of the sample the cumulation of the weights will make them very polarised. In the absence of resampling the weights assigned to particle $\mathbf{x}_t^{(j)}$ will by definition be

$$\prod_{i=0}^{t-1} w_{t-i}^{(j)} = \prod_{i=0}^{t-1} p(\mathbf{y}_{t-i}^o|\mathbf{x}_{t-i}^{(j)}) \tag{34}$$

These weights are such that after a while (at some t) the weight assigned to the particle even marginally most compatible with the observable data will be 1 and all the other particles will have zero weights. In other words in the absence of resampling the numerical accuracy of the filter quickly deteriorates.

One interesting way to monitor the numerical accuracy of the filter is to compute the sum of squares of the weights (NEFF, i.e. numerical efficiency index)

$$NEFF_t = \sum_{i=1}^{N}\left(w_t^{(i)}\right)^2 \tag{35}$$

As the Herfindhal-Hirschmann index, when all the weights are even this index is 1/N and this marks the most efficient working of the filter. In the opposite case (ie one weight equal to one) the index goes to 1.

Note that the unnormalised weights (33) are very important for inference: their sample mean is the t^{th} observation conditional density:

$$\begin{aligned}
& \frac{1}{N}\sum_{j=1}^{N} p(\mathbf{y}^o_{t+1}|\mathbf{x}^{(j)}_{t+1},\boldsymbol{\theta}) \\
\approx & \iint p(\mathbf{y}^o_{t+1}|\mathbf{x}_{t+1},\boldsymbol{\theta})p(\mathbf{x}_{t+1}|\mathbf{x}_t,\boldsymbol{\theta})p(\mathbf{x}_t|\underline{\mathbf{y}}^o_t,\boldsymbol{\theta})d\mathbf{x}_{t+1}d\mathbf{x}_t = \\
= & \; p(\mathbf{y}^o_{t+1}|\underline{\mathbf{y}}^o_t,\boldsymbol{\theta})
\end{aligned} \tag{36}$$

Therefore we can get easily the likelihood of the sample of observable variables. This likelihood can be used as a basis for full information inference (Bayesian or not) on the parameters of the model while the whole filtering procedure can be used for carrying out smoothed or filtered inference on the unobservable variables.

4.2 Inference on the parameters of the model

Once likelihood has been obtained, it can be used either in a ML estimation framework or in a Bayesian posterior simulation algorithm.

In this paper we use a random walk Metropolis Hastings algorithm (see Chib, 2001) which works by sequentially repeating the following steps:

- draw $\boldsymbol{\theta}^{(i)}$ from a candidate distribution $q_V(\boldsymbol{\theta}^{(i-1)})$;
- compute the solution of the DSGE model and the implied state space form;
- carry out the simulation filter which will produce also the likelihood of the model $p(\underline{\mathbf{y}}^o_T|\boldsymbol{\theta}^{(i)}) = \prod_{t=1}^{T-1} p(\mathbf{y}^o_{t+1}|\underline{\mathbf{y}}^o_t,\boldsymbol{\theta}^{(i)})$;
- accept $\boldsymbol{\theta}^{(i)}$ with probability

$$\frac{p(\boldsymbol{\theta}^{(i)})p(\underline{\mathbf{y}}^o_T|\boldsymbol{\theta}^{(i)})}{p(\boldsymbol{\theta}^{(i-1)})p(\underline{\mathbf{y}}^o_T|\boldsymbol{\theta}^{(i-1)})} \tag{37}$$

  if the draw is not accepted the MH simulator sets $\boldsymbol{\theta}^{(i)} = \boldsymbol{\theta}^{(i-1)}$.

4.3 Prior elicitation

One of the hardest parts in implementing Bayesian techniques is how to specify sensible priors. There are parameters for which this task is less difficult. For some others (typically the second order parameters) this task is more difficult.

For most of the macro parameters, we have adopted priors consistent with those of Smets and Wouters (2003). The main difference concerns the parameters which play a more determinant role for the second order approximation, namely the standard errors of the structural shocks.

For these latter parameters we have resorted to a *prior predictive* approach (see Geweke and McCausland, 2002): we draw parameter values from the joint prior, we solve the model and we compute the moments of the stationary distribution of the data. We obtain in this way a prior distribution of these model based features and we compare this distribution with the actual sample moments of the data available prior to the period used for estimation (the data for the 1970s). We calibrated the prior hyperparameters in order to have a prior distribution of the first and second moments of the model-based ergodic distribution centered around values on the same order of magnitude of the pre-sample data moments.

The prior distribution on the measurement standard errors is based on the assumption that their scale is believed to be small.

We have experienced that a bit of thought in the specification of the prior usually helps in eliminating some of the numerical problems encountered by the SMC filtering procedures.

5 Results

All our results are based on output, nominal interest rate and inflation data taken from the Area Wide Model database (see Fagan, Henry and Mestre, 2005). Following Smets and Wouters (2003), we remove a deterministic trend from the GDP series prior to estimation. No transformations are applied to inflation and interest rate data. The estimation period runs from 1980Q1 to 2004Q4. The data are shown in Figure 1.

In this section, we first present a comparison of the estimates obtained using first and second order approximations. We then focus on the implications of the second order results for inflation persistence. We conclude the section with a formal comparison of the

M1 and M2 estimates. All results are based on 26,000 replications, the first 6000 of which are discarded.

5.1 Linear vs. nonlinear estimates

Tables 1 and 2 present the results of the estimation of first and second order versions of M1 and M2. Three broad conclusions can be drawn from the tables.

First, estimates of the second order model tend to be more precise. This result, which is consistent with those reported by An and Schorfeide (2005) on the basis of simulated data, is particularly noticeable for M2. The posterior standard deviations of most parameter estimates are in fact lower in the quadratic case. A similar picture would be obtained looking at posterior marginal distributions. The distribution is much more concentrated for parameters such as the risk aversion coefficient and, in the case of M2, the habit parameter and the inflation response parameter in the Taylor rule.

Second, the nonlinear solutions do not produce sizable changes in parameter estimates. Variations in mean estimates tend to be quantitatively small and mostly insignificant from a statistical viewpoint. The most notable exception concerns the degree of inflation indexation in the M1 case. This is much higher in the nonlinear than in the linear model.

Third, estimates of the deep parameters are quite stable across models. Also here, however, there is a notable exception. The Calvo parameter is quite high and close to the results in Smets and Wouters (2003) for M1. With an integrated target, however, the parameter estimate falls considerably.

Focusing on M1, we observe that mean posteriors are consistent with a very high degree of price stickiness and of habit persistence (see Table 1). Monetary policy is characterised by an aggressive inflation response parameter in the Taylor rule, a mild response to the output gap and high interest rate smoothing.

For M2, mean posteriors are consistent with a much lower degree of price stickiness – prices are readjusted on average every 1.5 quarters (see Table 2). Monetary policy is characterised by a more moderate inflation response parameter in the Taylor rule, together with a negligible response to the output gap. Interest rate smoothing remains high. A few minor differences can be noted between posterior means, but these are not statistically significant. This is the case, for example, for the risk aversion paratemer, which is slightly lower in the second order case, or the elasticity of labour supply, which is higher. In this case, the similarity is to be expected, since the inflation target is estimated to follow actual inflation quite closely over the sample. Consequently, deviations of inflation and the nominal interest rate from the steady state are relativey small and second order terms in the solution play a negligible role.

In order to test the general stability of our results, we run several different estimation rounds. Estimates are quite stable for the M1 model, but less so for the M2 model. Across different estimation rounds, one observes quite some variability in the results. We plan to increase the lenght of the chains for the M2 model in future versions of this paper.

5.2 Euro area inflation dynamics

In this section, we discuss the implications of the model for the conditional dynamics of inflation. Various definition of inflation persistence are possible – see Angeloni *et al.* (2005). In the rest of this section, we focus on one main piece of evidence: the profile of the impulse response functions. We test whether responses starting from a high-inflation level are significantly more persistent than those starting from a low level.

For M1, the estimated steady state of the inflation objective is around 2.7 percent, so that the values of inflation observed at the beginning and at the end of our sample period (9.4 and 1.2 percent, respectively) do represent large deviations from the steady state. Figure 2-5 display the impulse responses of inflation to a standard error shock in all the structural disturbances, starting from the configuration of the state variables observed either at the beginning or at the end of the sample. Given these two starting points, the impulse responses are computed in deviation from the path which would have been observed in the absence of the impulse shocks.

With the exception of the cost-push shock, which in any case plays a minor role in term of variance decomposition, all other impulse responses demonstrate that the persistence of inflation tends to be higher when inflation is high.

More specifically, the response to a technology shock (Figure 2) follows the broad pattern typically observed in linearised models, if the shock occurs when inflation is low (the "Low inflation" line). Inflation falls for a few quarters and returns to baseline thereafter. The initial fall is also statistically significant at the 95% level for 4 quarters. Starting from a high inflation level (the "High inflation" line) however, the fall in inflation is reduced dramatically and ceases to be significant from a statistical viewpoint. In addition, the median response tends to have a much more persistent effects. Ten years after the shock, inflation remains around the level observed on impact.

The response to a positive inflation target shocks is even more markedly dependent on the starting point (Figure 3). The response is very similar to the linear case, if the shock hits when inflation is low. If inflation is already high, however, it tends to rise further on impact, after the shock, and to build up considerably thereafter (before being eventually reabsorbed). From a quantitative viewpoint, the impact of a given target shock is almost twice as big, and the difference is almost always significant from a statistical viewpoint. Nevertheless, in terms of persistence profile, the shock has similar effects independently of the starting point: inflation peaks after approximately 2 years and remains close to this higher level for the remaining 8 years displayed in the figure.

Finally, some quantitative differences between median responses are noticeable also in the cost-push and policy shock cases (Figures 4 and 5, respectively). However, these responses tend to be insignificant from a statistical viewpoint. More specifically, the cost-push shock has a truly negligible impact on inflation dynamics. As to the policy shock, after a surprising interest rate hike the median fall in inflation is smaller, when the starting point is a high inflation level. This would imply that "surprise" disinflations become increasingly costly the higher the level inflation is allowed to reach.

In the case of M2, we find that the estimated inflation objective tracks quite closely actual inflation. Since the objective represents the time-varying steady state of inflation, observed deviations of inflation from the steady state are estimated to be negligible. Consequently, there are no noticeable changes in the in-sample persistence of inflation.

Figures 6-9 illustrate this property. They display the impulse responses of inflation to a standard error shock in all the structural disturbances, starting from the configuration of the state variables observed either at the beginning or at the end of the sample. All impulse responses follow a similar profile independently of the point from which they are computed.

5.3 Model comparison

Since our results on inflation persistence are starkly different across the M1 or M2 models, we report in this section the results of a formal model comparison test based on the two Marginalised Likelihoods.

For reach model M_j we compute

$$\ln(p(\mathbf{y}|M_j) = \ln\left(\int p(\mathbf{y}|\boldsymbol{\theta}, M_j)p(\boldsymbol{\theta}|\mathbf{y}, M_j)d\boldsymbol{\theta}\right) \tag{38}$$

The difference between these two quantities gives the log Bayes Factor of one model vs the other. Computed values largely different from zero suggest dominance of one model vs the other. The marginal likelihoods are computed based on the modified Gelfand and Dey approach described in Geweke (1999). This method is very accurate when the posterior PDF is unimodal, a property which we cannot be sure of for the M2 model.

There are two issues of detail which have to be emphasised when comparing marginal likelihoods.

First, the two models do not use the same observable variables (differenced variables enter M2). Nevertheless, conditional on past information, the Jacobian determinant of the transformation from the variables in M1 to the ones in M2 is unity. This can be shown if we define $\mathbf{y}_t^{(M_1)} = [\pi_t, r_t, y_t]'$ and $\mathbf{y}_t^{(M_2)} = [\Delta\pi_t, r_t - \pi_t, y_t]'$. Then $\mathbf{y}_t^{(M_2)}$ can be rewritten as

$$\mathbf{y}_t^{(M_2)} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{y}_t^{(M_1)} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{y}_{t-1}^{(M_1)}$$

so that

$$\left| \frac{\partial \mathbf{y}_t^{(M_2)}}{\partial \mathbf{y}_t^{(M_1)'}} |\mathbf{y}_{t-1}^{(M_1)} \right| = \left| \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right| = 1$$

It follows that marginal likelihoods are directly comparable if we condition on $\mathbf{y}_0$.

The second issue concerns the sample size. As things stand, model M1 is estimated on a sample size that includes 1980:1 whereas M2 starts from 1980:2. To compare the two models we should rerun estimation of M1 excluding the first observation.

Table 3 summarises the results from the model comparison. The Bayes factor is so overwhelmingly in favour of M1 that no qualitative differences in the model comparison exercise can be expected in a comparison based on exactly the same sample size (i.e. if model M1 were also estimated as of 1980:2).

We conclude that inflation persistence has indeed varied significantly over the past 20 years. This variation can be explained through the nonlinear features of a standard microfounded macromodel, once those features are allowed to play a role in the solution.

6 Conclusions

Our results suggest that the nonlinearities in the dynamics of inflation can have pronounced and statistically significant effects in case of large swings in the inflation rate. In terms of the debate on the possible presence of shifts in the mean of euro area inflation, our evidence is in favour of the hypothesis of no shifts. Nevertheless, our nonlinear model without structural breaks can account for differences in the pesistence of the inflation response to shocks over time.

From a more general viewpoint, our results illustrate some of the advantages, and drawbacks, of estimating DSGE models based on a nonlinear approximation.

One clear advantage of nonlinear estimation is to increase the identifiability of parameters. Recent literature (eg. Canova and Sala, 2005) has emphasised the chronic under-identification of many DSGE models. It is possible to verify that resorting to higher order approximation induces sensibly more curvature in the likelihood function hence increases identifiability of the parameters. We have verified this feature also for the models that we estimate in this paper, both on the real dataset and on simulated data. An interesting qualification of this identification-enhancing feature of non-linear approximation is that it is very evident for those parameters for which the likelihood has some (even minimal) curvature already with the linear solution. If the likelihood is absolutely flat in the linear case, it tends to remain flat. We believe that this increase of likelihood curvature is generated by the fact that the nonlinear approximation generates many relationships involving first and second order moments of the distribution for the data.

The other main advantage of nonlinear models is that higher order terms can be theoretically and practically important. In our application, the higher order terms of the solution of the model cannot be neglected, without risking to draw the wrong conclusions on inflation persistence. Similar problems would occur, for example, if the model were used for asset pricing.

Nevertheless, the estimation on nonlinear models does have drawbacks. The first disadvantage (induced by the need of resorting to simulation filtering to carry out likelihood based inference) is essentially that much more computation time is required. This is clearly unpleasant, since Bayesian estimation of DSGE models typically requires running multiple, long chains after having tuned carefully (and painfully through the use of many chains) the candidate distribution for the parameters. The methodological "silver lining" of this feature is that the researcher can do data mining only with geological endowments of time.

Another drawback is the frequent occurrence of numerical problems. We have experienced a strong sensitivity of SMC methods to outliers and degeneracies which frequently arise in actual data. The researcher has to be very careful in monitoring the numerical efficiency indicators of the filter that is being used. Moreover the non linear state space system that arises often implies explosive behaviour for the state variables also for parameter values which would assign a stationary distribution to the state variables in a linear setting. Great care has to be to put in designing a prior distribution that rules out such occurrences.

Table 1: Estimation results: MODEL1

			I order approx		II order approx	
	prior mean	prior sd	post mean	post sd	post mean	post sd
γ	2.000	0.705	2.233	0.728	1.139	0.087
h	0.699	0.138	0.736	0.069	0.799	0.071
ϕ	3.982	1.993	5.265	2.091	2.290	0.899
θ	7.985	2.648	9.188	2.807	6.996	2.058
ζ	0.602	0.147	0.855	0.049	0.898	0.072
ι	0.666	0.179	0.304	0.157	0.771	0.111
ψ_π	0.500	0.092	0.445	0.085	0.332	0.061
ψ_y	0.050	0.035	0.017	0.012	0.045	0.025
ρ_i	0.799	0.100	0.925	0.041	0.913	0.056
ρ_τ	0.500	0.151	0.498	0.152	0.515	0.100
ρ_a	0.901	0.090	0.990	0.007	0.994	0.004
ρ_π	0.898	0.093	0.992	0.004	0.995	0.003
σ_τ	0.040	0.013	0.040	0.013	0.035	0.011
σ_a	0.003	0.001	0.009	0.002	0.008	0.001
σ_π	0.001	0.001	0.001	2.E-04	0.001	3.E-04
σ_i	0.001	3.E-04	0.002	1.E-04	0.002	1.E-04
τ	0.398	0.284	2.E-04	1.E-04	3.E-04	1.E-04
π	1.005	0.003	1.005	0.003	1.007	0.003
$\sigma_{m,\pi}$	0.010	0.014	0.003	4.E-04	0.003	0.001
$\sigma_{m,i}$	0.010	0.014	0.011	0.014	1.E-07	1.E-07
$\sigma_{m,y}$	0.010	0.014	0.010	0.014	0.021	0.015

Note: $\sigma_d = 0$, $\alpha = 0.76$, $\chi = 0.275$, $\beta = 0.9944$.

Table 2: Estimation results: MODEL2

			I order approx		II order approx	
	prior mean	prior sd	post mean	post sd	post mean	post sd
β	0.902	0.091	0.992	0.001	0.991	0.001
γ	2.000	0.708	2.064	0.737	1.469	0.055
h	0.700	0.139	0.632	0.104	0.721	0.029
ϕ	4.979	1.965	2.798	0.880	4.856	0.590
θ	7.996	2.646	6.991	2.288	7.514	1.737
ζ	0.601	0.146	0.404	0.093	0.312	0.080
ψ_π	1.502	0.352	1.524	0.289	1.556	0.078
ψ_y	0.050	0.035	0.053	0.035	0.035	0.008
ρ_i	0.701	0.137	0.947	0.020	0.926	0.021
ρ_τ	0.500	0.151	0.213	0.096	0.448	0.031
ρ_a	0.901	0.089	0.997	0.003	0.981	0.004
σ_τ	0.040	0.040	0.165	0.074	0.037	0.006
σ_a	0.003	0.005	0.011	0.003	0.011	0.001
σ_π	0.001	0.001	2.E-04	3.E-04	0.001	3.E-04
σ_i	0.001	0.002	0.001	2.E-04	0.001	4.E-04
$\overline{\tau}$	0.401	0.286	0.792	0.353	0.122	0.019
$\sigma_{m,\pi}$	0.010	0.014	1.E-03	1.E-04	0.001	0.001
$\sigma_{m,i}$	0.010	0.014	4.E-05	6.E-05	0.001	3.E-04
$\sigma_{m,y}$	0.010	0.014	0.002	5.E-04	0.002	4.E-04

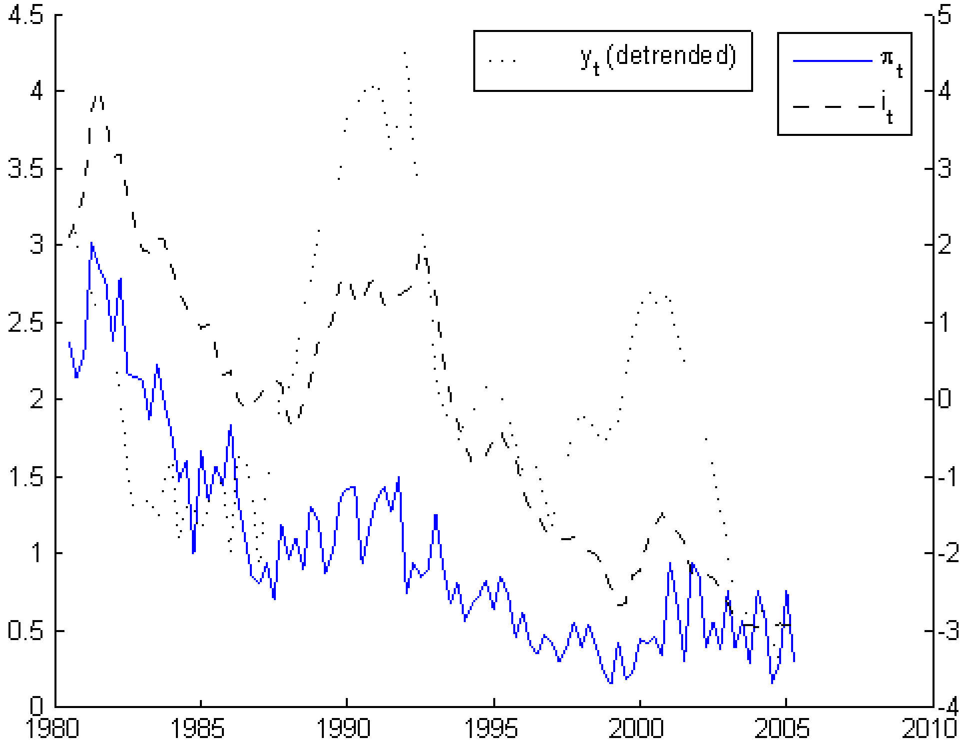
Note: $\sigma_d = 0$, $\alpha = 0.76$, $\chi = 0.275$, $\iota = 1.0$.

Table 3: ML comparison

Theoretical coverage	M1 Marg. Lik.	M1 Empirical coverage	M2 Marg. Lik.	M2 Empirical coverage	Log-Bayes factor M1 vs. M2
0.1	1160.989366	0.05705	1237.968087	0.0699	76.97872093
0.2	1160.270384	0.14895	1237.696806	0.1684	77.42642225
0.3	1159.985883	0.2553	1237.157829	0.28185	77.17194618
0.4	1159.50129	0.36225	1236.962517	0.41485	77.46122651
0.5	1159.423703	0.4731	1236.948869	0.51825	77.52516665
0.6	1159.144849	0.59175	1236.906854	0.63255	77.76200452
0.7	1158.13787	0.7187	1236.950353	0.7254	78.81248254
0.8	1157.803433	0.8389	1237.019148	0.8172	79.21571472
0.9	1157.794394	0.93685	1237.078919	0.9091	79.28452433

Note: based on the modified Gelfand and Dey approach (see Geweke, 1999): integral is approximated by fitting truncated (using theoretical coverage) multivariate Gaussian to posterior distribution. For computations to be reliable empirical coverage should match theoretical coverage. AT1 estimated using 1980:1-2004:4; AT2 estimated using 1980:2-2004:4.

Figure 1: Data

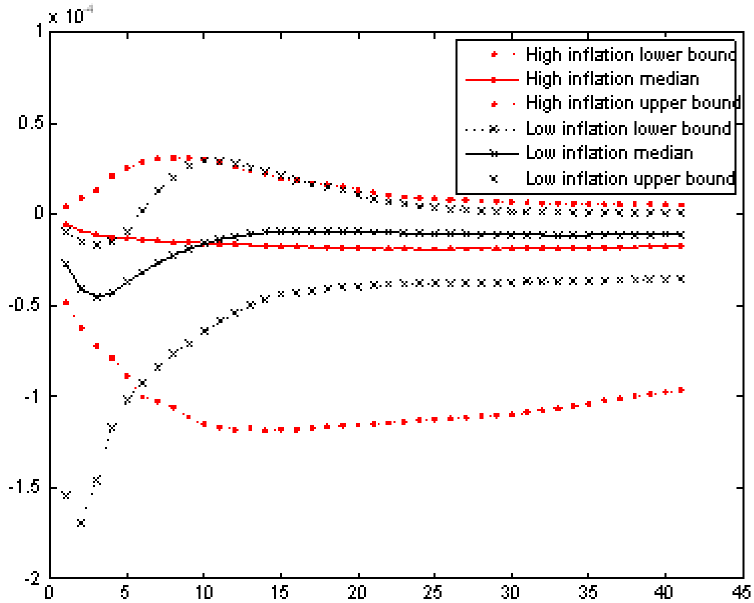


Source: Fagan, Henry and Mestre (2005). Inflation and the nominal interest rate are displayed on the left vertical axis; detrended output on the right vertical axis.

Figure 2: Response of inflation to a standard technology shock: MODEL1

responses of inflation to shock

1

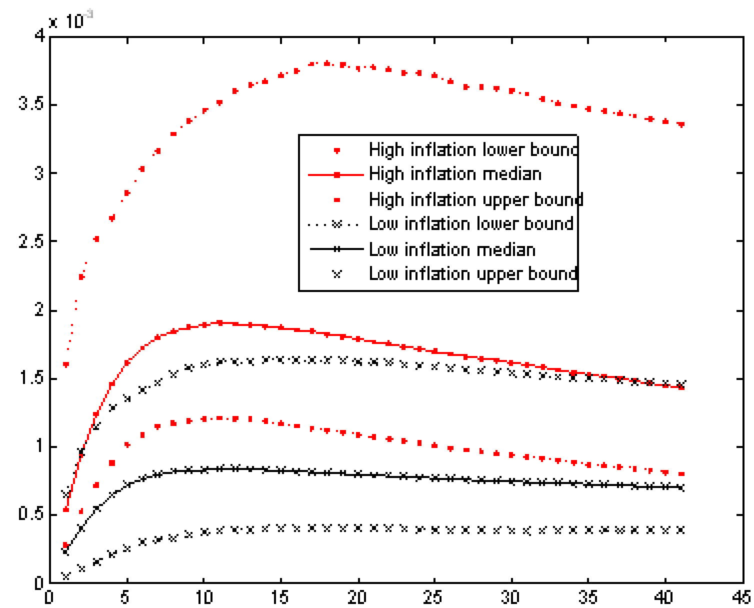


2.pdf

Figure 3: Response of inflation to a standard target shock: MODEL1

responses of inflation to shock

2



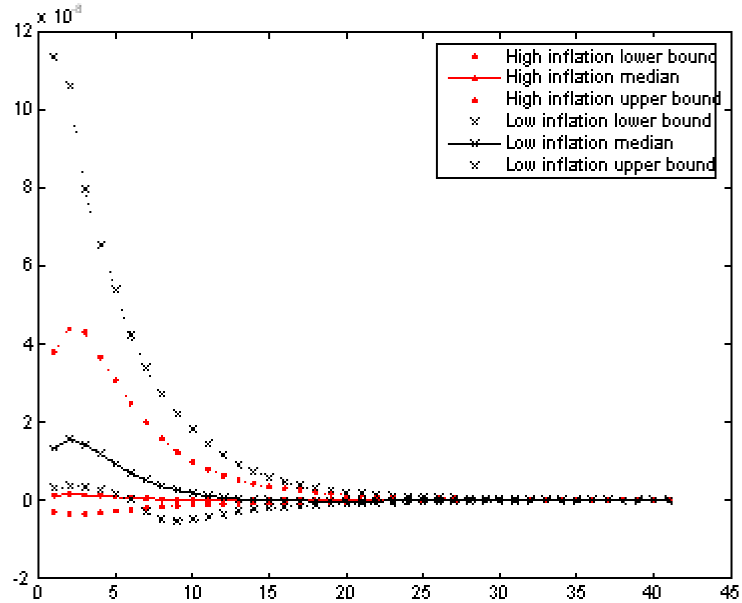
3.pdf

Note: "High inflation" indicates that the impulse response is computed starting from the values of the state vector observed in 1980Q1; "Low inflation" is computed starting from those observed in 2004Q4. Confidence bounds are are at the 95 percent level.

Figure 4: Response of inflation to a standard cost-push shock: MODEL1

responses of inflation to shock

3

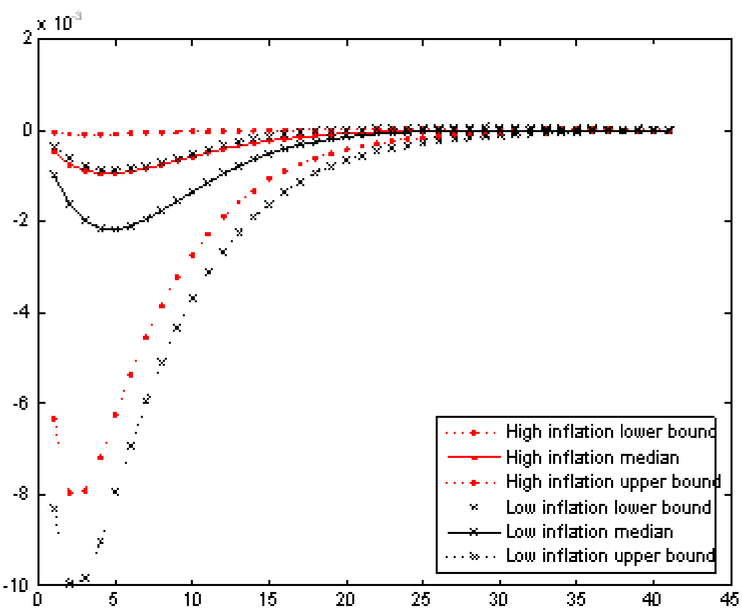


4.pdf

Figure 5: Response of inflation to a standard policy shock: MODEL1

responses of inflation to shock

4



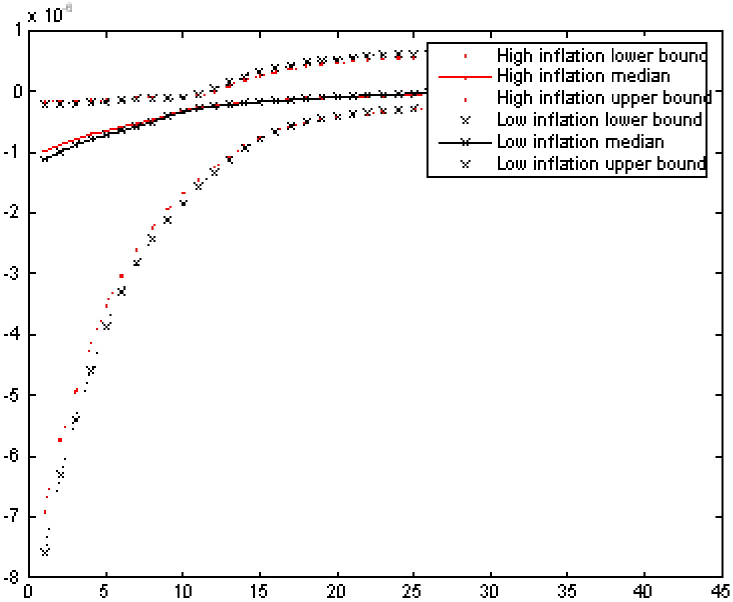
5.pdf

Note: "High inflation" indicates that the impulse response is computed starting from the values of the state vector observed in 1980Q1; "Low inflation" is computed starting from those observed in 2004Q4. Confidence bounds are are at the 95 percent level.

Figure 6: Response of inflation to a standard technology shock: MODEL2

responses of inflation to shock

1

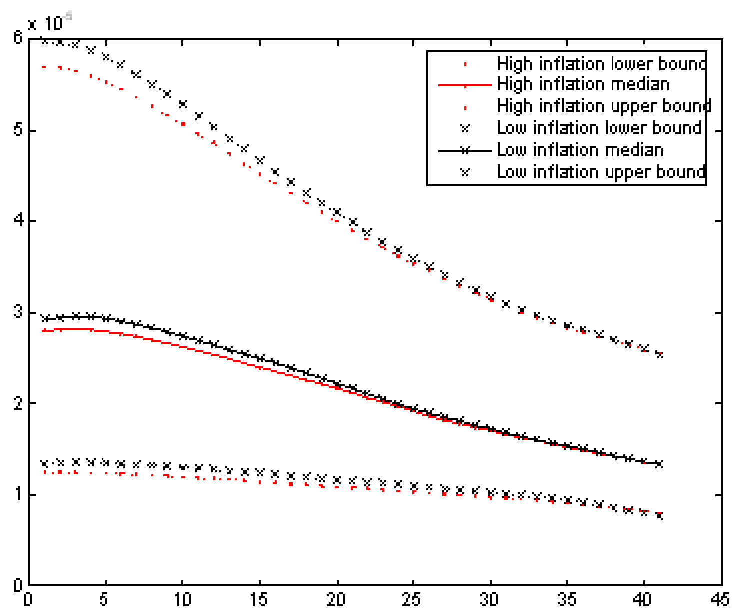


6.pdf

Figure 7: Response of inflation to a standard target shock: MODEL2

responses of inflation to shock

2



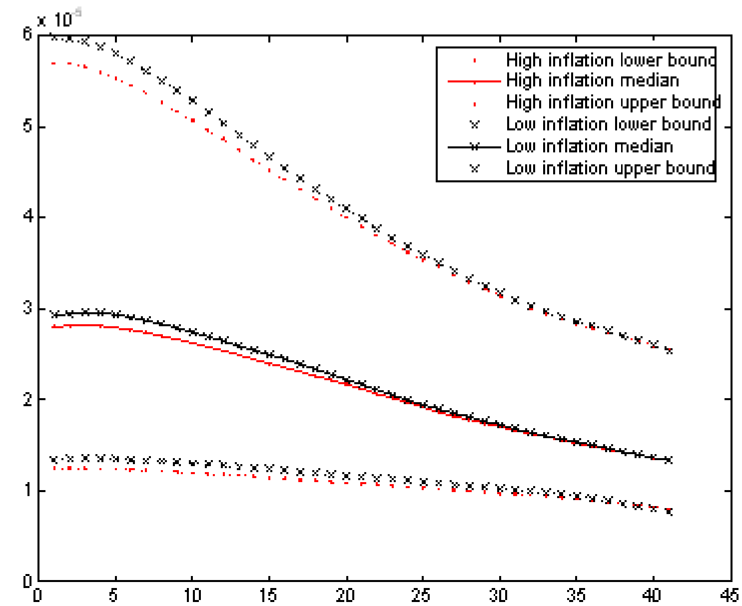
7.pdf

Note: "High inflation" indicates that the impulse response is computed starting from the values of the state vector observed in 1980Q1; "Low inflation" is computed starting from those observed in 2004Q4. Confidence bounds are are at the 95 percent level.

Figure 8: Response of inflation to a standard cost-push shock: MODEL2

responses of inflation to shock

3

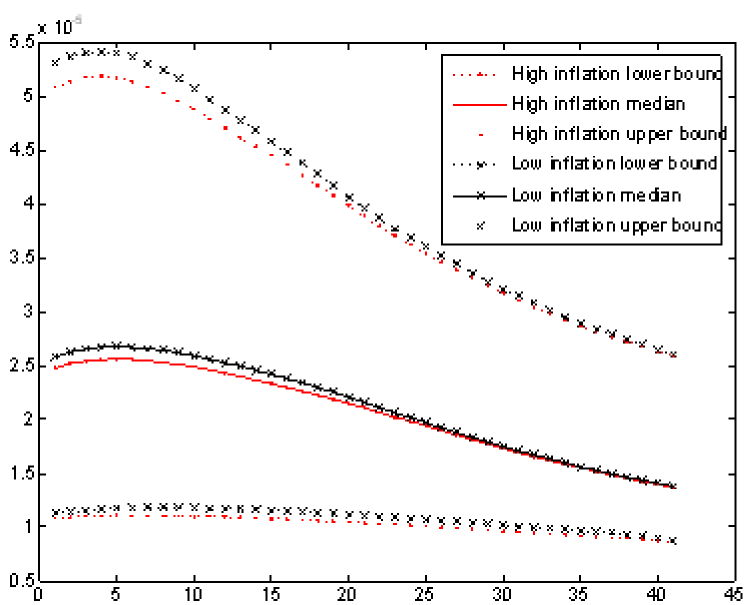


8.pdf

Figure 9: Response of inflation to a standard policy shock: MODEL2

responses of inflation to shock

4



9.pdf

Note: "High inflation" indicates that the impulse response is computed starting from the values of the state vector observed in 1980Q1; "Low inflation" is computed starting from those observed in 2004Q4. Confidence bounds are are at the 95 percent level.

A Appendix

A.1 The complete model(s)

The models are composed of the following equations

$$
\begin{aligned}
K_{2,t} &= \frac{\alpha(\theta-1)}{\phi\chi\theta}\left(\frac{1-\zeta\left(\frac{\Pi_{t-1}}{\Pi_t}\right)^{1-\theta}}{(1-\zeta)}\right)^{1+\frac{\theta}{1-\theta}\frac{\phi}{\alpha}} K_{1,t} \\
K_{2,t} &= (1+\tau_t)\frac{A_t^{-\frac{\phi}{\alpha}}}{\lambda_t}Y_t^{\frac{\phi}{\alpha}} + \mathrm{E}_t\zeta Q_{t,t+1}K_{2,t+1}\Pi_t^{-\theta\frac{\phi}{\alpha}}\Pi_{t+1}^{1+\theta\frac{\phi}{\alpha}} \\
K_{1,t} &= Y_t + \mathrm{E}_t\zeta Q_{t,t+1}K_{1,t+1}\Pi_t^{(1-\theta)}\Pi_{t+1}^{\theta} \\
(Y_t^n)^{\frac{\phi-\alpha}{\alpha}} &= \frac{\alpha}{\phi\chi\mu(1+\tau_t)}A_t^{\frac{\phi}{\alpha}}\left(\left(Y_t^n - hY_{t-1}^n\right)^{-\gamma} - \beta h\mathrm{E}_t\left[\left(Y_{t+1}^n - hY_t^n\right)^{-\gamma}\right]\right) \\
\frac{1}{R_t} &= E_t\left[\beta\left(\frac{Y_{t+1}^n}{Y_t^n}\right)^{\frac{\phi-\alpha}{\alpha}}\left(\frac{A_{t+1}}{A_t}\right)^{-\frac{\phi}{\alpha}}\frac{1+\tau_{t+1}}{1+\tau_t}\right] \\
\lambda_t &= (Y_t - hY_{t-1})^{-\gamma} - \beta h\mathrm{E}_t\left[(Y_{t+1} - hY_t)^{-\gamma}\right] \\
Q_{t,t+1} &= \beta\frac{1}{\Pi_{t+1}}\frac{\lambda_{t+1}}{\lambda_t} \\
\frac{1}{I_t} &= E_t\left(Q_{t,t+1}\right) \\
A_{t+1} &= \overline{A}^{1-\rho}A_t^{\rho}e^{v_{t+1}} \\
\tau_t &= (1-\rho_\tau)\overline{\tau} + \rho_\tau\tau_{t-1} + v_t^{\tau}
\end{aligned}
$$

plus either of the policy rules (12)-(13) or (14)-(15).

In the case of M1 the solution is standard. For M2, we first remove the stochastic trend from nominal variables. More precisely, we define the detrended variables

$$
\begin{aligned}
\widetilde{\Pi}_t^* &\equiv \frac{\Pi_t^*}{\Pi_t} \\
\widetilde{I}_t &\equiv \frac{I_t}{\Pi_t} \\
\widetilde{Q}_{t,t+1} &\equiv Q_{t,t+1}\Pi_{t+1}
\end{aligned}
$$

and rewrite the system as

$$
\begin{aligned}
K_{2,t} &= \frac{\alpha(\theta-1)}{\phi\chi\theta}\left(\frac{1-\zeta\left(\frac{1}{\Delta\Pi_t}\right)^{1-\theta}}{(1-\zeta)}\right)^{1+\frac{\theta}{1-\theta}\frac{\phi}{\alpha}} K_{1,t} \\
K_{2,t} &= (1+\tau_t)\frac{A_t^{-\frac{\phi}{\alpha}}}{\lambda_t}Y_t^{\frac{\phi}{\alpha}} + \mathrm{E}_t\zeta\widetilde{Q}_{t,t+1}K_{2,t+1}(\Delta\Pi_{t+1})^{\theta\frac{\phi}{\alpha}} \\
K_{1,t} &= Y_t + \mathrm{E}_t\widetilde{Q}_{t,t+1}K_{1,t+1}(\Delta\Pi_{t+1})^{\theta-1} \\
(Y_t^n)^{\frac{\phi-\alpha}{\alpha}} &= \frac{\alpha}{\phi\chi\mu(1+\tau_t)}A_t^{\frac{\phi}{\alpha}}\left(\left(Y_t^n-hY_{t-1}^n\right)^{-\gamma}-\beta h\mathrm{E}_t\left[\left(Y_{t+1}^n-hY_t^n\right)^{-\gamma}\right]\right) \\
\frac{1}{R_t} &= E_t\left[\beta\left(\frac{Y_{t+1}^n}{Y_t^n}\right)^{\frac{\phi-\alpha}{\alpha}}\left(\frac{A_{t+1}}{A_t}\right)^{-\frac{\phi}{\alpha}}\frac{1+\tau_{t+1}}{1+\tau_t}\right] \\
\lambda_t &= (Y_t-hY_{t-1})^{-\gamma}-\beta h\mathrm{E}_t\left[(Y_{t+1}-hY_t)^{-\gamma}\right] \\
\widetilde{Q}_{t,t+1} &= \beta\frac{\lambda_{t+1}}{\lambda_t} \\
\frac{1}{\widetilde{I}_t} &= E_t\left(\widetilde{Q}_{t,t+1}\frac{1}{\Delta\Pi_{t+1}}\right) \\
\widetilde{I}_t &= \left(\frac{1}{\beta}\left(\widetilde{\Pi}_t^*\right)^{1-\psi_\Pi}\left(\frac{Y_t}{Y_t^n}\right)^{\psi_Y}\right)^{1-\rho_I}\frac{\widetilde{I}_{t-1}^{\rho_I}}{(\Delta\Pi_t)^{\rho_I}}e^{\eta_t} \\
A_{t+1} &= \overline{A}^{1-\rho}A_t^{\rho}e^{v_{t+1}} \\
\widetilde{\Pi}_t^* &= \widetilde{\Pi}_{t-1}^*\frac{1}{\Delta\Pi_t}e^{v_t^\pi} \\
\tau_t &= (1-\rho_\tau)\overline{\tau}+\rho_\tau\tau_{t-1}+v_t^\tau
\end{aligned}
$$

A.2 Model solution

The approximate solution of the model is computed following Klein (2005). First, we collect all first order conditions in a vector function F such that

$$F_t(x_t,\sigma) \equiv \mathrm{E}_t f(z_{t+1}, z_t, x_{t+1}, x_t) = 0$$

where x_t is the vector of (natural logarithms of the) predetermined variables and z_t is the vector of (natural logarithms of the) non-predetermined variables. More specifically, in the case of M1 $x_t = \left[\pi_{t-1}, y_{t-1}^{nat}, y_{t-1}, i_{t-1}, a_t, \pi_t^*, \tau_t, v_t^i\right]'$ and $z_t = \left[y_t^{nat}, k_{1,t}, k_{2,t}, \pi_t, i_t, y_t, \lambda_t\right]'$, while for M2 $x_t = \left[\widetilde{\pi}_{t-1}^*, y_{t-1}^{nat}, y_{t-1}, \widetilde{i}_{t-1}, a_t, v_t^\pi, \tau_t, v_t^i\right]'$ and $z_t = \left[y_t^{nat}, k_{1,t}, k_{2,t}, \Delta\pi_t, \widetilde{i}_t, y_t, \lambda_t, \widetilde{\pi}_t^*\right]'$. In F_t, σ denotes a scalar perturbation parameter, such that the law of motion of the exogenous state variables x_t^{exog} (where $x_t^{exog} = \left[a_t, \pi_t^*, \tau_t, v_t^i\right]'$) can be written as $x_{t+1}^{exog} = c_1^{exog}x_t^{exog} + \sigma\mathbf{v}_{t+1}$, where the variance-covariance matrix of $\mathbf{v}_{t+1}$ is denoted as η.

References

[1] An, S. and F. Schorfeide (2005), "Bayesian analysis of DSGE models," CEPR Discussion Paper 5207.

[2] Angeloni, I., L. Aucremanne, M. Ehrmann, J Galí, A. Levin and F. Smets (2005) "Inflation persistence in the euro area: preliminary summary of findings," mimeo.

[3] Arulampalam, M.S., S. Maskell, N. Gordon and Tim Clapp (2002), "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking", IEEE Transactions on Signal Processing, 50, 2002, 174-188.

[4] Aruoba, Fernandez-Villaverde and Rubio-Ramirez (2006), "Comparing Solution Methods for Dynamic Equilibrium Economies", Journal of Economic Dynamics and Control, forthcoming.

[5] Ascari, G. (2004), "Staggered Prices and Trend Inflation: Some Nuisances," Review of Economic Dynamics 7, 642-667.

[6] Benigno, P. and M. Woodford (2005), "Inflation stabilisation and welfare: the case of a distorted steady state," mimeo.

[7] Bilke, L. (2004) "Break in the mean and persistence of inflation: a sectoral analysis of French CPI," mimeo.

[8] Calvo, G. A. (1983), "Staggered prices in a utility-maximising framework," Journal of Monetary Economics 12, 983-98.

[9] Canova, F. and L. Sala (2005), "Back to square one: identification issues in DSGE models," mimeo, June.

[10] Chib S. (2001), "Markov Chain Monte Carlo Methods: Computation and Inference", in Heckman, J.J. and E. Leamer (*eds.*): *Handbook of Econometrics*, Vol. 5, Noth Holland.",

[11] Christiano, L., M. Eichenbaum and C. L. Evans, (2005), "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," Journal of Political Economy 113, 1-45.

[12] Clarida, R., J. Galí and M. Gertler, 1998, Monetary policy rules in practice: Some international evidence, European Economic Review, 42, 1033-1067.

[13] Clarida, R., J. Galí and M. Gertler (2000), "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," Quarterly Journal of Economics, 147-180.

[14] Corvoisier, S. and B. Mojon (2004), "Breaks in the mean of inflation: how they happen and what to do with them," mimeo.

[15] Dossche, M. and G. Everaert (2004), "Measuring inflation persistence: a structural time-series approach," mimeo.

[16] Doucet, A., N. de Freitas and N. Gordon (2001), eds., Sequential Monte Carlo Methods in Practice, New York: Springer-Verlag.

[17] Fagan, G., J. Henry and R. Mestre (2005), "An area-wide model (AWM) for the euro area," Economic Modelling 22, 39-59.

[18] Fernandez-Villaverde J. and J. F. Rubio-Ramirez (2004), "Estimating macroeconomic models: a likelihood approach," mimeo.

[19] Fernandez-Villaverde J., J.F. Rubio-Ramirez and M.S. Santos (2006), "Convergence properties of the likelihood of computed dynamic models," Econometrica 74, 93-119.

[20] Geweke J. and W. McCausland (2002), ""

[21] Hamilton, J. (1994), Time series analysis, (Princeton University Press, Princeton).

[22] Kim, J., S. Kim, E. Schaumburg and C. Sims (2003), "Calculating and using second order accurate solutions of discrete time dynamic equilibrium models", *mimeo*, August.

[23] Klein, Paul (2000), "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model," Journal Of Economic Dynamics and Control 24, 1405-23.

[24] Klein, Paul (2005), "Second-order approximation of dynamic models without the use of tensors," University of Western Ontario, *mimeo*, August.

[25] Levin, A. and J. Piger (2004), "Is inflation persistence intrinsic in industrial economies," ECB Working Paper 334.

[26] Liu, J. and M. West (2001), "Combined parameter and state estimation in simulation-based filtering," in Doucet *et al.*, 2001, Chapter 10, 197-223.

[27] Schmitt-Grohé and Uribe (2004), "Solving dynamic general equilibrium models using a second-order approximation to the policy function," Journal of Economic Dynamics and Control 28, 755-75.

[28] Schorfeide, F. (2005), "Learning and Monetary Policy Shifts," Review of Economic Dynamics 8, 392-419.

[29] Smets, F. and R. Wouters (2003), "An estimated stochastic dynamic general equilibrium model of the euro area," Journal of European Economic Association 1, 1123-75

[30] Smets, F. and R. Wouters (2005), "Comparing shocks and frictions in US and euro area business cycles: a Bayesian DSGE approach," Journal of Applied Econometrics 20, forthcoming.

[31] Steinsson, J. (2003), "Optimal monetary policy in an economy with inflation persistence," Journal of Monetary Economics 50, 1425-56.

[32] Sutherland, A. (2002), "A simple second-order solution method for dynamic general equilibrium models," CEPR Discussion Paper No. 3554.

[33] Woodford, M. (2003), Interest and Prices, Princeton University Press.

[34] Yun, T. (1996), "Nominal price rigidity, money supply endogeneity, and business cycles," Journal of Monetary Economics 37, 345-370