Exact likelihood computation for nonlinear DSGE models with heteroskedastic innovations*

Gianni Amisano†
European Central Bank
Università di Brescia

Oreste Tristani‡
European Central Bank

Very preliminary: January 2010

Abstract

Phenomena such as the Great Moderation have increased the attention of macroeconomists towards models where shock processes are not (log-)normal. This paper studies a class of discrete-time rational expectations models where the variance of exogenous innovations is subjecto to stochastic regime shifts. We first show that, up to a second-order approximation using perturbation methods, regime switching has an impact only on the intercept coefficients of the decision rules. We then demonstrate how to derive the exact model likelihood when there are as many shocks as observable variables. We illustrate the applicability of the proposed solution and estimation methods in the case of a small DSGE model.

JEL classification: E0, C63
Keywords: DSGE Models, Second-Order Approximation, regime switching.A DSGE model of the term structure with regime shifts

*The opinions expressed are personal and should not be attributed to the European Central Bank.

†Address: European Central Bank, DG Research, Kaiserstrasse 29, D - 60311, Frankfurt, Germany. E-mail: amisano@eco.unibs.it, gianni.amisano@ecb.int.

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

1 Introduction

The estimation of the first-order approximations of the solution of nonlinear DSGE (dynamic stochastic general equilibrium) models is now commonplace in macroeconomics. By and large, applications tend to rely on the assumption that the state vector of the system is hit by i.i.d. innovations.

The estimation of higher order approximations of the solution of DSGE models has proven more difficult. The problem is that the shape of the likelihood function is unknown and must be approximated through numerical methods. While some general methods have been proposed in this context (see e.g. Fernandez-Villaverde and Rubio-Ramirez, 2006), these methods require a considerable effort from the computational viewpoint. The computational burder increases further when the assumption of i.i.d. innovations is relaxed.

A number of authors, however, have argued forcefully for the presence of heteroskedasticity in macroeconomic data. For example, the literature on the so-called Great moderation (see e.g. McConnell and Perez-Quiros, 2000) finds evidence for regime-switching in the variance of macroeconomic variables. Evidence for regime-switching is also found in Sims and Zha (2006), while Cogley and Sargent (2005), Primiceri (2005) and Fernandez-Villaverde and Rubio-Ramirez (2006), amongst others, argue in favour of stochastic volatility. With the exception of the last paper, all these applications are based on reduced form models.

In this paper, we propose an exact method to compute the likelihood of nonlinear DSGE models in which the variances of innovations to the state vector can be subject to stochastic regime switches.

We first demonstrate how this class of models can be solved using standard perturbation methods. In general, the solution of DSGE models with regime-switching coefficients requires computationally expensive methods (e.g. Coleman, 1991, Andolfatto and Gomme, 2003, Davig, Leeper and Chung, 2004, use an Euler equation iteration technique). In the case we focus on, however, regime switching only affects the variance of structural shocks. Our case can therefore be tackled using standard perturbation methods (see Judd, 1998, and the references therein, Schmitt-Grohé and Uribe, 2001; Kim, Kim, Schaumburg and Sims, 2003; Gomme and Klein, 2006; Lombardo and Sutherland, 2007). We show that, up to a second-order approximation, the coefficients on the linear and quadratic terms in the state vector of the decision rules are independent of the volatility of the exogenous shocks. The main impact of regime switching is on the constant terms of the decision rules, which become regime-dependent.

We then propose a methodology to compute the likelihood and which ultimately amounts to inverting the observation equation for the unobservable state variables of the model. This methodology can be applied subject to two conditions. The first one is that the number of variables used in estimation must be equal to the number of stochastic innovations (including both structural innovations and measurement errors). We believe that this is a soft restriction, since additional measurement errors can always be introduced when additional variables are included in the econometrician's information set. This restriction does however preclude the use of our methodology in cases where there are more shocks than variables. The second condition which needs to be met for the applicability of our method is that the economic system cannot include unobservable non-stochastic state variables. This implies, for example, that a general macroeconomic model can only be estimated using our proposed method if capital can be treated as an observable variable.

We illustrate the validity of our proposed method with an application to a small DSGE model with nominal rigidities.

2 A general model with heteroskedastic conditional variances

We are interested in a general nonlinear model of the form

$$\mathrm{E}_t \left[f \left(\mathbf{y}_{t+1}, \mathbf{y}_t, \mathbf{x}_{1t+1}, \mathbf{x}_{1t}, \mathbf{x}_{2t+1}, \mathbf{x}_{2t}; \mathbf{s}_t \right) \right] = 0 \tag{1}$$

where E_t is the expectational operator conditional on information available at time t, $\mathbf{y}_t$ represent a vector of non-predetermined variable, $\mathbf{x}_{1t}$ is a vector of endogenous predetermined variables, $\mathbf{x}_{2t}$ are exogenous variables with continuous support, and $\mathbf{s}_t$ is a vector including indicators of discrete regimes. The vectors have lenght n_y, n_{x1}, n_{x2} and n_s, respectively. Note that

$$\mathbf{x}_{2t+1} = \mathbf{A}\mathbf{x}_{2t} + \sigma \mathbf{\Sigma}_{s_t} \boldsymbol{\varepsilon}_{t+1} \tag{2}$$

$$\mathbf{s}_{t+1} = \mathbf{B}\mathbf{s}_t + \boldsymbol{\vartheta} \left(\mathbf{s}_t \right) \boldsymbol{\nu}_{t+1} \tag{3}$$

for known functions $\mathbf{A}$ and $\mathbf{\Sigma}_{s_t}$, mapping $R^{n_{x2}}$ into $R^{n_{x2}}$, and $\mathbf{B}$, mapping R^{n_s} into R^{n_s}. The function $\mathbf{A}$ is such that all the eigenvalues of its first derivative evaluated at the non-stochastic steady state lie within the unit circle. The innovation vector $\boldsymbol{\varepsilon}_{t+1}$ is independently and identically distributed, with zero mean and unit variance. The vector $\boldsymbol{\nu}_{t+1}$ has zero mean and heteroskedastic variance. We assume $\boldsymbol{\varepsilon}_{t+1}$ and $\boldsymbol{\nu}_{t+1}$ to be mutually uncorrelated. The scalar σ is the perturbation parameter.

The key distinguishing feature of the model in equation (1) is the presence of the discrete regimes $\mathbf{s}_t$. This generalises the standard model with Gaussian innovations typically studied in the macroeconomic literature. However, we only allow for discrete states to have an impact on the economy only through changes in the variance of the shocks $\boldsymbol{\varepsilon}_{t+1}$. This assumption allows us to compute the solution through a straightforward extension of standard perturbation methods. At the same time, a formulation with regime-switching conditional variances allows one to analyse a number of interesting extensions of the standard Gaussian model (see e.g. Amisano and Tristani, 2009b).

The states which switch across discrete regimes are grouped in vector $\mathbf{s}_t$. In order to write an n_s-state Markov chain $\mathbf{s}_t$ as in equation (3), we rely on Hamilton (1994) which shows that $\mathbf{s}_t$ must be a vector whose i-th element is equal to 1 if $s_t = i$ and zero otherwise, $\mathbf{B}$ is the transition matrix of the Markov chain, and $\boldsymbol{\nu}_{t+1} \equiv \mathbf{s}_{t+1} - \mathrm{E}\left(\mathbf{s}_{t+1}|\mathbf{s}_t\right)$.

Going back to the general model (1)-(3), we can define new vectors

$$\mathbf{x}_t' \equiv \left[\mathbf{x}_{1t}', \mathbf{x}_{2t}'\right], \mathbf{u}_t' \equiv \left[\mathbf{0}', \varepsilon_t'\right]$$

and matrix

$$\widetilde{\mathbf{\Sigma}}_{s_t} \equiv \begin{bmatrix} \underset{n_{x1}\times n_{x1}}{\mathbf{0}} & \underset{n_{x1}\times n_{x2}}{\mathbf{0}} \\ \underset{n_{x2}\times n_{x1}}{\mathbf{0}} & \mathbf{\Sigma}_{s_t} \end{bmatrix}$$

the model in equation (1) can be rewritten as

$$\mathrm{E}_t\left[f\left(\mathbf{y}_{t+1}, \mathbf{y}_t, \mathbf{x}_{t+1}, \mathbf{x}_t; \mathbf{s}_t\right)\right] = \mathbf{0}$$

and the solution is of the form

$$\mathbf{y}_t = g\left(\mathbf{x}_t, \sigma; \mathbf{s}_t\right) \quad (4)$$

$$\mathbf{x}_{t+1} = h\left(\mathbf{x}_t, \sigma; \mathbf{s}_t\right) + \sigma \widetilde{\mathbf{\Sigma}}_{s_t} \mathbf{u}_{t+1} \quad (5)$$

3 Approximating the solution

We seek a second-order approximation to the functions $g(\mathbf{x}_t, \sigma; s_t)$ and $h(\mathbf{x}_t, \sigma; s_t)$ around the non-stochastic steady state $\mathbf{x}_t = \overline{\mathbf{x}}_{s_t}$ and $\sigma = 0$. We leave the dependence on s_t outside the Taylor expansion. In other words, we seek for an expansion such that the coefficients of the second order approximate solution are potentially functions of s_t.

To write the approximation, we follow Gomme and Klein (2006) and use the following representation (from Magnus and Neudecker, 1999) of the second-order Taylor expansion of a twice-differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$

$$f(\mathbf{x}) \approx f(\overline{\mathbf{x}}; s_t) + [\mathrm{D}f(\overline{\mathbf{x}}; s_t)](\mathbf{x} - \overline{\mathbf{x}}_{s_t}) + \frac{1}{2}\left(\mathbf{I}_m \otimes (\mathbf{x} - \overline{\mathbf{x}}_{s_t})'\right)[\mathrm{H}f(\overline{\mathbf{x}}; s_t)](\mathbf{x} - \overline{\mathbf{x}}_{s_t})$$

where $\mathrm{D}f(\overline{\mathbf{x}}; s_t)$ and $\mathrm{H}f(\overline{\mathbf{x}}; s_t)$ are gradient and Hessian matrices structured as follows

$$\mathrm{D}f(\overline{\mathbf{x}}; s_t) \equiv \begin{bmatrix} \frac{\partial f_1(\overline{\mathbf{x}}; s_t)}{\partial x_1} & \frac{\partial f_1(\overline{\mathbf{x}}; s_t)}{\partial x_2} & \cdots & \frac{\partial f_n(\overline{\mathbf{x}}; s_t)}{\partial x_n} \\ \frac{\partial f_2(\overline{\mathbf{x}}; s_t)}{\partial x_1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_m(\overline{\mathbf{x}}; s_t)}{\partial x_1} & \cdots & \cdots & \frac{\partial f_m(\overline{\mathbf{x}}; s_t)}{\partial x_n} \end{bmatrix}$$

and

$$\mathrm{H}f(\overline{\mathbf{x}}; s_t) \equiv \mathrm{D}\ \mathrm{vec}\left[(\mathrm{D}f(\overline{\mathbf{x}}; s_t))'\right]$$

Using the solution (4)-(5), we can rewrite model (1) as the function

$$F(\mathbf{x}_t, \sigma; s_t) = \mathrm{E}_t\left[f\left\{\begin{array}{c} h(\mathbf{x}_t, \sigma; s_t) + \sigma\widetilde{\boldsymbol{\Sigma}}_{s_t}\mathbf{u}_{t+1}, g\left[h(\mathbf{x}_t, \sigma; s_t) + \sigma\widetilde{\boldsymbol{\Sigma}}_{s_t}\mathbf{u}_{t+1}, \sigma; \mathbf{B}\mathbf{s}_t + \boldsymbol{\vartheta}(\mathbf{s}_t)\boldsymbol{\nu}_{t+1}\right], \\ \mathbf{x}_t, g(\mathbf{x}_t, \sigma; s_t); s_t \end{array}\right\}\right] = \mathbf{0}$$

We analyse the second-order approximation to the functions h and g which can be represented as

$$g(\mathbf{x}_t, \sigma; s_t) = g(\overline{\mathbf{x}}; 0; s_t) + \mathbf{F}_{s_t}(\mathbf{x}_t - \overline{\mathbf{x}}_{s_t}) + \frac{1}{2}\left(\mathbf{I}_{n_y} \otimes (\mathbf{x}_t - \overline{\mathbf{x}}_{s_t})'\right)\mathbf{E}_{s_t}(\mathbf{x}_t - \overline{\mathbf{x}}_{s_t}) + \mathbf{k}_{y,s_t}\sigma^2$$

and

$$h(\mathbf{x}_t, \sigma; s_t) = h(\overline{\mathbf{x}}; 0; s_t) + \mathbf{P}_{s_t}(\mathbf{x}_t - \overline{\mathbf{x}}_{s_t}) + \frac{1}{2}\left(\mathbf{I}_{n_x} \otimes (\mathbf{x}_t - \overline{\mathbf{x}}_{s_t})'\right)\mathbf{G}_{s_t}(\mathbf{x}_t - \overline{\mathbf{x}}_{s_t}) + \mathbf{k}_{x,s_t}\sigma^2$$

for potentially state-dependent vectors and matrices $\mathbf{F}_{s_t}$, $\mathbf{E}_{s_t}$, $\mathbf{P}_{s_t}$, $\mathbf{G}_{s_t}$, $\mathbf{k}_{y,s_t}$, $\mathbf{k}_{x,s_t}$.

3.1 Steady state

By definition, $\overline{\mathbf{y}}(s_t) = g(\overline{\mathbf{x}}, 0; s_t)$ and $\overline{\mathbf{x}}_{s_t} = h(\overline{\mathbf{x}}, 0; s_t)$. In general, the steady state of a model with regime switches would be a function of the discrete regimes s_t. Given our assumption that discrete regimes only have an impact on (1) through the variance of the innovations, however, the steady state which arises when $\sigma = 0$ is not regime-dependent. Thus $\overline{\mathbf{x}}_{s_t} = \overline{\mathbf{x}}$ and $\overline{\mathbf{y}}(s_t) = \overline{\mathbf{y}}$. We can therefore simplify the form of the second order approximations as

$$g(\widehat{\mathbf{x}}_t, \sigma; s_t) = \overline{\mathbf{y}} + \mathbf{F}_{s_t}\widehat{\mathbf{x}}_t + \frac{1}{2}\left(\mathbf{I}_{n_y} \otimes \widehat{\mathbf{x}}_t^{\prime}\right)\mathbf{E}_{s_t}\widehat{\mathbf{x}}_t + \mathbf{k}_{y,s_t}\sigma^2$$

and

$$h(\widehat{\mathbf{x}}_t, \sigma; s_t) = \overline{\mathbf{x}} + \mathbf{P}_{s_t}\widehat{\mathbf{x}}_t + \frac{1}{2}\left(\mathbf{I}_{n_x} \otimes \widehat{\mathbf{x}}_t^{\prime}\right)\mathbf{G}_{s_t}\widehat{\mathbf{x}}_t + \mathbf{k}_{x,s_t}\sigma^2$$

where $\widehat{\mathbf{x}}_t \equiv \mathbf{x}_t - \overline{\mathbf{x}}$.

3.2 First-order approximation

The assumed form of the solution implies that

$$F(\mathbf{x}_t, \sigma; s_t) = \mathrm{E}_t\left[f\left\{\begin{array}{c} P_{s_t}\widehat{\mathbf{x}}_t + \sigma\widetilde{\mathbf{\Sigma}}_{s_t}\mathbf{u}_{t+1}, \mathbf{F}_{s_t}\left[\mathbf{P}_{s_t}\widehat{\mathbf{x}}_t + \sigma\widetilde{\mathbf{\Sigma}}_{s_t}\mathbf{u}_{t+1}, \sigma; \mathbf{B}\mathbf{s}_t + \boldsymbol{\vartheta}(\mathbf{s}_t)\boldsymbol{\nu}_{t+1}\right], \\ \mathbf{x}_t, \mathbf{F}_{s_t}\widehat{\mathbf{x}}_t; s_t \end{array}\right\}\right] = 0$$

In order to identify the coefficients $\mathbf{F}_{s_t}$ and $\mathbf{P}_{s_t}$, we exploit the property that the solution must be such that $\mathrm{D}F_1(\overline{\mathbf{x}}, 0; s_t) = 0$, where a subscript i indicates differentiation with respect to vector i in a given function. Using the definition of function $F(x, \sigma; s_t)$, we obtain

$$\mathrm{D}F_1(\overline{\mathbf{x}}, 0; s_t) = \mathbf{f}_1\mathbf{P}_{s_t} + \mathbf{f}_2\mathbf{F}_{s_t}\mathbf{P}_{s_t} + \mathbf{f}_3 + \mathbf{f}_4\mathbf{F}_{s_t} = [\mathbf{0}]$$

which is a (potentially regime dependent) system of quadratic equations in the elements of $\mathbf{F}_{s_t}$ and $\mathbf{P}_{s_t}$.

Note that this equation implies that

$$\mathbf{f}_1\mathbf{P}_{s_t}\mathrm{E}_t\widehat{\mathbf{x}}_{t+1} + \mathbf{f}_2\mathbf{F}_{s_t}\mathbf{P}_{s_t}\mathrm{E}_t\widehat{\mathbf{y}}_{t+1} + \mathbf{f}_3\widehat{\mathbf{x}}_t + \mathbf{f}_4\mathbf{F}_{s_t}\widehat{\mathbf{y}}_t = [\mathbf{0}]$$

or

$$\left[\begin{array}{cc} \mathbf{f}_1\mathbf{P}_{s_t} & \mathbf{f}_2\mathbf{F}_{s_t}\mathbf{P}_{s_t} \end{array}\right]\left[\begin{array}{c} \mathrm{E}_t\widehat{\mathbf{x}}_{t+1} \\ \mathrm{E}_t\widehat{\mathbf{y}}_{t+1} \end{array}\right] + \left[\begin{array}{cc} \mathbf{f}_3 & \mathbf{f}_4\mathbf{F}_{s_t} \end{array}\right]\left[\begin{array}{c} \widehat{\mathbf{x}}_t \\ \widehat{\mathbf{y}}_t \end{array}\right] = [\mathbf{0}]$$

Using

$$\begin{aligned} \mathrm{E}_t\widehat{\mathbf{x}}_{t+1} &= \begin{bmatrix} \mathrm{E}_t\widehat{\mathbf{x}}_{1t+1} \\ \mathrm{E}_t\widehat{\mathbf{x}}_{2t+1} \end{bmatrix} = \begin{bmatrix} \widehat{\mathbf{x}}_{1t+1} \\ \widehat{\mathbf{x}}_{2t+1} \end{bmatrix} - \sigma\widetilde{\mathbf{\Sigma}}_{s_t} \begin{bmatrix} \mathbf{0} \\ \varepsilon_{t+1} \end{bmatrix} \\ &= \widehat{\mathbf{x}}_{t+1} - \sigma\widetilde{\mathbf{\Sigma}}_{s_t}\mathbf{u}_t \end{aligned}$$

we can further rewrite this as

$$\begin{bmatrix} \mathbf{f}_1\mathbf{P}_{s_t} & \mathbf{f}_2\mathbf{F}_{s_t}\mathbf{P}_{s_t} \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{x}}_{t+1} \\ \mathrm{E}_t\widehat{\mathbf{y}}_{t+1} \end{bmatrix} = \begin{bmatrix} -\mathbf{f}_3 & -\mathbf{f}_4\mathbf{F}_{s_t} \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{x}}_t \\ \widehat{\mathbf{y}}_t \end{bmatrix} + \sigma \begin{bmatrix} \mathbf{f}_1\mathbf{P}_{s_t} & \mathbf{f}_2\mathbf{F}_{s_t}\mathbf{P}_{s_t} \end{bmatrix} \begin{bmatrix} \mathbf{u}_t \\ \mathbf{0} \end{bmatrix}$$

or

$$\mathbf{A}_{s_t} \begin{bmatrix} \widehat{\mathbf{x}}_{t+1} \\ \mathrm{E}_t\widehat{\mathbf{y}}_{t+1} \end{bmatrix} = \mathbf{B}_{s_t} \begin{bmatrix} \widehat{\mathbf{x}}_t \\ \widehat{\mathbf{y}}_t \end{bmatrix} + \sigma\mathbf{C}_{s_t} \begin{bmatrix} \mathbf{u}_t \\ \mathbf{0} \end{bmatrix}$$

By assumption, $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$ and $\mathbf{f}_4$ are not regime dependent in system (1). Hence, the solution such that the $\mathbf{A}$ and $\mathbf{B}$ matrices are regime independent is a solution of the system. This solution can be obtained using standard algorithms (see Blanchard and Kahn, 1980, Sims, 2001, Klein, 2000, Söderlind, 1999). The conditions for the local uniqueness of the solution are also unchanged.

Our first-order approximation should also include conditions for which $\mathrm{D}F_2(\overline{z}, 0) = [\mathbf{0}]$. As in Schmitt-Grohé and Uribe (2001), the resulting equations would be linear and homogeneous in the coefficients attached to σ, which implies that those coefficients must be zero.

Hence, to a first-order approximation, the solution is not affected by the variance of the shocks, and specifically the variance of the markov-switching states. This result is consistent with the loglinearisation used by Schorfheide (2005).

3.3 Second-order approximation

The assumed solutions imply that (exploiting the finding that $\mathbf{P}_{s_t} = \mathbf{P}$ and $\mathbf{F}_{s_t} = \mathbf{F}$)

$$\begin{aligned} &F(\mathbf{x}_t, \sigma; s_t) \\ = \;& \mathrm{E}_t \left[f \left\{ \begin{array}{c} \mathbf{P}\widehat{\mathbf{x}}_t + \frac{1}{2}(\mathbf{I}_{n_x} \otimes \widehat{\mathbf{x}}_t')\mathbf{G}_{s_t}\widehat{\mathbf{x}}_t + \mathbf{k}_{x,s_t}\sigma^2 + \sigma\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1}, \\ \mathbf{F}\left[\mathbf{P}\widehat{\mathbf{x}}_t + \frac{1}{2}(\mathbf{I}_{n_x} \otimes \widehat{\mathbf{x}}_t')\mathbf{G}_{s_t}\widehat{\mathbf{x}}_t + \mathbf{k}_{x,s_t}\sigma^2 + \sigma\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1}, \sigma; \mathbf{B}\mathbf{s}_t + \boldsymbol{\nu}_{t+1}\right] \\ +\frac{1}{2}\left(\mathbf{I}_{n_y} \otimes \left[\mathbf{P}\widehat{\mathbf{x}}_t + \frac{1}{2}(\mathbf{I}_{n_x} \otimes \widehat{\mathbf{x}}_t')\mathbf{G}_{s_t}\widehat{\mathbf{x}}_t + \mathbf{k}_{x,s_t}\sigma^2 + \sigma\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1}, \sigma; \mathbf{B}\mathbf{s}_t + \boldsymbol{\nu}_{t+1}\right]'\right) \cdot \\ \cdot\mathbf{E}_{s_t}\left[\mathbf{P}\widehat{\mathbf{x}}_t + \frac{1}{2}(\mathbf{I}_{n_x} \otimes \widehat{\mathbf{x}}_t')\mathbf{G}_{s_t}\widehat{\mathbf{x}}_t + \mathbf{k}_{x,s_t}\sigma^2 + \sigma\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1}, \sigma; \mathbf{B}\mathbf{s}_t + \boldsymbol{\nu}_{t+1}\right] + \mathbf{k}_{y,s_t}\sigma^2, \\ \mathbf{x}_t, F\widehat{\mathbf{x}}_t + \frac{1}{2}\left(I_{n_y} \otimes \widehat{\mathbf{x}}_t'\right)E_{s_t}\widehat{\mathbf{x}}_t + \mathbf{k}_{y,s_t}\sigma^2; s_t \end{array} \right\} \right] = [\mathbf{0}] \end{aligned}$$

Now evaluate $\mathrm{H}_{11}F(\overline{\mathbf{x}},0;s_t)=[\mathbf{0}]$. We obtain

$$\begin{aligned}
\mathrm{H}_{11}F(\overline{\mathbf{x}},\sigma;s_t) &= (\mathbf{I}_m\otimes\mathbf{P})'\,\mathbf{f}_{11}\mathbf{P}+2(\mathbf{I}_m\otimes\mathbf{P})'\,\mathbf{f}_{12}\mathrm{D}_1g(h(\overline{\mathbf{x}},\sigma;s_t)) \\
&\quad +2(\mathbf{I}_m\otimes\mathbf{P})'\,\mathbf{f}_{13}+2(\mathbf{I}_m\otimes\mathbf{P})'\,\mathbf{f}_{14}\mathbf{F} \\
&\quad +(\mathbf{I}_m\otimes\mathrm{D}_1g(h(\overline{\mathbf{x}},\sigma;s_t)))'\,\mathbf{f}_{22}\mathrm{D}_1g(y)+2(\mathbf{I}_m\otimes\mathrm{D}_1g(h(\overline{\mathbf{x}},\sigma;s_t)))'\,\mathbf{f}_{23} \\
&\quad +2(\mathbf{I}_m\otimes\mathrm{D}_1g(h(\overline{\mathbf{x}},\sigma;s_t)))'\,\mathbf{f}_{24}\mathbf{F}+\mathbf{f}_{33}+2\mathbf{f}_{34}\mathbf{F}+(\mathbf{I}_m\otimes\mathbf{F})'\,\mathbf{f}_{44}\mathbf{F} \\
&\quad +(\mathbf{f}_1\otimes\mathbf{I}_{n_x})(\mathbf{G}_{s_t}+\sigma\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1})+(\mathbf{f}_2\otimes\mathbf{I}_{n_x})\,\mathrm{H}_{11}g(h(\overline{\mathbf{x}},\sigma;s_t))+(\mathbf{f}_4\otimes\mathbf{I}_{n_x})\,\mathbf{E}_{s_t}
\end{aligned}$$

Now note that

$$\mathrm{H}_{11}g(h(\overline{\mathbf{x}},\sigma;s_t))=\left(\mathbf{I}_{n_y}\otimes\mathbf{P}'\right)\mathbf{E}_{s_t}\mathbf{P}+(\mathbf{F}\otimes\mathbf{I}_{n_x})\,\mathbf{G}_{s_t}$$

and

$$\mathrm{D}_1g(h(\overline{\mathbf{x}},\sigma;s_t))=\mathbf{F}\mathbf{P}$$

Once these expressions are substituted into $\mathrm{H}_{11}F(\overline{\mathbf{x}},\sigma;s_t)$ above and the latter is evaluated at $\sigma=0$, we find

$$\begin{aligned}
&(\mathbf{f}_1\otimes\mathbf{I}_{n_x})\,\mathbf{G}_{s_t}+(\mathbf{f}_2\otimes\mathbf{I}_{n_x})\left(\left(\mathbf{I}_{n_y}\otimes\mathbf{P}'\right)\mathbf{E}_{s_t}\mathbf{P}+(\mathbf{F}\otimes\mathbf{I}_{n_x})\,\mathbf{G}_{s_t}\right)+(\mathbf{f}_4\otimes\mathbf{I}_{n_x})\,\mathbf{E}_{s_t}+ \\
&+\left(\mathbf{I}_m\otimes\mathbf{P}'\right)\mathbf{f}_{11}\mathbf{P}+\left(\mathbf{I}_m\otimes\mathbf{P}'\mathbf{F}'\right)\mathbf{f}_{22}\mathbf{F}\mathbf{P}+\mathbf{f}_{33}+\left(\mathbf{I}_m\otimes\mathbf{F}'\right)\mathbf{f}_{44}\mathbf{F}+ \\
&+2\left(\mathbf{I}_m\otimes\mathbf{P}'\right)\mathbf{f}_{12}\mathbf{F}\mathbf{P}+2\left(\mathbf{I}_m\otimes\mathbf{P}'\right)\mathbf{f}_{13}+2\left(\mathbf{I}_m\otimes\mathbf{P}'\right)\mathbf{f}_{14}\mathbf{F}+ \\
&+2\left(\mathbf{I}_m\otimes\mathbf{P}'\mathbf{F}'\right)\mathbf{f}_{23}+2\left(\mathbf{I}_m\otimes\mathbf{P}'\mathbf{F}'\right)\mathbf{f}_{24}\mathbf{F}+2\mathbf{f}_{34}\mathbf{F} \\
&=[\mathbf{0}]
\end{aligned}$$

This is a linear equation which can be solved for $\mathbf{E}_{s_t}$ and $\mathbf{G}_{s_t}$. Note that, as in the case of the first-order approximation, all coefficients in the Hessians $\mathbf{f}_{ij}$ are constant by assumption. Hence, a solution such that the $\mathbf{E}$ and $\mathbf{G}$ matrices are regime independent is a solution of the system. In other words, $\mathbf{E}$ and $\mathbf{G}$ will be identical to the case with homoskedastic shocks.

Now consider the second derivative with respect to σ, namely $\mathrm{H}_{22}\mathbf{F}$. We obtain

$$\begin{aligned}
\mathrm{H}_{22}F &= (\mathbf{I}_m\otimes\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1})'\,\mathbf{f}_{11}\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1}+2(\mathbf{I}_m\otimes\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1})'\,\mathbf{f}_{12}\mathrm{D}_2g(h(\overline{\mathbf{x}},\sigma;s_t)) \\
&\quad +\left(\mathbf{I}_{n_y}\otimes\mathrm{D}_2g(h(\overline{\mathbf{x}},\sigma;s_t))\right)'\mathbf{f}_{22}\mathrm{D}_2g(h(\overline{\mathbf{x}},\sigma;s_t)) \\
&\quad +\mathbf{f}_1\mathbf{k}_z+\mathbf{f}_2\mathrm{H}_{22}g(h(\overline{\mathbf{x}},\sigma;s_t))+\mathbf{f}_4\mathbf{k}_y
\end{aligned}$$

where

$$\mathrm{D}_2g(h(\overline{\mathbf{x}},\sigma;s_t))=\mathbf{F}\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1}$$

and

$$\mathrm{H}_{22}g\left(h\left(\overline{\mathbf{x}},\sigma;s_t\right)\right)=\left(\mathbf{I}_{n_y}\otimes\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1}\right)'\mathbf{E}\left(\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1}\right)+\mathbf{k}_y+\mathbf{F}\mathbf{k}_z$$

Substitute the expressions for D_2g and $\mathrm{H}_{22}g$ into $\mathrm{H}_{22}F\left(\overline{\mathbf{x}},\sigma;s_t\right)$ and evaluate the result at $\sigma=0$ to find

$$\begin{aligned}\mathrm{H}_{22}F &= \mathbf{f}_1\mathbf{k}_z+\mathbf{f}_2\left(\mathbf{k}_y+\mathbf{F}\mathbf{k}_z+\mathrm{E}_t\left(\mathbf{I}_{n_y}\otimes\mathbf{u}'_{t+1}\mathbf{\Sigma}'_{s_t}\right)\mathbf{E}\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1}\right)+\mathbf{f}_4\mathbf{k}_y\\ &\quad+\mathrm{E}_t\left(\mathbf{I}_m\otimes\mathbf{\Sigma}_{s_t}\widetilde{\mathbf{u}}_{t+1}\right)'\mathbf{f}_{11}\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1}+2\mathrm{E}_t\left(\mathbf{I}_m\otimes\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1}\right)'\mathbf{f}_{12}\mathbf{F}\mathbf{\Sigma}_{s_t}\widetilde{\mathbf{u}}_{t+1}\\ &\quad+\mathrm{E}_t\left(\mathbf{I}_{n_y}\otimes\mathbf{u}'_{t+1}\mathbf{\Sigma}'_{s_t}\mathbf{F}'\right)\mathbf{f}_{22}\mathbf{F}\mathbf{\Sigma}_{s_t}\mathbf{u}_{t+1}\end{aligned}$$

Taking expectations, the elements of vectors $\mathbf{k}_y$ and $\mathbf{k}_z$ are the solution of the following linear equations

$$\begin{aligned}&\mathbf{f}_1\mathbf{k}_z+\mathbf{f}_2\mathbf{k}_y+\mathbf{f}_2\mathbf{F}\mathbf{k}_z+\mathbf{f}_2\mathrm{trm}\left(\mathbf{I}_{n_y}\otimes\left(\mathbf{\Sigma}'_{s_t}\mathbf{\Sigma}_{s_t}\right)\right)\mathbf{E}+\mathbf{f}_4\mathbf{k}_y\\ &+\mathrm{trm}\left(\mathbf{I}_m\otimes\left(\mathbf{\Sigma}'_{s_t}\mathbf{\Sigma}_{s_t}\right)\right)\mathbf{f}_{11}+2\mathrm{trm}\left(\mathbf{I}_m\otimes\left(\mathbf{\Sigma}'_{s_t}\mathbf{\Sigma}_{s_t}\right)\right)\mathbf{f}_{12}\mathbf{F}\\ &+\mathrm{trm}\left(\mathbf{I}_{n_y}\otimes\left(\mathbf{\Sigma}'_{s_t}\mathbf{\Sigma}_{s_t}\mathbf{F}'\right)\right)\mathbf{f}_{22}\mathbf{F}=[\mathbf{0}]\end{aligned}\tag{6}$$

where $\mathbf{\Sigma}'_{s_t}\mathbf{\Sigma}_{s_t}$ is the conditional variance covariance matrix of vector $\mathbf{x}_t$ and where, as in Gomme and Klein (2006), we define the matrix trace (trm) of an $(nm\times n)$ matrix

$$\begin{bmatrix}\mathbf{A}_1\\ \mathbf{A}_2\\ \vdots\\ \mathbf{A}_m\end{bmatrix}$$

as the $m\times 1$ vector

$$\begin{bmatrix}\mathrm{tr}\left(\mathbf{A}_1\right)\\ \mathrm{tr}\left(\mathbf{A}_2\right)\\ \vdots\\ \mathrm{tr}\left(\mathbf{A}_m\right)\end{bmatrix}$$

Due to the presence of the regime-switching terms, the variance-covariance matrix $\mathbf{\Sigma}'_{s_t}\mathbf{\Sigma}_{s_t}$ is regime dependent. Equation (6) can therefore be satisfied only if the coefficients $\mathbf{k}_{y,s_t}$, $\mathbf{k}_{x,s_t}$ assume different values depending on the realisation of the regime-switching states. If there are n_s possible regimes, n_s systems of the form (6) have to be solved, yielding n_s pairs $(\mathbf{k}_{y,s_t},\mathbf{k}_{z,s_t})$.

Finally, along the lines above, one can confirm that the coefficients of terms in $\sigma\widehat{\mathbf{x}}_t$ are zero solving the equation $\mathrm{H}_{12}F=[\mathbf{0}]$.

To summarise, the second order approximation of the policy functions $g(\widehat{\mathbf{x}}_t, \sigma; s_t)$ and $h(\widehat{\mathbf{x}}_t, \sigma; s_t)$ can be written as

$$g(\widehat{\mathbf{x}}_t, \sigma; s_t) = \overline{\mathbf{y}} + \mathbf{F}\widehat{\mathbf{x}}_t + \frac{1}{2}\left(\mathbf{I}_{n_y} \otimes \widehat{\mathbf{x}}_t'\right)\mathbf{E}\widehat{\mathbf{x}}_t + \mathbf{k}_{y,s_t}\sigma^2 \quad (7)$$

$$h(\widehat{\mathbf{x}}_t, \sigma; s_t) = \overline{\mathbf{x}} + \mathbf{P}\widehat{\mathbf{x}}_t + \frac{1}{2}\left(\mathbf{I}_{n_x} \otimes \widehat{\mathbf{x}}_t'\right)\mathbf{G}\widehat{\mathbf{x}}_t + \mathbf{k}_{x,s_t}\sigma^2 \quad (8)$$

Compared to the case with homoskedastic shocks, the only difference in the case of regime-switching is in the $\mathbf{k}_{x,s_t}$ and $\mathbf{k}_{y,s_t}$ vectors. A change in regime only affects the policy rule through a precautionary savings or risk premium component. To a first order approximation, the policy rules are exactly as in the case with homoskedastic shocks.

4 Estimation

In order to better explain the way in which we conduct likelihood inference, let us rewrite the model as

$$\mathbf{y}_{t+1}^o = \mathbf{c}_j + \mathbf{C}_1\mathbf{x}_{t+1} + \mathbf{C}_2\overline{vech}(\mathbf{x}_{t+1}\mathbf{x}_{t+1}') + \mathbf{D}\mathbf{v}_{t+1} \quad (9)$$

$$\mathbf{x}_{t+1} = \mathbf{a}_i + \mathbf{A}_1\mathbf{x}_t + \mathbf{A}_2\overline{vech}(\mathbf{x}_t\mathbf{x}_t') + \mathbf{B}_i\mathbf{w}_{t+1} \quad (10)$$

$$\mathbf{s}_t \backsim \text{Markov switching} \quad (11)$$

where the vector $\mathbf{y}_t^o$ includes all observable variables, and $\mathbf{v}_{t+1}$ and $\mathbf{w}_{t+1}$ are measurement and structural shocks, respectively. In this representation, the regime switching variables affect the system by changing the intercepts $\mathbf{a}_i$ and $\mathbf{c}_j$, and the loadings for the of the structural innovations $\mathbf{B}_i$.(we indicate here with i the value of the discrete state variables at t and with j the value of the discrete state variables at $t+1$).

If a linear approximation were used, we would be left with a linear state space model with Markov switching. See Kim (1994), Kim and Nelson (1999) and Schorfheide (2005). The likelihood cannot be obtained by recursive methods and it is approximated using a discrete mixture approach.

Things are easier when the number of continuous shocks (measurement and structural) is equal to the number of observables. In such a case the continuous latent variables can be obtained via inversion and the system can be written as a Markov Switching VAR. The likelihood can be obtained by using Hamilton's filter i.e. by integrating out the discrete latent variables

In the quadratic case, the likelihood cannot in general be obtained in closed form. One possible approach to compute the likelihood is to rely on sequential Monte Carlo techniques (for an application of these techniques in a DSGE setting, see e.g. Amisano and Tristani, 2009a). These methods, however, are computationally expensive in a case, such as the one of our model, in which both nonlinearities and non-Gaussianity of the shocks characterise the economy.

We thus adopt a simple extension of the filter employed in the linear case when the number of observed variables is equal to the sum of measurement and structural shocks. The only problem in this respect is the quadratic term in $\mathbf{x}_t$ in the observation equation (9). In particular, given that we assume that $\mathbf{x}_{2t}$, the vector of predetermined variables, contains only lagged endogenous variables, the problem in computing the likelihood is generated by the fact that $\mathbf{x}_{2t}$ enters in the reduced form state space representation in a nonlinear way.

We are interested in a particular case which renders likelihood computations quite easy. Let us define

$$\mathbf{z}_{t+1} = \begin{bmatrix} \mathbf{x}_{2t+1} \\ \mathbf{v}_{t+1} \end{bmatrix}$$

the $(n_{x_2}+n_{me})\times 1$ vector containing all continuous latent variables of the models which are not predetermined. When $n_z = n_{x_2} + n_{me} = n_y$, i.e. when there are as many observables as continuous latent variables in the system, the mapping

$$\mathbf{z}_{t+1}^{(j,:)} = \zeta(\mathbf{y}_{t+1}^{o}|s_{t+1} = j, \mathbf{x}_{1t+1}) = \zeta_j(\mathbf{y}_{t+1}^{o}) \tag{12}$$

has $K = 2^{n_{x_2}}$ solutions, some of them will be real and some of them complex. In fact, we can see that the reduced form state space representation describes a quadratic system of $(n_y - n_{x_2})$ equations for $\mathbf{x}_{2t+1}$ and n_{me} linear equations in $\mathbf{v}_{t+1}$.

Finding the solutions of this mapping entails filtering $\mathbf{z}_{t+1}$ out. When $n_{z_2} = 1$ or 2, it is very easy to find all roots. With higher dimensional problems, other methods can be used. As an example, we can use the the polynomial homotopy continuation (PHC) method (and its Matlab interface[1]) by Verschelde (1999). More details on the methodology can be found in Morgan (1987) and Judd (1998). The Matlab interface is documented in Guan and Verschelde (2008).

[1] See http://www.math.uic.edu/~jan/

With the solutions (12) in hand, it is possible to obtain the likelihood by simply running the following algorithm

$$
\begin{aligned}
p(\mathbf{y}^o_{t+1}|\underline{\mathbf{y}}^o_t) &= \sum_{j=1}^{m^*}\sum_{k=1}^{K_{jt+1}}\sum_{i=1}^{m^*}\sum_{h=1}^{K_{it}} p(\mathbf{z}^{(jk)}_{t+1}|\mathbf{z}^{(ih)}_t) \times \left\| \frac{\partial \mathbf{z}^{(jk)}_{t+1}}{\partial \mathbf{y}'_{t+1}} \right\| \times p_{ij} \times p(\mathbf{z}_t = \mathbf{z}^{(ih)}_t, s_t = i|\underline{\mathbf{y}}^o_t) \quad (13) \\
p(\mathbf{z}_{t+1} &= \mathbf{z}^{(jk)}_t, s_{t+1} = j|\underline{\mathbf{y}}^o_{t+1}) = \\
&= \frac{\sum_{i=1}^{m^*}\sum_{h=1}^{K_{it}} p(\mathbf{z}^{(jk)}_{t+1}|\mathbf{z}^{(ih)}_t) \times \left\| \frac{\partial \mathbf{z}^{(jk)}_{t+1}}{\partial \mathbf{y}'_{t+1}} \right\| \times p_{ij} \times p(\mathbf{z}_t = \mathbf{z}^{(ih)}_t, s_t = i|\underline{\mathbf{y}}^o_t)}{p(\mathbf{y}^o_{t+1}|\underline{\mathbf{y}}^o_t)} \quad (14) \\
\underline{\mathbf{y}}^o_t &= \{\mathbf{y}^o_\tau, \tau = 1, 2, ..., t\}, m^* = 2^{n_s}
\end{aligned}
$$

where K_{jt+1} is the number of real roots at time $t + 1$ conditional on $s_{t+1} = j$ and K_{it} is the number of real roots at time t conditional on $s_t = i$.

The recursion is initialised as follows

$$
\begin{aligned}
p(\mathbf{z}_1 &= \mathbf{z}^{(ih)}_1, s_1 = i|\mathbf{y}^o_1) = p(s_1 = i|\mathbf{y}^o_1) \times p(\mathbf{z}_1 = \mathbf{z}^{(ih)}_1|s_1 = i, \mathbf{y}^o_1) \\
p(s_1 &= i|\mathbf{y}^o_1) = \frac{p(\mathbf{y}^o_1|s_1 = i) \times p(s_1 = i)}{\sum_{i=1}^{m^*} p(\mathbf{y}^o_1|s_1 = i) \times p(s_1 = i)} \\
p(s_1 &= i) = \pi_i \text{ (ergodic probability)} \\
p(\mathbf{z}_1 &= \mathbf{z}^{(ih)}_1|s_1 = i, \mathbf{y}^o_1) = \frac{1}{K}
\end{aligned}
$$

Several remarks seem necessary. First of all, note that expression (13) is the t-th factor of the likelihood function.

Secondly, (14) describes a filtering recursion over a discrete set of realisations which allows us to marginalise to obtain marginal filtered probabilities of discrete states (s_{t+1}) and continuous latent variables ($\mathbf{z}_{t+1}$).

An interesting question is how we handle complex roots. If some roots at time $t + 1$ for some given discrete state j are complex, these will be assigned zero probability, since the support of the latent variables is the real line. If at time $t + 1$for some j all roots are complex, then the algorithm will anyway work and will assign probability zero to state j conditioned at time $t+1$. If at time $t+1$ all roots are complex for all j, then the likelihood of that observation is set to zero.

From the computational point of view, first of all note that the Jacobians $\left\| \frac{\partial \mathbf{z}^{(jk)}_{t+1}}{\partial \mathbf{y}'_{t+1}} \right\|$ are readily computable analytically. In addition, it is important to bear in mind that the only computationally expensive part of the algorithm is to obtain the roots $\mathbf{z}_{jt+1}$. Note that the problem can be fully parallelised, since at each $t+1$ and j the roots computations are functions only of $\mathbf{y}_{t+1}^{o}$ and of the coefficients of the reduced form measurement equations.

There is no conceptual problem in extending the algorithm to contexts in which the DSGE model is solved to third order. We need only to keep complying with the requirement that $n_{z_2} + n_{me} = n_y$ and that there are no unobservable predetermined state variables.

5 An application to a simple model

In order to highlight the marginal contribution of heteroskedasticity, we rely on a standard model in the spirit of Woodford (2003). The central feature is the assumption of nominal rigidities and the presence of discrete shifts in conditional volatility of the shocks. The model is kept deliberately simple in order to have a very low number of continuous unobserved non-predetermined state variables. A larger (and admittedly more satisfactory) model is described in Amisano and Tristani (2009b), where a quadratic model with heteroskedastic shocks is used in a context where the vector of observable variables includes term structure data.

5.1 Households

We assume that each household i provides $N(i)$ hours of differentiated labor services to firms in exchange for a labour income $w_t(i)\,N_t(i)$. Each household owns an equal share of all firms j and receives profits $\int_0^1 \Pi_t(j)\mathrm{d}j$.

As in Erceg, Henderson and Levin (2000), an employment agency combines households' labor hours in the same proportions as firms would choose. The agency's demand for each household's labour is therefore equal to the sum of firms' demands. The labor index L_t has the Dixit-Stiglitz form $L_t = \left[\int_0^1 N_t(i)^{\frac{\theta_{w,t}-1}{\theta_{w,t}}}\mathrm{d}i\right]^{\frac{\theta_{w,t}}{\theta_{w,t}-1}}$ where $\theta_{w,t} > 1$ is subject to exogenous shocks. At time t, the employment minimizes the cost of producing a given amount of the aggregate labor index, taking each household's wage rate $w_t(i)$ as given and then sells units of the labor index to the production sector at the aggregate wage index $w_t = \left[\int_0^1 w(i)^{1-\theta_{w,t}}\mathrm{d}i\right]^{\frac{1}{1-\theta_{w,t}}}$. The employment agency's demand for the labor hours of household i is given by

$$N_t(i) = L_t\left(\frac{w_t(i)}{w_t}\right)^{-\theta_{w,t}} \tag{15}$$

Each household i maximizes its intertemporal utility with respect to consumption, the wage rate and holdings of contingent claims, subject to its labor demand function (15) and the budget constraint

$$P_t C_t(i) + \mathrm{E}_t Q_{t,t+1} W_{t+1}(i) \leq W_t(i) + w_t(i) N_t(i) + \int_0^1 \Xi_t(j)\, \mathrm{d}j \qquad (16)$$

where C_t is a consumption index satisfying

$$C_t = \left(\int_0^1 C_t(z)^{\frac{\theta-1}{\theta}}\, \mathrm{d}z \right)^{\frac{\theta}{\theta-1}} \qquad (17)$$

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

$$P_t = \left(\int_0^1 p(z)^{1-\theta}\, \mathrm{d}z \right)^{\frac{1}{1-\theta}}. \qquad (18)$$

W_t denotes the beginning-of-period value of a complete portfolio of state contingent assets, $Q_{t,t+1}$ is their price, $w_t(i)$ is the nominal wage rate and $\Xi_t(j)$ are the profits received from investment in firm j.

Equation (16) states that each household can only consume or hold assets for amounts that must be less than or equal to its salary, the profits received from holding equity in all the existing firms and the revenues from holding a portfolio of state-contingent assets.

Households maximise the discounted sum of the period utility

$$u_t = \left[\left(C_t(i) - hC_{t-1}(i) \right) \left(\overline{N} - N_t^{\phi}(i) \right) \right]^{1-\gamma} \qquad (19)$$

subject to the budget contraint (16)

$$P_t C_t(i) + \mathrm{E}_t Q_{t,t+1} W_{t+1}(i) \leq W_t(i) + w_t(i) N_t(i) + \int_0^1 \Xi_t(j)\, \mathrm{d}j$$

and

$$N_t(i) = L_t \left(\frac{w_t(i)}{w_t} \right)^{-\theta_{w,t}}$$

where the choice variables are $w_t(i)$ and $C_t(i)$.

The first order conditions for this problem can be written as

$$\widetilde{w}_t = \phi \mu_w N_t^{\phi-1} \frac{\left(C_t - hC_{t-1} \right)^{1-\gamma} \left(\overline{N} - N_t^{\phi} \right)^{-\gamma}}{\widetilde{\Lambda}_t} \qquad (20)$$

$$Q_{t,t+1} = \beta \frac{\widetilde{\Lambda}_{t+1}}{\widetilde{\Lambda}_t} \frac{1}{\pi_{t+1}} \qquad (21)$$

$$\widetilde{\Lambda}_t = (C_t - hC_{t-1})^{-\gamma} \left(\overline{N} - N_t^{\phi}\right)^{1-\gamma} - \beta h \mathrm{E}_t (C_{t+1} - hC_t)^{-\gamma} \left(\overline{N} - N_{t+1}^{\phi}\right)^{1-\gamma} \tag{22}$$

where $\widetilde{w}_t \equiv w_t/P_t$ and $\mu_w \equiv \theta_w/(\theta_w - 1)$.

The gross interest rate, I_t, equals the conditional expectation of the stochastic discount factor, i.e.

$$I_t^{-1} = \mathrm{E}_t Q_{t,t+1} \tag{23}$$

Note that we will focus on a symmetric equilibrium in which nominal wage rates are all allowed to change optimally at each point in time, so that individual nominal wages will equal the average w_t.

5.2 Firms

We assume a continuum of monopolistically competitive firms (indexed on the unit interval by j), each of which produces a differentiated good. Demand arises from households' consumption and from government purchases G_t, which is an aggregate of differentiated goods of the same form as households' consumption. It follows that total demand for the output of firm i takes the form $Y_t(j) = \left(\frac{P_t(i)}{P_t}\right)^{-\theta} Y_t$. Y_t is an index of aggregate demand which satisfies $Y_t = C_t + G_t$.

Firms have the production function

$$Y_t(j) = A_t L_t^{\alpha}(j)$$

where L_t is the labour index L_t defined above.

Once aggregate demand is realised, the firm demands the labour necessary to satisfy it

$$L_t(j) = \left(\frac{Y_t(j)}{A_t}\right)^{\frac{1}{\alpha}}$$

The total nominal cost function for the firm will therefore be given by

$$TC_t(j) = w_t \left(\frac{Y_t(j)}{A_t}\right)^{\frac{1}{\alpha}}$$

where w_t is the wage index defined above. As a result, real marginal costs will be

$$mc_t(j) = \frac{1}{\alpha} \frac{w_t}{P_t} \frac{1}{A_t} \left(\frac{Y_t(j)}{A_t}\right)^{\frac{1-\alpha}{\alpha}}$$

where nominal costs are deflated using the aggregate price level (not the individual firm's price).

As in Rotemberg (1982), we assume the firms face quadratic costs in adjusting their prices. This assumption is also adopted, for example, by Schmitt-Grohé and Uribe (2004) and Ireland (1997). It is well-known to yield first-order inflation dynamics equivalent to those arising from the assumption of Calvo pricing.[2] From our viewpoint, it has the advantage of greater computational simplicity, as it allows us to avoid having to include an additional state variable in the model, i.e. the cross-sectional dispersion of price across firms.

The specific assumption we adopt is that firm j faces a quadratic cost when changing its prices in period t, compared to period $t-1$. Consistently with what is typically done in the Calvo literature, we modify the original Rotemberg (1982) formulation to allow for indexation of prices in part to lagged inflation, in part to the inflation objective

$$\frac{\zeta}{2}\left(\frac{P_t^j}{P_{t-1}^j} - (\Pi^*)^{1-\iota}\,\Pi_{t-1}^{\iota}\right)^2 Y_t$$

So, firms maximise their real profits

$$\max_{P_t^j} \mathrm{E}_t \sum_{s=t}^{\infty} Q_{t,s}\left[\frac{P_s^j Y_s^j\left(P_s^j\right)}{P_s} - \frac{TC_s\left(Y_s^j\left(P_s^j\right)\right)}{P_s} - \frac{\zeta}{2}\left(\frac{P_s^j}{P_{s-1}^j} - (\Pi^*)^{1-\iota}\,\Pi_{s-1}^{\iota}\right)^2 Y_s\right]$$

subject to

$$Y_t(j) = \left(\frac{P_t(j)}{P_t}\right)^{-\theta} Y_t$$

and to

$$Y_t(j) = A_t L_t^{\alpha}(j)$$

Focusing on a symmetric equilibrium in which all firms adjust their price at the same time, the first order condition for price setting can be written as

$$(\theta-1)\,Y_t + \zeta\left(\Pi_t - (\Pi^*)^{1-\iota}\,\Pi_{t-1}^{\iota}\right) Y_t \Pi_t = \frac{\theta}{\alpha}\widetilde{w}_t\left(\frac{Y_t}{A_t}\right)^{\frac{1}{\alpha}} + \mathrm{E}_t Q_{t,t+1}\zeta\left(\Pi_{t+1} - (\Pi^*)^{1-\iota}\,\Pi_t^{\iota}\right) Y_{t+1}\Pi_{t+1}$$

5.3 Monetary policy

We close the model with the simple Taylor-type policy rule

$$I_t = \left(\frac{\Pi^*}{\beta}\right)^{1-\rho_I}\left(\frac{\Pi_t}{\Pi^*}\right)^{\psi_\Pi}\left(\frac{Y_t}{Y}\right)^{\psi_Y} I_{t-1}^{\rho_I} \tag{24}$$

where Y_t is aggregate output and Π^* is a constant inflation target.

[2] The two pricing models, however, have in general different welfare implications – see Lombardo and Vestin (2008).

5.4 Market clearing

Market clearing in the goods market requires

$$Y_t = C_t + G_t$$

In the labour market, labour demand will have to equal labour supply. In addition, the total demand for hours worked in the economy must equal the sum of the hours worked by all individuals. Taking into account that at any point in time the nominal wage rate is identical across all labor markets because all wages are allowed to change optimally, individual wages will equal the average w_t. As a result, all households will chose to supply the same amount of labour and labour market equilibrium will require that

$$L_t = \left(\frac{Y_t}{A_t}\right)^{\frac{1}{\alpha}}$$

5.5 Exogenous shocks

In macroeconomic applications, exogenous shocks are almost always assumed to be (log-)normal, partly because models are typically log-linearised and researchers are mainly interested in characterising conditional means. However, Hamilton (2008) argues that a correct modelling of conditional variances is always necessary, for example because inference on conditional means can be inappropriately influenced by outliers and high-variance episodes. The need for an appropriate treatment of heteroskedasticity becomes even more compelling when models are solved nonlinearly, because conditional variances have a direct impact on conditional means.

In this paper, we assume that variances are subject to stochastic regime switches fall shocks other than the inflation target. More specifically, we assume a stochastic trend in techology growth

$$\begin{aligned} A_{t+1} &= A_t \Xi_{t+1} \\ \Xi_{t+1} &= \Xi^{1-\rho_\xi} \Xi_t^{\rho_\xi} e^{\varepsilon_{t+1}^{\xi}}, \qquad \varepsilon_{t+1}^{\xi} \tilde{} N\left(0, \sigma_{\xi, s_{\xi,t}}\right) \end{aligned}$$

where Ξ is the long run growth rate. We specify the exogenous government spending process in deviation from the stochastic trend, so that

$$\frac{G_t}{A_t} = \left(\frac{gY}{A}\right)^{1-\rho_g} \left(\frac{G_{t-1}}{A_{t-1}}\right)^{\rho_g} e^{\varepsilon_t^g} \qquad \varepsilon_{t+1}^{G} \tilde{} N\left(0, \sigma_{G, s_{G,t}}\right)$$

where the long run level g is specified in percent of output, so that $g \equiv G/Y$.

Both technology and government spending shocks have regime-switching variances, namely

$$\begin{aligned} \sigma_{\xi,s_{\xi,t}} &= \sigma_{\xi,L} s_{\xi,t} + \sigma_{\xi,H} (1 - s_{\xi,t}) \\ {}_{G,s_{G,t}} &= \sigma_{G,L} s_{G,t} + \sigma_{G,H} (1 - s_{G,t}) \end{aligned}$$

and the variables $s_{\xi,t}$, $s_{G,t}$ can assume the discrete values 0 and 1. For each variable $s_{j,t}$ ($j = \xi, G$), the probabilities of remaining in state 0 and 1 are constant and equal to $p_{j,0}$ and $p_{j,1}$, respectively.

We assume regime switches in these particular variances for the following reasons. The literature on the "Great moderation" (see e.g. McDonnell and Perez-Quiros, 2000) has emphasised the reduction in the volatility of real aggregate variables starting in the second half of the 1980s. We conjecture that this phenomenon could be captured by a reduction in the volatility of technology shocks in our structural setting. The literature has also often found a relationship between regimes and the business cycle. In our model, this relationship could be accounted for by regime switches of the volatility of demand (government spending) shocks.

5.6 Data and prior distributions

We estimate the model on quarterly US data over the sample period from 1966Q1 to 2009Q1. Our estimation sample starts in 1966, because this is often argued to be the date after which a Taylor rule provides a reasonable characterisation of Federal Reserve policy.[3]

The data included in the information set are real per-capita consumption growth, the consumption deflator and the 3-month nominal interest rate. Measurement error characterises the nominal interest rate.

Data described in figure 1. Note, that beside the differencing for log consumption, the data are not subjected to any transformation, such as de-trending.

[3] According to Fuhrer (1996), "since 1966, understanding the behaviour of the short rate has been equivalent to understanding the behaviour of the Fed, which has since that time essentially set the federal Funds rate at a target level, in response to movements in inflation and real activity". Goodfriend (1991) argues that even under the period of official reserves targeting, the Federal Reserve had in mind an implicit target for the Funds rate.

As argued by Giannoni and Woodford (2005), assuming habit persistence for the whole level of private expenditure is a reasonable assumption, given that models with capital typically need adjustment costs that imply inertia in the rate of investment spending. We therefore include in the information set total real personal consumption per-capita and the consumption deflator (from the FRED database of the St. Louis Fed). In addition, we use the 3-month nominal interest rate (from the Federal Reserve Board).

Prior distributions for our model are presented in Table 1.

Concerning regime switching processes, we assume beta priors for transition probabilities. The distributions imply that persistences in each state are symmetric and have high means. In the prior, we assume that the standard deviations of the structural shocks are identical in the various states.

The priors for the standard deviation and persistence of shocks, as well as for the long run growth rate of technology and for the long run inflation target, are centred on values which allow us to roughly match unconditional data moments in the first 10 years of the sample, given the other parameter.

For the policy rule, we use relatively loose priors centred around the classic Taylor (1993) parameters for the responses to inflation and output, but we also allow for a substantial interest rate smoothing. Finally, for the other parameters we use priors broadly in line with other macro studies.

Note that we impose priors which are completely symmetric across states. Therefore we sample from the posterior and ex post impose the constraint that state 1 for each of the discrete unobservable state variables is that with the lowest variance. This is the way in which we deal with the so-called label switching problem of Markov switching models. For a discussion of this problem and its bearing on posterior simulation, see Geweke and Amisano (2009).

5.7 Results

Bayesian estimation is performed by using a single block random walk Metropolis-Hastings using a multivariate Gaussian distribution as candidate density. We initialise the algorithm by finding the mode of the log posterior via simulated annealing and computing the covariance matrix of the candidate distribution by using a numerical Hessian of the posterior distribution at its mode.

We run several chains and we report results here based on 110,000 draws. The results from the posterior distribution are summarised in Table 2.

A few notable features are apparent from Table 2.

Looking at the transition probabilities of posterior distribution, we note that their marginal distributions are centered on mean values which are quite different from their prior counterpart (with the exception of the probability of staying in the regime of high variability for the G process).The state of low volatility seem to be more persistent than that of high volatility for the G process, whereas for the ξ process this conclusion is reversed.

Note also that the posterior distributions of the state specific standard errors are quite polarised: for instance, the standard deviation of G in the high volatility state is 3 times as high as the low volatility standard deviation, whereas this ratio is over 2 for the productivity process.

Looking at the policy rule parameters, we see that the interest smoothing parameter has a posterior mean which is higher than its prior mean (.90 versus .70), while the inflation coefficient posterior mean is lower than its prior mean.

It is interesting to consider one step ahead forecasting errors as a way to gauge model fit. These are illustrated in Figure 3. Looking at one step ahead forecasting errors, the fit seems not to be disappointing. One step ahed forecasts track actual variables quite well, with the exception of inflation at the beginning of the sample. At this stage there is no way to ascertain whether this feature is a consequence of the stripped down structural specification of the model, or a consequence of a bad initialisation of the filtering procedure. Further investigation is required on this front.

Figure 3 displays the filtered values for the continuous latent variables. Looking at the filtered values for the measurement error on interest rate, we see that it shows abnormal values in the early 90s, possibly in relationship with the so called "monetarist experiment". In this regard, see Amisano and Tristani (2009b).

Turning to the analysis of the discrete states indexing volatility regimes (Figure 4), we notice some interesting features.

First of all, given the posterior distribution of transition probabilities, we expect G to spend more time in low volatility state and ξ in high volatility state. This is clearly confirmed by looking at filtered and smoothed probabilities. G is considered to be in a volatile state only in the early 1970s and then from 2001 onwards.

Overall, it seems that the interpretation for the variances of the productivity growth and government spending processes is reversed, compared to our conjecture. Productivity growth appears to be highly cyclical: its variance increases during recessions. This is also true for the current recession. The variance of government spending shocks, on the contrary, displays a more persistent pattern. The variance is low from the second half of the seventies until the end of the nineties, while it was high in the early seventies and again at the beginning of the new millennium.

6 Conclusions

This paper shows that the second-order approximate solution of DSGE models with Markov switching states is characterised by coefficients on the linear and quadratic terms in the state vector of the decision rules that are independent of the volatility of the exogenous shocks. Up to second order, only the constant term of the decision rules is affected by the introduction of regime switching. We devise a procedure to compute the likelihood in situations in which the number of latent continuous non predetermined shocks match the number of observable variables and there are no unobservable predetermined variables. In such an environment we can compute the likelihood exactly via recursive methods withour resorting to approximations or to simulation filtering techniques. We believe that this is an interesting case worth analysing per se, and in order to provide benchmark when developing and refining simulation filtering algoritms.

References

[1] Amisano, G. and O. Tristani (2009a), "Euro area inflation persistence in an estimated nonlinear DSGE model," Journal of Economic Dynamics and Control, forthcoming.

[2] Amisano, G. and O. Tristani (2009b), "A DSGE model of the term structure with regime shifts", mimeo, European Central Bank.

[3] Andolfatto, D. and P. Gomme (2004), "Monetary policy regimes and beliefs", International Economic Review 44, 1-30.

[4] Aruoba, B., J. Fernandez-Villaverde and J.F. Rubio-Ramirez (2006), "Comparing Solution Methods for Dynamic Equilibrium Economies", Journal of Economic Dynamics and Control 30, 2477-508.

[5] Benigno, P. and M. Woodford (2006), "Linear-Quadratic Approximation of Optimal Policy Problems," mimeo.

[6] Blanchard, O. and C. Kahn (1980), “The solution of Linear Difference Models under Rational Expectations,” Econometrica 48, 1305-1311.

[7] Cecchetti, S.G., P. Lam and N.C.Mark (1993), “The equity premium and the risk-free rate,” Journal of Monetary Economics 31, 21-45.

[8] Cogley, T. and T.J. Sargent (2005), “Drifts and volatilities: monetary policies and outcomes in the post WWII U.S.,” Review of Economic Dynamics 8, 262-302.

[9] Coleman, W.J. (1991), “Equilibrium in a production economy with an income tax,” Econometrica 59, 1091-104.

[10] Davig, T. and E.M. Leeper (2006), “Generalising the Taylor principle,” mimeo.

[11] Davig, T., E.M. Leeper and H. Chung (2004), “Monetary and fiscal policy switching,” NBER WP No. 10362.

[12] J. Fernandez-Villaverde and J.F. Rubio-Ramirez (2006), “Estimating macroeconomic models: a likelihood approach”, mimeo.

[13] Geweke, J. and G. Amisano (2009): Hierarchical Markov mixture models and their applications in finance, forthcoming on Journal of Applied Econometrics.

[14] Giannoni, Marc P. and Michael Woodford (2005), "Optimal Inflation Targeting Rules," in: B.S. Bernanke and M. Woodford, eds., Inflation Targeting, University of Chicago Press.

[15] Gomme, P. and P. Klein (2006), “Second-order approximation of dynamic models without the use of tensors,” University of Western Ontario, *mimeo*, November.

[16] Guan, Yun and Jan Verschelde (2008), "PHClab: A MATLAB/Octave interface to PHCpack," in Michael E. Stillman, Nobuki Takayama, and Jan Verschelde, eds., IMA Volume 148: Software for Algebraic Geometry, pp. 15-32, Springer-Verlag.

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

[18] Judd, K. (1998), Numerical Methods in Economics, Cambridge (MA): MIT Press.

[19] 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.

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

[21] Lombardo, G. and A. Sutherland (2007), “Computing Second-Order-Accurate Solutions for Rational Expectation Models using Linear Solution Methods”, Journal of Economic Dynamics and Control 31, 515-30.

[22] Magnus, I. and H. Neudecker (1999), Matrix differential calculus with applications in statistics and econometrics, New York: Wiley.

[23] McConnell, M. and G. Perez-Quiros (2000),"Output Fluctuations in the United States: What Has Changed since the Early 1980s?" American Economic Review 90, 1464-76.

[24] A. Morgan (1987), Solving polynomial systems using continuation for engineering and scientific problems. Englewood Cliffs, N.J.: Prentice-Hall.

[25] Primiceri, G. (2005),"Time Varying Structural Vector Autoregressions and Monetary Policy," Review of Economic Studies 72, 821–852.

[26] 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.

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

[28] Sims, C.A. and T. Zha (2006), "Were there regime switches in U.S. monetary policy?," American Economic Review 96, 54-81.

[29] Sims, C.A. (2001), "Solving Linear Rational Expectations Models", Journal of Computational Economics 20,1-20.

[30] Söderlind, P. (1999), "Solution and Estimation of RE Macromodels with Optimal Policy," European Economic Review 43, 813-823.

[31] Verschelde, J. (1999): Algorithm 795: PHCpack: A general-purpose solver for polynomial systems by homotopy continuation. ACM Trans. Math. Softw., 25(2):251{276, 1999. Software available at http://www.math.uic.edu/~jan.)

Table 1: Prior specification

parameter name	**parameter role**	**prior type**	**mean**	**sd**	**lower q**	**upper q**
$p_{G,11}$	prob of G staying in state 1	Beta	0.9000	0.0655	0.7397	0.9870
$p_{G,22}$	prob of G staying in state 2	Beta	0.9000	0.0655	0.7397	0.9870
$p_{\xi,11}$	prob of XI staying in state 1	Beta	0.9000	0.0655	0.7397	0.9870
$p_{\xi,22}$	prob of XI staying in state 2	Beta	0.9000	0.0655	0.7397	0.9870
$\sigma_{G,1}$	Standard G shock state 1	Inv. Gamma	0.0081	0.0029	0.0045	0.0154
$\sigma_{G,2}$	Standard G shock state 2	Inv. Gamma	0.0081	0.0029	0.0045	0.0154
$\sigma_{\xi,1}$	Standard deviation of XI state 1	Inv. Gamma	0.0022	0.0007	0.0014	0.0038
$\sigma_{\xi,2}$	Standard deviation of XI state2	Inv. Gamma	0.0022	0.0007	0.0014	0.0038
ρ_G	Persistence of G shock	Beta	0.9000	0.0299	0.8344	0.9505
ξ	expected rate of growth of tech	Shifted Gamma	1.0040	0.0028	1.0005	1.0111
ρ_ξ	persistence of tech growth	Beta	0.2000	0.1206	0.0281	0.4825
$\overline{\pi}$	Long-run target	Shifted Gamma	1.0070	0.0026	1.0028	1.0131
ψ_π	Taylor rule inflation parameter	Gamma	0.5000	0.3536	0.0606	1.3929
ψ_y	Taylor rule output parameter	Gamma	0.0308	0.0218	0.0037	0.0857
ρ_I	Interest rate smoothing coefficient	Gaussian	0.7000	0.3000	0.1120	1.2880
ι	Inflation indexation	Beta	0.6000	0.2000	0.1941	0.9324
ϕ	elasticity of labour	Gaamma	2.0000	1.4142	0.2422	5.5716
γ	Relative risk aversion	shifted Gamma	3.0000	1.4142	1.2422	6.5716
σ	inverse of elasticity intertemp subst	shifted Gamma	2.0000	1.0000	1.0253	4.6889
ζ	Price adjustment cost	Gaussian	17.0000	2.0000	13.0800	20.9200
h	Habit parameter	Beta	0.7059	0.1074	0.4762	0.8898
θ	Elasticity of substitution across goods	shfted Gamma	8.0000	2.6458	3.8144	14.0595
β	Discount factor	Beta	0.9940	0.0042	0.9833	0.9993
$\sigma_{me,I}$	measurement error short rate	Inverted Gamma	0.0001	0.0001	0.0001	0.0003

Legend: "sd" denotes the standard deviation; "low q" and "up q" denote the 5th and 95th percentiles of the distribution. Note that shifted gamma distribution for z means that z-1 has Gamma distribution.

Table 2: Results from posterior simulation of model

	post mean	post sd	post low q	post up q	prior mean	prior sd	prior low q	prior up q
$p_{G,11}$	0.9693	0.0242	0.8937	0.9928	0.8997	0.0654	0.7391	0.9866
$p_{G,22}$	0.9145	0.0423	0.8228	0.9679	0.8994	0.0651	0.7385	0.9872
$p_{\xi,11}$	0.8130	0.0748	0.6603	0.9133	0.9004	0.0657	0.7397	0.9875
$p_{\xi,22}$	0.8829	0.0429	0.7681	0.9357	0.8989	0.0660	0.7396	0.9867
$\sigma_{G,1}$	0.0344	0.0024	0.0306	0.0394	0.0066	0.0016	0.0042	0.0104
$\sigma_{G,2}$	0.1046	0.0036	0.0975	0.1103	0.0095	0.0032	0.0056	0.0177
$\sigma_{\xi,1}$	0.0025	0.0004	0.0021	0.0034	0.0019	0.0004	0.0013	0.0028
$\sigma_{\xi,2}$	0.0057	0.0002	0.0053	0.0061	0.0025	0.0007	0.0017	0.0042
ρ_G	0.9942	0.0012	0.9920	0.9962	0.9005	0.0296	0.8356	0.9504
ξ	1.0042	0.0000	1.0042	1.0043	1.0040	0.0028	1.0005	1.0108
ρ_ξ	0.8957	0.0115	0.8739	0.9204	0.1995	0.1199	0.0281	0.4873
$\overline{\pi}$	1.0023	0.0003	1.0017	1.0031	1.0070	0.0027	1.0028	1.0131
ψ_π	0.2111	0.0084	0.1938	0.2235	0.5015	0.3522	0.0607	1.3882
ψ_y	0.0015	0.0011	0.0003	0.0039	0.0308	0.0220	0.0036	0.0857
ρ_I	0.9114	0.0085	0.8955	0.9280	0.7003	0.2985	0.1164	1.2999
ι	0.0706	0.0371	0.0232	0.1650	0.5996	0.1999	0.1916	0.9330
ϕ	3.5588	0.0583	3.4480	3.6551	2.0124	1.4194	0.2475	5.5981
γ	3.2432	1.2064	1.4828	5.8154	2.9600	1.4097	1.2412	6.6680
ζ	16.4775	1.3925	14.6676	19.7259	16.9787	1.9855	13.0238	20.8592
h	0.6963	0.0121	0.6747	0.7287	0.7057	0.1076	0.4764	0.8886
θ	3.4701	0.0933	3.2743	3.5937	8.0218	2.6644	3.8130	14.2124
β	0.9984	0.0008	0.9966	0.9995	0.9940	0.0042	0.9834	0.9993
$\sigma_{me,I}$	0.0024	0.0001	0.0022	0.0026	0.0001	0.0001	0.0001	0.0003

Estimation was carried out by MCMC using a random walk MH algorithm initialised at posterior mode. We run 110,000 draws, used the first 10,000 for burn-in and retained every 10th draw of the subsequent 100,000. The acceptance rate of the resulting chain was .35.

Figure 1: Series used for model estimation

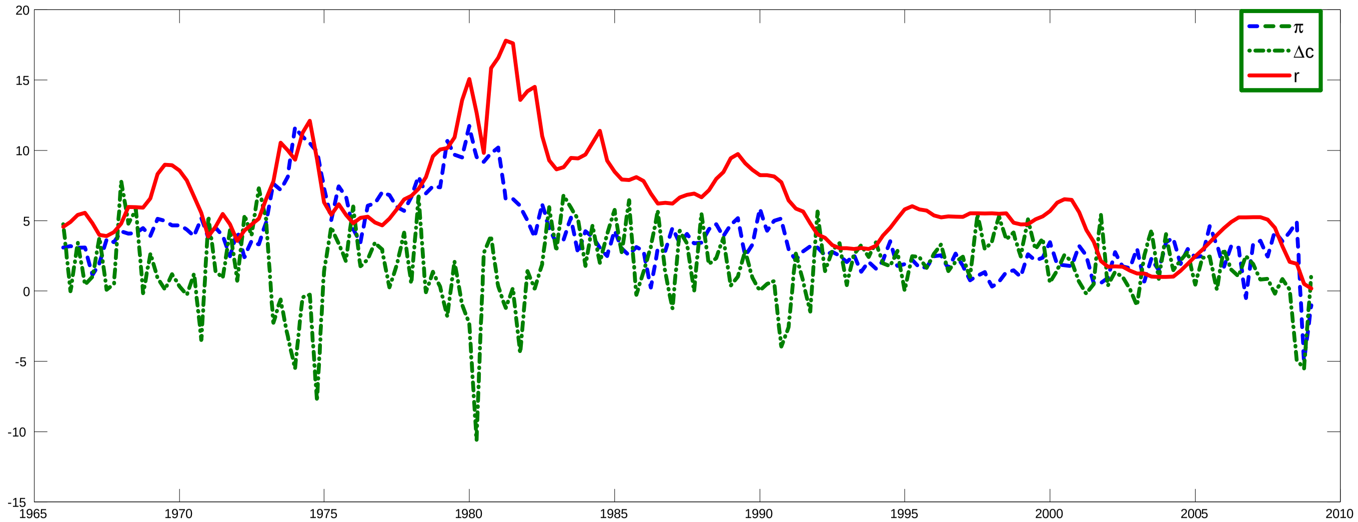
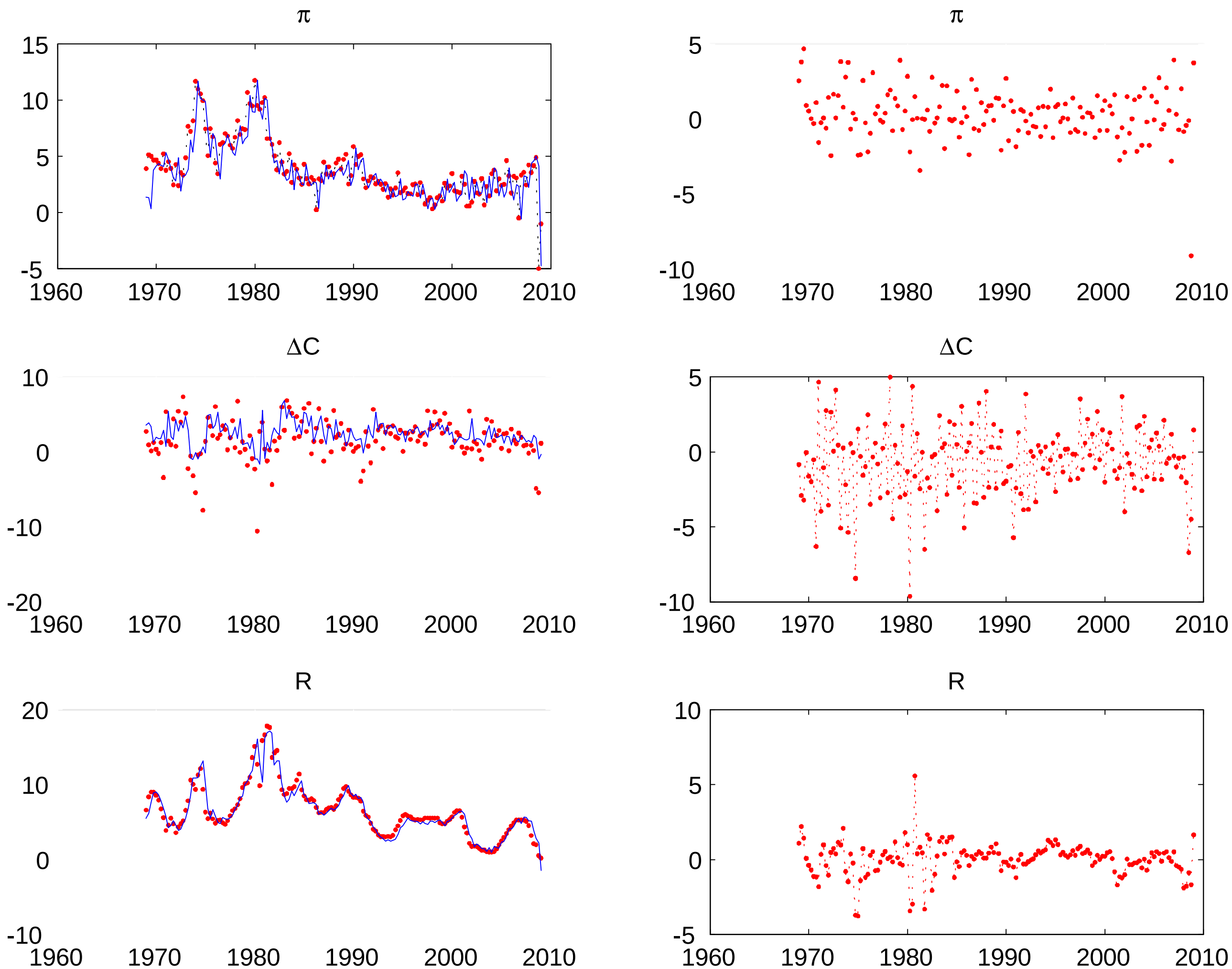
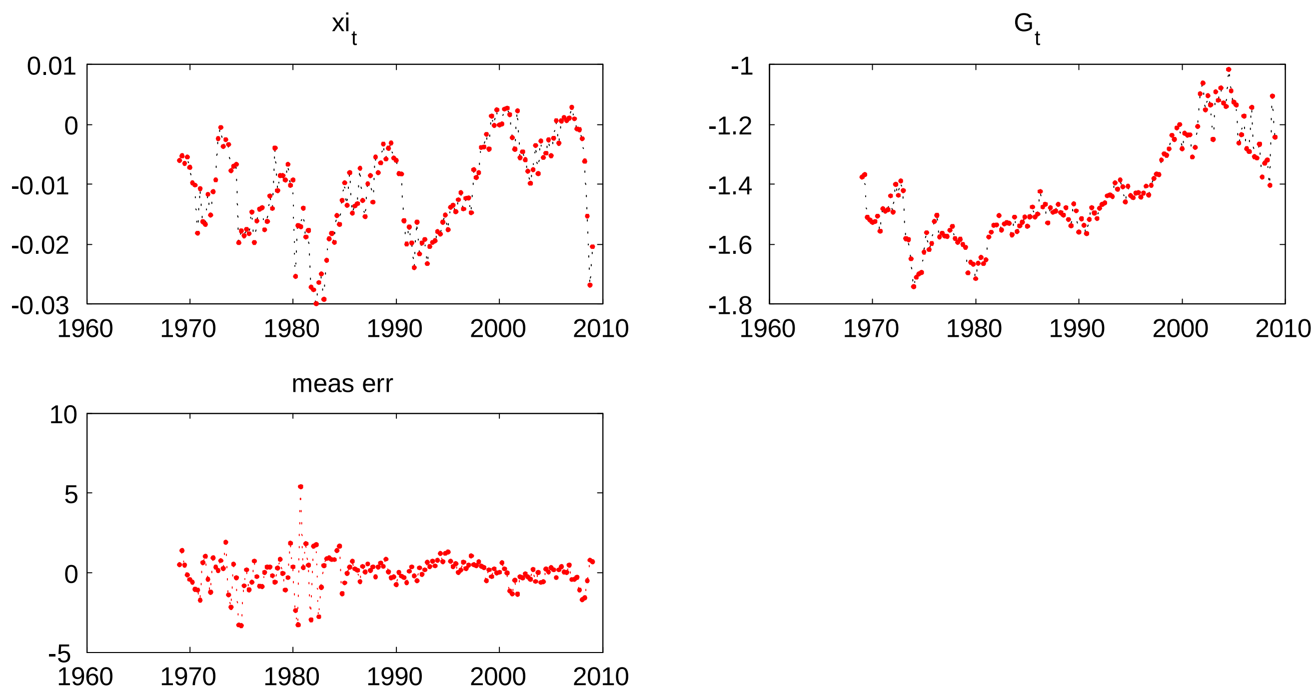


Figure 2: Model fit: actual series and one step ahead predictions



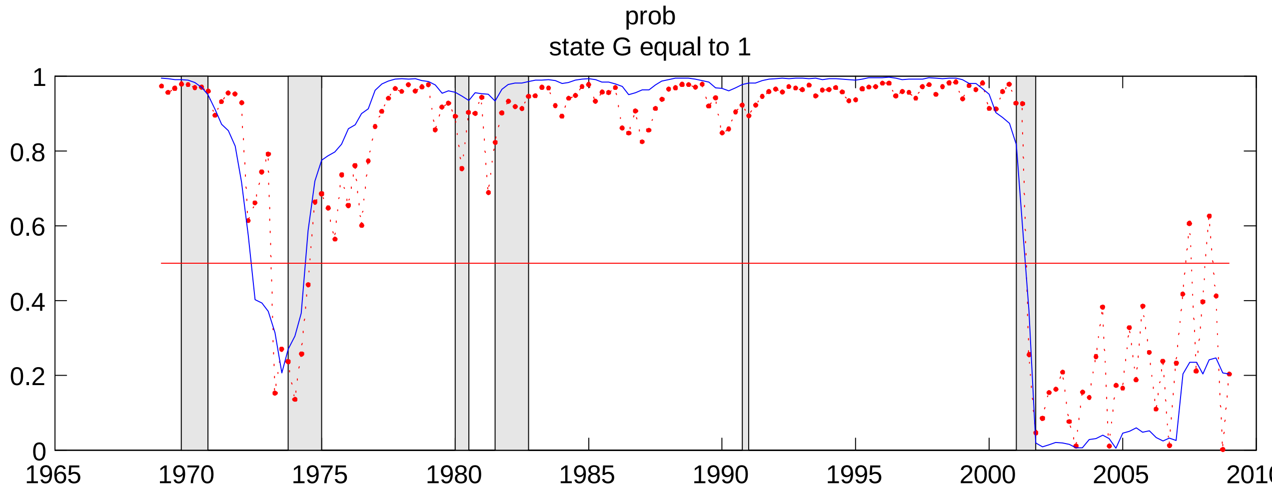
Legend: The left-hand side panels contain actual (red) and one step ahead forecasts of observed variables. The right-hand side panels contain one step ahead forecast errors. The reported values are posterior means of these quantities using the posterior distribution of estimates.

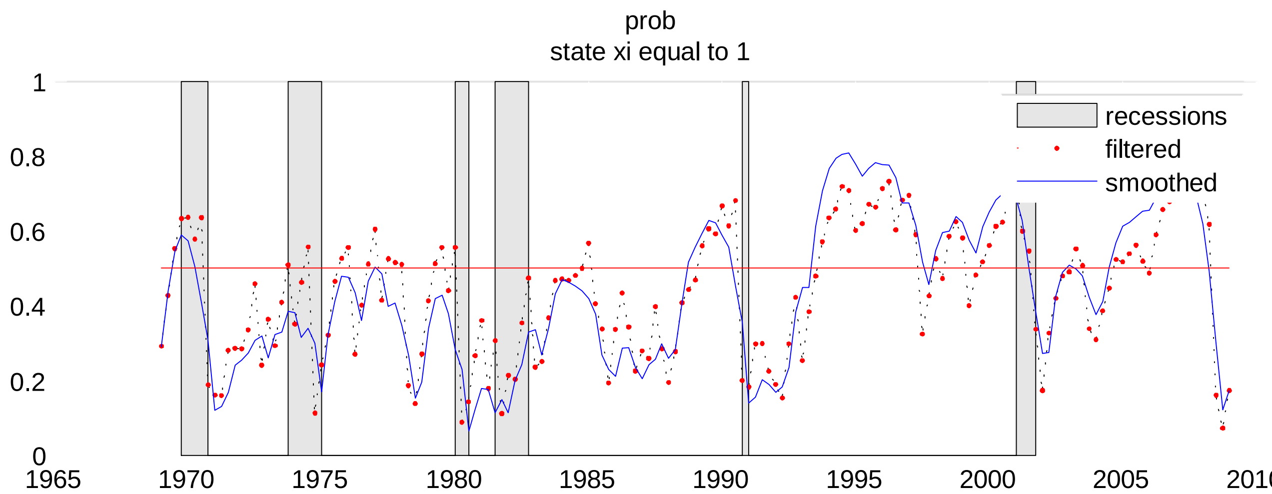
Figure 3: Continuous latent variables, filtered variables



Legend: Filtered values are computed by averaging over draws from posterior distribution of the parameters.

Figure 4: Filtered and smoothed values for discrete state variables





Legend: Filtered and smoothed probabilities are computed by averaging over draws from posterior distribution of the parameters..