A DSGE model of the term structure with regime shifts*

Gianni Amisano[†]
European Central Bank

Oreste Tristani[‡]
European Central Bank

13 June 2008

Abstract

We analyse the term structure implications of a small DSGE model with nominal rigidities in which the laws of motion of the structural shocks are subject to stochastic regime shifts. We first demonstrate that, to a second order approximation, switching regimes generate time-varying risk premia. We then estimate the model using sequential Monte Carlo methods and relying on information from both macroeconomic and term structure data. Our preliminary results, based on the linearised model, support the specification with regime switching. Shifts in the variance of technology shocks are clearly associated with the transition to the Great moderation; changes in the variance of policy shock identify the so-called monetarist experiment; switches in the variance of preference shocks have a cyclical nature.

JEL classification:

Keywords: DSGE models, term structure of interest rates, policy rules, particle filter, Bayesian estimation.

PRELIMINARY AND INCOMPLETE

*The opinions expressed are personal and should not be attributed to the European Central Bank. We wish to thank Michel Juillard for a precious suggestion on how to speed up the computation of the solution of the model.

[†]Email: gianni.amisano@ecb.int.

[‡]Email: oreste.tristani@ecb.int.

1 Introduction

The term structure of interest rates is a source of information for monetary policy. Many central banks analyse it to derive estimates of, *inter alia*, markets' expectations of future policy moves and perceptions of inflation expectations at future horizons. Since micro-founded general equilibrium models have traditionally had a hard time to match yield data, these estimates are often derived from finance-type models, where the relationship between interest rates, monetary policy and macroeconomic fundamentals is not explicitly accounted for. This strategy prevents a full undertanding of the determinants of risk premia and of their possible comovement with other economic variables. A fully structural explanation of the yield curve would be desirable.

At the same time, the yield curve plays implicitly a central role in macro (DSGE) models, because the expectations channel is a fundamental component of their monetary policy transmission mechanism. The central bank can often afford to react little, on impact, to deviations of inflation from its target value, because at the same time it promises – and private agents believe this promise – that it will keep reacting over a long time in the future. This type of monetary policy rule – often described as "inertial," or including a concern for "interest rate smoothing" – stabilises inflation because aggregate demand is affected by the whole expected future path of policy interest rates, not just the current rate. Given this central role of the yield curve in DSGE models, it would also be desirable to include bond prices in the information set of the econometrician when the models are taken to the data. Linearised DSGE models, however, appear to be inconsistent with yield data at a basic level. They imply that the unconditional slope of the term structure should be zero, contrary to the overwhelming evidence that the average term structure is positively sloped.

Finally, from a purely empirical viewpoint it is well-known that DSGE models are affected by partial and weak identification problems – see e.g. Canova and Sala (2006). These problems are particularly visible for some parameters of the monetary policy rule, which are often pinned down by the researcher's prior. Including information from the yield curve in the estimation process should help to mitigate these indentification problems. It should also help to filter more reliably certain unobservable variables, such as a time varying (perceived) inflation target.

In this paper, we explore the ability of a small microfounded model with nominal rigidities to match both macroeconomic and term structure data using a full-information estimation approach. However, we deviate from the DSGE literature in two respects.

First, we solve and estimate the second-order approximate solution of the model, rather than its log-linearised version. More specifically, we rely on perturbation methods to solve the model up to a second-order approximation. We then use a variant of the particle filter to estimate the nonlinear reduced form (see Amisano and Tristani, 2007a). The nonlinear solution has the advantage of being capable of generating non-negligible term-premia, which can explain the average positive slope of the yield curve. Linearised DSGE models, on the contrary, force the unconditional slope of the term structure to be zero, which is in blatant contrast with the available evidence.

The second deviation we take from the standard empirical DSGE literature is to allow for heteroskedasticity of macroeconomic shocks, due to the fact that selected parameters are assumed to be subject to regime switches. In terms of matching the dynamic features of the term structure, the assumption of heteroskedasiticy implies that the model is capable of generating time-variation in risk premia. We assume that heteroskedasiticy takes the specific form of regime switching, because this assumption has already been shown to help fit yields in the finance literature – see Hamilton (1988), Naik and Lee (1997), Ang and Bekaert (2002a,b), Bansal and Zhou (2002), Bansal, Tauchen and Zhou (2004), Ang, Bekaert and Wei (2008), Dai, Singleton and Yang (2008), Bikbov and Chernov (2007) – and is also increasingly used in macroeconomics following Sims and Zha (2007).

Our model is related to a growing literature exploring the term structure implications of new-Keynesian models. The closest papers to ours is Doh (2006), which also estimates a quadratic DSGE model of the term structure of interest rates with heteroskedastic shocks. However, Doh (2006) allows for additional non-structural parameters to model the unconditional slope of the yield curve, while our approach is fully theoretically consistent. Another difference between the two papers is that heteroskedasticity in Doh (2006) is modelled through ARCH shocks, while it is generated by regime switching in our case. Bekaert, Cho and Moreno (2006) and De Graeve, Emiris and Wouters (2007) estimate the loglinearised reduced form of DSGE models using both macroeconomic and term structure data. As in Doh (2006), these papers do not impose theoretical restrictions on the unconditional slope of the yield curve. In addition, they assume at the outset that risk-premia are constant. A different approach to generate time variation in risk premia, based on third order approximations, is pursued in Ravenna and Seppala (2007a, b), Rudebusch, Sack and Swanson (2007) and Rudebusch and Swanson (2007). However, these papers are purely theoretical: the estimation of DSGE models solved using third order approximations appears to be infeasible at this point in time.

Our preliminary results, based on the estimation of the first order approximation of the model, show considerable support for a specification with regime switches. The residuals of a model with gaussian shocks show clear signs of heteroskedasticity and serial correlation. Moreover, estimated regimes appear to bear an intuitively appealing structural interpretation: monetary policy shocks are normally in the low-variance regime, except for the so-called monetary experiment period at the beginning of the 1980s; technology shocks are in the high-variance regime before the 1980s, and switch persistently to the low-variance regime thereafter, consistently with the evidence of a Great moderation; finally, demand shocks have a cyclical connotation, with a variance which tends to be lower during expansions.

At the same time, linearised models – even when they include heteroskedastic shocks – display clear signs of mispecification when asked to match yields data. They can only do so at the cost of bending the parameter estimates towards regions that are not intuitively appealing from a macroeconomic viewpoint. For example, linearised models can explain yields only when policy interest rates become extremely persistent, to the extent that the real interest rate sensitivity of output must become negligible in order to avoid implausible repercussions on real variables. This implies an extremely high estimate of the long-term coefficient of relative risk aversion.

This problem should be mitigated when the model is estimated to a second order approximation, because risk premia could account for some of the yields dynamics which must otherwise be explained by expectations terms. We provide two pieces of evidence supporting this conjecture. First, we present estimates of the quadratic version of our model with homoskedastic shocks. While plagued by many of the problems relevant for the linearised, model, the quadratic version shows signs of improvements in its ability to explain yields dynamics wish a smaller degree of relative risk aversion. Its fitting errors are also smaller in absolute value. The second piece of evidence is based on the posterior mean of the paraters estimated with the linearised version of the heteroskedastic model. When we plug these parameter values in the quadratic version of the same model, we demonstrate that they are capable of generating non-negligible variability in yields premia.

The rest of the paper is organised as follows. Section 2 includes a brief description of the theoretical model, which is of the standard new-Keynesian type. This section also includes details on the solution method and on how a second order approximation of the model can generate time-variability in yields premia. The estimation methodology is then described in Section 3, which focuses on the problems introduces by non-normal shocks in a structural model. Section 4 presents our estimation results. For illustrative purposes, we estimate the model both with homoskedastic and heteroskedastic shocks. We draw some tentative conclusions in Section 5.

2 The model

The model we employ is in the spirit of Yun (1996) and Woodford (2003). The central feature is the assumption of nominal rigidities. Since the model features are quite standard, we only sketch its properties briefly.

Consumers maximise the discounted sum of the period utility

$$U\left(C_t, C_{t-1}, L_t\right) = \varepsilon_{Ct}^{\gamma} \frac{\left(C_t - hC_{t-1}\right)^{1-\gamma}}{1-\gamma} - \int_0^1 \chi \frac{L_{i,t}^{1+\nu}}{1+\nu} di \tag{1}$$

where C is a consumption index satisfying

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

workers provide L_i hours of labor to firm i and $\varepsilon_{Ct}^{\gamma}$ is a demand shock whose properties will be defined below. The presence of lagged consumption in utility captures households' internal habits.

The households' budget constraint is given by

$$P_t C_t + E_t\left(Q_{t,t+1} W_{t+1}\right) \leqslant \int_0^1 w_t\left(i\right) L_t\left(i\right) di + \int_0^1 \Xi_t\left(i\right) di + W_t \tag{3}$$

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

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

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

The first order conditions w.r.t. labour supply and intertemporal aggregate consumption allocation are

$$\frac{w_t(i)}{P_t} = \frac{L_{i,t}^{\nu}}{\widetilde{\Lambda}_t} \tag{5}$$

$$Q_{t,t+1} = \beta \frac{P_t}{P_{t+1}} \frac{\widetilde{\Lambda}_{t+1}}{\widetilde{\Lambda}_t} \tag{6}$$

where we define the marginal utility of consumption as

$$\widetilde{\Lambda}_t = \varepsilon_{Ct}^{\gamma} \left(C_t - hC_{t-1}\right)^{-\gamma} - \beta h \mathrm{E}_t \left[\varepsilon_{Ct+1}^{\gamma} \left(C_{t+1} - hC_t\right)^{-\gamma}\right] \tag{7}$$

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

$$I_t = \beta^{-1} \left\{ \mathrm{E}_t \left[\frac{P_t}{P_{t+1}} \frac{\widetilde{\Lambda}_{t+1}}{\widetilde{\Lambda}_t} \right] \right\}^{-1}. \tag{8}$$

The production function is given by

$$Y_t(i) = A_t L(i)^{\alpha} \tag{9}$$

where A_t is a technology shock.

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 maximise expected profits, namely

$$\max_{P_t^i} \mathrm{E}_t \sum_{s=t}^{\infty} \zeta^{s-t} Q_{t,s} \left(P_s^i Y_s^i - TC_s\right), \tag{10}$$

where TC denotes total costs. 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 the current inflation target Π_t^*. At time s, firms which set their price optimally at time t and have not been able to change it optimally since, will find themselves with a price $P_t^i \left(\frac{1}{\Pi_t^*} \prod_{j=t}^{s} \Pi_s^* \right)^{1-\iota} \left(\frac{P_{s-1}}{P_{t-1}} \right)^{\iota}$, where $0 \leq \iota \leq 1$.

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^*. Furthermore, the average level of prices in the group that does not change prices is partly indexed to the average price level from the last period so that

$$\frac{P_t^*}{P_t} = \left(\frac{1 - \zeta \left(\frac{(\Pi_t^*)^{1-\iota} \Pi_{t-1}^{\iota}}{\Pi_t} \right)^{1-\theta}}{1-\zeta} \right)^{\frac{1}{1-\theta}} \tag{11}$$

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

Firms' decisions can then be characterised as

$$\left(\frac{P_t^*}{P_t}\right)^{1-\theta\left(1-\frac{\nu+1}{\alpha}\right)} = \frac{\chi\theta}{\alpha(\theta-1)}\frac{K_{2,t}}{K_{1,t}} \tag{12}$$

$$K_{2,t} = \frac{A_t^{-\frac{\nu+1}{\alpha}}}{\widetilde{\Lambda}_t} Y_t^{\frac{\nu+1}{\alpha}} + \mathrm{E}_t \zeta Q_{t,t+1} \Pi_{t+1} \left(\frac{\Pi_{t+1}}{\Pi_t^{\iota}\left(\Pi_{t+1}^*\right)^{1-\iota}}\right)^{\theta\frac{\nu+1}{\alpha}} K_{2,t+1} \tag{13}$$

$$K_{1,t} = (1-\tau_t)\, Y_t + \mathrm{E}_t \zeta Q_{t,t+1} \Pi_{t+1}^{\theta} \left(\Pi_{t+1}^*\right)^{(1-\iota)(1-\theta)} \Pi_t^{\iota(1-\theta)} K_{1,t+1} \tag{14}$$

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

$$I_t = \left(\frac{\Pi_t^*}{\beta}\right)^{1-\rho_I} \left(\frac{\Pi_t}{\Pi_t^*}\right)^{\psi_\Pi} \left(\frac{Y_t}{Y_{t-1}}\right)^{\psi_Y} I_{t-1}^{\rho_I} e^{\varepsilon_{t+1}^I} \tag{15}$$

where Y_t is aggregate output, Π_t^* is a stochastic inflation target and ε_{t+1}^I is a serially uncorrelated policy shock.

Some authors, notably Clarida, Galí and Gertler (2000) and Lubik and Schorfheide (2004), have argued that the start of the Volcker era also signed a structural change in US monetary policy, which resulted in a much stronger anti-inflation determination of the Federal Reserve. The change allegedly manifests itself in an increase of the inflation reaction coefficient (ψ_Π in our notation above) in a simple Taylor rule characterisation of monetary policy. Until 1979Q2, monetary policy was allegedly such as to induce an indeterminate equilibrium.

Here, we propose a different interpretation of Federal Reserve behaviour. We maintain fixed the Taylor rule parameters, but allow for the possibility of changes in the inflation target Π_t^*. A lower anti-inflationary determination would therefore be captured by an upward drift of the target. This formulation allows us to abstract from issues of equilibrium determinacy when estimating the model.

Market clearing requires

$$Y_t = C_t. \tag{16}$$

Equilibrium dynamics are described by equations (6)-(8) and (11)-(16), plus the stochastic processes governing the motion of $\varepsilon_{Ct}^{\gamma}$, A_t, Π_t^* and ε_t^I. These are discussed below.

2.1 Solving the model

For all shocks other than the inflation target, we assume that variances are subject to stochastic regime switches. More specifically

$$
\begin{aligned}
A_{t+1} &= A_t^{\rho} e^{\varepsilon_{t+1}^{A}}, \qquad \varepsilon_{t+1}^{A} \,\tilde{}\, N\left(0, \sigma_{a,s_{Y,t}}\right) \\
\varepsilon_{I,t+1} &= e^{\varepsilon_{t+1}^{I}}, \qquad \varepsilon_{t+1}^{I} \,\tilde{}\, N\left(0, \sigma_{i,s_{I,t}}\right) \\
\varepsilon_{Ct+1}^{\gamma} &= \left(\varepsilon_{Ct}^{\gamma}\right)^{\rho_C} e^{\varepsilon_{t+1}^{C}} \qquad \varepsilon_{t+1}^{C} \,\tilde{}\, N\left(0, \sigma_{c,s_{C,t}}\right)
\end{aligned}
$$

where

$$
\begin{aligned}
\sigma_{a,s_{Y,t}} &= \sigma_{a,L} s_{Y,t} + \sigma_{a,H}\left(1 - s_{Y,t}\right) \\
\sigma_{i,s_{I,t}} &= \sigma_{i,L} s_{I,t} + \sigma_{i,H}\left(1 - s_{I,t}\right) \\
\sigma_{c,s_{C,t}} &= \sigma_{c,L} s_{C,t} + \sigma_{c,H}\left(1 - s_{C,t}\right)
\end{aligned}
$$

and the variables $s_{C,t}$, $s_{I,t}$, $s_{Y,t}$ can assume the discrete values 0 and 1. For each variable $s_{j,t}$ $(j = C, I, Y)$, 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 heteroskedasticity in policy shocks aims to capture the large increase in interest rate volatility in the early 1980s, the time of the so-called "monetarist experiment" of the Federal Reserve. Finally, the finance literature has found a relationship between regimes identified in term-structure models and the business cycle. In our model, this relationship could be accounted for by regime switches of the volatility of preferences (demand) shocks.

Concerning the process followed by the inflation target, we assume that

$$
\Pi_{t+1}^{*} = \left(\overline{\Pi}^{*}\right)^{1-\rho_\pi} \left(\Pi_t^{*}\right)^{\rho_\pi} e^{\varepsilon_{t+1}^{\pi}} \tag{17}
$$

so that the inflation target is allowed to change smoothly over time. We also plan to explore the alternative specification in which

$$
\Pi_{t+1}^{*} = \overline{\Pi}_L^{*} s_{\Pi,t} + \overline{\Pi}_H^{*}\left(1 - s_{\Pi,t}\right) \tag{18}
$$

where $s_{\Pi,t}$ can assume the discrete values 0 and 1 and the probabilities of remaining in state 0 and 1 are constant and equal to $p_{\Pi,0}$ and $p_{\Pi,1}$, respectively. As discussed below, specification (18) has the advantage of generating non-zero prices of regime-switching risk.

To solve the model, we exploit the recursive nature of bonds in equilibrium. We first solve for all macroeconomic variables and then construct the prices of bonds of various maturities.

We start by writing the macroeconomic system in compact form as

$$
\begin{aligned}
y_t &= g(z_t, \sigma) && (19) \\
z_{t+1} &= h(z_t, \sigma) + \widetilde{\zeta}(z_t)\, \sigma \widetilde{u}_{t+1} && (20)
\end{aligned}
$$

where $g(\cdot)$, $h(\cdot)$, and $\widetilde{\zeta}(\cdot)$ are matrix functions and we define the vectors: z_t, including the lagged endogenous predetermined variables, the state variables with continuous support and the state variables with discrete support; y_t, collecting all jump variables (excluding bond yields); and $\widetilde{u}_t$, containing all innovations. In order to write the law of motion of the discrete processes in the form implied in equation (20), we rely on Hamilton (1994). The law of motion of state $s_{C,t}$, for example, is written as $s_{C,t+1} = (1 - p_{C,0}) + (-1 + p_{C,1} + p_{C,0})\, s_{C,t} + \nu_{C,t+1}$, where $\nu_{C,t+1}$ is an innovation with mean zero and heteroskedastic variance.

We then seek a second-order approximation to the functions $g(z_t, \sigma)$ and $h(z_t, \sigma)$ around the non-stochastic steady state $z_t = \overline{z}$ and $\sigma = 0$. We define the non-stochastic steady-state as vectors $\overline{y}$ and $\overline{z}$ such that $f(\overline{y}, \overline{y}, \overline{z}, \overline{z})$.

For the continuous state variables, the non-stochastic steady state $\overline{z}$ corresponds to the value which they would eventually attain in the absence of further shocks. For the state variables with discrete support, the non-stochastic steady state is instead the ergodic mean of the Markov chain. Formally, when we take the limit as $\sigma = 0$ we shrink the support of the regime-switching processes, so that their two realisations become closer and closer to each other. Eventually, the two realisations coincide on the ergodic mean of the process.

Amisano and Tristani (2007b) show that the second-order approximate solution can be represented as

$$
\widehat{g}(z_t, \sigma) = F\widehat{z}_t + \frac{1}{2}\left(I_{n_y} \otimes \widehat{z}_t'\right) E\widehat{z}_t + k_{y,s}\sigma^2
$$

and

$$
\widehat{h}(z_t, \sigma) = P\widehat{z}_t + \frac{1}{2}\left(I_{n_z} \otimes \widehat{z}_t'\right) G\widehat{z}_t + k_{z,s}\sigma^2
$$

for vectors $k_{y,s}$, $k_{z,s}$ and matrices F, E, P and G to be determined. Note that $k_{y,s}$ and $k_{z,s}$ are vectors dependent on the realisation of the discrete states.

2.2 Regime switching and the variability of risk premia

Given the solution for inflation and the marginal utility of consumption, we compute bond prices using the method in Hördahl, Tristani and Vestin (2008). The building blocks are the processes followed by the state vector and the approximate solutions for inflation and the marginal utility of consumption, i.e.

$$\begin{aligned}
\widehat{z}_{t+1} &= P\widehat{z}_t + \frac{1}{2}\left(I_{n_z} \otimes \widehat{z}_t'\right) G\widehat{z}_t + k_{z,s}\sigma^2 + \widetilde{\zeta}\left(z_t\right)\sigma\widetilde{u}_{t+1} \\
\widehat{\lambda}_t &= F_\lambda \widehat{z}_t + \frac{1}{2}\widehat{z}_t' E_\lambda \widehat{z}_t + k_{\lambda,s}\sigma^2 \\
\widehat{\pi}_t &= F_\pi \widehat{z}_t + \frac{1}{2}\widehat{z}_t' E_\pi \widehat{z}_t + k_{\pi,s}\sigma^2
\end{aligned}$$

where F_λ and F_π are the appropriate rows of vector F and E_λ and E_π are appropriate sub-matrices of matrix E. In log-deviation from its deterministic steady state, the approximate price of a bond of maturity n, $\widehat{b}_{t,n}$, can then be written as

$$\widehat{b}_{t,n}\left(z_t, \sigma\right) = B_{z,n}\widehat{z}_t + \frac{1}{2}\widehat{z}_t' B_{zz,n}\widehat{z}_t + B_{n,s}\sigma^2$$

where $B_{z,n}$, $B_{zz,n}$ and $B_{n,s}$ are defined through a recursion. $B_{n,s}$ changes depending on the realisation of the discrete states, but matrices $B_{z,n}$ and $B_{zz,n}$ are state-independent.

The state-dependence of $B_{n,s}$ implies that bond risk premia will also become time-varying. In order to see this, it is useful to derive expected excess holding period returns, i.e. the expected return from holding a n-period bond for 1 period in excess of the return on a 1-period bond. To a second order approximation, the expected excess holding period return on an n-period bond can be written as

$$\widehat{hpr}_{t,n} - \widehat{i}_t = \mathrm{Cov}_t\left[\widehat{\pi}_{t+1}, \widehat{b}_{t+1,n-1}\right] - \mathrm{Cov}_t\left[\Delta\widehat{\lambda}_{t+1}, \widehat{b}_{t+1,n-1}\right]$$

This expression can be evaluated using the model solution to obtain

$$\widehat{hpr}_{t,n} - \widehat{i}_t = \sigma^2 B_{n-1,z}\widetilde{\zeta\zeta}'\left(F_\pi' - F_\lambda'\right) \tag{21}$$

where $\widetilde{\zeta\zeta}'$ is the conditional variance-covariance matrix of vector z_t, which depends on state s. In our model, therefore, risk premia change every time there is a switch in any of the discrete state variables.

Since the conditional variance of the price of a bond of maturity n can be written, to a second order approximation, as $\mathrm{E}_t\left[\widehat{b}_{t+1,n-1}\widehat{b}'_{t+1,n-1}\right] = \sigma^2 B_{z,n-1}\widetilde{\zeta\zeta}' B'_{z,n-1}$, it follows that we can define the (microfounded) price of risk for unit of volatility, or the "market prices of risk," in our model as

$$\xi_t \equiv \sigma\widetilde{\zeta}'\left(F'_\pi - F'_\lambda\right) \tag{22}$$

The market prices of risk are only affected by first-order terms in the reduced-form of the model. All terms in equation (22) would be constant in a world with a single regime. They becomes time-varying in our model due to the possibility of regime switches, because the variance-covariance matrix $\widetilde{\zeta\zeta}'$ is regime-dependent.

In the empirical finance literature, the market prices of risk are often postulated exogenously using slightly different specifications. For example, Naik and Lee (1997), Bansal and Zhou (2002) and Ang, Bekaert and Wei (2008) assume that the market prices of risk are regime dependent, but the risk of a regime-change is not priced. On the contrary, regime-switching risk is priced in Dai, Singleton and Yang (2008).

In our model, these specifications can arise endogenously depending on how the regime-switching processes affect the model. Based on the definition $z'_t = [x'_t, s'_t]'$, where vector $\mathbf{x}_t$ only includes the states with continuous support and vector $\mathbf{s}_t$ includes the states with discrete support, we can partition the matrix $\widetilde{\zeta}$ (recall that shocks with continuous and discrete support are all independently distributed) and the vectors F_π and F_λ conformably as

$$\widetilde{\zeta} \equiv \begin{bmatrix} \widetilde{\zeta}^x & 0 \\ 0 & \widetilde{\zeta}^s \end{bmatrix}, \qquad F_\pi \equiv \begin{bmatrix} F^x_\pi \\ F^s_\pi \end{bmatrix}, \qquad F_\lambda \equiv \begin{bmatrix} F^x_\lambda \\ F^s_\lambda \end{bmatrix}$$

As a result, equation (22) can be split into the vectors ξ^x_t and ξ^s_t such that $\xi'_t = \left[(\xi^x_t)', (\xi^s_t)'\right]'$ and

$$\xi^x_t = \sigma\left(\widetilde{\zeta}^x\right)'\left[(F^x_\pi)' - (F^x_\lambda)'\right] \tag{23}$$

$$\xi^s_t = \sigma\left(\widetilde{\zeta}^s\right)'\left[(F^s_\pi)' - (F^s_\lambda)'\right] \tag{24}$$

Vector ξ^x_t in equation (23) includes the prices of risk associated with variables with continuous support. These prices change across regimes. If, for example, technological risk were not diversifiable, then the price of risk associated with technology shocks would be higher in a high-variance regime for technology shocks (and lower in a low-variance regime). This is the regime-dependence of market prices of risk which is present in all the aforementioned finance models.

Vector ξ_t^s in equation (24) includes instead the market prices of regime-switching risk, i.e. the price of risk associated with the possibility of regime changes. These prices of risk are also regime-dependent, because they will be affected by the conditional variance of the discrete process, which depends on the regime prevailing at each point in time.

In our set-up, the prices of risk associated with variables with continuous support, ξ_t^x, will always be non-zero. Whether the prices of regime-switching risk are zero or not depends instead on the exact way in which regime-switching affects the economy. When only the variance of exogenous shocks is allowed to change regime stochastically, the market price of regime-switching risk is zero. The reason is that, as in a model with homoskedastic shocks, variances have no effect on the first order approximation of the model. The possibility that variances may change is therefore also irrelevant, to first order.

On the contrary, the prices of regime-switching risk are non-zero when regime-switching affects other structural elements of the model, as in our specification of the inflation target process in equation (18). In this case, a shift in the inflation target regime would have direct implications on, for example, inflation expectations. As a result, the possibility of such a regime-shift would also command a non-zero market price.

Our set-up can therefore offer a microfoundation for the different assumptions adopted in the finance literature. It should be emphasised, however, that papers in the finance literature also allow the prices of risk to be affine functions of the continuous states of the model. This would only be possible in our set-up if we solved the model to third order.

3 Estimation methodology

Looking at the system of equations (**??**) and (**??**), given that discrete state variables appear linearly and in a quadratic way, the system can be re-written as quadratic in the continuous state variables with interercept and linear terms changing according to the discrete state variables This alternative representation is particularly convenient for describing the estimation methodology. It is straigthforward to show that the model can be rewritten as

$$\mathbf{y}^o_{t+1} = \mathbf{c}_j + \mathbf{C}_{1,j}\mathbf{x}_{t+1} + \mathbf{C}_2\overline{vech}(\mathbf{x}_{t+1}\mathbf{x}'_{t+1}) + \mathbf{D}\mathbf{v}_{t+1} \tag{25}$$

$$\mathbf{x}_{t+1} = \mathbf{a}_i + \mathbf{A}_{1,i}\mathbf{x}_t + \mathbf{A}_2\overline{vech}(\mathbf{x}_t\mathbf{x}'_t) + \mathbf{B}_i\mathbf{w}_{t+1} \tag{26}$$

$$\mathbf{s}_t \backsim \text{Markov switching} \tag{27}$$

where the vector $\mathbf{y}^o_t$ includes all observable variables, vector $\mathbf{x}_t$ only includes the states with continuous support, vector $\mathbf{s}_t$ includes the states with discrete support, 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 $\mathbf{c}_j$, the slope coefficients $\mathbf{A}_{1,i}$ and $\mathbf{C}_{1,j}$, 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 the approximation of the state space form is truncated to the linear terms, then the system becomes

$$\mathbf{y}^o_{t+1} = \mathbf{c}_j + \mathbf{C}_1\mathbf{x}_{t+1} + \mathbf{D}\mathbf{v}_{t+1} \tag{28}$$

$$\mathbf{x}_{t+1} = \mathbf{a}_i + \mathbf{A}_1\mathbf{x}_t + \mathbf{B}_i\mathbf{w}_{t+1} \tag{29}$$

$$\mathbf{s}_t \backsim \text{Markov switching} \tag{30}$$

i.e. a linear system with (conditionally) Gaussian innovations and intercepts and loading factors which depend on the value of the discrete state variables. We describe how to obtain the likelihood of the model separately for the linear and the quadratic cases. With the likelihood in hand and a choice for prior specification, estimation is carried out by posterior simulation.

3.1 The linear case

In the linear case, we have 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

3.2 The quadratic case

In the quadratic case the likelihood cannot be obtained in closed form by using recursive methods. We therefore use sequential Monte Carlo techniques. See Amisano and Tristani (2007a) and the reference therein. We construct the likelihood using the conditional particle filter, as in Amisano and Tristani (2007a).. At each point in time t, this involves the following steps:

1. linearise the measurement equation around $\overline{\mathbf{x}}_{t+1|t} = E(\mathbf{x}_{t+1}|\underline{\mathbf{y}}_t^o, \boldsymbol{\theta})$;
2. use the linearised measurement equation to obtain $\widetilde{p}(\underline{\mathbf{y}}_{t+1}^o|\underline{\mathbf{y}}_t^o, \mathbf{s}_{t+1} = j, \underline{\mathbf{y}}_t^o, \mathbf{x}_t, \boldsymbol{\theta})$;
3. draw $\mathbf{s}_{t+1}$ from its marginal distribution (ie independently of $\mathbf{x}_{t+1}$);
4. draw $\mathbf{x}_{t+1}$ from the distribution conditional on $\mathbf{s}_{t+1}$
5. assign the resampling weight

$$w_{t+1} \propto \widetilde{p}(\mathbf{y}_{t+1}^o|\mathbf{x}_t, \boldsymbol{\theta}) = \sum_{j=1}^{m} \widetilde{p}(\mathbf{y}_{t+1}^o|\mathbf{x}_t, \mathbf{s}_{t+1} = j, \boldsymbol{\theta}) \times p_{ij}$$

The sample mean of the weights approximates the conditional likelihood.

To initialise the system, we proceed differently for discrete and continuous states.

For $\mathbf{s}_0$, we draw from the ergodic distribution $\boldsymbol{\pi}$, which is simple to compute state by state, given that each variable can only assume two values. For $\mathbf{x}_0$ the ergodic distribution is unknown. We approximate it with a Gaussian distribution matching the (analytically available) first and second moments of its ergodic distribution obtained by using the linear approximation.

4 Data and prior distributions

We estimate the model on quarterly US data over the sample period from 1966Q1 to 2006Q2. 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.[1]

[1] 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

The data included in the information set are real GDP, the GDP deflator, the 3-month nominal interest rate and yields on 3-year and 10-year zero-coupon bonds. Prior to estimation, GDP is de-meaned and detrended using a linear trend.

For most model parameters, we assume prior distributions broadly in line with the literature (see Tables 1-3). We only discuss here the priors for the parameters related to the regime switching processes. More specifically, we set the prior means for the standard deviations of policy, preference and technology shocks so as to induce an ordering in which state 0 is the high-volatility state.

Concerning transition probabilities, we assume beta priors such that the probabilities of persistence in each state are symmetric. We assume that they have relatively high means for regimes associated to monetary policy and technology shocks, a bit less high for preference shocks. This is consistent with the aforementioned conjecture that monetary policy shocks and technology shocks should be associated with highly persistent states, while preference shocks should be associated with an indicator of the business cycle. The variances of these prior distributions are relatively large, so as to extract as much information as possible from the data.

5 Empirical results

We have estimated our model under the simplifying assumption of absence of regime-switching and introducing incrementally regime switching in $s_{I,t}$, $s_{C,t}$ and $s_{Y,t}$. We refer to the model with a single regime as $M0$ and to the other models as $M1$, $M2$ and $M3$, where the digit refers to the number of discrete processes included in the specification. We denote the estimates of the first order (or linear) approximation of these models with $M0L$, $M1L$, $M2L$ and $M3L$; $M0Q$ will denote estimates of the second-order (or quadratic) approximation of the homoskedastic model $M0$.

Since $M3L$ dominates $M1L$ and $M2L$ in terms of marginal likelihood, we focus here on the comparison between $M0L$, $M0Q$ and $M3L$.

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.

5.1 Posterior distributions and goodness of fit

Tables 1-3 also report statistics on the posterior distributions of parameter estimates. The results highlight that all models find it hard to replicate macro and yields data at the same time.

The first sign of strain arises from the marked increase, compared to the prior mean, of the posterior mean of the standard deviations of almost all shocks. For example, compared to a prior mean around 1%, the standard deviation of preference shocks increase to 18% in $M0L$, to 24% in $M0Q$ and to between 15% and 26% in $M3L$. Large standard deviations tend to be necessary in order to produce movements in 10-year yields, which would otherwise tend to stay close to their long-run mean in an environment where the expectations hypothesis holds (see also Gürkaynak, Sack and Swanson, 2005).[2]

The increase is particularly large (more than 10-fold) for the standard deviation of the target shock. This increase must be interpreted jointly with the estimates of the posterior means of the policy rule coefficients. In all models, the policy rule becomes very aggressive against inflation deviations from target, with short-term reaction coeffients around 1.0 and a degree of interest rate smoothing which also hovers around 1. These coefficients imply that inflation is almost always kept on target by the central bank. All models are therefore forced to explain the inflation rates observed in our sample as induced by the central bank through a sequence of target shocks. This feature also explains the low posterior mean of the inflation indexation parameter.

In turn, the aggressiveness of the policy rule is related to the need of generating sufficient movements at the long-end of the yield curve. Very inertial (even superinertial, for the model with regime-switches) rules obviously help in this sense. At the same time, inertial rules tend to be associated with gradualism in interest rate setting. A large inflation response coefficient counters this tendency and induces sufficient volatility in the short-term rate.

Turning to the structural parameters, the most striking result is the large increase in the posterior mean of the coefficient of relative risk aversion. In the linear models $M0L$ and $M3L$ this is obviously not due to a standard, equity-premium-puzzle type of reason: these models explain observed variables entirely through expectations effects. The reason

[2] Even in model $M0Q$ a weak version of the expectations hypothesis holds because risk-premia are constant.

for the high coefficient risk aversion is rather related to the link between this parameter and the elasticity of intertemporal substitution ($1/\gamma$). To a first order approximation, the elasticity of intertemporal substitution shapes the sensitivity of output to changes in the real interest rate. Given the aforementioned estimates of the policy rule coefficients, γ must be high to shield output from the volatililty of the short-term.

Overall, the posterior distribution have unreasonable implications. For example, they would imply implausibly high values of the unconditional variances of all observed variables.

Nevertheless, there are two signals that the quadratic model with regime-switching could have a better performance. The first is that, in the low-variance state, the regime-switching standard deviations of the exogenous shocks tend to be smaller than the corresponding standard deviations of the $M0$ models. At the same time, the posterior estimates of the transition probabilities suggest that the low-variance states are more persistent than the high-variance ones. Overall, this implies that the ergodic variance of the shocks is not necessarily higher than in the homoskedastic case, even if, at the same time, the model with regime-switching would be able to occasionally generate bursts in volatility, hence in risk premia.

The second signal is that, while remaining high, the coefficient of relative risk-aversion estimated in the $M0Q$ model is lower than in the other models. This suggests that the additional restrictions imposed by the quadratic model, e.g. the influence of γ on risk premia, help to discipline its estimates. There are no signals that this worsens the model's ability to fit the data.

Turning to goodness of fit measures, there are clear signals that the model with regime-switches is superior to the homoskedastic models.

Table 4 reports estimates of the marginal likelihood of the three models discussed so far. The model with regime-switching is, to different extents, clearly superior to the two alternatives. This suggests that the need to introduce sources of heteroskedasticity in economic models is not linked to the desire to fit financial data, but rather necessary for a satisfactory explanation of macroeconomic data.

Figures 1-3 display 1-step-ahead forecasts and realised variables for each of the three models. The striking feature emerging from these figures is that all models are capable of fitting the data to a surprisingly good extent. What is particularly noticeable is that the level of yields can be matched by the linear models. Within linearised models, Bekaert, Cho and Moreno (2006) and De Graeve, Emiris and Wouters (2007) fit yields only by introducing exogenous parameters to explain their unconditional slope. In our case, however, the unconditional slope is zero. Nevertheless the models manage to replicate it in sample, thanks to the high persistence of the exogenous shocks.

The residuals from the $M0Q$ model are significantly smaller in absolute value than in the other models. This also reflects this model's ability to generate a non-zero unconditional slope of the term structure, which helps to match yields. Even in this case, however, residuals display clear signs of serial correlation.

A second feature which emerges from Figures 1-3 is the clear heteroskedasticity of the residuals. This is problematic for the $M0L$ and the $M0Q$ models, while it is explained by the model with regime-switching. A particularly visible increase in the variance of residuals is observed in the linear models for all interest rates at the beginning of the 1980s, the time of the so-called monetarist experiment of the Federal Reserve. Similarly, a reduction in the volatility of output shocks is clearly visible as of the mid-1980s, as highlighted in the literature on the Great moderation.

5.2 Implications of regime switching

Figure 4 displays smoothed and filtered estimates of the discrete states in model $M3L$, together with the official NBER recession dates. In all cases, 1 denotes the low-variance state, 0 the high-variance state.

The regimes associated with the policy shock clearly identify the Fed's monetarist experiment. This state jumps abruptly to the value 0 in 1980 and remains there until 1983; it then returns to the low-variance state over the rest of the sample (with a marginal exception around 1985). The identification of the two states is quite precise in real time. There are only small revisions noticeable in the smoothed estimates, compared to the filtered ones.

The regimes associated with preference shocks displays some association with the economic cycle. More specifically, it tends to move towards the low-variance state during prolonged expansions – e.g. at the end of the 1990s or around 2005. Overall, however, the association with the cycle is not strong.

Finally, the regimes associated with technology shocks clearly identifies the Great moderation period started in the mid-1980s. The switch to a low-variance regime occurs gradually over the 1980s and it is quite clearly identified also in real time. In previous years, however, only smoothed estimates confim that the economy was in a high-variance regime for technology shocks. Filtered estimates are much more volatile and tend to repeatedly move away towards the low-variance regime.

The various states can be composed to define 8 possible combinations of regimes. This is done to construct Figure 5, which displays excess holding period returns derived from the model. More precisely, the figure is based on the posterior means and on the filtered estimates of the regimes obtained from the estimation of the linearised $M3L$ model. These results are then used in the second order approximation of the same model to compute excess holding period returns. In so doing, we exploit the properties of these measures of risk premia highlighted in the discussion of equation (21). These premia vary over time only as a result of regime changes. They can therefore be computed without the need to filter the states with a continuous support.

Two notable features emerge from Figure 5. The first one is that the quadratic model is capable of generating sizable risk-premia. Premia are strictly increasing in the maturity of bonds and reach an average of 3-4 percentage points at the 10-year horizon. This value should obviously not be interpreted as a term premium, but it gives an indication that the model can go quite far in generating sizable premia. Some reduction in the variance of structural shocks appears to be possible when estimating the second order approximation of the model, without necessarily loosing much in terms of the model's ability to explain yields.

The second feature emerging from Figure 5 is that the premia are significantly variable over time, which is a desirable feature to explain observed deviations of the data from features consistent with the expectations hypothesis (see e.g. Dai and Singleton, 2002). A clear peak in risk-premia (up to over 5 percentage points at the 10-year horizon) is visible at the time of the monetarist experiment in the early 1980s. This is encouraging, because deviations of yields from values consistent with the expectations hypothesis are known to be particularly marked around this period. For example, Rudebusch and Wu (2006) note that the performance of the expectations hypothesis improves in the 1988-2002 period. At the 10-year maturity, premia tend to be more volatile in the first half of the sample than in the most recent half, consistently with the reduction in the variance of technology shocks.

In this model, variations in risk premia are associated with variations in the amount of uncertainty in the economy. Figure 6 displays a measure of such uncertainty: the conditional standard error of one-step-ahead forecasts for each observable series. This figure complements the information in Figure 5. It shows that the short-term rate and the 3-year yield are particularly difficult to forecast at the beginning of the 1980s. For 10-year bonds, however, volatility is also quite high during the recessions of the 1970s and the recession at the turn of the century.

6 Conclusions

Our preliminary results on the estimation of the first order approximation of a macro-yield curve model with regime switches show considerable support for this specification, compared to a model with homoskedastic shocks. Different regimes clearly help fitting macroeconomic variables, notably the heteroskedasticity of the model's residuals. Moreover, estimated regimes bear an intuitively appealing structural interpretation.

At the same time, the linearised model displays clear signs of mispecification when asked to match yields data. It can only do so at the cost of bending the parameter estimates towards implausible regions from a macroeconomic viewpoint. This problem should be mitigated when the model is estimated to a second order approximation.

References

[1] Amisano, G. and O. Tristani (2007a), Euro area inflation persistence in an estimated nonlinear DSGE model, ECB WP No. 754.

[2] Amisano, G. and O. Tristani (2007b), "Solving DSGE models with non-normal shocks using a second-order approximation to the policy function," presented at the CEF2007 conference in Montreal

[3] Ang, Andrew and G. Bekaert (2002a), "Regime Switches in Interest Rates", Journal of Business and Economic Statistics 20, 163-182.

[4] Ang, Andrew and G. Bekaert (2002b), "Short Rate Nonlinearities and Regime Switches", Journal of Economic Dynamics and Control 26, 1243-1274.

[5] Ang, Andrew, G. Bekaert and Min Wei (2008), "The Term Structure of Real Rates and Expected Inflation," Journal of Finance 63, 797-849.

[6] Bansal, R. and Hao Zhou (2002), "Term Structure of Interest Rates with Regime Shifts", Journal of Finance 57, 1997-2043.

[7] Bansal, R., G. Tauchen and Hao Zhou (2004), "Regime-Shifts, Risk Premiums in the Term Structure, and the Business Cycle", Journal of Business and Economic Statistics.

[8] Bekaert, G., Seonghoon Cho and Antonio Moreno (2006), "New-Keynesian Macroeconomics and the Term Structure," mimeo.

[9] Bikbov, R. and M. Chernov (2007), "Monetary Policy Regimes and The Term Structure of Interest Rates", mimeo, London Business School.

[10] Canova, Fabio and Luca Sala (2006), "Back to Square One: Identification Issues in DSGE Models", ECB Working Paper No. 583.

[11] Dai, G. and K. Singleton (2002), "Expectations puzzles, time-varying risk premia, and affine models of the term structure," *Journal of Financial Economics* 63, 415-441.

[12] Dai, Qiang, K. Singleton and Wei Yang (2008), "Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields," Review of Financial Studies, forthcoming.

[13] De Graeve, Ferre, Marina Emiris and Raf Wouters (2007), A Structural Decomposition of the US Yield Curve, mimeo.

[14] Doh, Taeyoung (2006), "What Moves the Yield Curve? Lessons from an Estimated Nonlinear Macro Model," manuscript, University of Pennsylvania.

[15] Fuhrer, J. (1996), "Monetary policy shifts and long-term interest rates," *Quarterly Journal of Economics* 111, pp. 1183-1209.

[16] Goodfriend, M. (1991), "Interest Rates and the Conduct of Monetary Policy," *Carnegie-Rochester Conference Series on Public Policy* 34, 7–30.

[17] Gürkaynak, Refet S., Brian Sack and Eric Swanson (2005), "The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models," American Economic Review 95, 425-436

[18] Hamilton, James D. (1988), "Rational-expectations econometric analysis of changes in regime : An investigation of the term structure of interest rates", Journal of Economic Dynamics and Control 12, 385-423

[19] Hamilton, James D. (1994), Time series analysis, (Princeton University Press, Princeton).

[20] Hördahl, P., O. Tristani and D. Vestin (2007), "The yield curve and macroeconomic dynamics", ECB Working Paper No 832.

[21] Lubik, Thomas and Frank Schorfheide (2004), "Testing for Indeterminacy: An Application to U.S. Monetary Policy," American Economic Review 94, 190-217

[22] McDonnell, M. M. and Perez-Quiros, G. (2000) "Output Fluctuations in the United States: What has Changed Since the Early 1980s?", American Economic Review 90, 1464-1467

[23] Naik, Vasant and Moon Hoe Lee (1997), "Yield curve dynamics with discrete shifts in economic regimes: theory and estimation", Working Paper Faculty of Commerce, University of British Columbia

[24] Ravenna, F. and J. Seppala (2007a), "Monetary Policy, Expected Inflation and Inflation Risk Premia", mimeo, University of California Santa Cruz

[25] Ravenna, F. and J. Seppala (2007b), "Monetary Policy and Rejections of the Expectations Hypothesis", mimeo, University of California Santa Cruz

[26] Rudebusch, Glenn D. (2002), "Term structure evidence on interest rate smoothing and monetary policy inertia", Journal of Monetary Economics 49, 1161-1187.

[27] Rudebusch, Glenn D. and Eric Swanson (2007), "Examining the bond premium puzzle with a DSGE model", mimeo.

[28] Rudebusch, Glenn D., Brian Sack, and Eric Swanson (2007), "Macroeconomic Implications of Changes in the Term Premium," Federal Reserve Bank of St. Louis, Review 89, 241–69.

[29] Rudebusch, G. and T. Wu (2006), "Accounting for a Shift in Term Structure Behavior with No-Arbitrage and Macro-Finance Models," *Journal of Money, Credit and Banking*, forthcoming.

[30] Sims, Christopher A. and Tao Zha (2007), Were There Regime Switches in U.S. Monetary Policy?, American Economic Review 96, 54 - 81.

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

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

Table 1: Parameter estimates: $M0L$ model

	post mean	**post sd**	**post low q**	**post up q**	**prior mean**	**prior sd**	**prior low q**	**prior up q**
Π^*	1.00938	0.00155	1.00631	1.01239	1.00499	0.00314	1.00085	1.01282
ρ_π	0.85337	0.02117	0.80886	0.89293	0.89983	0.02981	0.83414	0.95036
σ_π	0.00335	0.00043	0.00287	0.00396	0.00010	0.00014	0.00000	0.00050
ρ_c	0.93226	0.01834	0.88692	0.95799	0.90021	0.02982	0.83429	0.95042
σ_c	0.17910	0.02819	0.13961	0.22873	0.00989	0.01406	0.00001	0.04956
ρ_A	0.98373	0.00621	0.97913	0.98849	0.95004	0.02164	0.89932	0.98323
σ_A	0.02821	0.00588	0.01929	0.04078	0.01007	0.01419	0.00001	0.05091
σ_I	0.00284	0.00034	0.00250	0.00326	0.00050	0.00071	0.00000	0.00250
ψ_π	0.81235	0.17546	0.53993	1.20169	0.20124	0.16425	0.01459	0.62860
ψ_y	0.06872	0.02834	0.01931	0.12682	0.04987	0.04086	0.00350	0.15631
ρ_I	0.86999	0.05507	0.79881	0.94420	0.90204	0.20045	0.50909	1.29247
ι	0.04055	0.08648	0.00787	0.07837	0.59974	0.20009	0.19152	0.93249
ν	3.41566	1.02770	1.78588	5.75488	0.99836	0.81082	0.07361	3.08829
γ	14.24830	2.03230	10.65812	17.87543	2.00146	0.70443	1.12276	3.78020
ζ	0.70081	0.05782	0.57192	0.79694	0.60099	0.14774	0.29864	0.86489
h	0.35479	0.06377	0.24050	0.47937	0.70006	0.13799	0.39945	0.92471
θ	10.16136	2.86917	5.56182	16.39994	8.00409	2.66113	3.80330	14.08082
β	0.98999	0.00153	0.98695	0.99296	0.98020	0.01378	0.94547	0.99755
σ_{me40}	0.00110	0.00012	0.00097	0.00125	0.00100	0.00142	0.00000	0.00504

These results are based on 200000 simulations. Legend: "sd" denotes the standard deviation; "low q" and "up q" denote the 5th and 95th percentiles of the distribution. Acceptance rate: .73. Priors: beta distribution for β, h, ι, ζ, ρ_ξ, ρ_g, ρ_π; gamma distribution for ψ_π, ψ_y and all standard deviations; shifted gamma distribution (domain from 1 to ∞) for γ, ϕ, Ξ, Π^*; normal distribution for ρ_i.

Table 2: Parameter estimates: $M0Q$ model

	post mean	**post sd**	**post low q**	**post up q**	**prior mean**	**prior sd**	**prior low q**	**prior up q**
Π^*	1.009101	0.001246	1.00638	1.011464	1.004993	0.003157	1.000837	1.012855
ρ_π	0.812753	0.024431	0.762106	0.857434	0.900101	0.029681	0.834622	0.950422
σ_π	0.003145	0.000295	0.002681	0.003801	9.98E-05	0.000141	1.07E-07	0.000501
ρ_c	0.907751	0.019283	0.862941	0.938963	0.899835	0.029836	0.83455	0.950198
σ_c	0.241292	0.030861	0.187301	0.307759	0.010019	0.014217	9.13E-06	0.050201
ρ_A	0.979256	0.002939	0.972918	0.984342	0.949871	0.021628	0.900069	0.983264
σ_A	0.096721	0.009691	0.078797	0.115796	0.009926	0.013946	1.05E-05	0.048995
σ_I	0.003096	0.00025	0.002681	0.003649	0.000503	0.000712	4.99E-07	0.002536
ψ_π	0.867747	0.260221	0.526316	1.514195	0.20058	0.164109	0.014301	0.626527
ψ_y	0.089573	0.02534	0.040196	0.138254	0.049661	0.040797	0.003644	0.154901
ρ_I	0.91097	0.057745	0.815132	1.028609	0.89739	0.199708	0.505812	1.289396
ι	0.032308	0.017586	0.006726	0.073756	0.599276	0.200708	0.195229	0.932309
ν	0.00131	0.002452	3.11E-05	0.007132	0.992976	0.812001	0.071032	3.111754
γ	9.731058	1.648056	6.851732	13.32249	1.999478	0.704177	1.12047	3.776801
ζ	0.836738	0.050829	0.72467	0.909527	0.599237	0.147891	0.297063	0.861572
h	0.648677	0.053076	0.549132	0.749448	0.700138	0.138284	0.399005	0.924642
θ	13.05556	5.15664	6.475044	26.33228	7.996159	2.650672	3.794668	14.04297
β	0.990911	0.001326	0.988342	0.993546	0.980231	0.013764	0.945504	0.997551
σ_{me40}	0.001281	7.73E-05	0.001133	0.00143	0.001001	0.001421	1.01E-06	0.005081

These results are based on 2000000 simulations. Legend: "sd" denotes the standard deviation; "low q" and "up q" denote the 5th and 95th percentiles of the distribution. Acceptance rate: .78. Priors: beta distribution for β, h, ι, ζ, ρ_ξ, ρ_g, ρ_π; gamma distribution for ψ_π, ψ_y and all standard deviations; shifted gamma distribution (domain from 1 to ∞) for γ, ϕ, Ξ, Π^*; normal distribution for ρ_i.

Table 3: Parameter estimates: $M3L$ model

	post mean	**post sd**	**post low q**	**post up q**	**prior mean**	**prior sd**	**prior low q**	**prior up q**
$p_{x,11}$	0.9858	0.0090	0.9642	0.9978	0.8997	0.0909	0.6614	0.9973
$p_{x,11}$	0.8886	0.0744	0.7007	0.9869	0.9004	0.0907	0.6635	0.9972
$p_{c,11}$	0.8796	0.0797	0.6855	0.9867	0.7455	0.1303	0.4569	0.9507
$p_{c,11}$	0.8134	0.0956	0.5996	0.9733	0.7820	0.1266	0.4895	0.9671
$p_{A,11}$	0.9429	0.0571	0.7702	0.9920	0.9000	0.0907	0.6615	0.9972
$p_{A,11}$	0.9239	0.0624	0.7706	0.9912	0.9003	0.0902	0.6653	0.9971
$\sigma_{I,1}$	0.0026	0.0002	0.0023	0.0030	0.0096	0.0030	0.0056	0.0168
$\sigma_{I,0}$	0.0097	0.0020	0.0068	0.0142	0.0155	0.0076	0.0077	0.0344
$\sigma_{c,1}$	0.1572	0.0331	0.0858	0.2201	0.0096	0.0029	0.0056	0.0168
$\sigma_{c,0}$	0.2930	0.0704	0.1925	0.4684	0.0155	0.0080	0.0078	0.0346
$\sigma_{A,1}$	0.0276	0.0058	0.0183	0.0407	0.0287	0.0088	0.0169	0.0505
$\sigma_{A,0}$	0.0438	0.0114	0.0292	0.0736	0.0466	0.0231	0.0234	0.1045
Π^*	1.0080	0.0018	1.0043	1.0112	1.0050	0.0032	1.0008	1.0129
ρ_π	0.8833	0.0164	0.8483	0.9122	0.9800	0.0139	0.9457	0.9975
σ_π	0.0030	0.0003	0.0025	0.0038	0.0003	0.0002	0.0001	0.0007
ρ_x	0.9266	0.0245	0.8641	0.9579	0.9000	0.0299	0.8345	0.9505
ρ_A	0.9841	0.0029	0.9782	0.9891	0.9801	0.0138	0.9453	0.9975
ψ_π	1.1486	0.3679	0.5964	1.9699	0.2007	0.1656	0.0140	0.6333
ψ_y	0.0494	0.0312	0.0045	0.1191	0.0500	0.0411	0.0036	0.1565
ρ_I	1.0020	0.0811	0.8765	1.1855	0.8978	0.2000	0.5057	1.2901
ι	0.0322	0.0201	0.0058	0.0822	0.5992	0.2009	0.1899	0.9333
ν	3.2054	1.1303	1.5808	6.0579	1.0004	0.8199	0.0714	3.1490
γ	15.1056	2.0833	11.1973	19.4460	2.0023	0.7118	1.1209	3.8011
ζ	0.6662	0.0589	0.5470	0.7748	0.5998	0.1474	0.3005	0.8617
h	0.4363	0.0633	0.3086	0.5597	0.6994	0.1387	0.3985	0.9256
θ	9.8234	2.7239	5.6461	15.9354	7.9942	2.6310	3.8268	13.9742
β	0.9889	0.0018	0.9850	0.9921	0.9801	0.0138	0.9456	0.9975
$\sigma_{me,40}$	0.0011	0.0001	0.0010	0.0013	0.0014	0.0010	0.0006	0.0037

These results are based on 1000000 simulations. Legend: "sd" denotes the standard deviation; "low q" and "up q" denote the 5th and 95th percentiles of the distribution. Acceptance rate: .44 Priors: beta distribution for β, h, ι, ζ, ρ_ξ, ρ_g, ρ_π; gamma distribution for ψ_π, ψ_y and all standard deviations; shifted gamma distribution (domain from 1 to ∞) for γ, ϕ, Ξ, Π^*; normal distribution for ρ_i.

Table 4: Marginal likelihoods

model	**marginal likelihood**
M0L	3584.8
M0Q	3554.5
M1L	3594.9
M2L	3620.5
M3L	3626.2

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.

Figure 1: Actual variables and 1-step-ahead predictions: $M0L$ model

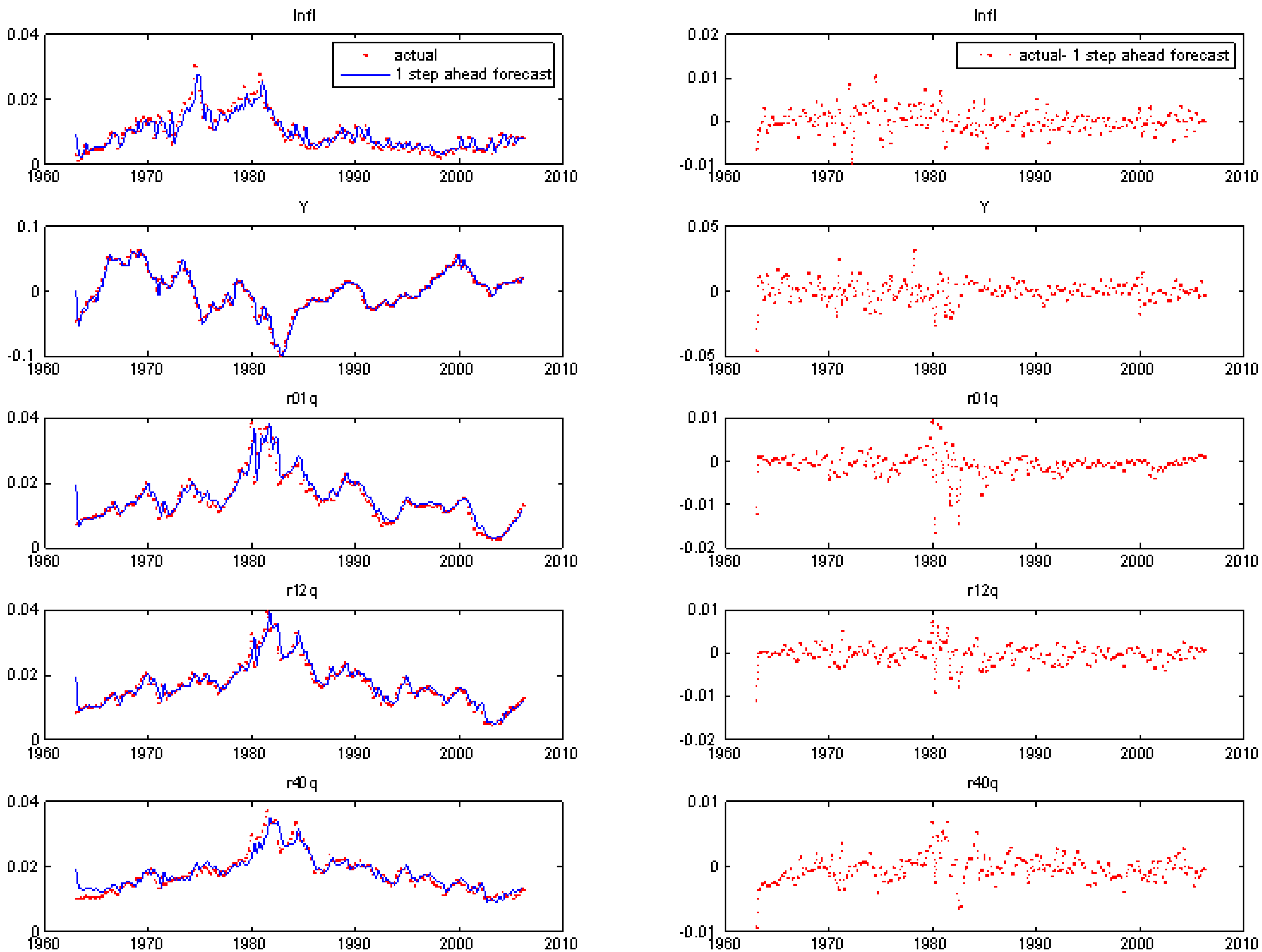


Figure 2: Actual variables and 1-step-ahead predictions: $M0Q$ model

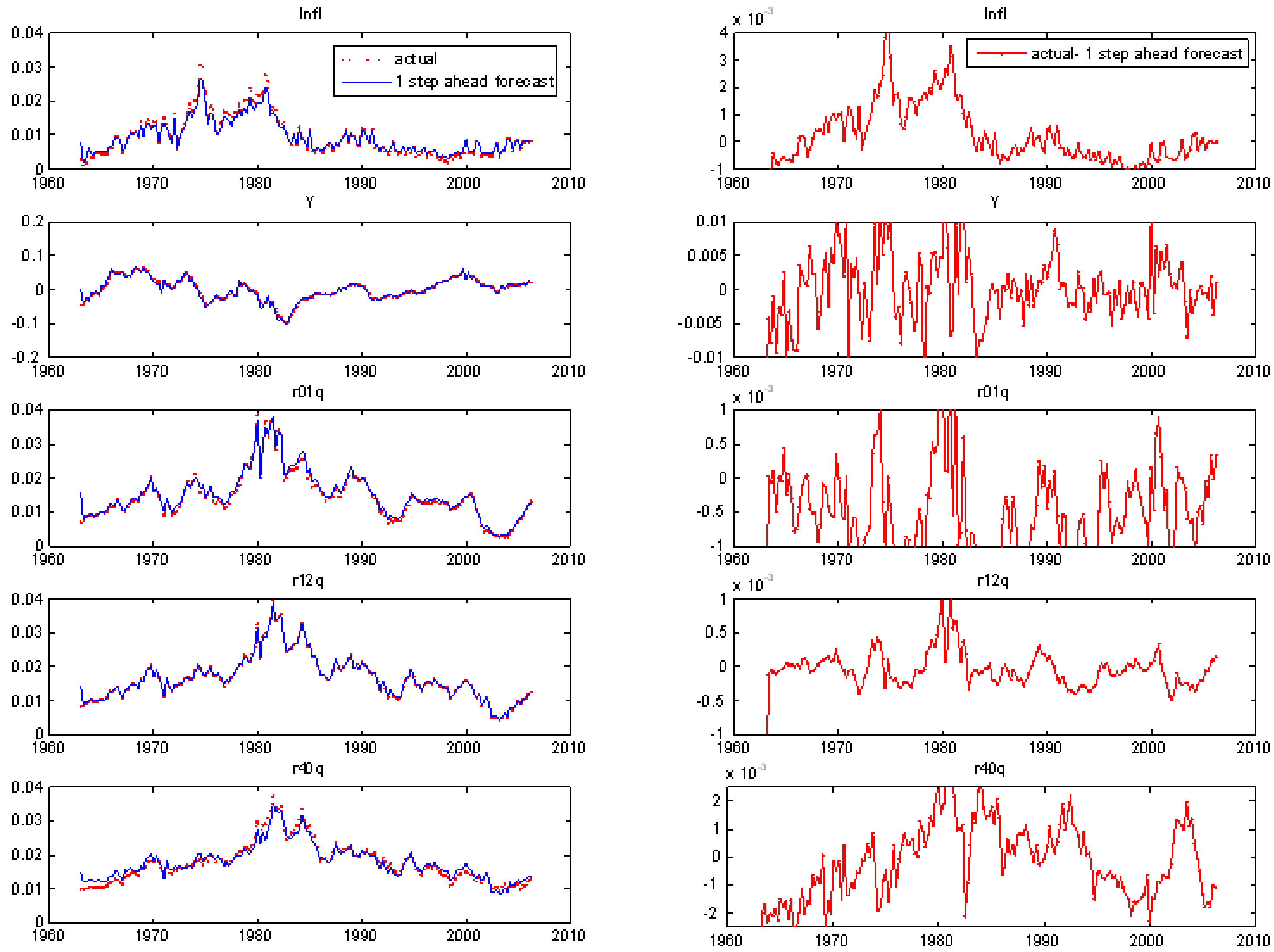


Figure 3: Actual variables and 1-step-ahead predictions: $M3L$ model

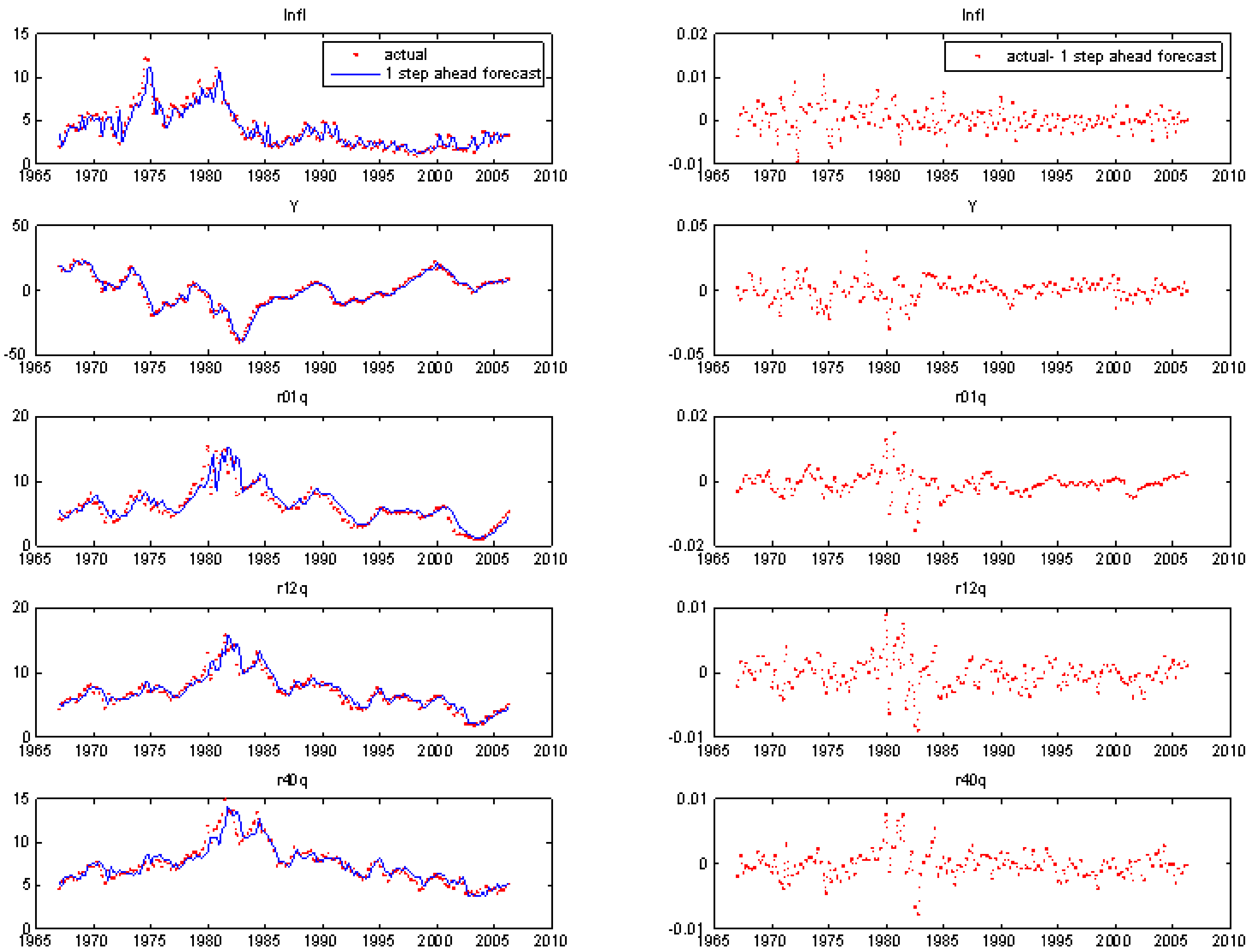
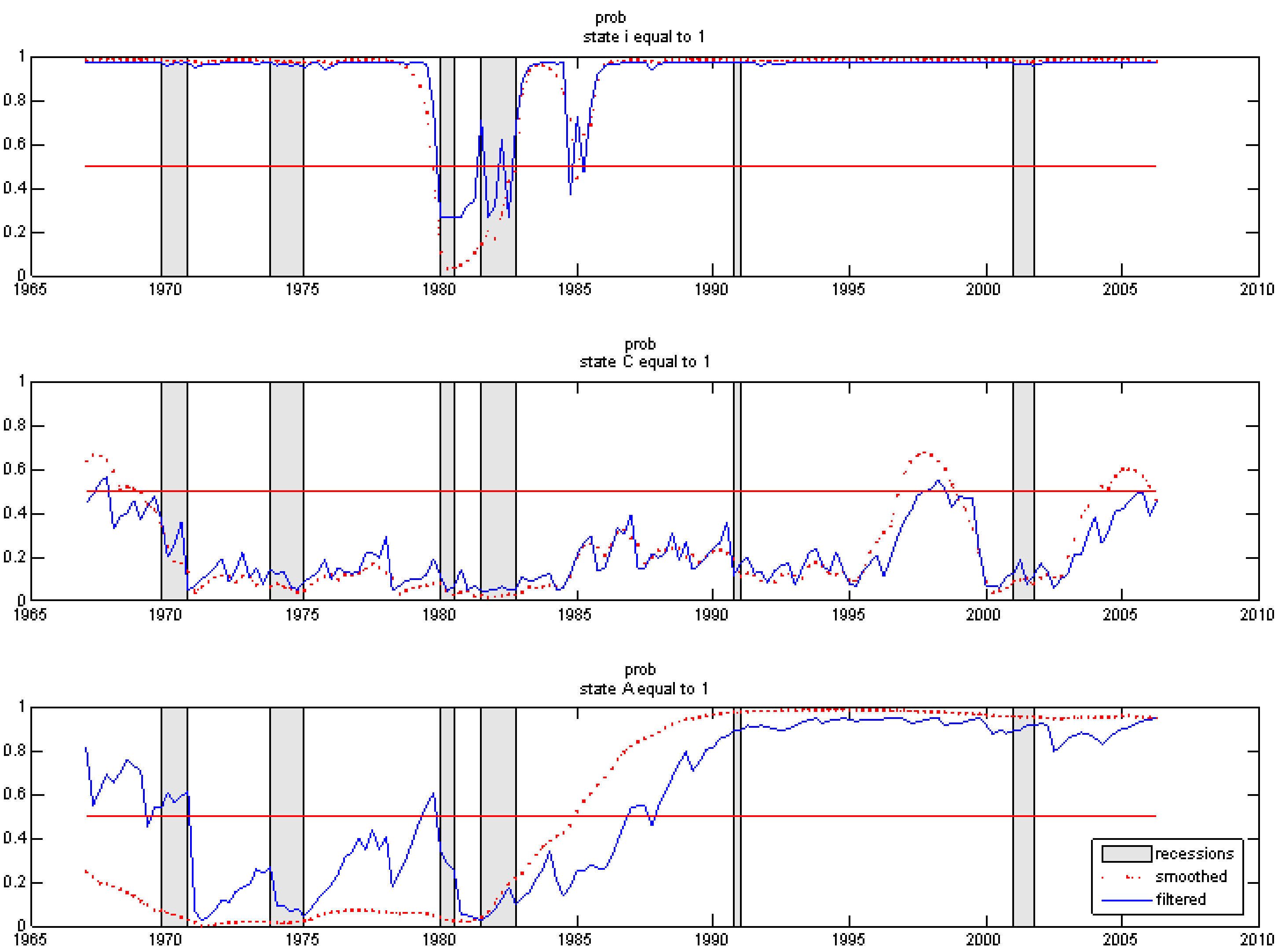
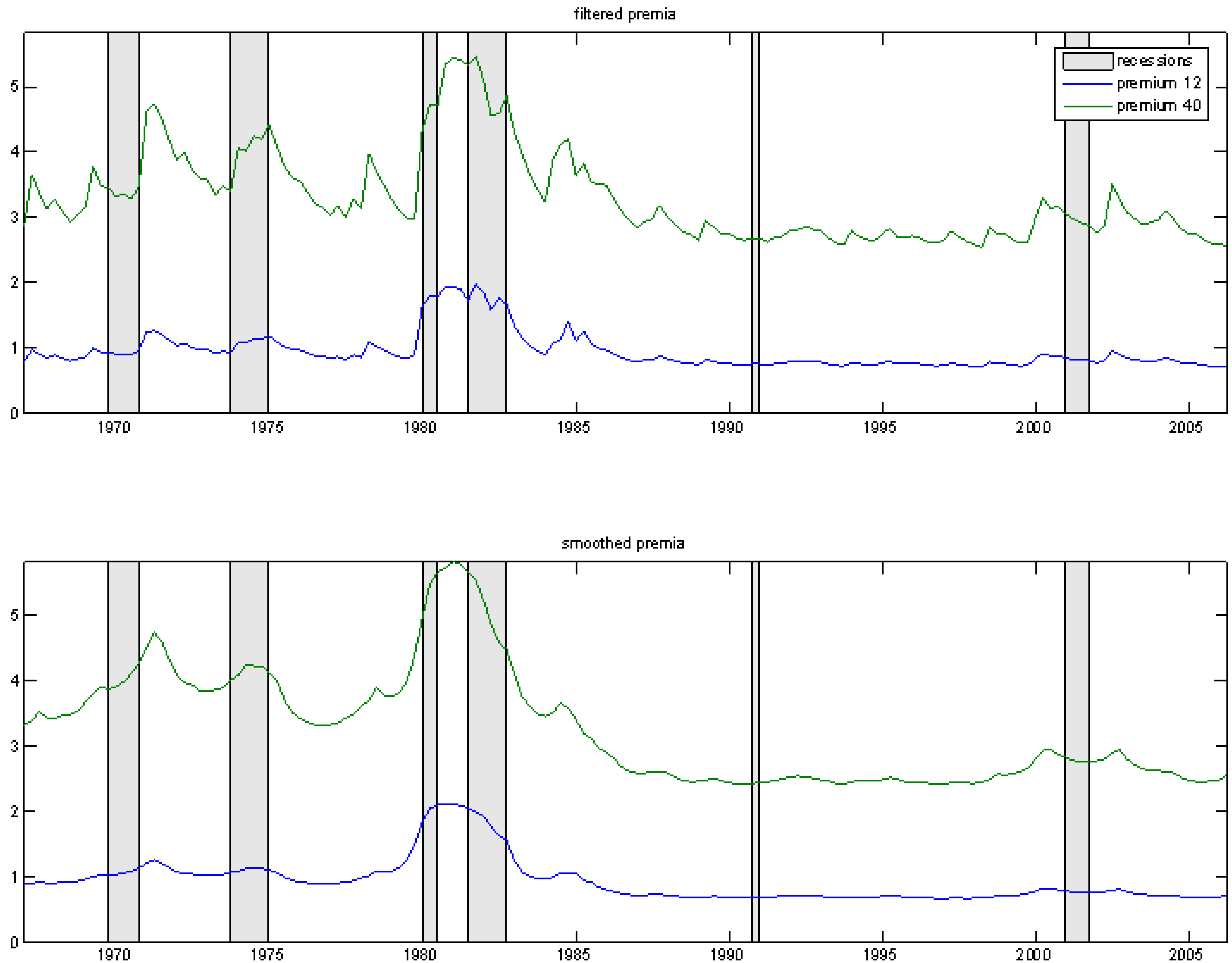


Figure 4: Filtered and smoothed estimates of the regime-variables: $M3L$ model



Legend: "state i" corresponds to the policy shock; "state C" corresponds to the preference shock; "state A" corresponds to the technology shock.

Figure 5: Term premia in the $M3Q$ model



The premia are computed in the second-order approximation of the model with regime-switches, using the parameter means and the filtered regimes obtained from the estimation of its first-order approximation.

Figure 6: Standard error of the 1-step ahead forecast in the $M3L$ model