A DSGE model of the term structure with regime shifts*

Gianni Amisano\textsuperscript{1} Oreste Tristani\textsuperscript{2}
European Central Bank European Central Bank

5 September 2008

Abstract

We test 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 on US data relying on information from both macroeconomic variables and the term structure. Our results support the specification with regime-switching: heteroskedasticity is a clear feature of the model’s residuals and the regimes have intuitively appealing features. The model is also capable of generating sizable time-variability in term premia. However, both the first and second order approximations of the model solutions can only match yield data for extreme values of the parameters.

JEL classification:

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

PRELIMINARY

\textsuperscript{*}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. We also thank Giorgio Primiceri, Ken Wallis and Paolo Zagaglia for useful comments and suggestions.

\textsuperscript{1}Email: gianni.amisano@ecb.int.

\textsuperscript{2}Email: oreste.tristani@ecb.int.
1 Introduction

The term structure of interest rates is a source of useful 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 microfounded 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 understanding 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 overwhelming evidence that the average term structure is positively sloped. Drawing on results from the finance literature, Atkeson and Kehoe (2008) argue that movements in risk premia should be allowed for in macro models to better understand the impact of monetary policy on the economy.

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 identification 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 and then estimate the nonlinear reduced form. 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 heteroskedasticity implies that the model is capable of generating time-variation in risk premia. We assume that heteroskedasticity 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. Bekaaert, 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 empirical results, based on US data from 1966 to 2006, show considerable support for a specification with regime switches, compared to a model with Gaussian shocks. This is the case for both linear and nonlinear approximations. The residuals of the models with Gaussian shocks show clear signs of heteroskedasticity and serial correlation. In the models with regime-switching, estimated regimes have intuitively appealing features: for example, monetary policy shocks are normally in the low-variance regime, except for the so-called monetary experiment period at the beginning of the 1980s. The regimes associated with the variance of technology shocks are also linked with the Great moderation in the linearised model, while the variance of preference shocks displays clear cyclical features in the model solved to second order.

Moreover, the quadratic model with regime switches is capable of generating considerable variations in risk premia. Premia are high especially in the early eighties and register a large drop in the second half of the 1990s.

At the same time, all versions of the model that we estimate – linearised and quadratic, homoskedastic and heteroskedastic – prove to be able to match yields under parameter values, especially for the monetary policy rule, which are vastly different from standard estimates obtained relying solely on macro-data. The monetary policy rule includes super-inertial characteristics and features exceptionally strong reactions to inflation deviations from target. Most of the variations in inflation are therefore attributed to variations in the inflation target. In the linear case, the real interest rate sensitivity of output must become negligible to avoid implausible repercussions of the high persistence of policy interest rates on real variables. This aspect of our results clearly deserves further analysis.

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

In order to highlight the marginal contribution of heteroskedasticity, we rely on a relatively standard model in the spirit of Yun (1996) and Woodford (2003). The central feature is the assumption of nominal rigidities. We only sketch the properties of the model briefly.

Consumers maximise the discounted sum of the period utility

$$U(C_t, C_{t-1}, L_t) = \varepsilon C_t \left( C_t - h C_{t-1} \right)^{1-\gamma} - \int_0^1 \frac{L_{t,t}^{1+\nu}}{1+\nu} \, dt$$

(1)

where $C$ is a consumption index satisfying

$$C = \left( \int_0^1 C(i) \frac{e^{i-1}}{\sigma} \, di \right)^{\frac{\sigma}{\sigma-1}},$$

(2)

workers provide $L_i$ hours of labor to firm $i$ and $\varepsilon C_t$ 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

$$P_t C_t + E_t(Q_{t,t+1} W_{t+1}) \leq \int_0^1 w_t(i) L_t(i) \, di + \int_0^1 \Xi_t(i) \, di + W_t$$

(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, $w_t(i)$ is the nominal wage rate and $\Xi_t(i)$ 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

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

(4)

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

$$\frac{w_t(i)}{P_t} = \frac{L_{t,t}^\nu}{\Lambda_{t}}$$

(5)

$$Q_{t,t+1} = \beta \frac{P_t}{P_{t+1}} \frac{\Lambda_{t+1}}{\Lambda_{t}}$$

(6)
where we define the marginal utility of consumption as

$$\overline{\Lambda}_t = e^{\gamma} (C_t - hC_{t-1})^{-\gamma} - \beta he^t \left[ e^{\gamma} (C_{t+1} - hC_t)^{-\gamma} \right]$$  \hspace{1cm} (7)

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

$$I_t = \beta^{-1} \left\{ E_t \left[ \frac{P_t}{P_{t+1}} \frac{\overline{\Lambda}_{t+1}}{\overline{\Lambda}_t} \right] \right\}^{-1} \hspace{1cm} (8)$$

The production function is given by

$$Y_t (i) = A_t L (i)^\alpha \hspace{1cm} (9)$$

where $A_t$ is a technology shock.

We assume Calvo (1983) contracts, so that firms face a constant probability $\zeta$ 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

$$\max P_i \sum_{s=t}^{\infty} \zeta^{s-t} Q_{t,s} (P_s^{i}y^{i}_{s} - TC_{s}) \hspace{1cm} (10)$$

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 $\Pi_t^*$. 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

$$P_t^i \left( \frac{1}{\Pi_t} \prod_{j=t}^{s} \Pi_j^* \right)^{1-\xi} \left( \frac{P_{t-1}}{P_{t-1}} \right)^{\xi}, \text{ where } 0 \leq \xi \leq 1.$$  

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 $P_t^i$. 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

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

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

6
Firms’ decisions can then be characterised as

\[
\left( \frac{P^*_t}{P_t} \right)^{1-\theta(1-\epsilon_{t+1})} = \frac{\chi^\theta}{\alpha (\theta - 1)} \frac{K_{2,t}}{K_{1,t}}
\]

(12)

\[
K_{2,t} = \frac{\lambda_t}{\lambda_t} Y_t^{1+1} + E_t \zeta Q_{t,t+1} \Pi_{t+1} \left( \frac{\Pi_{t+1}}{\Pi_{t+1}} \right)^{1-\epsilon_{t+1}} K_{2,t+1}
\]

(13)

\[
K_{1,t} = (1 - \tau_t) Y_t + E_t \zeta Q_{t,t+1} \Pi_{t+1} (\Pi_{t+1})^{1-\epsilon_{t+1}} K_{1,t+1}
\]

(14)

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

\[
I_t = \left( \frac{\Pi^*_t}{\beta} \right)^{1-\rho_t} \left( \frac{\Pi_t}{\Pi^*_t} \right)^{\psi_n} \left( \frac{Y_t}{Y_{t-1}} \right)^{\psi_y} \rho_t^{\epsilon_{t+1}}
\]

(15)

where \( Y_t \) is aggregate output, \( \Pi^*_t \) is a stochastic inflation target and \( \epsilon_{t+1} \) 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 (\( \psi_n \) 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 \( \Pi^*_t \). 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

\[
Y_t = C_t.
\]

(16)

Equilibrium dynamics are described by equations (6)-(8) and (11)-(16), plus the stochastic processes governing the motion of \( \epsilon_{C_t}, A_t, \Pi^*_t \) and \( \epsilon_t \). These are discussed below.

## 2.1 Solving the model

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

\[
\begin{align*}
A_{t+1} &= A_t^A e^{\varepsilon^A_{t+1}}, \quad \varepsilon^A_{t+1} \sim N(0, \sigma_{a,sY,t}) \\
\varepsilon_{I,I,t+1} &= \varepsilon^I_{t+1}, \quad \varepsilon^I_{t+1} \sim N(0, \sigma_{i,sI,t}) \\
\varepsilon^C_{C,t+1} &= (\varepsilon^C_{Ct})^{\rho_C} e^{\varepsilon^C_{t+1}} \quad \varepsilon^C_{t+1} \sim N(0, \sigma_{c,sC,t})
\end{align*}
\]

where

\[
\begin{align*}
\sigma_{a,sY,t} &= \sigma_{a,LsY,t} + \sigma_{a,H}(1-s_{Y,t}) \\
\sigma_{i,sI,t} &= \sigma_{i,LsI,t} + \sigma_{i,H}(1-s_{I,t}) \\
\sigma_{c,sC,t} &= \sigma_{c,LsC,t} + \sigma_{c,H}(1-s_{C,t})
\end{align*}
\]

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

\[
\Pi^*_{t+1} = \left(\Pi^*_{t}\right)^{1-\rho_o} (\Pi^*_{t})^{\rho_o} e^{\varepsilon^*_t_{t+1}} \tag{17}
\]
so that the inflation target is allowed to change smoothly over time.

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

\[
y_t = g(z_t, \sigma) \tag{18}
\]

\[
z_{t+1} = h(z_t, \sigma) + \tilde{\zeta}(z_t) \sigma \bar{u}_{t+1} \tag{19}
\]

where \( g(\cdot), h(\cdot), \) and \( \tilde{\zeta}(\cdot) \) 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 \( \bar{u}_t \), containing all innovations. In order to write the law of motion of the discrete processes in the form implied in equation (19), we rely on Hamilton (1994). The law of motion of state \( s_{C,t+1} \), for example, is written as

\[
s_{C,t+1} = (1 - p_{C,0}) + (-1 + p_{C,1} + p_{C,0}) s_{C,t} + \nu_{C,t+1}, \text{ where } \nu_{C,t+1} \text{ is an innovation with mean zero and heteroskedastic variance.}
\]

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 \( z_t = \bar{z} \) and \( \sigma = 0 \). We define the non-stochastic steady-state as vectors \( \bar{y} \) and \( \bar{z} \) such that \( f(\bar{y}, \bar{y}, \bar{z}, \bar{z}) \).

For the continuous state variables, the non-stochastic steady state \( \bar{z} \) 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 \( \sigma \to 0 \) 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

\[
\hat{g}(z_t, \sigma) = F \hat{z}_t + \frac{1}{2} (I_{ny} \otimes \hat{z}_t^T) E \hat{z}_t + k_{y,s} \sigma^2
\]

and

\[
\hat{h}(z_t, \sigma) = P \hat{z}_t + \frac{1}{2} (I_{nz} \otimes \hat{z}_t^T) G \hat{z}_t + k_{z,s} \sigma^2
\]

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.

\[
\begin{align*}
z_{t+1} &= P z_t + \frac{1}{2} \left( I_n \otimes \tilde{\zeta}_t \right) G \tilde{z}_t + k_{zz} \sigma^2 + \tilde{\zeta}_t \sigma \tilde{u}_{t+1} \\
\hat{\lambda}_t &= F_{\lambda} \tilde{z}_t + \frac{1}{2} \tilde{\zeta}_t E_{\lambda} \tilde{z}_t + k_{\lambda, s} \sigma^2 \\
\hat{\pi}_t &= F_{\pi} \tilde{z}_t + \frac{1}{2} \tilde{\zeta}_t E_{\pi} \tilde{z}_t + k_{\pi, s} \sigma^2
\end{align*}
\]

where \( F_{\lambda} \) and \( F_{\pi} \) are the appropriate rows of vector \( F \) and \( E_{\lambda} \) and \( E_{\pi} \) are appropriate sub-matrices of matrix \( E \). In log-deviation from its deterministic steady state, the approximate price of a bond of maturity \( n \), \( \hat{b}_{t,n} \), can then be written as

\[
\hat{b}_{t,n} (z_t, \sigma) = B_{z,n} \tilde{z}_t + \frac{1}{2} \tilde{z}_t B_{zz,n} \tilde{z}_t + B_{n, s} \sigma^2
\]

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

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

This expression can be evaluated using the model solution to obtain

\[
\hat{hpr}_{t,n} - \hat{i}_t = \sigma^2 B_{n-1,z} \zeta \zeta' \left( F_{\pi}' - F_{\lambda}' \right)
\]

(20)

where \( \zeta \zeta' \) 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 \( \text{E}_t \left[ b_{t+1,n-1} \hat{b}_{t+1,n-1} \right] = \sigma^2 B_{z,n-1} \zeta \zeta' B_{z,n-1} \), 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

$$\xi_t = \sigma \zeta' (F^x_t - F^x_0)$$  \hspace{1cm} (21)$$

The market prices of risk are only affected by first-order terms in the reduced-form of the model. All terms in equation (21) 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 $\zeta\zeta'$ 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 $x_t$ only includes the states with continuous support and vector $s_t$ includes the states with discrete support, we can partition the matrix $\tilde{\zeta}$ (recall that shocks with continuous and discrete support are all independently distributed) and the vectors $F_x$ and $F_\lambda$ conformably as

$$\tilde{\zeta} = \begin{bmatrix} \zeta^x & 0 \\ 0 & \zeta^s \end{bmatrix}, \quad F_x = \begin{bmatrix} F^x_x \\ F^x_\pi \end{bmatrix}, \quad F_\lambda = \begin{bmatrix} F^\lambda_x \\ F^\lambda_\pi \end{bmatrix}$$

As a result, equation (21) can be split into the vectors $\xi_t^x$ and $\xi_t^s$ such that $\xi_t' = [(\xi_t^x)', (\xi_t^s)']'$ and

$$\xi_t^x = \sigma \left( \begin{bmatrix} \zeta^x \\ 0 \end{bmatrix} \right)' \left[ (F^x_x)' - (F^x_\pi)' \right]$$

$$\xi_t^s = \sigma \left( \begin{bmatrix} 0 \\ \zeta^s \end{bmatrix} \right)' \left[ (F^\lambda_x)' - (F^\lambda_\pi)' \right]$$

Vector $\xi_t^x$ in equation (22) 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 $\xi_t^s$ in equation (23) 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,
$\xi_t$, 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 would be non-zero if regime-
switching affected other structural elements of the model. One obvious possibility would
be to replace the inflation target process in equation (17) with a specification allowing for
regime switching in the target mean. 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 (18) and (19), 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 intercept and linear terms changing according to
the discrete state variables This alternative representation is particularly convenient for
describing the estimation methodology. It is straightforward to show that the model can
be rewritten as

$$
y_{t+1} = c_j + C_{1,j}x_{t+1} + C_{2\text{vech}}(x_{t+1}x_{t+1}') + Dv_{t+1} \tag{24}
$$

$$
x_{t+1} = a_t + A_{1,t}x_t + A_{2\text{vech}}(x_tx_t') + B_tw_{t+1} \tag{25}
$$

$$
s_t \sim \text{Markov switching} \tag{26}
$$
where the vector $y_t^o$ includes all observable variables, vector $x_t$ only includes the states with continuous support, vector $s_t$ includes the states with discrete support, and $v_{t+1}$ and $w_{t+1}$ are measurement and structural shocks, respectively. In this representation, the regime switching variables affect the system by changing the intercepts $a_i$ and $c_j$, the slope coefficients $A_{1,i}$ and $C_{1,j}$, and the loadings for the structural innovations $B_i$.(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

\begin{align*}
y_{t+1}^o &= c_j + C_1x_{t+1} + Dv_{t+1} \\
x_{t+1} &= a_i + A_1x_t + B_iw_{t+1} \\
s_t &\sim \text{Markov switching}
\end{align*}

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. One possible approach to compute the likelihood is to rely on sequential Monte Carlo techniques (see
e.g. Amisano and Tristani, 2007a, and the reference therein). 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 we employ in the linear case. The only problem in this respect is the quadratic term in $x_t$ in the observation equation (24). For this reason, at each point in time $t$ we compute a linear approximation of this term around the estimate of $\hat{x}_{t-1}$. Relying on the assumption that the number of continuous shocks (measurement and structural) is equal to the number of observables, we then invert equation (24) to find a candidate $\hat{x}_t^c$ from the observation of $y_o$. We finally expand the quadratic term again around $\hat{x}_t^c$ and repeat this procedure until convergence.

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.\footnote{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.}

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 1-4). 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.

5 Empirical results

We have estimated our model under the simplifying assumption of absence of regime-switching and introducing incrementally regime switching in \( s_I, s_C \) and \( s_Y \). We refer to the model with a single regime as \( M0 \) and to the regime-switching models as \( M1, M2, M3 \), 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 an \( L \); estimates of the second-order (or quadratic) approximation with a \( Q \).

Since the linear model with three regime-switching processes dominates those with 1 or 2 in terms of marginal likelihood, we focus here on the comparison between \( M0L, M0Q, M3L \) and \( M3Q \).

5.1 Posterior distributions and goodness of fit

Tables 1-4 also report statistics on the posterior distributions of parameter estimates. The results highlight that all versions of the model must be stretched, albeit to different extents, 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, in the posterior mean of the standard deviations of almost all shocks. For example, compared to a prior mean around 1%, the estimated standard deviation of technology shocks increase to 3% in \( M0L \), to 10% in \( M0Q \), to between 3% and 4% in \( M3L \) and to between 2% and 5% in \( M3Q \). 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).\(^2\)

For all model, the posterior mean of the standard deviation of the target shock is particularly large. 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

\(^2\)Even in model \( M0Q \) a weak version of the expectations hypothesis holds because risk-premia are constant.
aggressive against inflation deviations from target, with short-term reaction coefficients above 1.0 and a degree of interest rate smoothing which is consistent with inertial or superinertial policy – in the sense of Woodford (2003). 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. The advantage of inertial rules is 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 $M3L$. The reason 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, $\gamma$ must be high to shield output from the volatility of the short-term. Posterior estimates of the coefficient of relative risk aversion are less extreme for the quadratic models and especially in the $M3Q$ model, where $\gamma$ is approximately equal to 6.

Overall, the posterior distribution have some puzzling implications, but less so for the quadratic model with regime-switching. The advantage of the regime switching specifications is to permit large standard deviations only in periods when explaining movements in long-yields is particularly difficult. In low-variance states, 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.

Turning to goodness of fit measures, Figures 1-4 display 1-step-ahead forecasts and realised variables for each of the three models.

The most striking feature emerging from these figures is probably 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.

A second feature which emerges from Figures 1-4 is the clear heteroskedasticity of the residuals. This is problematic for the models with Gaussian shocks, 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

Figures 5 and 6 display smoothed and filtered estimates of the discrete states in model $M_{3L}$ and $M_{3Q}$, together with the official NBER recession dates. In all cases, 1 denotes the low-variance state, 0 the high-variance state.

In both models, 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 most of the remaining the sample (with marginal exceptions). There are only small revisions noticeable in the smoothed estimates, compared to the filtered ones.

In the $M_{3Q}$ model, the regimes associated with preference shocks display some association with the economic cycle, with temporary drops to a high-variance state at the end of recessions (again, with an exception in the early 1980s). This is not the case in the $M_{3L}$ model, in which the probability of being in a low-variance state is almost never

17
above 0.5. This suggests that the cyclicity of the variance of preference shocks tend to reflect changes in risk-premia, which are ruled out by construction in the linear model.

Finally, the regimes associated with technology shocks clearly identifies the Great moderation period started in the mid-1980s in the linear model. 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 confirm 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 association with the Great moderation is much less clear in the quadratic model.

The various states can be composed to define 8 possible combinations of regimes. This is done to construct Figure 7, which displays excess holding period returns derived from the model. As discussed in relation to equation (20), these measures of risk premia vary over time only as a result of regime changes.

Two notable features emerge from Figure 7. 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 hover around a level of 5 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 7 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 around 8 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 after 1988 and until 2002.

Our estimated premia increase strongly again after the recession at the beginning of the new millennium, but then they drop sharply during the phase of moneatry tightening
which began in the middle of 2004 up to the lowest levels in the sample in 2006. This is the period which was characterised as a conundrum by Federal Reserve Chairman Greenspan in congressional testimony on 16 February 2005, because long-term rates did not rise as policymakers raised short-term rates. Our model explains the conundrum in terms of a drop in risk-premia on long-term bonds.

5.3 Impulse responses

Figure 8 shows impulse responses of our variables to shocks with continuous support in the $M3Q$ model.

Nonlinear impulse responses are defined as the difference between the expected future sample path of a variable conditional on a given initial state $x_t$, and the expected future path conditional on $x'_t$, where $x_t$ is equal to $x'_t$ except for an individual element which is perturbed by a known amount. The dependence of nonlinear impulse response functions on initial conditions is well-known (see e.g. Gallant, Rossi and Tauchen, 1993). Figure 8 shows simulations starting from the steady state of the model. Posterior median responses and the bounds corresponding to a 80% posterior coverage are reported in the figure. Both continuous and discrete states are simulated.

Compared to similar evidence based on estimates relying solely on information from macro-variables, the notable feature of Figure 8 is that the responses of the policy interest rate and, to a lesser extent, output and inflation, are much more persistent. In many cases, there is still no sign of a return to the baseline 3 years after the shock. Only for monetary policy shocks do endogenous variables go back to baseline quickly.

The high persistence of short term rates is responsible for movements in the yield curve. With the exception of policy shocks, the impulse response of 10-year yields is typically larger than the response of the short rate.

However, the impulse responses confirm that most of the movements in inflation are due to changes in the target. Technology and preference shocks affect output, but they have negligible effects on inflation.
6 Conclusions

Our results of 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, also models with regime-switching features must be stretched in order to match yields data. Parameter estimates are extreme compared to results based solely on macroeconomic information.
References


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


<table>
<thead>
<tr>
<th>Parameter</th>
<th>Post Mean</th>
<th>Post SD</th>
<th>Post Low Q</th>
<th>Post Up Q</th>
<th>Prior Mean</th>
<th>Prior SD</th>
<th>Prior Low Q</th>
<th>Prior Up Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Pi$</td>
<td>1.00938</td>
<td>0.00155</td>
<td>1.00631</td>
<td>1.01239</td>
<td>1.00499</td>
<td>0.00314</td>
<td>1.00085</td>
<td>1.01282</td>
</tr>
<tr>
<td>$\rho_\pi$</td>
<td>0.85337</td>
<td>0.02117</td>
<td>0.80886</td>
<td>0.89293</td>
<td>0.89983</td>
<td>0.02981</td>
<td>0.83414</td>
<td>0.95036</td>
</tr>
<tr>
<td>$\sigma_\pi$</td>
<td>0.00335</td>
<td>0.00043</td>
<td>0.00287</td>
<td>0.00396</td>
<td>0.00010</td>
<td>0.00014</td>
<td>0.00000</td>
<td>0.00050</td>
</tr>
<tr>
<td>$\rho_c$</td>
<td>0.93226</td>
<td>0.01834</td>
<td>0.88692</td>
<td>0.95799</td>
<td>0.90021</td>
<td>0.02982</td>
<td>0.83429</td>
<td>0.95042</td>
</tr>
<tr>
<td>$\sigma_c$</td>
<td>0.17910</td>
<td>0.02819</td>
<td>0.13961</td>
<td>0.22873</td>
<td>0.00989</td>
<td>0.01406</td>
<td>0.00001</td>
<td>0.04956</td>
</tr>
<tr>
<td>$\rho_A$</td>
<td>0.98373</td>
<td>0.00621</td>
<td>0.97913</td>
<td>0.98849</td>
<td>0.95004</td>
<td>0.02164</td>
<td>0.89329</td>
<td>0.98323</td>
</tr>
<tr>
<td>$\sigma_A$</td>
<td>0.02821</td>
<td>0.00588</td>
<td>0.01929</td>
<td>0.04078</td>
<td>0.01007</td>
<td>0.01419</td>
<td>0.00001</td>
<td>0.05091</td>
</tr>
<tr>
<td>$\sigma_I$</td>
<td>0.00284</td>
<td>0.00034</td>
<td>0.00250</td>
<td>0.00326</td>
<td>0.00050</td>
<td>0.00071</td>
<td>0.00000</td>
<td>0.00250</td>
</tr>
<tr>
<td>$\psi_\pi$</td>
<td>0.81235</td>
<td>0.17546</td>
<td>0.53993</td>
<td>1.20169</td>
<td>0.20124</td>
<td>0.16425</td>
<td>0.01459</td>
<td>0.62860</td>
</tr>
<tr>
<td>$\psi_y$</td>
<td>0.06872</td>
<td>0.02834</td>
<td>0.01931</td>
<td>0.12682</td>
<td>0.04987</td>
<td>0.04086</td>
<td>0.00350</td>
<td>0.15631</td>
</tr>
<tr>
<td>$\rho_I$</td>
<td>0.86999</td>
<td>0.05507</td>
<td>0.79881</td>
<td>0.94420</td>
<td>0.90204</td>
<td>0.20045</td>
<td>0.50809</td>
<td>1.29247</td>
</tr>
<tr>
<td>$\iota$</td>
<td>0.04055</td>
<td>0.08648</td>
<td>0.00787</td>
<td>0.07837</td>
<td>0.59974</td>
<td>0.20009</td>
<td>0.19152</td>
<td>0.93249</td>
</tr>
<tr>
<td>$\nu$</td>
<td>3.41566</td>
<td>1.02770</td>
<td>1.78588</td>
<td>5.75488</td>
<td>0.9836</td>
<td>0.81082</td>
<td>0.07361</td>
<td>3.08829</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>14.24830</td>
<td>2.03230</td>
<td>10.65812</td>
<td>17.87543</td>
<td>2.00146</td>
<td>0.70443</td>
<td>1.12276</td>
<td>3.78020</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>0.70081</td>
<td>0.05782</td>
<td>0.57192</td>
<td>0.79694</td>
<td>0.60099</td>
<td>0.14774</td>
<td>0.29864</td>
<td>0.86489</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.35479</td>
<td>0.06377</td>
<td>0.24050</td>
<td>0.47937</td>
<td>0.70006</td>
<td>0.13799</td>
<td>0.39945</td>
<td>0.92471</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.89899</td>
<td>0.00153</td>
<td>0.98695</td>
<td>0.99296</td>
<td>0.98020</td>
<td>0.01378</td>
<td>0.94547</td>
<td>0.99755</td>
</tr>
<tr>
<td>$\sigma_{me40}$</td>
<td>0.00110</td>
<td>0.00012</td>
<td>0.00097</td>
<td>0.00125</td>
<td>0.00100</td>
<td>0.00142</td>
<td>0.00000</td>
<td>0.00504</td>
</tr>
</tbody>
</table>

These results are based on 500,000 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 $\beta$, $h$, $\iota$, $\zeta$, $\rho_L$, $\rho_g$, $\rho_\pi$; gamma distribution for $\psi_\pi$, $\psi_y$ and all standard deviations; shifted gamma distribution (domain from 1 to $\infty$) for $\gamma$, $\phi$, $\Xi$, $\Pi$; normal distribution for $\rho_I$. 

Table 1: Parameter estimates: $M0L$ model
Table 2: Parameter estimates: M0Q model

<table>
<thead>
<tr>
<th></th>
<th>post mean</th>
<th>post sd</th>
<th>post low q</th>
<th>post up q</th>
<th>prior mean</th>
<th>prior sd</th>
<th>prior low q</th>
<th>prior up q</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Pi^*$</td>
<td>1.009101</td>
<td>0.001246</td>
<td>1.00638</td>
<td>1.011464</td>
<td>1.004993</td>
<td>0.003157</td>
<td>1.000837</td>
<td>1.012855</td>
</tr>
<tr>
<td>$\rho_\pi$</td>
<td>0.812753</td>
<td>0.024431</td>
<td>0.762106</td>
<td>0.857434</td>
<td>0.900101</td>
<td>0.029681</td>
<td>0.834622</td>
<td>0.950422</td>
</tr>
<tr>
<td>$\sigma_\pi$</td>
<td>0.003145</td>
<td>0.000295</td>
<td>0.002681</td>
<td>0.003801</td>
<td>9.9E-05</td>
<td>0.000141</td>
<td>1.07E-07</td>
<td>0.000501</td>
</tr>
<tr>
<td>$\rho_c$</td>
<td>0.907751</td>
<td>0.019283</td>
<td>0.86941</td>
<td>0.938963</td>
<td>0.899835</td>
<td>0.029836</td>
<td>0.83455</td>
<td>0.950198</td>
</tr>
<tr>
<td>$\sigma_c$</td>
<td>0.241292</td>
<td>0.030861</td>
<td>0.187301</td>
<td>0.307759</td>
<td>0.010019</td>
<td>0.014217</td>
<td>9.13E-06</td>
<td>0.050201</td>
</tr>
<tr>
<td>$\rho_A$</td>
<td>0.979256</td>
<td>0.002939</td>
<td>0.972918</td>
<td>0.984342</td>
<td>0.94871</td>
<td>0.021628</td>
<td>0.90069</td>
<td>0.983264</td>
</tr>
<tr>
<td>$\sigma_A$</td>
<td>0.096721</td>
<td>0.009691</td>
<td>0.078797</td>
<td>0.115796</td>
<td>0.009926</td>
<td>0.013946</td>
<td>1.05E-05</td>
<td>0.048995</td>
</tr>
<tr>
<td>$\sigma_I$</td>
<td>0.003096</td>
<td>0.00025</td>
<td>0.002681</td>
<td>0.003649</td>
<td>0.000503</td>
<td>0.000712</td>
<td>4.99E-07</td>
<td>0.002536</td>
</tr>
<tr>
<td>$\psi_\pi$</td>
<td>0.867747</td>
<td>0.260221</td>
<td>0.526316</td>
<td>1.514195</td>
<td>0.20058</td>
<td>0.164109</td>
<td>0.014301</td>
<td>0.626527</td>
</tr>
<tr>
<td>$\psi_y$</td>
<td>0.089573</td>
<td>0.02534</td>
<td>0.040196</td>
<td>0.138254</td>
<td>0.049661</td>
<td>0.040797</td>
<td>0.003644</td>
<td>0.154901</td>
</tr>
<tr>
<td>$\rho_I$</td>
<td>0.91097</td>
<td>0.057745</td>
<td>0.815312</td>
<td>1.028609</td>
<td>0.89739</td>
<td>0.199708</td>
<td>0.505812</td>
<td>1.289396</td>
</tr>
<tr>
<td>$\iota$</td>
<td>0.032308</td>
<td>0.017586</td>
<td>0.006726</td>
<td>0.073756</td>
<td>0.599276</td>
<td>0.200708</td>
<td>0.195229</td>
<td>0.932309</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.00131</td>
<td>0.002452</td>
<td>3.11E-05</td>
<td>0.007132</td>
<td>0.992976</td>
<td>0.812001</td>
<td>0.071032</td>
<td>3.111754</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>9.731058</td>
<td>1.648056</td>
<td>6.851732</td>
<td>13.32249</td>
<td>1.999478</td>
<td>0.704177</td>
<td>1.12047</td>
<td>3.776801</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>0.836738</td>
<td>0.050829</td>
<td>0.72467</td>
<td>0.909527</td>
<td>0.599237</td>
<td>0.147891</td>
<td>0.297063</td>
<td>0.861572</td>
</tr>
<tr>
<td>$h$</td>
<td>0.618677</td>
<td>0.053076</td>
<td>0.549132</td>
<td>0.749448</td>
<td>0.700138</td>
<td>0.138284</td>
<td>0.399005</td>
<td>0.924642</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.990011</td>
<td>0.001326</td>
<td>0.988342</td>
<td>0.993546</td>
<td>0.980231</td>
<td>0.013764</td>
<td>0.945504</td>
<td>0.997551</td>
</tr>
<tr>
<td>$\sigma_{me40}$</td>
<td>0.001281</td>
<td>7.73E-05</td>
<td>0.001133</td>
<td>0.00143</td>
<td>0.001001</td>
<td>0.001421</td>
<td>1.01E-06</td>
<td>0.005081</td>
</tr>
</tbody>
</table>

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 $\beta, h, \iota, \zeta, \rho_c, \rho_\pi$; gamma distribution for $\psi_\pi, \psi_y$ and all standard deviations; shifted gamma distribution (domain from 1 to $\infty$) for $\gamma, \phi, \Xi, \Pi^*$; normal distribution for $\rho_I$. 

### Table 3: Parameter estimates: $M3L$ model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Post Mean</th>
<th>Post SD</th>
<th>Post Low Q</th>
<th>Post Up Q</th>
<th>Prior Mean</th>
<th>Prior SD</th>
<th>Prior Low Q</th>
<th>Prior Up Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{x,1}$</td>
<td>0.9858</td>
<td>0.0090</td>
<td>0.9642</td>
<td>0.9978</td>
<td>0.8997</td>
<td>0.0909</td>
<td>0.8997</td>
<td>0.9978</td>
</tr>
<tr>
<td>$p_{x,00}$</td>
<td>0.8886</td>
<td>0.0744</td>
<td>0.7007</td>
<td>0.969</td>
<td>0.9004</td>
<td>0.0907</td>
<td>0.9004</td>
<td>0.9972</td>
</tr>
<tr>
<td>$p_{c,1}$</td>
<td>0.8796</td>
<td>0.0797</td>
<td>0.6855</td>
<td>0.9867</td>
<td>0.7455</td>
<td>0.1303</td>
<td>0.7455</td>
<td>0.9507</td>
</tr>
<tr>
<td>$p_{c,00}$</td>
<td>0.8134</td>
<td>0.0956</td>
<td>0.5996</td>
<td>0.9733</td>
<td>0.7820</td>
<td>0.1266</td>
<td>0.7820</td>
<td>0.9671</td>
</tr>
<tr>
<td>$p_{A,1}$</td>
<td>0.9429</td>
<td>0.0571</td>
<td>0.7702</td>
<td>0.9920</td>
<td>0.9000</td>
<td>0.0907</td>
<td>0.9000</td>
<td>0.9972</td>
</tr>
<tr>
<td>$p_{A,00}$</td>
<td>0.9239</td>
<td>0.0624</td>
<td>0.7706</td>
<td>0.9912</td>
<td>0.9003</td>
<td>0.0902</td>
<td>0.9003</td>
<td>0.9971</td>
</tr>
<tr>
<td>$\sigma_{1,1}$</td>
<td>0.0026</td>
<td>0.0002</td>
<td>0.0023</td>
<td>0.0030</td>
<td>0.0096</td>
<td>0.0030</td>
<td>0.0096</td>
<td>0.0168</td>
</tr>
<tr>
<td>$\sigma_{1,0}$</td>
<td>0.0097</td>
<td>0.0020</td>
<td>0.0068</td>
<td>0.0142</td>
<td>0.0155</td>
<td>0.0076</td>
<td>0.0155</td>
<td>0.0344</td>
</tr>
<tr>
<td>$\sigma_{c,1}$</td>
<td>0.1572</td>
<td>0.0331</td>
<td>0.0858</td>
<td>0.2201</td>
<td>0.0096</td>
<td>0.0029</td>
<td>0.0096</td>
<td>0.0168</td>
</tr>
<tr>
<td>$\sigma_{c,0}$</td>
<td>0.2930</td>
<td>0.0704</td>
<td>0.1925</td>
<td>0.4684</td>
<td>0.0155</td>
<td>0.0080</td>
<td>0.0155</td>
<td>0.0346</td>
</tr>
<tr>
<td>$\sigma_{A,1}$</td>
<td>0.0276</td>
<td>0.0058</td>
<td>0.0183</td>
<td>0.0407</td>
<td>0.0287</td>
<td>0.0088</td>
<td>0.0287</td>
<td>0.0505</td>
</tr>
<tr>
<td>$\sigma_{A,0}$</td>
<td>0.0438</td>
<td>0.0114</td>
<td>0.0292</td>
<td>0.0736</td>
<td>0.0466</td>
<td>0.0231</td>
<td>0.0466</td>
<td>0.1045</td>
</tr>
<tr>
<td>$\Pi^*$</td>
<td>1.0080</td>
<td>0.0018</td>
<td>1.0043</td>
<td>1.0112</td>
<td>1.0050</td>
<td>0.0032</td>
<td>1.0050</td>
<td>1.0129</td>
</tr>
<tr>
<td>$\rho_\pi$</td>
<td>0.8833</td>
<td>0.0164</td>
<td>0.8483</td>
<td>0.9122</td>
<td>0.9800</td>
<td>0.0139</td>
<td>0.9800</td>
<td>0.9975</td>
</tr>
<tr>
<td>$\sigma_\pi$</td>
<td>0.0030</td>
<td>0.0003</td>
<td>0.0025</td>
<td>0.0038</td>
<td>0.0003</td>
<td>0.0002</td>
<td>0.0003</td>
<td>0.0007</td>
</tr>
<tr>
<td>$\rho_x$</td>
<td>0.9266</td>
<td>0.0245</td>
<td>0.8641</td>
<td>0.9579</td>
<td>0.9000</td>
<td>0.0299</td>
<td>0.9000</td>
<td>0.9505</td>
</tr>
<tr>
<td>$\rho_A$</td>
<td>0.9841</td>
<td>0.0029</td>
<td>0.9782</td>
<td>0.9891</td>
<td>0.9801</td>
<td>0.0138</td>
<td>0.9801</td>
<td>0.9975</td>
</tr>
<tr>
<td>$\psi_\pi$</td>
<td>1.1486</td>
<td>0.3679</td>
<td>0.5964</td>
<td>1.9699</td>
<td>0.2007</td>
<td>0.1656</td>
<td>0.2007</td>
<td>0.6333</td>
</tr>
<tr>
<td>$\psi_y$</td>
<td>0.0494</td>
<td>0.0312</td>
<td>0.0045</td>
<td>0.1191</td>
<td>0.0500</td>
<td>0.0411</td>
<td>0.0500</td>
<td>0.1565</td>
</tr>
<tr>
<td>$\rho_i$</td>
<td>1.0020</td>
<td>0.0811</td>
<td>0.8765</td>
<td>1.1855</td>
<td>0.8978</td>
<td>0.2000</td>
<td>0.8978</td>
<td>1.2901</td>
</tr>
<tr>
<td>$\iota$</td>
<td>0.0322</td>
<td>0.0201</td>
<td>0.0058</td>
<td>0.0822</td>
<td>0.5992</td>
<td>0.2099</td>
<td>0.5992</td>
<td>0.9333</td>
</tr>
<tr>
<td>$\nu$</td>
<td>3.2054</td>
<td>1.1303</td>
<td>1.5808</td>
<td>6.0579</td>
<td>1.0004</td>
<td>0.8199</td>
<td>1.0004</td>
<td>3.1490</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>15.1056</td>
<td>2.0833</td>
<td>11.1973</td>
<td>19.4460</td>
<td>2.0023</td>
<td>0.7118</td>
<td>2.0023</td>
<td>3.8011</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>0.6662</td>
<td>0.0589</td>
<td>0.5470</td>
<td>0.7748</td>
<td>0.5998</td>
<td>0.1474</td>
<td>0.5998</td>
<td>0.8617</td>
</tr>
<tr>
<td>$h$</td>
<td>0.4363</td>
<td>0.0633</td>
<td>0.3086</td>
<td>0.5597</td>
<td>0.6994</td>
<td>0.1387</td>
<td>0.6994</td>
<td>0.9256</td>
</tr>
<tr>
<td>$\theta$</td>
<td>9.8234</td>
<td>2.7239</td>
<td>5.6461</td>
<td>15.9354</td>
<td>7.9942</td>
<td>2.6310</td>
<td>7.9942</td>
<td>13.9742</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.9889</td>
<td>0.0018</td>
<td>0.9850</td>
<td>0.9921</td>
<td>0.9801</td>
<td>0.0138</td>
<td>0.9801</td>
<td>0.9975</td>
</tr>
<tr>
<td>$\sigma_{mc,40}$</td>
<td>0.0011</td>
<td>0.0001</td>
<td>0.0010</td>
<td>0.0013</td>
<td>0.0014</td>
<td>0.0010</td>
<td>0.0014</td>
<td>0.0037</td>
</tr>
</tbody>
</table>

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 $\beta$, $h$, $\iota$, $\zeta$, $\rho_x$, $\rho_A$; gamma distribution for $\psi_\pi$, $\psi_y$, and all standard deviations; shifted gamma distribution (domain from 1 to $1$) for $\gamma$, $\phi$, $\Xi$, $\Pi^*$; normal distribution for $\rho_i$. 

38
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Post Mean</th>
<th>Post SD</th>
<th>Post Low Q</th>
<th>Post Up Q</th>
<th>Prior Mean</th>
<th>Prior SD</th>
<th>Prior Low Q</th>
<th>Prior Up Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{x,11}$</td>
<td>0.9858</td>
<td>0.0090</td>
<td>0.9642</td>
<td>0.9978</td>
<td>0.8997</td>
<td>0.0909</td>
<td>0.6614</td>
<td>0.9973</td>
</tr>
<tr>
<td>$p_{x,00}$</td>
<td>0.8886</td>
<td>0.0744</td>
<td>0.7007</td>
<td>0.9869</td>
<td>0.9004</td>
<td>0.0907</td>
<td>0.6635</td>
<td>0.9972</td>
</tr>
<tr>
<td>$p_{c,11}$</td>
<td>0.8796</td>
<td>0.0797</td>
<td>0.6855</td>
<td>0.9867</td>
<td>0.7455</td>
<td>0.1303</td>
<td>0.4569</td>
<td>0.9507</td>
</tr>
<tr>
<td>$p_{c,00}$</td>
<td>0.8134</td>
<td>0.0956</td>
<td>0.5996</td>
<td>0.9733</td>
<td>0.7820</td>
<td>0.1266</td>
<td>0.4895</td>
<td>0.9671</td>
</tr>
<tr>
<td>$p_{A,11}$</td>
<td>0.9429</td>
<td>0.0571</td>
<td>0.7702</td>
<td>0.9920</td>
<td>0.9000</td>
<td>0.0907</td>
<td>0.6615</td>
<td>0.9972</td>
</tr>
<tr>
<td>$p_{A,00}$</td>
<td>0.9239</td>
<td>0.0624</td>
<td>0.7706</td>
<td>0.9912</td>
<td>0.9003</td>
<td>0.0902</td>
<td>0.6653</td>
<td>0.9971</td>
</tr>
<tr>
<td>$\sigma_{I,1}$</td>
<td>0.0026</td>
<td>0.0002</td>
<td>0.0023</td>
<td>0.0030</td>
<td>0.0096</td>
<td>0.0030</td>
<td>0.0056</td>
<td>0.0168</td>
</tr>
<tr>
<td>$\sigma_{I,0}$</td>
<td>0.0097</td>
<td>0.0020</td>
<td>0.0068</td>
<td>0.0142</td>
<td>0.0155</td>
<td>0.0076</td>
<td>0.0077</td>
<td>0.0344</td>
</tr>
<tr>
<td>$\sigma_{c,1}$</td>
<td>0.1572</td>
<td>0.0331</td>
<td>0.0858</td>
<td>0.2201</td>
<td>0.0996</td>
<td>0.029</td>
<td>0.0056</td>
<td>0.0168</td>
</tr>
<tr>
<td>$\sigma_{c,0}$</td>
<td>0.2930</td>
<td>0.0704</td>
<td>0.1925</td>
<td>0.4684</td>
<td>0.1555</td>
<td>0.080</td>
<td>0.0078</td>
<td>0.0346</td>
</tr>
<tr>
<td>$\sigma_{A,1}$</td>
<td>0.0276</td>
<td>0.0058</td>
<td>0.0183</td>
<td>0.0407</td>
<td>0.0287</td>
<td>0.0088</td>
<td>0.0169</td>
<td>0.0505</td>
</tr>
<tr>
<td>$\sigma_{A,0}$</td>
<td>0.0438</td>
<td>0.0114</td>
<td>0.0292</td>
<td>0.0736</td>
<td>0.0466</td>
<td>0.0231</td>
<td>0.0234</td>
<td>0.1045</td>
</tr>
<tr>
<td>$\Pi^*$</td>
<td>1.0080</td>
<td>0.0018</td>
<td>1.0043</td>
<td>1.0112</td>
<td>1.0050</td>
<td>0.0032</td>
<td>1.0008</td>
<td>1.0129</td>
</tr>
<tr>
<td>$\rho_\pi$</td>
<td>0.9704</td>
<td>0.004274</td>
<td>0.960444</td>
<td>0.977641</td>
<td>0.980054</td>
<td>0.0139</td>
<td>0.945304</td>
<td>0.97465</td>
</tr>
<tr>
<td>$\sigma_\pi$</td>
<td>0.00262</td>
<td>0.000195</td>
<td>0.002316</td>
<td>0.003085</td>
<td>0.000275</td>
<td>0.00021</td>
<td>0.000114</td>
<td>0.000734</td>
</tr>
<tr>
<td>$\rho_{x}$</td>
<td>0.958199</td>
<td>0.005874</td>
<td>0.943336</td>
<td>0.967614</td>
<td>0.89541</td>
<td>0.03134</td>
<td>0.833658</td>
<td>0.95061</td>
</tr>
<tr>
<td>$\rho_{A}$</td>
<td>0.97226</td>
<td>0.003435</td>
<td>0.964226</td>
<td>0.978034</td>
<td>0.98011</td>
<td>0.013858</td>
<td>0.945375</td>
<td>0.97559</td>
</tr>
<tr>
<td>$\psi_{x}$</td>
<td>1.473756</td>
<td>0.021248</td>
<td>1.04904</td>
<td>1.969238</td>
<td>0.198679</td>
<td>0.160113</td>
<td>0.014325</td>
<td>0.60845</td>
</tr>
<tr>
<td>$\psi_{y}$</td>
<td>0.065758</td>
<td>0.032148</td>
<td>0.026878</td>
<td>0.142995</td>
<td>0.04527</td>
<td>0.040337</td>
<td>0.003615</td>
<td>0.15486</td>
</tr>
<tr>
<td>$\rho_{I}$</td>
<td>1.288397</td>
<td>0.069957</td>
<td>1.176666</td>
<td>1.446318</td>
<td>0.900755</td>
<td>0.10013</td>
<td>0.703863</td>
<td>1.06383</td>
</tr>
<tr>
<td>$t$</td>
<td>0.031046</td>
<td>0.03014</td>
<td>0.006939</td>
<td>0.099855</td>
<td>0.599823</td>
<td>0.200269</td>
<td>0.194846</td>
<td>0.93325</td>
</tr>
<tr>
<td>$\nu$</td>
<td>5.639422</td>
<td>1.466711</td>
<td>2.76844</td>
<td>8.278278</td>
<td>0.99735</td>
<td>0.814908</td>
<td>0.07387</td>
<td>3.12659</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>6.045877</td>
<td>2.053413</td>
<td>3.372705</td>
<td>10.76833</td>
<td>1.966331</td>
<td>0.70623</td>
<td>1.121729</td>
<td>3.75415</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>0.610444</td>
<td>0.052529</td>
<td>0.51324</td>
<td>0.706574</td>
<td>0.599456</td>
<td>0.14837</td>
<td>0.296311</td>
<td>0.82635</td>
</tr>
<tr>
<td>$h$</td>
<td>0.606648</td>
<td>0.078736</td>
<td>0.44843</td>
<td>0.739479</td>
<td>0.699271</td>
<td>0.138003</td>
<td>0.400271</td>
<td>0.92354</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.985656</td>
<td>0.001556</td>
<td>0.982711</td>
<td>0.988789</td>
<td>0.980213</td>
<td>0.013888</td>
<td>0.945258</td>
<td>0.997595</td>
</tr>
<tr>
<td>$\sigma_{mc,40}$</td>
<td>0.000826</td>
<td>0.000188</td>
<td>0.000603</td>
<td>0.001211</td>
<td>0.001381</td>
<td>0.00097</td>
<td>0.000569</td>
<td>0.003759</td>
</tr>
</tbody>
</table>

These results are based on 100000 simulations. Legend: "sd" denotes the standard deviation; "low q" and "up q" denote the 5th and 95th percentiles of the distribution. Acceptance rate: 0.54

Priors: beta distribution for $\beta$, $h$, $\nu$, $\gamma$, $\zeta$, $\rho_{\epsilon}$; gamma distribution for $\psi_{x}$, $\psi_{y}$ and all standard deviations; shifted gamma distribution (domain from 1 to $\infty$) for $\gamma$, $\phi$, $\Xi$; normal distribution for $\rho_{I}$. 

These results are based on 100000 simulations. Legend: "sd" denotes the standard deviation; "low q" and "up q" denote the 5th and 95th percentiles of the distribution. Acceptance rate: 0.54

Priors: beta distribution for $\beta$, $h$, $\nu$, $\gamma$, $\zeta$, $\rho_{\epsilon}$; gamma distribution for $\psi_{x}$, $\psi_{y}$ and all standard deviations; shifted gamma distribution (domain from 1 to $\infty$) for $\gamma$, $\phi$, $\Xi$; normal distribution for $\rho_{I}$. 


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: Actual variables and 1-step-ahead predictions: M3Q model
Figure 5: 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 6: Filtered and smoothed estimates of the regime-variables: MAQ model

Legend: "state i" corresponds to the policy shock; "state C" corresponds to the preference shock; "state A" corresponds to the technology shock.
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 second-order approximation.
Figure 8: Impulse responses in the M3Q model.