Monetary Policy Regime Changes and the Term Structure: Evidence from a DSGE Model*

SIDDHARThA CHIB†
KYU HO KANG‡
SRIKANTH RAMAMURTHY§

February 2011

Abstract

We develop and estimate a model of the term structure of interest rates within the context of a Dynamic Stochastic General Equilibrium framework. The model features multiple monetary policy and volatility regimes. We estimate this model by Bayesian methods. Our estimation results reveal that U.S. monetary policy has switched between more active and less active regimes since the middle of 1980’s, and that the average term premium has been lower during the more active regime. Also, the price of regime shift risk has been significantly positive over time. Finally, regime changes in monetary policy and the volatility of the technology shock have accounted for most of the variation in the term premium. (JEL C11, E43)

Keywords: Bayesian inference, Marginal likelihood, Markov switching process, Metropolis-Hastings, MCMC, Regime shifts, State space model, Term-premium

1 Introduction

This paper provides a detailed analysis of the link between the central bank’s monetary policy and the term structure of interest rates. The primary question here is whether an active interest rate feedback rule (Taylor 1993) that is widely considered to have

---

*We thank Scott Joslin, Taeyoung Doh, Hong Liu, James Morley, and the participants of the 2010 Western Finance Association meeting for their thoughtful and useful comments on the paper.
†Address for correspondence: Olin Business School, Washington University in St. Louis, Campus Box 1133, 1 Brookings Drive, St. Louis, MO 63130. E-mail: chib@wustl.edu.
‡Address for correspondence: Department of Financial Engineering, Ajou University, San 5, Woncheon-dong, Yeongtong-gu, Suweon, Gyeonggi-do, 443-749, South Korea. E-mail: kyu@kyukang.net.
§Address for correspondence: Sellinger School of Business, Loyola University Maryland, 4501 N. Charles St., Baltimore, MD 21210. E-mail: sramamurthy@loyola.edu.
stabilizing properties\textsuperscript{1} for the real economy has the corresponding effect on the yield curve. Based on a simple estimated New Keynesian DSGE model, this paper finds that the answer is, in fact, yes. If one allows for the possibility of monetary policy switches between “more active” and “less active” regimes, both the average term premium and its volatility are found to be smaller during the more active regimes.

In the context of general equilibrium models, this, somewhat counterintuitive, result arises from the impact of two opposing effects on the term premium. As in the partial equilibrium approach (Ang, Boivin, Dong, and Loo-Kung (2010), Bikbov and Chernov (2008) and Abdymomunov and Kang (2010)), the positive impact of a more aggressive monetary policy on the term premium perpetuates through an increase in the factor loadings. At the same time, however, because of a decrease in the market price of risk the net effect is that the average term premium decreases, leading to relatively flatter yield curve.

Our approach to modeling the term structure is motivated by the contemporaneous interaction between the bond market and the real economy. In this regard, the New Keynesian DSGE model provides a convenient framework for analyzing the joint dynamics of the term structure and macroeconomic variables. The prototypical model is comprised of a representative household that makes decisions on the optimal stream of consumption and a portfolio of bonds of various maturities, a continuum of intermediate goods producers, a representative final goods producer, the government sector (which issues bonds of various maturities) and the central bank. The various agents in the model are intertemporal optimizers that face uncertainty arising from exogenous shocks to productivity, monetary policy and government expenditure. To address the main question in this paper, we specify the central bank’s monetary policy function in terms of the generalized Taylor (1993) rule (Davig and Leeper (2007)) that features multiple policy regimes. Following this rule, the central bank adjusts the nominal short rate in response to deviations of inflation and output from their target levels. As in Davig and Doh (2009), Bikbov and Chernov (2008) and Abdymomunov and Kang (2010), an

\textsuperscript{1}Ignoring, in particular, the technical debate surrounding this issue raised by Benhabib, Schmitt-Grohe, and Uribe (2001)
important aspect of this policy function is that the inflation and output coefficients are
time varying in order to capture the possibility of policy changes between active and
less active regimes.

A notable feature of this general equilibrium orientation is that the nominal pric-
ing kernel and the no arbitrage conditions are derived within the model through the
agents’ optimization problem (as opposed to being specified exogenously as in a partial
equilibrium framework). Consequently, the macroeconomic aggregates (namely, output
and inflation) are determined endogenously in the model. The evolution of these quan-
tities depends in part on monetary policy, on expectations of future policy changes,
and the structural macroeconomic shocks. Because the pricing kernel is a function of
these quantities, any change in the policy regime impacts the entire term structure.
This information from the full term structure is in turn utilized by agents when forming
expectations of future outcomes of the macroeconomic aggregates and policy regimes.
This inextricable bidirectional feedback mechanism is what entails one to capture the
joint dynamics of the term structure and macroeconomic variables.

We complete our discussion of the term structure with the risk premium. Because
the risk premium is potentially determined by both the size of the structural shocks and
the sensitivity of the short rate to these shocks, as well as changes in the volatilities of
these shocks, we model these volatilities, following Ang, Bekaert, and Wei (2008) and
Bikbov and Chernov (2008), by a discrete time Markov switching process.

The empirical implications of our model are isolated by Bayesian techniques, which in
recent years have become the central workhorse for estimating DSGE models (Fernandez-
Villaverde and Rubio-Ramirez (2009), An and Schorfheide (2007), Smets and Wouters
(2007)). Despite the complex nature of the likelihood/posterior that is the characteristic
of such models, our fitting method, which is based on the MCMC simulation methods
in Chib and Ergashev (2009) and Chib and Ramamurthy (2010), is relatively efficient in
terms of the metrics that are used to evaluate MCMC procedures. Further, our inference
is based on realistic priors that reflect the assumption of a positive term-premium (Chib
and Ergashev (2009)).

The work in this paper can be viewed as a continuation of a recent line of inquiry into
general equilibrium modeling of the term structure, as exemplified in Wachter (2006), Rudebusch and Swanson (2008b), and Wu (2006). Unlike these papers, however, we allow for regime changes (a feature that has been shown to be important in the partial equilibrium models of Rudebusch and Wu (2007) and Chib and Kang (2010)), and employ econometric methods to estimate the model, as opposed to calibrating it by simulation methods.

Our estimation results for U.S. quarterly data from 1986:Q4 to 2010:Q3, with bonds of maturities up to 20 quarters, reveal that U.S. monetary policy has switched between “less active” and “more active” regimes since the mid 1980’s. More interestingly, as mentioned earlier, the average term premium and its volatility are lower during the more active regime. Secondly, the price of regime shift risk, while small compared to factor risk, is always significantly positive over time. Lastly, the changes in the monetary policy and the volatility of technology shock account for most of the variation in the term premium.

The rest of the paper is organized as follows. In Section 2 we develop the model, discuss the solution procedure and derive the bond prices. Section 3 provides the econometric details and Section 4 contains the empirical results. Concluding remarks are in Section 5.

2 Model

In this section we discuss the key aspects of our DSGE model with multiple monetary policy and volatility regimes. We present the model, derive the implied pricing kernel and compute the arbitrage-free $\tau$ maturity bond prices through the $\tau$-forward iterations of the log-linearized Euler equation.

The model economy comprises a representative household, a continuum of intermediate goods producers indexed by $j \in [0, 1]$, a representative final good producer, the government sector and the central bank. The household maximizes its utility by supplying labor to the intermediate goods sector, consuming the finished good and making a portfolio decision over bonds of various maturities issued by the government. All
firms maximize profits. A standard way of introducing market frictions in these models is to assume that the firms in the intermediate good sector face short run nominal rigidities in the form of quadratic price adjustment costs. In its goal to stabilize the economy, the central bank, following the Taylor (1993) rule, adjusts the short interest rate in response to output and inflation. As mentioned earlier, this policy function is time varying, switching between active and passive (or possibly “less active”) regimes.

The aggregate macroeconomic fluctuations in this model are driven by three structural shocks, namely a technology shock, a fiscal shock and a monetary policy shock. To capture the heteroskedastic nature of these shocks, we assume that their volatilities follow a two-state discrete time Markov switching process. As we show later in this section, these structural shocks play the analogous role of factors in the partial equilibrium framework.

In this economy, therefore, the agents’ behavior is shaped by three sources of uncertainty - the policy regime $s_t$, the volatility regime $v_t$ and the shocks themselves. The fundamental assumption regarding the agents’ expectation of the future realizations of the aggregate variables (which are functions of the underlying uncertainties) is that they are based on rational expectations. That is, their expectations at time $t$, denoted $E_t$, is based on the complete information set at time $t$ that includes current and past realizations of all decision variables in the model, the regime sequences, $\{s_t, s_{t-1}, s_{t-2}, \ldots\}$ and $\{v_t, v_{t-1}, v_{t-2}, \ldots\}$, and the shocks. We denote this period-$t$ information set as $I_t$ and use $E_t[X_{t+j}]$ and $E[X_{t+j}|I_t]$ interchangeably throughout the text to denote the $j$-period ahead expectation of $X$ conditioned on $I_t$. The agents also know the structural parameters of the model. The only unknowns in their information set are the future realizations of the shocks and the regimes. Given a specific stochastic process for the evolution of these regimes, the agents form one step ahead expectations of the regimes and thus solve for the growth path of the macroeconomic aggregates as a function of the shocks.

2.1 The Representative Household

The representative household faces a consumption-leisure choice, deriving utility from consuming $C_t$ units of the finished good purchased from the final good producer at
the nominal price $P_t$ and supplying $H_t$ units of labor to the intermediate goods sector in return for a real wage rate of $W_t$. In addition to the wage income, the household earns real profits $Q_t$ from the intermediate goods firms. Finally, the household carries a portfolio $\{B^\tau_t\}_{\tau=1}^{\tau^*}$ of nominal $\tau$-quarter maturity zero-coupon bonds $B^\tau_t$ with current prices $P^\tau_t$ at any time $t$. We assume that the agent cares only about the time to maturity of the various bonds and not the date at which the bonds are issued. In other words, at time $t$, she is indifferent between holding a $(\tau + 1)$ period maturity bond bought at time $t - 1$ and a $(\tau)$ period maturity bond bought at time $t$, so that $B^\tau_{t-1} = B^\tau_t$.

The government issues the multiple maturity bonds at a face value of unity. Current income and financial wealth brought over from the previous period $t - 1$ are allocated between consumption, purchases of new bonds and a lumpsum real tax $T_t$ levied by the government. The budget constraint of the household therefore satisfies

$$P_tC_t + \sum_{\tau=1}^{\tau^*} P^\tau_t B^\tau_t + T_t \leq P_t W_t H_t + \sum_{\tau=1}^{\tau^*} P^\tau_t B^\tau_{t-1} + B^1_{t-1} + P_t Q_t.$$  \hspace{1cm} (2.1)

The household then maximizes her expected utility function \(^{2}\)

$$\mathbb{E}_t \left[ \sum_{s=0}^{\infty} \delta^s \left( \frac{(C_{t+s}/A_{t+s})^{1-\gamma} - 1}{1 - \gamma} - H_{t+s} \right) \right]$$  \hspace{1cm} (2.2)

subject to the intertemporal budget constraint (2.1) and available information up to time $t$. Here the variable $A_t$ captures the general productivity level or aggregate technology, so that $C_t/A_t$ measures the effective consumption per unit of technology.\(^{3}\)

We assume that the growth rate of technology $a_t = A_t/A_{t-1}$ follows an autoregressive process

$$\ln a_t = (1 - \phi_0) \ln a^* + \phi_0 \ln a_{t-1} + \varepsilon_{a,t}$$  \hspace{1cm} (2.3)

\(^{2}\)The simpler log utility function (where $\gamma$ is fixed at 1) is not meaningful in this context because it generates a bond risk premium that is too small and stable relative to the data (Rudebusch and Swanson, 2008b).

\(^{3}\)Alternatively, preferences could display habit persistence (modeled through a lagged consumption variable), as in Buraschi and Jiltsov (2007), Wachter (2006) and Rudebusch and Swanson (2008b), which can potentially improve the model’s ability to fit the term premium and the nonlinearity of the spot rate process. However, currently there is no feasible method for solving a model with both habit persistence and multiple regimes. Another possibility is to allow for regime shifts in the target inflation rather than the reaction coefficients, as in Liu, Waggoner, and Zha (2010).
where $|\phi_a| < 1$ and the innovation $\varepsilon_{a,t}$ is normally distributed with mean 0 and a regime-switching volatility process $\sigma^2_{a,v_{t,a}}$. Specifically, we assume that the volatility regime $v_{t,a}$ follows a two-state discrete time Markov process (Hamilton, 1989, Albert and Chib, 1993, Fruhwirth-Schnatter, 2006). The economic interpretation of these two regimes is that the economy transits between high volatility and low volatility states. Accordingly, we impose the identification restriction $\sigma_{a,2} > \sigma_{a,1}$, so that $v_{t,a} = 2$ denotes the higher volatility regime. The associated transition probability matrix for the volatility process is given by

$$Q^a = \begin{bmatrix} q_{11}^0 & 1 - q_{11}^0 \\ 1 - q_{22}^0 & q_{22}^0 \end{bmatrix} \quad (2.4)$$

where $q_{ij}^a = \Pr[v_{t+1,a} = j | v_{t,a} = i]$.

### 2.2 The Final Good Sector

A representative firm in the finished goods sector combines a continuum of intermediate goods $Y_t(j)$ indexed by $j \in [0, 1]$ using the constant returns to scale production technology

$$\left( \int_0^1 Y_t(j)^{\frac{\zeta-1}{\zeta}} dj \right)^{\frac{\zeta}{\zeta-1}} \geq Y_t \quad (2.5)$$

where $\zeta > 1$ measures the elasticity of demand for each intermediate good. In each period $t = 0, 1, 2, \ldots$, it chooses the output level given the price $P_t$ of the finished good and input prices $P_t(j)$. Profit maximization implies that the demand for intermediate goods is given by

$$P_t(j) = \left( \frac{Y_t}{Y_t(j)} \right)^{1/\zeta} P_t. \quad (2.6)$$

The aggregate price level is determined by the zero profit condition under competitive equilibrium as

$$P_t = \left( \int_0^1 P_t(j)^{1-\zeta} dj \right)^{\frac{1}{1-\zeta}}. \quad (2.7)$$

### 2.3 The Intermediate Good Sector

The intermediate good sector is characterized by a continuum of monopolistically competitive firms. Each firm indexed by $j$ produces a unique, imperfectly substitutable,
perishable good $Y_t(j)$ using a linear production technology with respect to the labor input $N_t(j)$ given the exogenous aggregate technology $A_t$ in the economy

$$Y_t(j) = A_t N_t(j). \quad (2.8)$$

As mentioned earlier, the firms in the intermediate goods sector face nominal rigidities in the form of an explicit price adjustment cost. As is conventional in the literature, this price adjustment cost takes the quadratic form

$$AC_t(j) = \frac{\varphi}{2} \left( \frac{P_t(j)}{\pi^* P_{t-1}(j)} - 1 \right)^2 Y_t \quad (2.9)$$

where $\varphi > 0$ measures the degree of price stickiness, $\pi_t = P_t/P_{t-1}$ is the gross inflation and $\pi^*$ is the inflation target of the central bank in terms of the price of the final good. When selling its output to the final goods sector, each intermediate-good firm $j$ chooses a sequence of input prices $P_t(j)$ to maximize the expected profits

$$\mathbb{E}_t \left[ \sum_{s=0}^{\infty} \Lambda_{t,t+s} Q_t(j) \right] \quad (2.10)$$

where the real profit at time $t$ is

$$Q_t(j) = \frac{P_t(j)}{P_t} Y_t(j) - W_t N_t(j) - \frac{\varphi}{2} \left( \frac{P_t(j)}{\pi^* P_{t-1}(j)} - 1 \right)^2 Y_t \quad (2.11)$$

and

$$\Lambda_{t,t+s} = \delta^s \left( \frac{C_{t+s}}{A_{t+s}} \right)^{-\gamma} \left( \frac{C_t}{A_t} \right)^\gamma \frac{A_t}{A_{t+s}} \quad (2.12)$$

is the representative household’s “real” stochastic discount factor.

### 2.4 The Fiscal Authority

In addition to issuing bonds, the fiscal authority consumes a stochastic fraction $\rho_t$ of the aggregate output $Y_t$. The government also levies a lump-sum tax or issues a subsidy to finance any shortfalls in government revenues. The government’s (balanced) budget constraint is therefore given by

$$P_t G_t + \sum_{\tau=1}^{\tau^* - 1} P_t^\tau B_{t-\tau}^{\tau+1} + B_{t-1}^1 = T_t + \sum_{\tau=1}^{\tau^*} P_t^\tau B_t^\tau \quad (2.13)$$
where $G_t = \rho_t Y_t$ is the real government expenditure. Here, the aggregate government spending shock is modeled as

$$\ln g_t = (1 - \phi_g) \ln g^* + \phi_g \ln g_{t-1} + \varepsilon_{g,t}$$

(2.14)

where $g_t = 1/(1 - \rho_t)$, $|\phi_g| < 1$, and, as in the case of the technology shock $\varepsilon_{a,t}$, the fiscal innovation $\varepsilon_{g,t}$ is assumed to be normally distributed with mean 0 and a regime-switching volatility process $\sigma^2_{g,v_{t,g}}$. We denote the transition probability matrix for the volatility process of the fiscal shock as

$$Q_g = \begin{bmatrix} q_{11}^g & 1 - q_{11}^g \\ 1 - q_{22}^g & q_{22}^g \end{bmatrix}$$

(2.15)

where $q_{ij}^g = \Pr[v_{t+1,g} = j|v_{t,g} = i]$.

### 2.5 Symmetric Equilibrium, Nonstochastic Values and the Linearized Model

From the utility maximization problem, the first-order condition with respect to the short term bond $B^1_t$ is

$$P^1_t = \mathbb{E}_t [M_{t,t+1}]$$

(2.16)

where

$$M_{t,t+1} = \delta \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} \frac{1}{\alpha_{t+1}} \frac{1}{\pi_{t+1}}$$

(2.17)

is the nominal stochastic discount factor (SDF) and $c_t = C_t/A_t$ is the stochastically detrended consumption at time $t$. Given the form of the SDF derived from our model, we use this condition in section 2.9 to price bonds of various maturities.

The aggregate labor supply from the household’s problem is derived as

$$1 = \frac{W_t}{A_t} c_t^{-\gamma}$$

(2.18)

In this economy, each intermediate goods producer faces the same marginal cost. Hence, in a symmetric equilibrium, $Y_t(j) = Y_t$, $H_t(j) = H_t$, $P_t(j) = P_t$ and $Q_t(j) = Q_t$. Thus, the representative intermediate-goods firm’s first order condition for profit maximization implies

$$1 = \zeta - \zeta c_t^\gamma + \varphi \left( \frac{\pi_t}{\pi^*} - 1 \right) \left( \frac{\pi_t}{\pi^*} \right) - \varphi \mathbb{E}_t \left[ \Lambda_{t,t+1} \left( \frac{\pi_{t+1}}{\pi^*} \right) - 1 \right] \left( \frac{\pi_{t+1} Y_{t+1}}{\pi^* Y_t} \right)$$

(2.19)
Finally, the aggregate resource constraint must hold in equilibrium:

\[ Y_t = C_t + G_t + AC_t \quad \text{and} \quad H_t = N_t = \int_0^1 N_t(j) \, dj \]  

(2.20)

which implies that

\[ c_t = \left( \frac{1}{g_t} - \frac{\varphi}{2} \left( \frac{\pi_t}{\pi^*} - 1 \right)^2 \right) x_t \]  

(2.21)

where \( x_t = Y_t / A_t \). Further, from the Euler equation, the implied nonstochastic value of the gross nominal interest rate \( R_t = 1 / P^1_t \) denoted by \( R^* \) is

\[ R^* = a^* \pi^* / \delta \]  

(2.22)

Also the equation (2.21) implies that the nonstochastic value of the detrended output is determined by

\[ x^* = \frac{c^*}{(1 - \rho^*)} \]  

(2.23)

where the nonstochastic value of the detrended consumption, \( c^* \) is

\[ \left[ \frac{\zeta - 1}{\zeta} \right]^{\frac{1}{\gamma}} \]  

(2.24)

In the absence of shocks, the economy converges to a steady-state growth path along which all the stationary variables are constant over time. It is important to note that in this setup while the steady state values of the aggregated macroeconomic variables (namely inflation, output and the risk-free short rate) are not affected by regime shifts, the steady state values of the long term bond yields are regime specific. As we show in Section (2.10), this is because the term premium is a function of the monetary policy reaction coefficients and the volatilities, both of which are subject to regime shifts.

Letting hats denote the percentage deviation of the variables from their respective steady state levels, for instance, \( \hat{c}_t = \ln(c_t / c^*) \), the model whose equilibrium dynamics is summarized by the equations (2.16), (2.19) and (2.20) can be cast in its log-linearized form as follows

\[ \hat{x}_t = \hat{g}_t + \mathbb{E}_t [\hat{x}_{t+1}] - \mathbb{E}_t [\hat{g}_{t+1}] - \frac{1}{\gamma} \left( \hat{R}_t - \mathbb{E}_t [\hat{\pi}_{t+1}] - \mathbb{E}_t [\hat{a}_{t+1}] \right) \]  

(2.26)

where \( \kappa = \frac{\zeta \gamma (c^*)^{-\gamma}}{\varphi} \).
2.6 The Central Bank

Given that the primary focus of this paper is to analyze the impact of monetary policy regime changes on the dynamics of the bond prices, we model the central bank’s policy function in terms of the generalized Taylor rule (Davig and Leeper (2007)). According to this rule, the bank adjusts the short term nominal interest rate $R_t$ in response to deviations of inflation $\pi_t$ from the target $\pi^*$, and stochastically detrended output $x_t$ from its non stochastic value $x^*$

$$\ln R_t = \ln R^* + \alpha_s t (\ln \pi_t - \ln \pi^*) + \beta_s t (\ln x_t - \ln x^*) + \ln e_t. \quad (2.27)$$

Defining $\hat{e}_t = \ln (e_t)$ and $\hat{R}_t$, $\hat{\pi}_t$ and $\hat{x}_t$ as in linearized model above, this interest rate rule can be written as

$$\hat{R}_t = \alpha_s \hat{\pi}_t + \beta_s \hat{x}_t + \hat{e}_t \quad (2.28)$$

where $\hat{e}_t$ is assumed to follow a stationary AR(1) process

$$\hat{e}_t = \phi_e \hat{e}_{t-1} + \varepsilon_{e,t} \quad (2.29)$$

with $|\phi_e| < 1$ and $\varepsilon_{e,t} \sim \mathcal{N}(0, \sigma_{e,v}^2)$. That is, the volatility of the monetary policy shock $\varepsilon_{e,t}$ also follows a two-state Markov switching process. In keeping with the notation for the two other shock volatilities, we denote the transition probability matrix for the volatility process of the monetary shock as

$$Q^e = \begin{bmatrix} q_{11}^e & 1 - q_{11}^e \\ 1 - q_{22}^e & q_{22}^e \end{bmatrix} \quad (2.30)$$

where $q_{ij} = \text{Pr}[\nu_{t+1,e} = j|\nu_{t,e} = i]$. It is important to clarify that changes in the volatility of the policy shock $\sigma_{e,v}^2$ are distinct from the monetary policy regime change. The volatility regime shift captures the possibility of heteroskedasticity of the short rate process.

---

4As in standard DSGE models this Taylor rule is exogenously specified, and it can be viewed as a reduced form model which is not derived from a welfare optimization of the monetary authority. Palomino (2010) recently analyzed the bond yield dynamics implied by a welfare-maximizing monetary policy and its credibility within a general equilibrium. He finds that credibility improvements reduce the exposure to inflation risk and bond risk premium.
Notice that in the above short rate equation the target inflation is assumed to be constant over time whereas the monetary policy coefficients $\alpha$ and $\beta$ are regime dependent, as indicated by the subscript $s_t$.\footnote{In the empirical counterpart of this paper we deal with the time span since the great moderation - a period of relatively low and stable inflation. We therefore attribute the variation in inflation to inflation gap (rather than its trend) by assuming a constant target inflation. As we show below, the virtue of this simplifying assumption is that it allows us to isolate the effect of all monetary policy regime changes solely through changes in the reaction coefficients of inflation and output gaps. In contrast, if one were to analyze a longer time period including the 1960’s and 1970’s, regime shifts in the target inflation might be essential as in Schorfheide (2005), Bekoert, Cho, and Moreno (2010), Cogley and Sbordone (2008) and Davig and Doh (2009).} We interpret the regime dependency of the monetary policy coefficients as shifts between active and passive (or less active) regimes. A convenient way to model this is to assume a two-state Markov process for the policy regimes with the transition probability matrix

$$
P = \begin{bmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{bmatrix}
$$

with $p_{ij} = \Pr[s_{t+1} = j|s_t = i]$.

### 2.7 Summary of the Exogenous Shock Processes

Recall that there are three structural shocks in this model: the technology shock $\varepsilon_{a,t}$, the fiscal shock $\varepsilon_{g,t}$ and the monetary shock $\varepsilon_{e,t}$. We assume that these shocks are independent of one another. Combining this assumption with the notation for the regime-dependent volatilities introduced earlier, we summarize the shock processes as follows

$$
\bar{f}_t = \begin{bmatrix} \hat{a}_t \\ \hat{g}_t \\ \hat{e}_t \end{bmatrix} = \phi \bar{f}_{t-1} + \varepsilon_t
$$

where

$$
\phi = \begin{bmatrix} \phi_a & 0 & 0 \\ 0 & \phi_g & 0 \\ 0 & 0 & \phi_e \end{bmatrix}, \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{a,t} \\ \varepsilon_{g,t} \\ \varepsilon_{e,t} \end{bmatrix} \sim \mathcal{N} \left( 0_{3 \times 1}, \Omega_{\varepsilon_t} = \begin{bmatrix} \sigma^2_{a,v_t,a} \\ \sigma^2_{g,v_t,g} \\ \sigma^2_{e,v_t,e} \end{bmatrix} \right).
$$

We further assume that the Markov process for the policy regimes $s_t$ is independent of the volatility regimes $v_t = (v^a_t, v^g_t, v^e_t)$. For notational convenience, we aggregate the regime indicators comprising of both $s_t$ and $v_t$ into $d_t$ as follows (shown here for the number of policy regimes $m = 2$ and the number of volatility regimes $v = 8$).
This aggregation enables us to denote any possible distinct combination of the policy and volatility regimes with a single notation. For instance, \( d_t = 1 \) captures the first state for the policy regime as well as for each of the three volatility regimes. Thus, the total number of regimes \( d \) equals \( m \times v \). The corresponding “aggregated” transition probability matrix can therefore be written as \( Z = Q^s \otimes Q^g \otimes Q^a \otimes P \).

In section 2.10, we show that the recurrence of the volatility regimes, combined with the fact that \( v^a_t, v^g_t, v^e_t \) and \( s_t \) are independent, implies that both the model-implied term premium and the expected excess returns are time-varying in each monetary policy regime.

### 2.8 Model Solution and Determinacy Restrictions

For concerns of theoretical tractability, as well as econometric convenience, we focus on the (local) behavior of the economy around its deterministic, non-stochastic steady state. Our interest lies in the linearized system of equations (2.25)-(2.26), (2.28) and (2.32). On substituting (2.28) into (2.26), this system collapses to

\[
0 = \delta E_t [\hat{x}_{t+1}] - \hat{x}_t + \kappa (\hat{x}_t - \hat{y}_t) \tag{2.33}
\]

\[
0 = E_t [\hat{\pi}_{t+1}] + \gamma E_t [\hat{\pi}_{t+1}] - \alpha_s \hat{\pi}_t - (\beta_s + \gamma) \hat{x}_t + \phi_a \hat{a}_t - \gamma (\phi_g - 1) \hat{g}_t - \hat{\epsilon}_t \tag{2.34}
\]

We now have a simultaneous system of two equations in two key aggregated variables of interest (output deviation from its steady state, \( \hat{x}_t \), and, deviation of inflation from its target, \( \hat{\pi}_t \)) and three unobservable shocks (to technology \( \hat{a}_t \), government expenditure \( \hat{g}_t \) and monetary policy \( \hat{\epsilon}_t \)).

To analyze the evolution of the two variables of interest we first need to solve this model. For this purpose, we adopt the solution method of Davig and Leeper (2007). The solution process rids the system of the unobservable expectational terms by casting
them as a linear function of the underlying shock processes. In this paper we restrict our attention to the unique (determinate) solution.\(^6\) A full discussion of the solution algorithm is well beyond the scope of this paper. In terms of the computational details, we begin by casting the endogenous variables as a linear function of the shock processes

\[
\begin{bmatrix}
\hat{\pi}_{it} \\
\hat{x}_{it} \\
\hat{m}_{it}
\end{bmatrix}
= \begin{bmatrix}
h^a_p(s_t = i) \\
h^g_p(s_t = i) \\
h^e_p(s_t = i) \\
h^a_x(s_t = i) \\
h^g_x(s_t = i) \\
h^e_x(s_t = i)
\end{bmatrix}
H_{s_t = i}
\begin{bmatrix}
\bar{f}_t
\end{bmatrix}
\tag{2.35}
\]

where \(\hat{\pi}_{it}\) and \(\hat{x}_{it}\) denote the state-contingent \((s_t = i)\) values of inflation gap and output gap, respectively.

On inserting this linear solution into the system of equations (2.33)-(2.34), the conditional expectation of the one-period ahead inflation gap and output gap are

\[
\mathbb{E}_t \left[ (\hat{\pi}_{t+1} \quad \hat{x}_{t+1} )' | s_t = i \right] = \mathbb{E}_t [\bar{H}_{s_{t+1}} \bar{f}_{t+1} | s_t = i]
\]

\[
= p_{i1} \bar{H}_{s_{t+1} = 1} \phi \bar{f}_t + p_{i2} \bar{H}_{s_{t+1} = 2} \phi \bar{f}_t
\]

Equivalently, on letting \(h^j_{\pi,i} \equiv h^j_\pi(s_t = i)\) and \(h^j_{x,i} \equiv h^j_x(s_t = i), \; (j = a, g, e)\), \(\mathbb{E}_t [\hat{\pi}_{t+1} | s_t = i]\) can be expressed as

\[
p_{i1} \left[ h^a_{\pi,1} \phi_\pi \hat{a}_t + h^g_{\pi,1} \phi_g \hat{g}_t + h^e_{\pi,1} \phi_e \hat{e}_t \right] + p_{i2} \left[ h^a_{\pi,2} \phi_\pi \hat{a}_t + h^g_{\pi,2} \phi_g \hat{g}_t + h^e_{\pi,2} \phi_e \hat{e}_t \right]
\tag{2.37}
\]

and \(\mathbb{E}_t [\hat{x}_{t+1} | s_t = i]\) as

\[
p_{i1} \left[ h^a_{x,1} \phi_\pi \hat{a}_t + h^g_{x,1} \phi_g \hat{g}_t + h^e_{x,1} \phi_e \hat{e}_t \right] + p_{i2} \left[ h^a_{x,2} \phi_\pi \hat{a}_t + h^g_{x,2} \phi_g \hat{g}_t + h^e_{x,2} \phi_e \hat{e}_t \right]
\tag{2.38}
\]

Next, to compute the regime-dependent solutions \(\bar{H}_{s_t}\), one relies on the method of undetermined coefficients, setting the coefficients of \(\hat{a}_t\), \(\hat{g}_t\) and \(\hat{e}_t\) equal to zero and solving for the resulting solution in terms of the coefficients in \(\bar{H}_{s_t}\). Additional computational details of the solution are provided in Appendix A.

Note that because we work with a first-order approximation of the equilibrium conditions of the households and firms, the solution coefficients \(\bar{H}_{s_t}\) depend only on the

\(^6\)Farmer, Zha, and Waggoner (2009) show the existence of general forms of indeterminate equilibria in the quasi-linear system that depend not only on the structural shocks, but also on additional autoregressive shock driven by the structural shocks. The general forms include Davig and Leeper (2007) solutions as special cases.
monetary policy regime $s_t$ and not the volatility regimes $v_t$. In addition, recall that $\ln \pi_t = \hat{\pi}_t + \ln \pi^*$ and $\ln \left( Y_t / A_t \right) = \hat{x}_t + \ln x^*$. Hence, the solution for the DSGE model in equation (2.35) can be rewritten as

$$
\begin{bmatrix}
\ln \pi_t \\
\ln Y_t \\
\ln x^* \\
\ln A_t
\end{bmatrix} =
\begin{bmatrix}
\ln \pi^* \\
h^2_x(d_t = i) \\
h^2_x(d_t = i) \\
h^2_x(d_t = i)
\end{bmatrix} +
\begin{bmatrix}
h^2_x(d_t = i) & h^2_x(d_t = i) & h^2_x(d_t = i) & 0 \\
h^2_x(d_t = i) & h^2_x(d_t = i) & h^2_x(d_t = i) & 1
\end{bmatrix}
\begin{bmatrix}
m_t \\
J \\
H_{d_t=1} \\
f_t
\end{bmatrix}
$$

As can be seen in section 3.1 below, this representation of the solution turns out to be very convenient in the construction of the empirical model.

### 2.9 The Bond Prices

The first order conditions for the short and long term bonds $B^\tau_t$ ($1 \leq \tau \leq \tau^*$), which are absent in standard DSGE models without long term bonds, can be shown to have the form

$$P^\tau_t = \mathbb{E}_t \left[ M_{t,t+\tau} \right]$$

where

$$M_{t,t+\tau} = \delta \left( \frac{c_{t+\tau}}{c_t} \right)^{-\gamma} \frac{1}{a_{t+\tau} \pi_{t+\tau}}$$

is the intertemporal marginal rate of substitution between time $t$ and $t + \tau$. These first order conditions provides the demand function for long term bonds. Assuming that the supply of these bonds is perfectly elastic, and using the law of iterated expectation, one has the standard asset-pricing conclusion that

$$P^\tau_t = \mathbb{E}_t \left[ M_{t,t+1} \times M_{t+1,t+\tau} \right]$$

This equation implies that the equilibrium bond prices at time $t$, denoted by $P^{(\tau)}_{dt,t}$, satisfy the no-arbitrage condition

$$P^{(\tau)}_{dt,t} = \mathbb{E} \left[ M_{t,t+1} P^{(\tau-1)}_{d_{t+1},t+1} | \bar{F}_t, d_t \right]$$
and are a function of the model-determined pricing kernel which itself is a function of \(d_t\) and the exogenous shocks.

To calculate the form of these prices, we express the nominal pricing kernel in log-linearized form as

\[
\ln M_{t,t+1} = m_{t,t+1} + \lambda_{d_{t+1}} \bar{f}_{t} + L_{d_{t+1}} \varepsilon_{t+1}
\]

(2.44)

where

\[
c_{d_{t+1}} = -\ln R^{*} - \frac{1}{2} L_{d_{t+1}} \Omega_{d_{t+1}} L'_{d_{t+1}} \tag{2.45}
\]

\[
\lambda_{d_{t+1}} = - \begin{pmatrix} 1 & \gamma \end{pmatrix} \bar{H}_{d_{t+1}} \phi + \begin{pmatrix} 0 & \gamma \end{pmatrix} \bar{H}_{d_t} + \begin{pmatrix} -1 & \gamma & 0 \end{pmatrix} \phi - \begin{pmatrix} 0 & \gamma & 0 \end{pmatrix} \tag{2.46}
\]

\[
L_{d_{t+1}} = - \begin{pmatrix} 1 & \gamma \end{pmatrix} \bar{H}_{d_{t+1}} + \begin{pmatrix} -1 & \gamma & 0 \end{pmatrix} \tag{2.47}
\]

As in Ang et al. (2008), we assume that the one period bond is risk-free by augmenting the Jensen’s inequality term to equation (2.45). This assumption is necessary to generate a positive average term premium in our formulation. More importantly, due to the assumption the equilibrium short rate obtained from the recursions when \(\tau = 1\), which is discussed below, is exactly the same as the value of the short rate from the Taylor rule at equilibrium (obtained by substituting the equilibrium values of output and inflation into the Taylor rule). This agreement is a consequence of the fact that bond pricing as exemplified here comes from the dynamic general equilibrium solution of the model.

Also note that the market price of risk, which is associated with the structural shocks \(\varepsilon_{t+1}\), is given by the elements in \(L_{d_{t+1}} \Omega_{d_{t+1}}^{1/2}\).

Let \(p^{(\tau)}_{dt,t} \equiv \ln P^{(\tau)}_{dt,t}\) denote the log price of a \(\tau\)-period maturity bond at time \(t\) in regime \(d_t\) and suppose that

\[
-p^{(\tau)}_{dt,t} = a_{d_t}(\tau) + b_{d_t}(\tau)' \bar{f}_{t}. \tag{2.48}
\]

Under this guess and the form of the pricing kernel above we can use the method of undetermined coefficients to derive the recursive expressions for \(i \in \{1, 2, ..., d\}\)

\[
a_i(\tau) = \ln R^{*} + \sum_{j=1}^{d} p_{ij} \left( a_j(\tau - 1) + L_j \Omega_j b_j(\tau - 1)' - \frac{1}{2} b_j(\tau - 1)' \Omega_j b_j(\tau - 1) \right) \tag{2.49}
\]

\[
b_i(\tau)' = \sum_{j=1}^{d} p_{ij} \left( b_j(\tau - 1)' \phi - \lambda_{i,j} \right). \tag{2.50}
\]
Further details of this derivation are provided in Appendix B. These recursions are initialized by the no-arbitrage condition at \( \tau = 0 \)

\[
a_i(0) = b_i(0) = 0 \quad \text{for all } i
\]

Then, the continuously compounded yield to maturity \( r^{(\tau)}_{d_{t,t}} \) for the zero-coupon nominal bond is given by

\[
r^{(\tau)}_{d_{t,t}} = -\frac{\tilde{p}_{d_{t,t}}}{\tau} = \tilde{a}_{d_{t}}(\tau) + \tilde{b}_{d_{t}}(\tau)\tilde{f}_t
\]

with \( \tilde{a}_{d_{t}}(\tau) = \frac{a_{d_{t}}(\tau)}{\tau} \) and \( \tilde{b}_{d_{t}}(\tau) = \frac{b_{d_{t}}(\tau)}{\tau} \).

It is useful to note that the factor loadings \( \tilde{b}_{d_{t}}(\tau) \) are independent of the volatility regimes because \( \lambda_{i,j} \) is determined by the parameters in the linearized Euler equation (2.44).

### 2.10 Measures of Long-Term Bond Risk

We focus on three different measures of riskiness of long-term bonds in each regime: the term premium, the expected excess return on the long-term bond and the slope of the yield curve. We now discuss the characteristics of each of these measures.

The term spread is simply the difference between the long-term bond yield and the short rate. As is well-known, it can be rewritten as the sum of two components

\[
r^{(\tau)}_{d_{t,t}} - r^{(1)}_{d_{t,t}} = \left[ \frac{1}{\tau} \sum_{l=0}^{\tau-1} \mathbb{E}_t \left[ r^{(1)}_{d_{t+1,t+l}} - r^{(1)}_{d_{t,t}} \right] \right] + \left[ \frac{1}{\tau} \sum_{i=1}^{\tau-1} \text{ex}r^{(\tau+1-i)}_{d_{t,t}} \right],
\]

where \( \text{ex}r^{(\tau)}_{d_{t,t}} \) denotes the one-period expected excess return to holding the \( \tau \)-period bond. The first component on the right is the expectation hypothesis. Under risk-neutral pricing, after adjusting for risk, agents are indifferent between holding a long term bond and a one period risk-free bond. The risk adjustment is the term premium, captured by the second term on the right.

Two important points emerge from equation (2.53). First, the term spread depends on the expected excess returns as well as the expected average future short rate. Second,
the term premium reflects the expected excess return to all bonds of maturities less than \( \tau \)-periods, not just expected excess return to the \( \tau \)-period bond.

The one-period expected excess return of the \( \tau \)-period bond at time \( t \) is then defined as

\[
\text{exr}_{dt,t}^{(\tau)} = \left[ \mathbb{E}_t \left[ p_{dt+1,t+1}^{(\tau-1)} - p_{dt,t}^{(\tau)} \right] - (-p_{dt,t}^{(1)}) \right] - (\tau - 1) r_{dt+1,t+1}^{(\tau-1)} + \tau r_{dt,t}^{(\tau)} - r_{dt,t}^{(1)}
\] (2.54)

The first term on the right side of (2.54) is the expected one-period return to holding the bond and the second term is the one-period risk-free rate. Importantly, \( \text{exr}_{dt,t}^{(\tau)} \) can be expressed as a sum of the factor risk component \( \text{FR}_{dt=t}^{(\tau)} \) and the regime-shift risk component \( \text{RS}_{dt=t}^{(\tau)} \)

\[
\text{exr}_{dt=t}^{(\tau)} = \text{FR}_{dt=t}^{(\tau)} + \text{RS}_{dt=t}^{(\tau)}
\] (2.55)

where

\[
\text{FR}_{dt=t}^{(\tau)} = \sum_{j=1}^{d} p_{ij} L_j \Omega_j \beta_j (\tau - 1) - \frac{1}{2} \sum_{j=1}^{d} p_{ij} \beta_j (\tau - 1)' \Omega_j \beta_j (\tau - 1)
\] (2.56)

\[
\text{RS}_{dt=t}^{(\tau)} = \left[ \sum_{j=1}^{d} p_{ij} K_{j,t} \right] \left[ \sum_{j=1}^{d} p_{ij} W_{i,j,t} \right] - \sum_{j=1}^{d} p_{ij} W_{i,j,t} K_{j,t} - \frac{1}{2} \sum_{j=1}^{d} p_{ij} K_{j,t}^2 + \frac{1}{2} \left( \sum_{j=1}^{d} p_{ij} K_{j,t} \right)^2
\] (2.57)

and

\[
W_{dt,d_{t+1},t} = c_{d_{t+1}} + \lambda_{dt,d_{t+1}} \bar{f}_t
\]

\[
K_{d_{t+1},t} = -a_{d_{t+1}} - b_{d_{t+1}} (\tau - 1)' \bar{f}_t
\] (2.58)

Similarly, it is straightforward to decompose the term premium, denoted by \( \text{TP}_{dt=t}^{(\tau)} \), in equation (2.53) as the sum of two averages.

The proof of these results is given in Appendix C. Notice that the terms in the factor risk component \( \text{FR}_{dt=t}^{(\tau)} \) are all associated with the structural shocks in the following period. Not surprisingly, the compensation demanded for holding long term bonds depends largely on the size of the factor shocks \( \Omega_j^{1/2} \), the sensitivity of the yields to
the factor shocks \( b_j(\tau - 1) \) and the price of the risks \( L_j \Omega_j^{1/2} \). This market price of the risks is maturity-independent and determines how much one unit of risk translates into an expected excess return. Meanwhile, the regime-shift risk component \( RS_{d_t = i, t}^{(\tau)} \) will be absent under either a single regime model or a regime switching model with market price of regime shift risk equal to zero as pointed out by Dai, Singleton, and Yang (2007). Finally, it is interesting that \( FR_{d_t = i}^{(\tau)} \) is a regime-specific constant, whereas \( RS_{d_t = i, t}^{(\tau)} \) depends on the current values of the time-varying factors. Consequently, the expected excess return is time varying and so is the term premium\(^7\). Moreover, our regime-dependent factor loadings, generated by the monetary policy regime shifts, allow for the term premium to vary independently of factor volatility. As pointed out in Duffee (2002), this additional flexibility helps improve the forecast accuracy of future yields.

3 Estimation methodology

3.1 State Space Formulation

We begin by recalling the solution to the DSGE model in equation (2.39)

\[
\begin{bmatrix}
\ln \pi_t \\
\ln Y_t \\
\ln \pi^* \\
\ln x^*
\end{bmatrix} = 
\begin{bmatrix}
\ln \pi^* \\
\ln x^*
\end{bmatrix} + 
\begin{bmatrix}
h_a^*(d_t = i) & h_a^*(d_t = i) & h_a^*(d_t = i) & 0 \\
h_a^*(d_t = i) & h_a^*(d_t = i) & h_a^*(d_t = i) & 1 \\
h_a^*(d_t = i) & h_a^*(d_t = i) & h_a^*(d_t = i) & 1
\end{bmatrix}
\begin{bmatrix}
\bar{f}_t \\
\ln A_t
\end{bmatrix}
\]

Note that the short rate \( r_t^{(1)} \), which is set by the central bank according to the Taylor (1993) rule, incorporates the monetary policy shock. Thus, as in the estimation of standard DSGE models, we assume that the final outcomes \( (m_t, \hat{R}_t) \) are generated without additional (measurement) errors. As we show in Appendix D, the benefit of this assumption is that, given the regime process \( D_n \) and the initial value of the technology shock \( \ln A_0 \), the shock process \( \bar{f}_t \) can be solved entirely in terms of the observable quantities \( \ln (P_t/P_{t-1}) \), \( \ln Y_t \) and \( \hat{R}_t \), where \( \ln A_0 \) is treated as an additional parameter to be estimated. This, in turn, substantially simplifies the calculation of the likelihood function conditioned on the regimes.

\(^7\)An alternative way of achieving a time-varying term premium is to work with a second-order or third-order approximation of the optimality conditions (Doh (2009) and Bansal and Yaron (2004)). However, a suitable solution method for such non-linear models under a multi-regime specification currently does not exist.
We implement our model on a data set that comprises 5 yields of US T-bills measured on a quarterly basis. We denote these quarterly maturities of interest as
\[ \{\tau_1, \tau_2, \tau_3, \tau_4, \tau_5\} = \{1, 2, 4, 8, 20\} \]
and let
\[ R_t = \begin{pmatrix} r_t^{(\tau_1)} & r_t^{(\tau_2)} & r_t^{(\tau_3)} & r_t^{(\tau_4)} & r_t^{(\tau_5)} \end{pmatrix} \]
where \( r_t^{(\tau_i)} = r_{i,t} \). We assume that all bonds with maturity greater than 1 period are priced with errors - that is, the short rate is treated as a basis yield. Let \( \bar{a}_t = (\bar{a}_t(\tau_1), \bar{a}_t(\tau_2), \ldots, \bar{a}_t(\tau_1))' \) and \( \bar{b}_t = (\bar{b}_t(\tau_1), \bar{b}_t(\tau_2), \ldots, \bar{b}_t(\tau_5))' \). Then the observable quantities \( m_t \) and \( R_t \) are stacked to obtain the measurement equation
\[
\begin{bmatrix} m_t \\ R_t \end{bmatrix} = \begin{bmatrix} J \\ \bar{a}_t \end{bmatrix} + \begin{bmatrix} H_{dt} \\ \bar{b}_t \end{bmatrix} f_t + \begin{bmatrix} 0_{3 \times 4} \\ I_4 \end{bmatrix} e_t \tag{3.2}
\]
where \( e_t \sim \mathcal{N}_4(0, \Sigma); \Sigma = \text{diag}(\sigma_2^2, \sigma_3^2, \sigma_4^2, \sigma_5^2) \). We complete the state space formulation by combining equation (2.32) with the technology shock process \( \ln A_t = \ln a^* + \ln A_{t-1} + \hat{\alpha}_t \)
and write the transition equation as
\[
\begin{bmatrix} \bar{f}_t \\ \ln A_t \\ \ln a^* \end{bmatrix} = \begin{bmatrix} 0_{3 \times 1} \\ \phi_3 \times 3 \\ 0_{3 \times 1} \end{bmatrix} + \begin{bmatrix} \phi_a \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \bar{f}_{t-1} \\ \ln A_{t-1} \end{bmatrix} + \begin{bmatrix} I_3 \\ 1 \\ 0 \end{bmatrix} \varepsilon_t \tag{3.3}
\]
with \( \varepsilon_t \sim \mathcal{N}(0_{3 \times 1}, \Omega_{dt}) \). For notational convenience, we let \( \theta \) denote the free parameters in \( a_{dt}, b_{dt}, \Sigma, \mu, G \) and \( \Omega_{dt} \).

### 3.2 Prior Distribution

We formulate the prior on the parameters to reflect the belief that (under the prior) the average term premium is positive (Chib and Ergashev (2009)). This prior is, of course, restricted to the subset of the parameter space that implies a unique determinate solution to the model. Finally, various blocks of parameters are assumed to be a priori independent. Table I summarizes our prior.

Under this prior, the annual short interest rate is centered at 4.4% with a standard deviation of 0.32%. The steady state technology growth ranges from 1.13% to 2.17%.
Table I: Prior distribution for the 16-regime model parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>density</th>
<th>mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>beta</td>
<td>0.9992</td>
<td>0.0006</td>
</tr>
<tr>
<td>$\phi_a$</td>
<td>beta</td>
<td>0.3688</td>
<td>0.1189</td>
</tr>
<tr>
<td>$\phi_g$</td>
<td>beta</td>
<td>0.8472</td>
<td>0.1092</td>
</tr>
<tr>
<td>$\phi_e$</td>
<td>beta</td>
<td>0.6123</td>
<td>0.1293</td>
</tr>
<tr>
<td>$p_{11}$</td>
<td>beta</td>
<td>0.9745</td>
<td>0.0221</td>
</tr>
<tr>
<td>$p_{22}$</td>
<td>beta</td>
<td>0.9745</td>
<td>0.0221</td>
</tr>
<tr>
<td>$q_{11}$</td>
<td>beta</td>
<td>0.8995</td>
<td>0.0401</td>
</tr>
<tr>
<td>$q_{22}$</td>
<td>beta</td>
<td>0.8997</td>
<td>0.0401</td>
</tr>
<tr>
<td>$q_{11}$</td>
<td>beta</td>
<td>0.8997</td>
<td>0.0401</td>
</tr>
<tr>
<td>$q_{22}$</td>
<td>beta</td>
<td>0.8997</td>
<td>0.0401</td>
</tr>
<tr>
<td>$400 \times \ln R^*$</td>
<td>normal</td>
<td>4.4426</td>
<td>0.3141</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>gamma</td>
<td>0.4985</td>
<td>0.3036</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>normal</td>
<td>1.5154</td>
<td>0.2965</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>normal</td>
<td>1.9972</td>
<td>0.3161</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>normal</td>
<td>0.9968</td>
<td>0.3139</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>normal</td>
<td>1.0067</td>
<td>0.3117</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>gamma</td>
<td>39.952</td>
<td>10.011</td>
</tr>
<tr>
<td>$400 \times \ln \alpha^*$</td>
<td>normal</td>
<td>1.6575</td>
<td>0.3147</td>
</tr>
<tr>
<td>$\ln x^*$</td>
<td>gamma</td>
<td>0.9988</td>
<td>0.0978</td>
</tr>
<tr>
<td>$\ln A_0$</td>
<td>normal</td>
<td>2.3115</td>
<td>0.1009</td>
</tr>
<tr>
<td>$2.0 \times 10^4 \times \sigma_{\omega,1}^2$</td>
<td>inverse gamma</td>
<td>0.9539</td>
<td>0.1895</td>
</tr>
<tr>
<td>$2.0 \times 10^5 \times \sigma_{\eta,1}^2$</td>
<td>inverse gamma</td>
<td>0.9596</td>
<td>0.1937</td>
</tr>
<tr>
<td>$3.0 \times 10^4 \times \sigma_{\varepsilon,1}^2$</td>
<td>inverse gamma</td>
<td>0.9603</td>
<td>0.1951</td>
</tr>
<tr>
<td>$1.0 \times 10^4 \times \sigma_{\omega,2}^2$</td>
<td>inverse gamma</td>
<td>0.9635</td>
<td>0.1941</td>
</tr>
<tr>
<td>$1.0 \times 10^5 \times \sigma_{\eta,2}^2$</td>
<td>inverse gamma</td>
<td>0.9635</td>
<td>0.1956</td>
</tr>
<tr>
<td>$2.5 \times 10^3 \times \sigma_{\varepsilon,2}^2$</td>
<td>inverse gamma</td>
<td>0.9613</td>
<td>0.1943</td>
</tr>
<tr>
<td>$1.4 \times 10^7 \times \sigma_{\nu,1}^2$</td>
<td>inverse gamma</td>
<td>0.9623</td>
<td>0.1927</td>
</tr>
<tr>
<td>$3.0 \times 10^6 \times \sigma_{\nu,2}^2$</td>
<td>inverse gamma</td>
<td>0.9620</td>
<td>0.1948</td>
</tr>
<tr>
<td>$1.2 \times 10^6 \times \sigma_{\varepsilon,4}^2$</td>
<td>inverse gamma</td>
<td>0.9605</td>
<td>0.1942</td>
</tr>
<tr>
<td>$6.0 \times 10^5 \times \sigma_{\varepsilon,5}^2$</td>
<td>inverse gamma</td>
<td>0.9611</td>
<td>0.1979</td>
</tr>
</tbody>
</table>

For the variance of the structural shocks and the risk aversion parameters, the respective marginal prior distributions are set to generate an average positive term premium. The marginal prior distributions of the other parameters are set to be consistent with the existing empirical literature on the term structure and new Keynesian DSGE models. For example, the prior distribution of the slope parameter $\kappa$ in the Phillips curve is from Lubik and Schorfheide (2004) and the transition probabilities are consistent with Chib
and Kang (2010). It is important to note that the values of the hyperparameters in these marginal distributions are chosen to allow the parameters to vary considerably in the domain supported by the determinacy condition. Furthermore, in this Markov process for the policy regime, it is necessary to impose a restriction on the relative magnitudes of $\beta_{s_t=1}$, $\beta_{s_t=2}$, $\alpha_{s_t=1}$ and $\alpha_{s_t=2}$ for identification. We also normalize the labels for the volatility regimes by restricting that all diagonal elements in $\Omega_d$ are greater than those in $\Omega_1$. Finally, we note that our prior is quite symmetric across regimes in order to avoid the identification of the regimes through the prior information.

![Graphs illustrating time series plots for inflation and output growth](image)

**Figure 1: The prior-implied inflation and output growth dynamics** These graphs are based on 50,000 simulated draws of the parameters from the prior distribution. In the graphs on the left, the surfaces correspond to the 2.5%, 50%, and 97.5% quantile surfaces of the term structure dynamics in annualized percents implied by the prior distribution for each regime.

To understand what the prior distribution implies for the outcomes, we sample the parameters 20,000 times from the prior, and then for each drawing of the parameters, we simulate the shocks, macroeconomic variables and yields according to the structural model. The sampled sequences for each macroeconomic variable in annualized percents are shown in Figure 1. As one can see from those figures, this prior implies a deviation of roughly 5% for output growth and 7% for inflation. Similarly, the implied term structure in annualized percents for each time period is reproduced in Figure 2. As one can see, the implied average term structure is gently upward sloping in each regime with considerable a priori variation.
Figure 2: The prior-implied term structure dynamics. These graphs are based on 50,000 simulated draws of the parameters from the prior distribution. In the graphs on the left, the surfaces correspond to the 2.5%, 50%, and 97.5% quantile surfaces of the term structure dynamics in annualized percents implied by the prior distribution for each regime.

3.3 Posterior Distribution and MCMC Sampling

We now have the necessary ingredients to calculate the posterior distribution of the parameters. Let $D_n = \{d_t\}_{t=0,1,...,n}$ denote the sequence of the unobserved regime indicators, $F_n = \{f_t\}_{t=0,1,...,n}$ the sequence of the factors, $y = \{y_t\}_{t=0,1,...,n}$ the full set of observables (date set) and $\theta$ the collection of the model parameters. Then, the posterior distribution that we would like to analyze is given by

$$\pi(\theta, F_n, D_n | y) \propto f(y | \theta, F_n, D_n) p(F_n, D_n | \theta) \pi(\theta)$$

where $f(y | \theta, F_n, D_n)$ is the distribution of the data given the regime indicators and the parameters, $p(F_n, D_n | \theta)$ is the density of the latent factors and the regime-indicators given the parameters, and $\pi(\theta)$ is the prior density of $\theta$. Note that by conditioning on $D_n$ we avoid the calculation of the likelihood function $f(y | \theta)$ whose computation is more involved.

We summarize this complex posterior distribution by MCMC simulation methods. The basic idea behind the MCMC approach is to produce a (correlated) sequence of draws by simulating a Markov process whose invariant distribution is the target density (Chib and Greenberg (1995)). Practically, the sampled draws after a suitably specified
burn-in phase are taken as samples from the posterior density. We construct our simulation procedure by sampling various blocks of parameters and latent variables in turn within each MCMC iteration. The distributions of these various blocks of parameters are each proportional to the joint posterior $\pi(\theta, F_n, D_n|y)$. In particular, after initializing the model parameters $\theta$ and the regimes $D_n$, we go through an iterative sequence of steps in each MCMC cycle. First, we sample $\theta$ from the posterior distribution that is proportional to

$$f(y|\theta, D_n)\pi(\theta)$$

where $f(y|\theta, D_n)$ is obtained from the standard Kalman filtering recursions given the regime indicators $D_n$. The sampling of $\theta$ from the latter density is done by the tailored randomized block Metropolis-Hastings (TaRB-MH) method following Chib and Ramamurthy (2010). The use of this MCMC method is essential to improve the mixing of the draws when there is no natural way of grouping the parameters. In the next step we solve for $F_n$ in terms of the observable macro quantities and the short yield. Finally, we sample $D_n$ conditioned on $F_n$ and $\theta$ in one block by the algorithm of Chib (1996). These steps of the MCMC algorithm are summarized below. A more detailed description can be found in Appendix D.

**Algorithm: MCMC sampling**

**Step 1** Initialize $(\theta, D_n)$ and fix $n_0$ (the burn-in) and $n_1$ (the MCMC sample size)

**Step 2** Sample $\theta$ conditioned on $(y, D_n)$

**Step 3** Sample $F_n$ conditioned on $(y, \theta, D_n)$

**Step 4** Sample $D_n$ conditioned on $(y, \theta, F_n)$

**Step 5** Repeat Steps 2-4, discard the draws from the first $n_0$ iterations and save the subsequent $n_1$ draws.
3.4 Model Comparison

From the perspective of the data, we are interested in knowing whether a multi-regime model improves on a single regime model. Furthermore, we are also interested in learning which of these multi-regime specifications best describes the data. To address these questions, we compare the following models: a single regime model ($M_1$), a model with one regime change in monetary policy but no regime shifts in the shock volatilities (2 policy regimes, $M_2$), a model with one regime change in monetary policy together with simultaneous regime shifts in all three volatilities (2 policy regimes and 2 volatility regimes, $M_4$), and, finally, a model with one regime change in monetary policy together with independent regime shifts in each of the three volatilities (2 policy regimes and 8 volatility regimes, $M_{16}$).

<table>
<thead>
<tr>
<th>$M_d$</th>
<th># of monetary policy regimes($m$)</th>
<th># of volatility regimes($v$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$M_2$</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>$M_4$</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>$M_{16}$</td>
<td>2</td>
<td>8</td>
</tr>
</tbody>
</table>

Within the Bayesian context, these models are compared in terms of the marginal likelihoods $m(y|M_d)$ and their ratios (Bayes factors). Following Chib (1995) an estimate of the log marginal likelihood can be calculated from the following fundamental identity

$$\ln \hat{m}(y|M_d) = \ln f(y|\theta^*, M_d) + \ln \pi(\theta^*, M_d) - \ln \hat{\pi}(\theta^*|y,M_d) \quad (3.6)$$

where $d=1, 2, 4, \text{and } 16$, and $\theta^*$ is a high density point in the support of the parameter space. Notice that the first term on the right hand side of this expression is the likelihood ordinate. The second term is the prior ordinate. Both of these are readily available. The third term, the posterior ordinate $\pi(\theta^*|y,M_d)$, is estimated from a marginal-conditional decomposition (Chib (1995)). The specific implementation in this context requires the technique of Chib and Jeliazkov (2001) as modified by Chib and Ramamurthy (2010) for the case of randomized blocks. For details we refer the interested reader to these papers.
4 Results

Our empirical results are based on the collection of historical yields of treasury bills with maturities 1, 2, 4, 8 and 20 quarters, real GDP per capita and inflation for the sample period 1986:Q4 to 2010:Q3. The inflation is calculated as a quarterly decimal change in the GDP deflator. This data is available online from the Board of Governors of the Federal Reserve System (Gurkaynak, Sack, and Wright (2007)). From the DSGE model perspective, the relevance of this sample period is that it is known for its relative stability compared to the major oil price shocks during the 1970s, the monetary policy experiment and the Volcker disinflation period in the early 1980s.

4.1 Policy Regime Changes and Volatility of Structural Shocks

Table II unambiguously confirms the presence of a regime shift in monetary policy dated around the early 2000’s. In particular, it is interesting to note that both models $M_2$, that focuses purely on just the regime change in monetary policy without any regime shifts in the volatilities, and $M_4$ which incorporates simultaneous shifts in both policy and volatility regimes, provide the same estimate of the breakpoint as model $M_{16}$, which is the model with independent policy and volatility regimes. Based on marginal likelihoods, however, it is clear that the best fitting model is $M_{16}$.

As mentioned earlier, in this general equilibrium setup, both the structural shocks and the policy reaction coefficients drive output, inflation and the term premium dynamics, which is distinct from that in a partial equilibrium approach. This points to the fundamental notion that the macroeconomic fundamentals and the entire term structure, not just the short-term rate, contain valuable information about monetary policy regime shifts. Also it is important to note that this approach does not require one to estimate additional parameters in comparison to a DSGE model without multiple bonds. This distinction also helps explain why our finding of the regime-switching point positions is different from that in Ang et al. (2010), Bikbov and Chernov (2008) and Davig and Doh (2009).

Figure 3 shows the persistence of the policy regimes. Combining this figure with the
parameter estimates in Table III reveals the distinct change in policy from passive to active between the mid 1990s and the mid 2000s. In contrast, figures 4, 5 and 6 reveal that the volatility regime changes have been far less drastic than the policy regimes. Finally, Figure 7 plots the estimated exogenous shock processes $\hat{a}_t$, $\hat{g}_t$ and $\hat{e}_t$. The coincidence of the technology shock process $\hat{a}_t$ and business cycles is quite striking in this figure.

4.2 Model Parameters

We next discuss the posterior estimates of the parameters. Table III summarizes the posterior distribution of the parameters based on 20,000 of the MCMC algorithm beyond a burn-in of 5,000. We measure the efficiency of the MCMC sampling in terms of the
Figure 4: Model $\mathcal{M}_{16}$: The posterior probability of technology volatility regimes

These graphs are based on 20,000 simulated draws of the posterior simulation.

| model                        | lnL   | lnML   | n.s.e. | Pr[$\mathcal{M}_m|y$] |
|------------------------------|-------|--------|--------|-----------------------|
| Non-switching ($\mathcal{M}_1$) | 3278.56 | 3239.70 | 0.14   | 0.00                  |
| 2-Regime ($\mathcal{M}_2$)    | 3430.67 | 3420.18 | 0.42   | 0.00                  |
| 4-Regime ($\mathcal{M}_4$)    | 3541.56 | 3536.24 | 0.42   | 0.00                  |
| 16-Regime ($\mathcal{M}_{16}$) | 3591.30 | 3598.47 | 0.41   | 1.00                  |

Table II: Log likelihood (lnL), log marginal likelihood (lnML), numerical standard error (n.s.e) and the posterior probability of each model (Pr[$\mathcal{M}_m|y$]) under the assumption that the prior probability of each model is 1/4.

acceptance rate in the M-H step and the inefficiency factors\(^8\) (Chib (2001)). These values

\(^8\)The inefficiency factors approximate the ratio of the numerical variance of the estimate from the MCMC chain relative to that from hypothetical iid draws. For a given sequence of draws the inefficiency factor is computed as

$$1 + 2 \sum_{l=1}^{L} \rho_k(l)$$

where $\rho_k(l)$ is the autocorrelation at lag $l$ for the $k$th sequence, and $L$ is the value at which the autocorrelation function tapers off (the higher order autocorrelations are also downweighted by a windowing
on average are 26.2% and 28.3, respectively, indicating a well mixing, efficient sampler. Also, the sampler converges quickly to the same region of the parameter space regardless of the starting values. Finally, as one can see in Table III, the posterior densities of the parameters are mostly different from the prior given in Table I. This implies that the data carries information distinct from that contained in the prior distribution.

Two notable features emerge from the table. First, the estimates indicate that the Fed’s response to the macro fundamentals is markedly different across policy regimes. For instance, the posterior mean of the reaction coefficient $\beta$ for the output gap is 0.17 under the first policy regime whereas in the second policy regime it is 1.26. At the same time, the short rate adjustment to inflation gap is also more aggressive under the
second regime (2.02 vs. 0.85). One possible explanation for this is that because inflation has been reasonably stable during the sample period, the Fed’s reaction to output gap became relatively more aggressive, marking the regime shift points.

The second important point to note is that the risk-aversion parameter $\gamma$ has a large posterior mean of 38. This is closely related to the “bond premium puzzle”. Rudebusch and Swanson (2008b) show that many DSGE models with standard macroeconomic parameterizations fail to account for the magnitude of risk premium even with habit formation in the household’s utility function. This is often termed the “bond premium puzzle”. Like in the equity premium puzzle, one possible resolution is a very large value of risk-aversion parameter $\gamma$ to account for the level of the term premium.\footnote{In a standard CRRA preference, high risk aversion (low intertemporal elasticity of substitution) may lead to high real interest rates. However, the average annual real rate implied by our model is 1.884\%, which almost matches the observed annual real interest rates of 2.012\%. On the other hand, in a calibration exercise, Rudebusch and Swanson (2008a) show that Epstein-Zin preference with a relatively small risk aversion parameter can generate a large risk premium in the context of a single}

Figure 6: Model $\mathcal{M}_{16}$: The posterior probability of monetary policy volatility regimes These graphs are based on 20,000 simulated draws of the posterior simulation.
4.3 Changes in the Long Term Bond Risk

In this paper, the benchmark long-term bond is the five-year Treasury note. Its regime-specific risk is computed by the three different measures as discussed in the section 2.10. Figure 8 plots the posterior mean of the term premium for the long-term bond over time. regime DSGE model.
Table III: Posterior distribution for the 16-regime model parameters

This table presents the posterior mean, standard deviation, 90 percent interval and inefficiency factor based on 20,000 posterior draws beyond 5,000 burn-in.

Not surprisingly, this risk measure is strictly increasing in maturity (although it is not reported here). Recall that the time variation of the bond risk is mainly attributed by the change in the reaction coefficients and the shock volatilities. It clearly indicates that the monetary policy regime changes and the technology shock volatilities account for most

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>S.E.</th>
<th>90% credibility interval</th>
<th>Inefficiency factor</th>
<th>Acceptance rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
<td>0.9982</td>
<td>0.0019</td>
<td>[0.9954, 1.0000]</td>
<td>398.54</td>
<td>28.77</td>
</tr>
<tr>
<td>$\phi_a$</td>
<td>0.2875</td>
<td>0.0629</td>
<td>[0.1848, 0.3904]</td>
<td>70.37</td>
<td>33.87</td>
</tr>
<tr>
<td>$\phi_g$</td>
<td>0.9723</td>
<td>0.0040</td>
<td>[0.9657, 0.9788]</td>
<td>79.83</td>
<td>32.27</td>
</tr>
<tr>
<td>$\phi_e$</td>
<td>0.9076</td>
<td>0.0155</td>
<td>[0.8774, 0.9289]</td>
<td>424.92</td>
<td>26.03</td>
</tr>
<tr>
<td>$p_{11}$</td>
<td>0.9364</td>
<td>0.0130</td>
<td>[0.9144, 0.9564]</td>
<td>288.71</td>
<td>27.99</td>
</tr>
<tr>
<td>$p_{22}$</td>
<td>0.9997</td>
<td>0.0014</td>
<td>[0.9981, 1.0000]</td>
<td>115.68</td>
<td>26.66</td>
</tr>
<tr>
<td>$\eta_{11}$</td>
<td>0.9196</td>
<td>0.0767</td>
<td>[0.7761, 0.9938]</td>
<td>59.27</td>
<td>16.93</td>
</tr>
<tr>
<td>$\eta_{22}^a$</td>
<td>0.9342</td>
<td>0.0668</td>
<td>[0.8107, 0.9951]</td>
<td>26.64</td>
<td>18.38</td>
</tr>
<tr>
<td>$\eta_{11}^g$</td>
<td>0.9022</td>
<td>0.0801</td>
<td>[0.7395, 0.9931]</td>
<td>26.75</td>
<td>20.09</td>
</tr>
<tr>
<td>$\eta_{22}^g$</td>
<td>0.9143</td>
<td>0.1021</td>
<td>[0.6755, 0.9955]</td>
<td>84.25</td>
<td>16.12</td>
</tr>
<tr>
<td>$\eta_{11}^e$</td>
<td>0.7574</td>
<td>0.0239</td>
<td>[0.7198, 0.7969]</td>
<td>110.09</td>
<td>32.10</td>
</tr>
<tr>
<td>$\eta_{22}^e$</td>
<td>0.9703</td>
<td>0.0235</td>
<td>[0.9232, 0.9977]</td>
<td>114.17</td>
<td>27.67</td>
</tr>
<tr>
<td>$400 \times \ln R^*$</td>
<td>4.4159</td>
<td>0.1081</td>
<td>[4.2403, 4.5965]</td>
<td>28.33</td>
<td>33.81</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.2580</td>
<td>0.1095</td>
<td>[0.1254, 0.4689]</td>
<td>293.43</td>
<td>30.88</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.8584</td>
<td>0.1837</td>
<td>[0.5480, 1.1855]</td>
<td>253.41</td>
<td>27.46</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>2.0232</td>
<td>0.3403</td>
<td>[1.4564, 2.6164]</td>
<td>269.93</td>
<td>28.82</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.1716</td>
<td>0.3383</td>
<td>[-0.2816, 0.7401]</td>
<td>377.84</td>
<td>28.88</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>1.2646</td>
<td>0.3010</td>
<td>[0.7395, 1.7571]</td>
<td>143.44</td>
<td>29.18</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>38.255</td>
<td>10.104</td>
<td>[23.956, 56.600]</td>
<td>175.43</td>
<td>30.76</td>
</tr>
<tr>
<td>$400 \times \ln a^*$</td>
<td>1.6012</td>
<td>0.1012</td>
<td>[1.4340, 1.7665]</td>
<td>12.91</td>
<td>34.48</td>
</tr>
<tr>
<td>$\ln A_0$</td>
<td>2.3578</td>
<td>0.0804</td>
<td>[2.2151, 2.4791]</td>
<td>105.29</td>
<td>33.49</td>
</tr>
<tr>
<td>$2.0 \times 10^4 \times \sigma_{a1}^2$</td>
<td>1.7258</td>
<td>0.2805</td>
<td>[1.3129, 2.2123]</td>
<td>15.12</td>
<td>33.35</td>
</tr>
<tr>
<td>$2.0 \times 10^5 \times \sigma_{a1}^2$</td>
<td>0.6979</td>
<td>0.1130</td>
<td>[0.5352, 0.8906]</td>
<td>43.13</td>
<td>32.79</td>
</tr>
<tr>
<td>$3.0 \times 10^4 \times \sigma_{a1}^2$</td>
<td>0.6503</td>
<td>0.1128</td>
<td>[0.4892, 0.8590]</td>
<td>41.74</td>
<td>32.56</td>
</tr>
<tr>
<td>$1.0 \times 10^4 \times \sigma_{a2}^2$</td>
<td>0.8595</td>
<td>0.1495</td>
<td>[0.6433, 1.1232]</td>
<td>30.68</td>
<td>33.48</td>
</tr>
<tr>
<td>$1.0 \times 10^5 \times \sigma_{a2}^2$</td>
<td>0.7027</td>
<td>0.2181</td>
<td>[0.5352, 1.0370]</td>
<td>28.25</td>
<td>31.22</td>
</tr>
<tr>
<td>$2.5 \times 10^3 \times \sigma_{a2}^2$</td>
<td>0.5992</td>
<td>0.1539</td>
<td>[0.1561, 1.5003]</td>
<td>127.40</td>
<td>29.17</td>
</tr>
<tr>
<td>$1.4 \times 10^7 \times \sigma_{d2}^2$</td>
<td>0.2724</td>
<td>0.1739</td>
<td>[0.0476, 0.5648]</td>
<td>378.84</td>
<td>31.37</td>
</tr>
<tr>
<td>$5.0 \times 10^6 \times \sigma_{d2}^2$</td>
<td>3.5528</td>
<td>13.3158</td>
<td>[0.3604, 9.7903]</td>
<td>107.07</td>
<td>28.81</td>
</tr>
<tr>
<td>$1.2 \times 10^6 \times \sigma_{d4}^2$</td>
<td>0.4760</td>
<td>0.2417</td>
<td>[0.1762, 0.9640]</td>
<td>121.16</td>
<td>25.78</td>
</tr>
<tr>
<td>$6.0 \times 10^5 \times \sigma_{d5}^2$</td>
<td>10.0532</td>
<td>3.2517</td>
<td>[6.4321, 17.420]</td>
<td>361.72</td>
<td>26.01</td>
</tr>
</tbody>
</table>
of the variations in the term premium. However, these two driving sources differ in the way of influencing the term premium. The changes in the monetary policy affect the risk premium through the sensitivity of the yields (i.e. factor loadings $\bar{B}_{dt}$). Meanwhile, the regime switching shock volatility causes the variations in the term premium by changing the size of the risk. In addition, the average bond risk has diminished over time, which is consistent with the finding of Chib and Kang (2010) and Rudebusch, Sack, and Swanson (2007).

![Figure 8: The term premium and spread of the 5-year bond](image)

**Figure 8: The term premium and spread of the 5-year bond** These graphs are based on 20,000 simulated draws of the posterior simulation. The term premium, the expected excess return and the EH component are computed by 2.53 and 2.55, respectively. These are in annualized percents. The shaded area represents the less active policy regime. NBER recessions are shaded.

On the other hand, Figure 9 presents the result for the decomposition of the term premium of the 5-year bond over time. Interestingly, most of variation of the term premium is explained by the factor risk component. One possible explanation is that sizable factor shocks occur frequently whereas regime shifts happen relatively less frequently. Nevertheless, because the regime shift risk component is consistently positive over time, it should not be neglected.

Figure 10 indicates the regime-dependence of the factor loadings. The yields in the more active regime are more affected by the shocks to the government expenditure and the monetary policy in comparison with those in the less active regime.
Figure 9: Model $M_{16}$: Decomposition of the term premium of the 5-year bond

These graphs are based on 20,000 simulated draws of the posterior simulation. These two components in annualized percents are computed by (2.55).

4.4 Counterfactual Analysis

Because the Markov switching model enables us to estimate the parameters corresponding to each regime, it is straightforward to perform a counterfactual time series experiment. This exercise is very useful to measure the magnitude of the effect of the monetary policy change on the macro-economy and the asset prices.

Figure 11 plots the results for the short rate and the term spread. As seen in the figure, the short rate would have been more volatile and the slope of the yield curve
Figure 10: Model $\mathcal{M}_{16}$: The factor loadings These graphs plot the estimates of the factor loadings on each of the exogenous processes. These graphs are based on on 20,000 simulated draws of the posterior simulation.

steeper without the regime shifts. On the contrary, if the more active regime prevailed over the entire sample period, then the term spread in regime 1 would have been smaller. As a result, the average yield curve differs across regimes due to the policy change. Figure 12 confirms these findings. For instance, the graph on the top panel clearly shows that
Figure 11: Model $M_{16}$: Counterfactual analysis: interest rates The top panel graphs the results for the short rate and the bottom one is for the term spread of 20 quarter bond.

the parameters under the more active regime reproduce a much steeper average yield curve than the actual average during the period corresponding to the less active regime. This implies that a more active regime on average generates a flatter yield curve. A plausible argument here is that a more aggressive response by the monetary authority can potentially mitigate the effect of the (negative) shocks. This in turn leads the risk-averse agents to expect lower volatility in the macro variables. Hence they price bonds with a smaller market price of risk.

However, Figure 13 indicates that inflation and output growth exhibit little difference, regardless of which policy regime prevailed. This suggests that monetary policy regime changes mostly impact the term structure rather than inflation and output growth. This echoes the findings in Gallmeyer, Hollifield, Palomino, and Zin (2008) who also report, within the context of a partial equilibrium model, that the nominal term premium can be highly sensitive to the monetary policy regime.
5 Conclusion

In this paper we propose and estimate a general equilibrium model of the term structure of interest rates with regime changes. The main goal of our work is to examine the term structure of interest rates from a combined macro-finance perspective. Interest in such combined modeling is growing and the general equilibrium model we have described, the solution method we have used, and the econometrics we have employed, can all be adapted to the growing work in this area.

Our empirical results reveal that, in its goal of stabilizing the economy, monetary policy has been more responsive since 2003 with important effects on the dynamics of the term structure. Because in a more active regime, agents potentially anticipate less volatility in the macro variables, bonds are priced with a lower market price of risk. At the same time, the economy is less vulnerable to inflation risk. Consequently, investors require a lower compensation for the risk of holding long term bonds. This leads to a
lower average term premium, resulting in a flatter yield curve. Finally, we find that the volatility of technology shock plays an important role in explaining the time variations of the term premium within the policy regimes.

A Solution

When solving the model we enforce the condition that the stable solution is unique and bounded. Our model solution method relies on the approach of Davig and Leeper (2007). For this, we construct the auxiliary representation of the linearized equilibrium dynamics or the stacked system which is available for any purely forward-looking rational expectations model with regime changes. We begin by defining the state-contingent forecast error as

$$\eta^\pi_{jt+1} = \hat{\pi}_{jt+1} - \mathbb{E}_t(\hat{\pi}_{jt+1})$$ and $$\eta^x_{jt+1} = \hat{x}_{jt+1} - \mathbb{E}_t(\hat{x}_{jt+1}), \ j = 1, 2 \quad (A.1)$$
where \( \hat{y}_{jt+1} \) denotes the value of \( \hat{y}_{t+1} \) conditioned on \( s_{t+1} = j \). Then substituting the conditional expectations in equations (2.37) and (2.38) into the system of equations (2.33)-(2.34) yields the stacked system

\[
A \begin{bmatrix} \hat{\pi}_{1t+1} \\ \hat{\pi}_{2t+1} \\ \hat{x}_{1t+1} \\ \hat{x}_{2t+1} \end{bmatrix} = B \begin{bmatrix} \hat{\pi}_{1t} \\ \hat{\pi}_{2t} \\ \hat{x}_{1t} \\ \hat{x}_{2t} \end{bmatrix} + A \begin{bmatrix} \eta_{1t+1}^\tau \\ \eta_{2t+1}^\tau \\ \eta_{1t+1}^\pi \\ \eta_{2t+1}^\pi \end{bmatrix} + C \bar{f}_t 
\]

where

\[
A = \begin{bmatrix} \delta \otimes P & 0_{2 \times 2} \\ P & \gamma \otimes P \end{bmatrix}, 
\]

\[
B_{11} = I_{m+1}, \quad B_{12} = -\kappa \times I_{m+1}, \quad B_{21} = \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_{m+1}),
\]

\[
B_{22} = \text{diag}(\beta_1 + \gamma, \beta_2 + \gamma, \ldots, \beta_{m+1} + \gamma), 
\]

\[
B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, 
\]

and

\[
C = \begin{bmatrix} 0 & \kappa & 0 \\ 0 & \kappa & 0 \\ -\phi_a & \gamma (\phi_g - 1) & 1 \\ -\phi_a & \gamma (\phi_g - 1) & 1 \end{bmatrix}
\]

Uniqueness and boundedness of the MSV solution are equivalent to the determinacy restriction of the solution space of this stacked system (Davig and Leeper (2007)). In terms of the computational details, this restriction requires that all the generalized eigenvalues of \( A \) and \( B \) lie outside the unit circle.

**B   Bond Prices**

This section provides the details on the derivation of the bond prices in (2.49) and (2.50). We begin by letting \( E^{d_{t+1}} \) denote an expectation conditioned on \( d_{t+1} \). Then the equation (2.43) can be expressed as

\[
P_{d_{t:1}}^{(\tau)} = E^{d_{t+1}} \left[ P_{d_{t},d_{t+1},t}^{(\tau)} \right] \text{ where } P_{d_{t},d_{t+1},t}^{(\tau)} \equiv E \left[ M_{t,t+1} P_{d_{t+1},t+1}^{(\tau-1)} \bar{f}_t, d_t, d_{t+1} \right] 
\]

or

\[
1 = E^{s_{t+1}} \left[ E \left[ M_{t+1,t+1} \bar{h}_{r,t+1} \bar{f}_t, d_t, d_{t+1} \right] \right]
\]
where
\[ h_{\tau,t+1} = \frac{p_{d,t+1}^{(\tau-1)}}{p_{d,t}^{(\tau)}} \]
\[ = \exp \left[ -a_{d,t+1}(\tau - 1) - b_{d,t+1}(\tau - 1) \bar{f}_{t+1} + a_d(\tau) + b_d(\tau) \bar{f}_t \right]. \tag{B.3} \]

If we define
\[ \Theta_{d,t+1} = -a_{d,t+1}(\tau - 1) + a_d(\tau) + (b_d(\tau) - b_{d,t+1}(\tau - 1)) \phi \bar{f}_t \tag{B.4} \]
and \( \Gamma_{\tau,d,t+1} = L_{d,t+1} - b_{d,t+1}(\tau - 1)' \),
then \( M_{t,t+1} h_{\tau,t+1} \) can be rewritten as
\[ \exp \left[ -\ln R^* + \frac{1}{2} \Omega_{d,t+1} L'_{d,t+1} + \lambda_{d,t+1} \bar{f}_t + (L_{d,t+1} - b_{d,t+1}(\tau - 1)) H_{\tau,t+1} + \Theta_{d,t+1} \right] \]
\[ = \exp \left[ -\ln R^* + \frac{1}{2} \Omega_{d,t+1} L'_{d,t+1} + \lambda_{d,t+1} \bar{f}_t + \Gamma_{\tau,d,t+1} H_{\tau,t+1} + \Theta_{d,t+1} \right] \]
\[ = \exp \left[ -\ln R^* + \frac{1}{2} \Omega_{d,t+1} L'_{d,t+1} + \lambda_{d,t+1} \bar{f}_t + \frac{1}{2} \Gamma_{\tau,d,t+1} \Omega_{d,t+1} \Gamma'_{\tau,d,t+1} + \Theta_{d,t+1} \right] \]
\[ \times \exp \left[ -\frac{1}{2} \Gamma_{\tau,d,t+1} \Omega_{d,t+1} \Gamma'_{\tau,d,t+1} \right] \tag{B.5} \]

Since
\[ \mathbb{E} \left[ \exp \left[ -\frac{1}{2} \Gamma_{\tau,d,t+1} \Omega_{d,t+1} \Gamma'_{\tau,d,t+1} \right] \right] = 1 \tag{B.6} \]
the log-approximation gives
\[ \mathbb{E} \left[ M_{t,t+1} h_{\tau,t+1} | \bar{f}_t, d_t, d_{t+1} \right] \tag{B.7} \]
\[ \approx -\ln R^* + \lambda_{d,t+1} \bar{f}_t - L_{d,t+1} \Omega_{d,t+1} b_{d,t+1}(\tau - 1)' + \frac{1}{2} b_{d,t+1}(\tau - 1)' \Omega_{d,t+1} b_{d,t+1}(\tau - 1) + \Theta_{d,t+1} + 1 \]
The next step is integrating out \( d_{t+1} \) for \( d_t = i \) \((i = 1, 2, 3, 4)\). Then the equation (B.1) implies that
\[ 0 = \sum_{j=1}^{d} p_{ij} \left( -\ln R^* + \lambda_{i,j} \bar{f}_t - L_j \Omega_j b_j(\tau - 1)' + \frac{1}{2} b_j(\tau - 1)' \Omega_j b_j(\tau - 1) + \Theta_{i,j} \right) \tag{B.8} \]
Matching the coefficients for constant and \( \bar{f}_t \) completes the derivation of the bond prices.
C Proof of the Term Premium and the Expected Excess Return

This appendix provides the proof of the term premium and the expected excess return in the equation (2.53) and (2.55).

By definition, the term spread of $\tau$-period bond yield is given by

$$r_{\tau}^{(\tau)} - r_{\tau}^{(1)}$$  \hspace{1cm} (C.1)

Let $x_t^{(\tau)} = p_{\tau+1,t+1}^{\tau-1} - p_{\tau,t}^{\tau} - r_{\tau,t}^{(1)}$ denote the excess return. Then we have

$$r_{\tau,t}^{(\tau)} - r_{\tau,t}^{(1)} = \frac{1}{\tau} \sum_{l=0}^{\tau-1} \mathbb{E}_t \left[ r_{\tau,t+1,t+l}^{(1)} - r_{\tau,t}^{(1)} + \frac{1}{\tau} \sum_{i=1}^{\tau-1} \mathbb{E}_t \left[ x_t^{(\tau+1-i)} \right] \right]$$  \hspace{1cm} (C.2)

$$= \frac{1}{\tau} \sum_{l=0}^{\tau-1} \mathbb{E}_t \left[ r_{\tau,t+1,t+l}^{(1)} - r_{\tau,t}^{(1)} + \frac{1}{\tau} \sum_{i=2}^{\tau} \text{exr}_{\tau,t}^{(i)} \right]$$

$$= \frac{1}{\tau} \sum_{l=0}^{\tau-1} \mathbb{E}_t \left[ r_{\tau,t+1,t+l}^{(1)} - r_{\tau,t}^{(1)} + \text{TP}_{\tau,t} \right]$$

where

$$\text{TP}_{\tau,t} = \frac{1}{\tau} \sum_{i=2}^{\tau} \text{exr}_{\tau,t}^{(i)} = \frac{1}{\tau} \left( \text{exr}_{\tau,t}^{(2)} + \text{exr}_{\tau,t}^{(3)} + ... + \text{exr}_{\tau,t}^{(\tau)} \right)$$  \hspace{1cm} (C.3)

Now we prove the equation (2.55). We begin by noting that the risk-neutral pricing formula in the equation (2.43) implies

$$p_{\tau,t}^{(\tau)} = \mathbb{E}_t \left[ m_{t,t+1} + p_{\tau+1,t+1}^{(\tau-1)} \right] + \frac{1}{2} \mathbb{V}_t \left[ m_{t,t+1} + p_{\tau+1,t+1}^{(\tau-1)} \right]$$  \hspace{1cm} (C.4)

This equation holds exactly when the conditional distribution of bond prices and the pricing kernel are jointly log-normal. Then it follows that

$$p_{\tau,t}^{(\tau)} = \mathbb{E}_t \left[ m_{t,t+1} + p_{\tau+1,t+1}^{(\tau-1)} \right] + \frac{1}{2} \mathbb{V}_t \left[ m_{t,t+1} + p_{\tau+1,t+1}^{(\tau-1)} \right]$$

$$= \mathbb{E}_t [m_{t,t+1}] + \mathbb{E}_t \left[ p_{\tau+1,t+1}^{(\tau-1)} \right] + \frac{1}{2} \mathbb{V}_t \left[ m_{t,t+1} + p_{\tau+1,t+1}^{(\tau-1)} \right]$$

$$= p_{\tau,t}^{(1)} - \frac{1}{2} \mathbb{V}_t [m_{t,t+1}] + \mathbb{E}_t \left[ p_{\tau+1,t+1}^{(\tau-1)} \right] + \frac{1}{2} \mathbb{V}_t \left[ m_{t,t+1} + p_{\tau+1,t+1}^{(\tau-1)} \right]$$

since $p_{\tau,t}^{(1)} = \mathbb{E}_t [m_{t,t+1}] + \frac{1}{2} \mathbb{V}_t [m_{t,t+1}]$
and thus
\[
p_{d_{t,t}}^{(\tau)} = p_{d_{t,t}}^{(1)} + \mathbb{E}_t \left[ p_{d_{t+1,t+1}}^{(\tau-1)} \right] + \frac{1}{2} \mathcal{V}_t \left[ p_{d_{t+1,t+1}}^{(\tau-1)} \right] + \text{Cov}_t \left[ m_{t,t+1}, p_{d_{t+1,t+1}}^{(\tau-1)} \right] \tag{C.6}
\]

This implies that
\[
\text{exr}_{d_{t,t}}^{(\tau)} = \left[ \mathbb{E}_t \left[ p_{d_{t+1,t+1}}^{(\tau-1)} \right] - p_{d_{t,t}}^{(\tau)} \right] - \left( p_{d_{t,t}}^{(1)} \right)
= -\text{Cov}_t \left[ m_{t,t+1}, p_{d_{t+1,t+1}}^{(\tau-1)} \right] - \frac{1}{2} \mathcal{V}_t \left[ p_{d_{t+1,t+1}}^{(\tau-1)} \right] \tag{C.7}
\]

The covariance term is compensation for holding long term bond risk associated with the macro structural shocks, and the variance term is the convexity effect (Jensen's inequality).

The remaining is to compute the two terms in the equation (C.7). We begin by expressing the pricing kernel and the log of bond price as
\[
m_{t,t+1} \approx W_{d_{t+1,t+1}} + L_{d_{t+1}} \bar{\varepsilon}_{t+1} \tag{C.8}
\]

\[
p_{d_{t+1,t+1}}^{(\tau-1)} = -a_{d_{+1}} (\tau - 1) - b_{d_{t+1}} (\tau - 1)' (\bar{\phi}_t + \bar{\varepsilon}_{t+1}) \tag{C.9}
= K_{d_{t+1,t}} - b_{d_{t+1}} (\tau - 1)' \bar{\varepsilon}_{t+1}
\]

where
\[
W_{d_{t+1,t+1}} = c_{d_{t+1}} + \lambda_{d_{t+1}} \bar{f}_t \text{ and } K_{d_{t+1,t}} = -a_{d_{t+1}} - b_{d_{t+1}} (\tau - 1)' \bar{f}_t
\]

We first compute the conditional covariance between \( m_{t,t+1} \) and \( p_{d_{t+1,t+1}}^{(\tau-1)} \) using the law of iterative expectation as follows.

\[
\mathbb{E}_t[p_{d_{t+1,t+1}}^{(\tau-1)}] = \mathbb{E}_t \left( \mathbb{E}_t[p_{d_{t+1,t+1}}^{(\tau-1)} | d_{t+1}] \right) = \mathbb{E}_t \left( K_{d_{t+1,t}} \right) = \sum_{j=1}^{d} p_{ij} K_{j,t} \tag{C.10}
\]

\[
\mathbb{E}_t[m_{t,t+1}] = \mathbb{E}_t[W_{d_{t+1,t+1}}] = \sum_{j=1}^{d} p_{ij} W_{i,j,t} \tag{C.11}
\]

\[
\mathbb{E}_t[m_{t,t+1}p_{d_{t+1,t+1}}^{(\tau-1)}] = \mathbb{E}_t[(W_{d_{t+1,t+1}} + L_{d_{t+1}} \bar{\varepsilon}_{t+1})(K_{d_{t+1,t}} - b_{d_{t+1}} \varepsilon_{t+1})] = \sum_{j=1}^{d} p_{ij} (W_{i,j,t} K_{j,t} - b_{j} \varepsilon_{t+1}) \tag{C.12}
\]
- \text{Cov}(m_{t,t+1}, p_{dt+1,t+1}^{(r-1)}) = \mathbb{E}_t[p_{dt+1,t+1}^{(r-1)}\mathbb{E}_t[m_{t,t+1}^{(r-1)}] - \mathbb{E}_t[m_{t,t+1}^{(r-1)}p_{dt+1,t+1}^{(r-1)}] \\
= \left( \sum_{j=1}^{d} p_{ij} K_{j,t} \right) \left( \sum_{j=1}^{d} p_{ij} W_{i,j,t} \right) - \sum_{j=1}^{d} p_{ij} \left( W_{i,j,t} K_{j,t} - b_j \Omega_j L_j \right)
(C.13)

For the conditional variance of $p_{dt+1,t+1}^{(r-1)}$,
\begin{align*}
\mathbb{E}_t \left[ \left( p_{dt+1,t+1}^{(r-1)} \right)^2 \right] &= \mathbb{E}_t \left[ (K_{dt+1,t} - b_{dt+1}(\tau - 1) \varepsilon_{t+1})^2 \right] \\
&= \mathbb{E}_t \left[ K_{dt+1,t}^2 - 2 K_{dt+1,t} b_{dt+1} \varepsilon_{t+1} + b_{dt+1}^2 (\tau - 1) \varepsilon_{t+1}^2 \right] \\
&= \mathbb{E}_t \left[ K_{dt+1,t}^2 + b_{dt+1} (\tau - 1) \Omega_{dt+1} b_{dt+1} (\tau - 1) \right] \\
&= \sum_{j=1}^{d} p_{ij} \left( K_{j,t}^2 + b_j (\tau - 1) \Omega_j b_j (\tau - 1) \right)
(C.14)
\end{align*}
and thus
\begin{align*}
\mathbb{V}_t \left[ p_{dt+1,t+1}^{(r-1)} \right] &= \mathbb{E}_t \left[ \left( p_{dt+1,t+1}^{(r-1)} \right)^2 \right] - \left( \mathbb{E}_t \left[ p_{dt+1,t+1}^{(r-1)} \right] \right)^2 \\
&= \sum_{j=1}^{d} p_{ij} \left( K_{j,t}^2 + b_j (\tau - 1) \Omega_j b_j (\tau - 1) \right) - \left( \sum_{j=1}^{d} p_{ij} K_{j,t} \right)^2
(C.15)
\end{align*}
which completes the proof.

**D MCMC Sampling**

**Step 2 Sampling $\theta$**

Integrating out $F_n$, we sample $\theta$ conditioned on $D_n$ by using the tailored randomized block M-H (TaRB-MH) algorithm. In the $g$th iteration, we have $h_g$ sub-blocks of $\theta$

$$\theta_1, \theta_2, \ldots, \theta_{h_g}$$

The variance of pricing errors $\{\sigma_2^2, \sigma_3^2, \sigma_4^2, \sigma_5^2\}$ and the initial technology level $\ln A_0$ form two fixed blocks ($\theta_{h_g-1}$ and $\theta_{h_g}$), and the others are randomly grouped ($\theta_1, \theta_2, \ldots, \theta_{h_g-2}$). Then the proposal density $q(\theta_i|\theta_{-i}, y)$ for the $i$th block, conditioned
on the most current value of the remaining blocks $\theta_{-i}$, is constructed by a quadratic approximation at the mode of the current target density $\pi(\theta_i|\theta_{-i}, y)$. In our case, we let this proposal density take the form of a student $t$ distribution with 15 degrees of freedom

$$q(\theta_i|\theta_{-i}, y) = St\left(\theta_i|\hat{\theta}_i, V_{\hat{\theta}_i}, 15\right)$$  \hspace{1cm} (D.1)

where

$$\hat{\theta}_i = \arg \max_{\hat{\theta}_i} \ln \{f(y|\theta_i, \theta_{-i}, D_n)\pi(\theta_i)\}$$  \hspace{1cm} (D.2)

and $V_{\hat{\theta}_i} = \left(-\frac{\partial^2 \ln \{f(Y|\theta_i, \theta_{-i}, D_n)\pi(\theta_i)\}}{\partial \theta_i \partial \theta'_i}\right)_{\theta_i=\hat{\theta}_i}^{-1}$.

Because the likelihood function tends to be ill-behaved in these problems, we calculate $\hat{\theta}_i$ using a suitably designed version of the simulated annealing algorithm. In our experience, this stochastic optimization method works better than the standard Newton-Raphson class of deterministic optimizers.

We then generate a proposal value $\theta_i^\dagger$ which, upon satisfying all the constraints, is accepted as the next value in the chain with probability

$$\alpha \left(\theta_i^{(g-1)}, \theta_i^\dagger | \theta_{-i}, y\right)$$  \hspace{1cm} (D.3)

$$= \min \left\{ \frac{f \left( y | \theta_i^\dagger, \theta_{-i}, D_n \right) \pi \left( \theta_i^\dagger \right)}{f \left( y | \theta_i^{(g-1)}, \theta_{-i}, D_n \right) \pi \left( \theta_i^{(g-1)} \right)} \frac{St \left( \theta_i^{(g-1)} | \hat{\theta}_i, V_{\hat{\theta}_i}, 15 \right)}{St \left( \theta_i^\dagger | \hat{\theta}_i, V_{\hat{\theta}_i}, 15 \right)}, 1 \right\}.$$

If $\theta_i^\dagger$ violates any of the constraints in $\mathcal{R}$, it is immediately rejected. The simulation of $\theta$ is complete when all the sub-blocks

$$\pi \left( \theta_1 | \theta_{-1}, y, D_n \right), \pi \left( \theta_2 | \theta_{-2}, y, D_n \right), \ldots, \pi \left( \theta_{h_g} | \theta_{-h_g}, y, D_n \right)$$  \hspace{1cm} (D.4)

are sequentially updated as above.

Now we explain how to calculate $f \left( y | \theta, D_n \right)$ integrating out $F_n$ where $I_t$ is the history of the outcomes up to time $t$. The first step is to solve for the shock process $f_i$ in terms of the observable quantities, $\ln \left( P_t/P_{t-1} \right)$, $\ln Y_t$ and $R_t$ given $\theta$ and $D_n$. 
Since there is no measurement error for inflation, output and the short rate, we have

\[
\begin{bmatrix}
\ln \left( \frac{P_t}{P_{t-1}} \right) \\
\ln Y_t
\end{bmatrix}
= \begin{bmatrix}
\ln \pi^*
\ln x^* + \ln A_t
\end{bmatrix}
+ \begin{bmatrix}
h_n^a(d_t) & h_n^g(d_t) & h_n^e(d_t) \\
h_n^a(d_t) & h_n^g(d_t) & h_n^e(d_t)
\end{bmatrix}
\tilde{f}_t
\] (D.5)

\[
= \begin{bmatrix}
\ln \pi^*
\ln x^* + \ln a^* + \ln A_{t-1}
\end{bmatrix}
+ \begin{bmatrix}
h_n^a(d_t) & h_n^g(d_t) & h_n^e(d_t) \\
1 + h_n^a(d_t) & h_n^g(d_t) & h_n^e(d_t)
\end{bmatrix}
\tilde{f}_t
\] (D.6)

and thus

\[
\begin{bmatrix}
m_t \\
\gamma_t
\end{bmatrix}
= \begin{bmatrix}
\tilde{J}_t \\
\tilde{a}_{dt}(\tau_1)
\end{bmatrix}
+ \begin{bmatrix}
\tilde{H}_{dt} \\
\tilde{b}_{dt}(\tau_1)'
\end{bmatrix}
\tilde{f}_t
\] (D.7)

\[
= \begin{bmatrix}
\tilde{J}_{t-1} \\
\tilde{a}_{dt}(\tau_1)
\end{bmatrix}
+ \begin{bmatrix}
\tilde{H}_{dt} \\
\tilde{b}_{dt}(\tau_1)'
\end{bmatrix}
\tilde{f}_t
\] (D.8)

For \( t = 0 \), the vector of the initial state variables, \( \tilde{f}_0 \) is straightforwardly calculated by \( m_0 \) and \( r_{10} \) conditioned on \( \ln A_0 \) and \( s_0 \) where \( m_0 \) and \( r_{10} \) are observed in the data.

\[
\tilde{f}_0 = \begin{bmatrix}
\tilde{H}_{s_0} \\
\tilde{b}_{s_0}(\tau_1)'
\end{bmatrix}^{-1}
\begin{bmatrix}
m_0 \\
r_{10}
\end{bmatrix}
- \begin{bmatrix}
\tilde{J}_0 \\
\tilde{a}_{s_0}(\tau_1)
\end{bmatrix}
\] (D.9)

For \( t = 1, 2, ..., n - 1 \),

\[
f_t = \begin{bmatrix}
\tilde{f}_t \\
\ln A_t
\end{bmatrix}
\] (D.10)

where

\[
\tilde{f}_t = \begin{bmatrix}
\tilde{H}_{dt} \\
\tilde{b}_{dt}(\tau_1)'
\end{bmatrix}^{-1}
\begin{bmatrix}
m_t \\
r_{1t}
\end{bmatrix}
- \begin{bmatrix}
\tilde{J}_{t-1} \\
\tilde{a}_{dt}(\tau_1)
\end{bmatrix}
\] (D.11)

and

\[
\ln A_t = \ln A_{t-1} + \ln a^* + \begin{bmatrix}
1 & 0 & 0
\end{bmatrix} \tilde{f}_t
\] (D.12)

Notice that conditioned on \( y_t \), \( \tilde{f}_t \) (or \( \hat{a}_t \)) depends on \( \ln A_{t-1} \) and \( d_t \), and \( \ln A_{t-1} = (t - 1) \ln a^* + \sum_{i=1}^{t-1} \hat{a}_i \). Thus \( \ln A_{t-1} \) is affected by the path of regime process up to time \( (t - 1) \). Therefore, in the time updates of \( f_t \) it is very difficult to integrate out the regime path. This is the main reason for sampling \( \theta \) conditioned on \( D_n \).
The second step, which is prediction error decomposition, completes the likelihood function conditioned on $D_n$

$$\ln f (y|\theta, D_n) = \sum_{t=1}^{n} \ln f [y_t|I_{t-1}, d_t, \theta] \quad (D.13)$$

where

$$f [y_t|I_{t-1}, d_t, \theta] = -\frac{1}{2} \eta_{t|t-1}^{d_t} (\Lambda^{d_t})^{-1} \eta_{t|t-1}^{d_t}$$

$$f_{t|t-1} = \mu + G f_{t-1}$$

$$\eta_{t|t-1}^{d_t} = y_t - a_{d_t} - b_{d_t} f_{t|t-1}$$

and $\Lambda^{d_t} = b_{d_t} T_f \Omega_{d_t} T_f' b_{d_t}' + T_y \Sigma_{d_t} T_y'$

**Step 3 Sampling factors**

Conditioned on $\theta$ and $D_n$, the equations (D.9) - (D.12) give $F_n$.

**Step 4 Sampling regimes**

In this step one samples the states from $p[D_n|I_n, \theta]$. This is done according to the method of Chib (1996) by sampling $D_n$ in a single block from the output of one forward and backward pass through the data.

**References**


