

State Space Models with Endogenous Regime Switching[†]

Yoosoon Chang, Junior Maih, and Fei Tan*

[This Version: November 29, 2018]

ABSTRACT

This article studies the estimation of state space models whose parameters are switching endogenously between two regimes, depending on whether an autoregressive latent factor crosses some threshold level. Endogeneity stems from the sustained impacts of transition innovations on the latent factor, absent from which our model reduces to one with exogenous Markov switching. Due to the flexible form of state space representation, this class of models is vastly broad, including classical regression models and the popular dynamic stochastic general equilibrium (DSGE) models as special cases. We develop a computationally efficient filtering algorithm to estimate the non-linear model. Calculations are greatly simplified by appropriate augmentation of the transition equation and exploiting the conditionally linear and Gaussian structure. The algorithm is shown to be accurate in approximating both the likelihood function and filtered state variables. We also apply the filter to estimate a small-scale DSGE model with threshold-type switching in monetary policy rule, and find apparent empirical evidence of endogeneity in the U.S. monetary policy shifts. Overall, our approach provides a greater scope for understanding the complex interaction between regime switching and measured economic behavior.

Keywords: state space model; regime switching; endogenous feedback; filtering; DSGE model.

JEL Classification: C13, C32, E52

[†]This Working Paper should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank.

*Chang: Department of Economics, Indiana University; Maih: Norges Bank and BI Norwegian Business School; Tan: Department of Economics, Chaifetz School of Business, Saint Louis University and Center for Economic Behavior and Decision-Making, Zhejiang University of Finance and Economics. Send correspondence to email: tanf@slu.edu (F. Tan).

1 INTRODUCTION

In time series analysis, there has been a long tradition in modeling the structural changes in dependent data as the outcome of a regime switching process [Hamilton (1988, 1989)]. By introducing an unobserved discrete-state Markov chain governing the regime in place, this class of models affords a tractable framework for the empirical analysis of nonstationarity that is inherent in most economic and financial data. Among further developments of the approach, Kim (1994) made an important extension to the state space representation of dynamic linear models amenable to classical inference, whereas Chib (1996) presented a full Bayesian analysis for finite mixture models based on Gibbs sampling. An introductory exposition and overview of the related literature can be found in the monograph by Kim and Nelson (1999).

Yet despite the popularity and continued success of the Markov switching approach, its maintained hypothesis that the regime evolves exogenously and thereby falls completely apart from the rest of the model is neither realistic nor innocuous in many cases. As argued forcefully in Chang et al. (2017), the presence of endogeneity in regime switching is indeed ubiquitous and, if ignored, may yield substantial bias in the estimates of model parameters. It follows that a more promising approach to modeling occasional but recurrent regime shifts would admit some form of endogenous feedback from the behavior of underlying economic fundamentals to the regime generating process [Diebold et al. (1994), Kim (2004, 2009), Kim et al. (2008), Bazzi et al. (2014), Kang (2014), Kalliovirta et al. (2015), among others].

The purpose of this paper is to introduce a threshold-type endogenous regime switching into dynamic linear models that can be represented as state space forms. This class of models is vastly broad, including classical regression models and the popular dynamic stochastic general equilibrium (DSGE) models as special cases, and thus allows for a greater scope for understanding the complex interaction between regime switching and measured economic behavior.

Following Chang et al. (2017), an essential feature of the model is that the data generating process is switching between two regimes, depending on whether an autoregressive latent factor crosses some threshold level. In our approach, two sources of random innovations jointly drive the latent factor and hence the regime change: [i] the internal innovations from the transition equation that represent the fundamental shocks inside the model; [ii] an external innovation that captures all other shocks left outside of the model. The relative importance of the former source determines the degree of endogeneity in regime changes. The autoregressive nature of the latent factor, on the other hand, makes such endogenous effects long-lasting—a current shock to the transition equation will impact at a decaying rate on all future latent factors. Most importantly, regime switching of this type renders the transition probabilities time-varying as they are all functions of the model’s fundamentals. In the special case where regime shifts are purely driven by the external innovation, our model becomes observationally equivalent to one with conventional Markov switching.

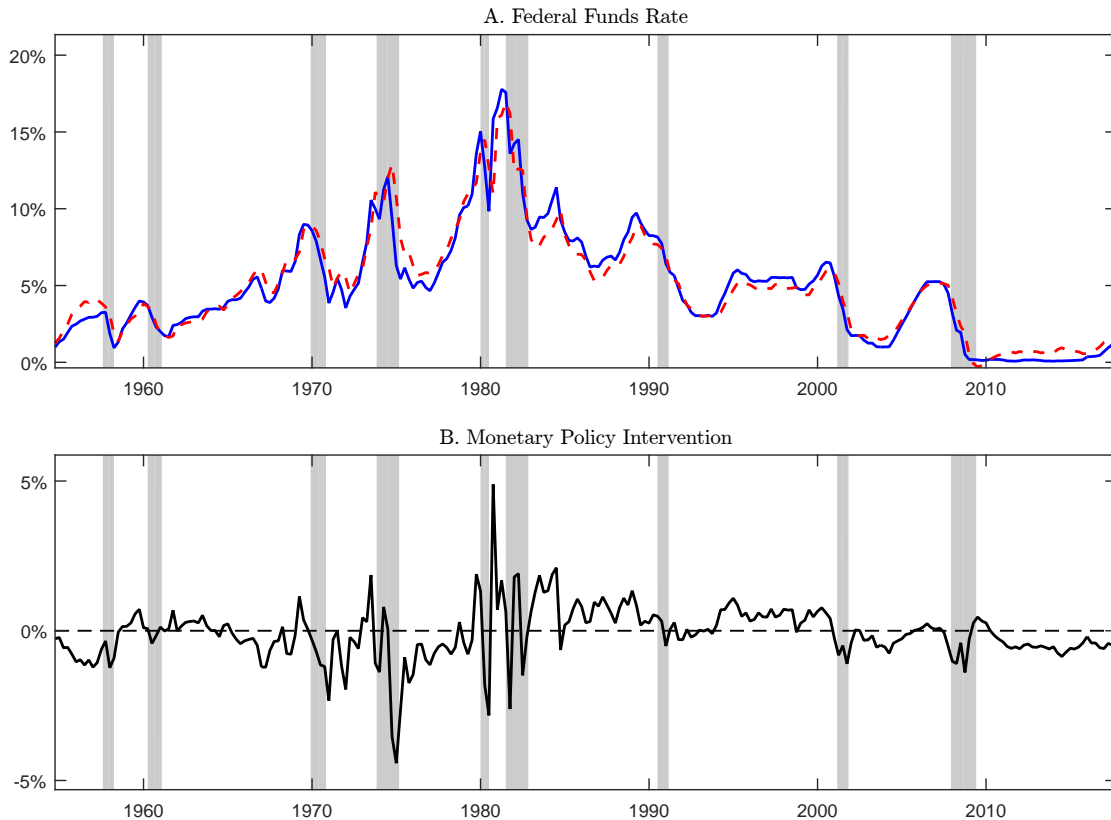


Figure 1: Federal funds rate and monetary policy intervention. Notes: Panel A plots the effective federal funds rate (blue solid line) and that implied by an inertial version of the Taylor (1993) rule (red dashed line), $i_t = \rho i_{t-1} + (1 - \rho)[4 + 1.5(\pi - 2) + 0.5y]$, where i denotes the federal funds rate, π the annual inflation rate, y the percentage deviation of real output from its potential, and $\rho = 0.75$. The black solid line in Panel B depicts their differential. Shaded bars indicate recessions as designated by the National Bureau of Economic Research.

The contributions of this paper are twofold, one methodological and the other substantive. First, we develop an endogenous-switching Kalman filter based on the algorithm of Kim (1994) to estimate the overall nonlinear state space model. Calculations are greatly simplified by appropriate augmentation of the transition equation and exploiting the conditionally linear and Gaussian structure. Unlike simulation-based filters, this avoids sequential Monte Carlo integration and therefore makes our filter computationally efficient. As a useful by-product of running the filter, the estimated autoregressive latent factor can be readily extracted from the augmented system. Simulation experiment indicates that our filtering algorithm is accurate in approximating both the likelihood function and filtered state variables.

Second, ever since the seminal work of Clarida et al. (2000), modeling the time-varying behavior of monetary policy has remained an active research agenda for macroeconomists. Figure 1 displays prima facie evidence of such time variation. Panel A makes clear that the Taylor rule for setting the federal funds rate provides by and large an accurate account of postwar

U.S. monetary policy. Nevertheless, there exist several persistent and sizable discrepancies as shown in Panel B—monetary interventions (i.e., surprise changes in the policy rate) reflecting policy considerations beyond what the Taylor rule mandates. Most evident is the sustained and dramatic loosening of policy under Federal Reserve chairmen Arthur Burns and G. William Miller in the early and late 1970s, followed by several severe tightening of policy to fight the Great Inflation under Paul Volcker in the early 1980s. Economic agents who observe this drastic policy change would shift their beliefs about monetary policy to a more aggressive regime for controlling the inflation. The ensuing well-anchored expectation of stable and low inflation in the near future, in conjunction with the actual monetary stance, ensures price stability thereafter.

While regime switching has emerged nowadays as perhaps the most promising approach to modeling the time variation in monetary policy, scant attention in the literature has been paid to the macroeconomic origins that give rise to monetary policy shifts over time.¹ Our paper takes a first step toward filling in this important gap; we apply the filter to estimate a threshold-type switching version of the prototypical new Keynesian model with the U.S. postwar data, and find that non-policy (i.e., aggregate demand and technology growth) shocks played a predominant role in triggering the historical regime changes in the U.S. monetary policy. To the best of our knowledge, modeling and quantifying such endogenous feedback channel are novel in the literature.

The rest of the paper is planned as follows. Section 2 describes the state space model and filtering algorithm. Section 3 illustrates the filter using two examples, a small state space model with simulated data and a monetary DSGE model with real data. Section 4 concludes. We also employ the following notation. Let $\mathbb{N}(\mu, \Sigma)$ denote the normal distribution with mean vector μ and covariance matrix Σ , $p_{\mathbb{N}}(\cdot|\mu, \Sigma)$ its probability density function, and $\Phi(\cdot)$ the cumulative distribution function of $\mathbb{N}(0, 1)$. In particular, $\mathbb{N}(0_{n \times 1}, \mathbf{I}_n)$ denotes the n -dimensional standard normal distribution. Moreover, $p(\cdot|\cdot)$ and $\mathbb{P}(\cdot|\cdot)$ denote the conditional density and probability functions, respectively. Lastly, $Y_{1:T}$ is a matrix that collects the sample for periods $t = 1, \dots, T$ with row observations y'_t .

2 MODEL AND ALGORITHM

This section introduces the threshold-type endogenous switching setup of Chang et al. (2017), which nests the conventional Markov switching as a special case, into the state space form of a general dynamic linear model. Like any regime switching model, the associated likelihood function depends on all possible histories of the entire regime path. This history-dependent nature creates a tight upper bound on the sample size that any exact recursive filter can comb through

¹As Chris Sims put it in his comment on Davig and Leeper (2006b), “Most people who think that policy changed dramatically and permanently in late 1979 in the United States believe that it did so because inflation appeared to be running out of control, not because an independently evolving switching process happened to call for a change at that date.”

within a reasonable amount of time.² Without appealing to the computationally expensive particle filter, some approximations would seem inevitable. Building on the ‘collapsing’ method of Kim (1994) to truncate the full history-dependence, we develop an endogenous switching version of the Kalman filter to approximate the likelihood function and estimate the unknown parameters as well as the state variables, including the autoregressive latent factor.

2.1 STATE SPACE MODEL Let y_t be a $l \times 1$ vector of observable variables, x_t a $m \times 1$ vector of latent state variables, and z_t a $k \times 1$ vector of predetermined explanatory variables. Consider the following regime-dependent linear state space model

$$y_t = D_{s_t} + Z_{s_t}x_t + F_{s_t}z_t + u_t, \quad u_t \sim \mathbb{N}(0, \Omega_{s_t}) \quad (2.1)$$

$$x_t = C_{s_t} + G_{s_t}x_{t-1} + E_{s_t}z_t + M_{s_t}\epsilon_t, \quad \epsilon_t \sim \mathbb{N}(0, \Sigma_{s_t}) \quad (2.2)$$

where the measurement equation (2.1) links the observable variables to the state variables subject to a $l \times 1$ vector of measurement errors u_t , the transition equation (2.2) describes the evolution of the state variables driven by a $n \times 1$ vector of exogenous innovations ϵ_t , and (u_t, ϵ_t) are mutually and serially uncorrelated at all leads and lags. The coefficient and covariance matrices $(D., Z., F., C., G., E., M., \Omega., \Sigma.)$ are allowed to depend on an index variable $s_t = \mathbb{1}\{w_t \geq \tau\}$ driven by a stationary autoregressive latent factor

$$w_t = \alpha w_{t-1} + v_t, \quad v_t \sim \mathbb{N}(0, 1) \quad (2.3)$$

where $0 \leq \alpha < 1$ controls the persistency of regime changes.³ As a result, the model is switching between regime-0 and regime-1, depending upon whether w_t takes a value below or above the threshold level τ . In what follows, we call w_t the regime factor.

Endogeneity in regime switching is introduced as in Chang et al. (2017), but we allow all current standardized transition innovations to jointly influence the next period regime through its correlation with the innovation to w_{t+1} . That is,

$$\begin{pmatrix} \tilde{\epsilon}_t \\ v_{t+1} \end{pmatrix} \sim \mathbb{N} \left(\begin{pmatrix} 0_{n \times 1} \\ 0 \end{pmatrix}, \begin{pmatrix} I_n & \rho \\ \rho' & 1 \end{pmatrix} \right), \quad \rho' \rho < 1 \quad (2.4)$$

where $\tilde{\epsilon}_t = [\tilde{\epsilon}_{1,t}, \dots, \tilde{\epsilon}_{n,t}]' = \Sigma_{s_t}^{-1/2} \epsilon_t$ and $\rho = [\rho_1, \dots, \rho_n]' = \text{corr}(\tilde{\epsilon}_t, v_{t+1})$ is a vector of correlation parameters that determines the degree of endogeneity in regime changes—as ρ approaches to one in modulus, today’s transition innovations impinge more forcefully on tomorrow’s regime factor.

²As noted by Kim (1994), even with two regimes, there would be over 1000 cases to consider by period $t = 10$.

³In linearized DSGE models, (2.2) represents the regime-dependent solution to the model variables, and the coefficient and covariance matrices in (2.1)–(2.2) become sophisticated functions of some structural parameters as well as the regime index. See Section 3.2 for details.

This type of endogenous impacts is not only sustained due to the autoregressive form of w_t , but also renders the transition probabilities time-varying because they are all functions of $\tilde{\epsilon}_t$ as will be shown subsequently. In the special case where $\tilde{\epsilon}_t$ and v_{t+1} are orthogonal (i.e., $\rho = 0_{n \times 1}$), the transition probabilities become constants and our model reduces to one with conventional Markov switching; in fact, there exists a one-to-one correspondence between the threshold-type switching specified by (α, τ) and the Markov switching specified by two transition probabilities [see Chang et al. (2017), Lemma 2.1].

Since $p(v_{t+1}|\tilde{\epsilon}_t)$ is normal, we can replace v_{t+1} by

$$v_{t+1} = \sum_{k=1}^n \rho_k \tilde{\epsilon}_{k,t} + \sqrt{1 - \sum_{k=1}^n \rho_k^2} \eta_{t+1}, \quad \eta_{t+1} \sim \mathbb{N}(0, 1) \quad (2.5)$$

where $\{\tilde{\epsilon}_{k,t}\}_{k=1}^n$ and the idiosyncratic innovation η_{t+1} are all orthogonal and have unit variance. On the surface, the residual η_{t+1} of projecting v_{t+1} onto $\tilde{\epsilon}_t$ appears to be a vague source of regime change in many economic applications where $\tilde{\epsilon}_t$ is interpreted as structural shocks with clear behavioral meanings. But it indeed captures potential misspecification of the transition equation—ideally one would expect the regime change to be fully driven by $\tilde{\epsilon}_t$ under the ‘true’ model—that leads to systematic disparities between model-implied and actual observables. To the extent that η_{t+1} picks up those missing components beyond what are incorporated in $\tilde{\epsilon}_t$, we may readily call $\tilde{\epsilon}_{k,t}$ and η_t the k -th internal and external innovation, respectively.

To quantify the importance of each source of regime change, iterate forward on (2.3) to obtain $w_{t+h} = \alpha^h w_t + \sum_{j=1}^h \alpha^{h-j} v_{t+j}$ for $h \geq 1$. Combining with (2.5), we have the conditional variance

$$\text{Var}_t(w_{t+h}) = \underbrace{\sum_{k=1}^n \rho_k^2 \sum_{j=1}^h \alpha^{2(h-j)}}_{\text{due to } \tilde{\epsilon}_k} + \underbrace{\left(1 - \sum_{k=1}^n \rho_k^2\right) \sum_{j=1}^h \alpha^{2(h-j)}}_{\text{due to } \eta} = \underbrace{\sum_{j=1}^h \alpha^{2(h-j)}}_{\text{total}}, \quad h \geq 1 \quad (2.6)$$

It follows directly that the percent of the h -step ahead forecast error variance of the regime factor due to the k -th internal (or external) innovation is given by ρ_k^2 (or $1 - \sum_{k=1}^n \rho_k^2$), which is independent of h . Letting $h \rightarrow \infty$, ρ_k^2 (or $1 - \sum_{k=1}^n \rho_k^2$) also measures the percentage contribution to the unconditional variance of the regime factor and hence the extent to which the k -th internal (or external) innovation contributes to the regime changes. For example, using a new Keynesian DSGE model with endogenous regime switching, Section 3.2 presents an empirical calculation on how much of the U.S. monetary policy shifts can be attributed to various internal innovations with distinct behavioral interpretations.

2.2 FILTERING ALGORITHM Estimating the state space model (2.1)–(2.2) entails the dual objectives of likelihood evaluation and filtering, both of which require the calculation of integrals

over the latent variables (i.e., x_t and s_t). While the system is linear in x_t and driven by Gaussian innovations, complication arises from the presence of s_t ; it introduces additional nonlinearities into the overall model structure that invalidate evaluating these integrals via the standard Kalman filter. Nevertheless, approximate analytical integration is still possible from a *marginalization-collapsing* procedure. In the marginalization step, we integrate out the state variables by exploiting the linear and Gaussian structure conditional on the most recent regime history, for which the standard Kalman filter can be applied. In the collapsing step, we approximate an otherwise exponentially growing number of history-dependent filtered distributions by a constant number of mixture Gaussian distributions in each period. This reduction effectively breaks the full history-dependence of the likelihood function and therefore makes the computation highly efficient. We call the resulting algorithm the *endogenous-switching Kalman filter*.

The key to operationalizing the above two-step procedure is an appropriate augmentation of the state space model. To that end, we introduce a dummy vector $d_t = \tilde{\epsilon}_t$ and augment the state vector $\varsigma_t = [x_t', d_t']'$. Accordingly, we rewrite the measurement and transition equations as

$$y_t = \underbrace{D_{s_t} + F_{s_t} z_t}_{\tilde{D}_{s_t}} + \underbrace{\begin{pmatrix} Z_{s_t} & 0_{l \times n} \end{pmatrix}}_{\tilde{Z}_{s_t}} \underbrace{\begin{pmatrix} x_t \\ d_t \end{pmatrix}}_{\varsigma_t} + u_t \quad (2.7)$$

$$\underbrace{\begin{pmatrix} x_t \\ d_t \end{pmatrix}}_{\varsigma_t} = \underbrace{\begin{pmatrix} C_{s_t} + E_{s_t} z_t \\ 0_{n \times 1} \end{pmatrix}}_{\tilde{C}_{s_t}} + \underbrace{\begin{pmatrix} G_{s_t} & 0_{m \times n} \\ 0_{n \times m} & 0_{n \times n} \end{pmatrix}}_{\tilde{G}_{s_t}} \underbrace{\begin{pmatrix} x_{t-1} \\ d_{t-1} \end{pmatrix}}_{\varsigma_{t-1}} + \underbrace{\begin{pmatrix} M_{s_t} \Sigma_{s_t}^{1/2} \\ I_n \end{pmatrix}}_{\tilde{M}_{s_t}} \tilde{\epsilon}_t \quad (2.8)$$

where the dependence of $(\tilde{D}_{s_t}, \tilde{C}_{s_t})$ on z_t has been suppressed for convenience. As will be shown in Algorithm 1 below, our main filtering algorithm, which is based on the augmented state space system (2.7)–(2.8), dates one period back to track regime indices in each recursion. At an exponentially rising computation cost though, one may improve the approximation by tracking even earlier regime history beyond the current and last periods and, in the end, recover the exact likelihood function. For notational ease, let $\mathcal{F}_t = \sigma(\{(z_s, y_s)\}_{s \leq t})$ denote the information available at period t . Define the predictive probability of regime- j at period t , joint with regime- i at period $t-1$, as $p_{t|t-1}^{(i,j)} = \mathbb{P}(s_{t-1} = i, s_t = j | \mathcal{F}_{t-1})$ and the filtered marginal probability of regime- j at period t as $p_{t|t}^j = \mathbb{P}(s_t = j | \mathcal{F}_t)$. Also define a battery of four conditional forecasts of ς_t and their forecast error covariances as

$$\begin{aligned} \varsigma_{t|t-1}^{(i,j)} &= \mathbb{E}[\varsigma_t | s_{t-1} = i, s_t = j, \mathcal{F}_{t-1}] \\ P_{t|t-1}^{(i,j)} &= \mathbb{E}[(\varsigma_t - \varsigma_{t|t-1})(\varsigma_t - \varsigma_{t|t-1})' | s_{t-1} = i, s_t = j, \mathcal{F}_{t-1}] \end{aligned}$$

where $\varsigma_{t|t-1} = \mathbb{E}[\varsigma_t | \mathcal{F}_{t-1}]$. Then the filter can be summarized by the following steps.

Algorithm 1. (Endogenous-Switching Kalman Filter)

1. **Initialization.** For $i = 0, 1$, initialize the conditional mean vector and covariance matrix of ς_0 , $(\varsigma_{0|0}^i, P_{0|0}^i)$, using the invariant distribution under regime- i . Also set $p_{0|0}^0 = \Phi(\tau\sqrt{1-\alpha^2})$ and $p_{0|0}^1 = 1 - p_{0|0}^0$ according to the invariant distribution of w_t (i.e., $\mathbb{N}(0, 1/(1-\alpha^2))$).
2. **Recursion.** For $t = 1, \dots, T$, the filter accepts two sets of triple inputs $\{(\varsigma_{t-1|t-1}^i, P_{t-1|t-1}^i, p_{t-1|t-1}^i)\}_{i=0}^1$, invokes the one-step Kalman filter to calculate the required integrals conditional on four possible mixes of the current and last period regimes, and returns two sets of updated triple outputs $\{(\varsigma_{t|t}^j, P_{t|t}^j, p_{t|t}^j)\}_{j=0}^1$.

- (a) **Forecasting.** First, apply the forecasting step of the Kalman filter for the state variables to obtain

$$\varsigma_{t|t-1}^{(i,j)} = \tilde{C}_j + \tilde{G}_j \varsigma_{t-1|t-1}^i \quad (2.9)$$

$$P_{t|t-1}^{(i,j)} = \tilde{G}_j P_{t-1|t-1}^i \tilde{G}_j' + \tilde{M}_j \tilde{M}_j' \quad (2.10)$$

for $i = 0, 1$ and $j = 0, 1$. Next, define $\lambda_t = \rho' \tilde{\epsilon}_t$ and compute the predictive joint probabilities

$$p_{t|t-1}^{(0,0)} = \int_{-\infty}^{\infty} \mathbb{P}(s_{t-1} = 0, s_t = 0 | \lambda_{t-1}, \mathcal{F}_{t-1}) p(\lambda_{t-1} | \mathcal{F}_{t-1}) d\lambda_{t-1} \quad (2.11)$$

and $p_{t|t-1}^{(0,1)} = p_{t-1|t-1}^0 - p_{t|t-1}^{(0,0)}$. To evaluate the integral in (2.11), note that

$$\begin{aligned} & \mathbb{P}(s_{t-1} = 0, s_t = 0 | \lambda_{t-1}, \mathcal{F}_{t-1}) \\ &= \mathbb{P}(s_t = 0 | s_{t-1} = 0, \lambda_{t-1}, \mathcal{F}_{t-1}) \mathbb{P}(s_{t-1} = 0 | \lambda_{t-1}, \mathcal{F}_{t-1}) \\ &= \mathbb{P}(s_t = 0 | s_{t-1} = 0, \lambda_{t-1}) \mathbb{P}(s_{t-1} = 0 | \mathcal{F}_{t-1}) \\ &= \frac{\int_{-\infty}^{\tau\sqrt{1-\alpha^2}} \Phi((\tau - \alpha x / \sqrt{1-\alpha^2} - \lambda_{t-1}) / \sqrt{1-\rho'\rho}) p_{\mathbb{N}}(x|0, 1) dx}{\Phi(\tau\sqrt{1-\alpha^2})} p_{t-1|t-1}^0 \end{aligned}$$

where the second equality holds since $p(w_t | w_{t-1}, \lambda_{t-1}, \mathcal{F}_{t-1}) = p(w_t | w_{t-1}, \lambda_{t-1})$ and w_{t-1} is independent of λ_{t-1} . Clearly, the transition probability $\mathbb{P}(s_t = 0 | s_{t-1} = 0, \lambda_{t-1})$ depends on the value of λ_{t-1} and hence $\tilde{\epsilon}_{t-1}$ but becomes a constant when $\rho = 0_{n \times 1}$. Moreover, we approximate

$$p(\lambda_{t-1} | \mathcal{F}_{t-1}) \approx p_{\mathbb{N}}(\lambda_{t-1} | \rho' \varsigma_{d,t-1|t-1}^0, \rho' P_{d,t-1|t-1}^0 \rho)$$

where $(\varsigma_{d,t-1|t-1}^0, P_{d,t-1|t-1}^0)$ can be extracted from $(\varsigma_{t-1|t-1}^0, P_{t-1|t-1}^0)$ corresponding to d_{t-1} . To the extent that the filtered distribution of $\tilde{\epsilon}_{t-1}$ serves as an essential input into the approximation of $p(\lambda_{t-1}|\mathcal{F}_{t-1})$, this justifies augmenting the state space system by the dummy vector $d = \tilde{\epsilon}$. Taken together, (2.11) can be rewritten as

$$p_{t|t-1}^{(0,0)} = \frac{p_{t-1|t-1}^0}{\Phi(\tau\sqrt{1-\alpha^2})} \int_{-\infty}^{\frac{\tau}{\sqrt{1-\rho'\rho}}} \int_{-\infty}^{\tau\sqrt{1-\alpha^2}} p_{\mathbb{N}}(x, y|\mu_0, \Sigma_0) dx dy \quad (2.12)$$

where

$$\mu_0 = \begin{pmatrix} 0 \\ \frac{\rho'\varsigma_{d,t-1|t-1}^0}{\sqrt{1-\rho'\rho}} \end{pmatrix}, \quad \Sigma_0 = \begin{pmatrix} 1 & \frac{\alpha}{\sqrt{1-\rho'\rho}\sqrt{1-\alpha^2}} \\ \frac{\alpha}{\sqrt{1-\rho'\rho}\sqrt{1-\alpha^2}} & 1 + \frac{\alpha^2}{(1-\rho'\rho)(1-\alpha^2)} + \frac{\rho'P_{d,t-1|t-1}^0\rho}{1-\rho'\rho} \end{pmatrix}$$

Similarly, we can compute

$$p_{t|t-1}^{(1,0)} = \frac{p_{t-1|t-1}^1}{1 - \Phi(\tau\sqrt{1-\alpha^2})} \int_{-\infty}^{\frac{\tau}{\sqrt{1-\rho'\rho}}} \int_{-\infty}^{-\tau\sqrt{1-\alpha^2}} p_{\mathbb{N}}(x, y|\mu_1, \Sigma_1) dx dy \quad (2.13)$$

and $p_{t|t-1}^{(1,1)} = p_{t-1|t-1}^1 - p_{t|t-1}^{(1,0)}$, where

$$\mu_1 = \begin{pmatrix} 0 \\ \frac{\rho'\varsigma_{d,t-1|t-1}^1}{\sqrt{1-\rho'\rho}} \end{pmatrix}, \quad \Sigma_1 = \begin{pmatrix} 1 & \frac{-\alpha}{\sqrt{1-\rho'\rho}\sqrt{1-\alpha^2}} \\ \frac{-\alpha}{\sqrt{1-\rho'\rho}\sqrt{1-\alpha^2}} & 1 + \frac{\alpha^2}{(1-\rho'\rho)(1-\alpha^2)} + \frac{\rho'P_{d,t-1|t-1}^1\rho}{1-\rho'\rho} \end{pmatrix}$$

Finally, the integrals in (2.12)–(2.13) can be easily evaluated using the cumulative bivariate normal distribution function. See the Online Appendix for derivation details.

- (b) **Likelihood evaluation.** Apply the forecasting step of the Kalman filter for the observable variables to obtain

$$y_{t|t-1}^{(i,j)} = \tilde{D}_j + \tilde{Z}_j \varsigma_{t|t-1}^{(i,j)} \quad (2.14)$$

$$F_{t|t-1}^{(i,j)} = \tilde{Z}_j P_{t|t-1}^{(i,j)} \tilde{Z}_j' + \Omega_j \quad (2.15)$$

for $i = 0, 1$ and $j = 0, 1$. Then the period- t likelihood contribution can be computed as

$$p(y_t|\mathcal{F}_{t-1}) = \sum_{j=0}^1 \sum_{i=0}^1 p_{\mathbb{N}}(y_t|y_{t|t-1}^{(i,j)}, F_{t|t-1}^{(i,j)}) p_{t|t-1}^{(i,j)} \quad (2.16)$$

(c) **Updating.** First, apply the Bayes formula to update

$$p_{t|t}^{(i,j)} = \frac{p_{\mathbb{N}}(y_t | y_{t|t-1}^{(i,j)}, F_{t|t-1}^{(i,j)}) p_{t|t-1}^{(i,j)}}{p(y_t | \mathcal{F}_{t-1})} \quad (2.17)$$

and calculate $p_{t|t}^j = \sum_{i=0}^1 p_{t|t}^{(i,j)}$. Next, apply the filtering step of the Kalman filter for the state variables to obtain

$$\varsigma_{t|t}^{(i,j)} = \varsigma_{t|t-1}^{(i,j)} + P_{t|t-1}^{(i,j)} \tilde{Z}'_j (F_{t|t-1}^{(i,j)})^{-1} (y_t - y_{t|t-1}^{(i,j)}) \quad (2.18)$$

$$P_{t|t}^{(i,j)} = P_{t|t-1}^{(i,j)} - P_{t|t-1}^{(i,j)} \tilde{Z}'_j (F_{t|t-1}^{(i,j)})^{-1} \tilde{Z}_j P_{t|t-1}^{(i,j)} \quad (2.19)$$

for $i = 0, 1$ and $j = 0, 1$. To avoid a twofold increment in the number of cases to consider for the next period, collapse $(\varsigma_{t|t}^{(i,j)}, P_{t|t}^{(i,j)})$ into⁴

$$\varsigma_{t|t}^j = \sum_{i=0}^1 \frac{p_{t|t}^{(i,j)}}{p_{t|t}^j} \varsigma_{t|t}^{(i,j)}, \quad P_{t|t}^j = \sum_{i=0}^1 \frac{p_{t|t}^{(i,j)}}{p_{t|t}^j} \left[P_{t|t}^{(i,j)} + (\varsigma_{t|t}^j - \varsigma_{t|t}^{(i,j)}) (\varsigma_{t|t}^j - \varsigma_{t|t}^{(i,j)})' \right] \quad (2.20)$$

Further collapsing $(\varsigma_{t|t}^j, P_{t|t}^j)$ into

$$\varsigma_{t|t} = \sum_{j=0}^1 p_{t|t}^j \varsigma_{t|t}^j, \quad P_{t|t} = \sum_{j=0}^1 p_{t|t}^j \left[P_{t|t}^j + (\varsigma_{t|t} - \varsigma_{t|t}^j) (\varsigma_{t|t} - \varsigma_{t|t}^j)' \right] \quad (2.21)$$

gives the filtered state variables.

3. Aggregation. The likelihood function is given by $p(Y_{1:T}) = \prod_{t=1}^T p(y_t | \mathcal{F}_{t-1})$.

Several remarks about this filtering algorithm are in order. First, while the general structure resembles that of the mixture Kalman filter in Chen and Liu (2000), our filter requires no sequential Monte Carlo integration and is thus computationally efficient. By analytically integrating out x_t and s_t , it also greatly simplifies estimating the model via classical or Bayesian approaches that would otherwise require a stochastic version of the expectation-maximization algorithm or Gibbs sampling, respectively [Wei and Tanner (1990), Tanner and Wong (1987)]. Second, in line with Kim (1994), the collapsing step (2.20) involves an approximation—its input $\varsigma_{t|t}^{(i,j)}$ does not calculate the conditional expectation $\mathbb{E}[\varsigma_t | s_{t-1} = i, s_t = j, \mathcal{F}_t]$ exactly since $p(\varsigma_t | s_{t-1} = i, s_t = j, \mathcal{F}_t)$ amounts to a mixture of Gaussian distributions for $t > 2$. Consequently, the period- t likelihood $p(y_t | \mathcal{F}_{t-1})$ and filtered states $\varsigma_{t|t}$ only approximately calculate their true values, whose accuracy will be examined in the next section. Third, an estimated autoregressive regime factor $w_{t|t}$ can be

⁴If $p_{t|t}^j = 0$, the conditional probability $p_{t|t}^{(i,j)} / p_{t|t}^j = \mathbb{P}(s_{t-1} = i | s_t = j, \mathcal{F}_t)$ in (2.20) is not well defined. In this case, we set $(\varsigma_{t|t}^j, P_{t|t}^j) = (\varsigma_{t|t}^{1-j}, P_{t|t}^{1-j})$.

easily extracted as a useful by-product of running the filter. Using the stored values of $\varsigma_{d,t-1|t-1}^i$, $P_{d,t-1|t-1}^i$, $p_{t-1|t-1}^i$, $p_{\mathbb{N}}(y_t|y_{t|t-1}^{(i,j)}, F_{t|t-1}^{(i,j)})$, and $p(y_t|\mathcal{F}_{t-1})$, it is straightforward to evaluate

$$\begin{aligned} p(w_t|\mathcal{F}_t) &= \sum_{i=0}^1 \int_{-\infty}^{\infty} p(w_t, s_{t-1} = i, \lambda_{t-1}|\mathcal{F}_t) d\lambda_{t-1} \\ &= \sum_{i=0}^1 \frac{p(y_t|w_t, s_{t-1} = i, \mathcal{F}_{t-1}) p_{t-1|t-1}^i}{p(y_t|\mathcal{F}_{t-1})} \int_{-\infty}^{\infty} p(w_t|s_{t-1} = i, \lambda_{t-1}) p(\lambda_{t-1}|\mathcal{F}_{t-1}) d\lambda_{t-1} \end{aligned} \quad (2.22)$$

where $p(y_t|w_t, s_{t-1} = i, \mathcal{F}_{t-1}) = p_{\mathbb{N}}(y_t|y_{t|t-1}^{(i,j)}, F_{t|t-1}^{(i,j)})$ for $j = \mathbb{1}\{w_t \geq \tau\}$ and

$$p(w_t|s_{t-1} = i, \lambda_{t-1}) = \begin{cases} \frac{\Phi\left(\sqrt{\frac{1-\rho'\rho+\alpha^2\rho'\rho}{1-\rho'\rho}}\left(\tau - \frac{\alpha(w_t-\lambda_{t-1})}{1-\rho'\rho+\alpha^2\rho'\rho}\right)\right)}{\Phi(\tau\sqrt{1-\alpha^2})} p_{\mathbb{N}}\left(w_t|\lambda_{t-1}, \frac{1-\rho'\rho+\alpha^2\rho'\rho}{1-\alpha^2}\right), & i = 0 \\ \frac{1-\Phi\left(\sqrt{\frac{1-\rho'\rho+\alpha^2\rho'\rho}{1-\rho'\rho}}\left(\tau - \frac{\alpha(w_t-\lambda_{t-1})}{1-\rho'\rho+\alpha^2\rho'\rho}\right)\right)}{1-\Phi(\tau\sqrt{1-\alpha^2})} p_{\mathbb{N}}\left(w_t|\lambda_{t-1}, \frac{1-\rho'\rho+\alpha^2\rho'\rho}{1-\alpha^2}\right), & i = 1 \end{cases} \quad (2.23)$$

is derived in Corollary 3.3 of Chang et al. (2017). Moreover, we again approximate $p(\lambda_{t-1}|\mathcal{F}_{t-1})$ by $p_{\mathbb{N}}(\lambda_{t-1}|\rho'\varsigma_{d,t-1|t-1}^i, \rho'P_{d,t-1|t-1}^i\rho)$ or simply the Dirac measure $\delta_{\rho'\varsigma_{d,t-1|t-1}^i}(\lambda_{t-1})$ for $s_{t-1} = i$. Then the filtered autoregressive regime factor can be computed as

$$w_t|t = \int_{-\infty}^{\infty} w_t p(w_t|\mathcal{F}_t) dw_t \approx \sum_{i=1}^N w_t^i \hat{p}(w_t^i|\mathcal{F}_t)$$

where we approximate $p(w_t|\mathcal{F}_t)$ by a discrete density function $\hat{p}(w_t|\mathcal{F}_t)$ defined on a swarm of grid points $\{w_t^i\}_{i=1}^N$ with their corresponding weights $\hat{p}(w_t^i|\mathcal{F}_t) = p(w_t^i|\mathcal{F}_t)/\sum_{i=1}^N p(w_t^i|\mathcal{F}_t)$. Lastly, given the filtered density (2.22), we also develop an algorithm to approximate the smoothed density $p(w_t|\mathcal{F}_T)$ and relegate the details of this algorithm to the Online Appendix.

3 EXAMPLES

We apply the endogenous-switching Kalman filter to two examples. First, we consider a small state space model with simulated observations that serves as a test bench for assessing the accuracy of our filter in approximating the likelihood function and filtered state variables. Second, we embed the filter into a Metropolis-Hastings posterior sampler and estimate a prototypical new Keynesian DSGE model with real data.

3.1 SIMULATION MODEL Consider the following small state space model that resembles the general structure of (2.1)–(2.2) for reduced-form DSGE models

$$y_t = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} + u_t, \quad u_t \sim \mathbb{N}(0, \omega^2) \quad (3.1)$$

$$\begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} = \begin{pmatrix} g_{11}(s_t) & g_{12}(s_t) \\ 0 & g_{22}(s_t) \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \epsilon_t, \quad \epsilon_t \sim \mathbb{N}(0, \sigma^2) \quad (3.2)$$

where three parameters in the transition equation, (g_{11}, g_{12}, g_{22}) , are allowed to switch between regime-0 and regime-1. We simulate 100 observations of y_t by setting $(g_{11}, g_{12}, g_{22}) = (0.8, 0.2, 0.1)$ if $s_t = 1$ and $(0.2, 0.8, 0.9)$ otherwise, $(\omega, \sigma) = (0.2, 0.5)$, and $(\alpha, \tau, \rho) = (0.7, -0.5, 0.9)$.⁵ Overall the model underwent frequent regime changes, on average about once every 3 periods, which poses a potential challenge to our filter in delivering a satisfactory approximation to the likelihood function. To evaluate its performance, we estimate the exact likelihood and filtered states from a regime switching version of the bootstrap particle filter in Gordon et al. (1993). It numerically integrates out $\{(x_t, w_t)\}_{t=1}^T$ using a discrete set of particles simulated from the transition equations in x_t and w_t , respectively.⁶ Our particle filtering algorithm, which is based on the original state space system (2.1)–(2.2), can be implemented through the following steps.

Algorithm 2. (Endogenous-Switching Particle Filter)

1. **Initialization.** For each regime $j = 0, 1$, initialize $(x_{0|0}^j, P_{0|0}^j, p_{0|0}^j)$ as in Algorithm 1 but based on the original state space form (2.1)–(2.2). For $i = 1, \dots, N$, draw an initial swarm of particles x_0^i from the mixture Gaussian distribution $p_{0|0}^0 \cdot \mathbb{N}(x_{0|0}^0, P_{0|0}^0) + p_{0|0}^1 \cdot \mathbb{N}(x_{0|0}^1, P_{0|0}^1)$, $\tilde{\epsilon}_0^i \sim \mathbb{N}(0_{n \times 1}, I_n)$, and $w_0^i \sim \mathbb{N}(0, 1/(1 - \alpha^2))$. Also, set the corresponding weight $W_0^i = 1/N$.
2. **Recursion.** For $t = 1, \dots, T$:

- (a) **Propagation.** Draw particles $\{w_t^i\}_{i=1}^N$ from

$$w_t^i = \alpha w_{t-1}^i + \sum_{k=1}^n \rho_k \tilde{\epsilon}_{k,t-1}^i + \sqrt{1 - \sum_{k=1}^n \rho_k^2} \eta_t^i, \quad \eta_t^i \sim \mathbb{N}(0, 1) \quad (3.3)$$

⁵A small measurement error is included so that the data density, conditioned on the states, remains non-degenerate in the particle filter.

⁶A complete tutorial on basic and advanced particle filtering methods can be found in Doucet and Johansen (2011). See also DeJong and Dave (2007) and Herbst and Schorfheide (2015) for textbook treatments of the particle filter in DSGE applications.

and compute $s_t^i = \mathbb{1}\{w_t^i \geq \tau\}$. Then draw particles $\{x_t^i\}_{i=1}^N$ from

$$x_t^i = C_{s_t^i} + G_{s_t^i} x_{t-1}^i + E_{s_t^i} z_t + M_{s_t^i} \Sigma_{s_t^i}^{1/2} \tilde{\epsilon}_t^i, \quad \tilde{\epsilon}_t^i \sim \mathbb{N}(0_{n \times 1}, \mathbf{I}_n) \quad (3.4)$$

(b) **Likelihood.** The period- t likelihood integral can be approximated as

$$\hat{p}(y_t | \mathcal{F}_{t-1}) = \sum_{i=1}^N W_{t-1}^i p(y_t | s_t^i, x_t^i) \quad (3.5)$$

where $p(y_t | s_t^i, x_t^i) = p_{\mathbb{N}}(y_t | D_{s_t^i} + Z_{s_t^i} x_t^i + F_{s_t^i} z_t, \Omega_{s_t^i})$.

(c) **Filtering.** Update the weights according to the Bayes Theorem

$$W_t^i = \frac{p(y_t | s_t^i, x_t^i) W_{t-1}^i}{\hat{p}(y_t | \mathcal{F}_{t-1})} \quad (3.6)$$

(d) **Resampling.** Define the effective sample size $\text{ESS} = 1 / \sum_{i=1}^N (W_t^i)^2$. If $\text{ESS} < N/2$, resample the particles $\{(x_t^i, \tilde{\epsilon}_t^i, w_t^i)\}_{i=1}^N$ with the systematic resampling scheme and set $W_t^i = 1/N$ for $i = 1, \dots, N$. Now the particle system $\{(x_t^i, \tilde{\epsilon}_t^i, w_t^i, W_t^i)\}_{i=1}^N$ approximates any filtered value by

$$\mathbb{E}[h(x_t, \tilde{\epsilon}_t, w_t) | \mathcal{F}_t] \approx \sum_{i=1}^N h(x_t^i, \tilde{\epsilon}_t^i, w_t^i) W_t^i \quad (3.7)$$

where h denotes some transformation function of interest.

3. **Aggregation.** The likelihood function is given by $\hat{p}(Y_{1:T}) = \prod_{t=1}^T \hat{p}(y_t | \mathcal{F}_{t-1})$.

Practically it turns out that $N = 100,000$ particles are more than sufficient to accurately approximate the density $p(Y_{1:T})$ dictated by the simulation model, which we take as ‘exact’. Evaluating at the true parameter values, Figure 2 depicts the filtered state variables (Panels A–C) and log likelihood decomposition by period (Panel D) computed from the endogenous-switching Kalman filter and particle filter. A visual comparison suggests that our filter performs fairly well as it delivers almost indistinguishable approximations to these quantities from their true values. That accuracy extends to a wide range of the parameter space as can be seen in Figure 3, though the particle filter has some difficulty evaluating the log likelihood functions of certain parameters near the boundaries.⁷ In addition, the log likelihood function of each

⁷This is because the particle generating distributions, $p(x_t^i | x_{t-1}^i)$ and $p(w_t^i | w_{t-1}^i)$, are simply based on the transition equations in x_t and w_t , which ignore the information in light of y_t . Refinements of the bootstrap particle filter abound in the literature. For example, one efficient choice is to generate particles from the filtered state distributions computed from our endogenous-switching Kalman filter and reweigh these particles through an importance sampling step, the so-called adaptation of the particle filter.

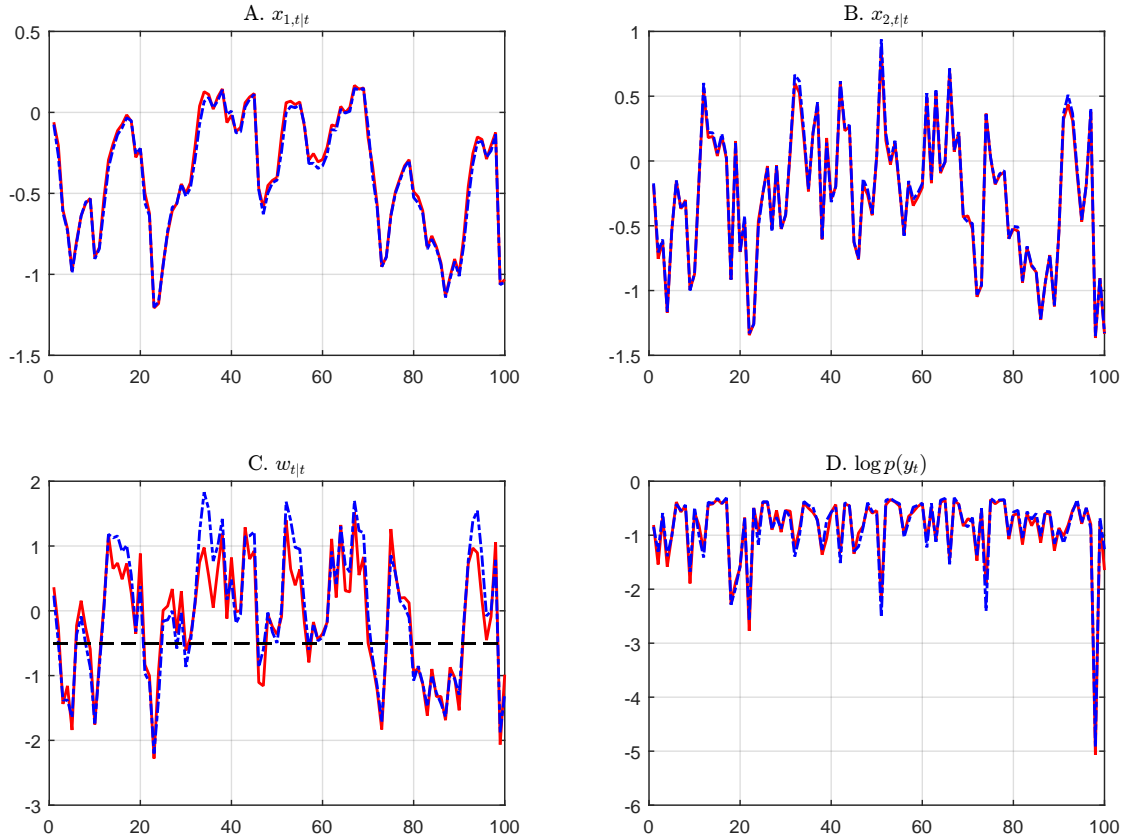


Figure 2: Filtered state variables and log likelihood decomposition at the true parameter values. Notes: The regime change occurs 31 times out of 100 periods. Panels A–C plot the filtered state variables computed from endogenous-switching Kalman filter (red solid line) and particle filter (blue dash-dotted line). The black dashed line in Panel C delineates the threshold level. Panel D plots the log likelihood contributions in each period evaluated with both filters (in aggregate -85.65 for Algorithm 1, -85.02 for Algorithm 2, and -92.15 for Algorithm 1 with $\rho = 0$).

individual parameter peaks in the immediate vicinity of its true value. Taken together, our filter ensures the overall likelihood surface is well preserved.

3.2 EMPIRICAL DSGE MODEL We consider the small-scale new Keynesian DSGE model presented in An and Schorfheide (2007), whose essential elements include: a representative household and a continuum of monopolistically competitive firms; each firm produces a differentiated good and faces nominal rigidity in terms of quadratic price adjustment cost; a cashless economy with one-period nominal bonds; a monetary authority that controls nominal interest rate as well as a fiscal authority that passively adjusts lump-sum taxes to ensure its budgetary solvency; a labor-augmenting technology that induces a stochastic trend in consumption and output. In the fixed regime benchmark, the model’s equilibrium conditions in terms of the detrended variables can be summarized as follows.

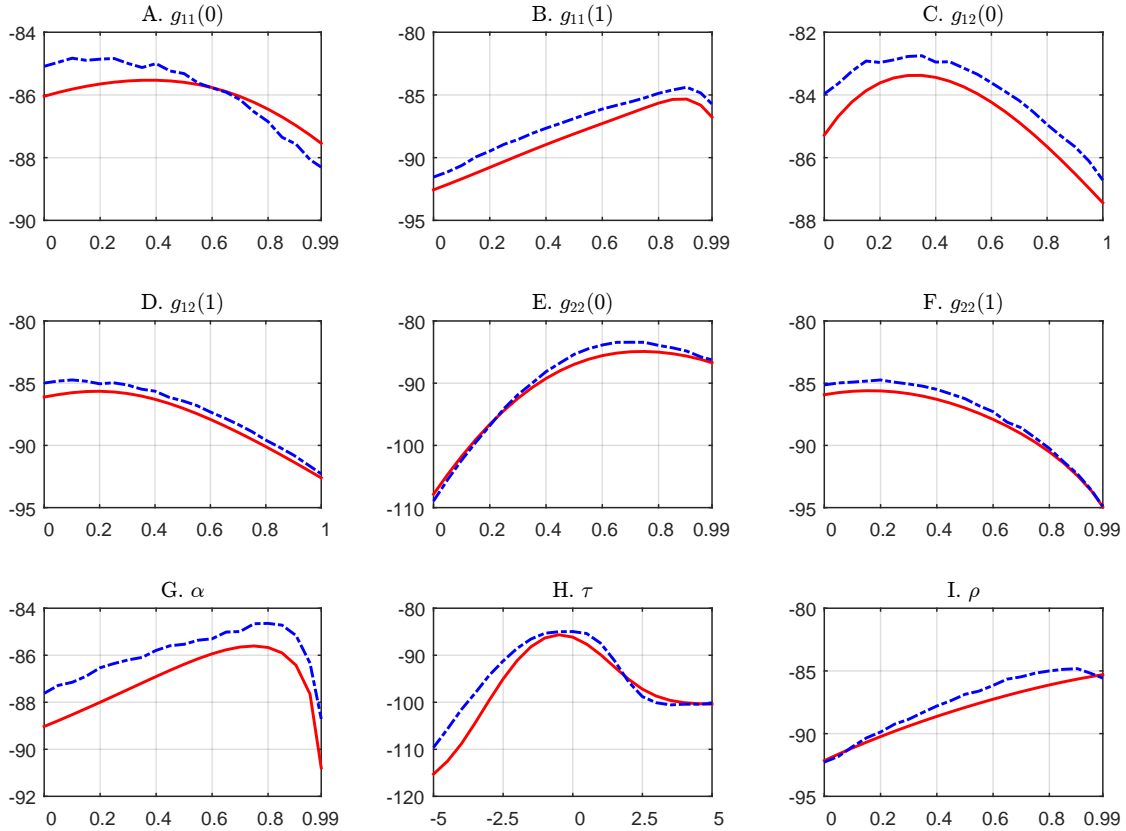


Figure 3: Profile of log likelihood functions. Notes: Each panel plots the log likelihood function of an individual parameter computed from Algorithm 1 (red solid line) and Algorithm 2 (blue dash-dotted line) while holding others at their true values.

First, the household's optimizing behavior implies

$$1 = \beta \mathbb{E}_t \left[\left(\frac{c_{t+1}}{c_t} \right)^{-\tau_c} \frac{R_t}{\gamma z_{t+1} \pi_{t+1}} \right] \quad (3.8)$$

where $0 < \beta < 1$ is the discount factor, $\tau_c > 0$ the coefficient of relative risk aversion, c_t the detrended consumption, R_t the nominal interest rate, π_t the inflation between periods $t - 1$ and t , z_t an exogenous shock to the labor-augmenting technology that grows on average at the rate γ , and \mathbb{E}_t represents conditional expectation given information available at time t . The firm's optimal price-setting behavior yields

$$1 = \frac{1 - c_t^{\tau_c}}{\nu} + \phi(\pi_t - \pi) \left[\left(1 - \frac{1}{2\nu} \right) \pi_t + \frac{\pi}{2\nu} \right] - \phi \beta \mathbb{E}_t \left[\left(\frac{c_{t+1}}{c_t} \right)^{-\tau_c} \frac{y_{t+1}}{y_t} (\pi_{t+1} - \pi) \pi_{t+1} \right] \quad (3.9)$$

where $1/\nu > 1$ is the elasticity of demand for each differentiated good, ϕ the degree of price stickiness that relates to the slope of the so-called new Keynesian Phillips curve κ via $\phi =$

$\tau_c(1 - \nu)/(\nu\pi^2\kappa)$, π the steady state inflation, and y_t the detrended output. Second, the goods market clearing condition is given by

$$y_t = c_t + \left(1 - \frac{1}{g_t}\right)y_t + \frac{\phi}{2}(\pi_t - \pi)^2 y_t \quad (3.10)$$

where g_t is an exogenous government spending shock. Third, the monetary authority follows an interest rate feedback rule that reacts to deviations of inflation from its steady state and output from its potential value

$$R_t = R_t^{*1-\rho_R} R_{t-1}^{\rho_R} e^{\epsilon_{R,t}}, \quad R_t^* = r\pi \left(\frac{\pi_t}{\pi}\right)^{\psi_\pi} \left(\frac{y_t}{y_t^*}\right)^{\psi_y} \quad (3.11)$$

where $0 \leq \rho_R < 1$ is the degree of interest rate smoothing, r the steady state real interest rate, $\psi_\pi > 0$ and $\psi_y > 0$ the policy rate responsive coefficients, $y_t^* = (1 - \nu)^{1/\tau_c} g_t$ the detrended potential output, and $\epsilon_{R,t}$ an exogenous policy shock. Finally, both $\ln g_t$ and $\ln z_t$ evolve as autoregressive processes

$$\ln g_t = (1 - \rho_g) \ln g + \rho_g \ln g_{t-1} + \epsilon_{g,t}, \quad \ln z_t = \rho_z \ln z_{t-1} + \epsilon_{z,t} \quad (3.12)$$

where $0 \leq \rho_g, \rho_z < 1$ and g is the steady state of g_t . The model is driven by the three innovations $\epsilon_t = [\epsilon_{R,t}, \epsilon_{g,t}, \epsilon_{z,t}]'$ that are serially uncorrelated, independent of each other at all leads and lags, and normally distributed with means zero and standard deviations $(\sigma_R, \sigma_g, \sigma_z)$, respectively.

There has been ample empirical evidence of time variation in estimated monetary policy rules documented in the literature. In that vein and for the purpose of illustrating our filter, we also consider a simple regime switching extension of the model.⁸ Specifically, the response of policy rate to inflation deviations is allowed to switch between, in Leeper's (1991) terminology, more 'active' and less 'active' (or possibly 'passive') monetary regimes

$$\psi_\pi(s_t) = \psi_\pi^0(1 - s_t) + \psi_\pi^1 s_t, \quad 0 \leq \psi_\pi^0 < \psi_\pi^1 \quad (3.13)$$

where the regime index s_t evolves according to $s_t = \mathbb{1}\{w_t \geq \tau\}$ and $w_t = \alpha w_{t-1} + v_t$. To keep the illustration simple and concrete, we abstract from sources of time variation in the model structure other than that from the policy parameter $\psi_\pi(s_t)$.⁹ It can be shown as in Chang et al. (2018b)

⁸Our paper complements recent likelihood-based estimation of DSGE models with exogenous Markov switching in monetary policy, including Schorfheide (2005), Liu et al. (2011), Bianchi (2013), Davig and Doh (2014), Bianchi and Ilut (2017), and Bianchi and Melosi (2017), among others. Using the same endogenous switching approach as in this paper, Chang et al. (2018a) also estimated a reduced-form model of monetary-fiscal regime changes.

⁹Initiated by Sims and Zha (2006), allowing for regime switching in both policy rules and shock volatilities has also come under scrutiny in DSGE models, but it would require introducing multiple latent factors to operationalize in our setup, which is beyond the scope of this paper.

that the implied time-varying transition probabilities to regime-0, which become an important part of the model solution in this case, are given by

$$p_{00}(\tilde{\epsilon}_t) = \mathbb{P}(s_{t+1} = 0 | s_t = 0, \tilde{\epsilon}_t) = \frac{\int_{-\infty}^{\tau\sqrt{1-\alpha^2}} \Phi_\rho(\tau - \alpha x / \sqrt{1-\alpha^2} - \rho'\tilde{\epsilon}_t) p_{\mathbb{N}}(x|0, 1) dx}{\Phi(\tau\sqrt{1-\alpha^2})} \quad (3.14)$$

$$p_{10}(\tilde{\epsilon}_t) = \mathbb{P}(s_{t+1} = 0 | s_t = 1, \tilde{\epsilon}_t) = \frac{\int_{\tau\sqrt{1-\alpha^2}}^{\infty} \Phi_\rho(\tau - \alpha x / \sqrt{1-\alpha^2} - \rho'\tilde{\epsilon}_t) p_{\mathbb{N}}(x|0, 1) dx}{1 - \Phi(\tau\sqrt{1-\alpha^2})} \quad (3.15)$$

where $\Phi_\rho(x) = \Phi(x/\sqrt{1-\rho'\rho})$, $\tilde{\epsilon}_t = [\epsilon_{R,t}/\sigma_R, \epsilon_{g,t}/\sigma_g, \epsilon_{z,t}/\sigma_z]'$, and $\rho = [\rho_{Rv}, \rho_{gv}, \rho_{zv}]' = \text{corr}(\tilde{\epsilon}_t, v_{t+1})$. Accordingly, we have $p_{01}(\tilde{\epsilon}_t) = 1 - p_{00}(\tilde{\epsilon}_t)$ and $p_{11}(\tilde{\epsilon}_t) = 1 - p_{10}(\tilde{\epsilon}_t)$. On the other hand, when $\tilde{\epsilon}_t$ and v_{t+1} are orthogonal (i.e., $\rho = 0_{3 \times 1}$), (3.14)–(3.15) become constants so that our model nests the exogenous Markov switching as a special case. We assume that private agents can observe all current and past values of the endogenous variables, exogenous shocks, and regime indices, but not the regime factors. Note that the presence of s_t poses keen computational challenges to solving the model. Subsequently, we obtain the first-order solution to the rational expectations system (3.8)–(3.15) using the efficient perturbation approach of Maih and Waggoner (2018) that is well-suited for solving and estimating nonlinear regime-switching DSGE models with state-dependent transition probabilities.¹⁰ An important precursor to our study is Davig and Leeper (2006a), who employ the projection method to solve and calibrate a new Keynesian model where monetary policy rule changes whenever lagged inflation crosses some threshold level. Interested readers are also referred to Chang et al. (2018b) for a more comprehensive investigation of the macroeconomic origins that give rise to monetary policy shifts.

Both the fixed regime model and its regime switching extension are estimated with Bayesian methods using a common set of quarterly observations, ranging from 1954:Q3 to 2007:Q4: per capita real output growth (YGR), annualized inflation rate (INF), and effective federal funds rate (INT). The actual data are constructed as in Appendix B of Herbst and Schorfheide (2015), available from the Federal Reserve Economic Data (FRED). Let $\hat{x}_t \equiv \ln x_t - \ln x$ denote the log-deviation of a variable x_t from its steady state x . The observable variables are linked to the model variables through the following measurement equations

$$\begin{pmatrix} \text{YGR}_t \\ \text{INF}_t \\ \text{INT}_t \end{pmatrix} = \begin{pmatrix} \gamma^{(Q)} \\ \pi^{(A)} \\ \pi^{(A)} + r^{(A)} + 4\gamma^{(Q)} \end{pmatrix} + 100 \begin{pmatrix} \hat{y}_t - \hat{y}_{t-1} + \hat{z}_t \\ 4\hat{\pi}_t \\ 4\hat{R}_t \end{pmatrix} \quad (3.16)$$

¹⁰All computations are executed in the MATLAB toolbox ‘RISE’. See the Online Appendix for a user guide. Barthélemy and Marx (2017) also generalize standard perturbation methods to solve a class of nonlinear rational expectations models with endogenous regime switching and establish conditions under which a unique bounded equilibrium exists.

Table 1: Prior Distributions of DSGE Parameters

Parameter	Density	Para (1)	Para (2)	[5%, 95%]
Fixed Regime				
τ_c , coefficient of relative risk aversion	\mathbb{G}	2.00	0.50	[1.25, 2.89]
κ , slope of new Keynesian Phillips curve	\mathbb{G}	0.20	0.10	[0.07, 0.39]
ψ_π , interest rate response to inflation	\mathbb{G}	1.50	0.25	[1.11, 1.93]
ψ_y , interest rate response to output	\mathbb{G}	0.50	0.25	[0.17, 0.97]
$r^{(A)}$, s.s. annualized real interest rate	\mathbb{G}	0.50	0.10	[0.35, 0.68]
$\pi^{(A)}$, s.s. annualized inflation rate	\mathbb{G}	3.84	2.00	[1.24, 7.61]
$\gamma^{(Q)}$, s.s. technology growth rate	\mathbb{N}	0.47	0.20	[0.14, 0.80]
ρ_R , persistency of monetary shock	\mathbb{B}	0.50	0.10	[0.34, 0.66]
ρ_g , persistency of spending shock	\mathbb{B}	0.50	0.10	[0.34, 0.66]
ρ_z , persistency of technology shock	\mathbb{B}	0.50	0.10	[0.34, 0.66]
$100\sigma_R$, scaled s.d. of monetary shock	$\mathbb{IG-1}$	0.40	4.00	[0.08, 1.12]
$100\sigma_g$, scaled s.d. of spending shock	$\mathbb{IG-1}$	0.40	4.00	[0.08, 1.12]
$100\sigma_z$, scaled s.d. of technology shock	$\mathbb{IG-1}$	0.40	4.00	[0.08, 1.12]
ν , inverse of demand elasticity	\mathbb{B}	0.10	0.05	[0.03, 0.19]
$1/g$, s.s. consumption-to-output ratio	\mathbb{B}	0.85	0.10	[0.66, 0.97]
Threshold Switching				
ψ_π^0, ψ_π under regime-0	\mathbb{G}	1.00	0.10	[0.84, 1.17]
ψ_π^1, ψ_π under regime-1	\mathbb{G}	2.00	0.25	[1.61, 2.43]
α , persistency of latent factor	\mathbb{B}	0.90	0.05	[0.81, 0.97]
τ , threshold level	\mathbb{N}	-1.00	0.50	[-1.82, -0.18]
ρ_{Rv} , endogeneity from monetary shock	\mathbb{U}	-1.00	1.00	[-0.90, 0.90]
ρ_{gv} , endogeneity from spending shock	\mathbb{U}	-1.00	1.00	[-0.90, 0.90]
ρ_{zv} , endogeneity from technology shock	\mathbb{U}	-1.00	1.00	[-0.90, 0.90]

NOTES: Para (1) and Para (2) refer to the means and standard deviations for Gamma (\mathbb{G}), Normal (\mathbb{N}), and Beta (\mathbb{B}) distributions; s and ν for the Inverse-Gamma Type-I ($\mathbb{IG-1}$) distribution with density $p(\sigma) \propto \sigma^{-\nu-1} \exp(-\frac{\nu s^2}{2\sigma^2})$; the lower and upper bounds of the support for the Uniform (\mathbb{U}) distribution. The effective prior is truncated at the boundary of the determinacy region.

where $(\gamma^{(Q)}, \pi^{(A)}, r^{(A)})$ are connected to the model's steady states via $\gamma = 1 + \gamma^{(Q)}/100$, $\beta = 1/(1 + r^{(A)}/400)$, and $\pi = 1 + \pi^{(A)}/400$. In conjunction with the model solutions under fixed regime and switching regime, this can be cast into the state space form (2.1)–(2.2) and evaluated with the standard Kalman filter and our endogenous-switching Kalman filter in Algorithm 1, respectively.

Table 1 summarizes the marginal prior distributions on the DSGE structural parameters. For

Table 2: Posterior Estimates of DSGE Parameters

Parameter	No Switching Model			Regime Switching Model		
	Mode	Median	[5%, 95%]	Mode	Median	[5%, 95%]
Fixed Regime						
τ_c	3.33	3.29	[2.50, 4.25]	1.09	2.47	[1.72, 3.34]
κ	0.43	0.49	[0.36, 0.65]	1.12	0.69	[0.49, 0.94]
ψ_π	1.00	1.01	[1.00, 1.04]	–	–	–
ψ_y	0.15	0.18	[0.07, 0.38]	0.25	0.27	[0.10, 0.53]
$r^{(A)}$	0.46	0.47	[0.34, 0.62]	0.46	0.47	[0.35, 0.61]
$\pi^{(A)}$	2.76	2.80	[1.65, 3.97]	2.62	2.72	[2.04, 3.38]
$\gamma^{(Q)}$	0.38	0.37	[0.30, 0.45]	0.32	0.37	[0.29, 0.44]
ρ_R	0.71	0.70	[0.66, 0.73]	0.63	0.70	[0.65, 0.74]
ρ_g	0.95	0.95	[0.94, 0.95]	0.95	0.95	[0.94, 0.95]
ρ_z	0.95	0.95	[0.94, 0.95]	0.95	0.95	[0.93, 0.95]
$100\sigma_R$	0.25	0.26	[0.23, 0.28]	0.31	0.27	[0.24, 0.30]
$100\sigma_g$	1.01	1.03	[0.95, 1.12]	1.07	1.04	[0.96, 1.13]
$100\sigma_z$	0.06	0.07	[0.05, 0.08]	0.08	0.06	[0.05, 0.08]
ν	0.08	0.09	[0.03, 0.18]	0.08	0.09	[0.03, 0.18]
$1/g$	0.92	0.87	[0.68, 0.98]	0.92	0.87	[0.69, 0.97]
Threshold Switching						
ψ_π^0	–	–	–	1.03	0.99	[0.92, 1.06]
ψ_π^1	–	–	–	1.56	1.25	[1.10, 1.43]
α	–	–	–	0.99	0.95	[0.91, 0.98]
τ	–	–	–	–1.19	–0.98	[–1.72, –0.23]
ρ_{Rv}	–	–	–	0.60	–0.05	[–0.38, 0.27]
ρ_{gv}	–	–	–	–0.17	0.11	[–0.29, 0.41]
ρ_{zv}	–	–	–	–0.77	–0.71	[–0.91, –0.39]

NOTES: The posterior medians and 90% equal-tailed intervals are computed using 10,000 posterior draws after thinning.

the steady state parameters, the prior means of $\gamma^{(Q)}$ and $\pi^{(A)}$ are calibrated to match the sample averages of YGR and INF, respectively, and that of $r^{(A)}$ translates into a β value of 0.998. The priors on the structural shock processes are harmonized: the autoregressive coefficients (ρ_R, ρ_g, ρ_z) are beta distributed with mean 0.5 and standard deviation 0.1; the standard deviation parameters $(\sigma_R, \sigma_g, \sigma_z)$, all scaled by 100, follow inverse-gamma type-I distribution with mean 0.5 and standard deviation 0.26. Furthermore, the priors on the private sector parame-

Table 3: Log Marginal Likelihood Estimates

Estimator	No Switching Model	Regime Switching Model	Bayes Factor
Meng-Wong	-1099.81	-1075.22*	exp(24.59)
Geweke	-1100.53	-1076.41*	exp(24.12)
Chib-Jeliazkov	-1099.97	-1076.14*	exp(23.83)
Sims-Waggoner-Zha	-1101.47	-1077.23*	exp(24.24)
Müeller	-1099.61	-1074.93*	exp(24.68)

NOTES: Asterisk (*) signifies decisive evidence in favor of the model with superior fit, leading to a log Bayes factor greater than 4.6 based on Jeffreys' (1961) criterion.

ters $(\tau_c, \kappa, \nu, 1/g)$ and the fixed regime policy responses (ψ_π, ψ_y) are largely adopted from An and Schorfheide (2007), whereas those on the switching parameters (ψ_π^0, ψ_π^1) closely follow the specification in Davig and Doh (2014), which a priori rule out the possibility of 'label switching'. Finally, turning to the parameters for the autoregressive regime factor, the prior on α centers at a rather persistent value that, together with the prior mean of τ , implies transition probabilities $(p_{00}, p_{11}) = (0.8, 0.9)$ under exogenous switching. On the other hand, the uniform distributions on $(\rho_{Rv}, \rho_{gv}, \rho_{zv})$ reflect an agnostic prior view about the sign and degree of endogeneity in regime switching.

For each model, we sample a total of 1.1 million draws from the posterior distribution using an adaptive version of the random walk Metropolis-Hastings algorithm, discard the first 100,000 draws as burn-in phase, and keep one every 100 draws afterwards.¹¹ The resulting 10,000 draws form the basis for performing the posterior inference. We highlight several aspects of the posterior estimates reported in Tables 2–3 as follows.

First, regarding the parameters shared by both models, allowing for regime switching does generate material impacts on some of their estimates, including the private sector parameters (τ_c, κ) and the interest rate response to output deviation ψ_y . Other parameters, however, remain more or less recalcitrant to regime switching. Second, turning to the switching parameters, there appears to be considerably less cross-regime difference in the policy response to inflation than assumed a priori—the 90% posterior interval of the more active response ψ_π^1 falls well below its prior mean, suggesting an overall mildly aggressive monetary policy stance. Third, moving on to the parameters (α, τ, ρ) unique to threshold switching, a prior-posterior comparison reveals that the data support an even more persistent process for the regime factor, a much longer expected duration of the more active regime, and attribute a significant portion of regime developments to

¹¹The acceptance rate is targeted within the range [25%, 45%] for both models. Diagnostics to check the convergence of Markov chains include graphical methods such as recursive means plot and the separated partial means test proposed by Geweke (1992). See Herbst and Schorfheide (2015) for a detailed textbook treatment of Bayesian estimation of DSGE models.

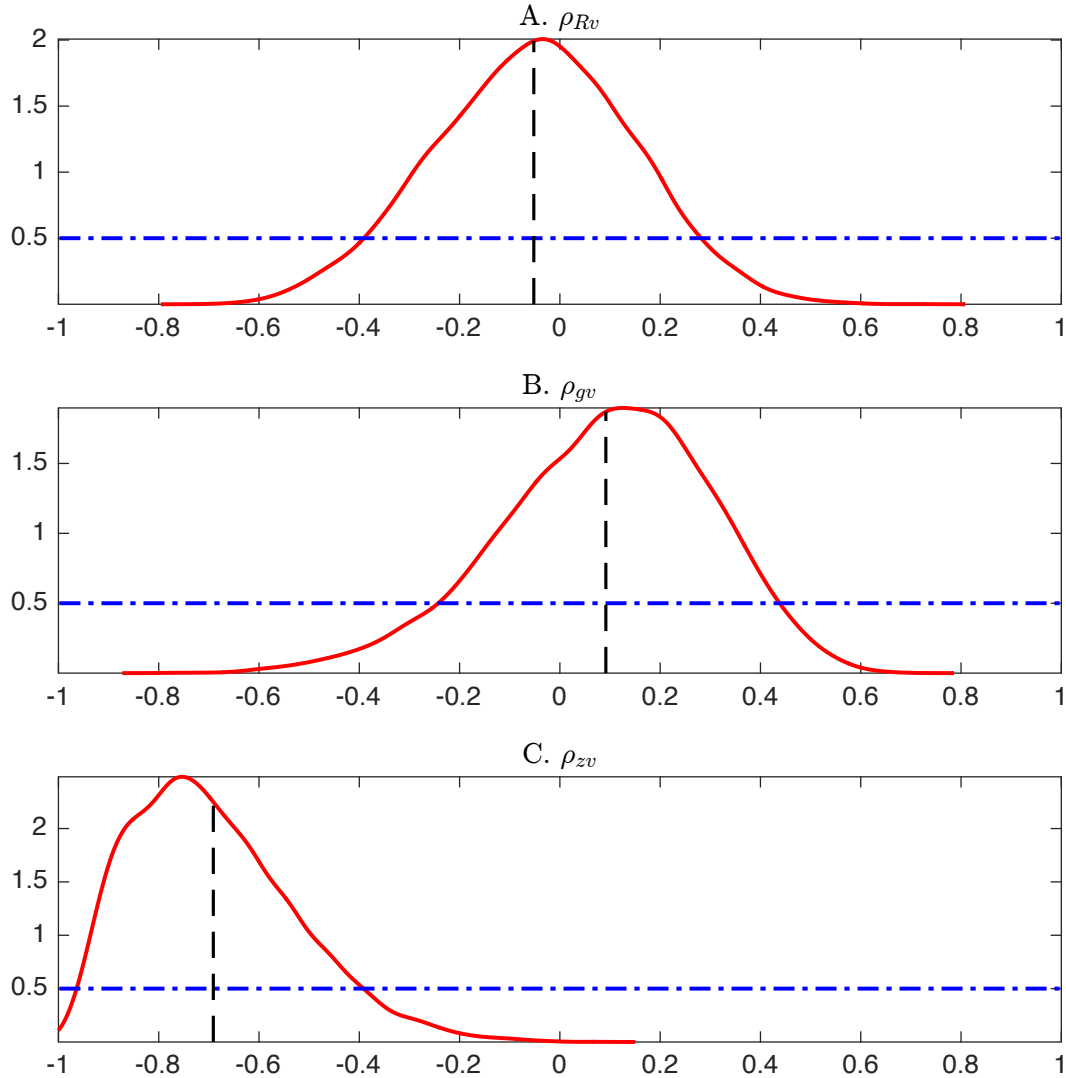


Figure 4: Prior and posterior density functions of correlation parameters. Notes: Each panel plots the prior (blue dash-dotted line) and posterior (red solid line) density functions of a correlation parameter. The black dashed lines delineate the posterior means. All posterior densities are estimated based on the normal kernel function.

past technology growth shocks (about 15%–83%) and, to a lesser extent, government spending shocks (about 0%–17%) as well as monetary policy shocks (about 0%–14%). Finally, Table 3 reports several popular estimators of the marginal likelihood proposed by a number of authors, including Meng and Wong (1996), Geweke (1999), Chib and Jeliazkov (2001), Sims et al. (2008), and Ulrich Müller.¹² All methods yield numerically very similar estimates that point to a unanimous conclusion: the data overwhelmingly favor the endogenous switching model over the fixed regime model.

¹²The Müller estimator is a variant of the bridge sampling method. It was suggested in an unpublished note by Ulrich Müller at Princeton University based on his collaboration with Chris Sims and Tao Zha.

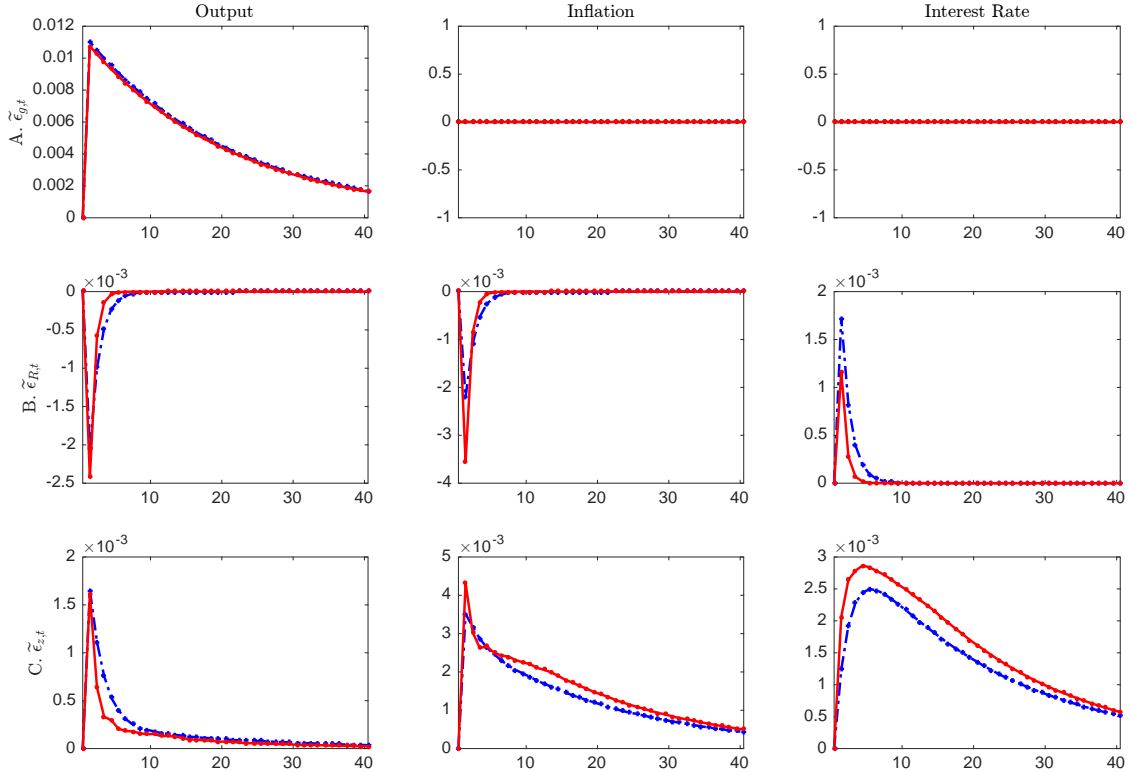


Figure 5: Dynamic responses to one percent positive shock. Notes: Responses are evaluated at the posterior mode of the regime switching model both with (red solid line) and without (blue dash-dotted line) endogenous feedback. The latter is obtained by setting $\rho_{Rv} = \rho_{gv} = \rho_{zv} = 0$ while keeping other parameters unchanged. All variables are in percentage deviations from steady state.

Figure 4 compares the prior and posterior densities of the correlation parameters $(\rho_{Rv}, \rho_{gv}, \rho_{zv})$, which furnish additional evidence on the relevance of accounting for the endogenous feedback from historical macroeconomic shocks to the prevailing policy regime. Despite the diffuse priors, the data turn out to be sufficiently informative to land the posteriors onto much narrower areas of the parameter space that deliver tightly estimated degrees of endogeneity in regime switching. Most noticeable is the endogenous feedback from technology growth shock, whose posterior mass falls entirely on the negative territory. As a result, favorable technological advancements unambiguously increase the likelihood of staying in or shifting to the less active regime, suggesting an accommodative monetary policy to promote long-term economic growth. Less evident is the posterior distribution of the endogenous feedback from government spending shock that concentrates somewhat more on the positive territory. Thus, on average, expansionary government expenditures tend to make the more active regime more likely, consistent with a countercyclical monetary policy that is ‘leaning against the wind’. Monetary policy shock, on the other hand, plays only an insignificant role in driving the regime changes. In sum, these patterns connect broadly to theoretical work and empirical observations about how central banks routinely act.

Compared to purely exogenous switching, endogenizing regime changes also generates quan-

titatively important expectations formation effects. Because the model is inherently nonlinear, Figure 5 displays the generalized impulse response functions of Koop et al. (1996) for key endogenous variables following a one percent positive shock $\tilde{\epsilon}_{i,t} = \epsilon_{i,t}/\sigma_i$, $i \in \{R, g, z\}$. To build intuition, consider first a conventional monetary contraction (Row B) that aims to lower inflation. A higher nominal rate not only translates into a higher real rate due to sticky prices, which dampens aggregate demand in the usual way, but triggers a positive feedback effect ($\rho_{Rv} = 0.60$) that induces agents to form a stronger belief in the more active regime next period. This shift in expectations is absent under exogenous switching (i.e., $\rho_{Rv} = 0$) and, therefore, inflation does not fall as much as under endogenous switching. Analogously, rising productivity (Row C), besides increasing contemporaneous and expected consumption, engenders a negative feedback effect ($\rho_{zv} = -0.77$) that leads agents to place a higher probability on the less active regime next period, thereby amplifying the inflationary impact of higher aggregate demand. Nevertheless, the expectations formation effect arising from a fiscal expansion (Row A) is not directly observable since in this simple model, the government spending shock only affects output but not output gap (hence consumption, inflation, and nominal rate).

Last but not least, unlike the Markov switching filter of Kim (1994), our endogenous switching filter produces an important by-product—an extracted time series of the autoregressive regime factor $w_{t|t}$ as depicted in Figure 6—that complements the information contained in its implied probability $p_{t|t}^1$ of being in the more active regime. In particular, it identifies the U.S. monetary policy as sluggishly fluctuating between the more active and less active regimes, the timing and nature of which are broadly consistent with the previous empirical findings. This pattern also aligns quite well with the narrative record of policymakers’ beliefs documented in Romer and Romer (2004): the more active stance of the late 1950s and most of the 1960s under chairman William McChesney Martin Jr. and of the early 1980s and beyond under Paul Volcker and Alan Greenspan stemmed from the conviction that inflation has high costs and few benefits, whereas the less active stance of most of the 1970s under Arthur Burns and G. William Miller stemmed from a highly optimistic estimate of the natural rate of unemployment and a highly pessimistic estimate of the sensitivity of inflation to economic slack.

4 CONCLUDING REMARKS

This paper aims at broadening the scope for understanding the complex interaction between recurrent structural changes and measured economic behavior. To that end, we introduce a threshold-type endogenous regime switching into an otherwise standard state space model, which is general enough to include many well-known dynamic linear models as special cases. In our approach, regime changes are, through an autoregressive latent factor, jointly driven by the internal innovations that represent the fundamental shocks inside the model and an external innovation that captures all other shocks left outside of the model. This allows the behavior of

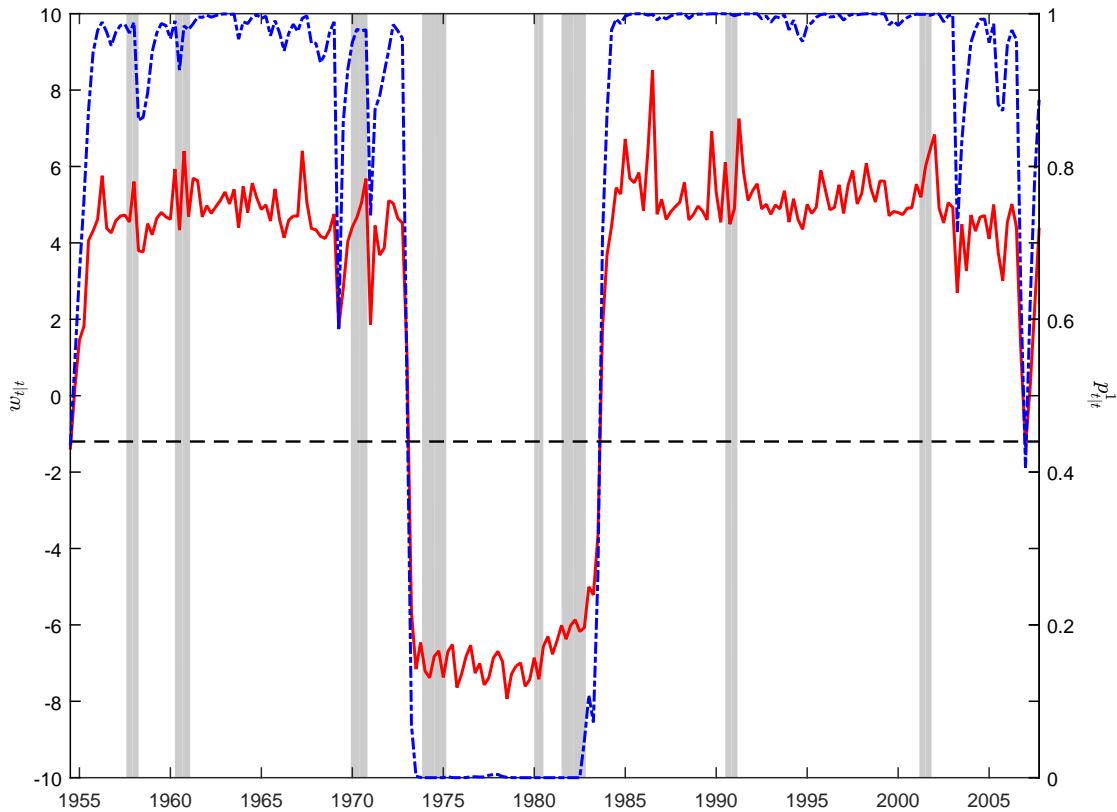


Figure 6: Filtered regime factor and regime-1 probability at the posterior mode. Notes: The left and right vertical axes measure the filtered autoregressive regime factor (red solid line) and the probability of being in regime-1 (blue dash-dotted line), respectively. The black solid line delineates the threshold level. Shaded bars indicate recessions as designated by the National Bureau of Economic Research.

underlying economic fundamentals to be brought to bear more directly on the regime generating process. When regime shifts are purely driven by the external innovation, our model reduces to one with exogenous Markov switching.

We develop an endogenous switching version of the Kalman filter to estimate the overall nonlinear state space model. The filter features augmenting the transition system with the internal innovations that, in conjunction with a collapsing procedure to truncate the regime history, makes the computation feasible and highly efficient. Monte Carlo experiment shows that our filtering algorithm is accurate in approximating both the likelihood function and filtered state variables. We also employ the filter to estimate a prototypical new Keynesian DSGE model with threshold-type switching in monetary policy rule, and find strong empirical support for the endogenous feedback from historical non-policy shocks to the prevailing policy regime. A natural extension of our framework is to permit multiple regimes and latent factors along with the development of both filtering and smoothing algorithms, which we defer to a sequel to this paper.

ACKNOWLEDGEMENTS

The authors would like to thank seminar and conference participants at Australian National University, Central University of Finance and Economics, Keio University, Korea Development Institute, MIT-Harvard, North Carolina State University, Renmin University of China, the Reserve Bank of Australia, Toulouse School of Economics, the University of California–Santa Barbara, the University of Chicago Booth School of Business, the University of Sydney, 2017 NBER Workshop on DSGE Models at Philadelphia Fed, 2018 Workshop on “Nonlinear Models in Macroeconomics and Finance for an Unstable World” at Norges Bank, the 26th Annual Symposium of the Society for Nonlinear Dynamics and Econometrics at Tokyo, 2018 International Association for Applied Econometrics Annual Conference at Montréal, and 2018 Taiwan Econometric Society Annual Meeting, as well as Francesco Bianchi, Yongok Choi, Edward Herbst, Eric Leeper, Joon Park, and Daniel Waggoner for very useful and stimulating comments. Xin Wei provided outstanding research assistance. Fei Tan acknowledges financial support from the Chaifetz School of Business summer research grant.

ONLINE APPENDIX

APPENDIX A: SMOOTHING ALGORITHM We approximate the smoothed density function $p(w_t|\mathcal{F}_T)$ by a discrete density function $\hat{p}(w_t|\mathcal{F}_T)$ defined on a swarm of grid points $\{w_t^i\}_{i=1}^N$ with their corresponding weights $W_t^i = \hat{p}(w_t^i|\mathcal{F}_T)$ such that $\sum_{i=1}^N W_t^i = 1$, which form the basis for computing the smoothed value

$$\mathbb{E}[h(w_t)|\mathcal{F}_T] \approx \sum_{i=1}^N h(w_t^i)W_t^i \quad (\text{A.1})$$

for some transformation function h . For example, setting $h(w_t) = w_t$ and $h(w_t) = \mathbb{1}\{w_t \geq \tau\}$ give $w_{t|T} = \mathbb{E}[w_t|\mathcal{F}_T]$ and $\mathbb{P}(s_t = 1|\mathcal{F}_T)$, respectively.

Our smoothing algorithm is initialized by $\{W_T^i\}_{i=1}^N$ available from (2.22) and then relies on the following backward recursion

$$\begin{aligned} p(w_t^j|\mathcal{F}_T) &= \int_{-\infty}^{\infty} p(w_t^j|w_{t+1}, \mathcal{F}_T)p(w_{t+1}|\mathcal{F}_T)dw_{t+1} \\ &\approx \sum_{i=1}^N p(w_t^j|w_{t+1}^i, \mathcal{F}_T)W_{t+1}^i \end{aligned} \quad (\text{A.2})$$

for $j = 1, \dots, N$ and $t = T - 1, \dots, 1$. By the Bayes formula

$$\begin{aligned}
 p(w_t^j | w_{t+1}^i, \mathcal{F}_T) &= \frac{p(Y_{t+1:T}, w_{t+1}^i | w_t^j, \mathcal{F}_t) p(w_t^j | \mathcal{F}_t)}{p(Y_{t+1:T}, w_{t+1}^i | \mathcal{F}_t)} \\
 &= \frac{p(Y_{t+1:T} | w_{t+1}^i, w_t^j, \mathcal{F}_t) p(w_{t+1}^i | w_t^j, \mathcal{F}_t) p(w_t^j | \mathcal{F}_t)}{p(Y_{t+1:T} | w_{t+1}^i, \mathcal{F}_t) p(w_{t+1}^i | \mathcal{F}_t)} \\
 &= \frac{p(w_{t+1}^i | w_t^j, \mathcal{F}_t) p(w_t^j | \mathcal{F}_t)}{p(w_{t+1}^i | \mathcal{F}_t)} \tag{A.3}
 \end{aligned}$$

since $p(Y_{t+1:T} | w_{t+1}^i, w_t^j, \mathcal{F}_t)$ is independent of w_t^j . To evaluate (A.3), we first approximate

$$p(w_{t+1}^i | w_t^j, \mathcal{F}_t) \approx p_{\mathbb{N}}(w_{t+1}^i | \alpha w_t^j + \rho' \varsigma_{d,t|t}^k, 1 - \rho' \rho), \quad k = \mathbb{1}\{w_t^j \geq \tau\}$$

Moreover, $p(w_t^j | \mathcal{F}_t)$ is available from (2.22). Lastly, we approximate

$$p(w_{t+1}^i | \mathcal{F}_t) \approx \sum_{k=0}^1 p(w_{t+1}^i | s_t = k, \rho' \varsigma_{d,t|t}^k) p(s_t = k | \mathcal{F}_t)$$

where $p(w_{t+1}^i | s_t = k, \rho' \varsigma_{d,t|t}^k)$ is given by (2.23) and $p(s_t = k | \mathcal{F}_t)$ is a by-product of running our endogenous-switching Kalman filter. Now set $W_t^j = p(w_t^j | \mathcal{F}_T) / \sum_{j=1}^N p(w_t^j | \mathcal{F}_T)$ for the next iteration. Evaluating at the true parameter values, Figure 7 depicts the smoothed autoregressive regime factor for the small simulation model of Section 3.1, which to a large extent recovers its true trajectory.

APPENDIX B: TRANSITION PROBABILITY We are interested in computing integrals of the form

$$\begin{aligned}
 F(a, b, c, d) &= \int_{-\infty}^{\infty} \int_{-\infty}^a \Phi(b + cx + dy) p_{\mathbb{N}}(x | 0, 1) p_{\mathbb{N}}(y | m, V) dx dy \\
 &= \mathbb{P}(Z_1 \leq a, Z_2 \leq b + cZ_1 + dZ_3) \tag{B.4}
 \end{aligned}$$

for some $a, b, c, d \in \mathbb{R}$, where $Z_1, Z_2 \sim \mathbb{N}(0, 1)$, $Z_3 \sim \mathbb{N}(m, V)$, and (Z_1, Z_2, Z_3) are independent of each other. Define $W_1 = Z_1 \sim \mathbb{N}(0, 1)$ and $W_2 = Z_2 - dZ_3 \sim \mathbb{N}(-dm, 1 + d^2V)$. Then (W_1, W_2) are independent and (B.4) can be rewritten as

$$\begin{aligned}
 F(a, b, c, d) &= \mathbb{P}(W_1 \leq a, W_2 - cW_1 \leq b) \\
 &= \int_{-\infty}^a \int_{-\infty}^b p_{\mathbb{N}}(x, y | \mu, \Sigma) dx dy \tag{B.5}
 \end{aligned}$$

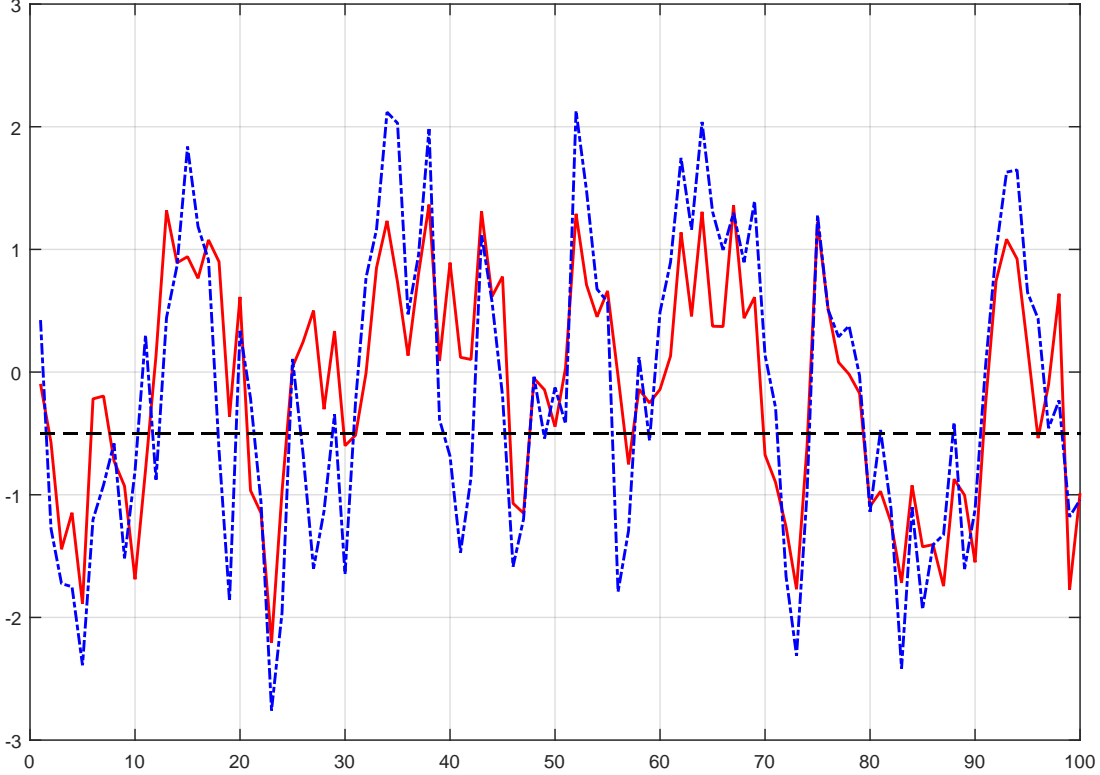


Figure 7: Smoothed regime factor at the true parameter values. Notes: The figure compares the smoothed autoregressive regime factor computed from the smoothing algorithm (red solid line) to its true value (blue dash-dotted line). The black dashed line delineates the threshold level.

where

$$\mu = \begin{pmatrix} 0 \\ -dm \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 1 & -c \\ -c & 1 + c^2 + d^2V \end{pmatrix}$$

Therefore, the transition probabilities can be calculated as

$$\mathbb{P}(s_t = 0 | s_{t-1} = 0) = \frac{1}{\Phi(\tau\sqrt{1-\alpha^2})} F\left(\tau\sqrt{1-\alpha^2}, \frac{\tau}{\sqrt{1-\rho'\rho}}, -\frac{\alpha}{\sqrt{1-\rho'\rho}\sqrt{1-\alpha^2}}, -\frac{1}{\sqrt{1-\rho'\rho}}\right)$$

with $m = \rho'\zeta_{d,t-1|t-1}^0$, $V = \rho'P_{d,t-1|t-1}^0\rho$, and

$$\mathbb{P}(s_t = 0 | s_{t-1} = 1) = \frac{1}{1 - \Phi(\tau\sqrt{1-\alpha^2})} F\left(-\tau\sqrt{1-\alpha^2}, \frac{\tau}{\sqrt{1-\rho'\rho}}, \frac{\alpha}{\sqrt{1-\rho'\rho}\sqrt{1-\alpha^2}}, -\frac{1}{\sqrt{1-\rho'\rho}}\right)$$

with $m = \rho'\zeta_{d,t-1|t-1}^1$ and $V = \rho'P_{d,t-1|t-1}^1\rho$.

APPENDIX C: USER GUIDE This brief tutorial provides information for replicating the empirical results of Section 3.2 in RISE, a flexible object-oriented MATLAB toolbox developed by Junior Maih. It is available, free of charge, at https://github.com/jmaih/RISE_toolbox.

1. Add the folders ‘ChangMaihTan2018/’ and ‘RISE_toolbox-master/’ to the MATLAB search path. In particular, ‘ChangMaihTan2018/’ serves as the main folder that contains two subfolders. The ‘core/’ subfolder includes the following files that can be replaced by the user’s own model and data set:

- as07_cmt.rs
- as07_cmt_para.m
- ass_cmt_ssfile.m
- as07_data.txt

The ‘routines/’ subfolder, on the other hand, includes those files that the user typically will not need to modify for different models.

2. To solve and estimate the model, run the function ‘as07_cmt_driver.m’.
3. To produce a summary report in PDF format, run the function ‘as07_cmt_report.m’.

REFERENCES

- AN, S., AND F. SCHORFHEIDE (2007): “Bayesian Analysis of DSGE Models,” *Econometric Reviews*, 26(2), 113–172.
- BARTHÉLEMY, J., AND M. MARX (2017): “Solving endogenous regime switching models,” *Journal of Economic Dynamics and Control*, 77, 1–25.
- BAZZI, M., F. BLASQUES, S. J. KOOPMAN, AND A. LUCAS (2014): “Time-Varying Transition Probabilities for Markov Regime Switching Models,” *Journal of Time Series Analysis*, 38(3), 458–478.
- BIANCHI, F. (2013): “Regime Switches, Agents’ Beliefs, and Post-World War II U.S. Macroeconomic Dynamics,” *The Review of Economic Studies*, 80(2), 463–490.
- BIANCHI, F., AND C. ILUT (2017): “Monetary/Fiscal Policy Mix and Agent’s Beliefs,” *Review of Economic Dynamics*, 26, 113–139.
- BIANCHI, F., AND L. MELOSI (2017): “Escaping the Great Recession,” *American Economic Review*, 107(4), 1030–58.

- CHANG, Y., Y. CHOI, AND J. Y. PARK (2017): “A New Approach to Model Regime Switching,” *Journal of Econometrics*, 196(1), 127 – 143.
- CHANG, Y., B. KWAK, AND S. QIU (2018a): “U.S. Monetary-Fiscal Regime Changes in the Presence of Endogenous Feedback in Policy Rules,” Manuscript.
- CHANG, Y., F. TAN, AND X. WEI (2018b): “What Drives Monetary Policy Shifts? A New Approach to Regime Switching in DSGE Models,” Manuscript.
- CHEN, R., AND J. S. LIU (2000): “Mixture Kalman Filters,” *J. R. Statist. Soc. B*, 62, 493–508.
- CHIB, S. (1996): “Calculating Posterior Distributions and Modal Estimates in Markov Mixture Models,” *Journal of Econometrics*, 75, 79–97.
- CHIB, S., AND I. JELIAZKOV (2001): “Marginal Likelihood from the Metropolis-Hastings Output,” *Journal of the American Statistical Association*, 96(453), 270–281.
- CLARIDA, R., J. GALÍ, AND M. GERTLER (2000): “Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory,” *Quarterly Journal of Economics*, 115(1), 147–180.
- DAVIG, T., AND T. DOH (2014): “Monetary Policy Regime Shifts and Inflation Persistence,” *The Review of Economics and Statistics*, 96(5), 862–875.
- DAVIG, T., AND E. M. LEEPER (2006a): “Endogenous Monetary Policy Regime Change,” in *NBER International Seminar on Macroeconomics 2006*, ed. by L. Reichlin, and K. D. West, pp. 345–377. MIT Press, Cambridge, MA.
- (2006b): “Fluctuating Macro Policies and the Fiscal Theory,” in *NBER Macroeconomics Annual 2006*, ed. by D. Acemoglu, K. Rogoff, and M. Woodford, vol. 21, pp. 247–298. MIT Press, Cambridge.
- DEJONG, D. N., AND C. DAVE (2007): *Structural Macroeconometrics*. Princeton University Press, Princeton, N.J.
- DIEBOLD, F. X., J.-H. LEE, AND G. WEINBACH (1994): “Regime switching with time-varying transition probabilities,” in *Nonstationary Time Series Analysis and Cointegration*, ed. by C. Hargreaves, pp. 283–302. Oxford University Press.
- DOUCET, A., AND A. M. JOHANSEN (2011): “A tutorial on particle filtering and smoothing: fifteen years later,” Manuscript.
- GEWEKE, J. (1992): “Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments,” in *Bayesian Statistics 4*, ed. by J. Bernardo, J. Berger, A. Dawid, and A. Smith. Clarendon Press, Oxford, UK.

- (1999): “Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication,” *Econometric Reviews*, 18, 1–73.
- GORDON, N., D. SALMOND, AND A. SMITH (1993): “Novel approach to nonlinear/non-Gaussian Bayesian state estimation,” *IEEE Proceedings F, Radar and Signal Processing*, 140(2), 107–113.
- HAMILTON, J. D. (1988): “Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates,” *Journal of Economic Dynamics and Control*, 12(2), 385–423.
- (1989): “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle,” *Econometrica*, 57(2), 357–384.
- HERBST, E. P., AND F. SCHORFHEIDE (2015): *Bayesian Estimation of DSGE Models*. Princeton University Press.
- JEFFREYS, J. (1961): *Theory of Probability*. Oxford University Press.
- KALLIOVIRTA, L., M. MEITZ, AND P. SAIKKONEN (2015): “A Gaussian Mixture Autoregressive Model for Univariate Time Series,” *Journal of Time Series Analysis*, 36(2), 247–266.
- KANG, K. H. (2014): “Estimation of state-space models with endogenous Markov regime-switching parameters,” *The Econometrics Journal*, 17(1), 56–82.
- KIM, C.-J. (1994): “Dynamic Linear Models with Markov-Switching,” *Journal of Econometrics*, 60(1), 1 – 22.
- (2004): “Markov-Switching Models with Endogenous Explanatory Variables,” *Journal of Econometrics*, 122(1), 127–136.
- (2009): “Markov-switching models with endogenous explanatory variables II: A two-step MLE procedure,” *Journal of Econometrics*, 148(1), 46–55.
- KIM, C.-J., AND C. R. NELSON (1999): *State-Space Models with Regime Switching*. MIT Press, Cambridge, MA.
- KIM, C.-J., J. PIGER, AND R. STARTZ (2008): “Estimation of Markov regime-switching regression models with endogenous switching,” *Journal of Econometrics*, 143(2), 263–273.
- KOOP, G., M. PESARAN, AND S. M. POTTER (1996): “Impulse Response Analysis in Nonlinear Multivariate Models,” *Journal of Econometrics*, 74(1), 119 – 147.

- LEEPER, E. M. (1991): “Equilibria Under ‘Active’ and ‘Passive’ Monetary and Fiscal Policies,” *Journal of Monetary Economics*, 27(1), 129–147.
- LIU, Z., D. F. WAGGONER, AND T. ZHA (2011): “Sources of macroeconomic fluctuations: A regime-switching DSGE approach,” *Quantitative Economics*, 2(2), 251–301.
- MAIH, J., AND D. WAGGONER (2018): “Perturbation Methods for DSGE Models with Time-Varying Coefficients and Transition Matrices,” Mimeograph, Norges Bank.
- MENG, X.-L., AND W. H. WONG (1996): “Simulating Ratios of Normalizing Constants via a Simple Identity: A Theoretical Exploration,” *Statistica Sinica*, 6(4), 831–860.
- ROMER, C. D., AND D. H. ROMER (2004): “Choosing the Federal Reserve Chair: Lessons from History,” *Journal of Economic Perspectives*, 18(1), 129–62.
- SCHORFHEIDE, F. (2005): “Learning and Monetary Policy Shifts,” *Review of Economic Dynamics*, 8(2), 392–419.
- SIMS, C. A., D. F. WAGGONER, AND T. ZHA (2008): “Methods for Inference in Large Multiple-Equation Markov-Switching Models,” *Journal of Econometrics*, 146(2), 255 – 274.
- SIMS, C. A., AND T. ZHA (2006): “Were There Regime Switches in US Monetary Policy?,” *American Economic Review*, 96(1), 54–81.
- TANNER, M. A., AND W. H. WONG (1987): “The Calculation of Posterior Distributions by Data Augmentation,” *Journal of the American Statistical Association*, 82(398), 528–540.
- TAYLOR, J. B. (1993): “Discretion versus Policy Rules in Practice,” *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.
- WEI, G. C. G., AND M. A. TANNER (1990): “A Monte Carlo Implementation of the EM Algorithm and the Poor Man’s Data Augmentation Algorithms,” *Journal of the American Statistical Association*, 85(411), 699–704.