CAEPR
Working
Paper
#2018-012

State Space Models with Endogenous Regime Switching

Yoosoon Chang
Indiana University

Fei Tan
Dept. of Economics
Chaifetz School of
Business, Saint Louis
University and Center
for Economic Behavior
and Decision-Making,
Zhejiang University of
Finance and Economics

Xin Wei
Indiana University

March 30, 2018

This paper can be downloaded without charge from the Social Science Research Network electronic library at https://papers.ssrn.com/sol3/abstract_id=3770135

The Center for Applied Economics and Policy Research resides in the Department of Economics at Indiana University Bloomington. CAEPR can be found on the Internet at: <http://www.indiana.edu/~caepr>. CAEPR can be reached via email at caepr@indiana.edu or via phone at 812-855-4050.

©2018 by Yoosoon Chang, Fei Tan and Xin Wei. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

State Space Models with Endogenous Regime Switching

Yoosoon Chang, Fei Tan*, Xin Wei

[This Version: March 30, 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 regime switching. Due to the flexible form of state space representation, this class of models is 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 nonlinear model. 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 that the Bayes factor decisively favors the endogenous switching version of the DSGE model over the fixed regime case. Overall, our approach provides a greater scope for understanding the complex interaction between regime switching and measured economic behavior.

JEL classification: C13, C32, E52

Keywords: state space model; regime switching; endogeneity; filtering; DSGE model.

*Chang: Department of Economics, Indiana University. Tan: Department of Economics, John Cook School of Business, Saint Louis University. Wei: Department of Economics, Indiana University. Send correspondence to: 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 in many cases nor innocuous. As argued forcefully in Chang, Choi, and Park (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, Lee, and Weinbach (1994), Kim (2004, 2009), Kim, Piger, and Startz (2008), Bazzi, Blasques, Koopman, and Lucas (2017), Kang (2014), Kalliovirta, Meitz, and Saikkonen (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, Choi, and Park (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 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. Apparently, 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 models fundamentals. In the special case where regime shifts are purely

Figure 1: Federal Funds Rate and Monetary Policy Intervention

Notes: Panel A plots the effective federal funds rate (blue line) and the rate implied by an inertial version of the Taylor (1993) rule (dashed red line) as shown in the plot, where i denotes the federal funds rate, π the annual inflation rate, y the percentage deviation of real output from its potential. Panel B depicts their differential. Shaded bars indicate recessions as designed by the National Bureau of Economic Research.

driven by the external innovation, our model reduces to one with exogenous switching.

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 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, we apply the filter to estimate an endogenous switching version of the new Keynesian DSGE model in An and Schorfheide (2007) with the U.S. postwar data, and find its strong statistical support over the fixed regime case. Ever since the seminal work of Clarida, Galí, and Gertler (2000), modeling the time variation in monetary policy has remained an active research agenda for macroeconomists. To that end, the regime switching approach has emerged nowadays as perhaps the most popular modeling choice in dynamic macro models. However, scant attention in the literature has been paid to why monetary policy regime has shifted over time. This paper takes a first step

toward filling in the important gap our estimated DSGE model implies that about 13% – 45% of the U.S. monetary policy shifts can be attributed to historical monetary interventions, the remainder of which to contemporaneous external shock.

While this attribution is model dependent and may be viewed as suggestive, Figure 1 displays some empirical evidence of such endogenous feedback mechanism. Panel A makes clear that the Taylor rule for setting the federal funds rate provides by and large an accurate account of the 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 dramatic loosening of policy under Federal Reserve chairmen Arthur Burns and G. William Miller in the early and late 1970s, followed by the sustained and 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.¹ To the best of our knowledge, modeling this channel is 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 $\mathbb{N}(a, b)$ denote the normal distribution with mean a and variance b , $p_{\mathbb{N}}(\cdot|a, b)$ its probability density function, and $\Phi(\cdot)$ the cumulative distribution function of $\mathbb{N}(0, 1)$. 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, Choi, and Park (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 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

¹As Chris Sims put it in his comment on Davig and Leeper (2007), "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."

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

(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

$$y_t = D(s_t) + Z(s_t)x_t + F(s_t)z_t + \Sigma_u^{1/2}u_t, \quad u_t \sim \mathbb{N}(0, I_l) \quad (2.1)$$

$$x_t = G(s_t)x_{t-1} + M(s_t)\Sigma_\epsilon^{1/2}\epsilon_t, \quad \epsilon_t \sim \mathbb{N}(0, I_n) \quad (2.2)$$

where the measurement equation (2.1) links the observable variables to the state variables subject to a $l \times 1$ vector of measurement errors $\Sigma_u^{1/2}u_t$, the transition equation (2.2) describes the evolution of the state variables driven by a $n \times 1$ vector of exogenous innovations $\Sigma_\epsilon^{1/2}\epsilon_t$, and (u_t, ϵ_t) are mutually and serially uncorrelated at all leads and lags. The coefficient matrices $D(\cdot)$, $Z(\cdot)$, $F(\cdot)$, $G(\cdot)$, and $M(\cdot)$ 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, Choi, and Park (2017), but we allow all current transition innovations to jointly influence the next period regime through its correlation with the innovation to w_{t+1} . That is,

$$\begin{pmatrix} \epsilon_t \\ v_{t+1} \end{pmatrix} \sim \mathbb{N} \left(\begin{pmatrix} 0_n \\ 0 \end{pmatrix}, \begin{pmatrix} I_n & \rho \\ \rho' & 1 \end{pmatrix} \right), \quad \rho'\rho < 1 \quad (2.4)$$

where $\epsilon_t = [\epsilon_{1,t}, \dots, \epsilon_{n,t}]'$ and $\rho = [\rho_1, \dots, \rho_n]'$ 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. This type of endogenous impacts is not only sustained due to the autoregressive form of w_t , but also renders the

³In linearized DSGE models, (2.2) represents the regime-dependent solution to the model variables, and the coefficient 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.

transition probabilities time-varying because they are all functions of ϵ_t . In particular, as shown in Chang, Tan, and Wei (2018), the transition probabilities to regime-0 are given by

$$p_{00}(\epsilon_t) = \mathbb{P}\{s_{t+1} = 0 | s_t = 0, \epsilon_t\} = \frac{\int_{-\infty}^{\tau\sqrt{1-\alpha^2}} \Phi_\rho\left(\tau - \frac{\alpha x}{\sqrt{1-\alpha^2}} - \rho'\epsilon_t\right) p_{\mathbb{N}}(x|0, 1) dx}{\Phi(\tau\sqrt{1-\alpha^2})} \quad (2.5)$$

$$p_{10}(\epsilon_t) = \mathbb{P}\{s_{t+1} = 0 | s_t = 1, \epsilon_t\} = \frac{\int_{\tau\sqrt{1-\alpha^2}}^{\infty} \Phi_\rho\left(\tau - \frac{\alpha x}{\sqrt{1-\alpha^2}} - \rho'\epsilon_t\right) p_{\mathbb{N}}(x|1, 1) dx}{1 - \Phi(\tau\sqrt{1-\alpha^2})} \quad (2.6)$$

where $\Phi_\rho(x) = \Phi(x/\sqrt{(1-\rho^2)})$. Accordingly, we have $p_{01}(\epsilon_t) = 1 - p_{00}(\epsilon_t)$ and $p_{11}(\epsilon_t) = 1 - p_{10}(\epsilon_t)$. In the special case where ϵ_t and v_{t+1} are orthogonal (i.e., $\rho = 0$), (2.5)–(2.6) become constants and our model reduces to one with exogenous regime switching.

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

$$v_{t+1} = \sum_{k=1}^n \rho_k \epsilon_{k,t} + \sqrt{1 - \sum_{k=1}^n \rho_k^2} \eta_{t+1}, \quad \eta_{t+1} \sim \mathbb{N}(0, 1) \quad (2.7)$$

where $\{\epsilon_{k,t}\}_{k=1}^n$ 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 ϵ_t appears to be a vague source of regime change in many economic applications where ϵ_t is endowed with certain behavioral interpretations. But it indeed captures potential misspecification of the transition equation ideally one would expect the regime change to be fully driven by ϵ_t under the true model that leads to systematic disparities between latent and observable variables. To some extent η_t plays a role reminiscent of that of the explanatory variables z_t . We therefore focus on the important case $z_t = \eta_t$ in the subsequent analysis, and treat η_t as an added regressor summarizing all external information beyond what is contained in (y_t, x_t) or their lags. Accordingly, we call $(\epsilon_{k,t}, \eta_t)$ 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 (2.7), we have the conditional variance

$$Var_t(w_{t+h}) = \underbrace{\sum_{k=1}^n \rho_k^2 \sum_{j=1}^h \alpha^{2(h-j)}}_{\text{due to } \epsilon_k} + \underbrace{\left(1 - \sum_{k=1}^n \rho_k^2\right) \sum_{j=1}^h \alpha^{2(h-j)}}_{\text{due to } \eta} = \underbrace{\sum_{j=1}^h \alpha^{2(h-j)}}_{\text{total}}, \quad h \geq 1 \quad (2.8)$$

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 past monetary interventions.

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 unobserved state variables (i.e., x_t and w_t). While the system is linear in x_t and driven by Gaussian innovations, complication arises from the presence of w_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 $d_t = \epsilon_t$ and augment the state vector x_t as $\varsigma_t = [x'_t, d'_t]'$. Accordingly, we rewrite the state space model as

$$y_t = \underbrace{D(s_t) + F(s_t)z_t}_{\tilde{D}(s_t)} + \underbrace{\begin{pmatrix} Z(s_t) & 0_{l \times n} \end{pmatrix}}_{\tilde{Z}(s_t)} \underbrace{\begin{pmatrix} x_t \\ d_t \end{pmatrix}}_{\varsigma_t} + \Sigma_u^{1/2} u_t \quad (2.9)$$

$$\underbrace{\begin{pmatrix} x_t \\ d_t \end{pmatrix}}_{\varsigma_t} = \underbrace{\begin{pmatrix} G(s_t) & 0_{m \times n} \\ 0_{n \times m} & 0_{n \times n} \end{pmatrix}}_{\tilde{G}(s_t)} \underbrace{\begin{pmatrix} x_{t-1} \\ d_{t-1} \end{pmatrix}}_{\varsigma_{t-1}} + \underbrace{\begin{pmatrix} M(s_t)\Sigma_\epsilon^{1/2} \\ I_n \end{pmatrix}}_{\tilde{M}(s_t)} \epsilon_t \quad (2.10)$$

where we suppress the dependence of $\tilde{D}(s_t)$ on z_t for convenience.

Our main filtering algorithm, which is based on the augmented state space system, 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 $\mathcal{F}_t = \sigma(\{z_s, y_s\}_{s \leq t})$. Define the predictive and filtered probabilities 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 $p_{t|t}^{(i,j)} = \mathbb{P}(s_{t-1} = i, s_t = j | \mathcal{F}_t)$, respectively. Also define the filtered marginal probability of regime- j at period t as $p_{t|t}^j = \mathbb{P}(s_t = j | \mathcal{F}_t)$. 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 and covariance of ς_0 , $(\varsigma_{0|0}^i, P_{0|0}^i)$, using the invariant distribution under regime i . Specify $p_{0|0}^0 = \Phi(\tau\sqrt{1-\alpha^2})$ according to the invariant distribution of w_t , i.e., $\mathbb{N}(0, 1/(1-\alpha^2))$, and compute $p_{0|0}^1 = 1 - p_{0|0}^0$.

2. **Recursion.** For each $t = 1, \dots, T$, the filter accepts two sets of triple inputs $\{\varsigma_{t-1|t-1}^i, P_{t-1|t-1}^i, p_{t-1|t-1}^i\}$ for $i = 0, 1$ and returns two sets of updated triple outputs $\{\varsigma_{t|t}^j, P_{t|t}^j, p_{t|t}^j\}$ for $j = 0, 1$.

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

$$\varsigma_{t|t-1}^{(i,j)} = \tilde{G}(s_t = j)\varsigma_{t-1|t-1}^i \quad (2.11)$$

$$P_{t|t-1}^{(i,j)} = \tilde{G}(s_t = j)P_{t-1|t-1}^i\tilde{G}(s_t = j)' + \tilde{M}(s_t = j)\tilde{M}(s_t = j)' \quad (2.12)$$

for $i = 0, 1$ and $j = 0, 1$. Next, let $\mu_t = \rho'\epsilon_t$. Compute

$$p_{t|t-1}^{(0,0)} = \mathbb{P}(s_t = 0 | s_{t-1} = 0, \mathcal{F}_{t-1})p_{t-1|t-1}^0 \quad (2.13)$$

and $p_{t|t-1}^{(0,1)} = p_{t-1|t-1}^0 - p_{t|t-1}^{(0,0)}$. Note that

$$\mathbb{P}(s_t = 0 | s_{t-1} = 0, \mathcal{F}_{t-1}) = \int_{\mathbb{R}} \mathbb{P}(s_t = 0 | s_{t-1} = 0, \mu_{t-1}, \mathcal{F}_{t-1})p(\mu_{t-1} | s_{t-1} = 0, \mathcal{F}_{t-1})d\mu_{t-1} \quad (2.14)$$

$$= \int_{\mathbb{R}} \mathbb{P}(s_t = 0 | s_{t-1} = 0, \mu_{t-1})p(\mu_{t-1} | s_{t-1} = 0, \mathcal{F}_{t-1})d\mu_{t-1} \quad (2.15)$$

$$= \frac{\int_{\mathbb{R}} \int_{-\infty}^{\tau\sqrt{1-\alpha^2}} \Phi_{\rho}\left(\tau - \frac{\alpha x}{\sqrt{1-\alpha^2}} - \mu_{t-1}\right) \varphi(x)p(\mu_{t-1} | s_{t-1} = 0, \mathcal{F}_{t-1})dx d\mu_{t-1}}{\Phi(\tau\sqrt{1-\alpha^2})} \quad (2.16)$$

$$= \frac{\int_{\mathbb{R}} \int_{-\infty}^{\frac{\tau}{\sqrt{1-\rho'\rho}}} \int_{-\infty}^{\tau\sqrt{1-\alpha^2}} f_0(x, y, \mu_{t-1})dx dy d\mu_{t-1}}{\Phi(\tau\sqrt{1-\alpha^2})} \quad (2.17)$$

with⁴

$$f_0(x, y, \mu_{t-1}) = \mathbb{N} \left(\begin{pmatrix} 0 \\ \frac{m^0}{\sqrt{1-\rho'\rho}} \\ m^0 \end{pmatrix}, \begin{pmatrix} 1 & \frac{\alpha}{\sqrt{1-\rho'\rho}\sqrt{1-\alpha^2}} & 0 \\ \frac{\alpha}{\sqrt{1-\rho'\rho}\sqrt{1-\alpha^2}} & \frac{\alpha^2}{(1-\alpha^2)(1-\rho'\rho)} + 1 + \frac{V^0}{1-\rho'\rho} & \frac{V^0}{\sqrt{1-\rho'\rho}} \\ 0 & \frac{V^0}{\sqrt{1-\rho'\rho}} & V^0 \end{pmatrix} \right),$$

where we approximate $p(\mu_{t-1}|s_{t-1} = 0, \mathcal{F}_{t-1})$ as a normal pdf with mean $m^0 = \rho'\zeta_{\epsilon,t-1|t-1}^0$ and variance $V^0 = \rho'P_{\epsilon,t-1|t-1}^0\rho$, where $(\zeta_{\epsilon,t-1|t-1}^0, P_{\epsilon,t-1|t-1}^0)$ can be extracted from $(\zeta_{t-1|t-1}^0, P_{t-1|t-1}^0)$ corresponding to the augmented shocks. Similarly, we compute

$$p_{t|t-1}^{(1,0)} = \mathbb{P}(s_t = 0|s_{t-1} = 1, \mathcal{F}_{t-1})p_{t-1|t-1}^1 \quad (2.18)$$

and $p_{t|t-1}^{(1,1)} = p_{t-1|t-1}^1 - p_{t|t-1}^{(1,0)}$. Note that

$$\mathbb{P}(s_t = 0|s_{t-1} = 1, \mathcal{F}_{t-1}) = \frac{\int_{\mathbb{R}} \int_{-\infty}^{\frac{\tau}{\sqrt{1-\rho'\rho}}} \int_{\tau\sqrt{1-\alpha^2}}^{\infty} f_1(x, y, \mu_{t-1}) dx dy d\mu_{t-1}}{1 - \Phi(\tau\sqrt{1-\alpha^2})} \quad (2.19)$$

with

$$f_1(x, y, \mu_{t-1}) = \mathbb{N} \left(\begin{pmatrix} 0 \\ \frac{m^1}{\sqrt{1-\rho'\rho}} \\ m^1 \end{pmatrix}, \begin{pmatrix} 1 & -\frac{\alpha}{\sqrt{1-\rho'\rho}\sqrt{1-\alpha^2}} & 0 \\ -\frac{\alpha}{\sqrt{1-\rho'\rho}\sqrt{1-\alpha^2}} & \frac{\alpha^2}{(1-\alpha^2)(1-\rho'\rho)} + 1 + \frac{V^1}{1-\rho'\rho} & \frac{V^1}{\sqrt{1-\rho'\rho}} \\ 0 & \frac{V^1}{\sqrt{1-\rho'\rho}} & V^1 \end{pmatrix} \right),$$

where we approximate $p(\mu_{t-1}|s_{t-1} = 1, \mathcal{F}_{t-1})$ as a normal pdf with mean $m^1 = \rho'\zeta_{\epsilon,t-1|t-1}^1$ and variance $V^1 = \rho'P_{\epsilon,t-1|t-1}^1\rho$.

(b) **Likelihood evaluation.** Apply the Kalman filter forecasting step for observables to obtain

$$\begin{aligned} y_{t|t-1}^{(i,j)} &= \tilde{D}(s_t = j) + \tilde{Z}(s_t = j)\zeta_{t|t-1}^{(i,j)}, \\ F_{t|t-1}^{(i,j)} &= \tilde{Z}(s_t = j)P_{t|t-1}^{(i,j)}\tilde{Z}(s_t = j)' + \Sigma_\nu. \end{aligned}$$

⁴Detailed derivations will be provided in online appendix.

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

$$p(y_t|\mathcal{F}_{t-1}) = \sum_{j=0}^1 \sum_{i=0}^1 p(y_t|s_{t-1}=i, s_t=j, \mathcal{F}_{t-1}) p_{t|t-1}^{(i,j)},$$

where $p(y_t|s_{t-1}=i, s_t=j, \mathcal{F}_{t-1})$ is a multivariate normal distribution with mean $y_{t|t-1}^{(i,j)}$ and variance $F_{t|t-1}^{(i,j)}$.

(c) **Updating.** First, use the Bayes formula to obtain

$$p_{t|t}^{(i,j)} = \frac{p(y_t|s_{t-1}=i, s_t=j, \mathcal{F}_{t-1}) p_{t|t-1}^{(i,j)}}{p(y_t|\mathcal{F}_{t-1})}$$

and compute $p_{t|t}^j = \sum_{i=0}^1 p_{t|t}^{(i,j)}$. Then, apply the updating step of the Kalman filter to obtain

$$\varsigma_{t|t}^{(i,j)} = \varsigma_{t|t-1}^{(i,j)} + P_{t|t-1}^{(i,j)} \tilde{Z}(s_t=j)' (F_{t|t-1}^{(i,j)})^{-1} (y_t - y_{t|t-1}^{(i,j)}), \quad (2.20)$$

$$P_{t|t}^{(i,j)} = P_{t|t-1}^{(i,j)} - P_{t|t-1}^{(i,j)} \tilde{Z}(s_t=j)' (F_{t|t-1}^{(i,j)})^{-1} \tilde{Z}(s_t=j) P_{t|t-1}^{(i,j)}. \quad (2.21)$$

for $i = 0, 1$ and $j = 0, 1$. Next, I collapse $(\varsigma_{t|t}^{(i,j)}, P_{t|t}^{(i,j)})$ into⁵

$$\varsigma_{t|t}^j = \frac{\sum_{i=0}^1 p_{t|t}^{(i,j)} \varsigma_{t|t}^{(i,j)}}{p_{t|t}^j}, \quad (2.22)$$

$$P_{t|t}^j = \frac{\sum_{i=0}^1 p_{t|t}^{(i,j)} \left\{ P_{t|t}^{(i,j)} + (\varsigma_{t|t}^j - \varsigma_{t|t}^{(i,j)}) (\varsigma_{t|t}^j - \varsigma_{t|t}^{(i,j)})' \right\}}{p_{t|t}^j}, \quad (2.23)$$

from which we can further compute the extracted filtered states as

$$\varsigma_{t|t} = \sum_{j=0}^1 p_{t|t}^j \varsigma_{t|t}^j, \quad (2.24)$$

$$P_{t|t} = \sum_{j=0}^1 p_{t|t}^j \left\{ P_{t|t}^j + (\varsigma_{t|t} - \varsigma_{t|t}^j) (\varsigma_{t|t} - \varsigma_{t|t}^j)' \right\}. \quad (2.25)$$

3. Aggregation. The log-likelihood function can be written as

$$\log p(Y_{1:T}) = \sum_{t=1}^T \log p(y_t|\mathcal{F}_{t-1}). \quad (2.26)$$

⁵ If $p_{t|t}^j = 0$, $(\varsigma_{t|t}^j, P_{t|t}^j)$ in (2.22) and (2.23) are not well defined. In this case, we set $(\varsigma_{t|t}^j, P_{t|t}^j) = (\varsigma_{t|t}^{1-j}, P_{t|t}^{1-j})$.

4. Extraction of latent regime factor. After the parameter estimates are obtained, the filter can be modified to easily extract the latent regime factor w_t . In the forecasting step, note that

$$p(w_t, s_{t-1}, \mu_{t-1} | \mathcal{F}_{t-1}) = p(w_t | s_{t-1}, \mu_{t-1}, \mathcal{F}_{t-1}) p(\mu_{t-1} | s_{t-1}, \mathcal{F}_{t-1}) p(s_{t-1} | \mathcal{F}_{t-1}) \quad (2.27)$$

$$= p(w_t | s_{t-1}, \mu_{t-1}) p(\mu_{t-1} | s_{t-1}, \mathcal{F}_{t-1}) p(s_{t-1} | \mathcal{F}_{t-1}) \quad (2.28)$$

Since $p(\mu_{t-1} | s_{t-1}, \mathcal{F}_{t-1})$ and $p(s_{t-1} | \mathcal{F}_{t-1})$ are both obtained from the previous updating step, we may compute $p(w_t, s_{t-1}, \mu_{t-1} | \mathcal{F}_{t-1})$ once we find $p(w_t | s_{t-1}, \mu_{t-1})$. In Chang, Choi, and Park (2017), they derived that ⁶

$$p(w_t | s_{t-1} = i, \mu_{t-1}) = \begin{cases} \frac{\Phi\left(\sqrt{\frac{1 - \rho'\rho + \alpha^2 \rho'\rho}{1 - \rho'\rho}} \left(\tau - \frac{\alpha(w_t - \mu_{t-1})}{1 - \rho'\rho + \alpha^2 \rho'\rho}\right)\right)}{\Phi(\tau\sqrt{1 - \alpha^2})} \mathbb{N}\left(\mu_{t-1}, \frac{1 - \rho'\rho + \alpha^2 \rho'\rho}{1 - \alpha^2}\right), & i = 0 \\ \frac{1 - \Phi\left(\sqrt{\frac{1 - \rho'\rho + \alpha^2 \rho'\rho}{1 - \rho'\rho}} \left(\tau - \frac{\alpha(w_t - \mu_{t-1})}{1 - \rho'\rho + \alpha^2 \rho'\rho}\right)\right)}{1 - \Phi(\tau\sqrt{1 - \alpha^2})} \mathbb{N}\left(\mu_{t-1}, \frac{1 - \rho'\rho + \alpha^2 \rho'\rho}{1 - \alpha^2}\right), & i = 1 \end{cases} \quad (2.29)$$

We may then obtain

$$p(w_t, s_{t-1}, \mu_{t-1} | \mathcal{F}_t) = \frac{p(y_t | w_t, s_{t-1}, \mu_{t-1}, \mathcal{F}_{t-1}) p(w_t, s_{t-1}, \mu_{t-1} | \mathcal{F}_{t-1})}{p(y_t | \mathcal{F}_{t-1})} \quad (2.30)$$

in the updating step. Note that

$$p(y_t | w_t, s_{t-1}, \mu_{t-1}, \mathcal{F}_{t-1}) = p(y_t | s_t = j, s_{t-1}, \mathcal{F}_{t-1}), \quad (2.31)$$

for $j = 1\{w_t \geq \tau\}$, and $p(y_t | \mathcal{F}_{t-1})$ are both obtained in the likelihood evaluation step. By marginalizing $p(w_t, s_{t-1}, \mu_{t-1} | \mathcal{F}_t)$, we get

$$p(w_t | \mathcal{F}_t) = \sum_{s_{t-1}=0}^1 \int_{\mathbb{R}} p(w_t, s_{t-1}, \mu_{t-1} | \mathcal{F}_t) d\mu_{t-1} \quad (2.32)$$

⁶ Detailed derivations can be found in Appendix A Proof of Corollary 3.3 of Chang, Choi, and Park (2017).

Finally, the latent regime factor w_t can be extracted as

$$\mathbb{E}(w_t|\mathcal{F}_t) = \int_{\mathbb{R}} w_t p(w_t|\mathcal{F}_t) dw_t \quad (2.33)$$

for all $t = 1, \dots, T$. Both (2.32) and (2.33) can be computed via numerical integration.

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 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 (2.22) and (2.23) 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 regime factor $w_{t|t}$ can be easily extracted as a useful by-product of running the filter.

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

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

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

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, $(\sigma_u, \sigma_\epsilon) = (0.2, 0.5)$, and $(\alpha, \tau, \rho) = (0.7, -0.5, 0.9)$.⁷ 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, Salmond, and Smith (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.⁸ Our particle filtering algorithm is implemented through the following steps.

Algorithm 2. (Endogenous-Switching Particle Filter)

1. Initialization. For $j = 0, 1$, 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 $p_{0|0}^0 \cdot \mathbb{N}(x_{0|0}^0, P_{0|0}^0) + p_{0|0}^1 \cdot \mathbb{N}(x_{0|0}^1, P_{0|0}^1)$, $\epsilon_0^i \sim \mathbb{N}(0, I_n)$, 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

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

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

$$x_t^i = G(s_t^i) x_{t-1}^i + M(s_t^i) \Sigma_\epsilon^{1/2} \epsilon_t^i, \quad \epsilon_t^i \sim \mathbb{N}(0, I_n) \quad (3.4)$$

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

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

where $p(y_t | s_t^i, x_t^i)$ can be evaluated from the measurement equation.

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

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

⁷ A small measurement error is included so that the data density, conditioned on the states, remains non-degenerate in the particle filter.

⁸ A complete tutorial on basic and advanced particle filtering methods can be found in Doucet and Johansen (2009). See also DeJong and Dave (2011) and Herbst and Schorfheide (2015) for textbook treatments of the particle filter in DSGE applications.

Figure 2: Filtered State Variables and 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 (-85.65 and -85.02 in aggregate, respectively).

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

$$\mathbb{E}[f(x_t, \epsilon_t, w_t)|Y_{1:t}] = \text{plim}_{N \rightarrow \infty} \sum_{i=1}^N W_t^i f(x_t^i, \epsilon_t^i, w_t^i) \quad (3.7)$$

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

Practically it turns out that $N = 100,000$ particles are more than sufficient to accurately approximate the density $p(Y_{1:T})$ implied 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

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.

filter performs fairly well as it delivers barely distinguishable 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.⁹ In addition, the log likelihood function of each individual parameter peaks in the immediate vicinity of its true value. Taken together, our filter ensures the overall likelihood surface is well preserved.

⁹This 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.

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 the nominal interest rate R_t as well as a fiscal authority that consumes g_t and passively adjusts lump-sum taxes to ensure its budgetary solvency; a labor-augmenting technology that induces a stochastic trend in output y_t . Let $\hat{x}_t \equiv \ln(x_t) - \ln(x^*)$ denote the log-deviation of a variable x_t from its steady state x^* . In the fixed-regime benchmark, a linear approximation to the models equilibrium conditions in terms of the detrended variables is summarized as follows.

First, the households optimizing behavior, when imposed by the goods market clearing condition, implies

$$\hat{y}_t = \mathbb{E}_t[\hat{y}_{t+1}] + \hat{g}_t - \mathbb{E}_t[\hat{g}_{t+1}] - \frac{1}{\phi}(\hat{R}_t - \mathbb{E}_t[\hat{\pi}_{t+1}] - \mathbb{E}_t[\hat{z}_{t+1}]) \quad (3.8)$$

where $\phi > 0$ is the relative risk aversion, π_t is the inflation between periods $t-1$ and t , z_t is an exogenous shock to the labor-augmenting technology that grows on average at the rate γ , and $\mathbb{E}$ represents mathematical expectation given information available at time t . The firms optimal price-setting behavior reduces to

$$\hat{\pi}_t = \beta \mathbb{E}_t[\hat{\pi}_{t+1}] + \kappa(\hat{y}_t - \hat{g}_t) \quad (3.9)$$

where $0 < \beta < 1$ is the discount factor and $\kappa > 0$ is the slope of the so-called new Keynesian Phillips curve. Second, 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

$$\hat{R}_t = \rho_R \hat{R}_{t-1} + (1 - \rho_R)\psi_\pi \hat{\pi}_t + (1 - \rho_R)\psi_y(\hat{y}_t - \hat{g}_t) + \epsilon_{R,t} \quad (3.10)$$

where $0 \leq \rho_R < 1$ determines the degree of interest rate smoothing, $\psi_\pi > 0$ and $\psi_y > 0$ are the policy rate responsive coefficients, and $\epsilon_{R,t}$ is an exogenous policy shock. Finally, both g_t and z_t evolve as autoregressive processes

$$\hat{g}_t = \rho_g \hat{g}_{t-1} + \epsilon_{g,t}, \quad \hat{z}_t = \rho_z \hat{z}_{t-1} + \epsilon_{z,t} \quad (3.11)$$

where $0 \leq \rho_g, \rho_z < 1$. 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. Subsequently, we solve the linear rational expectations system (3.8)–(3.11) in the state variables $x_t = [\hat{y}_t, \hat{\pi}_t, \hat{R}_t, \hat{g}_t, \hat{z}_t]'$ using

the procedure of Sims (2002).

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.¹⁰ In the conventional Markov switching case, the response of policy rate to inflation deviations is switching between, in Leeper (1991) terminology, more active and less active (or possibly passive) monetary regimes

$$\psi_\pi(s_t) = \psi_\pi^0(1 - s_t) + \psi_\pi^1 s_t, \quad 0 \leq \psi_\pi^0 < \psi_\pi^1 \quad (3.12)$$

where the regime index s_t evolves exogenously according to $p_{ij} = \mathbb{P}(s_t = j | s_{t-1} = i)$ for $i, j = 0, 1$ and $\sum_j p_{ij} = 1$. On the other hand, when $s_t = \mathbb{1}\{w_t \geq \tau\}$ so that the transition probabilities become an important part of the model solution, we assume the endogeneity in regime changes arises solely from the policy shock, $\text{corr}(\epsilon_{g,t}, v_{t+1}) = \text{corr}(\epsilon_{z,t}, v_{t+1}) = 0$, and a surprise monetary contraction tends to reinforce agents belief of staying in or switching to the more active regime in the future, $\text{corr}(\epsilon_{R,t}, v_{t+1}) = \rho \in [0, 1]$, which is motivated by the empirical evidence alluded to earlier. Interested readers are referred to Chang, Tan, and Wei (2018) for a more comprehensive investigation of the macroeconomic origins that give rise to monetary policy shifts. In both cases we abstract from sources of time variation in the model structure other than that from the policy parameter $\psi_\pi(s_t)$.¹¹ Note that the presence of s_t poses keen computational challenges to solving the model. In what follows, we obtain the first-order solution using the perturbation method of Barthélemy and Marx (2017) that is well-suited for solving regime-switching rational expectations models with state-dependent transition probabilities.¹² An important precursor to our study is Davig and Leeper (2006), 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.

Both models 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 is constructed as in Appendix B of Herbst and Schorfheide (2015), available from the Federal Reserve Economic Data (FRED). The observable variables are linked to the model variables through the following

¹⁰Our paper complements recent likelihood-based estimation of DSGE models with exogenous Markov switching in monetary policy, including Schorfheide (2005), Bianchi (2012), 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 and Kwak (2017) also estimated a reduced-form model of monetary-fiscal regime changes.

¹¹Initiated 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 would require introducing multiple latent factors to operationalize in our setup, which is beyond the scope of this paper.

¹²Chang, Tan, and Wei (2018) compute the time-varying transition probabilities implied by the threshold-type regime switching adopted in this paper, which is a prerequisite to applying the solution technique of Barthélemy and Marx (2017).

measurement equations

$$\begin{pmatrix} YGR_t \\ INF_t \\ INT_t \end{pmatrix} = \begin{pmatrix} \gamma^{(Q)} \\ \pi^{(A)} \\ 4\gamma^{(Q)} + r^{(A)} + \pi^{(A)} \end{pmatrix} + 100 \begin{pmatrix} \hat{y}_t - \hat{y}_{t-1} + \hat{z}_t \\ 4\hat{\pi}_t \\ 4\hat{R}_t \end{pmatrix} \quad (3.13)$$

where $(\gamma^{(Q)}, \pi^{(A)}, r^{(A)})$ are connected to the models 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, exogenous and endogenous switching, this can be cast into the state space form (2.1)–(2.2) and evaluated with the standard Kalman filter, the exogenous switching filter of Kim (1994), and our endogenous switching filter in Algorithm 1, respectively.

Table 1 summarizes the marginal prior distributions on the DSGE structural parameters. For 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 parameters (ϕ, κ) 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 a fair prior on τ , implies transition probabilities comparable to (p_{00}, p_{11}) under exogenous switching. On the other hand, the uniform distribution on ρ reflects an agnostic prior view about the degree of endogeneity in regime switching.

For each model, we sample a total of 550,000 draws from the posterior distribution using the random walk Metropolis-Hastings algorithm, discard the first 50,000 draws as burn-in phase, and keep one every 50 draws afterwards.¹³ The resulting 10,000 draws form the basis for the posterior inference. We highlight several aspects of the posterior estimates reported in Tables 2 as follows.

First, regarding the parameters shared by all models, allowing for regime switching does generate material impacts on some of their estimates, including the private sector parameters (ϕ, κ) , the steady state annual inflation $\pi^{(A)}$, as well as the government spending shock process (ρ_g, σ_g) , while others remain almost recalcitrant to that. Second, turning to the switching parameters, introducing endogeneity into regime switching seem to produce cross-equation restrictions that are

¹³For instance, for the regime switching model, the acceptance rate of 48%, in conjunction with low inefficiency factors (ranging from 7.5 to 21.8 with most below 15) suggest that the Markov chain converges well. Other 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 et al. (1991). See Herbst and Schorfheide (2015) for a detailed textbook treatment of Bayesian estimation of DSGE models.

Table 1: Prior Distributions of DSGE Parameters

Description	Parameter	Density	Para (1)	Para (2)
Fixed Regime				
relative risk aversion	ϵ	G	2.00	0.50
slope of the NKPC	κ	G	0.20	0.10
response to inflation	ψ_π	G	1.50	0.25
response to output gap	ψ_y	G	0.50	0.25
inflation target (annualized)	$\pi^{(A)}$	G	3.84	2.00
real interest rate (annualized)	$r^{(A)}$	G	0.50	0.10
technology growth rate (quarterly)	$\gamma^{(Q)}$	N	0.47	0.20
persistence of technology shock	ρ_a	B	0.50	0.10
persistence of government spending shock	ρ_g	B	0.50	0.10
persistence of monetary policy shock	ρ_R	B	0.50	0.10
scaled s.d. of technology shock	$100\sigma_a$	I	4.00	0.40
scaled s.d. of government spending shock	$100\sigma_g$	I	4.00	0.40
scaled s.d. of monetary policy shock	$100\sigma_R$	I	4.00	0.40
Threshold Switching				
response to inflation under regime-0	ψ_π^0	G	1.00	0.10
response to inflation under regime-1	ψ_π^1	G	2.00	0.25
persistence of regime factor	α	B	0.90	0.05
threshold level	τ	N	0.00	0.50
endogeneity from monetary policy shock	ρ_{Rv}	B	0.50	0.10

Notes: Para (1) and Para (2) are the means and standard deviations for Gamma, Beta and Normal distributions, the left side and right side for Uniform distribution, and the parameters ν and s for Inverse Gamma distribution with density $p(\sigma|\nu, s) \propto \sigma^{-\nu-1} \exp(-\nu s^2/2\sigma^2)$.

useful for identifying policy regimes compared to the Markov switching case, the posterior mean and 90% interval of the more (less) active policy response to inflation $\psi_\pi^1(\psi_\pi^0)$ are shifted upward (downward). Third, moving on to the parameters (α, τ, ρ) unique to endogenous switching, the posterior distributions suggest an even more persistent process for the regime factor relative to its prior, a longer expected duration of the more active regime, and attribute a significant portion (about 13% – 45%) of regime changes to past monetary interventions. The contemporaneous external shock, on the other hand, not only drives the rest of regime changes, but also behaves like a monetary policy shock outside of the model that, as evinced by the estimate of f_π , has a contractionary effect on inflation. Finally, the estimated Bayes factor, the ratio of marginal likelihood between two models, approximates to e^9 , indicating that the data decisively favor the regime switching model over the fixed regime model.¹⁴ This justifies the importance of accounting

¹⁴Log marginal likelihoods are approximated using the modified harmonic mean estimator of Geweke (1999) with a truncation parameter of 0.9. According to Kass and Raftery (1995) criterion, a Bayes factor greater than 100 signifies decisive evidence in favor of the model with superior fit.

Table 2: Posterior Estimates of DSGE Parameters

Parameters	No Switching Model			Regime Switching Model		
	Mode	Mean	[5%, 95%]	Mode	Mean	[5%, 95%]
Fixed Regime						
ϵ	4.14	3.74	[2.84, 4.72]	2.84	2.89	[2.05, 3.85]
κ	0.27	0.37	[0.27, 0.48]	0.41	0.48	[0.30, 0.70]
ψ_π	1.01	1.03	[1.00, 1.05]	—	—	—
ψ_y	0.39	0.28	[0.11, 0.45]	0.23	0.36	[0.13, 0.67]
$\pi^{(A)}$	0.32	3.21	[2.32, 4.08]	2.52	2.71	[2.06, 3.31]
$r^{(A)}$	0.58	0.49	[0.34, 0.65]	0.50	0.50	[0.36, 0.65]
$\gamma^{(Q)}$	0.13	0.36	[0.25, 0.45]	0.29	0.32	[0.20, 0.44]
ρ_a	0.95	0.93	[0.92, 0.95]	0.94	0.94	[0.92, 0.95]
ρ_g	0.97	0.97	[0.96, 0.98]	0.97	0.97	[0.96, 0.98]
ρ_R	0.76	0.73	[0.69, 0.77]	0.72	0.72	[0.66, 0.77]
$100\sigma_a$	0.12	0.13	[0.11, 0.16]	0.13	0.13	[0.11, 0.15]
$100\sigma_g$	1.07	1.10	[1.01, 1.20]	1.02	1.04	[0.95, 1.14]
$100\sigma_R$	0.25	0.26	[0.23, 0.28]	0.25	0.26	[0.22, 0.30]
Threshold Switching						
ψ_π^0	—	—	—	0.89	0.88	[0.78, 0.98]
ψ_π^1	—	—	—	2.01	2.05	[1.75, 2.37]
α	—	—	—	0.99	0.99	[0.98, 1.00]
τ	—	—	—	0.00	-0.16	[-0.89, 0.61]
ρ_{Rv}	—	—	—	0.50	0.51	[0.35, 0.67]
Log marginal likelihood	-1126.73 (0.02)			-1117.89 (0.00)		

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

for the endogenous feedback from historical monetary interventions to the prevailing policy regime.

Unlike the Markov switching filter of Kim (1994), our endogenous switching filter also produces an important by-product—an extracted time series of the autoregressive regime factor $w_{t|t}$ as depicted in Figure 4—that conveys richer information than its implied probability $p_{t|t}^1$ of being in the more active regime, most noticeably the regime strength. 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

Figure 4: 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.

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 from the transition equation that represents the fundamental shocks inside

the model and an external innovation that captures all other shocks left outside of the model. This allows for an endogenous feedback from the behavior of underlying economic fundamentals to the regime generating process. When regime shifts are purely driven by the external disturbance, the model reduces to one with exogenous switching.

We develop an endogenous switching version of the Kalman filter to estimate the overall nonlinear state space model. This 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 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 regime switching in monetary policy, and find decisive statistical support for the endogenous feedback from historical monetary interventions to the prevailing policy regime. A natural extension of our framework is to allow for 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, Fall 2017 Midwest Macroeconomics Meeting at University of Pittsburg, 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, 13th Annual Jordan River Economics Conference at Indiana University, 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, 2018 International Association for Applied Econometrics Annual Conference at Montreal, 2018 Taiwan Econometric Society Annual Meeting, the 27th Annual Symposium of the Society for Nonlinear Dynamics and Econometrics at Dallas, and the 27th Annual Meeting of the Midwest Econometrics Group at Texas A&M University, as well as Francesco Bianchi, Yongok Choi, Andrew T. Foester, Grey Gordon, Luca Guerrieri, Edward Herbst, Eric Leeper, Chaojun Li, Junior Maih, Christian Matthes, Joon Park, Shi Qiu, Daniel Waggoner, Todd B. Walker, and Tao Zha for very useful and stimulating comments. Fei Tan acknowledges financial support from the Chaifetz School of Business summer research grant.

References

- S. An and F. Schorfheide. Bayesian analysis of DSGE models. *Econometric Reviews*, 26(2-4): 113–172, 2007.
- J. Barthélemy and M. Marx. Solving endogenous regime switching models. *Journal of Economic Dynamics and Control*, 77:1–25, 2017.
- M. Bazzi, F. Blasques, S. J. Koopman, and A. Lucas. Time-varying transition probabilities for Markov regime switching models. *Journal of Time Series Analysis*, 38(3):458–478, 2017.
- F. Bianchi. Regime switches, agents’ beliefs, and post-World War II US macroeconomic dynamics. *Review of Economic Studies*, 80(2):463–490, 2012.
- F. Bianchi and C. Ilut. Monetary/fiscal policy mix and agents’ beliefs. *Review of Economic Dynamics*, 26:113–139, 2017.
- F. Bianchi and L. Melosi. Escaping the great recession. *American Economic Review*, 107(4): 1030–58, 2017.
- Y. Chang and B. Kwak. Endogenous monetary-fiscal regime change in the United States. *Working Paper, Indiana University*, 2017.
- Y. Chang, Y. Choi, and J. Y. Park. A new approach to model regime switching. *Journal of Econometrics*, 196(1):127–143, 2017.
- Y. Chang, F. Tan, and X. Wei. A structural investigation of monetary policy shifts. *Manuscript, Indiana University*, 2018.
- R. Chen and J. S. Liu. Mixture kalman filters. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(3):493–508, 2000.
- S. Chib. Calculating posterior distributions and modal estimates in Markov mixture models. *Journal of Econometrics*, 75(1):79–97, 1996.
- R. Clarida, J. Galí, and M. Gertler. Monetary policy rules and macroeconomic stability: Evidence and some theory. *The Quarterly Journal of Economics*, 115(1):147–180, 2000.
- T. Davig and T. Doh. Monetary policy regime shifts and inflation persistence. *Review of Economics and Statistics*, 96(5):862–875, 2014.
- T. Davig and E. M. Leeper. Endogenous monetary policy regime change. In *NBER International Seminar on Macroeconomics 2006*, pages 345–377. MIT Press, Cambridge, MA, 2006.

- T. Davig and E. M. Leeper. Fluctuating macro policies and the fiscal theory. In *NBER Macroeconomics Annual 2006, Volume 21*, pages 247–316. MIT Press, Cambridge, MA, 2007.
- D. N. DeJong and C. Dave. *Structural macroeconometrics*. Princeton University Press, 2011.
- F. X. Diebold, J.-H. Lee, and G. C. Weinbach. Regime switching with time-varying transition probabilities. *Business Cycles: Durations, Dynamics, and Forecasting*, pages 144–165, 1994.
- A. Doucet and A. M. Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of Nonlinear Filtering*, 12(656-704):3, 2009.
- J. Geweke. Using simulation methods for Bayesian econometric models: Inference, development, and communication. *Econometric Reviews*, 18(1):1–73, 1999.
- J. Geweke et al. *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*, volume 196. Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, 1991.
- N. J. Gordon, D. J. Salmond, and A. F. Smith. Novel approach to nonlinear/non-Gaussian bayesian state estimation. In *IEE proceedings F (radar and signal processing)*, volume 140, pages 107–113. IET, 1993.
- J. D. Hamilton. Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. *Journal of Economic Dynamics and Control*, 12(2-3): 385–423, 1988.
- J. D. Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384, 1989.
- E. P. Herbst and F. Schorfheide. *Bayesian estimation of DSGE models*. Princeton University Press, 2015.
- L. Kalliovirta, M. Meitz, and P. Saikkonen. A Gaussian mixture autoregressive model for univariate time series. *Journal of Time Series Analysis*, 36(2):247–266, 2015.
- K. H. Kang. Estimation of state-space models with endogenous markov regime-switching parameters. *The Econometrics Journal*, 17(1):56–82, 2014.
- R. E. Kass and A. E. Raftery. Bayes factors. *Journal of the American Statistical Association*, 90 (430):773–795, 1995.
- C.-J. Kim. Dynamic linear models with Markov-switching. *Journal of Econometrics*, 60(1-2):1–22, 1994.

- C.-J. Kim. Markov-switching models with endogenous explanatory variables. *Journal of Econometrics*, 122(1):127–136, 2004.
- C.-J. Kim. Markov-switching models with endogenous explanatory variables II: A two-step MLE procedure. *Journal of Econometrics*, 148(1):46–55, 2009.
- C.-J. Kim and C. R. Nelson. *State-space models with regime switching: Classical and Gibbs-sampling approaches with applications*, volume 2. MIT Press, Cambridge, MA, 1999.
- C.-J. Kim, J. Piger, and R. Startz. Estimation of Markov regime-switching regression models with endogenous switching. *Journal of Econometrics*, 143(2):263–273, 2008.
- E. M. Leeper. Equilibria under ‘active’ and ‘passive’ monetary and fiscal policies. *Journal of Monetary Economics*, 27(1):129–147, 1991.
- C. D. Romer and D. H. Romer. Choosing the federal reserve chair: lessons from history. *Journal of Economic Perspectives*, 18(1):129–162, 2004.
- F. Schorfheide. Learning and monetary policy shifts. *Review of Economic Dynamics*, 8(2):392–419, 2005.
- C. A. Sims. Solving linear rational expectations models. *Computational Economics*, 20(1):1–20, 2002.
- C. A. Sims and T. Zha. Were there regime switches in US monetary policy? *The American Economic Review*, 96(1):54–81, 2006.
- M. A. Tanner and W. H. Wong. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398):528–540, 1987.
- J. B. Taylor. Discretion versus policy rules in practice. In *Carnegie-Rochester conference series on public policy*, volume 39, pages 195–214. North-Holland, Amsterdam, 1993.
- G. C. Wei and M. A. Tanner. 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, 1990.