State Space Models with Endogenous Regime Change: Estimation and DSGE Application†

Yoosoon Chang, Junior Maih, and Fei Tan*

[This Version: September 8, 2019]

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 historical 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 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 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; endogeneity; filtering; DSGE model.

JEL Classification: E52, C13, C32

†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. An earlier draft of the paper was circulated under the title “State Space Models with Endogenous Regime Switching”.

*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: yoosoon@indiana.edu (Y. Chang).
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 endemic to many 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 of the Markov switching approach, its maintained hypothesis that regime evolves exogenously and thereby falls completely apart from the rest of the model is rather dubious in many cases. As argued convincingly in Chang et al. (2017a), 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 desirable 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), Chib and Dueker (2004), Kim (2004, 2009), Kim et al. (2008), Bazzi et al. (2014), Kang (2014), Kalliovirta et al. (2015), Kim and Kim (2018), among others].

The purpose of this paper is to introduce a threshold-type endogenous regime switching into dynamic linear models that can be represented in state space forms. This class of models is 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. (2017a), 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 (solid line) and the rate implied by an inertial version of the Taylor (1993) rule (dashed line), \( i_t = \rho i_{t-1} + (1 - \rho)[4 + 1.5(\pi_t - 2) + 0.5y_t] \), where \( i \) denotes the federal funds rate, \( \pi \) the annual inflation rate, \( y \) the percentage deviation of real output from its potential, and \( \rho = 0.75 \). 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 algorithms of Kim (1994) and Chang et al. (2017a) 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 as such makes our filter computationally efficient. As a useful by-product of running the filter, the estimated autoregressive latent factor can be readily constructed from the filter outputs. 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 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 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 tightenings 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. Our paper takes a first step toward filling in this important gap. We first employ a Fisherian model of inflation determination to endogenize regime change in monetary policy. The specification is simple enough to admit an analytical solution that makes transparent the mechanism at work, but also rich enough to highlight the general features of a rational expectations model with threshold-type endogenous switching. We then extend the simple model to a prototypical new Keynesian DSGE model whose state space form can be analyzed with our filter, and find that price markup shocks played a predominant role in triggering the historical regime changes in the postwar 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 organized as follows. Section 2 describes the state space model and filtering algorithm. Section 3 adopts a simple analytical model to illustrate how to endogenize regime change in monetary policy using our endogenous-switching framework. Section 4 extends the simple model to an empirical DSGE model and derives its state space form that can be analyzed with our filtering algorithm. Section 5 concludes. We also employ the following notation. Let $N(\mu, \Sigma)$ denote the normal distribution with mean vector $\mu$ and covariance matrix $\Sigma$, $p_{N}(\cdot | \mu, \Sigma)$ its probability density function, and $\Phi(\cdot)$ the cumulative distribution function of $N(0, 1)$. In particular, $N(0_{n \times 1}, I_{n})$ 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 $t = 1, \ldots, T$ with row observations $y_{t}$.

---

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.”
2 Model and Algorithm

This section introduces the threshold-type endogenous switching framework of Chang et al. (2017a), 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. Without using 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. We also document the numerical performance of our filter against the particle filter using a small state space model with simulated data.

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

\begin{align}
  y_t &= D_{s_t} z_t + Z_{s_t} x_t + F_{s_t} z_t + \Omega_{s_t}^{1/2} u_t, \quad u_t \sim \mathcal{N}(0_{l \times 1}, I_l) \\
  x_t &= C_{s_t} z_t + G_{s_t} x_{t-1} + E_{s_t} z_t + M_{s_t} \Sigma_{s_t}^{1/2} \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0_{n \times 1}, I_n)
\end{align}

(2.1) (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 $\Omega_{s_t}^{1/2} u_t$, the transition equation (2.2) describes the evolution of the state variables driven by a $n \times 1$ vector of exogenous innovations $\Sigma_{s_t}^{1/2} \epsilon_t$, and $(u_t, \epsilon_t)$ are mutually and serially uncorrelated at all leads and lags. The coefficient matrices $(D_{s_t}, Z_{s_t}, F_{s_t}, C_{s_t}, G_{s_t}, E_{s_t}, M_{s_t})$ and the covariance matrices $(\Omega_{s_t}, \Sigma_{s_t})$ are both allowed to depend on an index variable $s_t = 1\{w_t \geq \tau\}$ driven by a stationary autoregressive latent factor

\begin{align}
  w_t &= \alpha w_{t-1} + \nu_t, \quad \nu_t \sim \mathcal{N}(0, 1)
\end{align}

(2.3)

where $-1 < \alpha < 1$ controls the persistency of $w_t$. 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 $\tau$. In what follows, we call $w_t$ the regime factor.

Endogeneity in regime switching is introduced as in Chang et al. (2017a), but we allow all...
current standardized transition innovations $\epsilon_t$ to jointly influence the next period regime through their correlations with the innovation $v_{t+1}$ to $w_{t+1}$. To wit,

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

(2.4)

where $\rho = [\rho_1, \ldots, \rho_n]' = \text{corr}(\epsilon_t, v_{t+1})$ is a vector of correlation parameters that determines the degree of endogeneity in regime changes—as $\rho$ 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 transition probabilities time-varying because they are all functions of $\epsilon_t$ as will be shown subsequently. In the special case where $\epsilon_t$ and $v_{t+1}$ are orthogonal (i.e., $\rho = 0_{n \times 1}$), the transition probabilities become constants and the model reduces to one with conventional Markov switching; in fact, there exists a one-to-one correspondence between our threshold-type switching specified by $(\alpha, \tau)$ and the Markov switching specified by two transition probabilities [see Chang et al. (2017a), Lemma 2.1].

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

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

(2.5)

where $\{\epsilon_{k,t}\}_{k=1}^{n}$ and the idiosyncratic innovation $\eta_{t+1}$ are all orthogonal and have unit variance. On the surface, the residual $\eta_{t+1}$ of projecting $v_{t+1}$ onto $\epsilon_t$ appears to be a vague source of regime change in many economic applications where $\epsilon_t$ 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 $\epsilon_t$ under the ‘true’ model—that leads to systematic disparities between model-implied and actual observables. To the extent that $\eta_{t+1}$ picks up those missing components beyond what are incorporated in $\epsilon_t$, we may readily call $\epsilon_{k,t}$ and $\eta_t$ 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 (2.5), we have the conditional variance

$$
\text{Var}_t(w_{t+h}) = \sum_{k=1}^{n} \rho_k^2 \sum_{j=1}^{h} \alpha^{2(h-j)} + \left(1 - \sum_{k=1}^{n} \rho_k^2\right) \sum_{j=1}^{h} \alpha^{2(h-j)} = \sum_{j=1}^{h} \alpha^{2(h-j)}, \quad h \geq 1
$$

(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 $\rho_k^2$ (or $1 - \sum_{k=1}^{n} \rho_k^2$), which is
independent of $h$. Letting $h \to \infty$, $\rho_k^2$ (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 4 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 through 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 two mixture Gaussian distributions in each period. This reduction effectively breaks the full history-dependence of the likelihood function and therefore makes the computation feasible and 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 $z_t = [x_t', d_t']'$. Accordingly, we rewrite the measurement and transition equations as

\[
\begin{align*}
    y_t &= \begin{pmatrix} D_{st} + F_{st} & Z_{st} \\ D_{st} & 0_{l \times n} \end{pmatrix} \begin{pmatrix} x_t \\ d_t \end{pmatrix} + \Omega_{st}^{1/2} u_t \\
    \begin{pmatrix} x_t \\ d_t \end{pmatrix} &= \begin{pmatrix} C_{st} + E_{st} & G_{st} \\ 0_{n \times 1} & 0_{n \times m} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ d_{t-1} \end{pmatrix} + \begin{pmatrix} M_{st} \Sigma_{st}^{1/2} \\ I_n \end{pmatrix} \epsilon_t
\end{align*}
\]

(2.7) \hspace{1cm} (2.8)

where the dependence of $(\tilde{D}_{st}, \tilde{C}_{st})$ 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), tracks the regime indices of both the current period and its preceding period in each recursion. At an exponentially rising computation cost though, one may improve the approximation by tracking even earlier regime history 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 \( p^{(i,j)}_{t|t-1} = \mathbb{P}(s_{t-1} = i, s_t = j|\mathcal{F}_{t-1}) \) and the filtered marginal probability of regime-\( j \) at period \( t \) as \( p^j_t = \mathbb{P}(s_t = j|\mathcal{F}_t) \). Also define a battery of four conditional forecasts of \( \varsigma_t \) and their forecast error covariances as

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

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 \( s_0 \), \( (\varsigma^0_0, P^0_0) \), using the invariant distribution under regime-\( i \). Also set \( p^0_0 = \Phi(\tau \sqrt{1 - \alpha^2}) \)
and \( p^j_0 = 1 - p^0_0 \) according to the invariant distribution of \( w_t \), i.e., \( \mathbb{N}(0, 1/(1 - \alpha^2)) \).

2. **Recursion.** For \( t = 1, \ldots, T \), the filter accepts two sets of triple inputs \( \{(\varsigma^i_{t-1|t-1}, P^i_{t-1|t-1}, p^i_{t-1|t-1})\}_{i=0}^1 \), invokes the one-step Kalman filter to calculate the required integrals conditional on four possible mixes of the regimes in the current period and its preceding period, and returns two sets of updated triple outputs \( \{(\varsigma^i_t, P^i_t, p^i_t)\}_{i=0}^1 \).

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

\[
\begin{align*}
\varsigma^{(i,j)}_{t|t-1} &= \tilde{C}_j + \tilde{G}_j \varsigma^i_{t-1|t-1} \\
P^{(i,j)}_{t|t-1} &= \tilde{C}_j P^i_{t-1|t-1} \tilde{C}_j' + \tilde{M}_j \tilde{M}_j'
\end{align*}
\]

for \( i = 0, 1 \) and \( j = 0, 1 \). Next, define \( \lambda_t = \rho' \epsilon_t \) and compute the predictive joint probabilities

\[
p^{(0,0)}_{t|t-1} = \mathbb{P}(s_t = 0|s_{t-1} = 0, \mathcal{F}_{t-1}) \mathbb{P}(s_{t-1} = 0|\mathcal{F}_{t-1})
\]

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

and \( p^{(0,1)}_{t|t-1} = p^0_{t-1|t-1} - p^{(0,0)}_{t|t-1} \). To evaluate the integral in (2.11), note that the predictive transition probability of remaining in regime-0 between periods \( t - 1 \) and \( t \) can be
computed as

\[ \mathbb{P}(s_t = 0|s_{t-1} = 0, \lambda_{t-1}, \mathcal{F}_{t-1}) = \mathbb{P}(s_t = 0|s_{t-1} = 0, \lambda_{t-1}) = \int_{-\infty}^{\tau\sqrt{1-\alpha^2}} \frac{\Phi((\tau - \alpha x)/\sqrt{1-\alpha^2} - \lambda_{t-1})p_N(x|0,1)dx}{\Phi(\tau\sqrt{1-\alpha^2})} \]

where \( \Phi(x) = \Phi(x/\sqrt{1-\rho'\rho}) \), the first 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 the second equality will be derived in Section 3.2. Clearly, the transition probability \( \mathbb{P}(s_t = 0|s_{t-1} = 0, \lambda_{t-1}) \) depends on the value of \( \lambda_{t-1} \) and hence \( \epsilon_{t-1} \) but becomes a constant when \( \rho = 0_{n\times1} \). Moreover, we approximate

\[ p(\lambda_{t-1}|s_{t-1} = 0, \mathcal{F}_{t-1}) \approx p_N(\lambda_{t-1}|\rho'\xi_{d,t-1|t-1}^0, \rho'P_{d,t-1|t-1}^0) \]

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

\[ p_{t|t-1}^{(0,0)} = \frac{p_{t-1|t-1}^{(0,0)}}{\Phi(\tau\sqrt{1-\alpha^2})} \int_{-\infty}^{\tau\sqrt{1-\alpha^2}} \int_{-\infty}^{\tau\sqrt{1-\alpha^2}} p_N(x,y|\mu_0, \Sigma_0)dx dy \tag{2.12} \]

where

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

Similarly, we can compute

\[ p_{t|t-1}^{(1,0)} = \frac{p_{t-1|t-1}^{(1,0)}}{1 - \Phi(\tau\sqrt{1-\alpha^2})} \int_{-\infty}^{\tau\sqrt{1-\alpha^2}} \int_{-\infty}^{\tau\sqrt{1-\alpha^2}} p_N(x,y|\mu_1, \Sigma_1)dx dy \tag{2.13} \]

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

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

Finally, the integrals in (2.12)–(2.13) can be easily evaluated using the cumulative bivariate normal distribution function. Formulas for calculating all transition proba-
bilities considered in this paper appear in the Online Appendix.

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

\[
y_{t|t-1}^{(i,j)} = \tilde{D}_j + \tilde{Z}_j s_{t|t-1}^{(i,j)}
\]

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

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

\[
p(y_t|\mathcal{F}_{t-1}) = \sum_{j=0}^{1} \sum_{i=0}^{1} p_{t|t-1}(y_t^{(i,j)}|y_{t|t-1}^{(i,j)}, F_{t|t-1}^{(i,j)})p_{t|t-1}^{(i,j)}
\]

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

\[
p_{t|t}^{(i,j)} = \frac{p_{t|t-1}(y_t^{(i,j)}|y_{t|t-1}^{(i,j)}, F_{t|t-1}^{(i,j)})p_{t|t-1}^{(i,j)}}{p(y_t|\mathcal{F}_{t-1})}
\]

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 the updated conditional forecasts of \(\varsigma_t\) and their forecast error covariances

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

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

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 \((\varsigma_{t|t}^{(i,j)}, P_{t|t}^{(i,j)})\) into\(^4\)

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

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

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

\(^4\)If \(p_{t|t}^j = 0\), 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 \((\varsigma_{t|t}^j, P_{t|t}^j) = (\varsigma_{t|t}^{1-j}, P_{t|t}^{1-j})\).
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 approach 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 $E[\varsigma_t | s_{t-1} = i, s_t = j, \mathcal{F}_t]$ 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. Using the stored values of $\varsigma_{d,t-1|t-1}^i$, $P_{d,t-1|t-1}^i$, $p_n(y_t | y_{t,t-1}^{(i,j)}, F_{t|t-1}^{(i,j)})$, and $p(y_t | \mathcal{F}_{t-1})$, it is straightforward to evaluate

$$p(w_t | \mathcal{F}_t) = \frac{1}{\tau} \int_{-\infty}^{\infty} p(w_t, s_{t-1} = i, \lambda_{t-1} | \mathcal{F}_t) d\lambda_{t-1}$$

$$= \frac{1}{\tau} \int_{-\infty}^{\infty} \frac{p(y_t | w_t, s_{t-1} = i, \mathcal{F}_t) p_{t|t-1}^{(i,j)}}{p(y_t | \mathcal{F}_{t-1})} \int_{-\infty}^{\infty} p(w_t | s_{t-1} = i, \lambda_{t-1}) p(\lambda_{t-1} | \mathcal{F}_{t-1}) d\lambda_{t-1}$$

(2.22)

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

$$p(w_t | s_{t-1} = i, \lambda_{t-1}) = \begin{cases} \frac{\Phi\left(\frac{1-\rho^2+\alpha^2\rho^2}{1-\rho^2} \left(\frac{1}{\sqrt{1-\alpha^2}} - \alpha^2\rho\right)\right)\Phi\left(\frac{1-\rho^2+\alpha^2\rho^2}{1-\rho^2} \left(\frac{1}{\sqrt{1-\alpha^2}} + \alpha^2\rho\right)\right)}{\Phi(\tau \sqrt{1-\alpha^2})} P_n\left(w_t | \lambda_{t-1}, 1-\rho^2+\alpha^2\rho^2 \right), & i = 0 \\ \frac{1-\Phi\left(\frac{1-\rho^2+\alpha^2\rho^2}{1-\rho^2} \left(\frac{1}{\sqrt{1-\alpha^2}} - \alpha^2\rho\right)\right)P_n\left(w_t | \lambda_{t-1}, 1-\rho^2+\alpha^2\rho^2 \right)}{\Phi(\tau \sqrt{1-\alpha^2})} & i = 1 \end{cases}$$

(2.23)

is derived in Corollary 3.3 of Chang et al. (2017a). Moreover, we again approximate $p(\lambda_{t-1} | \mathcal{F}_{t-1})$ by $p_n(\lambda_{t-1} | \lambda_{d,t-1|t-1}^i, \rho P_{d,t-1|t-1}^i)$ or simply the Dirac measure $\delta_{\lambda_{d,t-1|t-1}^i}$ for $s_{t-1} = i$. Then the filtered regime factor can be computed as

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

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 $\{w_{k|t}\}_{k=1}^{N}$ with their corresponding weights $\hat{p}(w_{k|t} | \mathcal{F}_t) = p(w_{k|t} | \mathcal{F}_t) / \sum_{k=1}^{N} p(w_{k|t} | \mathcal{F}_t)$. 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.

### 2.3 Comparison with Particle Filter

Before the empirical DSGE application, 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. The model structure resembles that of (2.1)–(2.2) for reduced-form DSGE models

\[
y_t = \begin{pmatrix} 1 & 1 \\ \end{pmatrix} \begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} + \omega u_t, \quad u_t \sim \mathcal{N}(0, 1) \tag{2.24}
\]

\[
\begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} = \begin{pmatrix} g_{11}(s_t) & g_{12}(s_t) \\ 0 & g_{22}(s_t) \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \sigma \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, 1) \tag{2.25}
\]

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 \((0.2, 0.8, 0.9)\) if \( s_t = 0 \), \((\omega, \sigma) = (0.2, 0.5)\), and \((\alpha, \tau, \rho) = (0.7, -0.5, 0.9)\).\(^5\) 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.

As a benchmark, we estimate the exact likelihood and obtain the 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 \( i = 0, 1 \), initialize \( \{x_{i,0}^k, P_{0,0}^i, p_{0,0}^i\} \) as in Algorithm 1 but based on the original state space form (2.1)–(2.2). For \( k = 1, \ldots, N \), draw an initial swarm of particles \( x_0^k \) from the mixture Gaussian distribution \( p_0^0 \cdot \mathcal{N}(x_{0,0}^0, P_{0,0}^0) + p_0^1 \cdot \mathcal{N}(x_{0,1}^0, P_{0,0}^1) \), \( \epsilon_0^k \sim \mathcal{N}(0_n \times 1, I_n) \), and \( w_0^k \sim \mathcal{N}(0, 1/(1-\alpha^2)) \). Also, set the corresponding weight \( W_0^k = 1/N \).

2. **Recursion.** For \( t = 1, \ldots, T \):

---

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

\(^6\)A 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.
(a) **Propagation.** Draw particles \( \{w^k_t\}_{k=1}^N \) from

\[
w^k_t = \alpha w^k_{t-1} + \sum_{j=1}^{n} \rho_j x^k_{t-1} + \sqrt{1 - \sum_{j=1}^{n} \rho_j^2} \eta^k_t, \quad \eta^k_t \sim \text{N}(0, 1) \tag{2.26}
\]

and compute \( s^k_t = \mathbb{1}\{w^k_t \geq \tau\} \). Then draw particles \( \{x^k_t\}_{k=1}^N \) from

\[
x^k_t = C_s^k + G_s^k x^k_{t-1} + E_s^k z_t + M_s^k \Sigma^{1/2} s^k_t \epsilon^k_t, \quad \epsilon^k_t \sim \text{N}(0_{n \times 1}, I_n) \tag{2.27}
\]

(b) **Likelihood evaluation.** The period-\( t \) likelihood can be approximated via numerical integration

\[
p(y_t | \mathcal{F}_{t-1}) = \sum_{k=1}^{N} W^k_{t-1} p(y_t | s^k_t, x^k_t) \tag{2.28}
\]

where \( p(y_t | s^k_t, x^k_t) = p_n(y_t | D_{s^k_t} + Z_{s^k_t} x^k_t + F_{s^k_t} z_t, \Omega_{s^k_t}) \).

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

\[
W^k_t = \frac{p(y_t | s^k_t, x^k_t) W^k_{t-1}}{p(y_t | \mathcal{F}_{t-1})} \tag{2.29}
\]

(d) **Resampling.** Define the effective sample size \( \text{ESS}_t = 1/ \sum_{k=1}^{N} (W^k_t)^2 \). If \( \text{ESS}_t < N/2 \), resample the particles \( \{x^k_t, \epsilon^k_t, w^k_t\}_{k=1}^N \) with the systematic resampling scheme introduced in Kitagawa (1996) and set \( W^k_t = 1/N \) for \( k = 1, \ldots, N \). Now the particle system \( \{x^k_t, \epsilon^k_t, w^k_t, W^k_t\}_{k=1}^N \) approximates any filtered value by

\[
\mathbb{E}[h(x_t, \epsilon_t, w_t) | \mathcal{F}_t] \approx \sum_{k=1}^{N} h(x^k_t, \epsilon^k_t, w^k_t) W^k_t \tag{2.30}
\]

where \( h \) denotes some transformation function of interest.

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

Practically it turns out that \( N = 100,000 \) particles are more than sufficient to deliver a numerically stable estimate of the likelihood function 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 Algorithms 1 and 2. A visual comparison suggests that our filter performs fairly well as it produces virtually indistinguishable approximations to these quantities from their particle filter counterparts. That accuracy extends to a wide range of the parameter space as can be seen in
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 (solid line) and particle filter (dash-dotted line). The horizontal 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 $\rho = 0$).

Figure 3, though the particle filter creates some non-smoothness in the log likelihood functions of certain parameters (Panel A). With one exception (Panel C), 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.

Another look at the performance of our filter can be achieved through examining the sampling properties of the differentials between log likelihoods approximated by various filters. Define

\[
\Delta_1 = \ln p_{\text{KF}}(Y_{1:T}) - \ln p_{\text{PF}}(Y_{1:T})
\]
\[
\Delta_2 = \ln p_{\text{KF}}(Y_{1:T}) - \ln p_{\text{Kim}}(Y_{1:T})
\]

This 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.
where KF, PF, and Kim refer to the endogenous-switching Kalman filter (Algorithm 1), particle filter (Algorithm 2), and Kim’s (1994) filter (Algorithm 1 with $\rho^*$), respectively. Figure 4 displays the empirical distributions of $\Delta_1$ (solid line) and $\Delta_2$ (dashed line) based on 500 simulated data sets of equal length $T = 100$. Most of the simulated values for $\Delta_1$ concentrate in the neighborhood of its sample mean (−1.65), which is reasonably close to zero so that our filter is competitive with the particle filter in terms of approximating the true log likelihoods. On the other hand, the probability mass of $\Delta_2$ falls mostly on the positive domain, encompassing values from −3 up to 13 in log likelihood unit. To put these differentials into perspective, compare twice the sample mean of $\Delta_2$ (11.46) to the 5% critical value of $\chi^2$ limit distribution with one degree of freedom (3.84) for testing the null hypothesis $\rho = 0$ against the alternative $\rho = 0.9$. It follows that ignoring the presence of endogeneity in regime switching can on average lead to a significant deterioration of model fit.
2.4 DSGE Application: The Road Ahead  
Over the past 20 years, DSGE models have become a useful tool for quantitative macroeconomic analysis in both academia and policymaking institutions. One particularly important development is incorporating the possibility of recurrent regime shifts (e.g., changes in monetary policy) into the model specification. Due to the substantial improvement in model fit, a multitude of empirical studies have proposed to estimate the state space representation of regime-switching DSGE models using likelihood-based econometric approaches [Schorfheide (2005), Liu et al. (2011), Bi and Traum (2012, 2014), Bianchi (2013), Davig and Doh (2014), Bianchi and Ilut (2017), Bianchi and Melosi (2017), Best and Hur (2019), among others].

We complement the recent literature on likelihood-based estimation of DSGE models with exogenous Markov switching by making regime change endogenous. At the core of our analysis is the endogenous feedback effect of underlying structural shocks on the regime generating process. As a result, economic agents update their beliefs each period about future regimes conditional on the realizations of shocks disturbing the economy. For pedagogical purposes, Section 3 employs a simple model adopted from Chang et al. (2017b) to endogenize regime switching in monetary policy, which admits analytical characterizations of the mechanism at work. Section 4 extends the simple model to a prototypical new Keynesian DSGE model, and derives its state space form that can be analyzed with our endogenous-switching Kalman filter introduced earlier in conjunction with a posterior sampler.
An important precursor to our study is Davig and Leeper (2006a), who applied the projection method to solve and calibrate a new Keynesian model where monetary policy rule changes whenever its target variables (e.g., inflation and output gap) cross some thresholds. More recently, Guerrieri and Iacoviello (2015, 2017) developed piecewise linear solution toolkit and likelihood-based estimation method for DSGE models subject to an occasionally binding constraint (e.g., the zero lower bound on nominal interest rates). In their setup, each state of the constraint—slack or binding—is handled as one of two distinct regimes under the same model. Like these studies, one may argue that it is more natural to assign the immediate triggers of regime switch to the state variables rather than the structural shocks. However, dynamics of all state variables are ultimately driven by a small number of structural shocks. We therefore view the structural shocks that generate aggregate fluctuations as the macroeconomic origins of regime shifts, and establish a novel feedback channel by which they contribute to regime switching. Our analytical and empirical examples below illustrate how the underlying structural shocks impact agents’ expectations formation and monetary policy regimes through this endogenous feedback channel.

Throughout the rest of this paper, we solve the model based on the assumption that economic agents, when forming rational expectations about future endogenous variables, exogenous shocks, and regime states, can observe their current and all past realizations. The regime factor, however, remains latent to agents as well as econometricians. As will be shown subsequently, it merely serves as an auxiliary variable that rationalizes the specific functional forms of time-varying transition probabilities. Consequently, the model can be solved without reference to the regime factor. From an empirical perspective, though, it is interesting to extract the latent regime factor from the data, which may be correlated with measured economic behavior in a meaningful way. For instance, using the same regime switching approach as in this paper, Chang et al. (2019) estimated a reduced-form model of monetary-fiscal regime changes and found that fiscal variables, particularly the tax to GDP ratio and net interest payment to government spending ratio, are among the most important variables in explaining the monetary regime factor.

3 Analytical Example

We first consider the simple frictionless model of inflation determination studied in Davig and Leeper (2006a), comprising a standard Fisher relation and an interest rate rule for monetary policy. In what follows, all variables are written in terms of log-deviations from their steady states.

---

8Using the same solution method, Bi and Traum (2012, 2014) estimated a real business cycle model where the government partially defaults on its debt whenever the debt level rises beyond a ‘fiscal limit’.

9Allowing agents to observe the regime factor poses keen computational challenges to solving the model. Also, incorporating unsynchronized regime changes in other parameters requires multiple regime factors to operationalize in our setup. These extensions, though theoretically appealing, are beyond the consideration of this paper.
3.1 The Setup  First, given a perfectly competitive endowment environment with flexible prices and one-period nominal bonds, the Fisher relation arises from the bond pricing equation and is, in its linearized form, given by

\[ i_t = E_t \pi_{t+1} + E_t r_{t+1} \]  

(3.1)

where \( i_t \) denotes the short-term nominal interest rate, \( \pi_t \) the inflation rate between periods \( t-1 \) and \( t \), and \( E_t \) the conditional expectation given information available through period \( t \). The real interest rate \( r_t \) evolves as an autoregressive process

\[ r_t = \rho_r r_{t-1} + \sigma_r \epsilon_{r,t}, \quad \epsilon_{r,t} \sim N(0, 1) \]  

(3.2)

where \( 0 \leq \rho_r < 1 \) and \( \sigma_r > 0 \).

Second, the monetary authority follows an interest rate feedback rule that systematically varies its response to contemporaneous inflation, depending on the underlying policy regime

\[ i_t = \phi_{s_t} \pi_t + \sigma_i \epsilon_{i,t}, \quad \phi_{s_t} = \phi_0 (1 - s_t) + \phi_1 s_t, \quad \epsilon_{i,t} \sim N(0, 1) \]  

(3.3)

where \( 1 < \phi_0 < \phi_1 \) and \( \sigma_i > 0 \).\(^{10}\) Here the response of policy rate to inflation is allowed to switch between, in the spirit of Leeper’s (1991) terminology, ‘more active’ and ‘less active’ monetary regimes. The regime index evolves according to \( s_t = 1 \{ w_t \geq \tau \} \) and the autoregressive regime factor follows \( w_t = \alpha w_{t-1} + v_t \). In the case of \( w_t = \pi_{t-1} \), our model reproduces that of Davig and Leeper (2006a) where the monetary authority responds systematically more aggressively when lagged inflation exceeds a particular threshold, and less aggressively when it is below the threshold.

Finally, we introduce an endogenous feedback channel from the current structural shocks to the future regime changes. There are two standardized shocks—the real rate shock \( \epsilon_{r,t} \) and the monetary policy shock \( \epsilon_{i,t} \)—driving this simple economy, but for illustration purposes, we only consider the feedback from the current monetary policy shock \( \epsilon_{i,t} \) through its potential correlation with the next period regime factor innovation \( v_{t+1} \). That is,

\[
\left( \begin{array}{c} \epsilon_{i,t} \\ v_{t+1} \end{array} \right) \sim N \left( \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} 1 & \rho \\ \rho & 1 \end{array} \right) \right), \quad -1 < \rho < 1
\]  

(3.4)

where \( \rho = \text{corr}(\epsilon_{i,t}, v_{t+1}) \) is a correlation parameter that measures the strength of endogeneity in regime switching. The above specification is simple enough to admit an analytical solution, yet rich enough to highlight the general features of a rational expectations model with endogenous

\(^{10}\)For analytical tractability, we assume that the monetary authority does not respond to output gap.
regime change in monetary policy.

It follows from (3.4) that

$$w_{t+1} = \alpha w_t + \rho \epsilon_{i,t} + \sqrt{1 - \rho^2} \eta_{t+1}, \quad \eta_{t+1} \sim \mathbb{N}(0, 1) \tag{3.5}$$

where the internal innovation $\epsilon_{i,t}$ and the external innovation $\eta_{t+1}$ are orthogonal to each other. This alternative representation of the regime factor points to the endogenous feedback from monetary interventions to the regime generating process. For example, when $\epsilon_{i,t}$ and $v_{t+1}$ are orthogonal (i.e., $\rho = 0$), regime shifts become the outcome of an exogenous process driven entirely by the non-structural shock $\eta_{t+1}$; as $\rho$ approaches to one in absolute value, today’s monetary shocks bear more directly on tomorrow’s regime factor; when $|\rho| = 1$, future regimes turn out to be completely predetermined. In general, one would expect $0 < |\rho| < 1$.\(^{11}\) The autoregressive coefficient $\alpha$, on the other hand, determines the persistency and hence the expected duration of each regime—as $\alpha$ takes values towards positive (negative) unity, the model will on average undergo less (more) frequent regime shifts.

### 3.2 Transition Probability

Like any regime switching model, it is essential to compute the associated transition probabilities. From a modeling point of view, it can be helpful to treat the regime factor as a computational device that produces the specific functional forms of transition probabilities adopted by this paper. To see that, first note

$$P(w_{t+1} < \tau | w_t, \epsilon_{i,t}) = P\left(\eta_{t+1} < \frac{\tau - \alpha w_t - \rho \epsilon_{i,t}}{\sqrt{1 - \rho^2}} \bigg| w_t, \epsilon_{i,t}\right) = \Phi_{\rho}(\tau - \alpha w_t - \rho \epsilon_{i,t})$$

where $\Phi_{\rho}(x) = \Phi(x/\sqrt{1 - \rho^2})$. Moreover, $w_t$ is independent of $\epsilon_{i,t}$ and follows $\mathbb{N}(0, 1/(1 - \alpha^2))$. Therefore, we can obtain the transition probability of staying in regime-0 (i.e., the less active regime) between periods $t$ and $t+1$ explicitly as

$$p_{00}(\epsilon_{i,t}) = P(s_{t+1} = 0 | s_t = 0, \epsilon_{i,t}) = \frac{P(w_{t+1} < \tau, w_t < \tau | \epsilon_{i,t})}{P(w_t < \tau)} = \frac{\int_{-\infty}^{\tau \sqrt{1 - \alpha^2}} \Phi_{\rho}(\tau - \alpha x/\sqrt{1 - \alpha^2} - \rho \epsilon_{i,t}) p_{\mathbb{N}}(x|0,1) dx}{\Phi(\tau \sqrt{1 - \alpha^2})}$$

$$= \frac{\int_{-\infty}^{\tau \sqrt{1 - \rho^2}} \int_{-\infty}^{\tau \sqrt{1 - \alpha^2}} p_{\mathbb{N}}(x,y|\mu_0, \Sigma_0) dx dy}{\Phi(\tau \sqrt{1 - \alpha^2})} \tag{3.6}$$

\(^{11}\)From now on, we dispense with $|\rho| = 1$ because predetermined regimes, though theoretically possible, make less economic sense.
where $\mu_0 = [0, 0]^\prime$, $\Sigma_0 = [1, c; c, 1+c^2]$, and $c = \alpha/(\sqrt{1-\rho^2}\sqrt{1-\alpha^2})$. Analogously, the transition probability from regime-1 (i.e., the more active regime) in period $t$ to regime-0 in period $t+1$ can be computed as

$$
p_{10}(\epsilon_{i,t}) = \mathbb{P}(s_{t+1} = 0|s_t = 1, \epsilon_{i,t})
= \frac{\mathbb{P}(w_{t+1} < \tau, w_t \geq \tau|\epsilon_{i,t})}{\mathbb{P}(w_t \geq \tau)}
= \int_{\tau\sqrt{1-\alpha^2}}^{\infty} \Phi_p(\tau - \alpha x/\sqrt{1-\alpha^2} - \rho \epsilon_{i,t})p_{\eta}(x|0,1)dx
= \frac{\int_{-\infty}^{\tau\sqrt{1-\alpha^2}} \int_{-\infty}^{-\tau\sqrt{1-\alpha^2}} p_{\eta}(x,y|\mu_1, \Sigma_1)dxdy}{1 - \Phi(\tau\sqrt{1-\alpha^2})} \tag{3.7}
$$

where $\mu_1 = [0,0]^\prime$ and $\Sigma_1 = [1, -c; -c, 1+c^2]$. Accordingly, we have $p_{01}(\epsilon_{i,t}) = 1 - p_{00}(\epsilon_{i,t})$ and $p_{11}(\epsilon_{i,t}) = 1 - p_{10}(\epsilon_{i,t})$. Finally, the integrals in (3.6)–(3.7) can be easily evaluated using the cumulative bivariate normal distribution function. See the Online Appendix for derivation details.

In sum, our endogenous feedback mechanism renders the transition probabilities, which now become an integral part of the model solution, time-varying because they are all functions of $\epsilon_{i,t}$. A key difference from Chang et al. (2017a), though, lies in that their $\epsilon_{i,t}$ corresponds to a univariate regression error that can be readily computed given the data and parameter values, whereas ours represents structural shocks whose values remain latent to the econometrician and hence must be inferred from the data. In the special case of $\rho = 0$, transition probabilities (3.6)–(3.7) become constants and our model reduces to one with conventional Markov switching.
Figure 5 plots transition probabilities as functions of monetary policy shock at selected values of \((\alpha, \tau, \rho)\), which is intended to highlight the distinct role of each parameter in shaping these functions while holding other parameters fixed. Overall, \((\alpha, \tau)\) uniquely determine the constant levels of \((p_{00}, p_{11})\) under exogenous Markov switching (solid and dash-dotted lines) owing to their one-to-one correspondence. Specifically, increasing the value of \(\alpha\) from \(-0.9\) to \(0.9\), for example, raises the likelihood of remaining in the current regime by making the regime factor more persistent. Meanwhile, decreasing the value of \(\tau\) from 0 to \(-1\) favors the more active regime by making it relatively easier for the regime factor to stay above the threshold. On the other hand, the endogeneity parameter \(\rho\) introduces shock-specific variations into \((p_{00}, p_{11})\) (dashed line). Due to the positive feedback effect \((\rho < 0.9)\), for instance, a one-time unanticipated tightening \(\epsilon_{i,t} > 0\) (loosening \(\epsilon_{i,t} < 0\)) of policy today increases the probability of staying in or shifting to the systematically tighter (looser) policy in the next period. Of course, the overall shapes of transition probability functions rest on all three parameters.

3.3 Equilibrium Characteristics

Together with the transition probabilities (3.6)–(3.7), equations (3.1)–(3.3) constitute a nonlinear rational expectations system in the endogenous variables \((\pi_t, i_t)\) that is driven by the exogenous variables \((r_t, \epsilon_{i,t})\). Substituting (3.2) and (3.3) into (3.1) delivers a regime-specific expectational difference equation in inflation

\[
\phi_s t \pi_t + \sigma_i \epsilon_{i,t} = E_t \pi_{t+1} + \rho_r r_t
\]

(3.8)

Now solving the model entails mapping the minimum set of state variables \((r_t, \epsilon_{i,t})\) into the endogenous variable \((\pi_t)\). We find such a minimum state variable solution with the method of undetermined coefficients by postulating a regime-specific solution that takes an additive form

\[
\pi_t = A_{st} r_t + B_{st} \epsilon_{i,t}
\]

(3.9)

The Online Appendix shows that these coefficients can be expressed as

\[
A_{st} = \frac{\rho_r}{\phi_{st}} \left( \phi_1 - \phi_0 \right) p_{st,0}(\epsilon_{i,t}) + \phi_1 \left( \frac{\phi_0}{\rho_r} - E p_{00}(\epsilon_{i,t}) \right) + \phi_0 E p_{10}(\epsilon_{i,t})
\]

(3.10)

\[
B_{st} = -\frac{\sigma_i}{\phi_{st}}
\]

where \(A_{st}\) also depends on \(\epsilon_{i,t}\) and the unconditional expectations \(E p_{00}(\epsilon_{i,t})\) and \(E p_{10}(\epsilon_{i,t})\) are constant terms.

Two special cases arise from the general solution (3.10). When regime changes are purely
exogenous (i.e., $\rho = 0$), (3.10) reduces to

$$A_{st} = \frac{\rho_r}{\phi_{st}} \left( \frac{\phi_1 - \phi_0}{\phi_{st}} \right) p_{st,0} + \frac{\phi_1}{\phi_{st}} \left( \frac{\phi_0}{\rho_r} - p_{00} \right) + \phi_0 p_{10} \quad B_{st} = -\frac{\sigma_i}{\phi_{st}} \quad (3.11)$$

where $p_{st,0}$ and hence $A_{st}$ become independent of $\epsilon_{i,t}$. Further imposing the restriction $\phi_0 = \phi_1 = \phi > 1$ gives the equilibrium inflation under fixed regime

$$A_{st} = \frac{\rho_r}{\phi - \rho_r}, \quad B_{st} = -\frac{\sigma_i}{\phi} \quad (3.12)$$

where both $A_{st}$ and $B_{st}$ are deterministic constants. It is clear from (3.12) that a more aggressive monetary stance (i.e., higher $\phi$) can effectively insulate inflation against exogenous disturbances.

In all cases, we have $A_{st} > 0$ so that a positive real rate shock, as it does in a fixed regime, raises the contemporaneous demand for consumption and thus inflation. With endogenous feedback in regime change, it immediately follows from (3.10) that two distinct effects on inflation emerge after a monetary policy intervention. First, conditioning on the prevailing policy regime, a monetary contraction tends to curtail inflation through its linear and direct effect captured by the $B_{st} < 0$ term. Second, and more importantly, a monetary contraction also generates an endogenous expectations-formation effect—the difference between the impacts of a shock when the regime switches endogenously and when it switches exogenously—that is captured by the transition probability $p_{st,0}(\epsilon_{i,t})$ in the $A_{st}$ term. This nonlinear effect arises in that the intervention induces a change in agents’ beliefs about the future policy regime. The resultant adjustment in agents’ behavior can shift the projected path and probability distribution of equilibrium inflation in economically meaningful ways. As opposed to the endogenous switching case, such forward-looking effect vanishes in (3.11) under exogenous switching and thus there is no channel by which monetary interventions can alter agents’ expectations about future regime.

To make the analytics more concrete, Figure 6 isolates the endogenous expectations-formation effect by comparing the contemporaneous responses of inflation to exogenous shocks under endogenous ($\rho = 0.9$) and exogenous switching ($\rho = 0$). We consider a policy process that adjusts nominal rate only ‘mildly’ in the less active regime ($\phi_0 = 1.1$), but to a degree more consistent with the standard Taylor rule specification when the more active regime is in place ($\phi_1 = 1.5$). We also have the regime factor relatively persistent ($\alpha = 0.9$) and somewhat favor the more active regime ($\tau = -1$). Panel B illustrates a slice of the impulse response surface reported in Panel A for a given positive real rate $r_t = 0.5\%$. Starting with the less active policy, contractionary monetary shocks trigger a positive feedback effect with endogenous switching, which leads agents

---

12The implied transition probabilities under exogenous switching are given by $(p_{00}, p_{11}) = (0.8, 0.9)$. See Figure 5 for a visualization of the transition probability functions associated with endogenous switching.
to revise their beliefs towards a tighter policy in the subsequent period (as evinced by Figure 5). In comparison with the exogenous switching case, this shift in expectations about future policy helps to further mitigate the inflationary effect of a positive real rate on impact. Analogously, expansionary monetary shocks trigger a negative feedback effect that bolsters agents’ beliefs in the looser policy for the next period. Although less noticeable, the same positive real rate thus has larger impacts on current inflation.\(^{13}\)

Formally, we measure expectations-formation effects from a policy intervention based on conditional inflation forecasts along the lines of Leeper and Zha (2003). Since private agents can observe current and all past realizations of endogenous variables \((\pi_t, i_t)\), exogenous shocks \((r_t, \epsilon_{i,t})\), and regime states \(s_t\), they formulate rational expectations about future inflation based on the information set \(F_T = \sigma(\{\pi_t, i_t, r_t, \epsilon_{i,t}, s_t\}_{t=0}^T)\). Let \(I_T\) be a hypothetical intervention at period \(T\), specified as a \(K\)-period sequence of exogenous policy actions \(I_T = \{\epsilon_{i,T+1}, \ldots, \epsilon_{i,T+K}\}\). Given the analytical solution under endogenous switching (3.10), it is straightforward to evaluate the forecast of \(\pi_{T+K}\) conditional on \(I_T\) as

\[
\mathbb{E}[\pi_{T+K}|I_T, F_T] = (p_{s_T,0}A_0(\epsilon_{i,T+K}) + p_{s_T,1}A_1(\epsilon_{i,T+K}))\rho_r^K r_T + (p_{s_T,0}B_0 + p_{s_T,1}B_1)\epsilon_{i,T+K}
\]

where \(p_{s_T,s_{T+K}}\) for \(s_{T+K} = 0, 1\) are the \(K\)-period-ahead transition probabilities that depend on the sequence \(\{\epsilon_{i,T}, \ldots, \epsilon_{i,T+K-1}\}\). Likewise, forecasts under exogenous switching \(\mathbb{E}[\pi_{T+K}|I_T, F_T, \rho =

\footnote{The analysis with a negative real rate is similar and therefore omitted here to conserve space.}
Figure 7: Expectations-formation effects on inflation. Notes: Panel A plots the inflation forecasts conditional on the intervention $I_T = \{3, 2, 1, 0.5, 0, \ldots, 0\}$. Panel B plots their differentials as in the definitions of expectations-formation effects. Panel C reports the ratio of endogenous to total effects. Panel D plots the $K$-period-ahead probabilities of remaining in the less active regime. The current set of state variables is set to $(r_T, \epsilon_i, T, s_T) = (0.5\%, 0, 0)$. See Figure 6 notes for parameter settings.

0] and fixed regime $\mathbb{E}[\pi_{T+K}|I_T, \mathcal{F}_T, s_t = 0, t = T+1, \ldots, T+K]$ can be computed using solutions (3.11) and (3.12), respectively. The total expectations-formation effect—a term coined by Leeper and Zha (2003)—refers to the difference between the impacts of a policy intervention when regime can and cannot switch, respectively

$$\text{Total effect} \equiv \mathbb{E}[\pi_{T+K}|I_T, \mathcal{F}_T] - \mathbb{E}[\pi_{T+K}|I_T, \mathcal{F}_T, s_t = 0, t = T+1, \ldots, T+K] \quad (3.13)$$

We focus on its endogenous component which is more relevant in our context

$$\text{Endogenous effect} \equiv \mathbb{E}[\pi_{T+K}|I_T, \mathcal{F}_T] - \mathbb{E}[\pi_{T+K}|I_T, \mathcal{F}_T, \rho = 0] \quad (3.14)$$

Of course, the difference between (3.13) and (3.14) quantifies the exogenous component.

For illustration purpose, suppose the monetary authority undertakes a one-year contractionary intervention $I_T = \{3, 2, 1, 0.5, 0, \ldots, 0\}$ with gradually declining magnitudes from 3 to 1/2 stan-
standard deviations $\sigma_i$ of the monetary policy shock. Panel A of Figure 7 records the conditional forecasts of inflation obtained from the three models.\textsuperscript{14} As expected, the intervention further tempers the initial inflationary impact of $r_T = 0.5\%$ by about 1\% with endogenous switching and 0.7\% with exogenous switching relative to the fixed regime baseline. These discrepancies among the forecasts translate into sizable expectations-formation effects in Panel B, which slowly taper off over the 10-year forecasting horizon. More importantly, over 30\% of the total effect during the intervention can be attributed to its endogenous component (see Panel C). At very short horizons, the consecutive spikes in policy rate put forth a significantly positive feedback effect with endogenous switching, placing nearly all probability weight on the more active regime (see Panel D). Once the intervention ends after one year, the probability of remaining in the less active regime eventually reverts to a permanently lower level relative to the exogenous switching case.

4 Empirical Illustration

We now consider the small-scale new Keynesian DSGE model presented in An and Schorfheide (2007) with the following features: 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.

4.1 The Setup

In the fixed regime benchmark, the model’s equilibrium conditions in terms of the detrended variables can be summarized as follows. First, the household’s optimizing behavior implies

\begin{equation}
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]
\end{equation}

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, $\pi_t$ 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 $\gamma$, and $\mathbb{E}_t$ represents the conditional expectation given information available at time $t$. The firm’s

\textsuperscript{14}In the special case of one-period intervention $I_T = \{1, 0, \ldots, 0\}$, these forecasts correspond to the conventional impulse response analysis.
optimal price-setting behavior yields
\begin{equation}
1 = \frac{1 - \frac{c_t^{\tau_c}}{\nu_t}}{\nu_t} + \phi(\pi_t - \pi) \left[ \left(1 - \frac{1}{2\nu_t}\right) \tau_c + \frac{\pi}{2\nu_t} \right] - \phi \beta \mathbb{E}_t \left[ \left( \frac{c_{t+1}}{c_t} \right)^{\frac{1-\tau_c}{\nu_t}} \frac{y_{t+1}}{y_t} (\pi_{t+1} - \pi) \pi_{t+1} \right]
\end{equation}
(4.2)

where \(\phi\) is the degree of price stickiness that relates to the slope of the so-called new Keynesian Phillips curve \(\kappa\) via \(\phi = \kappa (1 - \nu) / (\nu \pi^2 \kappa)\), \(\pi\) the steady state inflation, \(1/\nu\) the steady state of the demand elasticity \(1/\nu_t = u_t / (u_t - 1)\), \(u_t\) an exogenous price markup shock with the steady state \(u\), and \(y_t\) the detrended output.\(^\text{15}\)

Second, the goods market clearing condition is given by
\begin{equation}
y_t = c_t + \left(1 - \frac{1}{g_t}\right) y_t + \frac{\phi}{2} (\pi_t - \pi)^2 y_t
\end{equation}
(4.3)
where \(g_t\) is an exogenous government spending shock with the steady state \(g\). 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
\begin{equation}
R_t = R_t^{\ast 1 - \rho_R} R_{t-1}^{\rho_R} \epsilon_R R_t, \quad R_t^{\ast} = R \left( \frac{\pi_t}{\pi} \right)^{\psi_{\pi}} \left( \frac{y_t}{y_t^{\ast}} \right)^{\psi_{y}}
\end{equation}
(4.4)
where \(0 \leq \rho_R < 1\) is the degree of interest rate smoothing, \(\sigma_R > 0\), \(R\) the steady state nominal interest rate, \(\psi_{\pi} > 0\) and \(\psi_{y} > 0\) the policy rate responsive coefficients, \(y_t^{\ast} = (1 - \nu_t)^{1/\tau_c} g_t\) the detrended potential output that would prevail in the absence of nominal rigidities (i.e., \(\phi = 0\), and \(\epsilon_{R,t}\) an exogenous policy shock.\(^\text{16}\)

Finally, each of \((\ln z_t, \ln u_t, \ln g_t)\) evolves as an autoregressive process
\begin{align}
\ln z_t &= \rho_z \ln z_{t-1} + \sigma_z \epsilon_{z,t} \quad (4.5) \\
\ln u_t &= (1 - \rho_u) \ln u + \rho_u \ln u_{t-1} + \sigma_u \epsilon_{u,t} \quad (4.6) \\
\ln g_t &= (1 - \rho_g) \ln g + \rho_g \ln g_{t-1} + \sigma_g \epsilon_{g,t} \quad (4.7)
\end{align}
where \(0 \leq \rho_z, \rho_u, \rho_g < 1\) and \(\sigma_z, \sigma_u, \sigma_g > 0\). The model is driven by the four innovations \((\epsilon_{z,t}, \epsilon_{u,t}, \epsilon_{g,t}, \epsilon_{R,t})\) that are serially uncorrelated, independent of each other at all leads and lags, and normally distributed with zero mean and unit standard deviation.

\(^\text{15}\)Because our sample spans the 1970s oil crisis when the U.S. experienced severe petroleum shortages with elevated prices, we extend the original An and Schorfheide (2007) model to incorporate a shock to the price markup that serves the dual role of aggregate supply disturbance and potential trigger of monetary regime change.\(^\text{16}\)Alternatively, \(y_t^{\ast}\) could be defined as the level of detrended output when both nominal rigidities and price markup shocks vanish, i.e., \(y_t^{\ast} = (1 - \nu_t)^{1/\tau_c} g_t\). In this case, the monetary authority faces a trade-off between inflation stabilization and output gap stabilization.
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. As in the analytical model, the response of policy rate to inflation deviations is allowed to switch between more active and less active (or possibly ‘passive’) monetary regimes

$$\psi_{\pi,s,t} = \psi_{\pi,0}(1 - s_t) + \psi_{\pi,1} s_t, \quad 0 \leq \psi_{\pi,0} < \psi_{\pi,1}$$

(4.8)

where the regime index evolves according to

$$s_t = \mathbb{1}\{w_t \geq \tau\}$$

and the regime factor follows

$$w_t = \alpha w_{t-1} + v_t.$$ To introduce the sources of endogeneity in regime change, we allow all current structural shocks to jointly influence the next period regime through their correlations with the innovation $v_{t+1}$. That is,

$$
\begin{pmatrix}
\epsilon_t \\
\eta_{t+1}
\end{pmatrix} \sim \mathcal{N}
\begin{pmatrix}
\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
\begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{pmatrix},
\begin{pmatrix}
1 & \rho \\
\rho' & 1
\end{pmatrix}
\rho \rho' < 1
(4.9)
$$

where $\rho = [\rho_{zv}, \rho_{uv}, \rho_{gv}, \rho_{Rv}] = \text{corr}(\epsilon_t, v_{t+1})$. To keep the illustration simple and concrete, we abstract from time variation in the model structure other than that from the policy parameter $\psi_{\pi,s,t}$.\footnote{Initiated 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 regime factors in our setup, which is beyond the scope of this paper.}

Again it follows from (4.9) that a more explicit representation of the regime factor can be written as

$$w_{t+1} = \alpha w_t + \rho_{zv} \epsilon_{z,t} + \rho_{uv} \epsilon_{u,t} + \rho_{gv} \epsilon_{g,t} + \rho_{Rv} \epsilon_{R,t} + \sqrt{1 - \rho' \rho} \eta_{t+1}, \quad \eta_{t+1} \sim \mathcal{N}(0, 1)$$

(4.10)

where the internal innovations $(\epsilon_{z,t}, \epsilon_{u,t}, \epsilon_{g,t}, \epsilon_{R,t})$ and the external innovation $\eta_{t+1}$ are all orthogonal and have unit variance. Equation (4.10) asserts a complete separation between the four individual endogenous drivers $(\epsilon_{z,t}, \epsilon_{u,t}, \epsilon_{g,t}, \epsilon_{R,t})$ and the exogenous driver $\eta_{t+1}$ of the regime factor. Also recall that $(\rho_{zv}^2, \rho_{uv}^2, \rho_{gv}^2, \rho_{Rv}^2, 1 - \rho' \rho)$ measure the percentage contributions of $(\epsilon_{z,t}, \epsilon_{u,t}, \epsilon_{g,t}, \epsilon_{R,t}, \eta_{t+1})$ to the unconditional variance of $w$ and hence the extents to which these drivers trigger historical regime changes. In what follows, we quantify how much of the U.S. monetary policy shifts can be attributed, respectively, to each of technology growth, price markup, government spending, and monetary policy shocks.

### 4.2 Solution Method

In the regime switching context, the model is completed by deriving the probabilities governing regime transition from one period to the next. Analogous to the
scalar case in the analytical model, the implied (endogenous) transition probabilities to the less active regime, which become an important part of the model solution, are given by

\[ p_{00}(\epsilon_t) = P(s_{t+1} = 0|s_t = 0, \epsilon_t) = \frac{\int_{-\infty}^{\tau\sqrt{1-\alpha^2}} \Phi_p(\tau - \alpha x/\sqrt{1-\alpha^2} - \rho' \epsilon_t) p_\eta(x|0, 1) dx}{\Phi(\tau\sqrt{1-\alpha^2})} \] (4.11)

\[ p_{10}(\epsilon_t) = P(s_{t+1} = 0|s_t = 1, \epsilon_t) = \frac{\int_{-\infty}^{\tau\sqrt{1-\alpha^2}} \Phi_p(\tau - \alpha x/\sqrt{1-\alpha^2} - \rho' \epsilon_t) p_\eta(x|0, 1) dx}{1 - \Phi(\tau\sqrt{1-\alpha^2})} \] (4.12)

where \( \Phi_p(x) = \Phi(x/\sqrt{1-\rho^2}) \). See the Online Appendix for derivation details. When \( \epsilon_t \) and \( v_{t+1} \) are orthogonal (i.e., \( \rho = 0_{4 \times 1} \)), transition probabilities (4.11)–(4.12) become constants so that our model nests the exogenous Markov switching as a special case.

Conditional on the current regime \( s_t = 0, 1 \), equations (4.1)–(4.7) constitute a rational expectations system that can be cast into the generic form

\[ E[f_{s_t}(x_{t+1}, x_t, x_{t-1}, \epsilon_t)|F_t] = 0 \] (4.13)

where \( f_{s_t} \) is a vector of nonlinear functions, and the vectors of model variables \( x_t \) and shock innovations \( \epsilon_t \) are explicitly given by

\[ x_t = [c_t, y_t^*, y_t^*, R_t^*, R_t^*, \pi_t, \nu_t, z_t, u_t, g_t]' \quad \epsilon_t = [\epsilon_z, \epsilon_u, \epsilon_g, \epsilon_R]' \]

respectively. As mentioned earlier, private agents can observe current and all past realizations of endogenous variables, exogenous shocks, and regime states, but not the regime factors. Accordingly, they formulate rational expectations about future variables on the basis of the information set \( F_t = \sigma(\{x_k, \epsilon_k, s_k\}_{k=0}^t) \).

System (4.13) has to be solved before the model can be taken to data. To that end, a spate of theoretical and empirical efforts have managed to solve regime-switching rational expectations models using numerical techniques. One strand of the literature embraces the projection method to iteratively construct policy functions over a discretized state space [Davig (2004), Davig and Leeper (2006a,b), Bi and Traum (2012, 2014), Davig et al. (2010, 2011), Richter et al. (2014)]. Nevertheless, global approximations suffer from, among other problems, the curse of dimensionality that renders the practical implementation computationally costly even for small-scale models. The second strand begins with a linear or linearized model as if its parameters were constant and then annexes Markov switching to certain parameters [Svensson and Williams (2007), Farmer et al. (2011), Bianchi (2013), Cho (2016), Bianchi and Ilut (2017), Bianchi and Melosi (2017)]. While this approach is not as limited by the model size as the projection method, linearization without accounting for the switching parameters may be inconsistent with the optimizing behavior of agents who are aware of the switching process in the original nonlinear model. The third
strand circumvents the above problems by embedding regime switching in perturbation solutions whose accuracy can be enhanced with higher-order terms [Maih (2015), Foerster et al. (2016), Barthélémy and Marx (2017), Bjørnland et al. (2018), Maih and Waggoner (2018)].

We obtain the model solution to the nonlinear rational expectations system (4.13) using the perturbation approach of Maih and Waggoner (2018), which is more general than the ones found in the earlier literature. As opposed to Foerster et al. (2016), for example, it requires no partition of the switching parameters, allows for the possibility of multiple steady states and, more importantly, can handle models with endogenous transition probabilities. To fix ideas, we seek a regime-specific policy function for $x_t$ that depends on the minimum set of state variables $(x_{t-1}, \epsilon_t)$

$$x_t = g_{s_t}(x_{t-1}, \epsilon_t)$$

Substituting (4.14) into (4.13) and integrating out the future regime yield

$$\mathbb{E} \left[ \sum_{j=0}^{1} p_{i,j}(\epsilon_t) f_i(g_j(x_{t-1}, \epsilon_t), \epsilon_{t+1}), g_i(x_{t-1}, \epsilon_t); x_{t-1}, \epsilon_t \bigg| \mathcal{F}_t \right] = 0$$

(4.15)

where the notational convention that $s_t = i$ and $s_{t+1} = j$ is followed. In general, there is no analytical solution to (4.15) even when $f_i$ is linear. Maih and Waggoner (2018) proposed to obtain a Taylor series approximation to (4.14) by introducing an auxiliary argument $\sigma$ (i.e., the perturbation parameter)

$$x_t = g_i(x_{t-1}, \epsilon_t, \sigma)$$

(4.16)

that solves a perturbed version of (4.15)

$$\mathbb{E} \left[ \sum_{j=0}^{1} q_{i,j}(\epsilon_t, \sigma) f_i(g_j(h_i(x_{t-1}, \epsilon_t, \sigma), \epsilon_{t+1}, \sigma), g_i(x_{t-1}, \epsilon_t, \sigma); x_{t-1}, \epsilon_t \bigg| \mathcal{F}_t \right] = 0$$

(4.17)

with perturbed transition probability $q_{i,j}(\epsilon_t, \sigma)$ and perturbed policy function $h_i(x_{t-1}, \epsilon_t, \sigma)$.

The principle of perturbation is to choose $q_{i,j}(\epsilon_t, \sigma)$ and $h_i(x_{t-1}, \epsilon_t, \sigma)$ such that (4.17) becomes the original system (4.15) when $\sigma = 1$, but it reduces to a tractable and interpretable system when $\sigma = 0$. While (4.17) is already successful in eliminating the stochastic disturbances when $\sigma = 0$, choices of $q_{i,j}(\epsilon_t, \sigma)$ and $h_i(x_{t-1}, \epsilon_t, \sigma)$ still play a key role in interpreting the model’s steady state $x_i = g_i(x_i, 0, 0)$ around which the solution is expanded. In the context of exogenous Markov switching, Foerster et al. (2016) set $q_{i,j}(\epsilon_t, \sigma) = p_{i,j}$ and $h_i(x_{t-1}, \epsilon_t, \sigma) = g_i(x_{t-1}, \epsilon_t, \sigma)$.

\footnote{Barthélémy and Marx (2017) also generalized standard perturbation methods to solve a class of nonlinear rational expectations models with endogenous regime switching.}
Plugging their choice into (4.17) and evaluating it at the steady state give

\[ \sum_{j=0}^{1} p_{i,j} f_i(g_j(x_i, 0, 0), x_i, x_i, 0) = 0 \]

Since the only point at which we know how to evaluate \( g_j \) is \((x_j, 0, 0)\), this implies that \( x_i = x_j \) for each \( j \) and hence the steady state must be independent of any regime.\(^{19}\) By contrast, recognizing that switching parameters may imply distinct steady states, Maih and Waggoner’s (2018) choice of

\[
q_{i,j}(\epsilon_t, \sigma) = \begin{cases} 
\sigma p_{i,j}(\epsilon_t), & i \neq j \\
\sigma(p_{i,i}(\epsilon_t) - 1) + 1, & i = j
\end{cases}
\]

and \( h_i(x_{t-1}, \epsilon_t, \sigma) = g_i(x_{t-1}, \epsilon_t, \sigma) + (1 - \sigma)(x_j - x_i) \) implies

\[ f_i(x_i, x_i, x_i, 0) = 0 \]

Consequently, \( x_i \) can be readily interpreted as the deterministic steady state that would prevail in regime-\( i \) when it is considered in isolation.

In practice, the solution algorithm of Maih and Waggoner (2018) has been coded up in RISE, a flexible object-oriented MATLAB toolbox developed by Junior Maih, that is well-suited for solving and estimating a general class of regime-switching DSGE models.\(^{20}\) See Maih (2015) for a detailed description of the theory behind the routines implemented in RISE.

### 4.3 Econometric Method

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).\(^{21}\) The actual data are constructed as in Appendix B of Herbst and Schorfheide (2015) and available from the Federal Reserve Economic Data (FRED). The observable variables are linked to the model variables through the following measurement

---

\(^{19}\)In this regard, the literature typically defines the steady state as the one associated with the ergodic mean values of Markov switching parameters. Such unique steady state, however, need not be an attractor—a resting point towards which the model tends to converge in the absence of further shocks. For example, Aruoba et al. (2017) presented a model that exhibits two distinct attractors, i.e., a targeted-inflation steady state and a deflationary steady state.

\(^{20}\)The toolbox is available, free of charge, at https://github.com/jmaih/RISE_toolbox.

\(^{21}\)Our sample begins when the federal funds rate data first became available and ends before the federal funds rate nearly hit its effective lower bound.
equations

\[
\begin{pmatrix}
Y_{GR_t} \\
INF_t \\
INT_t
\end{pmatrix} = \begin{pmatrix}
\gamma^{(Q)} \\
\pi^{(A)} \\
\pi^{(A)} + r^{(A)} + 4\gamma^{(Q)}
\end{pmatrix} + 100 \begin{pmatrix}
\ln(z_t y_t / y_{t-1}) \\
4 \ln(\pi_t / \pi) \\
4 \ln(R_t / R)
\end{pmatrix}
\] (4.18)

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 \(\pi = 1 + \pi^{(A)}/400.\) Let \(\theta\) be a vector collecting all model parameters. In conjunction with the model solutions under fixed regime and switching regime, a first-order approximation to equations (4.18) and (4.16) around steady states maps directly into the general state space form (2.1)–(2.2), whose likelihood function \(p(Y_{1:T}|\theta)\) can be evaluated with the standard Kalman filter and our endogenous-switching Kalman filter.\(^{22}\)

In the Bayesian paradigm, the state space model is completed with a prior distribution \(p(\theta)\) summarizing the researcher’s initial views of the model parameters. This prior information is updated with the sample information via Bayes’ theorem

\[
p(\theta|Y_{1:T}) \propto p(Y_{1:T}|\theta)p(\theta)
\] (4.19)

where the posterior distribution \(p(\theta|Y_{1:T})\) characterizing the researcher’s updated parameter beliefs is calculated up to the normalization constant (i.e., the marginal likelihood \(p(Y_{1:T})\)). Since the posterior surface can be highly irregular (e.g., non-elliptical and multimodal) in sophisticated DSGE models, we adopt the tailored randomized block Metropolis-Hastings (TaRB-MH) algorithm of Chib and Ramamurthy (2010) to sample the parameters from their joint posterior distribution. Two defining features of this algorithm, which are designed to overcome the difficulties in irregular problems, are worth mentioning in relation to the more standard posterior samplers (e.g., the random-walk MH algorithm) in empirical macroeconomics. One feature is the random clustering of the parameters into an arbitrary number of blocks at every iteration. Each block is then sequentially updated through an MH step. This proves to be particularly useful when the researcher does not have a priori knowledge about the correlation pattern of the parameters so that the grouping by correlation principle becomes infeasible. Another feature is the local tailoring of the proposal density to the location and curvature of the posterior distribution for a given block using an optimization routine. This allows for sizable moves from the neighborhood of the current parameter draw.

\(^{22}\)The standard stability concept for constant-parameter linear rational expectations models does not extend to the regime switching case. Instead, following the lead of Svensson and Williams (2007) and Farmer et al. (2011) among others, we adopt the concept of mean square stability to characterize first-order stable solutions.
4.4 Priors  Following standard practice, we assume that all structural parameters are a priori independent. Table 1 summarizes their marginal prior distributions. 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 the prior mean of $r^{(A)}$ translates into a $\beta$ value of 0.998. The priors on the structural shock processes are harmonized: the autoregressive coefficients $(\rho_z, \rho_u, \rho_g, \rho_R)$ are beta distributed with mean 0.5 and standard deviation 0.1, and the standard deviation parameters $(\sigma_z, \sigma_u, \sigma_g, \sigma_R)$, all scaled by 100, follow inverse-gamma type-I distribution whose hyperparameters imply a mean of 0.5 and standard deviation of 0.26. Furthermore, the priors on the private sector parameters $(\tau_c, \kappa, \nu, 1/g)$ and the fixed regime policy responses $(\psi_\pi, \psi_y)$ are largely adopted from An and Schorfheide (2007), whereas those on the switching parameters $(\psi_{\pi,0}, \psi_{\pi,1})$ closely follow the specification in Davig and Doh (2014), which a priori rule out the possibility of ‘label switching’ by imposing $\psi_{\pi,0} < \psi_{\pi,1}$ and hence achieve regime identification. Finally, turning to the parameters for the autoregressive regime factor, the prior on $\alpha$ centers at a rather persistent value that, together with the prior mean of $\tau$, implies symmetric transition probabilities $p_{00}, p_{11} = (0.85, 0.85)$ under exogenous switching. On the other hand, the relatively diffuse priors on $(\rho_{zv}, \rho_{uv}, \rho_{gv}, \rho_{Rv})$ centering at zero reflect an agnostic view about the sign and degree of endogeneity in regime switching.

4.5 Posterior Estimates For each model, we sample a total of 11,000 draws from the posterior distribution using the TaRB-MH algorithm and discard the first 1,000 draws as burn-in phase. The resulting 10,000 draws form the basis for performing our posterior inference. Owing to the efficiency gains achieved by the TaRB-MH algorithm, the number of draws is substantially smaller than the number typically used for the conventional random-walk MH algorithm. Evidence of such gains can be seen from the very low inefficiency factors (ranging from 2.3 to 27.9, with most values below 10) in Table 2.24 In conjunction with a rejection rate of approximately 50% in the MH step for each model, the small inefficiency factors suggest that the Markov chain mixes well. We highlight several aspects of the posterior estimates reported in Table 2 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 $(\tau_c, \kappa)$, the interest rate response to output deviation $\psi_y$, the steady state parameters $(\pi^{(A)}, \gamma^{(Q)})$, and the standard deviation of markup shock $100\sigma_u$. Other parameters, however, remain more or less recalcitrant to regime switching. Second, turning to the switching parameters, there remains considerable cross-regime difference in the policy response to inflation, although the
Table 2: Posterior Estimates of DSGE Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>No Switching Model</th>
<th>Regime Switching Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mode</td>
<td>90% HPD</td>
</tr>
<tr>
<td>Fixed Regime</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau_c$</td>
<td>4.73</td>
<td>[4.02, 5.95]</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.24</td>
<td>[0.15, 0.38]</td>
</tr>
<tr>
<td>$\psi_\pi$</td>
<td>1.10</td>
<td>[1.00, 1.23]</td>
</tr>
<tr>
<td>$\psi_y$</td>
<td>0.14</td>
<td>[0.03, 0.29]</td>
</tr>
<tr>
<td>$\rho^{(A)}$</td>
<td>0.46</td>
<td>[0.32, 0.62]</td>
</tr>
<tr>
<td>$\pi^{(A)}$</td>
<td>3.75</td>
<td>[2.96, 4.67]</td>
</tr>
<tr>
<td>$\gamma^{(Q)}$</td>
<td>0.40</td>
<td>[0.31, 0.49]</td>
</tr>
<tr>
<td>$\rho_z$</td>
<td>0.52</td>
<td>[0.42, 0.62]</td>
</tr>
<tr>
<td>$\rho_\nu$</td>
<td>0.86</td>
<td>[0.82, 0.89]</td>
</tr>
<tr>
<td>$\rho_g$</td>
<td>0.53</td>
<td>[0.38, 0.69]</td>
</tr>
<tr>
<td>$\rho_R$</td>
<td>0.75</td>
<td>[0.72, 0.79]</td>
</tr>
<tr>
<td>$100\sigma_{\pi}$</td>
<td>0.69</td>
<td>[0.58, 0.78]</td>
</tr>
<tr>
<td>$100\sigma_u$</td>
<td>2.86</td>
<td>[2.02, 3.76]</td>
</tr>
<tr>
<td>$100\sigma_g$</td>
<td>0.31</td>
<td>[0.21, 0.43]</td>
</tr>
<tr>
<td>$100\sigma_R$</td>
<td>0.24</td>
<td>[0.22, 0.27]</td>
</tr>
<tr>
<td>$\nu$</td>
<td>0.08</td>
<td>[0.02, 0.16]</td>
</tr>
<tr>
<td>$1/g$</td>
<td>0.92</td>
<td>[0.74, 0.99]</td>
</tr>
<tr>
<td>Threshold Switching</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{\pi,0}$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\psi_{\pi,1}$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\tau$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\rho_{zv}$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\rho_{uv}$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\rho_{gv}$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\rho_{Rv}$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\ln p(Y_{1:T})$</td>
<td>–1050.42 (0.06)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The 90% highest probability density (HPD) intervals are constructed as in Chen and Shao (1999). The numerical standard errors of log marginal likelihood estimates are reported in parentheses.

Overall monetary policy stance appears to be somewhat less aggressive than assumed a priori—
the 90% posterior intervals of \((\psi_{\pi,0}, \psi_{\pi,1})\) are both shifted downward. Third, moving on to the parameters \((\alpha, \tau)\) unique to our threshold switching, a prior-posterior comparison reveals that the data support an even more persistent process for the regime factor, and slightly favor the more active regime by making it relatively easier for the regime factor to remain above the threshold. Furthermore, the posterior mode of \(\rho\) attributes a significant portion of regime developments to past price markup shocks (about 50%) and, to a lesser extent, technology growth shocks (about 4%) as well as monetary policy shocks (about 3%). Finally, the estimated marginal likelihoods (in logarithm scale) translate into a Bayes factor of approximately \(e^{34}\), indicating that the data overwhelmingly favor the regime switching model over the fixed regime model.\(^{25}\)

It is worth noting that, up to first-order approximation, the endogenous and exogenous switching solutions consist of identical policy function \((4.14)\) and differ only in terms of the transition probabilities \((4.11)\)–\((4.12)\), irrespective of the perturbation method used.\(^{26}\) Although not reported here, the exogenous switching model yields comparable parameter estimates and a marginally smaller log marginal likelihood. The associated log likelihoods, however, can still be significantly different. To wit, let
\[
\Delta = \ln p_{\text{KF}}(Y_{1:T}|\hat{\theta}) - \ln p_{\text{KF}}(Y_{1:T}|\hat{\theta}, \rho = 0_{1 \times 1}),
\]
where \(\hat{\theta}\) denotes the posterior mode under each model, and compare twice the differential \(\Delta\) (10.99) to the 5% critical value of \(\chi^2\) limit distribution with four degrees of freedom (9.49). It follows that the data favor the model that accounts for the endogeneity in regime changes.

Figure 8 compares the prior and posterior densities of the correlation parameters \((\rho_{zv}, \rho_{uv}, \rho_{gv}, \rho_{Rv})\), which further substantiate 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 informative to land the posteriors onto narrower areas of the parameter space that deliver tightly estimated degrees of endogeneity in regime switching. Most noticeable is the endogenous feedback from price markup shock, whose posterior mass falls almost entirely on the negative territory. As a result, adverse supply disturbances unambiguously increase the likelihood of staying in or shifting to the less active regime, consistent with a countercyclical monetary policy that is ‘leaning against the wind’. Less evident is the posterior distribution of the endogenous feedback from technology growth shock that concentrates somewhat more on the negative territory. Thus, on balance, favorable technological advancements tend to make the less active regime more likely, suggesting an accommodative monetary policy to promote long-term economic growth. These patterns connect broadly to theoretical work and empirical observations about how central banks routinely act.\(^{27}\) Government spending shock, on the other hand, plays no observable role in driving the regime changes since in this simple model it only affects output

\(^{25}\)Log marginal likelihoods are approximated using the modified harmonic mean estimator of Geweke (1999) with a truncation parameter of 0.5. According to Jeffreys’ (1961) criterion, a log Bayes factor greater than 4.6 signifies decisive evidence in favor of the model with superior fit.

\(^{26}\)This property can be seen from the fact that in system \((4.15)\) the model equations are multiplied by the transition probabilities that sum to unity, regardless of whether the model is in steady state or not.

\(^{27}\)See Section 3 for an analysis of the endogenous feedback caused by monetary policy shock.
but not output gap (hence consumption, inflation, and nominal rate).

Just as the analytical example of Section 3 demonstrates, endogenizing regime changes here can generate important expectations-formation effects beyond what the exogenous switching can do. For example, consider a sequence of positive markup shocks $\{2, 2.5, 3, 2.5, 2, 1.5, 1, 0.5, 0, \ldots, 0\}$ that aims to mimic the sharp increases in oil prices beginning in 1973 (see Panel A of Figure 10 below) and reproduces the dire inflationary scenario afterward. This exogenous ‘intervention’ persists for two years and peaks in magnitude after two quarters since it hits the economy. Figure 9 displays the conditional forecasts of key model variables for the three cases.\textsuperscript{28} With the more

\textsuperscript{28}Because the model is inherently nonlinear and does not admit an analytical solution, sequential Monte Carlo methods are required to numerically integrate out the regime index in each period. To compute the conditional forecasts under endogenous switching, for instance, we simulate $N = 500$ trajectories of regime states $(s_{T+1}^{(i)}, \ldots, s_{T+K}^{(i)})_{i=1}^{N}$ according to the time-varying transition probabilities (4.11)–(4.12), apply the regime-specific policy function (4.14) to obtain the corresponding model variables $(x_{T+1}^{(i)}, \ldots, x_{T+K}^{(i)})_{i=1}^{N}$ conditional on a pre-specified sequence of markup shocks $(\epsilon_{u,T+1}, \ldots, \epsilon_{u,T+K})$ but zeroing out all other shocks, and take the sample average of $(x_{t}^{(i)})_{i=1}^{N}$ for $t = T + 1, \ldots, T + K$. In the special case of one-period shock $\{1, 0, \ldots, 0\}$, these
active regime initially in place, higher price markups not only dampen aggregate supply and hence engender the usual stagflation—a mix of rising inflation (Panel B) and falling economic activity (Panel A)—as in the fixed regime, but also trigger a negative feedback effect ($\rho_{uv} = -0.71$) that induces agents to form a stronger belief in the less active regime under which inflation rises by more and output falls by less. This shift in agents’ expectations spells out the discrepancies in the forecasts under exogenous and endogenous switching (i.e., with and without the restriction $\rho_{uv} = 0$). In response to the inflationary pressure that outweighs the recessionary pressure, the monetary authority raises the nominal rate according to its policy rule (Panel C).

Unlike the Markov switching filter of Kim (1994), our filter also produces an important by-product—an estimated time series of the regime factor $w_{1t}$ as portrayed in Panel A of Figure 10—that complements the information contained in the estimated regime-1 probability $p_{1t}$ (Panel forecasts are equivalent to the generalized impulse response functions of Koop et al. (1996).
C). In particular, this series exhibits prolonged swings that identify 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. Such a 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 mid-1980s and beyond under Paul Volcker and Alan Greenspan stemmed from the conviction that inflation has high costs and few benefits; the less active stance of most of the 1970s under Arthur Burns and G. William Miller stemmed from an overestimate of the natural rate of unemployment as well as an underestimate of the sensitivity of inflation to economic slack.

We conclude with a final remark on the regime factor. Loosely speaking, it conveniently packs,
through a reduced-form mechanism, various structural and extraneous sources of regime changes into one single index that is likely to proxy for some observables relevant to the regime generating process. To corroborate this interpretation, Panel A of Figure 10 relates the estimated regime factor to a common measure of oil price inflation. The apparent negative correlation since 1973 stems from the fact that the latter variable potentially identifies the price markup shock, which is found to be the most important structural driver of monetary regime changes according to our estimated DSGE model. For example, the precipitous run-ups in oil prices during the 1970s are well captured by the continuous upticks in the estimated markup shocks (Panel B).

5 Concluding Remarks

This paper aims at broadening the scope for understanding the complex interaction between recurrent regime 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 encompass many well-known dynamic linear models, including classical regression models and the popular DSGE 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 underlying economic fundamentals to bear more directly on the regime generating process. When regime switches 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. 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. Our 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 the state space representation of a prototypical new Keynesian DSGE model with threshold switching in monetary policy rule, and illustrate how the underlying structural shocks impact agents' expectations formation and monetary policy regimes through our novel endogenous feedback channel. We find compelling statistical support for the endogenous feedback from historical price markup shocks to the prevailing policy regime, but observe limited evidence for other structural shocks as the drivers of the U.S. monetary policy changes.

We remind the reader that quantifying such feedback mechanism depends on nearly every aspect of private and policy behavior. A comprehensive investigation calls for a richly structured medium-scale DSGE model with fiscal details in both the model specification and the observable data. Another natural extension of our framework is to permit multiple regimes and latent factors along with the development of its filtering algorithm. We defer these extensions 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, 2018 International Association for Applied Econometrics Annual Conference at Montréal, 2018 Taiwan Econometric Society Annual Meeting, and the 27th Annual Symposium of the Society for Nonlinear Dynamics and Econometrics at Dallas, as well as Francesco Bianchi, Yongok Choi, Luca Guerrieri, Edward Herbst, Eric Leeper, Chaojun Li, Joon Park, Shi Qiu, Daniel Waggoner, and Tao Zha 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: Transition Probability**

The prerequisite to computing transition probabilities is to evaluate integrals of the form

\[
F(a, b, c, d, m, V) = \int_{-\infty}^{\infty} \int_{-\infty}^{a} \Phi(b + cx + dy)p_{N}(x|0, 1)p_{N}(y|m, V)dx\,dy
= \mathbb{P}(Z_1 \leq a, Z_2 \leq b + cZ_1 + dZ_3) \tag{A.1}
\]

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

\[
F(a, b, c, d, m, V) = \mathbb{P}(W_1 \leq a, W_2 - cW_1 \leq b)
= \int_{-\infty}^{b} \int_{-\infty}^{a} p_{\mathcal{N}}(x, y|\mu, \Sigma)dx\,dy \tag{A.2}
\]

where

\[
\mu = \begin{pmatrix} 0 \\ -dm \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 1 & -c \\ -c & 1 + c^2 + d^2V \end{pmatrix}
\]
Therefore, the expected transition probabilities in Algorithm 1 of Section 2.2 can be calculated as
\[
\int_{-\infty}^{\infty} \mathbb{P}(s_t = 0|s_{t-1} = 0, \lambda_{t-1}, \mathcal{F}_{t-1}) p(\lambda_{t-1}|s_{t-1} = 0, \mathcal{F}_{t-1}) d\lambda_{t-1}
\]
\[= \frac{1}{\phi(\tau \sqrt{1 - \alpha^2})} F \left( \tau \sqrt{1 - \alpha^2}, \frac{\tau}{\sqrt{1 - \rho^2}}, \frac{-\alpha}{\sqrt{1 - \rho^2 \sqrt{1 - \alpha^2}}}, -\frac{1}{\sqrt{1 - \rho^2 \sqrt{1 - \alpha^2}}}, \rho' s_{d,t-1|t-1}, \rho' P_{d,t-1|t-1} \right) \]
and
\[
\int_{-\infty}^{\infty} \mathbb{P}(s_t = 0|s_{t-1} = 1, \lambda_{t-1}, \mathcal{F}_{t-1}) p(\lambda_{t-1}|s_{t-1} = 1, \mathcal{F}_{t-1}) d\lambda_{t-1}
\]
\[= \frac{1}{1 - \phi(\tau \sqrt{1 - \alpha^2})} F \left( -\tau \sqrt{1 - \alpha^2}, \frac{\tau}{\sqrt{1 - \rho^2}}, \frac{-\alpha}{\sqrt{1 - \rho^2 \sqrt{1 - \alpha^2}}}, -\frac{1}{\sqrt{1 - \rho^2 \sqrt{1 - \alpha^2}}}, \rho' s_{d,t-1|t-1}, \rho' P_{d,t-1|t-1} \right) \]
Moreover, the expected transition probabilities in Section 3.3 can be calculated as
\[\mathbb{E}_{\rho_{00}(\epsilon_{i,t})} = \frac{1}{\phi(\tau \sqrt{1 - \alpha^2})} F \left( \tau \sqrt{1 - \alpha^2}, \frac{\tau}{\sqrt{1 - \rho^2}}, \frac{-\alpha}{\sqrt{1 - \rho^2 \sqrt{1 - \alpha^2}}}, -\frac{\rho}{\sqrt{1 - \rho^2}}, 0, 1 \right) \]
and
\[\mathbb{E}_{\rho_{10}(\epsilon_{i,t})} = \frac{1}{1 - \phi(\tau \sqrt{1 - \alpha^2})} F \left( -\tau \sqrt{1 - \alpha^2}, \frac{\tau}{\sqrt{1 - \rho^2}}, \frac{-\alpha}{\sqrt{1 - \rho^2 \sqrt{1 - \alpha^2}}}, -\frac{\rho}{\sqrt{1 - \rho^2}}, 0, 1 \right) \]
Lastly, the transition probabilities in Sections 3.2 and 4.2 can be calculated as
\[p_{00}(\epsilon_t) = \frac{1}{\phi(\tau \sqrt{1 - \alpha^2})} F \left( \tau \sqrt{1 - \alpha^2}, \frac{\tau - \rho' \epsilon_t}{\sqrt{1 - \rho^2 \rho}}, \frac{-\alpha}{\sqrt{1 - \rho^2 \rho \sqrt{1 - \alpha^2}}}, 0, \cdot, \cdot \) \]
and
\[p_{10}(\epsilon_t) = \frac{1}{1 - \phi(\tau \sqrt{1 - \alpha^2})} F \left( -\tau \sqrt{1 - \alpha^2}, \frac{\tau - \rho' \epsilon_t}{\sqrt{1 - \rho^2 \rho}}, \frac{-\alpha}{\sqrt{1 - \rho^2 \rho \sqrt{1 - \alpha^2}}}, 0, \cdot, \cdot \) \]

**Appendix B: 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 \(\{w_t^k\}_{k=1}^N\) with their corresponding weights \(W_t^k = \hat{p}(w_t^k|\mathcal{F}_T)\) such that \(\sum_{k=1}^N W_t^k = 1\), which form the basis for computing the smoothed value
\[\mathbb{E}[h(w_t)|\mathcal{F}_T] \approx \sum_{k=1}^N h(w_t^k) W_t^k \]  
(B.1)
for some transformation function h. For example, setting $h(w_t) = w_t$ and $h(w_t) = 1\{w_t \geq \tau\}$ give $w_{t|T} = \mathbb{E}[w_t|\mathcal{F}_T]$ and $p(s_t = 1|\mathcal{F}_T)$, respectively.

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

$$p(w_t^j|\mathcal{F}_T) = \int_{-\infty}^{\infty} p(w_t^j|w_{t+1}^k,\mathcal{F}_T)p(w_{t+1}^k|\mathcal{F}_T)dw_{t+1}$$

$$\approx \sum_{k=1}^N p(w_t^j|w_{t+1}^k,\mathcal{F}_T)W_{t+1}^k \quad (B.2)$$

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

$$p(w_t^j|w_{t+1}^k,\mathcal{F}_T) = \frac{p(Y_{t+1:T}|w_{t+1}^k, w_t^j, \mathcal{F}_T)p(w_t^j|\mathcal{F}_t)}{p(Y_{t+1:T}, w_{t+1}^k|\mathcal{F}_t)}$$

$$= \frac{p(Y_{t+1:T}|w_{t+1}^k, w_t^j, \mathcal{F}_T)p(w_t^j|\mathcal{F}_t)p(w_{t+1}^k|\mathcal{F}_t)}{p(Y_{t+1:T}|w_{t+1}^k, \mathcal{F}_T)p(w_{t+1}^k|\mathcal{F}_t)}$$

$$= \frac{p(w_{t+1}^k|w_t^j, \mathcal{F}_t)p(w_t^j|\mathcal{F}_t)}{p(w_{t+1}^k|\mathcal{F}_t)} \quad (B.3)$$
To evaluate (B.3), we first define the regime index $i = 1 \{ w_{i}^{j} \geq \tau \}$ and approximate

$$p(w_{i+1}^{k} | w_{i}^{j}, F_{t}) \approx p_{N}(w_{i+1}^{k} | \alpha w_{i}^{j} + \rho' \xi_{d,t}^{i}, 1 - \rho' \rho)$$

Moreover, $p(w_{i}^{j} | F_{t})$ is available from (2.22). Lastly, we approximate

$$p(w_{i+1}^{k} | F_{t}) \approx \sum_{i=0}^{1} p(w_{i+1}^{k} | s_{t} = i, \rho' \xi_{d,t}^{i})p(s_{t} = i | F_{t})$$

where $p(w_{i}^{j} | s_{t} = i, \rho' \xi_{d,t}^{i})$ is given by (2.23) and $p(s_{t} = i | F_{t})$ is a by-product of running our endogenous-switching Kalman filter. Now set $W_{i}^{j} = p(w_{i}^{j} | F_{t}) / \sum_{j=1}^{N} p(w_{j}^{j} | F_{t})$ for the next iteration. Evaluating at the true parameter values, Figure 11 depicts the smoothed autoregressive regime factor for the small simulation model of Section 2.3, which to a large extent recovers its true trajectory.

**APPENDIX C: ANALYTICAL SOLUTION** We posit an additive solution to the analytical model of Section 3.1

$$\pi_{t} = A_{s_{t}}(\epsilon_{i,t})r_{t} + B_{s_{t}}\epsilon_{i,t}$$

(C.1)

with undetermined coefficients $A_{s_{t}}(\epsilon_{i,t})$ and $B_{s_{t}}$. Since private agents can observe current and all past realizations of endogenous variables ($\pi_{t}, i_{t}$), exogenous shocks ($r_{t}, \epsilon_{i,t}$), and regime states $s_{t}$, they formulate rational expectations about future inflation according to $E_{t}\pi_{t+1} = E[\pi_{t+1} | F_{t}]$ with $F_{t} = \sigma(\{\pi_{k}, i_{k}, r_{k}, \epsilon_{i,k}, s_{k} \}_{k=0}^{t})$. Under this information structure, we can obtain

$$E_{t}\pi_{t+1} = E[A_{s_{t+1}}(\epsilon_{i,t+1})r_{t+1} | F_{t}]$$

$$= E[A_{s_{t+1}}(\epsilon_{i,t+1}) | F_{t}]E[r_{t+1} | F_{t}]$$

$$= (p_{s_{t},0}(\epsilon_{i,t})E[A_{0}(\epsilon_{i,t+1}) | F_{t}] + p_{s_{t},1}(\epsilon_{i,t})E[A_{1}(\epsilon_{i,t+1}) | F_{t}])\rho_{r}r_{t}$$

$$= (p_{s_{t},0}(\epsilon_{i,t})E_{0}(\epsilon_{i,t+1}) + p_{s_{t},1}(\epsilon_{i,t})E_{1}(\epsilon_{i,t+1}))\rho_{r}r_{t}$$

(C.2)

where the first equality holds by the independence of $s_{t+1}, \epsilon_{i,t+1}$, and $F_{t}$, the second equality holds because $r_{t+1}$ is independent of $s_{t+1}$ and $\epsilon_{i,t+1}$, and the last equality holds by the independence of $\epsilon_{i,t}$ and $F_{t}$.

Now substituting (C.1) and (C.2) into (3.8) and solving for $\pi_{t}$ yield

$$\pi_{t} = \frac{\rho_{r}}{\phi_{s_{t}}}(p_{s_{t},0}(\epsilon_{i,t})E_{0}(\epsilon_{i,t+1}) + p_{s_{t},1}(\epsilon_{i,t})E_{1}(\epsilon_{i,t+1}) + 1)r_{t} - \frac{\sigma_{t}}{\phi_{s_{t}}}\epsilon_{i,t}$$

43
Matching the coefficients of $r_t, \epsilon_{i,t}$ with those in (C.1), we can deduce that

$$A_{st}(\epsilon_{i,t}) = \frac{\rho_r}{\phi_{st}} (p_{st,0}(\epsilon_{i,t})C_0 + p_{st,1}(\epsilon_{i,t})C_1 + 1)$$  \hspace{1cm} (C.3)$$

$$B_{st} = -\frac{\sigma_i}{\phi_{st}}$$  \hspace{1cm} (C.4)$$

where $C_0 = \mathbb{E}A_0(\epsilon_{i,t+1})$ and $C_1 = \mathbb{E}A_1(\epsilon_{i,t+1})$ are two undetermined constants. For $s_t = 0$ and $s_t = 1$, respectively, take unconditional expectation of (C.3) to obtain a system of two equations in two unknowns $(C_0, C_1)$

$$C_0 = \mathbb{E}A_0(\epsilon_{i,t}) = \frac{\rho_r}{\phi_0} (\mathbb{E}p_{00}(\epsilon_{i,t})C_0 + \mathbb{E}p_{01}(\epsilon_{i,t})C_1 + 1)$$

$$C_1 = \mathbb{E}A_1(\epsilon_{i,t}) = \frac{\rho_r}{\phi_1} (\mathbb{E}p_{10}(\epsilon_{i,t})C_0 + \mathbb{E}p_{11}(\epsilon_{i,t})C_1 + 1)$$

from which we can solve for $(C_0, C_1)$. Plugging the expressions of $(C_0, C_1)$ into (C.3) gives

$$A_{st}(\epsilon_{i,t}) = \frac{\rho_r}{\phi_{st}} \left( \phi_1 - \phi_0 \right) p_{st,0}(\epsilon_{i,t}) + \phi_1 \left( \frac{\phi_0}{\rho_r} - \mathbb{E}p_{00}(\epsilon_{i,t}) \right) + \phi_0 \mathbb{E}p_{10}(\epsilon_{i,t})$$

which confirms our initial conjecture about the dependence of $A_{st}$ on $\epsilon_{i,t}$.

**References**


CHANG, MAIH & TAN: ENDogenous-SWITCHING KALMAN FILTER


