Timing and Signals of Monetary Regime Switching

Daniel Soques

September 27, 2019

Abstract

This study investigates if the reaction function of the Federal Reserve switches between two distinct policy rules. By using a time-varying transition probability framework, we also determine if forward-looking macroeconomic or financial co-variates signal an impending monetary policy switch. We find that U.S. monetary policy is best described by a Markov-switching model with two regime processes, one of which controls for heteroskedasticity in the shocks to the policy rule. We find that the Fed switches between an aggressive regime with a high weight on inflation and a dovish regime that is less responsive to inflationary pressures. We find that an increase in private forecasters’ expectations of an impending recession signals a switch from the more aggressive policy regime to the less aggressive regime. A recovery in equity returns signals a return back to the more aggressive regime.

Keywords: Monetary Policy, Markov-Switching, Time-Varying Transition Probabilities

JEL Codes: C24, E32, E52

*Jordan Fonville provided research assistance. The author benefitted from conversations with Yoosoon Chang, Adam Check, Taeyoung Doh, Andrew Foerster, Neville Francis, Ming Chien Lo, Michael Owyang, Laura Jackson Young, and participants at the 2017 Conference of the Society for Economic Measurement, the 2018 Symposium of the Society of Nonlinear Dynamics and Econometrics, the 2018 IAAE Annual Conference, and the 2018 Midwestern Econometrics Group Conference. The author is indebted to Athanasios Orphanides for providing data on the real-time output gap.

†Department of Economics and Finance, University of North Carolina Wilmington. soquesd@uncw.edu
1 Introduction

Modern macroeconomic models illustrate the importance of monetary policy in smoothing business cycle fluctuations and ensuring price stability. The most popular way to describe monetary policy in these models is through a simple interest rate rule similar to the one outlined by Taylor (1993). These Taylor-type rules are intuitive in that they embody the Fed’s dual mandate by explaining movements in the short-term policy rate through movements in inflation and the output gap. In addition to modern macroeconomic models, the Fed explicitly considers an array of interest rate rules when setting monetary policy.¹

The assumption in the standard New Keynesian model is that the interest rate rule followed by the central bank does not change across the business cycle. Theoretical studies such as Barthélémy and Marx (2017), Best and Hur (2017), Bianchi (2013), Chang and Kwak (2017), Davig and Doh (2014), Foerster (2016), Lhuissier and Zabelina (2015), Liu et al. (2009), Liu et al. (2011), Schorfheide (2005) and others now consider switches in monetary policy and their macroeconomic implications in DSGE frameworks. Notably, Choi and Foerster (2016) finds that monetary policy rules with switching can provide welfare gains relative to constant rules.

Numerous empirical studies investigate whether the assumption of parameter stability in interest rate rules is historically valid, or if the Fed’s behavior differs across time. Common approaches to identifying switches in policy include time-varying parameter models² and Markov-switching models,³ which only identify the timing of switches and not potential indicators of these policy switches. A third approach is smooth-transition models which allow for smoothed regime changes in the interest rate reaction function to

¹For example, the Monetary Policy Report submitted to Congress on July 5, 2019 states, “In addition to weighing a wide range of economic and financial data and information received from business contacts and other informed parties around the country, policymakers regularly consult prescriptions for the interest rate arising from various monetary policy rules. These rule prescriptions can serve as useful guidelines to the FOMC in the course of arriving at its policy decisions.”


depend on an observable transition variable.\textsuperscript{4} This feature is useful when the researcher knows which single covariate informs the evolution of the switching process. However, these frameworks are limited in that they can only consider a single factor influencing the policy transition, and the transition between regimes is deterministic as it relates to the transition variable.\textsuperscript{5}

This study investigates if the Fed’s reaction function can be described by different monetary regimes over the past 43 years and if there are indicators of switches between these regimes. We identify monetary policy regimes using a Markov-switching model of a Taylor-type rule which allows for a separate regime-switching process for the coefficients and the variance of the shock to the policy rule. Using a time-varying transition probability framework, we consider a host of transition covariates that influence the switching process, including measures of private market uncertainty of future inflation and output growth, expectations of an impending recession, and equity price movements. Additionally, we estimate the model using real-time, forward-looking data to account for the Fed’s information set at the time of setting monetary policy.

We find that the Fed switches between two distinct policy regimes. Both regimes abide by the Taylor Principle – the idea that the monetary policy rate reacts more than one-for-one with inflation.\textsuperscript{6} However, the regimes vary in their degree of aggressiveness to meet the price stability mandate. The more aggressive regime is characterized by a higher inflation coefficient and a lower inflation target compared to the less aggressive regime. In terms of timing, the Fed was in the less aggressive regime during recessionary times and stayed in the less aggressive regime for some time after the recessions of the early 1990s and 2001, most likely due to the economy’s slow recoveries coming out of these recessions.

Our main contribution is identifying which covariates signal an impending switch of


\textsuperscript{5}Ahmad (2016) considers two transition variables by allowing for more than two regimes, however the switches are still deterministic as it relates to the transition covariates.

\textsuperscript{6}This finding of active monetary policy throughout the sample is in line with Orphanides (2004) and Davig and Doh (2014), however it runs counter to Murray et al. (2015). Similar to Davig and Doh (2014), our model is relatively more flexible than that of Murray et al. (2015) since it allows for a separate regime-switching process for the variance of the interest rate shock.
the monetary policy regime. We find that an increase in the Anxious Index (private forecasters’ expectation of a recession) considerably increases the transition probability from the more aggressive regime to the less aggressive regime. Additionally, a switch back from the less aggressive regime to the more aggressive regime increases in probability as equity returns rise. We do not find evidence that changes in private forecasters’ uncertainty about inflation or output growth signal switches in policy.

We compare the baseline model to a number of alternative specifications. Our baseline model with two separate regime processes fits the data better than a simple Taylor-type rule with smoothing, a recession-based switching rule, and a number of Markov-switching models with alternative regime dynamics. For robustness, we consider an extended sample which includes the zero lower bound period and find qualitatively similar results for our baseline model estimation and model comparison.

Our paper is most similar to that of Davig and Doh (2014). Their paper estimates a New Keynesian model with two regime processes, one for the Taylor rule coefficients and one for shock volatilities. Similar to our results, they find that monetary policy differs across regimes in its response to inflation but not to the output gap. However, our paper differs from theirs in a number of aspects. First, their model assumes the monetary regime process follows constant transition probabilities and therefore it is unable to identify variables which may indicate impending switches in monetary policy. Our approach is able to identify these signals by using a time-varying transition probability framework. Second, Davig and Doh (2014) used revised data whereas our paper uses real-time data to account for the Fed’s true information set.

Other similar papers include Murray et al. (2015) and Assenmacher-Wesche (2006). Murray et al. (2015) estimates a regime-switching model of the Taylor rule using real time data, but their model is not able to identify signals of switching since they use constant transition probabilities. Additionally, their paper only considers a single regime processes which we show does not describe the data as well as a two regime model. Assenmacher-Wesche (2006) addresses this second concern by estimating a two regime model where the coefficients and volatility switch independently. However, their paper uses revised
data and uses constant transition probabilities which again does not allow us to identify signals of monetary switching.

The paper proceeds as follows: Section 2 outlines the various interest rate models to be estimated. Section 3 explains the Bayesian estimation methods. Section 4 details the data sources and coverage period. Section 5 overviews the results for the various interest rate rules. We offer some robustness checks in section 6. Section 7 concludes.

2 Empirical Model

In a standard Taylor-type rule, the central bank adjusts its target short-term policy rate, $i_T^t$ based on movements in inflation, $\pi_t$ and the output gap, $y_t$. Explicitly the equation is:

$$i_T^t = r^* + \pi^T + \phi^\pi(\pi_t - \pi^T) + \phi^y y_t + \epsilon_t,$$

where $r^*$ is the real neutral rate of interest and $\pi^T$ is the central bank’s target inflation rate. This simple rule is a popular description of the stance of U.S. monetary policy since it encapsulates the dual mandate of price stability and full employment. For example, when inflation goes above the central bank’s target (i.e., $\pi_t > \pi^T$) the central bank increases its target policy rate according to the coefficient $\phi^\pi$. Similarly when actual output is above the potential level of output (i.e., $y_t > 0$) the Fed increases its target policy rate by a factor of $\phi^y$.

The Fed tends to adjust the policy rate smoothly towards its target.\footnote{See Coibion and Gorodnichenko (2012), Goodfriend (1991), Rudebusch (1995), and Woodford (1999), among others.} Therefore, we model the observed policy rate $i_t$ as a weighted average between the lagged policy rate $i_{t-1}$ and the target policy rate $i_T^t$:

$$i_t = \rho i_{t-1} + (1 - \rho)i_T^t$$

where the parameter $\rho$ reflects the degree of smoothing done by the central bank. Plugging
in the target reaction function (1) gives us:

\[ i_t = \rho i_{t-1} + (1 - \rho) \left[ C + \phi^\pi \pi_t + \phi^y y_t \right] + \varepsilon_t \quad (2) \]

where

\[ C = r^* + (1 - \phi^\pi) \pi^T, \quad (3) \]

and \( \varepsilon_t \sim N(0, \sigma^2) \) reflects shocks to the implied policy rule. We will refer to equation (2) as the simple interest rate rule.

One objective of this paper is to identify if and/or when the monetary policy reaction function has switched across time. Therefore, we augment the simple rule by allowing the coefficients and volatility parameter to differ across time. We adopt a Markov-switching framework in the vein of Hamilton (1989) to capture discrete switches between different reaction functions. Evidence from Assenmacher-Wesche (2006), Banaian and Lo (2010), Best and Hur (2017), Davig and Doh (2014), and Perruchoud (2009) suggests the timing of switches in the coefficients and the variance parameter may differ from each other. Notably, Davig and Doh (2014) find that a model with two separate regime processes provides the best model fit in a Markov-switching New Keynesian framework.

Given this evidence, we opt for a Markov-switching model with two regime processes. The coefficient parameters (\( \rho, C, \phi^\pi, \phi^y \)) depend upon the regime variable \( S^\theta_t \), whereas the error variance (\( \sigma^2 \)) depends upon a separate regime variable \( S^\sigma_t \). Formally, the baseline model with two regime variables is:

\[ i_t = \rho_{S^\theta_t} i_{t-1} + (1 - \rho_{S^\theta_t}) \left( C_{S^\theta_t} + \phi_{S^\theta_t}^\pi \pi_t + \phi_{S^\theta_t}^y y_t \right) + \varepsilon_t, \quad (4) \]

where \( \varepsilon \sim N(0, \sigma^2_{S^\sigma_t}) \).

In Markov-switching models with multiple regime processes like ours, we must make an assumption about how the two regime variables interact. We assume the two regime variables are independent. A similar model using revised data was first outlined by Assenmacher-Wesche (2006), which imposes the two switching processes are independent. Perruchoud (2009) estimated a similar model on revised data without time-varying transition probabilities using Bayesian methods to describe Switzerland’s monetary policy reaction function.
variables \( S_t^\theta \) and \( S_t^\sigma \) are independent in our baseline model. In addition to being more parsimonious, this assumption allows us to disentangle switches in interest rate volatility that are not explained by shifts in monetary policy preferences. However, one could allow full- or partial-dependence in the two regime processes, which we consider as alternative model specifications in section 6.

Additionally, we must outline how each of the regime variables evolve across time. As in Hamilton (1989), Markov-switching models typically assume the aggregate regime \( S_t \) evolves according to constant transition probabilities (CTP). CTP accounts for regime persistence but lacks insight as to what factors might be driving or signaling a switch in the regime variable. In Markov-switching models with CTP, the characterization of regimes is done \textit{ex-post} by comparing the parameter estimates governing the reaction function across the two regimes. However, this setup does not allow us to determine if there are certain variables which signal a switch between regimes. Therefore, we assume the coefficient regime \( S_t^\theta \) follows a first-order, two-state Markov-process with time-varying transition probabilities (TVTP):

\[
p^\theta_{\theta,i,t} = Pr(S_t^\theta = j|S_{t-1}^\theta = i, Z_t)
\]

where \( Z_t \) is a \((L \times 1)\) vector of observable covariates. As in Kaufmann (2015), the TVTP take the logistic form:

\[
p^\theta_{1,i,t} = \frac{1}{1 + \exp (\bar{\gamma}_{2i} + \gamma'_{2i} Z_t)},
\]
\[
p^\theta_{2,i,t} = \frac{\exp (\bar{\gamma}_{2i} + \gamma'_{2i} Z_t)}{1 + \exp (\bar{\gamma}_{2i} + \gamma'_{2i} Z_t)}.
\]

for \( i = 1, 2 \). Under the TVTP assumption first outlined by Filardo (1994) and Diebold et al. (1994), the coefficient regime process depends not only on the previous state (represented in \( \bar{\gamma}_{ji} \)) but also on the additional covariates \( Z_t \).

Due to the assumption of independence between the two regime processes, we must specify a separate evolution for \( S_t^\sigma \). We assume the variance regime variable follows a
separate first-order, two-state Markov process with transition matrix:

\[
P^\sigma = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix},
\]

and constant transition probabilities:

\[
p^\sigma_{ji} = \Pr(S_t^\sigma = j | S_{t-1}^\sigma = i),
\]

where \(p_{1i}^\sigma + p_{2i}^\sigma = 1\) is imposed. Note that we assume the variance regime \(S_t^\sigma\) follows CTP for parsimony and because the goal of this paper is to explain switches in the Fed’s reaction to the inflation and output gaps, and not switches in the volatility of the error term.

It is useful to summarize the state of the two regime variables \(S_t^\theta\) and \(S_t^\sigma\) with an aggregate regime variable \(S_t\). At any given time period, the aggregate regime variable takes one of four values:

\[
S_t = \begin{cases} 
1 & \text{if } S_t^\theta = 1 \text{ and } S_t^\sigma = 1 \\
2 & \text{if } S_t^\theta = 1 \text{ and } S_t^\sigma = 2 \\
3 & \text{if } S_t^\theta = 2 \text{ and } S_t^\sigma = 1 \\
4 & \text{if } S_t^\theta = 2 \text{ and } S_t^\sigma = 2
\end{cases},
\]

with time-varying transition matrix

\[
P_t = \begin{bmatrix} p_{11,t}^\theta p_{11}^\sigma & p_{11,t}^\theta p_{12}^\sigma & p_{12,t}^\theta p_{11}^\sigma & p_{12,t}^\theta p_{12}^\sigma \\ p_{11,t}^\theta p_{21}^\sigma & p_{11,t}^\theta p_{22}^\sigma & p_{12,t}^\theta p_{21}^\sigma & p_{12,t}^\theta p_{22}^\sigma \\ p_{21,t}^\theta p_{11}^\sigma & p_{21,t}^\theta p_{12}^\sigma & p_{22,t}^\theta p_{11}^\sigma & p_{22,t}^\theta p_{12}^\sigma \\ p_{21,t}^\theta p_{21}^\sigma & p_{21,t}^\theta p_{22}^\sigma & p_{22,t}^\theta p_{21}^\sigma & p_{22,t}^\theta p_{22}^\sigma \end{bmatrix}.
\]

The regime switching rule outlined in this section allows us to address a number of interesting questions about the Fed’s behavior. First, the model allows us to estimate
when the Fed has switched between different reaction functions while controlling for phases of high- and low-interest rate volatility. Second, it allows us to determine if the Fed has consistently followed the Taylor principle (i.e., $\phi^\pi > 1$) across different monetary regimes. Lastly, the TVTP specification determines if certain macroeconomic or financial factors signal an imminent switch between the monetary regimes.

### 3 Estimation

We estimate the model using Bayesian methods. In particular, we use Gibbs sampling to construct the full joint posterior distribution of a given model’s parameters and the latent regime time series $S^T = \{S_t\}_{t=1}^T$. Table 1 shows the prior distributions for each of the model parameters. The full estimation technique is outlined in Appendix A.

We partition the set of parameters and the latent variable to be estimated into distinct blocks to be drawn separately. There are four blocks: (1) the state-dependent coefficients $\theta = [\theta'_1, \theta'_2]'$ where $\theta_i = [\rho_i, C_i, \phi_i^x, \phi_i^y]'$, and (2) the state-dependent variances $\sigma^2 = [\sigma_1^2, \sigma_2^2]'$, (3) the aggregate regime variable $S^T$, and (4) the time series of transition probabilities in $P^T = \{P_t\}_{t=1}^T$. The draws for the coefficient and variance parameters are standard conditional on the aggregate regime variable as outlined by Kim and Nelson (1999). To draw the regime variable, we utilize the filter first outlined by Hamilton (1989) and extended by Carter and Kohn (1994). Following Chib (1996), we draw the parameters determining the constant transition probabilities from a Dirichlet distribution. We draw the time-varying transition parameters $\gamma_2 = [\hat{\gamma}'_{21}, \hat{\gamma}'_{22}]$ where $\hat{\gamma}_{2i} = [\hat{\gamma}_{2i}, \bar{\gamma}_{2i}]'$ using the difference random utility model from Kaufmann (2015).

We run the sampler for 10,000 iterations after a burn-in period of 20,000 iterations. Convergence of the sampler is confirmed by analyzing running means and autocorrelation functions of the draws.
4 Data

The baseline coverage period of our sample is 1968:Q4 - 2007:Q4. We consider an extended sample from 1968:Q4 to 2012:Q4 in section 6.2. We use real-time data for both the output gap and inflation. Data on the output gap comes from two sources. For the period 1968:Q4 - 1995:Q4, we use the real-time output gap from Orphanides (2004). For 1996:Q1 - 2007:Q4, we use the real-time output gap compiled by the Philadelphia Fed’s Real Time Data Research Center, which is based on the Greenbook estimates from the Federal Reserve Board. We use the Greenbook projections of the annualized quarterly change in the GDP Price Deflator to measure inflation. To calculate $\pi_t$, we take the average of the projected inflation rates from periods $t = t - 1, ..., t + 3$.

For the short-term policy rate, we use the observed federal funds rate averaged over the entire quarter. Due to the zero lower bound period, we replace the observed federal funds rate from 2008Q4 to 2012:Q4 with the estimated shadow rate from Wu and Xia (2016). The shadow rate acts as a proxy for the stance of monetary policy when the observed policy rate does not reflect unconventional actions taken by the Fed at the zero lower bound.

In addition to the variables that directly enter the Fed’s reaction function, we need to specify transition covariates $Z_t$ which influence the switching process for the coefficient regime $S^\theta_t$. Following the literature review of Gnabo and Moccero (2015), we choose four transition covariates suggested by previous studies: (1) dispersion in inflation forecasts, (2) dispersion in output growth forecasts, (3) the anxious index, and (4) equity market returns. The two dispersion measures account for private sector uncertainty about either the outlook for inflation or economic growth, respectively. When the dispersion of either inflation or output growth forecasts is high, the Fed may switch its policy response due to heightened uncertainty about macroeconomic conditions. Dispersion for variable $x$ at time $t$ is calculated as:

$$D^x_t = \frac{x^75_{t+4|t} - x^{25}_{t+4|t}}{x^{90}_{t+4|t}}$$

where $x^{p}_{t+q|t}$ is the $p$th forecast percentile for variable $x$ over the next $q$ quarters. For the
inflation forecast distribution, we use the GDP Price Deflator forecasts from the Survey of Professional Forecasters (SPF). Similarly, we use real GDP growth forecasts from the SPF for the output growth forecast distribution.

The anxious index comes from the SPF and is forecasters’ estimated probability that there will be a decline in Real GDP in the quarter the survey is taken. Therefore the anxious index measures expectations of a recession. From a Phillips curve perspective, the Fed may switch its reaction function between expansions and recessions based on which of the dual mandates requires attention at the time.

Lastly, we use equity returns from Ken French’s website\textsuperscript{9} which is the value-weighted return of all US firms listed on the NYSE, AMEX, or NASDAQ. We consider equity returns since the Fed might change policy rules between bull and bear markets to both alleviate financial pressures and due to their forward-looking nature.

5 Results

5.1 Baseline Parameter Estimates

Table 2 presents the estimation results for the baseline model using the baseline sample period 1968:Q4 - 2007:Q4. For each parameter, we show the posterior median along with the 67\% highest posterior density (HPD) interval. To obtain an estimate for $\pi^T$ we utilize equation (3) assuming $r^* = 2.5\%$, which is the average effective federal funds rate minus realized inflation (as measured by the GDP Deflator) across the baseline sample period.\textsuperscript{10}

In comparing the parameter estimates across the two regimes, we find no marked differences in the Fed’s smoothing parameter or their weight on the output gap. However, we find that the Fed has a substantially higher weight on the inflation gap in regime 1 relative to regime 2. In response to a change in inflation, the Fed changes its target policy rate nearly 67\% more in regime 1 compared to regime 2. Additionally, the Fed

\textsuperscript{9}http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
\textsuperscript{10}We also estimated the target inflation rate using the time-varying real rate $r^*_t$ estimates from Laubach and Williams (2003). This implies the target inflation rate varies across time due to either preference changes represented in the regime variable $S_t^\theta$ or changes across time in $r^*$. The results are qualitatively similar and available upon request.
has a notably lower inflation target in regime 1 than regime 2. A higher $\phi^\pi$ and lower $\pi^T$ implies the Fed is relatively more aggressive at keeping both the level and volatility of inflation lower in regime 1. Therefore, we refer to regime 1 as the “more aggressive” regime and regime 2 as the “less aggressive” regime.

Note that in both regimes the long-run response to an increase in inflation is substantially more than one-for-one (i.e., $\phi^\pi > 1$), signifying that the Taylor Principle has held throughout the entire sample period. This result is in line with Orphanides (2004) and Davig and Doh (2014) but stands in stark contrast to the results of Murray et al. (2015), which finds one stabilizing regime where the Taylor Principle holds and one destabilizing regime where it does not hold. We also find notable differences from Murray et al. (2015) in the estimated inflation targets for each regime. Their estimate for the Fed’s inflation target in the destabilizing regime (where the Taylor Principle is violated) ranges from 21% to 23%, which they suggest implies that “the Fed did not have target level of inflation during those periods.” However, a misidentification of $\pi^T$ would contaminate the estimates for the weight on inflation. We find more reasonable estimates for the inflation target using both the baseline and extended sample in our more general model. There are a number of possible explanations for the contrasting findings of this study and that of Murray et al. (2015). Their model assumes full dependence in switching for the coefficient and volatility parameters, whereas we allow the timing of switches to be different for the two sets of parameters. Also, Murray et al. (2015) estimates a regime-switching rule using frequentist techniques where this paper uses Bayesian methods. Finally, there are slight differences in data across the two studies.

We should note that $\phi^\pi$ and $\phi^\gamma$ reported in table 2 are the long-term responses of the policy rate to a change in inflation or the output gap, respectively. Given different values for the smoothing parameter $\rho$, it is possible that the Fed reacts more to inflation in the short-run despite a relatively low value for $\phi^\pi$. Therefore, we calculate the change in the policy rate $\Delta i_{t+q}$ in the different regimes across different horizons $q$ due to a change in either inflation or the output gap. Figure 1 shows the posterior distribution for the regime-specific responses in the policy rate to a 100 basis point increase in either the
inflation gap (the top panel) or the output gap (the bottom panel). Across all horizons, the change in the policy rate due to inflation is substantially higher in regime 1 relative to regime 2. We find no considerable difference in the response to the output gap, though the posterior medians suggest the Fed responds more aggressively towards movements in the output gap in the short-run when following the parameterization outlined by regime 2.

5.2 Timing of Monetary Policy Regimes

Figure 2 plots the posterior regime probabilities for the two regime processes. The top two panels show the posterior means for the the coefficient regime variable $S^\theta_t$, while the bottom two panels show the posterior means for the the variance regime variable $S^\sigma_t$. One noticeable trend is that the Fed tends to be in the more aggressive regime during expansionary periods and switched to the less active regime during recessions. In the 1970s, we find that the Fed was in the less aggressive regime for 25 of the 40 quarters despite relatively high inflation at the time. The Fed did switch briefly to the more aggressive regime (e.g., 1971Q2 - 1973Q3) in order to combat rising inflation. Notably we find the Fed did switch to the more aggressive regime during the quarter when Volcker was appointed and he delivered his “Saturday Night Special.”

We describe the Fed’s behavior in the 1980s in three parts. In the early part of the decade during the double-dip recession, the Fed was in the less aggressive regime as it cut the policy rate despite high inflation. Monetary policy switched briefly to the more aggressive policy regime to combat inflation which increased above 19% between the double-dip. We find the Fed was relatively stable in the more aggressive regime in the mid-1980s as the Great Inflation ended and the Great Moderation began. In the late 1980s, the Fed switched to the less aggressive regime during the stock market crash associated with Black Monday and briefly switched back to the more aggressive regime as inflation began to rise once again.

From 1989Q3 - 1994Q1, the Fed entered a less aggressive phase as it cut rates after the savings and loan crisis and subsequent recession, despite inflation hitting 5%. The Fed
began issuing public statements in 1994Q1 to properly signal a policy change to raising rates as the economy recovered. Monetary policy remained in the more aggressive regime until 1996Q1 - 1997Q3 when it only increased the policy rate by a total of 25 bps despite relatively high average annualized quarterly output growth of 4.1% during this period and inflation picking up to 2.4%. The Fed kept policy stable until the LTCM crisis when it began cutting the policy rate.

In the early 2000s, the Fed began in the more aggressive regime but switched to a less aggressive phase at the onset of the dot-com recession. Interestingly, after the recessions of the early 1990s and 2001, we find the Fed remained in the less aggressive regime for a considerable amount of time after the recession ended. One possible explanation for this result is the Fed acting cautiously in response to the slow recovery in employment following these recessions as noted by Galí et al. (2012) and Jaimovich and Siu (2018), among others. They switched to the more aggressive regime in 2003Q4 as they expected inflation to substantially undershoot its target (the Greenbook forecast for inflation at this time was 0.99%). Despite the common belief that monetary policy was providing too much accommodation at this time, the Fed believed the output gap was negative and inflation would remain subdued. Thus based on the real time data in hand at the time, the model characterizes the Fed as being in the relatively aggressive regime as it consistently increased the policy rate by 25 bps at every meeting from June 2004 to June 2006. Leading up to the Financial Crisis, the Fed switched to the less aggressive regime as it began cutting rates despite a positive output gap and above target inflation.

Next, we discuss the timing of the volatility regime variable. The policy rate experiences high volatility during the early 1970s, the early 1980s, and the 2001 recession. This estimated volatility timing is similar to that of Davig and Doh (2014) which finds a low volatility regime present during the Great Moderation with the exception of 2001. However, their model solely attributes the period of Volcker disinflation to the high volatility regime rather than the more aggressive regime as one might expect. Our results corroborate their finding of a high volatility regime during this period, but suggest the Fed
entered the more aggressive regime between the two recessions of the early 1980s.\textsuperscript{11}

Another way to characterize the Fed’s historical behavior is by looking at the implied target rate of the estimated rule compared to the actual policy rate. Figure 4 presents the posterior median of the implied target rate $i_t^T$ compared to the actual effective federal funds rate $i_t$. It appears the Fed was acting in line with the estimated switching rule during the early 1970s and began deviating during the period when it targeted monetary aggregates (the late 1970s and early 1980s), holding the actual policy rate below the implied target. In the mid 1980s, the Fed began holding the policy rate above the implied target rate, possibly to combat any lingering inflationary fears. Since the beginning of the Great Moderation period, it appears the Fed set policy approximately in line with the estimated target policy rate. A notable exception occurs during the mid 2000s when the Fed set the policy rate below the implied target. Taylor (2011) describes this as one of the principle actions of what he calls the “Great Deviation”.

5.3 Signals of Monetary Policy Regimes

Are there variables that indicate a switch between the two policy regimes? The TVTP framework allows us to address this question by looking at how the covariates influence the transition probabilities. The TVTP in equation (5) are logistic functions, thus we must calculate marginal effects. We calculate the marginal effect for covariate $n$, $\delta_{ji}^n$, as the difference between (1) the transition probability $p_{ji,t}(Z_t)$ when covariate $n$ is one standard deviation above its average value (i.e., $Z_{nt} = \bar{Z}_n + \sigma_n Z_n$), and (2) the transition probability when covariate $n$ is one standard deviation below its average value (i.e., $Z_{nt} = \bar{Z}_n - \sigma_n Z_n$):

$$\delta_{ji}^n = p_{ji,t}(\bar{Z}_n + \sigma_n Z_n | Z_{-n} = \bar{Z}_{-n}) - p_{ji,t}(\bar{Z}_n - \sigma_n Z_n | Z_{-n} = \bar{Z}_{-n}).$$

(6)

while holding all other covariates besides $n$, $Z_{-n}$, at their average level $\bar{Z}_{-n}$. At each iteration of the Gibbs sampler, we calculate each covariate’s marginal effect based on the draw of $\gamma$.

\textsuperscript{11}These differences in results could come from a number of differences between this study and Davig and Doh (2014), including model specification, variables used, or sample size.
Table 3 presents the posterior median for each of the marginal effects along with the associated 67% HPD interval. We find no evidence that the dispersion in forecasters’ expectations for either inflation or output growth inform monetary policy switches. The primary indicators of regime switches are the anxious index and equity returns. When the anxious index rises (i.e., more forecasters expect a decline in real GDP), the Fed is more likely to switch from the more aggressive regime to the less aggressive regime. This result is intuitive as it suggests the Fed shifts attention away from its price stability mandate towards its full employment mandate when there is an impending recession. To put the estimated marginal effect in perspective, an increase in the anxious index from 0 (i.e., no forecasters expect a decline in real GDP) to 41.16 (i.e., 41.16% of forecasters expect a decline in real GDP) increases the probability of the Fed switching from the more aggressive regime 1 to the less aggressive regime 2 by 38 percentage points.

The persistence of the less aggressive regime is signaled by movements in equity returns. As equity returns increase, the probability of the Fed staying in the less aggressive regime falls and, by definition, they are more likely to switch to the more aggressive regime. Conversely, a fall in equity returns increases the probability the Fed will stay in the less aggressive regime. This result suggests that the Fed waits until equity prices recover to become more aggressive towards inflation. In context, the estimated marginal effects suggest that when equity returns go from -2.46% to +7.51% then the probability of switching from the less aggressive regime to the more aggressive regime goes up by 19 percentage points.

In order to show the importance of the TVTP framework, we plot the transition probabilities across time in figure 3. We opt to plot the persistence probabilities $p_{11,t}^\theta$ and $p_{22,t}^\theta$, since the switching probabilities can be inferred from the plots as $p_{21,t}^\theta = 1 - p_{11,t}^\theta$ and $p_{12,t}^\theta = 1 - p_{22,t}^\theta$. The persistence probability of the more aggressive regime (regime 1) stays relatively high during expansionary periods and falls considerably during recessions. This result illustrates the effect of the anxious index signaling a policy switch from the more aggressive regime to the less aggressive regime when a recession is expected. Conversely, the persistence probability of the less aggressive regime is relatively high during recessions.
and falls once the economy recovers. However, the persistence probability of the less aggressive regime is less volatile than the persistence probability of the more aggressive regime. This result implies that using TVTP offers more information for switches from the more aggressive regime to the less aggressive regime rather than for switches from the less aggressive regime to the more aggressive regime.

6 Robustness Checks

6.1 Alternative Specifications

In this section, we compare the baseline model to a number of alternative specifications. Table 4 shows the posterior distributions for the baseline model and four alternative models. The first column restates the posterior distribution for the baseline model discussed in the previous section.

The second column of table 4 shows the parameter estimates for the simple rule with persistence represented in equation (2). According to BIC, the baseline model fits the data better than the simple rule despite the penalty for a higher degree of sophistication. This finding suggests that there is some degree of switching in the monetary rule that is not captured by simple linear rules with persistence.

Since the timing of the regime switches in the baseline model appears to coincide somewhat with the NBER recession dates, we consider an alternative rule that switches based on expansionary and recessionary periods. The third column of table 4 displays the estimates for an interest rate rule which switches during recessions. Specifically, we estimate a recession-based interest rate rule which allows the parameters to switch depending upon if the economy is in expansion, $S^R_t = 1$, or recession, $S^R_t = 2$:

$$i_t = \rho^R i_{t-1} + (1 - \rho^R) \left[ C^R + \phi^R \pi_t + \phi^y y_t \right] + \varepsilon_t. \quad (7)$$

The error variance is also allowed to be stochastic across business cycle phases: $\varepsilon \sim N(0, \sigma^2_{S^R_t})$. This recession model differs from the baseline model since it imposes the
Fed’s reaction function switches based on recession dates, whereas the regime-switching of the baseline model reflects periods where Fed behavior differs the most according to the observed data.\textsuperscript{12}

In some respects, the parameter estimates for the recession rule mimic the results of the baseline model. During recessions, the Fed has a appreciably lower weight on inflation which intuitively follows from the central bank focusing less on inflation when unemployment is high. The recession rule estimates also show that the variance of the error term is larger during recessions relative to expansions. There are a number of possible explanations for this result. One possibility is that the Fed uses more discretion during recessions as it deviates from the target implied by the rule. Similarly, the Fed could be reacting to (potentially higher-frequency) data not included in our standard specification of the policy function. Finally, the higher error variance could be a result of either a time-varying neutral real interest rate or increased variability of factors outside the Fed’s control that move the short-term policy rate.

Despite the simplified framework, the recession rule fits the data better than the simple model but worse than the baseline model. This result suggests that the timing of monetary policy switches is not simply a function of the state of the business cycle.

We also consider two alternative Markov-switching specifications. The first alternative is a Markov-switching rule that assumes full dependence between the coefficient regime $S^\theta_t$ and the variance regime $S^\sigma_t$ which we label the MS-FD interest rate rule. The MS-FD model simplifies the baseline model by imposing that the coefficient and variance parameters switch contemporaneously. The fourth column of table 4 displays the parameter estimates for MS-FD model. The only considerable difference across the two regimes is the variance parameter, which is markedly lower in regime 1 relative to regime 2. The estimated timing of regime 2 in the MS-FD model matches well with the periods of high volatility in the baseline model. Using a analogous framework, Murray et al. (2015) finds similar timing in a Markov-switching model with full dependence. This

\textsuperscript{12}It is possible that the estimated monetary regimes switch in tandem with the business cycle, but the Markov-switching framework does not impose this timing. Therefore, the Markov-switching framework can be considered a more general case of the recession-based rule.
result leads us to conclude that the timing of switches will be dominated by the dynamics of variance in models estimated with a single regime variable governing both the coefficients and variance parameters.\textsuperscript{13} Thus, in order to determine if there are changes in the coefficients (and therefore the response to inflation and output), a model must control for heteroskedasticity either through a separate regime switching process or alternative time-varying volatility framework as in Check (2017). In terms of model fit, the MS-FD rule does a better job at matching the data than the simple rule and recession rule but performs worse than the baseline rule.

The second Markov-switching alternative that we consider is a model which loosens the restriction that the regime variables $S_t^\theta$ and $S_t^\sigma$ are fully independent and therefore allows them to be (imperfectly) correlated. This Markov-switching model with correlation, which we call the MS-C interest rate rule, can be thought of as an intermediate case between the extremes of full independence (as in the baseline model) and full dependence (as in the MS-FD model). Explicitly, the MS-C model assumes four aggregate regimes given by (2) which evolve according to TVTP given by

$$p_{ji,t}^\theta = \frac{\exp(\tilde{\gamma}_{ji} + \gamma_{ji}^Z t)}{\sum_{k=1}^{4} \exp(\tilde{\gamma}_{ki} + \gamma_{ki}^Z t)},$$

for $j = 1, \ldots, 4$.\textsuperscript{14} The estimates for the MS-C model are shown in the last column of table 4. The only appreciable difference across the coefficient regimes is the smoothing parameter $\rho$, which is smaller in the first regime. In results not reported, we calculate responses across different horizons and find that the short-run response to output is substantially higher in regime 1 and this regime is in effect for the entire sample except the mid-1980s.\textsuperscript{15} The MS-C model is the most generalized model considered and also the most parameterized. Due to this high degree of sophistication, BIC for the MS-C model shows that it is a poor fit compared to the baseline model. In fact the MS-C rule is outperformed by both the recession rule and MS-FD rule which are more parsimonious.

\textsuperscript{13}In preliminary work, Banaian and Lo (2010) finds a similar result that the variance switches dominate the regime timing when assumed to be contemporaneous with the coefficient switches.

\textsuperscript{14}In order to identify the model, we set $\tilde{\gamma}_{1i} = 0$ and $\gamma_{1i}^Z = 0$ for all $i = 1, \ldots, 4$.

\textsuperscript{15}Full results are available from the author upon request.
Overall, the baseline model with two regime variables that are fully independent fits the data the best compared to the four alternative models. Additionally, the simple rule with persistence is outperformed by both the recession-based rule and the Markov-switching rule with full dependence. This result suggests monetary policy studies which only use a simple Taylor-type rule do not capture the apparent parameter instability captured by nonlinear rules such as the ones considered in this paper.

6.2 Extended Sample

In our baseline results, the time period considered is prior to the Great Recession since the zero lower bound (ZLB) period may distort the regime estimates. One might be interested in how these baseline results compare to a sample including more recent periods. In this section we estimate the baseline and alternative models using data during the ZLB by substituting the observed federal funds rate with the shadow rate estimates from Wu and Xia (2016).

Table 5 presents the estimation results for the various rules using the extended sample from 1968:Q4 to 2012:Q4. The results for the simple, recession-based, and MS-FD rule are broadly similar to those using the baseline sample. The MS-C estimates no longer show a considerable difference in the smoothing parameter across the two coefficient regimes. In terms of model fit, the baseline model still fits the data the best according to BIC.

Figure 5 presents the posterior regime probabilities for both the coefficient regimes and variance regimes using the extended sample. Not surprisingly, the timing of each regime variable is highly correlated with the estimated timing using the baseline sample period.\(^{16}\) The high variance regime does not include the 2001 recession as in the estimated baseline timing, but it does include the Great Recession period. Additionally, the coefficient regime variable coincides with the NBER dates marginally more than when using the baseline sample.\(^{17}\)

The first column of table 5 displays the parameter estimates for the baseline model

---

\(^{16}\)The correlation of the coefficient regime across sample periods is 0.77 and the correlation of the variance regimes is 0.82.

\(^{17}\)The correlation of the NBER dates with regime 2 of the coefficient variable is 0.35 when using the baseline sample and 0.43 when using the extended sample.
using the extended sample. Regime 1 is again characterized by more aggressiveness towards inflation with a substantially higher $\phi^\pi$ and lower $\pi^T$. The primary difference in the results between using the extended sample compared to the baseline sample is in the Fed’s response to changes in the output gap. The Fed’s reaction to an increase in the output gap is considerably larger in regime 1 than regime 2. Figure 6 illustrates this result by plotting the implied change in the policy rate in response to changes in the inflation gap and the output gap across different horizons. The estimated response to an increase in the output gap is larger across all time periods except the first two quarters. Therefore when using the extended sample, we characterize regime 1 as overall more aggressive monetary policy since it reacts relatively strongly to both the inflation gap and the output gap, whereas the baseline sample suggest they only react more aggressively to inflation.

Lastly, table 6 shows the marginal effects of each transition covariate on the transition dynamics for $S_\theta$. We find qualitatively similar results to our baseline sample estimation, though the HPD intervals have widened a bit for some of the marginal effects. Nonetheless, the anxious index and equity prices are still important signals of monetary policy switching.

7 Conclusion

Describing monetary policy through simple interest rates rules is standard in modern macroeconomic models. This study investigated whether or not the assumption of a stable interest rate rule is appropriate, and if certain variables signal a switch in monetary policy.

We found that U.S. monetary policy is better described by regime-switching rules overall. Allowing for policy switches based on recession dating improved model fit compared to simple rules, but both were outperformed by the more general Markov-switching model. In particular, we found the best description of the Fed’s reaction function over the past 43 years is a Markov-switching model with two regime processes. One regime variable allows us to control for heteroskedasticity of the policy rate, while the other cap-
tures changes in the interest rate rule’s coefficients. Other regime-switching specifications failed to outperform the model which assumes full independence between the two regime variables.

The Fed appears to switch between an aggressive regime with a high coefficient on inflation, and a more dovish regime that is relatively less responsive to inflationary movements. Using a time-varying transition probability framework, we identified two variables which signal a switch between these two monetary policy regimes. A rise in the anxious index, which captures forecasters’ expectations of an impending recession, signals the end of aggressive monetary policy towards a more dovish stance. As equity returns recover, the Fed is more likely to switch from the dovish regime back to the more aggressive regime. For robustness, we estimate the model using an extended sample which includes the Great Recession and find that the more aggressive regime is more active against movements in both the inflation gap and output gap.

The results of this paper have a number of implications for future research. First, studies on monetary policy need to account for switches in rules-based policy. In particular, one must control for heteroskedasticity in the policy rule in order to obtain proper identification of evolving coefficients. Second, forecasters of Fed policy should focus on recession forecasts (i.e., the anxious index) and equity returns when determining a impending shift in monetary policy. Lastly, the Fed is transparent in considering multiple rules when setting the short-term policy rate.\textsuperscript{18} It is possible they act as this study suggests by switching between individual rules. Or, the Fed could use a weighted average of implied policy rates from various rules at any given time period.

References


Brüggemann, Ralf, and Jana Riedel (2011) ‘Nonlinear interest rate reaction functions for the uk.’ Economic Modelling 28(3), 1174 – 1185


Castro, Vítor (2011) ‘Can central banks’ monetary policy be described by a linear (augmented) taylor rule or by a nonlinear rule?’ Journal of Financial Stability 7(4), 228 – 246


Choi, Jason, and Andrew Foerster (2016) ‘Optimal monetary policy regime switches.’ Federal Reserve Bank of Kansas City working paper


Creel, Jérôme, and Paul Hubert (2015) ‘Has inflation targeting changed the conduct of monetary policy?’ Macroeconomic Dynamics 19(1), 1–21


Galí, Jordi, Frank Smets, and Rafael Wouters (2012) ‘Slow recoveries: A structural interpretation.’ *Journal of Money, Credit and Banking* 44(s2), 9–30


Kazanas, Thanasssis, Apostolis Philippopoulos, and Elias Tzavalis (2011) ‘Monetary policy rules and business cycle conditions*.’ *The Manchester School* 79(s2), 73–97


Wu, Jing Cynthia, and Fan Dora Xia (2016) ‘Measuring the macroeconomic impact of monetary policy at the zero lower bound.’ *Journal of Money, Credit and Banking* 48(2-3), 253–291
A Estimation Details

This appendix outlines the draws for the Gibbs sampler of the baseline model. It is useful to define the entire time series of observed variables as $Y^T = \{i_t, \pi_t, y_t\}_{t=1}^T$ and the covariates as $Z^T = \{Z_t\}_{t=1}^T$. Similarly we define The parameters and latent variables are partitioned into four blocks to be drawn separately:

1. Draw the coefficient parameters $\theta$ conditional on $Y^T, S^T, \sigma^2$
2. Draw the variance parameters $\sigma^2$ conditional on $Y^T, S^T, \theta$
3. Draw the latent regime variable $S^T$ conditional on $Y^T, \theta, \sigma^2, P^T$
4. Draw the transition matrices $P^T$ conditional on $S^T, Z^T$

A.1 Draw of $\theta$ given $Y^T, S^T, \sigma^2$

In order to draw the set of regime-dependent coefficients $\theta = [\rho, C, \phi^\pi, \phi^y]'$, we rewrite equation (4) as:

$$ i_t = \rho S_t i_{t-1} + C_S \pi_t + \phi^\pi_S \pi_t + \phi^y_S y_t + \varepsilon_t, $$

where the short run coefficients are $\hat{C}_k = C_k(1-\rho_k), \hat{\phi}_k^\pi = \phi_k^\pi(1-\rho_k)$, and $\hat{\phi}_k^y = \phi_k^y(1-\rho_k)$. The entire set of coefficients for both regimes is given by

$$ \hat{\theta} = \left[ \begin{array}{cccc} \rho_1 & \hat{C}_1 & \hat{\phi}_1^\pi & \hat{\phi}_1^y \\ \rho_2 & \hat{C}_2 & \hat{\phi}_2^\pi & \hat{\phi}_2^y \end{array} \right]. $$

Let

$$ x_t = \frac{1}{\sigma S_t} \left[ \tilde{i}_{1,t-1} I(S^\theta_t = 1) \quad \tilde{\pi}_{1,t} \quad \tilde{y}_{1,t} \quad \tilde{i}_{2,t-1} \quad I(S^\theta_t = 2) \quad \tilde{\pi}_{2,t} \quad \tilde{y}_{2,t} \right], $$

where $\tilde{i}_{k,t-1} = i_{t-1} I(S^\theta_t = k), \tilde{\pi}_{k,t} = \pi_t I(S^\theta_t = k)$, and $\tilde{y}_{k,t} = y_t I(S^\theta_t = k)$. Assuming a prior distribution given by $\hat{\theta} \sim N(t_0, T_0)$, the posterior distribution is:

$$ \hat{\theta} \sim N(t_1, T_1), $$

27
where

\[ t_1 = T_1(T_0^{-1}t_0 + X^T\hat{i}_T), \]
\[ T_1 = (T_0^{-1} + X^TX), \]

and \( X^T = [x_1, ..., x_T]' \) and \( \hat{i}^T = \left[ \frac{i_1}{\sigma^2_1}, ..., \frac{i_T}{\sigma^2_T} \right] \). We redraw any draw of \( \hat{\theta} \) which does not satisfy the stationarity condition \( |\rho_k| < 1 \) for \( k = 1, 2 \).

### A.2 Draw of \( \sigma^2 \) given \( Y^T, S^T, \theta \)

Following Kim and Nelson (1999), we set \( \sigma^2_2 = \sigma^2_1(1 + h) \). We proceed by drawing \( \sigma^2_1 \) given \( h \), then drawing \( h \) given the new draw for \( \sigma^2_1 \). To draw \( \sigma^2_1 \), we begin by defining

\[ \tilde{x}_t = \frac{1}{\sqrt{1 + hI(S^T_{\sigma} = 2)}} \left[ \tilde{i}_{1,t-1} I(S^\theta_t = 1) \hat{\pi}_{1,t} \tilde{y}_{1,t} \tilde{i}_{2,t-1} I(S^\theta_t = 2) \hat{\pi}_{2,t} \tilde{y}_{2,t} \right]', \]

Assuming the prior \( \sigma^2_1 \sim IG(\nu_{10}/2, \tau_{10}/2) \), the posterior is

\[ \sigma^2_1 \sim IG \left( \nu_{10} + T, \tau_{10} + \sum_{t=1}^T (\tilde{i}_t - \hat{\theta}'\tilde{x}_t)^2 \right). \]

Given the draw for \( \sigma^2_1 \), we draw \( h \) by first defining \( \bar{h} = 1 + h \) and

\[ \tilde{x}_t = \frac{1}{\sigma_1} \left[ \tilde{i}_{1,t-1} I(S^\theta_t = 1) \hat{\pi}_{1,t} \tilde{y}_{1,t} \tilde{i}_{2,t-1} I(S^\theta_t = 2) \hat{\pi}_{2,t} \tilde{y}_{2,t} \right]' \]

Assuming the prior \( \bar{h} \sim IG(\nu_{20}/2, \tau_{20}/2) \), the posterior for \( \bar{h} \) is given by

\[ \bar{h} \sim IG \left( \nu_{20} + |T^\sigma_2|, \tau_{20} + \sum_{T^\sigma_2} (\tilde{i}_t - \hat{\theta}'\tilde{x}_t)^2 \right). \]

where the set \( T^\sigma_2 \) contains the time periods where \( S^\sigma_t = 2 \) and \( |T^\sigma_2| \) is the cardinality of that set.
A.3 Draw of $S^T$ given $Y^T$, $\theta$, $\sigma^2$

In order to draw the time series of the latent variable $S^T$, we utilize the filter first outlined by Hamilton (1989) and augment it for time-varying transition probabilities similar to Kaufmann (2015).

We first run the filter forward for $t = 1, \ldots, T$, obtaining

$$p(S_t|Y^t) = \frac{f(i_t|S_t, Y^{t-1})p(S_t|Y^{t-1})}{f(i_t|Y^{t-1})},$$

where

$$f(i_t|S_t, Y^{t-1}) = \frac{1}{\sqrt{2\pi\sigma_S^2}} \exp \left[ -\frac{(i_t - \theta_S' \tilde{x}_t)^2}{2\sigma_S^2} \right],$$

$$p(S_t|Y^{t-1}) = \sum_{j=1}^{4} p(S_t|S_{t-1} = j, Z_t)p(S_{t-1} = j|Y^{t-1}),$$

$$f(i_t|Y^{t-1}) = \sum_{k=1}^{4} f(i_t|S_t = k, Y^{t-1})p(S_t = k|Y^{t-1}),$$

and $\tilde{x}_t = [i_{t-1}, 1, \pi_t, y_t]'$. The last iteration of the filter gives us $p(S_T|Y^T)$ from which we draw the terminal state $S_T$. Running the smoother outlined by Chib (1996), we draw $S_{T-1}, \ldots, S_1$ recursively from

$$p(S_t|S_{t+1}) = \frac{p(S_{t+1}|S_t, Z_{t+1})p(S_t|Y^t)}{\sum_{k=1}^{4} p(S_{t+1}|S_t = k, Z_{t+1})p(S_t = k|Y^t)}.$$ 

A.4 Draw of $P^T$ given $S^T$, $Z^T$

We draw the transition matrices in two steps. First, we draw the constant transition probability parameters that determine $P^\sigma$. Then, we draw the parameters in the time-varying transition probability matrix $P^\theta_t$.

We assume a prior for $P^\sigma_k = [p_{1k}, p_{2k}]'$ given by

$$P^\sigma_k \sim D(\bar{p}_{1k}, \bar{p}_{2k}).$$
The posterior distribution is then

$$p_k^\sigma \sim D(\bar{p}_{1k} + N_{1k}(S^\sigma), \bar{p}_{2k} + N_{2k}(S^\sigma)),$$

where $N_{jk}(S^\sigma)$ counts the number of transitions from variance regime $k$ to regime $j$ in $S^\sigma$).

The draw for the TVTP parameters $\gamma$ follows directly from Appendix A.2 of Kaufmann (2015). We set regime 1 of $S_t^\theta$ to be the normalizing state, which sets $\gamma_{1k} = 0$ for $k = 1, 2$. We adopt their centered parameterization for $Z_t$ implying the covariates are demeaned.
B Tables and Figures

Table 1: Prior Distributions for Estimation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Prior Distribution</th>
<th>Hyperparameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>$N(r_0, R_0)$</td>
<td>$r_0 = [0, 0]'$, $R_0 = I_2$</td>
</tr>
<tr>
<td>$\phi^x$</td>
<td>$N(p_0, P_0)$</td>
<td>$p_0 = [0, 0]'$, $P_0 = I_2$</td>
</tr>
<tr>
<td>$\phi^y$</td>
<td>$N(q_0, Q_0)$</td>
<td>$q_0 = [0, 0]'$, $Q_0 = I_2$</td>
</tr>
<tr>
<td>$\sigma^2_1$</td>
<td>$IG(\nu_{10}/2, \tau_{10}/2)$</td>
<td>$\nu_{10} = 0.10$, $\tau_{10} = 0.10$</td>
</tr>
<tr>
<td>$\bar{h}$</td>
<td>$IG(\nu_{20}/2, \tau_{20}/2)$</td>
<td>$\nu_{20} = 0.10$, $\tau_{20} = 0.10$</td>
</tr>
<tr>
<td>$\hat{\gamma}_{2i}$</td>
<td>$N(g_0, G_0)$</td>
<td>$g_0 = 0$, $G_0 = I_{L+1}$</td>
</tr>
<tr>
<td>$p^*_k$</td>
<td>$D(\bar{p}<em>{1k}, \bar{p}</em>{2k})$</td>
<td>$\bar{p}_{jk}$ for $j = 1, 2$</td>
</tr>
</tbody>
</table>
Table 2: Posterior Distribution for Baseline Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Median</th>
<th>68% Posterior Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_1$</td>
<td>0.81</td>
<td>[0.76, 0.85]</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>0.77</td>
<td>[0.73, 0.81]</td>
</tr>
<tr>
<td>$C_1$</td>
<td>-0.36</td>
<td>[-1.22, 0.43]</td>
</tr>
<tr>
<td>$C_2$</td>
<td>0.34</td>
<td>[-0.31, 0.96]</td>
</tr>
<tr>
<td>$\phi^\pi_1$</td>
<td>2.80</td>
<td>[2.48, 3.24]</td>
</tr>
<tr>
<td>$\phi^\pi_2$</td>
<td>1.67</td>
<td>[1.52, 1.85]</td>
</tr>
<tr>
<td>$\phi^y_1$</td>
<td>0.62</td>
<td>[0.45, 0.77]</td>
</tr>
<tr>
<td>$\phi^y_2$</td>
<td>0.60</td>
<td>[0.51, 0.69]</td>
</tr>
<tr>
<td>$\pi^T_1$</td>
<td>1.58</td>
<td>[1.29, 1.80]</td>
</tr>
<tr>
<td>$\pi^T_2$</td>
<td>3.17</td>
<td>[2.66, 3.68]</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.24</td>
<td>[0.21, 0.28]</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>1.35</td>
<td>[1.12, 1.62]</td>
</tr>
</tbody>
</table>

Note: This table shows the posterior distributions of the model parameters for the baseline Markov-switching model with two regime variables and TVTP. These estimates were constructed using the sample period 1967Q1 - 2007:Q4.
Table 3: Determinants of Switching for the Coefficient Regime

<table>
<thead>
<tr>
<th></th>
<th>$\sigma_{\pi,t}$</th>
<th>$\sigma_{y,t}$</th>
<th>$Anx_t$</th>
<th>$ER_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{11}^n$</td>
<td>0.10</td>
<td>-0.26</td>
<td>-0.39</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>[-0.07,0.29]</td>
<td>[-0.59,0.07]</td>
<td>[-0.66,-0.11]</td>
<td>[-0.18,0.19]</td>
</tr>
<tr>
<td>$\delta_{21}^n$</td>
<td>-0.10</td>
<td>0.26</td>
<td>0.39</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>[-0.29,0.07]</td>
<td>[-0.07,0.59]</td>
<td>[0.11,0.66]</td>
<td>[-0.19,0.18]</td>
</tr>
<tr>
<td>$\delta_{12}^n$</td>
<td>0.04</td>
<td>0.02</td>
<td>-0.02</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>[-0.08,0.17]</td>
<td>[-0.10,0.10]</td>
<td>[-0.15,0.10]</td>
<td>[0.06,0.34]</td>
</tr>
<tr>
<td>$\delta_{22}^n$</td>
<td>-0.04</td>
<td>-0.02</td>
<td>0.02</td>
<td>-0.19</td>
</tr>
<tr>
<td></td>
<td>[-0.17,0.08]</td>
<td>[-0.10,0.10]</td>
<td>[0.10,0.15]</td>
<td>[-0.34,-0.06]</td>
</tr>
</tbody>
</table>

Note: This table shows the posterior distributions of the marginal effects of each transition covariate for the regime process governing the coefficient parameters using the baseline sample period 1968:Q4 - 2007:Q4. The marginal effect shown in equation (6) is the difference in the transition probability assuming a transition covariate $n$ is one standard deviation above the average ($p_{ji}(\bar{Z}+\sigma_{Z})$) and one standard deviation below the average ($p_{ji}(\bar{Z}-\sigma_{Z})$). The 68% posterior coverage is shown in parenthesis below the median parameter draw. Bold indicates the posterior coverage interval does not include 0.
Table 4: Posterior Distributions for the Baseline and Alternative Models

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baseline</th>
<th>Simple</th>
<th>Recession</th>
<th>MS-FD</th>
<th>MS-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho_1 )</td>
<td>0.81</td>
<td>0.82</td>
<td>0.88</td>
<td>0.74</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>[0.76, 0.85]</td>
<td>[0.79, 0.85]</td>
<td>[0.85, 0.91]</td>
<td>[0.68, 0.81]</td>
<td>[0.74, 0.82]</td>
</tr>
<tr>
<td>( \rho_2 )</td>
<td>0.77</td>
<td>-</td>
<td>0.84</td>
<td>0.82</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>[0.73, 0.81]</td>
<td></td>
<td>[0.72, 0.94]</td>
<td></td>
<td>[0.77, 0.88]</td>
</tr>
<tr>
<td>( C_1 )</td>
<td>-0.36</td>
<td>0.64</td>
<td>-2.10</td>
<td>0.87</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>[-1.22, 0.43]</td>
<td>[-0.35, 1.50]</td>
<td>[-3.92, -0.84]</td>
<td>[0.19, 1.28]</td>
<td>[0.14, 1.34]</td>
</tr>
<tr>
<td>( C_2 )</td>
<td>0.34</td>
<td>-</td>
<td>0.80</td>
<td>2.78</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>[-0.31, 0.96]</td>
<td></td>
<td>[-5.12, 4.80]</td>
<td></td>
<td>[0.20, 5.38]</td>
</tr>
<tr>
<td>( \phi_1^T )</td>
<td>2.80</td>
<td>2.05</td>
<td>3.25</td>
<td>1.77</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td>[2.48, 3.24]</td>
<td>[1.79, 2.37]</td>
<td>[2.77, 4.00]</td>
<td>[1.58, 2.09]</td>
<td>[1.58, 2.09]</td>
</tr>
<tr>
<td>( \phi_2^T )</td>
<td>1.67</td>
<td>-</td>
<td>0.81</td>
<td>1.80</td>
<td>3.12</td>
</tr>
<tr>
<td></td>
<td>[1.52, 1.85]</td>
<td></td>
<td>[-0.47, 1.78]</td>
<td></td>
<td>[1.35, 2.44]</td>
</tr>
<tr>
<td>( \phi_1^y )</td>
<td>0.62</td>
<td>0.64</td>
<td>0.78</td>
<td>0.76</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>[0.45, 0.77]</td>
<td>[0.50, 0.83]</td>
<td>[0.60, 1.03]</td>
<td>[0.68, 0.84]</td>
<td>[0.59, 0.79]</td>
</tr>
<tr>
<td>( \phi_2^y )</td>
<td>0.60</td>
<td>-</td>
<td>0.44</td>
<td>0.67</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>[0.51, 0.69]</td>
<td></td>
<td>[-0.19, 2.01]</td>
<td></td>
<td>[0.39, 1.10]</td>
</tr>
<tr>
<td>( \pi_1^T )</td>
<td>1.58</td>
<td>1.78</td>
<td>2.04</td>
<td>2.15</td>
<td>2.11</td>
</tr>
<tr>
<td></td>
<td>[1.29, 1.80]</td>
<td>[1.14, 2.30]</td>
<td>[1.74, 2.32]</td>
<td>[1.91, 2.35]</td>
<td>[1.82, 2.38]</td>
</tr>
<tr>
<td>( \pi_2^T )</td>
<td>3.17</td>
<td>-</td>
<td>1.73</td>
<td>-0.11</td>
<td>0.98</td>
</tr>
<tr>
<td>( \sigma_1 )</td>
<td>0.24</td>
<td>0.95</td>
<td>0.68</td>
<td>0.29</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>[0.21, 0.28]</td>
<td>[0.88, 0.98]</td>
<td>[0.64, 0.72]</td>
<td>[0.24, 0.34]</td>
<td>[0.28, 0.37]</td>
</tr>
<tr>
<td>( \sigma_2 )</td>
<td>1.35</td>
<td>-</td>
<td>1.50</td>
<td>1.34</td>
<td>1.49</td>
</tr>
<tr>
<td></td>
<td>[1.12, 1.62]</td>
<td></td>
<td>[1.26, 1.82]</td>
<td></td>
<td>[1.16, 1.56]</td>
</tr>
</tbody>
</table>

Note: This table shows the posterior distributions of the model parameters for: [1] the baseline Markov-switching rule with two regime variables shown in (4), [2] the simple rule with persistence shown in (2), [3] the recession-based rule shown in (7), [4] the Markov-switching rule with full dependence between \( S_t^R \) and \( S_t^q \), [5] the Markov-switching rule with correlation between \( S_t^R \) and \( S_t^q \). For the recession-based rule, \( S_t \) reflects NBER business cycle dating, with \( S_t^R = 1 \) representing expansions and \( S_t^R = 2 \) representing recessions. These estimates were constructed using the sample period 1967Q1 - 2007:Q4. The 68% posterior coverage is shown in parenthesis below the median parameter draw.
The 68% posterior coverage is shown in parenthesis below the median parameter draw. These estimates were constructed using the sample period 1968Q4 - 2012:Q4. For the recession-based rule, \( S_t \) reflects NBER business cycle dating, with \( S_t^R \) = 1 representing expansions and \( S_t^R \) = 2 representing recessions. These estimates were constructed using the sample period 1968Q4 - 2012:Q4. The 68% posterior coverage is shown in parenthesis below the median parameter draw.

Table 5: Posterior Distributions for the Baseline and Alternative Models – Extended Sample

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baseline</th>
<th>Simple</th>
<th>Recession</th>
<th>MS-FD</th>
<th>MS-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho_1 )</td>
<td>0.73</td>
<td>0.83</td>
<td>0.85</td>
<td>0.76</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>[0.62,0.83]</td>
<td>[0.79,0.86]</td>
<td>[0.82,0.88]</td>
<td>[0.72,0.79]</td>
<td>[0.74,0.80]</td>
</tr>
<tr>
<td>( \rho_2 )</td>
<td>0.81</td>
<td>-</td>
<td>0.86</td>
<td>0.81</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>[0.75,0.85]</td>
<td></td>
<td>[0.76,0.95]</td>
<td>[0.74,0.88]</td>
<td>[0.75,0.97]</td>
</tr>
<tr>
<td>( C_1 )</td>
<td>-0.39</td>
<td>0.25</td>
<td>-1.19</td>
<td>0.55</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>[-1.49,0.18]</td>
<td>[-0.73,1.17]</td>
<td>[-2.43,-0.20]</td>
<td>[0.08,0.99]</td>
<td>[0.03,0.99]</td>
</tr>
<tr>
<td>( C_2 )</td>
<td>-1.47</td>
<td>-</td>
<td>-0.30</td>
<td>2.26</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>[-2.86,-0.07]</td>
<td></td>
<td>[-5.51,3.54]</td>
<td>[-0.61,4.88]</td>
<td>[-7.55,8.34]</td>
</tr>
<tr>
<td>( \phi_1^T )</td>
<td>2.57</td>
<td>2.07</td>
<td>2.78</td>
<td>1.88</td>
<td>1.87</td>
</tr>
<tr>
<td></td>
<td>[2.29,3.32]</td>
<td>[1.82,2.35]</td>
<td>[2.45,3.21]</td>
<td>[1.73,2.05]</td>
<td>[1.71,2.06]</td>
</tr>
<tr>
<td>( \phi_2^T )</td>
<td>1.86</td>
<td>-</td>
<td>0.30</td>
<td>1.60</td>
<td>1.82</td>
</tr>
<tr>
<td></td>
<td>[1.64,2.07]</td>
<td></td>
<td>[-1.88,1.22]</td>
<td>[1.17,2.22]</td>
<td>[0.55,7.15]</td>
</tr>
<tr>
<td>( \phi_1^Y )</td>
<td>0.72</td>
<td>0.57</td>
<td>0.57</td>
<td>0.70</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>[0.61,0.85]</td>
<td>[0.43,0.74]</td>
<td>[0.44,0.73]</td>
<td>[0.63,0.77]</td>
<td>[0.66,0.74]</td>
</tr>
<tr>
<td>( \phi_2^Y )</td>
<td>0.40</td>
<td>-</td>
<td>0.00</td>
<td>0.35</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>[0.25,0.54]</td>
<td></td>
<td>[-0.69,0.90]</td>
<td>[0.02,0.79]</td>
<td>[-0.89,2.05]</td>
</tr>
<tr>
<td>( \pi_1^T )</td>
<td>1.82</td>
<td>2.11</td>
<td>2.08</td>
<td>2.20</td>
<td>2.25</td>
</tr>
<tr>
<td></td>
<td>[1.56,2.01]</td>
<td>[1.44,2.68]</td>
<td>[1.72,2.42]</td>
<td>[1.95,2.45]</td>
<td>[1.95,2.56]</td>
</tr>
<tr>
<td>( \pi_2^T )</td>
<td>4.60</td>
<td>-</td>
<td>-0.77</td>
<td>1.02</td>
<td>1.11</td>
</tr>
<tr>
<td>( \sigma_1 )</td>
<td>0.31</td>
<td>1.09</td>
<td>0.89</td>
<td>0.41</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>[0.28,0.35]</td>
<td>[1.03,1.15]</td>
<td>[0.84,0.94]</td>
<td>[0.36,0.46]</td>
<td>[0.38,0.48]</td>
</tr>
<tr>
<td>( \sigma_2 )</td>
<td>1.71</td>
<td>-</td>
<td>1.52</td>
<td>1.74</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>[1.49,2.01]</td>
<td></td>
<td>[1.29,1.79]</td>
<td>[1.53,2.00]</td>
<td>[1.63,2.26]</td>
</tr>
<tr>
<td><strong>BIC</strong></td>
<td>478.40</td>
<td>550.88</td>
<td>527.79</td>
<td>497.74</td>
<td>527.89</td>
</tr>
</tbody>
</table>

Note: This table shows the posterior distributions of the model parameters for: [1] the baseline Markov-switching rule with two regime variables shown in (4), [2] the simple rule with persistence shown in (2), [3] the recession-based rule shown in (7), [4] the Markov-switching rule with full dependence between \( S_t^R \) and \( S_t^Y \), [5] the Markov-switching rule with correlation between \( S_t^R \) and \( S_t^Y \). For the recession-based rule, \( S_t \) reflects NBER business cycle dating, with \( S_t^R = 1 \) representing expansions and \( S_t^R = 2 \) representing recessions. These estimates were constructed using the sample period 1968Q4 - 2012:Q4. The 68% posterior coverage is shown in parenthesis below the median parameter draw.
Table 6: Determinants of Switching for the Coefficient Regime – Extended Sample

<table>
<thead>
<tr>
<th></th>
<th>$\sigma_{\pi,t}$</th>
<th>$\sigma_{y,t}$</th>
<th>$Anx_t$</th>
<th>$ER_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{11}^n$</td>
<td>0.09</td>
<td>-0.24</td>
<td><strong>0.38</strong></td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>[-0.14,0.31]</td>
<td>[-0.58,0.05]</td>
<td>[-0.66,-0.11]</td>
<td>[-0.20,0.17]</td>
</tr>
<tr>
<td>$\delta_{21}^n$</td>
<td>-0.09</td>
<td>0.24</td>
<td><strong>0.38</strong></td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>[-0.31,0.14]</td>
<td>[-0.05,0.58]</td>
<td>[0.11,0.66]</td>
<td>[-0.17,0.20]</td>
</tr>
<tr>
<td>$\delta_{12}^n$</td>
<td>0.00</td>
<td>0.03</td>
<td>-0.07</td>
<td><strong>0.15</strong></td>
</tr>
<tr>
<td></td>
<td>[-0.16,0.29]</td>
<td>[-0.08,0.14]</td>
<td>[-0.27,0.07]</td>
<td>[0.02,0.33]</td>
</tr>
<tr>
<td>$\delta_{22}^n$</td>
<td>0.00</td>
<td>-0.03</td>
<td>0.07</td>
<td><strong>-0.15</strong></td>
</tr>
<tr>
<td></td>
<td>[-0.29,0.16]</td>
<td>[-0.14,0.08]</td>
<td>[-0.07,0.27]</td>
<td>[-0.33,-0.02]</td>
</tr>
</tbody>
</table>

Note: This table shows the posterior distributions of the marginal effects of each transition covariate for the regime process governing the coefficient parameters using the baseline sample period 1968Q4 - 2012:Q4. The marginal effect shown in equation (6) is the difference in the transition probability assuming a transition covariate $n$ is one standard deviation above the average ($p_{ji}(\bar{Z} + \sigma_Z)$) and one standard deviation below the average ($p_{ji}(\bar{Z} - \sigma_Z)$). The 68% posterior coverage is shown in parenthesis below the median parameter draw. Bold indicates the posterior coverage interval does not include 0.
Figure 1: Policy Response to Inflation and Output Gap

(a) $\Delta i_{t+q}$ due to $\Delta \pi_t = 100$ bps

(b) $\Delta i_{t+q}$ due to $\Delta y_t = 100$ bps

Note: This figure shows the marginal change in the policy rate ($i$) at different time horizons ($q$) to a change at time $t$ in either inflation ($\pi$) or the output gap ($y$). The thin dotted lines represent the 68% highest posterior density interval for each marginal effect.
Figure 2: Regime Probabilities – Baseline Sample

(a) Coefficient Regime \( (S^θ_t) \)

(b) Variance Regime \( (S^σ_t) \)

Note: This figure shows the posterior probability of each regime for the benchmark Markov-switching rule with two different regime processes and time-varying transition probabilities. The first regime variable \( S^θ_t \) determines the coefficients in the interest rate rule, and the second regime variable \( S^σ_t \) determines the error variance. Gray bars represent NBER recession dates for the U.S. The estimation period is 1968:Q4 - 2007:Q4.
Figure 3: Time-Varying Transition Probabilities for Coefficient Regime $S_t^\theta$

Note: This figure shows the posterior median (solid line) and 68% highest posterior density interval (dotted lines) for the time-varying transition probabilities given by equation (5). Gray bars represent NBER recession dates for the U.S. The estimation period is the baseline sample of 1968:Q4 - 2007:Q4.
Figure 4: Actual Policy Rate ($i_t$) and Implied Target Rate ($i_t^*$)

Note: This figure shows the actual federal funds rate and the implied target rate from the benchmark model. Gray bars represent NBER recession dates for the U.S. The estimation period is 1968:Q4 - 2007:Q4.
Figure 5: Regime Probabilities – Extended Sample

(a) Coefficient Regime ($S_\theta^t$)

(b) Variance Regime ($S_\sigma^t$)

Note: This figure shows the posterior probability of each regime for the benchmark Markov-switching rule with two different regime processes and time-varying transition probabilities. The first regime variable $S_\theta^t$ determines the coefficients in the interest rate rule, and the second regime variable $S_\sigma^t$ determines the error variance. Gray bars represent NBER recession dates for the U.S. The estimation period is 1968:Q4 - 2012:Q4.
Figure 6: Policy Response to Inflation and Output Gap – Extended Sample

(a) $\Delta i_{t+q}$ due to $\Delta \pi_t = 100$ bps

(b) $\Delta i_{t+q}$ due to $\Delta y_t = 100$ bps

Note: This figure shows the marginal change in the policy rate ($i$) at different time horizons $q$ to a change at time $t$ in either inflation ($\pi$) or the output gap ($y$). The thin dotted lines represent the 68% highest posterior density interval for each marginal effect. The estimation period is 1968:Q4 - 2012:Q4.