Identifying the Stance of Monetary Policy at the Zero Lower Bound: A Markov-switching Estimation Exploiting Monetary-Fiscal Policy Interdependence

Manuel Gonzalez-Astudillo *
Federal Reserve Board
Washington, D.C., USA
manuel.p.gonzalez-astudillo@frb.gov

February 14, 2014

Abstract

In this paper, I propose an econometric technique to estimate a Markov-switching Taylor rule subject to the zero lower bound of interest rates. I show that the estimation of a Markov-switching regression when the dependent variable is censored yields estimates of the coefficients of the model and the transition probabilities that are inconsistent, while the prevalent regime may not be identified. The incorporation of a Tobit specification to the Markov-switching regression allows for a consistent estimation of the switching coefficients and the transition probabilities, but the prevalent regime remains unidentified. Linking the switching of the parameters of interest to the switching of the coefficients of an auxiliary uncensored Markov-switching regression, in addition to the Tobit specification, improves the identification of the prevalent regime. For the estimation, I use U.S. quarterly data spanning 1960:1-2013:1. The chosen auxiliary Markov-switching regression is a fiscal policy rule where taxes respond to debt and the output gap. Results show that at the end of the sample the economy is likely in a regime of passive monetary policy and debt-stabilizing fiscal policy. Additionally, results show that there is evidence of policy co-movements with debt-stabilizing fiscal policy more likely accompanying active monetary policy, and vice versa.

Keywords: Markov-switching coefficients, zero lower bound, monetary-fiscal policy interactions

JEL Classification Numbers: C34, E52, E63

---

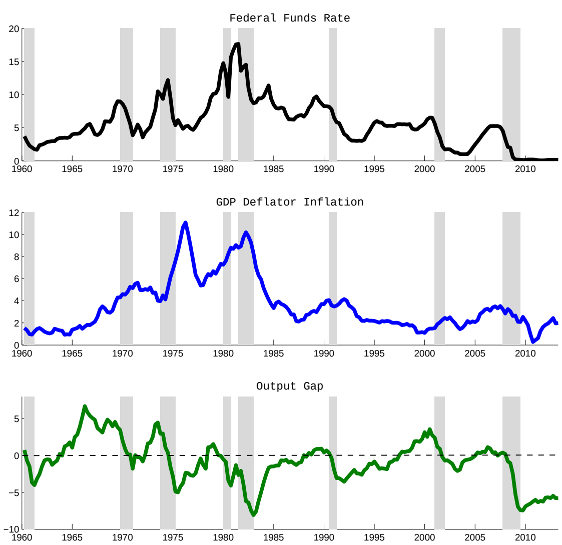
*The views expressed in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System.

1 Introduction

The forward guidance provided by the Federal Open Market Committee in its most recent statements indicates that "a highly accommodative stance of monetary policy will remain appropriate for a considerable time after the asset purchase program ends and the economic recovery strengthens." In particular, the forward guidance in place sets exceptionally low federal funds rates between 0 and 1/4 percent, defining an effective lower bound.

At least since Clarida et al. (2000), we have known that the monetary-policy regime can change. One could infer the stance of monetary policy, as measured by the strength of the reaction of the federal funds rate with respect to inflation deviations from target, by estimating a Markov-switching coefficients Taylor rule and obtaining the prevalent regime. Unfortunately, as shown in Figure 1, because of the current effective lower bound the federal funds rate does not react to fluctuations in the inflation rate and the output gap. This introduces an important censoring problem in the estimation of monetary-policy rules and poses identification problems to the estimation of the prevalent regime.

Figure 1: Evolution of Interest Rate, Inflation and the Output Gap



Notes: Shaded areas are NBER recession periods.

In this paper, I exploit monetary-fiscal policy interdependence and develop an estimation method for an interest rate rule with Markov-switching coefficients subject to the effective lower bound. The devised estimation technique provides the probability that, at or just after exiting the effective lower bound, the central bank adopts a hawkish or a dovish regime, hence providing an estimate of the current stance of monetary policy.

Since the work of Tobin (1958), it is known that the inadequate estimation of a censored regression produces inconsistent estimators. In this paper, I show that estimating a Markov-switching regression using the Hamilton (1989) filter ignoring the censoring problem produces inconsistent estimators of the Markov-switching regression coefficients and the transition probabilities. Moreover, I show that, even when a censored regression specification is introduced in the estimation, the filtered probabilities fail to identify the prevalent regime over the censored part of the sample.

There is a way to solve, at least partially, the lack of identification problem of the prevalent regime over the censored part of the sample. The solution involves the joint estimation of the censored Markov-switching regression and an uncensored auxiliary Markov-switching regression whose switching is correlated with the switching of the coefficients of the censored equation. In particular, I show that as the correlation between the states driving the switching of the coefficients of the two regressions increases, identification of the prevalent regime of the censored Markov-switching regression is more precise.

The present work fits in the literature of estimating Taylor rules with Markov-switching coefficients. Bae et al. (2012), for example, estimate a forward-looking Taylor rule for the period spanning 1956 to 2005 and identify regimes that roughly corresponded to the terms of the Federal Reserve chairs. Murray et al. (2013) estimate a real-time forward-looking two-state Markov-switching Taylor rule to make inference about the periods when the Taylor principle was present. They find that the Fed consistently adhered to the Taylor principle before 1973 and after 1984, but did not follow the Taylor principle from 1980 to 1984.

Markov-switching monetary policy regimes have also been considered within the context of dynamic stochastic general equilibrium (DSGE) models. Eo (2009) estimates a Markov-switching DSGE model with recurring regime changes in the monetary policy rule coefficients, the technology coefficients, and the coefficients characterizing nominal price rigidities. In an application to postwar U.S. data, he finds stronger support for regime switching in monetary policy than in technology or nominal rigidities. Davig and Doh (2009) estimate a Markov-switching New Keynesian model that allows shifts in the monetary policy reaction coefficients and shock volatilities. Using U.S. data, they find that a more-aggressive monetary policy regime was in place after the Volcker disinflation and before 1970 than during the Great Inflation of the 1970s. Bianchi (2013) estimates a two-state model and finds that monetary policy has fluctuations between a Hawk and a Dove regime, with the latter prevalent in the 1970s and during the recent crisis.

Another strand of the literature estimates the monetary policy rule along with a fiscal policy rule. For example, Davig and Leeper (2006, 2011) estimate two-state Markov-switching monetary and fiscal policy rules to evaluate the presence of regimes of monetary or fiscal dominance. In their specification of the Markov-switching processes, two independent states drive the evolution of the monetary and fiscal policy rule coefficients. They find that monetary and fiscal policies fluctuate between active and passive behavior. In a Markov-switching DSGE framework, Bianchi (2012) specifies and estimates a model with monetary and fiscal policy rules whose coefficients' switching is driven by a single state. His estimates show that the monetary/fiscal policy mix has evolved over time and identifies three distinct regimes.

I apply the proposed estimation technique to a two-state Markov-switching forward-looking Taylor rule using quarterly data spanning 1960:1-2013:1. Interest rates at or below 0.25 percent are classified as censored, and the lower bound is set to that value. For the Markov-switching uncensored auxiliary regression, I take a fiscal policy rule where taxes respond to debt deviations from target and to the output gap.

Results imply that the estimated correlation between the switching states of the two policy rules is 0.81. Moreover, the null hypothesis of independent switching between the coefficients of the monetary and fiscal policy rules is rejected at conventional significance levels. The estimated coefficients allow us to classify the monetary/fiscal policy mix into four regimes according to the response of the interest rate to inflation and the response of taxes to debt: (i) a regime of weak interest rate response to inflation and weak tax response to debt, that I denominate regime F, for fiscal; (ii) a regime of weak interest rate response to inflation and strong tax response to debt, that I denominate regime I, for indeterminate; (iii) a regime of strong interest rate response to inflation and weak tax response to debt, that I denominate regime E, for explosive; (iv) a regime of strong interest rate response to inflation and strong tax response to debt, that I denominate regime M, for monetary. The estimated ergodic regime probabilities are: 39% for regime F, 2% for regime I, 8% for regime E, and 51% for regime M. The transition probabilities for the policy rule coefficients imply that regime M is expected to last about 12.5 quarters, regime F, about 7.1 quarters, and regimes I and E, 1 quarter each.

The model's smoothed probabilities imply that in the first quarter of 2013 the economy was more likely in regime I, where the stance of monetary policy was accommodative while fiscal policy was trying to stabilize debt deviations from target.

This document is structured as follows: in Section 2, I present the specification of a Markov-switching Taylor Rule at the zero lower bound. Section 3 develops the estimation procedure and the Monte Carlo exercise that justifies it. Results of the estimation appear in Section 4. Section 5 puts the results in context with the historical narrative on monetary and fiscal policy. Section 6 concludes.

2 A Markov-switching Taylor Rule at the Zero Lower Bound

I am interested in estimating the following Markov-switching regression model of a monetary policy rule with a smoothing component:

$$R_t^* = \rho_{S_{m,t}} R_{t-1} + \left(1 - \rho_{S_{m,t}}\right) \left(R_{S_{m,t}} + \alpha^{\pi}_{S_{m,t}} \pi_t + \alpha^{y}_{S_{m,t}} y_t\right) + \sigma_{S_{\sigma_R,t}} u_t \tag{1}$$

$$R_t = \max\left(\underline{R}, R_t^*\right) \tag{2}$$

$$S_{m,t} = 1, 2, \ldots, J_m; \qquad S_{\sigma_R,t} = 1, 2, \ldots, J_{\sigma_R},$$

where R_t^* is the underlying policy rate in period t, R_t is the observed policy rate in period t, π_t is a measure of the inflation rate in period t, y_t is a measure of the output gap in period t, and $u_t \sim \mathbb{N}(0,1)$ is a monetary policy shock to the policy rate. The observed interest rate is bounded from below by $\underline{R} \geq 0$.

$S_{m,t}$ and $S_{j_{\sigma_R},t}$ are J_m-state and J_{σ_R}-state, possibly correlated, first-order Markov switching processes, respectively. Their transition probabilities are

$$\mathbb{P}\left(S_{m,t} = j_m | S_{m,t-1} = j'_m\right) = p_{j_m j'_m} \tag{3}$$

$$\mathbb{P}\left(S_{\sigma_R,t} = j_{\sigma_R} | S_{\sigma_R,t-1} = j'_{\sigma_R}\right) = p_{j_{\sigma_R} j'_{\sigma_R}}. \tag{4}$$

Bae et al. (2012) show that Equation (1) is the empirical model of the federal funds rate with a smoothing component assuming that the forward-looking monetary policy rule is subject to regime changes. A result of this specification is that the inflation rate and the output gap are correlated with the error term.

I show that, to consistently estimate this Markov-switching regression model, it is not enough to incorporate in the estimation the censored part of the process. In particular, inference about the prevalent regime over the censored period is inaccurate. The next section specifies a system of equations with interdependent Markov-switching coefficients. Interdependent switching is the key to identification of the prevalent regime of the economy over the censored part of the sample.

3 Estimation Procedure

This section shows that estimation of the monetary policy rule as specified in Equations (1)-(4) yields biased estimates if censoring is not considered, and may be unable to discriminate the prevalent regime over the censored part of the sample even if censoring is incorporated. I then implement an estimation procedure that allows resolving the lack of identification problem.

3.1 Setup

Consider the following Markov-switching regression model with a censored dependent variable:[1]

$$\begin{aligned}
y_{1t}^* &= x_{1t}'\beta_{1,S_{1t}} + \sigma_{1,S_{1t}}u_{1t}, \quad S_{1t} = 1, 2, ..., J_1, && (5)\\
y_{1t} &= \max(y_{1L}, y_{1t}^*), && (6)\\
y_{2t} &= x_{2t}'\beta_{2,S_{2t}} + \sigma_{2,S_{2t}}u_{2t}, \quad S_{2t} = 1, 2, ..., J_2 && (7)\\
\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} &\sim \mathbb{N}(\mathbf{0}_{2\times 1}, I_2) && (8)\\
\beta_{1,S_{1t}} &= \sum_{j_1=1}^{J_1} \beta_{1,j_1}\tilde{S}_{1,j_1,t}; \qquad \sigma_{1,S_{1t}} = \sum_{j_1}^{J_1} \sigma_{1,j_1}\tilde{S}_{1,j_1,t} && (9)\\
\beta_{2,S_{2t}} &= \sum_{j_2=1}^{J_2} \beta_{2,j_2}\tilde{S}_{2,j_2,t}; \qquad \sigma_{2,S_{2t}} = \sum_{j_2}^{J_2} \sigma_{2,j_2}\tilde{S}_{2,j_2,t}, && (10)
\end{aligned}$$

where

$$\tilde{S}_{i,j_i,t} = \begin{cases} 1, & \text{if } S_{it} = j_i; j_i = 1, 2, \ldots, J_i; i = 1, 2 \\ 0, & \text{otherwise,} \end{cases} \qquad (11)$$

and where y_{1t} and y_{2t} are 1×1; x_{1t} and x_{2t} are $k_1 \times 1$ and $k_2 \times 1$, respectively, vectors of explanatory variables. I assume that y_{1t} conditional on S_{1t}, and x_{1t} are covariance stationary. The same holds for y_{2t} conditional on S_{2t}, and x_{2t}. Following Kim (2009), to allow for non-

[1]For simplicity of exposition, I assume that the state that drives the switching in the conditional mean parameters also drives the switching in the standard deviation of the shocks. I will relax this assumption to conduct the estimation of the Markov-switching model of the federal funds rate given in specification (1)-(4).

zero correlation between S_{1t} and S_{2t} I introduce the following $J = J_1 \times J_2$-state Markov-switching process S_t:

$$S_t = (S_{2t} - 1)\, J_1 + S_{1t}, \qquad S_{it} = 1, 2, \ldots, J_i, i = 1, 2, \tag{12}$$

where the transition probabilities are given by

$$\begin{aligned} \mathbb{P}\left(S_t = j | S_{t-1} = j'\right) &= \mathbb{P}\left(S_{1t} = j_1, S_{2t} = j_2 | S_{1,t-1} = j_1', S_{2,t-1} = j_2'\right) \\ &= p_{jj'}, \end{aligned} \tag{13}$$

and

$$\begin{aligned} j &= (j_2 - 1) J_1 + j_1, \\ j' &= (j_2' - 1) J_1 + j_1', \end{aligned}$$

with $\sum_{j=1}^{J} p_{jj'} = 1$. The marginalized transition probabilities for S_{1t} and S_{2t} are given by

$$p_{1,j_1 j_1'} = \mathbb{P}\left(S_{1t} = j_1 | S_{1,t-1} = j_1'\right), \tag{14}$$

$$p_{2,j_2 j_2'} = \mathbb{P}\left(S_{2t} = j_2 | S_{2,t-1} = j_2'\right), \tag{15}$$

which can be obtained using the derivation in Kim (2009).

I assume that the explanatory variables x_{1t} and x_{2t} are uncorrelated with the error terms of their respective equations, u_{1t} and u_{2t}. In case of correlation with the error terms, the approaches in Kim (2004) or Kim (2009) can be added to the system above.

Notice that the errors of Equations (5) and (7) are uncorrelated. Hence, conditional on S_t, y_{1t} and y_{2t} are independent. The dependence between y_{1t} and y_{2t} occurs only through the dependent switching of the coefficients of both equations.

3.2 Maximum Likelihood Estimation

Let $y_t = [y_{1t} \quad y_{2t}]'$, $x_t = [x_{1t}' \quad x_{2t}']'$. Let $\mathfrak{F}_{i,t-1} = \sigma\left(x_{i1}, x_{i2}, \ldots, x_{it}, y_{i0}, y_{i1}, \ldots, y_{i,t-1}\right)$ for $i = 1, 2$ be the sigma-algebras generated by the vectors of exogenous random variables of Equations (5)-(7), and let $\mathfrak{F}_{t-1} = \sigma\left(x_1, x_2, \ldots, x_t, y_0, y_1, \ldots, y_{t-1}\right)$ be the sigma-algebra generated by the vectors of all exogenous random variables. Let $\theta = [\theta_1' \quad \theta_2' \quad vec(\tilde{p})']'$ be the vector of parameters of the model, where

$$\begin{aligned} \theta_1 &= \left[\beta_{1,1}' \ldots \beta_{1,J_1}' \quad \sigma_{1,1} \ldots \sigma_{1,J_1}\right]', \\ \theta_2 &= \left[\beta_{2,1}' \ldots \beta_{2,J_2}' \quad \sigma_{2,1} \ldots \sigma_{2,J_2}\right]', \end{aligned}$$

and $\tilde{p}$ is a $J \times J$ matrix of transition probabilities given in (13). For consistent and efficient estimation of the model (5)-(13), I maximize the log-likelihood function, $\mathfrak{L}_\theta(\theta; Y_T) = \ln f_Y(Y_T; \theta)$, with respect to θ by applying the conventional Hamilton (1989) filter, where

$Y_t = \{y_s\}_{s=1}^t$. The filter allows obtaining $f_Y(Y_T;\theta)$ as follows:

$$\begin{aligned} f_Y(Y_T;\theta) &= \prod_{t=1}^{T} f_y(y_t|\mathfrak{F}_{t-1};\theta) \\ &= \prod_{t=1}^{T}\sum_{j=1}^{J} f_{y|S}(y_t|S_t = j, \mathfrak{F}_{t-1};\theta)\mathbb{P}(S_t = j|\mathfrak{F}_{t-1};\theta), \end{aligned}$$

where

$$f_{y|S}(y_t|S_t = j, \mathfrak{F}_{t-1};\theta) = f_{y_1|S_1}(y_{1t}|S_{1t} = j_1, \mathfrak{F}_{1,t-1};\theta_1) f_{y_2|S_2}(y_{2t}|S_{2t} = j_2, \mathfrak{F}_{2,t-1};\theta_2),$$

with

$$\begin{aligned} f_{y_1|S_1}(y_{1t}|S_{1t} = j_1, \mathfrak{F}_{1,t-1};\theta_1) &= \left[\Phi\left(\frac{y_{1L} - x_{1t}'\beta_{1,j_1}}{\sigma_{1,j_1}}\right)\right]^{1[y_{1t}=y_{1L}]} \times \\ &\times \left[\frac{1}{\sigma_{1,j_1}}\phi\left(\frac{y_{1t} - x_{1t}'\beta_{1,j_1}}{\sigma_{1,j_1}}\right)\right]^{1[y_{1t}>y_{1L}]}, \\ f_{y_2|S_2}(y_{2t}|S_{2t} = j_2, \mathfrak{F}_{2,t-1};\theta_2) &= \frac{1}{\sigma_{2,j_2}}\phi\left(\frac{y_{2t} - x_{2t}'\beta_{2,j_2}}{\sigma_{2,j_2}}\right), \end{aligned}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ denote the distribution and density functions, respectively, of the standard normal distribution, and

$$\mathbb{P}(S_t = j|\mathfrak{F}_{t-1};\theta) = \sum_{j'=1}^{J} p_{jj'}\mathbb{P}(S_{t-1} = j'|\mathfrak{F}_{t-1};\theta). \tag{16}$$

Once y_t is realized at the end of time t, the filtered probability of S_t in (16) is updated as

$$\mathbb{P}(S_t = j|\mathfrak{F}_t;\theta) = \frac{f_{y|S}(y_t|S_t = j, \mathfrak{F}_{t-1};\theta)\mathbb{P}(S_t = j|\mathfrak{F}_{t-1};\theta)}{f_y(y_t|\mathfrak{F}_{t-1};\theta)}. \tag{17}$$

3.3 Why Is Interdependent Switching Necessary?

In this section, I discuss the need of introducing the auxiliary Equation (7) and correlated states S_{1t} and S_{2t}. I show that if censoring is ignored, the estimates of $\beta_{1,S_{1t}}$, $\sigma_{1,S_{1t}}$, $p_{j_1,j_1'}$ are biased, and that inference about the prevalent regime S_{1t} is not feasible. When censoring is introduced in the specification, the biases in $\beta_{1,S_{1t}}$, $\sigma_{1,S_{1t}}$, and $p_{j_1,j_1'}$ are corrected, but inference about the prevalent regime remains infeasible. I finally show that when the system (5)-(13) is estimated jointly, discrimination of the prevalent regime is possible.

To show the potential estimation problem and the features of the proposed solution, I perform a Monte Carlo experiment where the model is specified as in (5)-(13) with $J_1 =$

$J_2 = 2$, $k_1 = k_2 = 1$, $T = 200$, and

$$\begin{aligned}
\beta_{1,j_1} &= \begin{cases} 0.5 & \text{if } j_1 = 1 \\ 1.5 & \text{if } j_1 = 2 \end{cases}, \qquad \sigma_{1,j_1} = \begin{cases} 0.05 & \text{if } j_1 = 1 \\ 0.05 & \text{if } j_1 = 2 \end{cases}, \\
\beta_{2,j_2} &= \begin{cases} 0 & \text{if } j_2 = 1 \\ 0.1 & \text{if } j_2 = 2 \end{cases}, \qquad \sigma_{2,j_2} = \begin{cases} 0.005 & \text{if } j_2 = 1 \\ 0.005 & \text{if } j_2 = 2 \end{cases}, \\
x_{1t} &\sim \begin{cases} \mathcal{U}(1,2) & \text{if } t \leq 150 \\ \mathcal{U}(-1,0) & \text{if } 151 \leq t \leq 200 \end{cases}, \\
x_{2t} &\sim \mathcal{U}(0,1), \\
\tilde{p} &= \begin{bmatrix} 0.3 & 0.2 & 0.2 & 0.1 \\ 0.05 & 0.05 & 0.05 & 0.05 \\ 0.05 & 0.05 & 0.05 & 0.05 \\ 0.6 & 0.7 & 0.7 & 0.8 \end{bmatrix}, \\
\operatorname{corr}(S_{1t}, S_{2t}) &= 0.67, \\
y_{1L} &= 0.
\end{aligned} \tag{18}$$

In the benchmark specification, censoring of y_{1t}^* occurs over the final 25% of the sample. I implement this censoring by switching x_{1t} to a different distribution, as shown in (18). I choose a cluster of periods where censoring occurs to illustrate the severity of the problem at obtaining the estimates of the prevalent regime.

To perform the Monte Carlo analysis, I need to obtain estimates of β_{1,j_1}, σ_{1,j_1}, $p_{1,j_1 j_1'}$ for $j_1 = 1, 2$, and the smoothed estimate of $\mathbb{P}(S_{1t} = j_1)$ for $t = 1, 2, \ldots, 200$ under three scenarios:

(i) Ignoring both censoring of y_{1t} and joint switching between S_{1t} and S_{2t}.

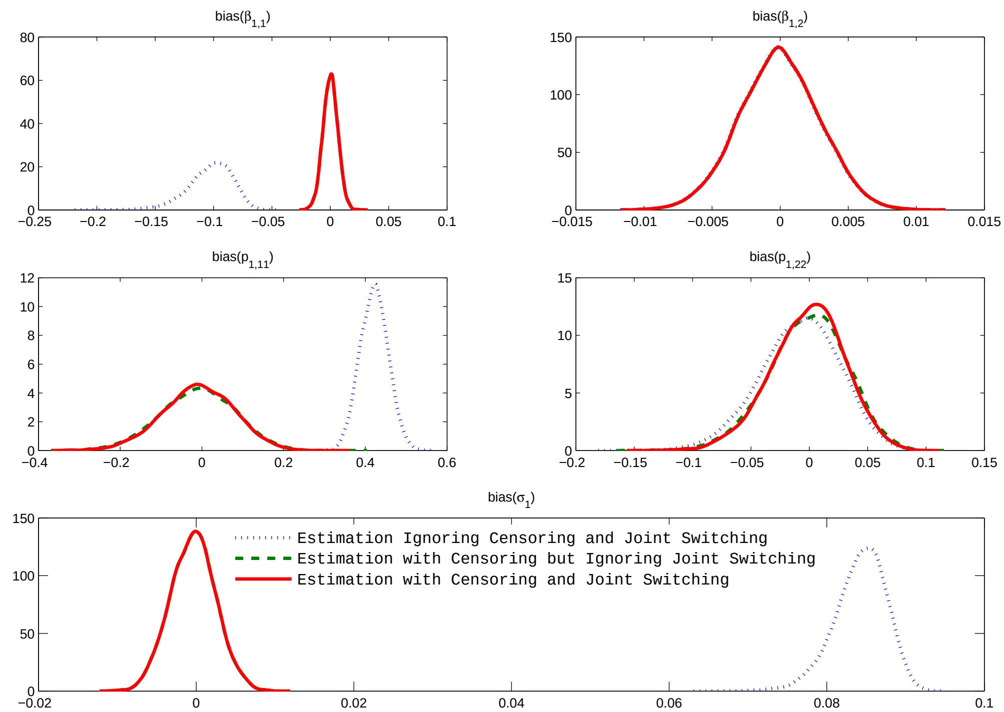
(ii) Allowing for censoring of y_{1t} but ignoring joint switching between S_{1t} and S_{2t}.

(iii) Allowing for both censoring of y_{1t} and joint switching between S_{1t} and S_{2t}.

Appendices A and B obtain the likelihood functions for cases (i) and (ii), respectively. The likelihood function for case (iii) was obtained in Section 3.2. I simulate and estimate the model 10,000 times.

Figure 2 shows the bias in the estimates of the parameters under the three scenarios listed above. The results show that the estimation under scenario (i) yields biased estimates as would have been expected. In particular, the estimates of $\beta_{1,1}$ is downward biased, while the estimates of σ_1 and $p_{1,11}$ are upward biased. The downward bias in $\beta_{1,1}$ is due to the estimation attributing to a low slope coefficient the fact that $y_{1t} = 0$ in the final 25% of the sample. The persistent censoring implies an upward bias in $p_{1,11}$, the probability of remaining in the low-β state. A higher standard deviation of shocks is also needed to reconcile the fact that $y_{1t} = 0$ while x_{1t} takes negative values. On the other hand, $\beta_{1,2}$ and $p_{1,22}$ do not seem to suffer from a bias problem. The information that the estimation obtains from the uncensored part of the sample seems enough to obtain accurate estimates of these parameters.

Figure 2: Parameter Bias in the Benchmark Case



Estimation under scenarios (ii), which incorporates censoring, and (iii), which incorporates censoring and joint switching, yield unbiased parameter estimates, as expected. I point out that estimation under scenario (ii) is sufficient to achieve unbiasedness. Estimation under scenario (iii) is not necessary for unbiasedness.

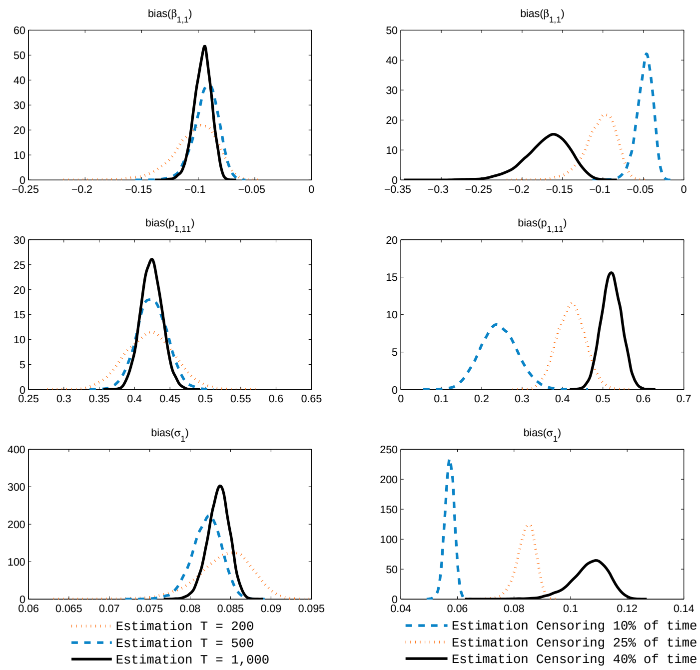
I also analyze the effect of increasing the sample size and of changing the length of censoring at the end of the sample. Figure 3 reports the changes on the biases of $\beta_{1,1}$, σ_1, and $p_{1,11}$ estimated under scenario (i). Increasing the sample size does not reduce the bias of the estimates, suggesting that there is a problem of consistency. On the other hand, and as expected, reducing the length of censoring over the final part of the sample reduces the biases.

3.3.1 Discriminating the Prevalent Regime

I now investigate the ability of the estimation strategies to identify correctly the prevalent regime. In the Monte Carlo exercise, I have set the standard deviation, σ_1, so that there is an almost perfect discrimination of the states over the uncensored part of the sample. Hence, to evaluate the capabilities of the three estimation scenarios at discriminating correctly the prevalent regimes, I focus on the censored part of the sample only.

To measure the ability of the estimation techniques to identify the prevalent regimes, I use the area under the Receiver Operating Characteristic (ROC) curve. The ROC curve is a plot that assesses the performance of a binary classifier system as its discrimination threshold is changed. The ROC curve was first developed by electrical engineers and radar engineers during World War II for detecting enemy objects in battlefields and was soon introduced in

Figure 3: Parameters Biases in the Benchmark Case Ignoring Censoring and Joint Switching - Effects of Sample Size and Censoring Length



psychology to account for perceptual detection of signals (see Peterson et al., 1954; Swets, 1979). The use of ROC curves in medicine to assess diagnostic test performance has been described by Lusted (1971). In our case, I obtain smoothed estimates of $\mathbb{P}\left(S_{1t}=1\right)$, vary the discrimination threshold between 0 and 1, and evaluate the ability of the smoothed estimates of $\mathbb{P}\left(S_{1t}=1\right)$ to classify correctly the prevalent regime, which is given by the simulated states.

The ROC curve plots the fraction of true positives out of the total of actual positives, called true positive rate (TPR), against the fraction of false positives, called false positive rate (FPR), at various threshold settings. TPR is also known as Sensitivity, and FPR is known as one minus the Specificity or true negative rate. Given a cut-off value $q \in [0,1]$, a realization of $\{S_{1t}\}_{t=1}^{200}$, and smoothed estimates of $\mathbb{P}\left(S_{1t}=1\right)$, I can tabulate a contingency table like Table 1. Varying the cut-off value $q \in [0,1]$ allows obtaining Sensitivity values to plot against 1-Specificity values, which is the ROC curve.

A perfectly discriminating variable would have Sensitivity and Specificity both equal to 1. If a cut-off value existed to produce such a test, then Sensitivity would be 1 for any non-zero values of 1Specificity. The ROC curve would start at the origin (0,0), go vertically up the y-axis to (0,1), and then horizontally across to (1,1) (see Bewick et al., 2004). On the other hand, a completely random guess would give a point along a diagonal line (the so-called line of no-discrimination). In that case, the discriminating variable would produce a TPR equal to its FPR, or Sensitivity = 1-Specificity. The ROC curve would start at the origin (0,0) and go diagonally to (1,1).

The performance of a discriminating variable can be quantified by calculating the area

Table 1: Contingency Table

		State of Nature	
		$S_{1t} = 1$	$S_{1t} = 2$
Predicted State	$\mathbb{P}\left(S_{1t} = 1\right) \geq q$	True Positive	False Positive
	$\mathbb{P}\left(S_{1t} = 1\right) < q$	False Negative	True Negative
		Sensitivity $= \frac{\sum_{t=1}^{200} \mathbb{1}_{\{\mathbb{P}(S_{1t}=1)\geq q\}}}{\sum_{t=1}^{200} \mathbb{1}_{\{S_{1t}=1\}}}$	*Specificity* $= \frac{\sum_{t=1}^{200} \mathbb{1}_{\{\mathbb{P}(S_{1t}=1)<q\}}}{\sum_{t=1}^{200} \mathbb{1}_{\{S_{1t}=2\}}}$

under the ROC curve. An ideal discriminating variable would have an area under the ROC curve of 1, whereas a random guess would have an area under the ROC curve of 0.5.

Figure 4 plots the ROC curves and reports the areas under the ROC curves for the estimation of model (18) under the three scenarios mentioned above: (i) Ignoring both censoring of y_{1t} and joint switching between S_{1t} and S_{2t}; (ii) Allowing for censoring of y_{1t} but ignoring joint switching between S_{1t} and S_{2t}; (iii) Allowing for both censoring of y_{1t} and joint switching between S_{1t} and S_{2t}. The figure shows that, over the censored part of the sample, the only estimation scenario that allows for some degree of discrimination of the prevalent regime is the one with censoring of y_{1t} and joint switching between S_{1t} and S_{2t}. The areas under the ROC curves for the scenarios that do not allow for joint switching are very close to 0.5, whereas the area under the ROC curve for the joint switching scenario is about 0.84. This indicates that, to identify the prevalent regime over the censored part of the sample, an auxiliary uncensored Markov-switching regression whose coefficients switch in a correlated manner with the coefficients of the variable of interest is needed.

Figure 5 plots the ROC curves and reports the areas under the ROC curves for the effects of changing the sample size, T, the frequency of censoring, and the correlation between S_{1t} and S_{2t}. The results show that the estimation procedure that incorporates censoring and dependent switching is not affected in its discrimination ability when the sample size increases to $T = 500$ or $T = 1,000$. Changing the proportion of the sample that is subject to censoring to 10% or 40% does not change the performance of the estimation technique, either. Finally, the results show that eliminating the correlation between S_{1t} and S_{2t} annihilates the ability of the estimation technique to identify the prevalent regime, while a perfect correlation between the latent states implies a nearly perfect discrimination, as it would have been expected. This exercise highlights the importance of an auxiliary regression with Markov-switching coefficients whose state is correlated with the state of the coefficients of the censored variable. The higher the correlation between switching states, the better the discrimination of the prevalent regime of the coefficients of interest over the censoring period.

As a final exercise, I show that the lack of discrimination is also present when the variance of the variable subject to censoring switches. In specification (18) of the Monte Carlo exercise,

Figure 4: Area under the ROC Curve in the Benchmark Case

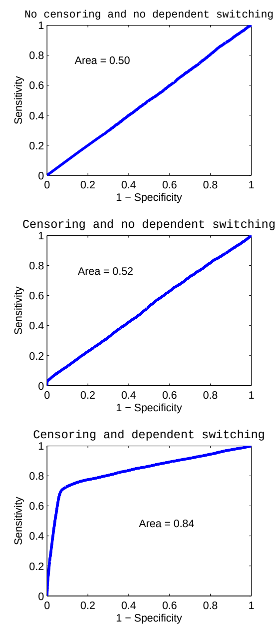


Figure 5: Area under the ROC Curve in the Benchmark Case - Effects of Sample Size, Censoring Frequency and Correlation between States

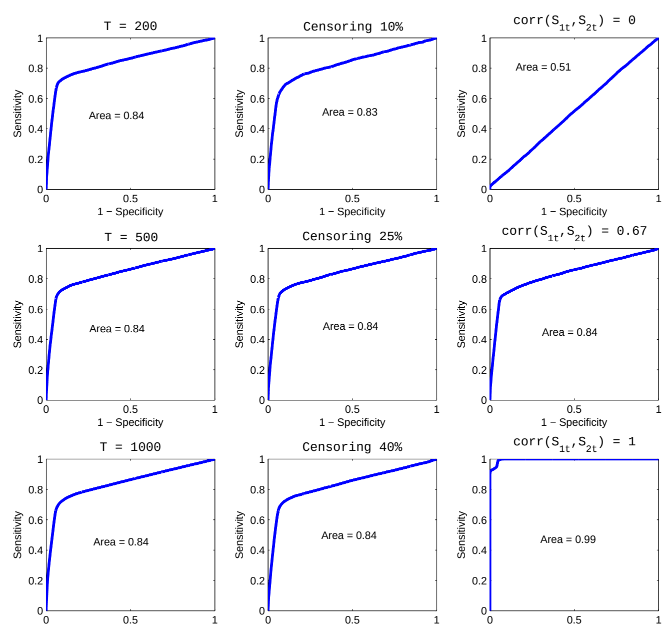
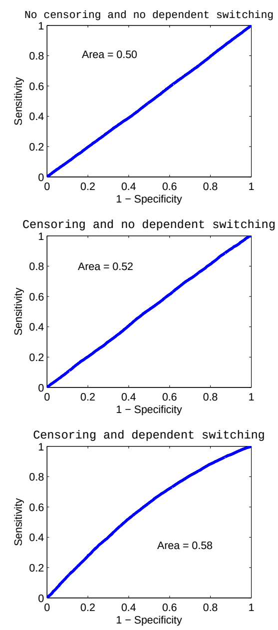


Figure 6: Area under the ROC Curve in the Switching Variance Case



I now allow the standard deviation σ_1 to switch between regimes as a function of a latent state that drives the switching of β_1. Figure 6 shows the ROC curves for the discrimination of the prevalent regime of σ_1 over the censoring period for the three estimation scenarios described before. As can be seen from the figure, discrimination of the prevalent regime for the standard deviation of the censored variable improves when a joint estimation where an auxiliary Markov-switching regression is included. Additional simulations (not shown here) illustrate that the higher the correlation between the latent state driving the standard deviation, σ_1, and the latent states S_1 or S_2, the better the discrimination of the prevalent regime of the standard deviation.

3.3.2 Why Does Joint Switching Help Identify the Prevalent Regime over the Censoring Period?

Under case (ii), that is, allowing for censoring but not for joint switching, I use the Tobit-like specification for the density function $f_{y_1|S_1}(y_{1t}|S_{1t}=j_1,\mathfrak{F}_{1,t-1};\theta_1)$ that appears in Appendix B. In this case, as shown in the previous section, the Hamilton filter is unable to identify the prevalent regime S_{1t} over the censoring period. The cause for the lack of discrimination lies in the fact that

$$\mathbb{P}(S_{1t}=j_1|\mathfrak{F}_{1t};\theta_1)=\frac{f_{y_1|S_1}(y_{1t}|S_{1t}=j_1,\mathfrak{F}_{1,t-1};\theta_1)\mathbb{P}(S_{1t}=j_1|\mathfrak{F}_{1,t-1};\theta_1)}{f_{y_1}(y_{1t}|\mathfrak{F}_{1,t-1};\theta_1)} \tag{19}$$

may not vary enough over the censored sample. In particular,

$$f_{y_1|S_1}(y_{1t}|S_{1t}=j_1,\mathfrak{F}_{1,t-1};\theta_1)=\Phi\left(\frac{y_{1L}-x_{1t}'\beta_{1,j_1}}{\sigma_{1,j_1}}\right)$$

will show little variation over the censored sample if $\frac{y_{1L}-x'_{1t}\beta_{1,j_1}}{\sigma_{1,j_1}}$ is too small or too large. If that is the case, $\mathbb{P}(S_{1t}=j_1|\mathfrak{F}_{1t};\theta_1)$ will have very little variation.

In contrast, incorporating joint switching in the estimation allows us to write the updated probability $\mathbb{P}(S_{1t}=j_1|\mathfrak{F}_t;\theta)$, using (17), as

$$
\begin{aligned}
\mathbb{P}(S_{1t}=j_1|\mathfrak{F}_t;\theta) &= \sum_{j_2=1}^{J_2}\mathbb{P}(S_{1t}=j_1,S_{2t}=j_2|\mathfrak{F}_t;\theta)\\
&= \sum_{j_2=1}^{J_2}\frac{f_{y|S_1,S_2}(y_t|\mathfrak{F}_{t-1},S_{1t}=j_1,S_{2t}=j_2;\theta)\mathbb{P}(S_{1t}=j_1,S_{2t}=j_2|\mathfrak{F}_{t-1};\theta)}{f_y(y_t|\mathfrak{F}_{t-1};\theta)}\\
&= \sum_{j_2=1}^{J_2}\frac{f_{y_1|S_1}(\cdot)f_{y_2|S_2}(\cdot)\mathbb{P}(S_{1t}=j_1,S_{2t}=j_2|\mathfrak{F}_{t-1};\theta)}{f_y(y_t|\mathfrak{F}_{t-1};\theta)}\\
&= \frac{f_{y_1|S_1}(y_{1t}|\mathfrak{F}_{1,t-1},S_{1t}=j_1;\theta_1)\mathbb{P}(S_{1t}=j_1|\mathfrak{F}_{t-1};\theta)}{f_{y_1}(y_{1t}|\mathfrak{F}_{t-1};\theta)}\\
&\times\sum_{j_2=1}^{J_2}\frac{f_{y_2|S_2}(y_{2t}|\mathfrak{F}_{2,t-1},S_{2t}=j_2;\theta_2)\mathbb{P}(S_{2t}=j_2|S_{1t}=j_1,\mathfrak{F}_{2,t-1};\theta)}{f_{y_2|y_1}(y_{2t}|y_{1t},\mathfrak{F}_{t-1};\theta)}\\
&= \frac{f_{y_1|S_1}(y_{1t}|\mathfrak{F}_{1,t-1},S_{1t}=j_1;\theta_1)\mathbb{P}(S_{1t}=j_1|\mathfrak{F}_{t-1};\theta)}{f_{y_1}(y_{1t}|\mathfrak{F}_{t-1};\theta)}\\
&\times\sum_{j_2=1}^{J_2}\mathbb{P}\left(S_{2t}=j_2|\mathfrak{F}_{2t};\theta_2\right)\frac{\mathbb{P}(S_{2t}=j_2|S_{1t}=j_1,\mathfrak{F}_{2,t-1};\theta)}{\mathbb{P}(S_{2t}=j_2|\mathfrak{F}_{2,t-1};\theta_2)}\frac{f_{y_2}(y_{2t}|\mathfrak{F}_{2,t-1};\theta_2)}{f_{y_2|y_1}(y_{2t}|y_{1t},\mathfrak{F}_{t-1};\theta)},
\end{aligned}
\tag{20}
$$

where the step from the next-to-last to the last equation uses the definition

$$
\mathbb{P}(S_{2t}=j_2|\mathfrak{F}_{2,t};\theta_2)=\frac{f_{y_2|S_2}(y_{2t}|\mathfrak{F}_{2,t-1},S_{2t}=j_2;\theta_2)\mathbb{P}(S_{2t}=j_2|\mathfrak{F}_{2,t-1};\theta_2)}{f_{y_2}(y_{2t}|\mathfrak{F}_{2,t-1};\theta_2)}.
$$

If S_{1t} and S_{2t} are uncorrelated, the last line of (20) is equal to one. In that case, inference about S_{1t} obtained from (20) would be the same as inference about S_{1t} obtained from (19). Hence, discrimination about the prevalent regime would be unfeasible. It is the additional information given by the degree of interdependence between the latent states what allows a better inference about the prevalent regime of S_{1t}.

4 Estimating a Markov-Switching Taylor Rule at the Zero Lower Bound

In this section I apply the proposed estimation technique to estimate a Taylor rule with Markov-switching coefficients including the sample period after the financial crisis, where the federal funds rate has been at the ZLB.

4.1 Selecting the Auxiliary Regression

To implement the estimation procedure presented in Section 3, I need an auxiliary Markov-switching regression that is not subject to censoring and whose switching could be correlated with the switching of the coefficients of the Taylor rule.

Gonzalez-Astudillo (2013) estimates time-varying monetary and fiscal policy rules whose coefficients are driven by correlated latent factors and finds a non-negligible degree of interdependence between the coefficients of the policy rules. This finding is related to the literature on monetary-fiscal policy interactions initiated by Leeper (1991) and followed by Davig and Leeper (2006) and Chung et al. (2007), among others. Along these lines, I propose a fiscal policy rule with Markov-switching coefficients to be the auxiliary regression. I will test for interdependence between the switching of the Taylor rule coefficients and the coefficients of the proposed fiscal policy rule to confirm that this is an adequate choice.

4.2 Setting up the System to be Estimated

The system to be estimated in order to consistently estimate the Markov-switching coefficients of the Taylor rule, as well as to make inference about the prevalent regime, is given by

$$R_t^* = \rho_{S_{m,t}}^R R_{t-1} + \left(1 - \rho_{S_{m,t}}^R\right)\left(R_{S_{m,t}} + \alpha_{S_{m,t}}^\pi \pi_t + \alpha_{S_{m,t}}^y y_t\right) + \sigma_{S_{\sigma^R,t}}^R u_t^R, \tag{I}$$

$$R_t = \max\left(\underline{R}, R_t^*\right) \tag{II}$$

$$\tau_t = \rho_{S_{f,t}}^\tau \tau_{t-1} + \left(1 - \rho_{S_{f,t}}^\tau\right)\left(\tau_{S_{f,t}} + \gamma_{S_{f,t}}^b b_{t-1} + \gamma_{S_{f,t}}^y y_t\right) + \sigma_{S_{\sigma^\tau},t}^\tau u_t^\tau \tag{III}$$

where R_t is the policy rate t, π_t is the inflation rate in period t, y_t is the output gap in period t, and $u_t^R \sim \mathbb{N}(0,1)$ is a monetary policy shock to the federal funds rate. The observed interest rate is bounded from below by $\underline{R} = 0.25$. In the auxiliary equation, τ_t is a measure of tax receipts net of transfers in period t, b_{t-1} is a measure of federal government debt in period $t-1$, and $u_t^\tau \sim \mathbb{N}(0,1)$ is a fiscal policy shock to taxes net of transfers.

I introduce dependent switching between $S_{m,t}$, $S_{\sigma^R,t}$, and $S_{f,t}$ by specifying the following $J = J_m \times J_{\sigma^R} \times J_f$-state Markov-switching process S_t:

$$S_t = \left(S_{f,t} - 1\right) J_m J_{\sigma^R} + \left(S_{\sigma^R,t} - 1\right) J_m + S_{m,t}, \qquad S_{i,t} = 1, 2, \ldots, J_i, \qquad \text{for } i = f, m, \sigma^R,$$

where the transition probabilities are given by:

$$\begin{aligned}\mathbb{P}\left(S_t = j | S_{t-1} = j'\right) &= \mathbb{P}\left(S_{m,t} = j_m, S_{\sigma^R,t} = j_{\sigma^R}, S_{f,t} = j_f | S_{m,t-1} = j'_m, S_{\sigma^R,t-1} = j'_{\sigma^R}, S_{f,t-1} = j'_f\right)\\ &= p_{jj'},\end{aligned}$$

and

$$\begin{aligned} j &= (j_f - 1) J_m J_{\sigma^R} + \left(j_{\sigma^R} - 1\right) J_m + j_m,\\ j' &= (j'_f - 1) J_m J_{\sigma^R} + \left(j'_{\sigma^R} - 1\right) J_m + j'_m,\end{aligned}$$

with $\sum_{j=1}^{J} p_{jj'} = 1$. I denote as P the $J \times J$ transition probability matrix of S_t. To simplify the estimation, I assume that $S_{\sigma^\tau,t}$ switches independently from $S_{m,t}$, $S_{\sigma^R,t}$, and $S_{f,t}$ with a $J_{\sigma^\tau} \times J_{\sigma^\tau}$ transition probability matrix P_{σ^τ}.

I set $J_m = J_f = J_{\sigma^R} = J_{\sigma^\tau} = 2$. This yields 58 transition probabilities to be estimated: $56 = (64 - 8)$ in P, and $2 = (4 - 2)$ in P_{σ^τ}. Allowing for more states for S_m, S_f, S_{σ^R}, or S_{σ^τ} would imply an increasing number of transition probabilities to be estimated that could result in an unfeasible estimation, in particular if the latent states are correlated.

In this setup, the policy rules have endogenous explanatory variables, namely the inflation rate and the output gap, so that I implement the two-step MLE procedure proposed by Kim (2009).

4.3 Data

I use quarterly data from 1960:1 to 2013:1. The policy rate is the federal funds rate. Inflation is the percentage change over the last four quarters of the price level given by the GDP price deflator. The output gap is the log difference between real GDP and the Congressional Budget Office's measure of potential real GDP. These variables are obtained from FRED. Tax receipts net of transfers corresponds to the seasonally adjusted quarterly current receipts of the federal government from which the current transfer payments have been deducted. This variable is obtained from the NIPA Table 3.2. Debt is the market value of privately held gross federal debt at the end of the quarter. This variable comes from the Federal Reserve Bank of Dallas. To perform the correction for endogeneity, I use a set of instrumental variables that includes M2 growth given by the percentage change over the last four quarters of seasonally adjusted M2, commodity price inflation given by the percentage change over the last four quarters of the commodity price index, and government spending given by the federal consumption expenditures and gross investment. These variables are obtained from FRED. Some of the variables need to be transformed before proceeding to the estimation, namely the interest rate, tax receipts, debt, and government spending. Appendix C describes the transformations.

4.4 Estimation Results

This section analyzes the results of the estimation by performing a set of hypothesis tests that allows to find the most parsimonious model in terms of switching parameters. Then, it presents a test of joint switching between the coefficients of the two policy rules, and between the coefficients of the conditional mean and the conditional variance of the monetary policy rule.

Before proceeding to the presentation of results, I emphasize that the estimation proposed here incorporates a correction for endogeneity of regressors. To correct for endogeneity, I perform two-stage constant-parameter estimations where the inflation rate and the output gap are regressed against a set of instruments that include: four lags of the inflation rate, four lags of the output gap, four lags of M2 growth, four lags of inflation of the commodity price index, and four lags of the ratio of government spending to GDP. The (standardized) residuals from these regressions appear as additional regressors with Markov-switching coefficients in the specifications of Equations (I) and (III).

Table 2: Likelihood Ratio Test

		Log-likelihood	**LR test statistic**
	Unrestricted Model	-26.73	–
Null Hypotheses	$R_1 = R_2,\ \tau_1 = \tau_2$	-29.51	5.56
	$R_1 = R_2,\ \tau_1 = \tau_2$ $\rho_1^R = \rho_2^R,\ \rho_1^\tau = \rho_2^\tau$	-30.09	6.72

Critical values: $\chi^2_{0.95,(2)} = 5.99$, $\chi^2_{0.99,(2)} = 9.21$, $\chi^2_{0.95,(4)} = 9.49$, $\chi^2_{0.99,(4)} = 13.28$

4.4.1 Finding the Most Parsimonious Specification

Results of the hypotheses tests to find the most parsimonious model in terms of switching parameters appears in Table 2. The test statistics of the Likelihood Ratio test do not reject policy rules whose intercepts and persistence coefficients are invariant between regimes. To test for switching regimes on the remaining coefficients, I use the z-statistics of the difference between the coefficients of the two regimes. The results appear in Table 4 of Appendix D. The z-statistics reject the null hypothesis that each of the coefficients is the same between regimes. Hence, I choose the specification with constant intercepts and persistence coefficients as the most parsimonious.

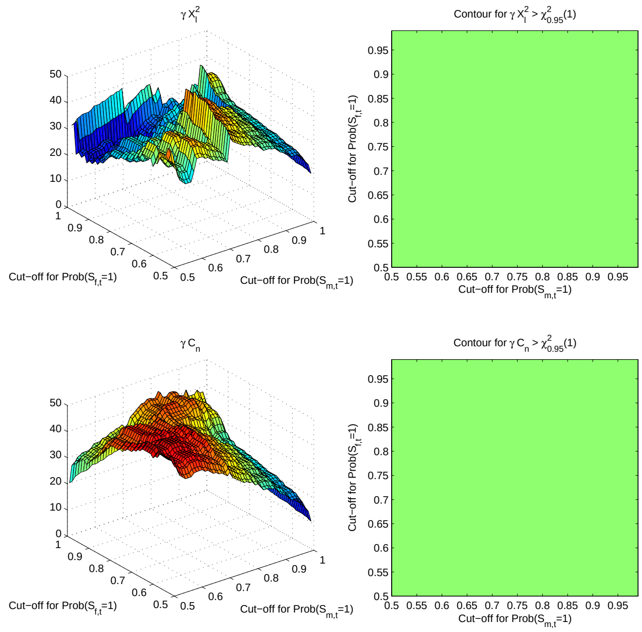
4.4.2 Testing for Independence Between Switching States

With the most parsimonious specification found above, I test for independence between S_m and S_f, and between S_m and S_{σ^R}. That is, I test for independence between the switching of the monetary policy and the fiscal policy rule coefficients, and between the switching of the coefficients of the conditional mean of the monetary policy rule and its variance. To test the independence hypotheses, I use the conventional independence tests based on 2×2 contingency tables. Tavaré and Altham (1983) modify the conventional chi-square test of independence based on contingency tables for the case when the data are generated by first-order Markov sequences.

To implement the independence tests, I obtain the smoothed probabilities $\mathbb{P}(S_{m,t} = 1)$ and $\mathbb{P}(S_{f,t} = 1)$ and write 2×2 contingency tables varying, in the range $[0.5, 1]$, the threshold at which it is decided that $S_{m,t} = 1$ or $S_{f,t} = 1$. With each of these contingency tables, I calculate the two tests statistics for independence between S_m and S_f obtained by Tavaré and Altham (1983) which, under the null, are distributed as a chi-square with one degree of freedom. I follow the same procedure for testing independence between S_m and S_{σ^R}.

Figure 7 shows the value of the statistics for testing the null hypothesis of independence between S_m and S_f for different thresholds of $\mathbb{P}(S_{m,t} = 1)$ and $\mathbb{P}(S_{f,t} = 1)$ on the left hand side, and the contour of the figure on the left for values of the statistics greater than the critical value that corresponds to a chi-square with one degree of freedom. Both test statistics reject the null hypothesis of independence between S_m and S_f at the 5% level of significance.

Figure 7: Independence Test S_m and S_f



Notes: The tests statistics are defined as:

$$X_l^2 = \left\{ \ln \left(\frac{n_1 n_4}{n_3 n_2} \right) \right\}^2 \Bigg/ \left(\frac{1}{np_1} + \frac{1}{np_2} + \frac{1}{np_3} + \frac{1}{np_4} \right)$$

$$C_n = \frac{n \left(n_1 n_4 / n^2 - n_2 n_3 / n^2 \right)^2}{n^{-4} \left(n_1 + n_2 \right) \left(n_1 + n_3 \right) \left(n_2 + n_4 \right) \left(n_3 + n_4 \right)},$$

where $p_{j_{mf}} = \mathbb{P}\left(S_{mf} = j_{mf}\right)$, and $S_{mf} = (S_f - 1)J_m + S_m$ and $j_{mf} = (j_f - 1)J_m + j_m$. Also, $n_{j_{mf}} = \sum_{t=1}^{n} \mathbb{1}_{\{S_{mf,t} = j_{mf}\}}$. Additionally, $\gamma = (1 - \mu\lambda)(1 + \mu\lambda)$, where μ and λ are the nonunit eigenvalues of P_m and P_f, the transition probability matrices of S_m and S_f, respectively.

I conclude that S_m and S_f should be specified with a joint 4×4 transition probability matrix that needs to be estimated.

Figure 12 in Appendix E shows the value of the statistics for testing the null hypothesis of independence between $S_{m,t}$ and $S_{\sigma^R,t}$ for different thresholds of $\mathbb{P}\left(S_{m,t} = 1\right)$ and $\mathbb{P}\left(S_{\sigma^R,t} = 1\right)$ on the left hand side, and the contour of the figure on the left for values of the statistics greater than the critical value that corresponds to a chi-square with one degree of freedom. Both test statistics fail to reject the null hypothesis of independence between S_m and S_{σ^R} at the 5% level of significance for a vast majority of possible thresholds. Hence, S_m and S_{σ^R} can be specified with separate transition probability matrices that need to be estimated.

Since the switching of the variance of the monetary policy rule shock cannot be associated with the switching of its conditional mean coefficients, identification of the prevalent regime for the variance will not be feasible over the censoring period. Further exploration is needed to find a switching variable that can be associated to the switching of the variance of the shock.

4.4.3 Obtaining Estimates for the Most Parsimonious Specification

The final specification re-estimates the system of Markov-switching Equations (I)-(III) with intercepts and smoothing coefficients fixed between regimes and four switching states: $S_m = 1, 2$, $S_f = 1, 2$, $S_{\sigma^R} = 1, 2$, and $S_{\sigma^\tau} = 1, 2$, where S_m and S_f have a joint transition probability matrix, denoted as P_{mf}, that corresponds to the four-regime state $S_{mf} = (S_f - 1)J_m + S_m$, with $S_i = 1, 2$ for $i = m, f$ and $J_m = 2$. Results of the estimation under this specification appear in Table 3.

Table 3: Parameter Estimates

Monetary Policy Rule

Parameters	$j_m = 1$	$j_m = 2$
R_{j_m}	2.52 (4.85)	
$\rho^R_{j_m}$	0.89 (58.63)	
$\alpha^\pi_{j_m}$	0.53 (3.66)	1.71 (3.94)
$\alpha^y_{j_m}$	1.61 (6.70)	0.40 (−7.01)
	$j_{\sigma^R} = 1$	$j_{\sigma^R} = 2$
$\sigma^R_{j_{\sigma^R}}$	0.24 (11.10)	1.22 (8.40)

For $j_m = 1$ and $j_{\sigma^R} = 1$, values in parenthesis are z-statistics of the null hypothesis that the coefficient is zero. For $j_m = 2$ and $j_{\sigma^R} = 2$, values in parenthesis are z-statistics of the null hypothesis that the difference between the coefficients of the two regimes is zero.

Fiscal Policy Rule

Parameters	$j_f = 1$	$j_f = 2$
τ_{j_f}	1.61 (10.46)	
$\rho^\tau_{j_f}$	0.89 (63.70)	
$\gamma^b_{j_f}$	0.05 (2.78)	0.13 (4.26)
$\gamma^y_{j_f}$	0.29 (5.10)	0.07 (−6.98)
	$j_{\sigma^\tau} = 1$	$j_{\sigma^\tau} = 2$
$\sigma^\tau_{j_{\sigma^\tau}}$	0.08 (14.18)	0.45 (5.31)

For $j_f = 1$ and $j_{\sigma^\tau} = 1$, values in parenthesis are z-statistics of the null hypothesis that the coefficient is zero. For $j_f = 2$ and $j_{\sigma^\tau} = 2$, values in parenthesis are z-statistics of the null hypothesis that the difference between the coefficients of the two regimes is zero.

The results show that the average underlying interest rate in absence of inflation and a zero output gap is 2.52%. The smoothing coefficient implies that about 11% of the adjustment, according to the target, occurs every quarter. The monetary policy rule coefficients on inflation take the values 0.53 and 1.71, depending on the state. The monetary policy rule coefficients on the output gap take the values 1.61 and 0.40, depending on the state. It is important to notice that when the monetary authority is hawkish, less attention is given to the output gap in comparison to the regime when the monetary authority is dovish. With respect to volatility, the standard deviation of the interest rate takes the values 0.24% in the low volatility regime and 1.22% in the high volatility regime.

In regard to the fiscal policy rule, the average real per capita quarterly receipts net of transfers is $1,610 of 2005, in absence of debt and a zero output gap. The smoothing coefficient implies that about 11% of the adjustment, according to the target, occurs every quarter. The fiscal policy rule coefficients on debt take the values 0.05 and 0.13, depending on the state. The fiscal policy rule coefficients on the output gap take the values 0.29 and 0.07, depending on the state. A result to emphasize is that the fiscal authority would pay more attention to the output gap when paying less attention to debt dynamics. With respect to volatility, the standard deviation of the real per capita quarterly tax receipts net of transfers takes the values \$80 of 2005 in the low volatility regime and \$450 of 2005 in the high volatility regime.

Before analyzing the estimates of the transition probabilities and the smoothed probabilities, I make precisions about the labels of the switching states of the model, namely S_m, S_f, S_{σ^R}, and S_{σ^τ}. The labels correspond to the identification conditions imposed in the estimation of the Markov-switching regressions: $\alpha_2^\pi \geq \alpha_1^\pi$, $\gamma_2^b \geq \gamma_1^b$, $\sigma_2^R \geq \sigma_1^R$, and $\sigma_2^\tau \geq \sigma_1^\tau$.

Leeper (1991) labels the monetary and fiscal policy regimes according to the strength of the response of the policy instrument to the targets. Roughly speaking, a strong (weak) response of interest rates to inflation is called an 'Active' ('Passive') monetary policy regime, while a strong (weak) response of taxes to debt is called a 'Passive' ('Active') fiscal policy regime. Hence, there are four possible combinations of regimes, depending on the strength of the response of the policy instruments to their targets. I label the four possible regimes as follows:

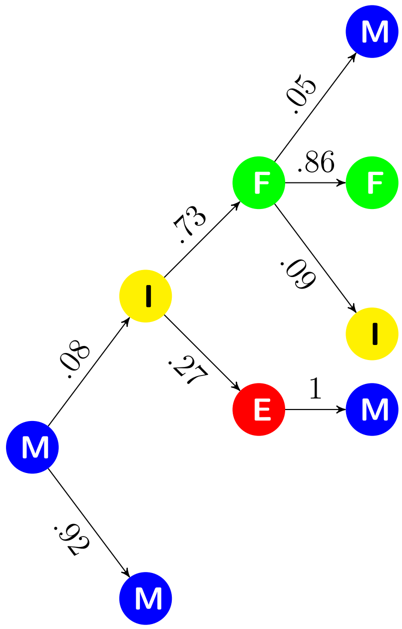
- $S_{mf} = 1$: $(\alpha_1^\pi, \gamma_1^b) \Leftrightarrow$ **F** regime,
- $S_{mf} = 2$: $(\alpha_2^\pi, \gamma_1^b) \Leftrightarrow$ **E** regime,
- $S_{mf} = 3$: $(\alpha_1^\pi, \gamma_2^b) \Leftrightarrow$ **I** regime,
- $S_{mf} = 4$: $(\alpha_2^\pi, \gamma_2^b) \Leftrightarrow$ **M** regime,

Here F stands for 'fiscal', a regime where the fiscal authority is reacting weakly to debt deviations from target and the monetary authority is reacting weakly to inflation deviations from target. M stands for 'monetary', a regime where the monetary authority is reacting strongly to inflation deviations from target and the fiscal authority is reacting strongly to debt deviations from target. I stands for 'indeterminate', a regime where the monetary authority is reacting weakly to inflation deviations from target and the fiscal authority is reacting strongly to debt deviations from target. E stands for 'explosive', a regime where the monetary authority is reacting strongly to inflation deviations from target and the fiscal authority is reacting weakly to debt deviations from target. According to Leeper (1991), regimes M and F could deliver determinacy of the equilibrium in a local-linear version of a dynamic stochastic general equilibrium model, depending on the values of the coefficients. Along the same lines, in regime I there would be indeterminacy of the equilibrium, while in regime E, except for a particular case, there would be no equilibrium with bounded debt.

To better understand the transitional dynamics between the four regimes described above, Figure 8 presents the probability tree implied by the estimated transition probability matrix, along with the ergodic regime probabilities, and the implied correlation between S_m and S_f.

The estimated transition probabilities imply that regime M is expected to last about 12.5 quarters, regime F, about 7.1 quarters, and regimes I and E, 1 quarter each. The ergodic probability of regime M is about 51%, while the ergodic probability of regime F is about 39%. Taken together, the ergodic probabilities of regimes I and E add to about 10%. The probability tree shows that if the economy starts in regime M, the only possibilities would be to stay in regime M with probability 92%, or to move to regime I with probability 8%. If the economy moves to regime I, the only possibilities would be to move to regime E with probability 27%, or to move to regime F with probability 73%. If the economy moves to

Figure 8: Estimated Transition Probabilities, Ergodic Probabilities and Correlation between States



$$P_{mf} = \begin{bmatrix} 0.86 & 0 & 0.73 & 0 \\ 0 & 0 & 0.27 & 0 \\ 0.09 & 0 & 0 & 0.08 \\ 0.05 & 1 & 0 & 0.92 \end{bmatrix}$$

$$\begin{aligned} \mathbb{P}(F) &= .39 \\ \mathbb{P}(E) &= .02 \\ \mathbb{P}(I) &= .08 \\ \mathbb{P}(M) &= .51 \end{aligned}$$

$$\operatorname{corr}(S_m, S_f) = 0.81$$

regime E, the only possibility is to then move to regime M. If the economy moves to regime F, it can stay in regime F with probability 86%, it can move to regime I with probability 9%, or it can move to regime M with probability 5%. Notice that the results rule out the possibility of moving from regime M to F directly. Finally, the implied correlation between the state driving the switching of the monetary policy rule coefficients and the state driving the switching of the fiscal policy rule coefficients is 0.81.

Using the smoothing algorithm of Kim (1994), I obtain the smoothed probabilities for each of the four states. The evolutions of the smoothed probabilities appear in Figure 9. The results show a high complementarity between regimes M and F. According to the smoothed probabilities, the regime M was more likely in place during the 1960s, the 1980s, the second half of the 1990s, and a short period between 2005 and 2007. On the other hand, regime F was more likely in place during a large portion of the 1970s, the first half of the 1990s, a short period between 2003 and 2005, and a period between 2008 and the end of the sample. With respect to regimes I and E, there are short periods in the first half of the 1980s and the end of the sample for the former, and at the beginning of the 1980s for the latter. I will put these results in context with the narrative of monetary-fiscal policymaking in the next section.

In regard to the smoothed probabilities for interest rate and tax volatilities, Figure 10 plots the evolution of these probabilities along with showing the transition probability matrices. The transition probability matrix for interest rate volatility indicates that the low volatility regime is expected to last about 17 quarters, while the high volatility regime, about 9 quarters. On the other hand, the transition probability matrix for taxes indicates that the low volatility regime is expected to last about 14 quarters, while the high volatility regime, about 2 quarters. The smoothed probability for the high volatility regime of interest rates

Figure 9: Smoothed Probabilities - Policy Regimes

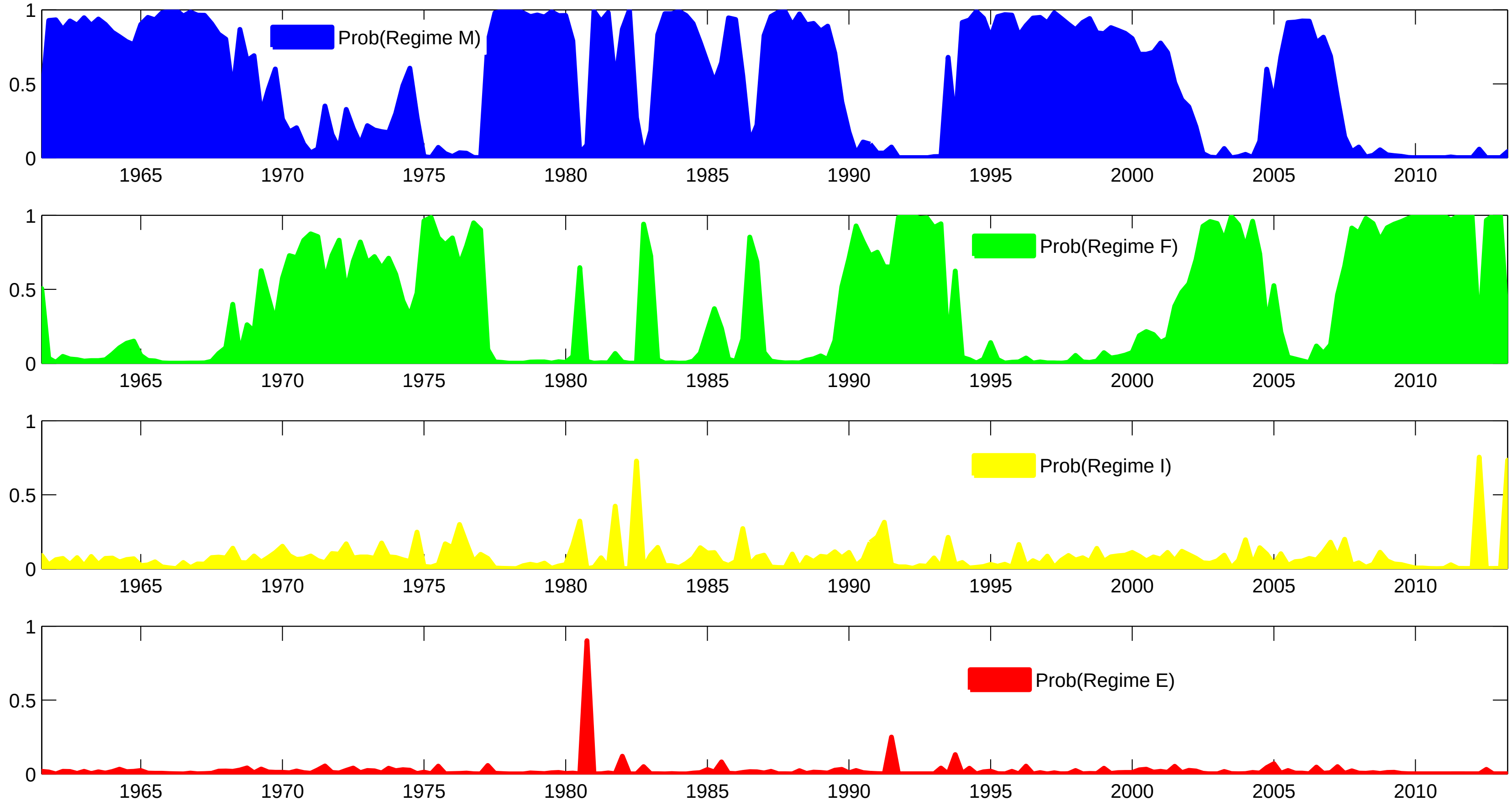
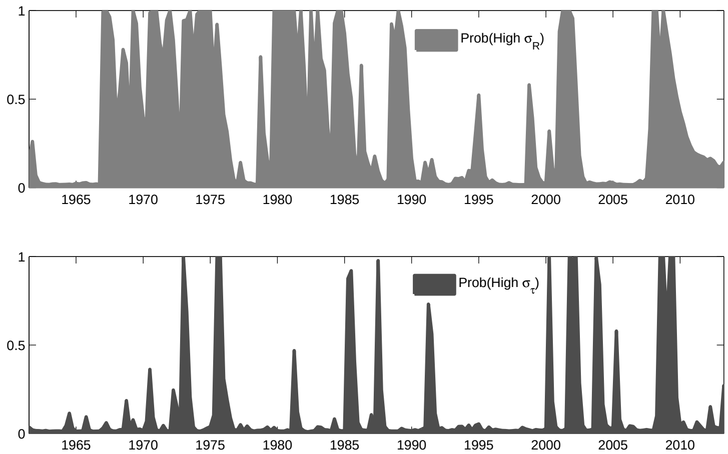


Figure 10: Smoothed Probabilities - Volatility Regimes



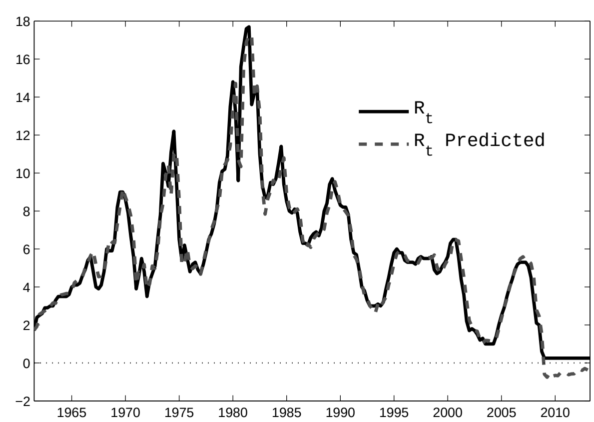
$$P_{\sigma^R} = \begin{bmatrix} 0.94 & 0.11 \\ 0.06 & 0.89 \end{bmatrix}$$

$$P_{\sigma^\tau} = \begin{bmatrix} 0.93 & 0.44 \\ 0.07 & 0.56 \end{bmatrix}$$

indicates that highly volatile interest rates were in place between around 1965 and 1975, the first half of and the end of the 1980s, a few years during the first half of the 2000s, and the 2008-2009 years. One has to recall, however, that identification of the volatility regime over the censored part of the sample is not feasible, given that I have not been able to find dependency between the switching of its state and another switching state of the system. On the other hand, the smoothed probability for the high volatility regime of taxes net of transfers shows a few spikes. The estimates indicate that two short high volatility regimes were likely to be present in the 1970s, a couple more in the 1980s, one at the beginning of the 1990s, at least three in the first half of the 2000s, and a final one during the year 2009. I will put these results in context with the narrative of monetary-fiscal policymaking in the next section.

Finally, Figure 11 shows the evolution of realized and predicted interest rates. The model performs reasonably well to predict the interest rate. In particular, at the end of the sample, the underlying interest rate is below zero and increases gradually.

Figure 11: Observed and Predicted Interest Rate



5 Narrative of the Results

This section puts the results from Figures 9 and 10 in context with the historical narrative on monetary and fiscal policy.

The good conduct of monetary policy dominated the policy mix during the 1960s. Hetzel (2008) compares Fed Chairman William Martin to Fed Chairmen Paul Volcker and Alan Greenspan in that Martin believed that raising short-term interest rates in an expansion was a way to preempt inflation. Despite the Tax Reduction Act of 1964 that cut income tax rates across the board by approximately 20%, fiscal policy remained supporting monetary policy during the 1960s. A fiscal regime starts to take place during the 1970s, possibly due to the expansionary tax reforms of 1971, 1975 and 1976. Hetzel (2008) emphasizes the weak reaction of interest rates to inflation during the 1970s due to the focus of the central bank to promote employment and the belief that inflation was a nonmonetary phenomenon. According to Hetzel, the 1980s saw the commitment of the Federal Reserve to money targets allowing the FOMC to raise interest rates by whatever extent necessary to lower inflation. In general, a monetary regime was in place during this decade, except for a couple of very short fiscal regimes due, most likely, to the expansionary tax reforms of 1981 and 1986. After tightening monetary policy at the end of the 1980s to counteract concerns about inflation, the results show the prevalence of a fiscal regime at the beginning of the 1990s due, possibly, to the combination of policies in reaction to the early 1990s recession. The "covert inflation targeting" of the 1990s (see Mankiw, 2001) and the deficit reduction act of 1993 make a monetary regime more likely during the second half of this decade. The rapid decline in interest rates during the first half of the 2000s and the expansionary tax reforms during that period put the economy, most likely, in a fiscal regime. A monetary regime starts to take place, most likely, after 2005 to avoid inflation pressures and the fact that economic activity was boosting tax revenues. This monetary regime lasts until the second half of 2007 when the central bank adopts a more dovish regime due to recessionary concerns. Once the recession hit in 2008, a fiscal regime is much more likely to have taken place. The budget sequestration and the recovery of tax revenues puts, most likely, a regime of debt-stabilizing fiscal policy and passive monetary policy at the end of the sample.

With respect to volatility, interest rates experienced, most likely, a long period of high interest rate volatility between 1970 and the first half of the 1980s. Then, interest rate volatility decreases except for the stock market crashes of 1989 (Black Monday) and 2000 (Dot-com Bubble). Finally, volatility increases during the recent financial crisis and has, eventually, declined. On the other hand, taxes net of transfers experience spikes in volatility that coincide with some of the tax reforms that we listed in the previous paragraph, and are of very short durations. In particular, there are spikes that coincide with the tax reforms of the 1970s, the 1980s, the deficit reduction act of 1993, the numerous tax reforms of the 2000s, and the recovery act of 2009.

6 Concluding Remarks

This paper devised an estimation technique for a Markov-switching Taylor rule at the zero lower bound. The estimation method allows to obtain consistent estimates of the switching coefficients, the transition probabilities, and, most importantly, identify the prevalent regime of monetary policy. Results show that monetary and fiscal authorities switch between policy regimes in a correlated way and that the policy regime in the first quarter of 2013 leaned towards a regime of passive monetary policy and debt-stabilizing fiscal policy.

The results of the paper suggest that, in modeling monetary policy at the zero lower bound, it is necessary to endow agents with information from fiscal policymaking so that they can draw reasonable inferences on the monetary policy regime. Inferring the monetary-fiscal policy regime right after the lift-off date for the federal funds rate has implications on inflation expectations, as pointed out by Melosi and Bianchi (2013).

Future research goes along the lines of incorporating interdependent monetary and fiscal policy switching in Markov-switching DSGE models to explore the consequences of interdependent switching on the determinacy conditions of the model. Additionally, another possibility is to endow agents that have imperfect information about the prevalent monetary regime at the zero lower bound with a fiscal policy rule to evaluate the implications of forward guidance.

Appendix

A Estimation Ignoring Censoring and Joint Switching

Here the estimation consists of maximizing the log-likelihood function $\mathfrak{L}_{\theta_1}(\theta_1; Y_{1T}) = \ln g_{Y_1}(Y_{1T}; \theta_1)$, with respect to θ_1 by applying the Hamilton filter. The filter allows obtaining $g_{Y_1}(Y_{1T}; \theta_1)$ as follows:

$$\begin{aligned} g_{Y_1}(Y_{1T}; \theta_1) &= \prod_{t=1}^{T} g_{y_1}(y_{1t}|\mathfrak{F}_{1,t-1}; \theta_1) \\ &= \prod_{t=1}^{T} \sum_{j_1=1}^{J_1} g_{y_1|S_1}(y_{1t}|S_{1t} = j_1, \mathfrak{F}_{1,t-1}; \theta_1)\mathbb{P}(S_{1t} = j_1|\mathfrak{F}_{1,t-1}; \theta_1), \end{aligned}$$

where

$$g_{y_1|S_1}(y_{1t}|S_{1t} = j_1, \mathfrak{F}_{1,t-1}; \theta_1) = \frac{1}{\sigma_{1,j_1}} \phi\left(\frac{y_{1t} - x_{1t}'\beta_{1,j_1}}{\sigma_{1,j_1}}\right), \tag{21}$$

and

$$\mathbb{P}(S_{1t} = j_1|\mathfrak{F}_{1,t-1}; \theta_1) = \sum_{j_1'=1}^{J} p_{j_1 j_1'} \mathbb{P}(S_{1,t-1} = j_1'|\mathfrak{F}_{1,t-1}; \theta_1). \tag{22}$$

Once y_{1t} is realized at the end of time t, the filtered probability of S_{1t} in (22) is updated as

$$\mathbb{P}(S_{1t} = j_1|\mathfrak{F}_{1t}; \theta_1) = \frac{g_{y_1|S_1}(y_{1t}|S_{1t} = j_1, \mathfrak{F}_{1,t-1}; \theta_1)\mathbb{P}(S_{1t} = j_1|\mathfrak{F}_{1,t-1}; \theta_1)}{g_{y_1}(y_{1t}|\mathfrak{F}_{1,t-1}; \theta_1)}.$$

To obtain the smoothed probabilities $\mathbb{P}(S_{1t}|\mathfrak{F}_T; \theta_1)$ of the prevalent regime S_{1t} I use the smoothing algorithm in Kim (1994).

B Estimation Ignoring Joint Switching

Here the estimation consists of maximizing the log-likelihood function $\mathfrak{L}_{\theta_1}(\theta_1; Y_{1T}) = \ln f_{Y_1}(Y_{1T}; \theta_1)$, with respect to θ_1 by applying the Hamilton filter. The filter allows obtaining $f_{Y_1}(Y_{1T}; \theta_1)$ as in Appendix A, where I replace $g_{y_1|S_1}(y_{1t}|S_{1t} = j_1, \mathfrak{F}_{1,t-1}; \theta_1)$ in (21) with:

$$\begin{aligned} f_{y_1|S_1}(y_{1t}|S_{1t} = j_1, \mathfrak{F}_{1,t-1}; \theta_1) = &\left[\Phi\left(\frac{y_{1L} - x_{1t}'\beta_{1,j_1}}{\sigma_{1,j_1}}\right)\right]^{1[y_{1t} = y_{1L}]} \\ &\times \left[\frac{1}{\sigma_{1,j_1}}\phi\left(\frac{y_{1t} - x_{1t}'\beta_{1,j_1}}{\sigma_{1,j_1}}\right)\right]^{1[y_{1t} > y_{1L}]}. \end{aligned}$$

C Data Transformation

The transformation of the data is as follows:

- R_t: It is the quarterly federal funds rate until 2008:3. Starting 2008:4, the rate is fixed at 0.25%.
- τ_t: There are two choices for this variable:
  - The real per capita tax receipts net of transfers. The GDP deflator is used to deflate the series to (thousand) dollars of 2005, and the total population is used to transform the series to per capita terms.
  - The ratio of tax receipts net of transfers to GDP.
- b_{t-1}: There are two choices for this variable, which correspond with the two choices of the tax series above:
  - The average over the last four quarters of the real per capita market value of privately held gross federal debt. The GDP deflator is used to deflate the series to dollars of 2005, and the total population is used to transform the series to per capita terms.
  - The average over the last four quarters of the ratio of the market value of privately held gross federal debt to GDP.
- Among the instrumental variables, government spending is transformed to the ratio over GDP.

D Estimation with Correlated S_m and S_{σ^R}

Estimates of the specification where S_m and S_{σ^R} are correlated appear in Table 4.

Table 4: Parameter Estimates

Monetary Policy Rule

Parameters	$j_m = 1$	$j_m = 2$
R_{j_m}	2.59 (5.08)	
$\rho^R_{j_m}$	0.89 (63.51)	
$\alpha^\pi_{j_m}$	0.53 (3.62)	1.73 (3.62)
$\alpha^y_{j_m}$	1.71 (7.38)	0.40 (-8.41)
	$j_{\sigma^R} = 1$	$j_{\sigma^R} = 2$
$\sigma^R_{j_{\sigma R}}$	0.25 (11.51)	1.21 (8.71)

For $j_m = 1$ and $j_{\sigma R} = 1$, values in parenthesis are z-statistics of the null hypothesis that the coefficient is zero. For $j_m = 2$ and $j_{\sigma R} = 2$, the values in parenthesis are z-statistics of the null hypothesis that the difference between the coefficient of the two regimes is zero.

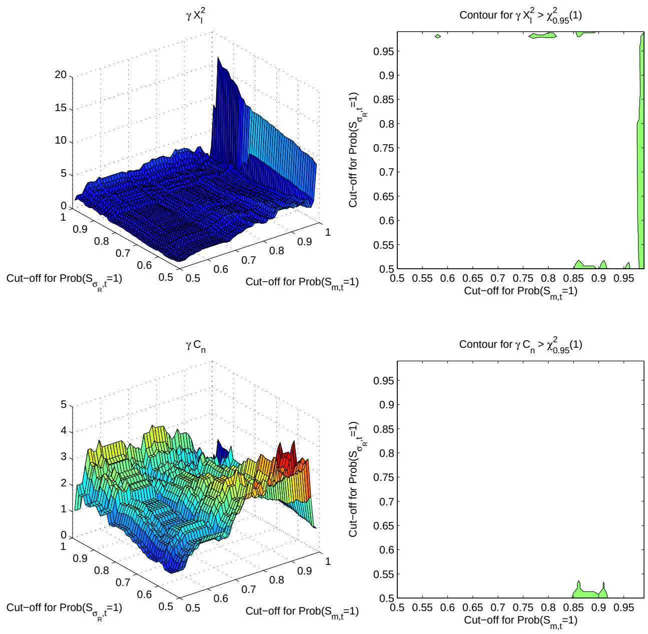
Fiscal Policy Rule

Parameters	$j_f = 1$	$j_f = 2$
τ_{j_f}	1.64 (11.30)	
$\rho^\tau_{j_f}$	0.89 (68.91)	
$\gamma^b_{j_f}$	0.04 (2.48)	0.12 (5.19)
$\gamma^y_{j_f}$	0.25 (5.69)	0.07 (-6.20)
	$j_{\sigma^\tau} = 1$	$j_{\sigma^\tau} = 2$
$\sigma^\tau_{j_{\sigma^\tau}}$	0.08 (14.12)	0.45 (5.28)

For $j_f = 1$ and $j_{\sigma^\tau} = 1$, values in parenthesis are z-statistics of the null hypothesis that the coefficient is zero. For $j_f = 2$ and $j_{\sigma^\tau} = 2$, the values in parenthesis are z-statistics of the null hypothesis that the difference between the coefficient of the two regimes is zero.

E Results of the Independence Test between S_m and S_{σ_R}

Figure 12: Independence Test S_m and S_{σ_R}



Notes: The tests statistics are defined as:

$$X_l^2 = \left\{ \ln\left(\frac{n_1 n_4}{n_3 n_2}\right)\right\}^2 \Big/ \left(\frac{1}{np_1} + \frac{1}{np_2} + \frac{1}{np_3} + \frac{1}{np_4}\right)$$

$$C_n = \frac{n\left(n_1 n_4/n^2 - n_2 n_3/n^2\right)^2}{n^{-4}\left(n_1+n_2\right)\left(n_1+n_3\right)\left(n_2+n_4\right)\left(n_3+n_4\right)},$$

where $p_{j_{m\sigma R}} = \mathbb{P}\left(S_{m\sigma R} = j_{m\sigma R}\right)$, and $S_{m\sigma R} = (S_{\sigma R} - 1)J_m + S_m$ and $j_{m\sigma R} = (j_{\sigma R} - 1)J_m + j_m$. Also, $n_{j_{m\sigma R}} = \sum_{t=1}^{n} \mathbb{1}_{\{S_{m\sigma R,t} = j_{m\sigma R}\}}$. Additionally, $\gamma = (1 - \mu\lambda)(1 + \mu\lambda)$, where μ and λ are the nonunit eigenvalues of P_m and $P_{\sigma R}$, the transition probability matrices of S_m and $S_{\sigma R}$, respectively.

References

Bae, Jinho, Chang-Jin Kim, and Dong Kim (2012) 'The evolution of the monetary policy regimes in the u.s.' *Empirical Economics* 43(2), 617–649

Bewick, Viv, Liz Cheek, and Jonathan Ball (2004) 'Statistics review 13: receiver operating characteristic curves.' *Critical care* 8(6), 508

Bianchi, Francesco (2012) 'Evolving monetary/fiscal policy mix in the united states.' *American Economic Review* 102(3), 167–72

Bianchi, Francesco (2013) 'Regime switches, agents' beliefs, and post-world war ii us macroeconomic dynamics.' *The Review of Economic Studies* 80(2), 463–490

Chung, Hess, Troy Davig, and Eric Leeper (2007) 'Monetary and fiscal policy switching.' *Journal of Money, Credit and Banking* 39(4), 809–842

Clarida, Richard H., Jordi Gali, and Mark Gertler (2000) 'Monetary policy rules and macroeconomic stability: Evidence and some theory.' *The Quarterly Journal of Economics* 115(1), 147–180

Davig, Troy, and Eric Leeper (2006) 'Fluctuating macro policies and the fiscal theory.' In 'NBER Macroeconomics Annual 2006, Volume 21' (National Bureau of Economic Research, Inc) pp. 247–316

_ (2011) 'Monetary-fiscal policy interactions and fiscal stimulus.' *European Economic Review* 55(2), 211–227

Davig, Troy, and Taeyoung Doh (2009) 'Monetary policy regime shifts and inflation persistence.' Research Working Paper RWP 08-16, Federal Reserve Bank of Kansas City

Eo, Yunjong (2009) 'Bayesian analysis of dsge models with regime switching.' MPRA Paper, University Library of Munich, Germany

Gonzalez-Astudillo, Manuel (2013) 'Monetary-fiscal policy interactions: interdependent policy rule coefficients.' Finance and Economics Discussion Series 2013-58, Board of Governors of the Federal Reserve System (U.S.)

Hamilton, James D. (1989) 'A new approach to the economic analysis of nonstationary time series and the business cycle.' *Econometrica* 57(2), pp. 357–384

Hetzel, Robert L. (2008) *The Monetary Policy of the Federal Reserve: A History (Studies in Macroeconomic History),* 1 ed. (Cambridge University Press)

Kim, Chang-Jin (1994) 'Dynamic linear models with markov-switching.' *Journal of Econometrics* 60(12), 1 – 22

_ (2004) 'Markov-switching models with endogenous explanatory variables.' *Journal of Econometrics* 122(1), 127 – 136

— (2009) ‘Markov-switching models with endogenous explanatory variables ii: A two-step mle procedure.’ *Journal of Econometrics* 148(1), 46 – 55

Leeper, Eric (1991) ‘Equilibria under ’active’ and ’passive’ monetary and fiscal policies.’ *Journal of Monetary Economics* 27(1), 129–147

Lusted, Lee (1971) ‘Signal detectability and medical decision-making.’ *Science* 171(3977), 1217–1219

Mankiw, N. Gregory (2001) ‘U.s. monetary policy during the 1990s.’ NBER Working Papers 8471, National Bureau of Economic Research, Inc

Melosi, Leonardo, and Francesco Bianchi (2013) ‘Escaping the great recession.’ 2013 Meeting Papers 203, Society for Economic Dynamics

Murray, Christian, Alex Oleksandr Nikolsko-Rzhevskyy, and David H. Papell (2013) ‘Markov switching and the taylor principle.’ Working Papers 2013-219-06, Department of Economics, University of Houston

Peterson, WWTG, T Birdsall, and We Fox (1954) ‘The theory of signal detectability.’ *Information Theory, IRE Professional Group on* 4(4), 171–212

Swets, John A (1979) ‘Roc analysis applied to the evaluation of medical imaging techniques.’ *Investigative radiology* 14(2), 109–121

Tavaré, Simon, and Patricia ME Altham (1983) ‘Serial dependence of observations leading to contingency tables, and corrections to chi-squared statistics.’ *Biometrika* 70(1), 139–144

Tobin, James (1958) ‘Estimation of relationships for limited dependent variables.’ *Econometrica: Journal of the Econometric Society* pp. 24–36