Regime Switching Models with Multiple Dynamic Factors

Yoosoon Chang*
Department of Economics
Indiana University

Joon Y. Park†
Department of Economics
Indiana University
and Sungkyunkwan University

Shi Qiu‡
Department of Economics
Indiana University

June 5, 2018

Very Preliminary. Please Do Not Quote.

Abstract

This paper introduces a regime switching model driven by bivariate dynamic latent factors, which allows for unsynchronized regime shifts and endogenous feedbacks. The regimes are determined by whether the autoregressive latent factors signifying the strength of regimes are realized above or below thresholds. On the other hand, the endogenous feedback is modeled by a correlation between current innovation to latent regime factors and innovations to observed time series next period. A modified Markov switching filter is developed to estimate the model by the maximum likelihood method, which can also be used to extract the latent regime factors. The extracted latent factors show, more directly and clearly, how unsynchronized regime shifts are interrelated with each other. As an application, we examine U.S. monetary and fiscal policy interactions. In our framework, policy interactions are effectively characterized by the joint dynamics of the latent regime factors that we interpret as the policy factors.

JEL Classification: C13, C22

Keywords: regime switching, multiple dynamic factors, latent autoregressive factors, regime switching in regressions

*Bloomington, IN 47405-7104. E-mail: yoosoon@indiana.edu.
†Bloomington, IN 47405-7104. E-mail: joon@indiana.edu.
‡Bloomington, IN 47405-7104. E-mail: shiqiu@indiana.edu.

1 Introduction

Regime switching model is widely applicable when modeling abrupt shifts in macroeconomic time series. For example, we consider an application a two-equation state space model of regime switching Taylor rule and fiscal policy rule, with two dynamic latent factors. The positions of the bivariate latent factors relative to a threshold vector define the strength of regimes as well as the regimes themselves. It is noted in the literature that events such as the sizeable monetary policy shift of U.S. in late 1979 are unlikely to be driven by exogenous regime change. Moreover, despite the commonly presumed independence of policy authorities, a substantive question pertains to the characterization of unobserved (de)synchronization of these policy regime shifts. In this paper, we propose a threshold-type regime-switching model with bivariate vector autoregressive dynamic factors, allowing "endogenous feedback" modeled as an intertemporal correlation between shocks to the observed time-series and those to the latent factors. Then upon the model, we devise a modified Markov-switching filter.

The convential Markov-switching model proposed in the seminal work of Hamilton [1989] is widely used and proven fruitful in many applications. It stipulates a sequence of regimes following an exogenous, homogeneous Markov chain. Nonetheless, economic considerations often render it more plausible and desirable to allow time-varying transition probabilities. Many authors, including Kim [1994, 2004, 2009], Diebold et al. [1994], Chib [1996], Chib and Dueker [2004], Kim et al. [2008], Bazzi et al. [2014], Kang [2014], Kalliovirta et al. [2015], and Chang et al. [2017] (CCP, hereafter), among many others, explore along this direction and introduce time-varying transition probabilities. For an overview of regime switching models, the reader is referred to the monograph by Kim and Nelson [1999].

Our model is a vector extension of CCP, in which an autoregressive process of order one (AR(1)) drives the two-state regime as it crosses a threshold parameter on the real line $\mathbb{R}$, with current shocks to measurement equation correlating with next periods innovations to factors. We extend by considering stationary bivariate latent factors following vector autoregression of order one (VAR(1)), with a vector of shocks to measurement correlated with the vector of next periods factor innovations. CCP also allows $I(1)$ latent factor, which we will not explore in this paper. Instead, we argue stationarity is not a restrictive assumption because one should otherwise expect very few switches: a data feature that is better captured using other methods, such as a linear model with structural breaks. The regime in our setup is a bivariate binary vector, determined as the latent factor moves across quadrants defined by a pair of threshold parameters on the real plane $\mathbb{R} \times \mathbb{R}$. Our model nests CCP in the sense that stacking two unrelated univariate models corresponds to diagonal factor loading matrix and a restricted correlation matrix between vectors of shocks and vectors of innovations.

A key feature of our model is the endogenously determined time-varying transition of regimes. Rather than relying upon a transition equation augmented by exogenous variables, this time-variation stems purely from the error terms in the regression, propagating through intertemporal feedback. For a statistical model, we call this mechanism the endogenous feedback.[1] It is straightforward to observe that endogenous feedback results to the state of the regime being a function of the entire history of shocks.

In addition to the apparent extension of methodology, vector specification has two essential advantages in application over univariate modeling.

First, we identify coordination in the latent factors through the off-diagonal elements of factor loading matrix as they represent the effects of the lagged factors onto current ones, as well as through correlation of factor innovations. If interpreting policy factors as authorities' implicit positions toward a particular regime, then the factor loading matrix describes the extent to which their positions comove systematically, and the correlation of innovation describes how much of their discretion of regime is based on shared information. The univariate framework of CCP cannot consistently account for this level of policy interaction. Indeed, factor dynamics and feedback together characterize the synchronization in regime shifts, whereby our approach is flexible enough to capture a spectrum of synchronizations between fully synchronized and unsynchronized switching. In contrast, CCP allows only fully (un)synchronized regime switching, depending on model specification. One can effectively impose unsynchronized switching by considering two unrelated univariate models. If one instead considers a vector of measurements with a single latent factor, then regime switching must be fully synchronized. We argue this flexibility is vital as both extremities as assumptions are dubious in applications.

Second, vector specification naturally introduces cross-equation feedback. The intertemporal correlation stipulated in CCP may be viewed as within-equation feedback because shocks to the observed time series transmit only to its corresponding regime factors. In comparison, we allow current shocks to the observed time series to transmit to both regime factors in the next period. Empirically, one often interprets these shocks to measurement equations as omitted and perhaps quantitatively important information orthogonal to the independent variables included in the regression. Cross-equation feedback is thus desirable in a broad class of applications since these shocks are more plausibly stipulated to be measurable with respect to a common information set of all equations due to certain

[1] For a dynamic stochastic general equilibrium (DSGE) model, this design in its strict form is still generally considered exogenous unless regimes are shifted in part by agents' optimal choice.

misspecification.

A large literature studies various statistical properties of regime switching models such as Hansen [1992], Hamilton [1996], Garcia [1998], Timmermann [2000], and Cho and White [2007]. However, identification is often unclear. We report, to our knowledge, a novel identification result for the regime-switching model with endogenous feedback. Specifically, we collect the standard identification results for mixture models and establish a one-to-one correspondence between our model parameters and the identified mixing distribution and mixing components. Normality of factor innovations appears to be critical in our proof, but the normality of shocks to measurements is less so. We acknowledge that a large class of shock distributions also falls within the scope of our result.

As an illustration, we analyze a slight extension of Chang and Kwak [2017] to study U.S. monetary-fiscal policy interactions. On top of rediscovering substantial size of within-equation feedback,[2] we report evidence of non-trivial synchronization between monetary and fiscal factors, and non-trivial cross-equation feedback. Specifically, we report three primary results. First, the impact of fiscal factor on monetary factor is twice in size compared to the impact of monetary factor on the fiscal factor. Second, a common factor that drives the shocks of both policy rules. Third, the common factor is fed mainly to next period monetary factor innovation, whereas the fiscal factor is mainly influenced by its equation-specific policy shocks. We also find the extracted latent factors from our model to be highly positively correlated over the course of history with occasional exceptions in the 1950s, 1970s and early 2000s. The coherence of extracted latent factors suggests a substantial correlation between policy factors over the typical business cycle frequency. These results indicate a tighter connection between the monetary and fiscal regimes than those reported by Chang and Kwak [2017].

We simulate at ML estimates to evaluate the performance of our approach against CCP. Our results suggest that ignoring cross-equation interaction inflicts substantial bias for the threshold values, the size of interaction and feedback channels, and the correlation of shocks. Nevertheless, we find relatively small bias for the coefficients in measurement equations when the latent factors extracted from restricted and unrestricted models imply similar regimes in timing and length. To supplement empirical exercises, we obtain standard errors of parameter estimates by simulation since it is computationally challenging to evaluate Hessian matrix of the log-likelihood function. Using the simulated standard error, we conclude improved efficiency after allowing cross-equation temporal dynamics and feedback.

[2] Chang et al. [2017] document ubiquitous and robust feedback in both macro and financial time-series. Chang and Kwak [2017] also report strong feedback in the series of monetary and fiscal policy instruments.

The rest of the paper is organized as follows. Section 2 introduces the model and characterize transition probabilities. Section 3 presents our modified Markov filter and characterizes the likelihood function. In Section 4, we establish identification. In Section 5, we examine a reduced form bivariate system of monetary and fiscal policy rules and analyze the policy regime interactions. Moreover, Section 6 reports our simulation results and supplements empirical section by evaluating the performance of our model relative to the model without cross-equation interactions. Finally, Section 8 concludes the paper. Appendix collects computational details, omitted proofs, additional tables, and figures.

2 The Model and Preliminaries

We consider

$$y_{1t} = x_{1t}'\beta_{1t} + \sigma_1 u_{1t}$$
$$y_{2t} = x_{2t}'\beta_{2t} + \sigma_2 u_{2t}$$

with exogenous variable $x_t = (x_{1t}, x_{2t})$ and level functions

$$\begin{aligned} \beta_{it} &= \underline{\beta}_i(1 - s_{it}) + \overline{\beta}_i s_{it} \\ \sigma_{it} &= \underline{\sigma}_i(1 - s_{it}) + \overline{\sigma}_i s_{it} \end{aligned}$$

for $i = 1, 2$, where (s_{it}) for $i = 1, 2$ are binary state variables taking values 0 and 1, denoting respectively the low and high states and are generated as

$$s_{it} = 1\{w_{it} \geq \tau_i\}$$

with

$$w_t = Aw_{t-1} + v_t,$$

where $w_t = (w_{1t}, w_{2t})'$, $v_t = (v_{1t}, v_{2t})'$ and

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$$

such that both of its eigenvalues lie inside unit disk. With stationarity, $w_t \sim \mathbb{N}(0, \Sigma_w)$ such that

$$\Sigma_{ww} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}$$

Defining $u_t = (u_{1t}, u_{2t})'$, we further specify

$$(u_t', v_{t+1}')' =_d \mathbb{N}(0, P)$$

with

$$P = \begin{pmatrix} P_{uu} & P_{uv} \\ P_{vu} & P_{vv} \end{pmatrix} = \begin{pmatrix} 1 & & & \\ \rho_{u_2u_1} & 1 & & \\ \rho_{v_1u_1} & \rho_{v_1u_2} & 1 & \\ \rho_{v_2u_1} & \rho_{v_2u_2} & \rho_{v_2v_1} & 1 \end{pmatrix},$$

and assume that (u_t) and (v_t) are i.i.d. and independent of each other at all other leads and lags.

We say there is endogenous feedback in the regime switching if $P_{vu} \neq 0$, and say the regime switching is exogenous if otherwise. Note in particular that u_t is correlated with v_{t+1}, not v_t, for all $t = 1, 2, \ldots$. The unit variance in P is necessary because when $A = 0$, the pair (τ, P_{vv}) and $(c\tau, cP_{vv})$ for any $c > 0$ imply identical state transition probabilities.

If we let

$$z_t = v_t - P_{vu}P_{uu}^{-1}u_{t-1},$$

it follows that z_t is independent of $u_{t-1}, u_{t-2}, \ldots$, and of $v_{t-1}, v_{t-2}, \ldots$, and therefore, also of $w_{t-1}, w_{t-2}, \ldots$, and that

$$z_t =_d \mathbb{N}\left(0, P_{vv\cdot u}\right)$$

with

$$P_{vv\cdot u} = P_{vv} - P_{vu}P_{uu}^{-1}P_{uv}, \tag{2.1}$$

i.e., the conditional variance of v_{t+1} given u_t for each $t = 1, 2, \ldots$. Therefore, we have

$$\begin{aligned} &\mathbb{P}\left\{w_t < \tau \middle| w_{t-1}, y_{t-1}, x_{t-1}\right\} \\ &= \mathbb{P}\left\{z_t < \tau - P_{vu}P_{uu}^{-1}u_{t-1} - Aw_{t-1} \middle| w_{t-1}, y_{t-1}, x_{t-1}\right\} \\ &= \Phi_{v|u}\left(\tau - P_{vu}P_{uu}^{-1}u_{t-1} - Aw_{t-1}\right), \end{aligned}$$

where $\Phi_{v|u}$ is the distribution function of bivariate normal distribution with covariance matrix $P_{vv\cdot u}$ defined in (2.1).

In what follows, we show more explicitly how we may obtain the probabilities required for our modified Markov filter. We may easily deduce that

$$
\begin{aligned}
&\mathbb{P}\left\{w_t < \tau \middle| w_{t-1} < \tau, y_{t-1}, x_{t-1}\right\} \\
&= \mathbb{P}\left\{w_t < \tau, w_{t-1} < \tau | y_{t-1}, x_{t-1}\right\} \Big/ \mathbb{P}\left\{w_{t-1} < \tau \middle| y_{t-1}, x_{t-1}\right\} \\
&= \left[\int_{-\infty}^{\tau} \Phi_{v|u}\left(\tau - P_{vu}P_{uu}^{-1}u_{t-1} - Aw_{t-1}\right)\phi(w_{t-1})dw_{t-1}\right] \Bigg/ \Phi(\tau),
\end{aligned} \tag{2.2}
$$

where $\Phi_{v|u}$ is defined as above in (2.1), and ϕ and Φ are respectively the density and distribution functions of (w_t), which is bivariate normal distribution with covariance matrix Σ_{ww}, which is given by

$$\mathrm{vec}\,\Sigma_{ww} = (I - A \otimes A)^{-1}\mathrm{vec}\,P_{vv},$$

where $\mathrm{vec}(\cdot)$ is the operator stacking rows of a matrix and transforming it into a column vector.

The transition probability (2.2) is time-invariant if $P_{vu} = 0$. In this case, our model reduces to the conventional Markov switching model after relabeling the states $(0,0)'$, $(0,1)'$, $(1,0)'$ and $(1,1)'$ to be $1, 2, 3$ and 4.

For the actual computations of transition probabilities, we need to calculate

$$N(A, b, c, \Sigma_1, \Sigma_2) = \int_{-\infty}^{c_2}\int_{-\infty}^{c_1} \Phi_1\left(b - Ax\right)\phi_2(x)dx_1dx_2 \tag{2.3}$$

for various sets of values of 2×2 matrix A, and two dimensional vectors b and c, where Φ_1 and ϕ_2 are respectively the bivariate normal distribution and density functions with covariance matrices $\Sigma_1 = P_{vv\cdot u}$ and $\Sigma_2 = \Sigma_{ww}$. For this purpose, we note that

$$N(A, b, c, \Sigma_1, \Sigma_2) = \mathbb{P}\left\{\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \leq \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \leq b - AX\right\},$$

where $X = (X_1, X_2)'$ is a bivariate normal random variate with mean zero and covariance matrix given by Σ_2, and $Y = (Y_1, Y_2)'$ is also a bivariate normal random variate, independent of X, with mean zero and covariance matrix Σ_1. However, we may rewrite

$$\mathbb{P}\left\{\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \leq \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \leq b - AX\right\} = \mathbb{P}\{X \leq c, Y + AX \leq b\},$$

and therefore, it follows that

$$N(A,b,c,\Sigma_1,\Sigma_2) = \int_{-\infty}^{b_2}\int_{-\infty}^{b_1}\int_{-\infty}^{c_2}\int_{-\infty}^{c_1} p_{A,\Sigma_1,\Sigma_2}(z_1,z_2,z_3,z_4)dz_1dz_2dz_3dz_4,$$

where p_{A,Σ_1,Σ_2} is the density function of a four-dimensional normal random variate with zero mean and covariance matrix

$$\begin{pmatrix} \Sigma_2 & \Sigma_2 A' \\ A\Sigma_2 & \Sigma_1 + A\Sigma_2 A' \end{pmatrix}$$

for given Σ_1 and Σ_2.

Note that

$$\begin{aligned}
&\int_{c_2}^{\infty}\int_{-\infty}^{c_1} \Phi_1\left(b - Ax\right)\phi_2(x)dx_1dx_2 \\
&\qquad = N(A,b,\begin{pmatrix} c_1 \\ \infty \end{pmatrix},\Sigma_1,\Sigma_2) - N(A,b,c,\Sigma_1,\Sigma_2), \\
&\int_{-\infty}^{c_2}\int_{c_1}^{\infty} \Phi_1\left(b - Ax\right)\phi_2(x)dx_1dx_2 \\
&\qquad = N(A,b,\begin{pmatrix} \infty \\ c_2 \end{pmatrix},\Sigma_1,\Sigma_2) - N(A,b,c,\Sigma_1,\Sigma_2), \\
&\int_{c_2}^{\infty}\int_{c_1}^{\infty} \Phi_1\left(b - Ax\right)\phi_2(x)dx_1dx_2 \\
&\qquad = N(A,b,\begin{pmatrix} \infty \\ \infty \end{pmatrix},\Sigma_1,\Sigma_2) - N(A,b,\begin{pmatrix} \infty \\ c_2 \end{pmatrix},\Sigma_1,\Sigma_2) - N(A,b,\begin{pmatrix} c_1 \\ \infty \end{pmatrix},\Sigma_1,\Sigma_2) \\
&\qquad\quad + N(A,b,c,\Sigma_1,\Sigma_2),
\end{aligned}$$

which can also be easily obtained, once we compute the integral in (2.3).[3]

3 A Modified Markov-Switching Filter

In this section, we characterize the likelihood function and extract latent factors by developing a modified Markov-switching filter.

[3] Calculation of (2.3) involves non-trivial randomness, since efficient implementations such as the Matlab [Genz, 1992] employ Monte-Carlo integration techniques. The issue can be practically resolved by fixing the random seed of Monte-Carlo integration. The disadvantage is that we lose control of precision. But our limited experience suggests the resulting precision is roughly to the level of 10^{-4}.

3.1 Likelihood Function

Let information set $\mathcal{F}_0 = \emptyset$ and $\mathcal{F}_t = \sigma(y_{1:t}, x_{1:T})$ for $t = 1, \cdots, T$ and suppress x_t in all future notation for it is exogenous and will not change our arguments. The likelihood function for a vector of parameters $\theta \in \Theta$ takes form

$$\ell(\theta) = p(Y_{1:T}|\theta) \quad = \quad \prod_{t=1}^{T} \left(\sum_{s_t} p(y_t|s_t, \mathcal{F}_{t-1}) p(s_t|\mathcal{F}_{t-1}) \right) \tag{3.1}$$

Equation (3.1) can be evaluated sequentially with following predict-update recursion. An equivalent but parallelizable algorithm is proposed in Appendix A.

Algorithm 3.1 (Modified Markov-Switching Filter)**.**

1. ***Initialization.*** *If $t = 0$, set $p(s_t)$ to be the unconditional state probabilities such that $\mathbb{P}\{s_0 = (0,0)'\} = \mathbb{P}\big\{w_{1,0} < \tau_1, w_{2,0} < \tau_2\big\}$, in which $(w_{1,0}, w_{2,0})' \sim \mathbb{N}(0, \Sigma_{ww})$.*

2. ***Recursion.*** *If $t \geq 1$, repeat steps (a) - (c). Stop if $t = T + 1$.*

   (a) ***Forecasting.*** *Predict the distribution of s_t given information set $\mathcal{F}_{t-1}$*

   $$p(s_t|\mathcal{F}_{t-1}) = \sum_{s_{t-1}} p(s_t|s_{t-1}, \mathcal{F}_{t-1}) p(s_{t-1}|\mathcal{F}_{t-1})$$

   with time-varying transition probability $p(s_t|s_{t-1}, \mathcal{F}_{t-1})$ characterized by (2.2).

   (b) ***Evaluation.*** *Calculate the conditional density of y_t given information set $\mathcal{F}_{t-1}$*

   $$p(y_t|\mathcal{F}_{t-1}) = \sum_{s_t} p(y_t|s_t, \mathcal{F}_{t-1}) p(s_t|\mathcal{F}_{t-1}).$$

   (c) ***Updating.*** *Update the distribution of s_t given information set $\mathcal{F}_t$*

   $$\begin{aligned} p(s_t|\mathcal{F}_t) &= \sum_{s_{t-1}} p(s_t, s_{t-1}|y_t, \mathcal{F}_{t-1}) \\ &= \sum_{s_{t-1}} \frac{p(y_t|s_t, \mathcal{F}_{t-1}) p(s_t|s_{t-1}, \mathcal{F}_{t-1}) p(s_{t-1}|\mathcal{F}_{t-1})}{p(y_t|\mathcal{F}_{t-1})}. \end{aligned}$$

Several remarks are in order. First, immediately useful byproducts of Algorithm 3.1 are the time-varying transition probabilities $p(s_t|s_{t-1}, \mathcal{F}_{t-1})$ and the filtered state probabilities $p(s_t|\mathcal{F}_t)$, with which inferences of current and future states can be drawn. Second, an optional step parallel in sequence to *Updating* can be easily implemented to extract conditional moments of latent factor w_t given information $\mathcal{F}_t$ in each iteration of *Recursion.* In particular, we show in Subsection 3.2 the conditional density $p(w_t|\mathcal{F}_t)$ is characterized by densities and distributions computed in previous iterations. Finally, Algorithm 3.1 can be extended to admit regime switching autoregression of order m by replacing conditional probabilities $p(s_t|\mathcal{F}_{t-1})$, $p(s_t|s_{t-1}, \mathcal{F}_{t-1})$, $p(s_{t-1}|\mathcal{F}_{t-1})$ and $p(y_t|s_t, \mathcal{F}_{t-1})$ for each t by $p(s_{t-m:t}|\mathcal{F}_{t-1})$, $p(s_t|s_{t-m-1:t-1}, \mathcal{F}_{t-1})$, $p(s_{t-m-1:t-1}|\mathcal{F}_{t-1})$ and $p(y_t|s_{t-m:t}, \mathcal{F}_{t-1})$, respectively. Accordingly, the summations in *Forecasting* and *Updating* are to be adjusted to marginalize out s_{t-m-1}, and the summation in *Evaluation* marginalizes out $s_{t:t-m}$. The characterization of transition probability is invariant to this extension by *Theorem 3.1* in CCP.

3.2 Extraction of Latent Factor

We now characterize the conditional density $p(w_t|\mathcal{F}_t)$, with which the extraction of, for example, $\mathbb{E}(w_t|\mathcal{F}_t)$ becomes a standard integration exercise. Applying Bayes formula to write

$$p(w_t|\mathcal{F}_t) \quad = \quad \sum_{s_{t-1}} \frac{p(w_t|s_{t-1}, \mathcal{F}_{t-1}) p(y_t|s_t, \mathcal{F}_{t-1}) p(s_{t-1}|\mathcal{F}_{t-1})}{p(y_t|\mathcal{F}_{t-1})}, \tag{3.2}$$

in which all densities but $p(w_t|s_{t-1}, \mathcal{F}_{t-1})$ have been specified above.

For $s_{t-1} = (0,0)'$, there is

$$\begin{aligned} p(w_t|s_{t-1} = (0,0)', \mathcal{F}_{t-1}) \quad = \quad & \left(\det \Omega^{-1} Q^{-1} P_{vv\cdot u} \Sigma_{ww}\right)^{-1/2} \\ & \times \frac{\Phi_Q\left(\tau - QA' P_{vv\cdot u}^{-1} (w_t - P_{vu} P_{uu}^{-1} u_{t-1})\right)}{\Phi_{\Sigma_{ww}}(\tau)} \\ & \times \phi\left(w_t; P_{vu} P_{uu}^{-1} u_{t-1}, \Omega\right) \end{aligned} \tag{3.3}$$

in which

$$\begin{aligned} Q &= (A' P_{vv\cdot u}^{-1} A + \Sigma_{ww}^{-1})^{-1}, \\ \Omega &= (P_{vv\cdot u}^{-1} - P_{vv\cdot u}^{-1} A Q A' P_{vv\cdot u}^{-1})^{-1}, \end{aligned}$$

with generic notation $\Phi_M(\cdot)$ denoting zero-mean bivariate normal distribution function with covariance matrix M, and $\phi(\cdot; \mu, \Omega)$ denoting bivariate normal density with mean μ and covariance Ω. For a proof of (3.2) and (3.3), see Appendix C.1. Similar calculation easily delivers the conditional densities $p(w_t|s_{t-1}, \mathcal{F}_{t-1})$ in which $s_{t-1} = (0,1)', (1,0)'$ and $(1,1)'$.

The calculation of $\mathbb{E}(w_t|\mathcal{F}_t)$ is computationally demanding due to both the non-standard shape of underlying distribution and high dimension. Many methods are available for efficient calculation. For simplicity, we apply self-normalized importance sampler. The detail of our sampler is reported in Appendix B.1 as the treatment is now standard in literature.

4 Identification

In this section, we first bridge the standard theory of mixture distribution to show identification of parameters in state transition. Then we show the parameters characterizing feedback are identified through the interaction between transition and measurement. Finally, the parameter identification in the measurement equation follows trivially from Gaussianity of regression errors.

4.1 Identification of Mixture Distribution

Following standard terminologies of mixture distribution, the likelihood function

$$\begin{aligned} p(y_1, y_2, \cdots, y_T) &= \sum_{s_1,\cdots,s_T} p(y_1, y_2, \cdots, y_T | s_1, s_2, \cdots, s_T) p(s_1, s_2, \cdots, s_T) \\ &= \sum_{s_1,\cdots,s_T} \left(\prod_{t=1}^{T} p(y_t | s_t, \mathcal{F}_{t-1}) \right) p(s_1, s_2, \cdots, s_T) \end{aligned} \tag{4.1}$$

is a finite mixture of normal densities with mixing distribution $p(s_1, s_2, \cdots, s_T)$ and mixing components $\prod_t p(y_t | s_t, \mathcal{F}_{t-1})$.

The joint distribution $p(s_{t-1}, s_t)$ is identifiable following a two-step argument.[4] First, the finite mixture

$$\sum_{s_t} p(y_t | s_t, \mathcal{F}_{t-1}) q(s_t),$$

with generic mixing distribution $q(s_t)$ and mixing components $p(y_t | s_t, \mathcal{F}_{t-1})$, are identifiable by *Proposition 1* of Teicher [1963] provided a total ordering for the mixing components. Following proof of *Proposition 1*, we may impose a lexicographical order such that

$$p(y_t | (a_1, a_2)', \mathcal{F}_{t-1}) \quad \prec \quad p(y_t | (b_1, b_2)', \mathcal{F}_{t-1})$$

if and only if $a_1 < b_1$ or $a_1 = b_1$ but $a_2 < b_2$ where $a_i, b_i \in \{0, 1\}$ for $i = 1, 2$ by assuming $\underline{\sigma} \neq \overline{\sigma}$ and/or $x'\underline{\beta} < x'\overline{\beta}$ almost surely X in each equation. Second, the mixture (4.1) is

[4] See Krolzig [1997], *Section 6.2*, for a relevant discussion

identifiable by *Theorem 1* of Teicher [1967] since its mixing component $\prod_{t=1}^{T} p(y_t|s_t, \mathcal{F}_{t-1})$ is a T-fold product of $p(y_t|s_t, \mathcal{F}_{t-1})$. Then by *Definition 12.4.3* of Cappé et al. [2005], the mixing distribution $p(s_1, s_2, \cdots, s_T)$ of (4.1) is identifiable. It follows immediately that for each s_t, mixing component $\prod_t p(y_t|s_t, \mathcal{F}_{t-1})$ is identifiable. By normality of the mixing component, parameters in our measurement equations, (β, P_{uu}), are identifiable.

Our argument may be extended to include Markov switching autoregression of order m with a total ordering of mixing components such that

$$p\left(y_t \middle| \begin{pmatrix} a_{1,1} \\ a_{1,2} \end{pmatrix}, \cdots, \begin{pmatrix} a_{m,1} \\ a_{m,2} \end{pmatrix}, \mathcal{F}_{t-1}\right) \prec p\left(y_t \middle| \begin{pmatrix} b_{1,1} \\ b_{1,2} \end{pmatrix}, \cdots, \begin{pmatrix} b_{m,1} \\ b_{m,2} \end{pmatrix}, \mathcal{F}_{t-1}\right)$$

if and only if $(a_{1,1}, a_{1,2}, \cdots, a_{m,1}, a_{m,2})' \prec (b_{1,1}, b_{1,2}, \cdots, b_{m,1}, b_{m,2})'$.

4.2 Identification of State Transition

In the sequel, we characterize parameters of regime switching, (τ, A, P_{vv}), in terms of joint state distribution $p(s_t, s_{t-1})$. Note the transition of latent factor *per se* is a homogeneous Markov process, thereby the sequence of state variables (s_t) by itself is an exogenous homogeneous Markov chain. Therefore the state transition probability $p(s_t|s_{t-1})$ carries the same information of $p(s_t, s_{t-1})$.

4.2.1 Identification of Thresholds and Factor Correlation

From state distribution $p(s_t)$, we identify threshold parameters τ up to constant multiples, as well as the correlation matrix between latent factors.

Let D_w be the diagonal of Σ_{ww} such that

$$D_w^{1/2} = \begin{pmatrix} \sigma_1 & \\ & \sigma_2 \end{pmatrix},$$

in which σ_1, σ_2 are reparameterizations of (A, P_{vv}) given by

$$\begin{aligned} \sigma_1^2 &= \Big[-\alpha_{12}^3\alpha_{21} + \alpha_{12}^2(1+\alpha_{11}\alpha_{22}) + (1-\alpha_{11}\alpha_{22})(1-\alpha_{22}^2) - \alpha_{12}\alpha_{21}(1+\alpha_{22}^2) \\ &\quad +2\alpha_{12}(\alpha_{11} - \alpha_{22}\det A)\rho_{v_1,v_2}\Big] \Big/ \pi(\lambda_1, \lambda_2), \qquad (4.2) \\ \sigma_2^2 &= \Big[\alpha_{11}^3\alpha_{22} - \alpha_{11}^2(1+\alpha_{12}\alpha_{21}) + (1-\alpha_{12}\alpha_{21})(1+\alpha_{21}^2) - \alpha_{11}\alpha_{22}(1-\alpha_{21}^2) \\ &\quad +2\alpha_{21}(\alpha_{22} - \alpha_{11}\det A)\rho_{v_1,v_2}\Big] \Big/ \pi(\lambda_1, \lambda_2), \qquad (4.3) \end{aligned}$$

with

$$\pi(\lambda_1, \lambda_2) = (1 - \lambda_1\lambda_2)(1 - \lambda_1^2)(1 - \lambda_2^2),$$

where λ_1, λ_2 are eigenvalues of A. Note the stability of A guarantees $\pi(\lambda_1, \lambda_2) \neq 0$. The distribution of $D_w^{-1/2} w_t$ is standard bivariate normal with correlation ρ_{12}, whose density is denoted by $\phi_{\rho_{12}}(z_1, z_2)$. Then the state probability

$$\mathbb{P}(s_t = (0,0)') = \mathbb{P}\left\{D_w^{-1/2} w_t < D_w^{-1/2}\tau\right\}$$

by construction.

Marginalizing $p(s_t)$ along s_2 to have $p(s_{1,t})$ and

$$\frac{\tau_1}{\sigma_1} = \Phi_1^{-1}(\mathbb{P}(s_{1,t} = 0))$$

in which $\Phi_1^{-1}(\cdot)$ is the inverse of standard normal distribution function. Then the normalized thresholds are

$$D_w^{-1/2}\tau = \begin{pmatrix} \Phi_1^{-1}(\mathbb{P}(s_{1,t} = 0)) & \\ & \Phi_1^{-1}(\mathbb{P}(s_{2,t} = 0)) \end{pmatrix}.$$

For each $D_w^{-1/2}\tau$, the function $\Psi(\rho_{12}) = \mathbb{P}\left\{D_w^{-1/2} w_t < D_w^{-1/2}\tau\right\}$ is strictly increasing in ρ_{12} by the fact that

$$\begin{aligned}
\frac{d\Psi(\rho_{12})}{d\rho_{12}} &= \frac{\partial}{\partial \rho_{12}} \int_{-\infty}^{\tau_2/\sigma_2} \int_{-\infty}^{\tau_1/\sigma_1} \phi_{\rho_{12}}(z_1, z_2) dz_1 dz_2 \\
&= \int_{-\infty}^{\tau_2/\sigma_2} \int_{-\infty}^{\tau_1/\sigma_1} \frac{\partial^2}{\partial z_1 \partial z_2} \phi_{\rho_{12}}(z_1, z_2) dz_1 dz_2 \\
&= \phi_{\rho_{12}}(\tau_1/\sigma_1, \tau_2/\sigma_2) > 0.
\end{aligned}$$

We thus have

$$\rho_{12} = \Psi^{-1}\left(\mathbb{P}(s_t = (0,0)')\right).$$

This result alone establishes global identification for the stationary case of CCP: it is directly recovered from $p(s_t, s_{t-1})$ the scaled threshold parameter $\tau\sqrt{1-\alpha^2}$, and the correlation coefficient α for $(w_t, w_{t-1})'$.

4.2.2 Identification of Transition Equation

Extending results in the preceding section, we may identify all parameters in the state transition. We first show the correlation matrix of $(w_{t-1}', w_t')'$ is identified through the joint state distribution $p(s_t, s_{t-1})$. Then we characterize (A, P_{vv}) in terms of the identified correlation matrix.

To begin with, observe $(w_{t-1}', w_t')' \sim \mathbb{N}(0, \Sigma)$ such that

$$\Sigma = \begin{pmatrix} \Sigma_{ww} & \Sigma_{ww}A' \\ A\Sigma_{ww} & \Sigma_{ww} \end{pmatrix}.$$

We may write

$$\Sigma = D_{ww}^{1/2}\Gamma(w_{t-1}, w_t)D_{ww}^{1/2},$$

such that $D_{ww}^{1/2} = \operatorname{diag}\left(D_w^{1/2}, D_w^{1/2}\right)$ and

$$\Gamma(w_{t-1}, w_t) = \begin{pmatrix} \Gamma_0 & \Gamma_1' \\ \Gamma_1 & \Gamma_0 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ \rho_{12} & 1 & & \\ \rho_{13} & \rho_{23} & 1 & \\ \rho_{14} & \rho_{24} & \rho_{34} & 1 \end{pmatrix},$$

denoting the correlation matrix of $(w_{t-1}', w_t')'$ with $\rho_{12} = \rho_{34}$. As is clear, $(D_w^{-1/2}\tau, \rho_{12})$ is identified through $p(s_{t-1})$. With identical arguments, the correlation matrix $\Gamma(w_{t-1}, w_t)$ is identified through $p(s_t, s_{t-1})$. Specifically, we obtain the marginal distributions $p(s_{1,t-1}, s_{1,t})$, $p(s_{1,t-1}, s_{2,t})$, $p(s_{2,t-1}, s_{1,t})$ and $p(s_{2,t-1}, s_{2,t})$ and identify $\rho_{13}, \rho_{14}, \rho_{23}$, and ρ_{24}, respectively.

The correlation matrix $\Gamma(w_{t-1}, w_t)$ provides 5 equations, i.e. $\rho_{12}, \rho_{13}, \rho_{14}, \rho_{23}$ and ρ_{24}, for 5 unknowns $(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}, \rho_{v_1,v_2})$. We solve this system in part easily using the block matrices of Σ. Note that

$$A\Sigma_{ww} = D_w^{1/2}\Gamma_1 D_w^{1/2},$$

with

$$\Sigma_{ww} = D_w^{1/2}\Gamma_0 D_w^{1/2}.$$

Therefore,

$$\begin{aligned}
\alpha_{11} &= \frac{\rho_{13} - \rho_{23}\rho_{12}}{1 - \rho_{12}^2} \\
\alpha_{12}^* &= \frac{\rho_{23} - \rho_{13}\rho_{12}}{1 - \rho_{12}^2}, \quad \alpha_{12}^* = \alpha_{12}\frac{\sigma_2}{\sigma_1} \\
\alpha_{21}^* &= \frac{\rho_{14} - \rho_{24}\rho_{12}}{1 - \rho_{12}^2}, \quad \alpha_{21}^* = \alpha_{21}\frac{\sigma_1}{\sigma_2} \\
\alpha_{22} &= \frac{\rho_{24} - \rho_{14}\rho_{12}}{1 - \rho_{12}^2}
\end{aligned}$$

Note α_{11}, α_{22} and $\alpha_{12}\alpha_{21}$ are identified with the sign of α_{12} and α_{21}, followed by tr$A$ and det A. In particular, α_{12}, α_{21} are directly identified at 0.

We may substitute α_{21}, since the only non-trivial case concerns $\alpha_{12}\alpha_{21} \neq 0$. Hence the remaining parameters $(\alpha_{12}, \rho_{v_1 v_2})$ are characterized by the solution of the bivariate system regarding ρ_{12} and α_{12}^*. Note that $\rho_{12} = \frac{\sigma_{12}}{\sigma_1\sigma_2}$, in which

$$
\begin{aligned}
\sigma_{12} \quad = \quad & \Big[-\alpha_{11}^2\alpha_{12}\alpha_{22} + \alpha_{12}(1+\alpha_{21}^2)\alpha_{22} + \alpha_{11}\alpha_{21}(1+\alpha_{12}^2-\alpha_{22}^2) \\
& + (1-\alpha_{11}^2-\alpha_{22}^2+(\alpha_{11}\alpha_{22})^2-(\alpha_{12}\alpha_{21})^2)\rho_{v_1,v_2}\Big] \Big/ \pi(\lambda_1,\lambda_2).
\end{aligned}
$$

It is easily solved that

$$
\begin{aligned}
\alpha_{12} \quad &= \quad \mathrm{sign}(\alpha_{12})\sqrt{\frac{\left[(\alpha_{12}\alpha_{21})^2-(1-\alpha_{22}^2){\alpha_{12}^*}^2\right](1+\mathrm{tr}A+\det A)+2\alpha_{22}\alpha_{12}^*\rho_{12}\alpha_{12}\alpha_{21}}{(\alpha_{11}^2-1+{\alpha_{12}^*}^2)(1+\mathrm{tr}A+\det A)+2\alpha_{11}\alpha_{12}^*\rho_{12}}}, \\
\rho_{v_1,v_2} \quad &= \quad \Big[\alpha_{12}^2\left((1+\det A)(\alpha_{21}-\alpha_{12}^*)^2+(1-\det A)(1-\alpha_{11}^2)-2\alpha_{12}\alpha_{21}\right) \\
& \qquad -{\alpha_{12}^*}^2\left(1-\alpha_{12}\alpha_{21}-\alpha_{22}(\mathrm{tr}A-\alpha_{22}\det A)\right)\Big] \\
& \qquad \Big/\Big[2\alpha_{12}\Big(\alpha_{11}(1-\alpha_{22}^2){\alpha_{12}^*}^2-\alpha_{22}(1-\alpha_{11}^2-{\alpha_{12}^*}^2)(\alpha_{12}\alpha_{21})-\alpha_{11}(\alpha_{12}\alpha_{21})^2\Big)\Big].
\end{aligned}
$$

Finally, we conclude that τ_1 and τ_2 are identified since σ_1^2, σ_2^2 are specified by (4.2) and (4.3).

4.3 Identification of Endogenous Feedback

The identification of state transition implies the identification of endogenous feedback. By the theory of mixture distribution, the joint density

$$
p(y_1,\cdots,y_T,s_1,\cdots,s_T) = p(y_1,\cdots,y_T|s_1,\cdots,s_T)p(s_1,\cdots,s_T),
$$

is identifiable, followed by

$$
p(s_t|s_{t-1}) = \frac{p(s_t,s_{t-1})}{p(s_{t-1})}
$$

and time-varying transition

$$
p(s_t|s_{t-1},\mathcal{F}_{t-1}) = \frac{p(y_1,\cdots,y_{t-1},s_t,s_{t-1})}{p(y_1,\cdots,y_{t-1},s_{t-1})}.
$$

On the one hand, all parameters except for P_{vu} are identifiable through the unconditional transition probability $p(s_t|s_{t-1})$. On the other hand, suppose there exist P_{vu} and P_{vu}^* such that both induce the same time-varying transition probability $p(s_t|s_{t-1}, \mathcal{F}_{t-1})$, then by (2.2), it must be true that

$$P_{vu}P_{uu}^{-1}u_{t-1} =_{a.s.} P_{vu}^{*}P_{uu}^{-1}u_{t-1},$$

followed by $P_{vu} = P_{vu}^{*}$.

5 Empirical Illustration

In this section, we examine a vector model for U.S. monetary and fiscal policy rules with regime switching and feedback in policy regimes and analyzing the unobserved regime interactions.

Main theoretical works on policy interactions include Sargent and Wallace [1981], Wallace [1981], Aiyagari and Gertler [1985] and Leeper [1991]. The overriding message is that monetary and fiscal policies together stabilizes real government debt and determines price level. There are two main competing theories. The conventional theory describes an active monetary authority, who systematically raise nominal interest rate more than one-for-one in response to current inflation while the fiscal authority adjusts tax and spending passively to maintain solvency. The alternative theory describes an active fiscal authority who spend on its agenda, and a passive monetary authority generates seigniorage to maintain solvency.

Empirical works devote much attention to the dynamics of policy interactions. For example, Favero and Monacelli [2005] consider regime switching policy rules and find little evidence of synchronization in regime-switching; Bianchi and Ilut [2017] embed regime-switching policy rules into a New-Keynesian dynamic stochastic general equilibrium (DSGE) model and explain inflation drop in the 1980s concerning a policy shift. Two works are more closely related to our exercise. Davig and Leeper [2006] study monetary and fiscal policy rules jointly with exogenous Markov regime process. Moreover, Chang and Kwak [2017] examine these rules separately and demonstrate the importance of allowing for endogenous feedback to the latent regime factors, and in which it is argued that latent factor provide a more plausible interpretation of sudden policy shifts because rarely do policymakers choose to shift discretely to a new regime.

The primary advantage of our approach is that we provide a consistent framework for analyzing policy interaction with the presence of feedback because the channels of interaction are explicitly modeled. In what follows, we refer to the direct channel as the factor loading A of latent factors, and the indirect channel as the feedback to regime factors from a joint policy innovation, characterized by correlation matrix P.

5.1 Interactions of U.S. Monetary and Fiscal Policies

We consider regime-switching monetary and fiscal policy rules

$$
\begin{aligned}
i_t &= \alpha_c(s_t^M) + \alpha_\pi(s_t^M)\pi_t + \sigma_M u_t^M && (5.1)\\
\tau_t &= \beta_c(s_t^F) + \beta_b(s_t^F)b_{t-1} + \beta_g(s_t^F)g_t + \sigma_F u_t^F && (5.2)
\end{aligned}
$$

with latent regime factor

$$
\begin{aligned}
w_t &= Aw_{t-1} + v_t\\
s_t^i &= 1\{w_t^i \geq \tau_i\}
\end{aligned}
$$

for $i = M, F$ such that $w_t = (w_t^M, w_t^F)', v_t = (v_t^M, v_t^F)'$. In the policy rules, i_t and π_t represent nominal interest rate and inflation rate at time t; τ_t, b_{t-1} and g_t represent revenues net of transfer payments, government spending and debt held by public. The main model we consider assumes both general stable A and unrestricted correlation matrix P, each characterizing a channel for policy interactions. In what follows, we use "main" and "unrestricted" interchangeably.

We hereby view regression error $u_t = (u_t^M, u_t^F)'$ as fundamental shocks or functions of variables omitted by econometrician. Within-equation feedback thus arises naturally when policy regime responds systematically to information available to policymakers. In such case, size of responses are functions of both historical policy disturbances and exogenous shocks. Also, since both authorities routinely project variables in the information set of the other branch, the cross-equation feedback arises naturally as well.

Following Leeper [1991], regimes for monetary policy and fiscal policy depend upon the parameter values in monetary and fiscal policy rules. The monetary policy is said to be active if $\alpha_\pi > 1$, i.e., the policy rate respond more than one for one to inflation, and passive if $0 \leq \alpha_\pi < 1$. The fiscal policy is said to be active if β_b is less than the real interest rate, and passive if otherwise. We follow Leeper [1991] and call Regime M and Regime F to be "active-passive" and "passive-active" combinations of monetary and fiscal regimes, respectively. Both regimes imply the existence of a determinate bounded rational expectations equilibrium.

Additional identification issues may arise, because, in addition to Regime M/F, our model allows both doubly active and passive regimes. Leeper [1991] argues that doubly active regime leads to the nonexistence of a money-growth process that ensures consumer will hold government debt unless policy innovations are correlated; and doubly passive regime leads to multiple money-growth processes satisfying equilibrium conditions and, hence, in-determinate pricing function. Nonetheless, our model assumptions are different. First, we allow shocks $(u_t^M, u_t^F)'$ to be correlated and assume independence over time, whereas Leeper [1991] assumes uncorrelated innovations across equation, and AR(1) specification for u_t^F. Second, provided a passive fiscal rule, Davig and Leeper [2007] show that monetary policy can satisfy the Taylor principle in the long run, even while deviating from it substantially for brief periods or modestly for prolonged periods. Empirically, we find doubly passives regimes to be all relatively short-lived. We thus argue indeterminacy (hence, identification) is not a major concern for our empirical exercise.

Davig and Leeper [2006] include output gaps and switching volatility in the regression, but can otherwise be considered as a special case of our model since they estimate policy rules with, equivalently,

$$A = \begin{pmatrix} \alpha_{11} & \\ & \alpha_{22} \end{pmatrix}, \quad P = I_4.$$

Chang and Kwak [2017] is also a special case of our model since separate estimation of (5.1) and (5.2) is equivalent to assuming only within-equation feedback, characterized by

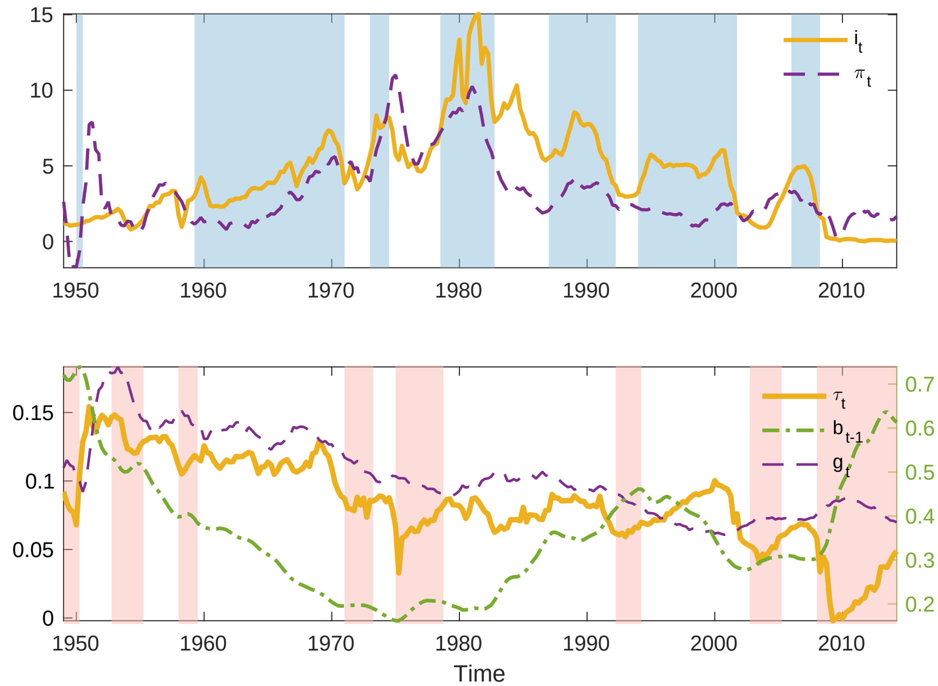
$$A = \begin{pmatrix} \alpha_{11} & \\ & \alpha_{22} \end{pmatrix}, \quad P = \begin{pmatrix} 1 & & & \\ 0 & 1 & & \\ \rho_{v_1 u_1} & 0 & 1 & \\ 0 & \rho_{v_2 u_2} & 0 & 1 \end{pmatrix}.$$

In what follows, we call Chang and Kwak [2017] the "restricted" model.

5.2 Data and Identified Regimes

We use quarterly U.S. data from 1949:2 to 2014:2 for the empirical exercise. Figure 1 plots these time-series for policy instruments, on which we superimpose the Regime M and Regime F identified using extracted latent factors $\mathbb{E}(w_t|\mathcal{F}_t)$ and estimated threshold τ from the main model. The regimes are primarily M or F over the course of history with a brief doubly active period in the 1980s and multiple brief doubly passive epochs in 1950s and 2000s.

For monetary policy (5.1), we define π_t to be inflation rate over contemporaneous and prior three quarters as in Taylor [1993] and obtain inflation each period as log difference of GDP deflator. We use three-month Treasury bill (T-bill) rate in the secondary market for nominal interest rate i_t. We choose T-bill rate over federal funds rate (FFR) mainly due to its short series. On the other hand, T-bill rate is highly correlated with FFR with sample correlation 0.988 over the period 1954:1-2014:2, and is available since 1949:1. Using



Note: Blue indicates estimated Regime M, and Melon indicates estimated Regime F. The series for b_{t-1} corresponds to right vertical axis in the lower panel. All other series correspond to left vertical axis.

Figure 1: Historical Data and Estimated Policy Regimes

T-bill rate allows us to study meaningful regime changes in monetary and fiscal policy rules before 1954 which include critical historical episodes such as Treasury Accord of March 1951 leading to passive monetary policy and the wartime fiscal financing for Korean war leading to active fiscal policy.

For fiscal policy (5.2), all variables are for the federal government only. τ_t is federal tax receipts net of total federal transfer payments to GDP ratio, and b_{t-1} is the market value of gross marketable federal debt held by the public to GDP ratio and g_t is federal government consumption plus investment expenditures to GDP ratio. Finally, we use average debt-output ratio over previous four quarters as a measure of b_{t-1}.

5.3 Estimation and Results

We perform ML estimation for a sequel of specifications with a detailed description of optimization routine in Appendix B.2 and results in Table 1. Standard errors are reported for only the major cases, and are obtained by simulation since the hessian of log-likelihood is difficult to compute. The regime-switching coefficients for both policies appear to be identified in all specifications. In the unrestricted model, the monetary rule responds actively to inflation with coefficient 1.049(0.147) and passively with coefficient 0.640(0.044), while the fiscal rule responds actively to debt with coefficient -0.029(0.004) and passively with coefficient 0.050(0.011).

In estimation, we propose an error component parametrization for $(u_t', v_{t+1}')'$ with exactly 6 parameters to identify structure of shocks and give proper correlation matrix P. Let

$$u_{it} = \lambda_i \xi_t + \sqrt{1-\lambda_i^2}\zeta_{it} \tag{5.3}$$

$$v_{i,t+1} = \phi_i \xi_t + \psi_i \zeta_{it} + \sqrt{1-\phi_i^2-\psi_i^2}\epsilon_{it} \tag{5.4}$$

for $i = 1, 2$ such that $(\xi_t), (\zeta_{it})$ and (ϵ_{it}) are independent standard normal, and $-1 < \lambda_i, \phi_i, \psi_i < 1$. We further assume $\lambda_1 \geq 0$ and $\phi_i^2 + \psi_i^2 < 1$ for each $i = 1, 2$. The former restriction is necessary since (ξ_t) paired with (λ_i, ϕ_i) gives identical transition as by $(-\xi_t)$ paired with $(-\lambda_i, -\phi_i)$. Channels for endogenous feedback are characterized by λ_i, ϕ_i and ψ_i for $i = 1$ and 2. The within-equation feedback of equation i is determined by $\lambda_i\phi_i + \psi_i\sqrt{1-\lambda_i^2}$ for $i = 1$ and 2, whereas the cross-equation feedback is characterized by products $\lambda_1\phi_2$ and $\lambda_2\phi_1$.

We report two main findings from ML estimation in addition to the improved efficiency of the unrestricted model compared to the restricted one. On the one hand, the likelihood ratio test gives evidence that restricted model is misspecified, and the estimates suggest that the misspecifications lie in both channels of policy interaction we consider. Data favors the direct channel of interaction, characterized by off-diagonal entries of A. Notably, the impact of fiscal factor on monetary factor, α_{12}, is twice in size compared to the effect in the opposite direction, α_{21}. Data also favors a non-trivial positive correlation between fiscal and monetary information, ρ_{u_1,u_2}, and positive cross-equation feedback from fiscal information to monetary regime factor, ρ_{u_2,v_1}, but not the opposite direction of cross-equation feedback, ρ_{u_1,v_2}. Both channels are essential in the sense that severing one leads to non-rejection in likelihood ratio test (see (3) in Table 1). On the other hand, however, coefficient estimates for the regime-switching policy rules are similar across all specifications considered.

Granger causality test is easily constructed for the direct channel of factor interaction. The fiscal (monetary) factor Granger causes shifts in monetary (fiscal) factor if $\alpha_{12} \neq 0$ ($\alpha_{21} \neq 0$). Since the sampling distribution for our model is unclear, we formally test the null hypothesis $\alpha_{12} = 0$ ($\alpha_{21} = 0$) with a bootstrap test. Data, however, generates

Parameter	(1)-Unres.	S.E.	(2) -Res.	S.E.	(3)	(4)
			Thresholds			
τ_m	0.435	1.705	-0.389	1.648	0.802	0.533
τ_f	-0.582	1.407	-0.600	1.551	-0.026	-0.462
			Transition of Latent Factor			
α_{11}	0.956	0.248	0.984	0.420	0.975	0.961
α_{21}	0.023	0.161	-	-	-	0.012
α_{12}	0.056	0.249	-	-	-	0.050
α_{22}	0.938	0.177	0.968	0.283	0.961	0.953
			Endogenous Feedback			
$\rho_{u_1u_2}$	0.178	0.066	-	-	0.180	0.146
$\rho_{u_1v_1}$	0.997	0.418	0.999	0.470	0.943	0.982
$\rho_{u_2v_1}$	0.165	0.206	-	-	0.322	-
$\rho_{u_1v_2}$	0.000	0.243	-	-	0.225	0.136
$\rho_{u_2v_2}$	0.970	0.242	0.999	0.368	0.988	0.996
$\rho_{v_1v_2}$	0.000	0.245	-	-	0.403	-
			Regime-Switching Monetary Policy			
$\alpha_c(s_m = 0)$	0.533	0.265	0.443	0.529	0.456	0.528
$\alpha_c(s_m = 1)$	2.524	0.477	2.601	0.607	2.531	2.564
$\alpha_\pi(s_m = 0)$	0.640	0.044	0.661	0.095	0.658	0.647
$\alpha_\pi(s_m = 1)$	1.049	0.147	1.039	0.183	1.049	1.047
σ_m	1.310	0.064	1.306	0.068	1.309	1.309
			Regime-Switching Fiscal Policy			
$\beta_c(s_f = 0)$	-0.028	0.001	-0.028	0.003	-0.028	-0.028
$\beta_c(s_f = 1)$	0.011	0.003	0.014	0.011	0.011	0.012
$\beta_b(s_f = 0)$	-0.029	0.004	-0.033	0.009	-0.028	-0.029
$\beta_b(s_f = 1)$	0.050	0.011	0.052	0.020	0.054	0.051
$\beta_g(s_f = 0)$	1.016	0.024	1.020	0.051	1.017	1.018
$\beta_g(s_f = 1)$	0.644	0.056	0.603	0.107	0.634	0.637
σ_f	0.014	0.001	0.014	0.001	0.014	0.014
log-likelihood	275.137		270.409		273.332	274.629
p-value (vs (2))	0.051*		-		0.211	0.077

Note:* indicates $df = 4$ in the likelihood ratio test against restricted (Res.) model. Missing values are 0 in the shaded rows, and undefined if otherwise. Cyan highlights key model differences in transition and feedback. Rose highlights identified regimes.

Table 1: Maximum Likelihood Estimates (1949:Q2-2014:Q2)

insufficient evidence for interaction between policy factors (see Table 2).

Interestingly, the estimated main model suggests that fiscal information leads monetary regime determination. Nonetheless, data is ambiguous about the leading role of monetary and fiscal information since the monetary leading model also fits significantly better than the restricted model (see (4) in Table 1). We formalize the statement by testing null hypothesis that $\rho_{u_2v_1} = \rho_{v_1v_2} = 0$ and monetary information leads using a bootstrap test. The test result is reported in Table 2.

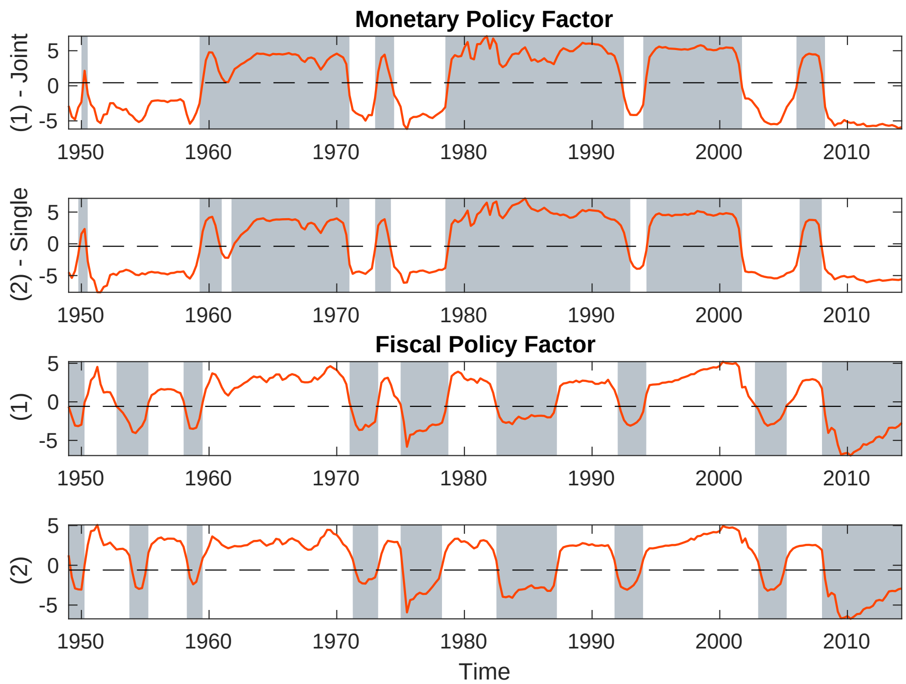
Parameter	Leading FP	90% Band		Leading MP	90% Band	
α_{11}	0.956	0.230	0.987	0.961	0.031	0.993
α_{21}	0.023	-0.087	0.085	0.012	-0.066	0.134
α_{12}	0.056	-0.008	0.158	0.050	-0.029	0.209
α_{22}	0.938	0.814	0.981	0.952	0.812	0.992
$\rho_{u_1u_2}$	0.178	0.069	0.288	0.146	0.025	0.253
$\rho_{u_1v_1}$	0.997	-0.186	1.000	0.983	-0.305	0.997
$\rho_{u_2v_1}$	0.165	-0.205	0.519	-	-0.554	0.751
$\rho_{u_1v_2}$	-	-0.393	0.465	0.137	-0.035	0.470
$\rho_{u_2v_2}$	0.970	0.390	0.996	0.996	0.488	1.000
$\rho_{v_1v_2}$	-	-0.455	0.421	-	-0.562	0.747

Note: The Granger causality tests are constructed using Cyan. All missing values are 0 unless noted otherwise. And Rose is for bootstrap test with $H_0 : \rho_{u_2v_1} = \rho_{v_1v_2} = 0$ and $\rho_{u_1v_2} = 0.137$.

Table 2: Leading FP vs. Leading MP

We also report the extracted latent factors for different specifications. Following Chang and Kwak [2017], the monetary regime is defined to be active if $w_t^M \geq \tau_m$, whereas the fiscal regime is active if $w_t^F < \tau_f$, and the inferred regimes are determined by their estimates and extracted conditional expectation. Figure 2 presents the extracted latent dynamic factors with inferred active regimes indicated by the shaded area. The result suggests implied regime strength are different between specification (1) and (2). However, nonetheless, the implied regimes are similar in timing and length.

Without a suitable framework, Chang and Kwak [2017] use the extracted latent factors to perform a second-stage inference on policy interactions by fitting them to a time-varying coefficient vector autoregression (TVC-VAR). Their approach, however, is inconsistent with the notion of interaction because none is actually accounted for. Figure 3 reports the impulse responses of regime factors corresponds to MLE and second-stage inference, respectively. There is notable difference between MLE implied impulse responses and estimated response



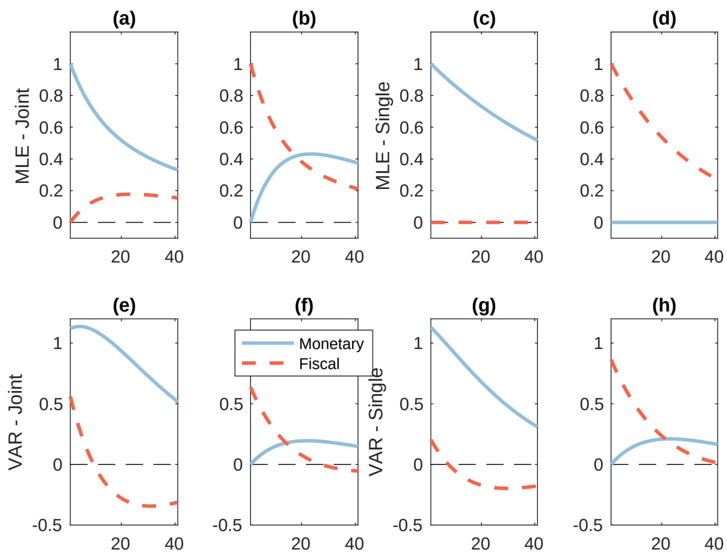
Note: The extracted factor (solid line) $\mathbb{E}(w_t|\mathcal{F}_t)$ and threshold (dashed line) τ defines the active regime (shade). In the panels of monetary factor, the regime is identified as active if $\mathbb{E}(w_t^M|\mathcal{F}_t) \geq \tau_M$. In the panels of fiscal factor, active regime is identified if $\mathbb{E}(w_t^F|\mathcal{F}_t) < \tau_F$.

Figure 2: Extracted Latent Factor (MLE)

from latent factors in the restricted model suggesting such practice is more appropriate in our model since the interaction is reflected in the extracted latent factor.

The correlation of extracted factors is 0.43 for the restricted model, whereas it is 0.78 for the unrestricted model (see Table 7). Figure 4 examines the pattern of comovements for extracted factor by comparing the 6-year rolling window correlation. This result shows the regime factors move more closely in the unrestricted model than in the restricted model. In terms of historical events, the restricted model suggests non-cooperative regime factors in the 1950s and 1980s. In contrast, the unrestricted model presents evidence of non-trivial cooperation through out the whole sample period with several exceptions. Nonetheless, It must be noted such cooperation does not guarantee synchronization in the policy regimes.

We also characterize correlations of data and factors in frequency domain. Figure 5 compares the coherence of policy rates and those implied by filtered latent factors in frequency domain for different models, from which we summarize two results. First, our unrestricted



Note: Panels (a) - (d) plot impulse responses of factors implied by MLE. And panels (e) - (h) plot estimated impulse responses implied by extracted latent factors. Panels (a), (c), (e) and (g) assume 1 standard deviation shock to monetary factor. The rest assume 1 standard deviation shock to fiscal factor.

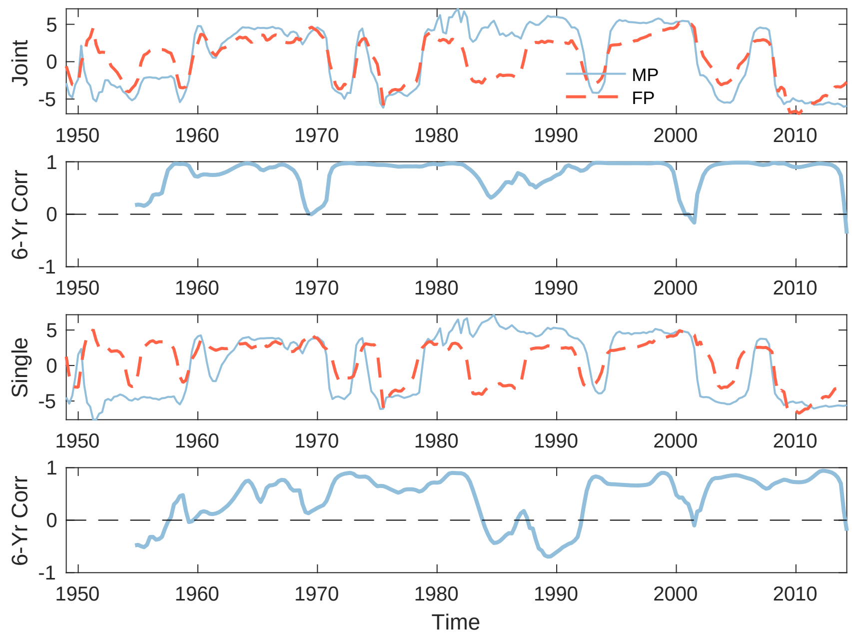
Figure 3: Impulse Responses of Factors

model implies strong coherence between factors in business cycle frequency. In contrast, policy rates display strong correlation at lower frequencies while restricted model attribute rather weak correlation across the entire spectrum. Second, combining with parameter es-

Parameter	Unres.	S.E.	Res.	S.E.
λ_1	0.744	0.234	-	-
λ_2	0.239	0.190	-	-
ϕ_1	0.691	0.469	-	-
ϕ_2	0.000	0.368	-	-
ψ_1	0.723	0.462	0.999	0.470
ψ_2	0.999	0.283	0.999	0.368

Table 3: Selected Parameter Estimates for Error Component Model (1949:Q2-2014:Q2)

timates in Table 2, there is evidence of direct but weak interaction between factors. In addition, our unrestricted model identifies shocks that propagate very differently to monetary and fiscal factors. It is tempted to regard $(\lambda_i, \phi_i, \psi_i)$ in Equation 5.3 and 5.4 as relative



Note: The panels demonstrate the comovement between extracted monetary factor (blue solid) and fiscal factor (red dashed). "6-Yr Corr" reports rolling window correlation between extracted latent factors with window size 24. The upper panels pertain to the unrestricted model, and the lower panels to the restricted model.

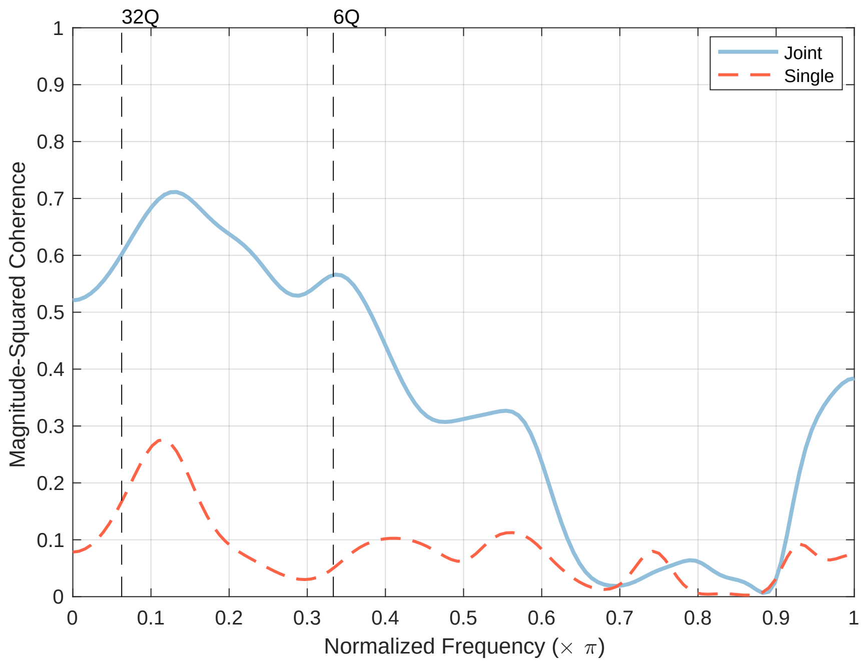
Figure 4: Comovements and Correlation of Extracted Latent Factors

importance of components identified in shocks. The common factor of shocks accounts for a large proportion in monetary rule and is much weaker in fiscal rule (Table 3). It is then largely propagated to monetary factor in next period, whereas its impact to fiscal factor is negligible. Fiscal factor innovation, on the other hand, is mainly driven by last period equation-specific factor to fiscal rule.

6 Simulation

6.1 Simultaneity Bias Reduction

Simultaneity bias is well expected if the conventional approach is applied when the data generating process (DGP) entails interacting latent factors, correlated error terms or cross-equation feedback. In this section, we draw a comparison between the unrestricted model and the restricted model by simulation. Using this result, we supplement the empirical



Note: Magnitude-squared coherence at frequency λ is defined to be $|f_{xy}(\lambda)|^2/[f_{xx}(\lambda)f_{yy}(\lambda)]$, with $f_{xy}(\lambda)$ denoting the cross-spectral-density, and estimated using Welch's method with window size 24. The dashed vertical lines indicate the normalized frequencies associated with 6Q and 32Q.

Figure 5: Magnitude-Squared Coherence of Extracted Latent Factors

exercise in the preceding section.

We simulate data from our empirical model (5.1) and (5.2) at its ML estimates. And as before, the correlation matrix P is parameterized by error component model (5.3) and (5.4). The exogenous variables are fixed to be the full sample data with sample size is 262. The number of iteration is set to be 1,000.

Table 4 reports the simulation results. We measure bias reduction with the restricted model as benchmark using a simple metric

$$\%\text{Bias Reduction} = |\%\text{Bias}_{Res.}| - |\%\text{Bias}_{Unres.}|.$$

Column 5 of Table 4 shows that the estimates in the restricted model suffer heavy simultaneity bias when ignoring the off-diagonal elements of A and the non-trivial cross-equation endogenous feedback, and the correlation in the shocks. Individually, ignoring cross-equation interaction inflicts (i) sizable bias for threshold values, (ii) substantial bias in the size of

Parameter	DGP	% Bias		%Bias Reduction
		(1)-Unres.	(2)-Res.	
		Thresholds		
τ_m	0.435	-28.627	74.016	45.389
τ_f	-0.582	-1.589	-133.315	131.726
		Transition of Latent Factor		
α_{11}	0.956	-7.210	-15.945	8.734
α_{21}	0.023	6.138	-100.000	93.862
α_{12}	0.056	79.508	-100.000	20.492
α_{22}	0.938	-2.641	-6.339	3.699
		Endogenous Feedback		
$\rho_{u_1 u_2}$	0.178	-0.504	-100.000	99.496
$\rho_{u_1 v_1}$	0.997	-27.336	-28.332	0.996
$\rho_{u_2 v_1}$	0.165	-5.163	-100.000	94.837
$\rho_{u_1 v_2}$	0.000	-	-	-
$\rho_{u_2 v_2}$	0.970	-14.159	-19.089	4.930
$\rho_{v_1 v_2}$	0.000	-	-	-
		Regime-Switching Monetary Policy		
$\alpha_c(s_m = 0)$	0.533	11.008	25.428	14.419
$\alpha_c(s_m = 1)$	2.524	-2.889	-6.482	3.593
$\alpha_\pi(s_m = 0)$	0.640	0.401	1.238	0.837
$\alpha_\pi(s_m = 1)$	1.049	-1.906	-3.524	1.618
σ_m	1.310	-1.227	-1.159	-0.068
		Regime-Switching Fiscal Policy		
$\beta_c(s_f = 0)$	-0.028	-0.188	-0.107	-0.080
$\beta_c(s_f = 1)$	0.011	0.048	-1.540	1.492
$\beta_b(s_f = 0)$	-0.029	-0.443	1.516	1.073
$\beta_b(s_f = 1)$	0.050	-0.284	-0.938	0.654
$\beta_g(s_f = 0)$	1.016	0.382	1.233	0.852
$\beta_g(s_f = 1)$	0.644	-1.062	-2.581	1.519
σ_f	0.014	-1.213	-1.238	0.026

Note: We suppress results for ρ_{u_1,v_2} and ρ_{v_1,v_2} for their relative biases are not defined.

Table 4: Relative Bias against DGP at Joint MLE

interaction and feedback channels, (iii) substantial bias in the correlation of shocks, but lead to (iv) relatively small bias for the coefficients in the measurement equation.

The implication is twofold. First, one should consider unrestricted model whenever possible since, otherwise, the channels for interaction may be distorted. In consequence, making an inference of interaction by examining latent factors from individual rules is inconsistent with the notion of interactions. Second, estimating a single equation model may be worthwhile because the parameters in the measurement equation are not very sensitive to the dynamics of latent factors and it is much less computationally intensive than estimating the unrestricted model. An efficient strategy is to learn from data incrementally by starting from estimating the model equation by equation, followed by a joint estimation with the first-stage ML estimates as the initial guess.

7 Extension

8 Conclusion

We have shown the regime-switching model driven by latent VAR(1) factors is a powerful tool. We allow rich temporal dynamics of factors and both within-equation and between-equation feedback, so a shock to one observed time series or factor may transmit to both regime factors. Our model provides a proper treatment for empirical works in which we explicitly account for channels of regime interactions. In our exercises, we find evidence suggests the presence of regime factor coordination and cross-equation feedback in U.S. monetary and fiscal policy regimes. Moreover, our simulation makes it clear that neglecting simultaneity in regime-switching incurs substantial bias for parameters of regime switching, but relatively small bias for parameters in the measurement equation.

References

S. R. Aiyagari and M. Gertler. The backing of government bonds and monetarism. *Journal of Monetary Economics*, 16(1):19–44, 1985.

M. Bazzi, F. Blasques, S. J. Koopman, and A. Lucas. Time varying transition probabilities for Markov regime switching models. Tinbergen Institute Discussion Paper TI 2014-072/III, 2014.

F. Bianchi and C. Ilut. Monetary/fiscal policy mix and agent's beliefs. *Review of Economic Dynamics*, 26:113–139, 2017.

O. Cappé, E. Moulines, and T. Rydén. *Inference in Hidden Markov Models*. Spinger Science+Business Media, Inc., 2005.

Y. Chang and B. Kwak. U. S. monetary-fiscal regime change in the presence of endogeneous feedback in policy rules. Working Paper, 2017.

Y. Chang, Y. Choi, and J. Y. Park. A new approach to model regime switching. *Journal of Econometrics*, 196:127–143, 2017.

S. Chib. Calculating posterior distributions and modal estimates in Markov mixture models. *Journal of Econometrics*, 75:79–97, 1996.

S. Chib and M. Dueker. Non-Markovian regime switching with endogenous states and time-varying state strengths. Working Paper 2008-31, Federal Reserve Bank of St. Louis, 2004.

J. S. Cho and H. White. Testing for regime switching. *Econometrica*, 75:1671–1720, 2007.

T. Davig and E. M. Leeper. Fluctuating Macro Policies and the Fiscal Theory. *NBER Macroeconomics Annual 2006*, 21:247–316, 2006.

T. Davig and E. M. Leeper. Generalizing the Taylor Principle. *American Economic Review*, 97(3):607–635, 2007.

F. Diebold, J.-H. Lee, and G. Weinbach. Regime switching with time-varying transition probabilities. In C. Hargreaves, editor, *Nonstationary Time Series Analysis and Cointegration*, pages 283–302. Oxford University Press, Oxford, UK, 1994.

C. A. Favero and T. Monacelli. Fiscal Policy Rules and Regime (In)Stability: Evidence from the U.S. *IGIER Working Paper*, (No. 282.), 2005.

R. Garcia. Asymptotic null distribution of the likelihood ratio test in Markov switching models. *International Economic Review*, 39:763–788, 1998.

A. Genz. Numerical computation of multivariate normal probabilities. *Journal of Computational and Graphical Statistics*, 1(2):141–149, 1992. doi: 10.1080/10618600.1992.10477010. URL http://amstat.tandfonline.com/doi/abs/10.1080/10618600.1992.10477010.

J. Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57:357–384, 1989.

J. Hamilton. Specification testing in Markov-switching time-series models. *Journal of Econometrics*, 70:127–157, 1996.

B. E. Hansen. The likelihood ratio test under non-standard conditions. *Journal of Applied Econometrics*, 7:S61–82, 1992.

L. Kalliovirta, M. Meitz, and P. Saikkonen. A gaussian mixture autoregressive model. *Journal of Time Series Analysis*, 36:247–266, 2015.

K. H. Kang. State-space models with endogenous Markov regime switching parameters. *Econometrics Journal*, 17:56–82, 2014.

C.-J. Kim. Dynamic linear models with Markov-switching. *Journal of Econometrics*, 60: 1–22, 1994.

C.-J. Kim. Markov-switching models with endogenous explanatory variables. *Journal of Econometrics*, 122:127–136, 2004.

C.-J. Kim. Markov-switching models with endogenous explanatory variables II: A two-step MLE procedure. *Journal of Econometrics*, 148:46–55, 2009.

C.-J. Kim and C. Nelson. *State-Space Models with Regime Switching*. MIT Press, Cambridge, MA, 1999.

C.-J. Kim, J. Piger, and R. Startz. Estimation of Markov regime-switching regression models with endogeneous switching. *Journal of Econometrics*, 143:263–273, 2008.

H.-M. Krolzig. *Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis*. Springer, 1997.

E. M. Leeper. Equilibria under ‘Active’ and ‘Passive’ Monetary and Fiscal Policies. *Journal of Monetary Economics*, 27(1):129–147, February 1991.

T. J. Sargent and N. Wallace. Some unpleasant monetarist arithmetic. *Quarterly Review*, (Fall):v. 5, no. 3, 1981.

J. B. Taylor. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, 39:195 – 214, 1993.

H. Teicher. Identifiability of Finite Mixtures. *The Annals of Mathematical Statistics*, 34(4): 1265–1269, 1963.

H. Teicher. Identifiability of Mixture of Product Measures. *The Annals of Mathematical Statistics*, 38:1300–1302, 1967.

A. Timmermann. Moments of Markov switching models. *Journal of Econometrics*, 96: 75–111, 2000.

N. Wallace. A modigliani-miller theorem for open-market operations. *The American Economic Review*, 71(3):267–274, 1981.

Appendices

A A Parallelizable Algorithm for Likelihood Evaluation

With 2 regimes for each equation, the computational complexity for Equation (3.1) is $\mathcal{O}(4^2T)$. However, practical issue arises from the calculation of (2.3). Its time-varying nature entails heavy and repeated evaluations of 4-dimensional multivariate normal distribution. This motivates a simpler and parallelizable characterization of the log-likelihood function. Here, we devise a parallelizable algorithm of complexity $\mathcal{O}(4^2T)$, in which calculations of conditional density $p(y_t|\mathcal{F}_{t-1})$ for each t can be distributed to separate CPUs.

Note the likelihood function can be written in form

$$\ell(\theta) \quad = \quad \prod_{t=1}^{T}\left(\sum_{s_t,s_{t-1}} p(y_t|s_t,s_{t-1},\mathcal{F}_{t-1})p(s_t,s_{t-1}|\mathcal{F}_{t-1})\right) \tag{A.1}$$

with $p(y_t|s_t,s_{t-1},\mathcal{F}_{t-1}) = p(y_t|s_t,\mathcal{F}_{t-1})$ by construction. From Equation (A.1),

Algorithm A.1 (Parallelizable Modified Markov-Switching Filter)**.**

1. ***Initialization.*** *For $t=0$, set $p(s_t,s_{t-1})$ to be the unconditional state probabilities such that*

$$\mathbb{P}\{s_0 = (0,0)', s_{-1} = (0,0)'\} = \mathbb{P}\big\{w_{1,0} < \tau_1, w_{2,0} < \tau_2, w_{1,-1} < \tau_1, w_{2,-1} < \tau_2\big\}$$

in which

$$(w_0', w_{-1}')' \sim \mathbb{N}\left(0, \begin{pmatrix} \Sigma_{ww} & \Sigma_{ww}A' \\ A\Sigma_{ww} & \Sigma_{ww} \end{pmatrix}\right)$$

2. ***Parallelization.*** *For each $t \geq 0$, calculate $p(s_t,s_{t-1}|\mathcal{F}_{t-1})$ using (2.3) for each possible realization of (s_t,s_{t-1}), and evaluate $p(y_t|s_t,\mathcal{F}_{t-1})$ for each possible realization of s_t. Then*

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

3. ***Combination.*** *Combining conditional density for each $t \geq 0$ to have the log-likelihood*

$$\log \ell(\theta) = \sum_{t=1}^{T} \log\left(\sum_{s_t,s_{t-1}} p(y_t|s_t,\mathcal{F}_{t-1})p(s_t,s_{t-1}|\mathcal{F}_{t-1})\right)$$

B Computation

B.1 Computation of Latent Factor

The computation of latent factor w_t is costly by integration techniques such as Newton-Cotes method (e.g. Riemann sum) due to curse of dimensionality. A standard alternative is Monte Carlo integration by importance sampling. The extracted latent factor at time t can be written as

$$\mathbb{E}(w_t|\mathcal{F}_t) \quad = \quad \int_{\mathbb{R}^2} w_t \frac{p(w_t|\mathcal{F}_t)}{q(w_t)} q(w_t) dw_t$$

where $q(w_t)$ is any proposal such that $p(w_t|\mathcal{F}_t) > 0$ implies $q(w_t) > 0$. An obvious candidate for $q(w_t)$ is the stationary distribution of w_t. The self-normalized importance sampling estimator of $\mathbb{E}(w_t|\mathcal{F}_t)$ is thus given by

$$\widehat{\mathbb{E}}(w_t|\mathcal{F}_t) \quad = \quad \frac{\sum_{i=1}^n w_t^i \omega_t^i}{\sum_{i=1}^n \omega_t^i}$$

where (w_t^i) is a sequence of draws generated from the proposal $q(w_t)$, and $\omega_t^i = p(w_t^i|\mathcal{F}_t)/q(w_t^i)$ is the importance weight associated with the i-th draw. It is obvious that the $p(w_t|\mathcal{F}_t)/q(w_t)$ is a proper density with respect to the measure $q(w_t)dw_t$. We use self-normalized estimator to control the finite sample behavior. The conditions above are sufficient to imply $\widehat{E}(w_t|\mathcal{F}_t) \rightarrow_{a.s.} E(w_t|\mathcal{F}_t)$. Central limit theorem holds for $\widehat{\mathbb{E}}(w_t|\mathcal{F}_t)$. We thus can measure the approximation precision of $\widehat{\mathbb{E}}(w_t|\mathcal{F}_t)$ by constructing a confidence interval around it using approximated variance

$$\widehat{\text{var}}(\widehat{\mathbb{E}}(w_t|\mathcal{F}_t)) = \frac{1}{n} \frac{\frac{1}{n}\sum_{i=1}^n {\omega_t^i}^2 (w_t^i - \widehat{\mathbb{E}}(w_t|\mathcal{F}_t)^2)}{\left(\frac{1}{n}\sum_{i=1}^n \omega_t^i\right)^2}.$$

Note this confidence interval is different in concept to the confidence interval for parameter inference. It merely provides a measurement of approximation precision.

B.2 Optimization

In ML estimation, we propose the use of pattern search optimization. There are three arguments that favors pattern search over derivative-based methods (who are usually more efficient).

1. The numerical likelihood surface is expected to be rough. This roughness renders derivative-based optimizer less robust. In contrast, patterns search is documented to have more robustness since it is derivative-free.

2. We can easily impose nonlinear constraints such as $\|A\| < 1$ in pattern search. This is not so easy in other methods.

3. Pattern search is a global method as the likelihood surface is known to have multiple local maxima.

Our exercises as of now suggests that pattern search can robustly reproduce estimates in Chang and Kwak [2017]. It also appears that pattern search is more viable than other methods we considered in joint estimation. For instance, our exercise shows joint estimation with pattern search yields estimates close to those of the single equation models, whereas derivative-based methods failed to find any local optimum.

We also proposed a strategy to reduce the chance of being stranded at a local minimum. Specifically, pattern search is called recursively with initial guess being the last stage local minimum $\hat{\theta}_n$ and stopping criteria that sup-norm $\|\hat{\theta}_n - \hat{\theta}_{n+1}\| < \epsilon$ for some prescribed ϵ. In our exercises, we set $\epsilon = 10^{-5}$.

C Omitted Proofs

C.1 Extraction of Latent Factor

Apply Bayes formula to $p(w_t|\mathcal{F}_t)$ and have

$$
\begin{aligned}
p(w_t|\mathcal{F}_t) &= p(w_t|y_t, \mathcal{F}_{t-1}) \\
&= \frac{p(y_t|s_t(w_t), \mathcal{F}_{t-1}) \sum_{s_{t-1}} p(w_t, s_{t-1}|\mathcal{F}_{t-1})}{p(y_t|\mathcal{F}_{t-1})} \\
&= \sum_{s_{t-1}} \frac{p(y_t|s_t(w_t), \mathcal{F}_{t-1}) p(w_t|s_{t-1}, \mathcal{F}_{t-1}) p(s_{t-1}|\mathcal{F}_{t-1})}{p(y_t|\mathcal{F}_{t-1})}.
\end{aligned}
$$

It thus amounts to specify conditional density function $p(w_t|s_{t-1}, \mathcal{F}_{t-1})$. By the decomposition of v_t,

$$
\begin{aligned}
p(w_t|s_{t-1} = (0,0), \mathcal{F}_{t-1}) &= p(w_t|s_{t-1} = (0,0), u_{t-1}) \\
&= \left[\int_{-\infty}^{\tau} \phi_{v|u}(w_t - P_{vu}P_{uu}^{-1}u_{t-1} - Aw_{t-1})\phi(w_{t-1})dw_{t-1}\right] \Big/ \Phi(\tau) \\
&= \Big[\int_{-\infty}^{\tau} (2\pi)^{-1}(\det P_{vv\cdot u})^{-1/2} \\
&\quad \times \exp\Big(-\frac{1}{2}(\underbrace{w_t - P_{vu}P_{uu}^{-1}u_{t-1}}_{w_{u,t}} - Aw_{t-1})' P_{vv\cdot u}^{-1}(w_t - P_{vu}P_{uu}^{-1}u_{t-1} - Aw_{t-1})\Big) \\
&\quad \times (2\pi)^{-1}(\det \Sigma_{ww})^{-1/2} \exp\left(-\frac{1}{2}w_{t-1}'\Sigma_{ww}^{-1}w_{t-1}\right) dw_{t-1}\Big] \Big/ \Phi(\tau) \\
&= (2\pi)^{-2}(\det P_{vv\cdot u}\Sigma_{ww})^{-1/2} \\
&\quad \times \Big[\int_{-\infty}^{\tau} \exp\Big(-\frac{1}{2}(w_{t-1} - QA'P_{vv\cdot u}^{-1}w_{u,t})'Q^{-1}(w_{t-1} - QA'P_{vv\cdot u}^{-1}w_{u,t})\Big) \\
&\quad \exp\left(-\frac{1}{2}\left(w_{u,t}'P_{vv\cdot u}^{-1}w_{u,t} - w_{u,t}'P_{vv\cdot u}^{-1}AQA'P_{vv\cdot u}^{-1}w_{u,t}\right)\right) dw_{t-1}\Big] \Big/ \Phi(\tau) \\
&= \left(\frac{\det \Omega^{-1}Q^{-1}}{\det P_{vv\cdot u}\Sigma_{ww}}\right)^{1/2} \phi(w_t; P_{vu}P_{uu}^{-1}u_{t-1}, \Omega) \\
&\quad \times \Big[\int_{-\infty}^{\tau} \phi_Q(w_{t-1} - QA'P_{vv\cdot u}^{-1}w_{u,t})dw_{t-1}\Big] \Big/ \Phi(\tau) \\
&= \left(\frac{\det \Omega^{-1}Q^{-1}}{\det P_{vv\cdot u}\Sigma_{ww}}\right)^{1/2} \left[\frac{\Phi_Q(\tau - QA'P_{vv\cdot u}^{-1}w_{u,t})}{\Phi(\tau)}\right] \\
&\quad \times \phi(w_t; P_{vu}P_{uu}^{-1}u_{t-1}, \Omega)
\end{aligned} \tag{C.1}
$$

where

$$
\begin{aligned}
Q &= (A' P_{vv\cdot u}^{-1} A + \Sigma_{ww}^{-1})^{-1} \\
\Omega &= (P_{vv\cdot u}^{-1} - P_{vv\cdot u}^{-1} A Q A' P_{vv\cdot u}^{-1})^{-1}
\end{aligned}
$$

and Φ is the unconditional distribution function for (w_t), whereas ϕ_Q (Φ_Q) denotes the zero-mean bivariate normal density (distribution) function with covariance matrix Q. Similar results are easily obtained for cases in which $s_{t-1} = (0,1),(1,0)$ and $(1,1)$.

D Additional Tables and Graphs

D.1 ML Estimates of Error Component

The error component form estimates of different specifications are reported in Table 5. And we report quasi-Bayesian MLE results (Table 6) as a mean to test sensitivity.

Parameter	(1) - Unres. .	(2) - Res.	(3)	(4)
τ_m	0.435	-0.389	0.802	0.533
τ_f	-0.582	-0.600	-0.026	-0.462
λ_1	0.744	-	0.364	0.182
λ_2	0.239	-	0.494	0.803
ϕ_1	0.691	-	0.652	0.000
ϕ_2	0.000	-	0.618	0.747
ψ_1	0.723	0.999	0.758	0.999
ψ_2	0.999	0.999	0.785	0.665
α_{11}	0.956	0.984	0.975	0.961
α_{21}	0.023	-	-	0.012
α_{12}	0.056	-	-	0.050
α_{22}	0.938	0.968	0.961	0.953
$\alpha_c(s_m = 0)$	0.533	0.443	0.456	0.528
$\alpha_c(s_m = 1)$	2.524	2.601	2.531	2.564
$\alpha_\pi(s_m = 0)$	0.640	0.661	0.658	0.647
$\alpha_\pi(s_m = 1)$	1.049	1.039	1.049	1.047
σ_m	1.310	1.306	1.309	1.309
$\beta_c(s_f = 0)$	-0.028	-0.028	-0.028	-0.028
$\beta_c(s_f = 1)$	0.011	0.014	0.011	0.012
$\beta_b(s_f = 0)$	-0.029	-0.033	-0.028	-0.029
$\beta_b(s_f = 1)$	0.050	0.052	0.054	0.051
$\beta_g(s_f = 0)$	1.016	1.020	1.017	1.018
$\beta_g(s_f = 1)$	0.644	0.603	0.634	0.637
σ_f	0.014	0.014	0.014	0.014
log-likelihood	275.137	270.409	273.332	274.629

Table 5: MLE for Error Component Model (1949:Q2-2014:Q2)

D.2 Quasi-Bayesian Estimates

Parameter	Prior	Mean	Var	(1)	(2)	(3)
τ_m	Normal	0.000	1.000	0.541	0.427	0.090
τ_f	Normal	0.000	1.000	-0.467	-0.323	0.121
λ_1	Beta	0.330	0.100	0.357	-	0.750
λ_2	Beta	0.540	0.100	0.430	-	0.728
ϕ_1	Beta	0.240	0.100	0.235	-	0.267
ϕ_2	Beta	0.290	0.100	0.306	-	0.318
ψ_1	Beta	0.998	0.100	0.972	0.999	0.960
ψ_2	Beta	0.998	0.100	0.952	0.999	0.946
α_{11}	Normal	0.940	0.300	0.946	0.957	0.999
α_{21}	Normal	0.000	0.300	-0.007	0.004	-
α_{12}	Normal	0.100	0.300	0.069	0.049	-
α_{22}	Normal	0.970	0.300	0.974	0.964	0.990
$a_c(s_m = 0)$	Gamma	0.570	0.300	0.548	0.547	1.297
$a_c(s_m = 1) - a_c(s_m = 0)$	Gamma	1.500	1.000	2.029	2.030	2.446
$a_\pi(s_m = 0)$	Gamma	0.640	0.300	0.637	0.645	0.615
$a_\pi(s_m = 1) - a_\pi(s_m = 0)$	Gamma	0.410	0.300	0.404	0.399	0.401
σ_m	Gamma	1.320	0.200	1.310	1.311	1.439
$b_c(s_f = 0)$	Normal	0.000	0.300	-0.010	-0.009	0.001
$b_c(s_f = 1) - b_c(s_f = 0)$	Gamma	0.040	0.030	0.031	0.030	0.022
$b_b(s_f = 0)$	Normal	0.000	0.300	-0.037	-0.043	-0.019
$b_b(s_f = 1) - b_b(s_f = 0)$	Gamma	0.100	0.300	0.078	0.086	0.112
$b_g(s_f = 0)$	Normal	1.010	0.300	0.854	0.871	0.717
$b_g(s_f = 1) - b_g(s_f = 0)$	-Gamma	-0.380	0.100	-0.276	-0.304	-0.307
σ_f	Gamma	0.014	0.003	0.014	0.014	0.016

Table 6: Quasi-Bayesian Maximum Likelihood Estimates (1949:Q2-2014:Q2)

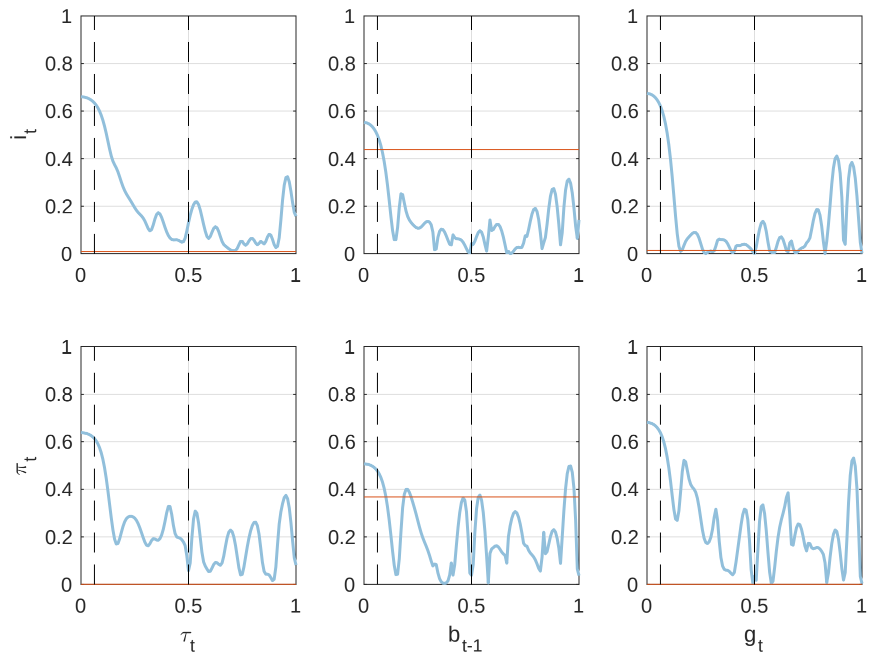
D.3 VAR(1) Estimates of Extracted Latent Factor

Table 7: VAR(1) Fit for Extracted Latent Factor

Parameter	(1)-Unres.	(2)-Res.
α_{11}	0.989 (0.026)	0.975 (0.017)
α_{21}	0.037 (0.020)	0.028 (0.014)
α_{12}	-0.038 (0.036)	-0.023 (0.025)
α_{22}	0.921 (0.028)	0.942 (0.020)
corr b/w factors	0.780	0.428

D.4 Coherence and Correlation of Data

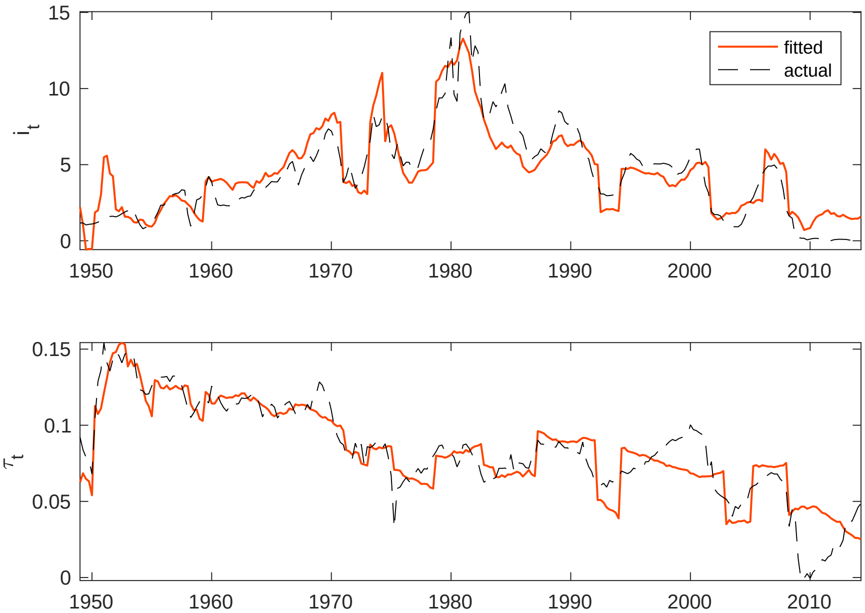
We collect a panel of coherence estimated pairwise for variables considered in our empirical exercise Figure 6. And we report the filtered probability of active regimes in Figure 8.



Note: Magnitude-squared coherence at frequency λ is defined to be $|f_{xy}(\lambda)|^2/[f_{xx}(\lambda)f_{yy}(\lambda)]$, with $f_{xy}(\lambda)$ denoting the cross-spectral-density, and estimated using Welch's method with window size 24. The dashed vertical lines indicate the frequencies associated with 6Q and 32Q. The red line represents squared correlation coefficient.

Figure 6: Coherence of Variables

D.5 Fitted Time-Series



Note: Policy regimes are identified using $\mathbb{E}(w_t|\mathcal{F}_t)$.

Figure 7: Fitted Policy Instruments

D.6 Filtered State Probability

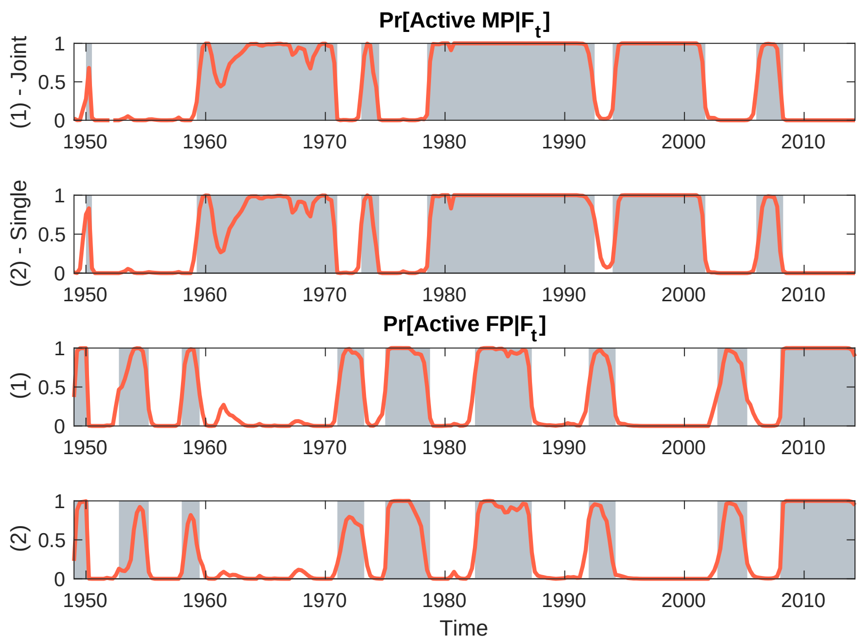


Figure 8: Filtered Probability of Active Regime