A Time-Varying Threshold STAR Model with Applications

A Time-Varying Threshold STAR Model with Applications*

Michael Dueker†
Laura E. Jackson‡
Michael T. Owyang§
Martin Sola¶

November 16, 2022

Abstract

Smooth-transition autoregressive (STAR) models, competitors of Markov-switching models, are limited by an assumed time-invariant threshold level. We augment the STAR model with a time-varying threshold that can be interpreted as a "tipping level" where the mean and dynamics of the VAR shift. Thus, the time-varying latent threshold level serves as a demarcation between regimes. We show how to estimate the model in a Bayesian framework using a Metropolis step and an unscented Kalman filter proposal. To show how allowing time variation in the threshold can affect the results, we present two applications: a model of the natural rate of unemployment and a model of regime-dependent government spending.

Keywords: Regime switching, smooth-transition autoregressive model, unemployment, nonlinear models.
JEL Classification: C22; E31; G12.

Abbreviations: STAR: Smooth-transition autoregressive; MCMC: Markov Chain Monte Carlo; MH: Metropolis-Hastings; HP: Hodrick-Prescott-filtered; VAR: Vector autoregression; UKF: Unscented Kalman Filter

*The authors thank Kristie Engemann, Kate Vermann, Hannah Shell, Julie Bennett, and Ashley Stewart for research assistance. The authors benefitted from conversations with Ana Galvão, Jim Hamilton, Tatevik Sekhposyan, and Tara Sinclair, as well as comments from participants at the Conference on Business Cycle Models at the UC Riverside and the SNDE. The views expressed here do not represent those of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

†Russell Investments
‡Department of Economics. Bentley University. Corresponding Author. ljackson@bentley.edu
§Research Division. Federal Reserve Bank of St. Louis.
¶Department of Economics. Universidad Torcuato di Tella

1 Introduction

Since Hamilton's (1989) Markov-switching model, time series models with nonlinearities have become increasingly popular. These models allow econometricians to characterize differences in economic dynamics across states, whether they be expansions and recessions or high and low uncertainty regimes. In addition to the standard Markov-switching model, other models can characterize differences in regimes. Threshold autoregressions and smooth transition autoregressions [henceforth STARs; see Granger and Teräsvirta, 1993] also have regime-switching dynamics, along with the added benefit that the state of the world can be determined by variables in the VAR. In particular, STAR models (which can nest threshold autoregressions) have recently been used for a number of applications: a time-varying natural rate of unemployment [Ball and Mankiw, 2002 and others]; fiscal policy [Auerbach and Gorodnichenko, 2012; Ramey and Zubairy, 2018 (henceforth RZ); etc]; and financial stress [Galvão and Owyang, 2018], among others.

In these applications, the threshold at which the model dynamics switch is constant.[1] For long samples (as in Ramey and Zubairy, 2018) or for particular types of variables (such as relative financial stress), one might want the threshold level of the regime-driving variable to vary over time. In this paper, we propose an extension of the STAR model that allows for time-variation of the latent threshold. In our model, the regime-driving variable is in the VAR, allowing shocks to any variable to affect future regime dynamics. We assume that the time-varying latent threshold follows an AR independent of the other variables in the VAR.

Because the type of time-varying threshold we consider makes direct maximum likelihood estimation intractable, we estimate the model in a Bayesian environment. Lopes and Salazar (2006) estimate a STAR model using a Metropolis-Hastings (MH) step to estimate the hyperparameters of the transition function. We take a similar approach but split the estimation of the transition hyperparameters. This becomes necessary both because of the time-variation in the threshold and because the threshold appears as the latent attractor in the error-correction

[1] Multiple-regime models with multiple thresholds do exist in the literature, but the thresholds themselves are constant across time. Only special cases of a time-varying threshold have appeared in the literature so far. Dueker et al. (2010) present a model in which the threshold level varies as a function of another observable variable.

term. In addition, because of the nonlinearity in the transition function, we use the unscented Kalman filter—a nonlinear version of the classic Kalman filter—to obtain the hyperparameters of a normal proposal.

We present two applications of the time-varying threshold STAR model: (i) a Phillips-curve-type application where the threshold is analogous to the natural rate of unemployment and (ii) a fiscal policy application where government spending multiplier varies across states of the economy.

In our first application, our time-varying threshold tracks some of the movements but is smoother than observed unemployment data and varies more over time than the CBO natural rate. For the most part, the qualitative responses to shocks are the same across the threshold. However, the responses are quantitatively larger in the low unemployment regime. In our second application, we find that fiscal policy shocks seem to produce different effects depending on the prevailing economic conditions at the time of the shock. Consistent with Auerbach and Gorodnichenko (2012), our government spending multipliers are much larger for shocks occurring in economic downturns than those occurring during expansions, where the latter take on values less than 1.

The rest of the paper develops as follows: Section 2 presents the STAR model with the time-varying threshold. Section 3 outlines the sampler used for the MCMC estimation. In particular, we provide details of the MH step using the unscented Kalman filter proposal for the time-varying threshold. Sections 4 and 5 present the empirical results for the two applications. Section 6 concludes.

2 The Time-Varying Threshold STAR Model

The STAR model is a regime switching model in which the state of the economy is a convex combination of two latent regimes. Let $\mathbf{y}_t$ be an $(n \times 1)$ vector of variables of interest. Define a state-dependent P_i-lag process, $\mathbf{y}_{it}$, as

$$\mathbf{y}_{it} = \alpha_i + \sum_{p=1}^{P_i} \theta_{ip} \mathbf{y}_{t-p} + \Gamma_i \left(z_{t-1} - z^* \right) + \boldsymbol{\varepsilon}_{it}, \tag{1}$$

where α_i is a vector of state-dependent intercepts, θ_{ip} are matrices of state-dependent VAR coefficients, Γ_i is a matrix of state-dependent error correction coefficients, and $\boldsymbol{\varepsilon}_{it} \sim N\left(\mathbf{0}_{n\times 1}, \Omega_i\right)$ for $i = 0, 1$.[2] Then, the STAR(P), $P = \max\left\{P_0, P_1\right\}$, model can be thought of as a combination of two error correction models:

$$\mathbf{y}_t = \left[1 - \pi\left(z_{t-1}\right)\right] \mathbf{y}_{0t} + \pi\left(z_{t-1}\right) \mathbf{y}_{1t}, \tag{2}$$

where z_{t-1} is the threshold variable and $\pi\left(z_{t-1}\right)$ is the transition function.

The threshold variable and the transition function determine the weights of each autoregression on the path of $\mathbf{y}_t$ in terms of the two latent regimes defined by $\pi\left(z_{t-1}\right) = 0$ and $\pi\left(z_{t-1}\right) = 1$. The path of the economy is determined by $\pi\left(z_{t-1}\right)$, which is bounded by zero and one and a function of past values of z_{t-1}. When $\pi\left(z_{t-1}\right) = 0$, the steady-state $\mathbf{y}_t$ is $\alpha_0 \left[1 - \sum_{p=1}^{P_0} \theta_{0p}\right]^{-1}$; conversely, when $\pi\left(z_{t-1}\right) = 1$, the steady-state $\mathbf{y}_t$ is $\alpha_1 \left[1 - \sum_{p=1}^{P_1} \theta_{1p}\right]^{-1}$. The transition function can take a number of forms and depends on the threshold variable, z_{t-1}. We assume a first-order logistic transition function[3]:

$$\pi\left(z_{t-1}\right) = \left[1 - \exp\left(-\gamma\left(z_{t-1} - z^*\right)\right)\right]^{-1}, \tag{3}$$

where $\gamma \geq 0$ is the speed of transition, z^* is a (fixed) threshold. To normalize the regimes, we impose $\gamma > 0$. **In principle, the threshold variable could consist of any data series (or combination of data series) inside or outside the model. Here, and in the empirical applications, z_t is also the variable inside the error correction term. Therefore, the threshold z^* measures the long-run value of z_t.**

In (3), the regime process is determined by the sign and magnitude of the deviation of

[2] See Chan and Tong (1986) and Granger and Teräsvirta (1993) for the baseline model. Rothman, van Dijk, and Franses (2001) extended the model to include error correction terms.

[3] We refer the reader to van Dijk, Teräsvirta, and Franses (2002) for a review of alternative representations for the transition function. The estimation algorithm outlined below generalizes to any number of alternative transition functions.

z_{t-1} from the threshold z^*. If z_{t-1} is less than z^*, the transition function, $\pi(z_{t-1})$, moves further toward the first regime (in this case, normalized to $\pi(z_{t-1}) = 0$). The coefficient γ determines the speed of adjustment: as $|\gamma| \to \infty$, the transition becomes sharper. At $\gamma = 0$, the model has only one regime, leaving the model parameters unidentified.

We are interested in the case in which the threshold parameter, z^*, varies over time. This might occur, for example, in a model of unemployment where the natural rate is time varying. We can rewrite (3) to account for time-variation:

$$\pi\left(z_{t-1}; z^*_{t-1}\right) = \left[1 - \exp\left(-\gamma\left(z_{t-1} - z^*_{t-1}\right)\right)\right]^{-1}. \tag{4}$$

In addition, we need to specify a process for the evolution of the time-varying threshold. Suppose that the time-varying threshold is an autoregressive process:

$$z^*_t = \lambda_0 + \sum_{j=1}^{m} \lambda_j z^*_{t-j} + u_t, \tag{5}$$

where u_t is normalized to have a small, fixed variance.[4] This assumption implies that—while the threshold varies over time—the movements are gradual and the process displays a strong sense of inertia. The threshold innovation may be correlated with the shocks to the observable data,

$$Cov(\boldsymbol{\varepsilon}_t, u_t) = \rho,$$

where ρ is a vector of parameters that governs this cross-correlation. For simplicity, we abstract from potential regime-dependence in the error terms so that $\boldsymbol{\varepsilon}_t \sim N(0, \Omega)$ and $v_t = [\boldsymbol{\varepsilon}_t, u_t]'$, where

$$v_t \sim N\left(0, \begin{bmatrix} \Omega & \rho \\ \rho & \sigma^2 \end{bmatrix}\right).$$

[4]The threshold is not restricted to an AR process. It could also be a function of exogenous shifters as in Dueker et al. (2010).

In practice, we will normalize the threshold innovation, $\sigma^2 = 0.01$, for identification.

3 Estimation

Because the time-varying latent threshold enters the regime weights, $\pi(z_{t-1})$ and $1-\pi(z_{t-1})$, the model may be difficult to estimate by maximum likelihood. It is also not possible to derive exact conditional distributions for Gibbs sampling. Instead, an MH algorithm [Chib and Greenberg, 1995] is needed for Bayesian estimation of this model. The MH algorithm generates a draw from the target distribution by first drawing a candidate from a proposal density and then accepting or rejecting the candidate based on a probability determined, in part, by the model likelihood.

The algorithm partitions the set of model parameters into five blocks, including one block that draws the time-varying threshold. The parameter groupings for Markov Chain Monte Carlo estimation of the model are: (1) $\mathbf{\Psi} = [\boldsymbol{\alpha}, \boldsymbol{\theta}, \mathbf{\Gamma}]$, the coefficients in the observation equation; (2) λ, the coefficients in the latent variable equation; (3) $\mathbf{\Sigma} = [\Omega, \rho]$, the elements in the variance-covariance matrix for joint system; (4) γ, the coefficient in the transition function; and (5) $\{z_t^*\}_{t=1}^T$, the set of latent time-varying thresholds. We can combine the estimation of the first and second blocks. Posteriors are constructed from 5,000 iterations after discarding the first 5,000 to achieve convergence.

The priors for each of the parameter blocks, their prior distributions, and their prior hyperparameters are given in Table 1, where $\kappa = n(4 + P_0 + P_1) + m + 1$. The VAR coefficients, the coefficient on the lag in the latent variable equation, the covariance between the VAR parameters, and the initial value of the threshold have normal priors. The covariance matrix for the VAR has an inverse Wishart prior. The rate of transition parameter, γ, has a gamma prior.

Most of the draws are conjugate with details in Section A in the Appendix. The time-varying thresholds, $\{z_t^*\}_{t=1}^T$, are drawn from an MH-in-Gibbs step. To speed the convergence of the algorithm, we want a tractable proposal density close to the target distribution. Because the system is nonlinear, the standard linear Kalman filter is not appropriate. Instead, we

obtain the MH proposal densities from the unscented Kalman filter (UKF).[5] The UKF is an alternative to the extended Kalman filter that tracks the state variable by computing its distribution across a set of deterministic points called sigma points. Because the UKF may not be familiar to some practitioners, we outline the algorithm and its use in the MH step here.

3.1 The Unscented Kalman Filter

The model can be recast in its state-space representation, with a linear measurement equation and a nonlinear state equation:

$$\mathbf{x_t} = \mathbf{g}(\mathbf{x_{t-1}}) + \epsilon_t, \tag{6}$$

where $\mathbf{x}_t = [\mathbf{y}_t, z_t^*]'$ is the state variable and $\mathbf{g}(.) = [\mathbf{g}_1(.); \mathbf{g}_2(.)]$ is the state equation, where $\mathbf{g}_1(.)$ is eq. (2) and $\mathbf{g}_2(.)$ is eq. (5).

The state vector is estimated using the UKF, a nonlinear filter that serves as an alternative to the extended Kalman filter (EKF). The EKF uses first-order Taylor-series approximations to any nonlinear functions in the measurement and transition equations. The UKF tracks the state variable by computing its distribution across a set of L deterministic points called sigma points, $\boldsymbol{\chi}_t^n$, that are functions of $\mathbf{x}_t$ described below.[6] Instead of propagating the state directly through eq. (6), the UKF propagates the sigma points and uses a weighted sum of the propagated sigma points in the prediction or update step.

Define the sigma points as

$$\boldsymbol{\chi}^n = \begin{cases} \mathbf{x} & , \quad for \quad n = 0 \\ \mathbf{x} + \left(\sqrt{(L+\zeta)\,\mathbf{Q}}\right)_n & , \quad for \quad n = 1, ..., L \\ \mathbf{x} - \left(\sqrt{(L+\zeta)\,\mathbf{Q}}\right)_{n-L} & , \quad for \quad n = L+1, ..., 2L \end{cases}, \tag{7}$$

where $\zeta = a^2\,(L + \kappa) - L$ and the sigma point retains the time index of $\mathbf{x}$. Here, a and κ

[5] See Julier and Uhlmann (1997) and Wan and van der Merwe (2001).
[6] See Julier and Uhlmann (1997) and Wan and van der Merwe (2001).

are user-chosen parameters that govern the spread and scale of the cloud of sigma points, respectively, and $\left(\sqrt{X}\right)_i$ is the ith column of the lower triangular Cholesky factorization of the square matrix X. The matrix $\mathbf{Q}$ is the uncertainty surrounding the state vector $\mathbf{x}$.

Given the set of sigma points, we propagate $\boldsymbol{\chi}^n$ through the function $\mathbf{g}(.)$ to recover a *prediction* for the period t value of the sigma points:

$$\boldsymbol{\chi}^n_{t|t-1} = \mathbf{g}\left(\boldsymbol{\chi}^n_{t-1}\right), \quad \text{for} \quad n = 0, ..., 2L. \tag{8}$$

The state prediction, $\widehat{\mathbf{x}}_{t|t-1}$, and its covariances can then be extrapolated from a weighted sum of the propagated sigma points:

$$\widehat{\mathbf{x}}_{t|t-1} = \sum_{n=0}^{2L} w_s^n \boldsymbol{\chi}^n_{t|t-1}$$

and

$$\mathbf{Q}_{t|t-1} = \sum_{n=0}^{2L} w_c^n \left[\boldsymbol{\chi}^n_{t|t-1} - \widehat{\mathbf{x}}_{t|t-1}\right]\left[\boldsymbol{\chi}^n_{t|t-1} - \widehat{\mathbf{x}}_{t|t-1}\right]',$$

where $w_s^0 = \zeta\left(L+\zeta\right)^{-1}$, $w_c^0 = \zeta\left(L+\zeta\right)^{-1} + \left(1-\alpha^2+\beta\right)$, and $w_s^n = w_c^n = \frac{1}{2}\left(L+\zeta\right)^{-1}$ for all other n. The *updated* state is computed similarly to the standard Kalman filter:

$$\widehat{\mathbf{x}}_{t|t} = \widehat{\mathbf{x}}_{t|t-1} + K_t\left(\mathbf{y}_t - \widehat{\mathbf{y}}_t\right), \tag{9}$$

where K_t is the Kalman gain:

$$K_t = \mathbf{Q}_{yz}^{-1}\mathbf{Q}_{yy}.$$

Here, $\mathbf{Q}_{yz}$ defines the cross-covariance:

$$\mathbf{Q}_{yz} = \sum_{n=0}^{2L} w_c^n \left[\boldsymbol{\chi}^n_{t|t-1} - \widehat{\mathbf{x}}_{t|t-1}\right]\left[\boldsymbol{\zeta}^n_{t|t} - \widehat{\mathbf{y}}_t\right]',$$

where $\boldsymbol{\zeta}_{t|t}^n$ are obtained from transforming the sigma points, $\boldsymbol{\chi}_{t|t-1}^n$, using the state-space's measurement equation. $\mathbf{Q}_{yy}$ is the predicted covariance:

$$\mathbf{Q}_{yy} = \sum_{n=0}^{2L} w_c^n \left[\boldsymbol{\zeta}_{t|t}^n - \widehat{\mathbf{y}}_t\right] \left[\boldsymbol{\zeta}_{t|t}^n - \widehat{\mathbf{y}}_t\right]'$$

and updated state covariance is:

$$\mathbf{Q}_{t|t} = \mathbf{Q}_{t|t-1} - K_t \mathbf{Q}_{yy} K_t'.$$

3.2 Unscented Smoothing

Multi-move sampling of the latent attractor, $\{\mathbf{z}^*\}_{t=1}^T$, requires backwards sampling from one-period smoothed inferences of the state vector. A typical smoothed Kalman filter uses a forward filter and a backward smoother. Because the filter is linear, it is always possible to employ a backward smoother for the Kalman filter. The unscented transformation, however, is not always invertible, making a backward-looking smoother harder to implement in all cases. The Rauch-Tung-Striebel (RTS) filter is an alternative smoother for the Kalman filter. Särkkä (2008) constructed a smoother using principles similar to RTS. The unscented RTS smoother augments the unscented Kalman filter with a step that recomputes the state estimate:

$$\widehat{\mathbf{x}}_t^s = \widehat{\mathbf{x}}_t + D_t \left[\widehat{\mathbf{x}}_{t+1}^s - \widehat{\mathbf{x}}_{t+1|t}\right] \tag{10}$$

with covariance matrix

$$\mathbf{Q}_t^s = \mathbf{Q}_t + D_t \left[\mathbf{Q}_{t+1}^s - \mathbf{Q}_{t+1|t}\right] D_t'.$$

The smoother gain, D_t, is defined by

$$D_t = \mathbf{Q}_{z_t, z_{t+1}}^{-1} \mathbf{Q}_{t+1|t},$$

where

$$\mathbf{Q}_{z_t,z_t+1} = \sum_{n=0}^{2L} w_c^n \left[\boldsymbol{\chi}_{t+1|t}^n - \widehat{\mathbf{x}}_{t+1|t}\right]\left[\boldsymbol{\chi}_{t|t}^n - \widehat{\mathbf{x}}_{t|t}\right]'$$

and

$$\boldsymbol{\chi}_{t|t} = \boldsymbol{\chi}_{t|t-1} + K_t(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}). \tag{11}$$

The series of latent thresholds, $\{z_t^*\}_{t=1}^T$, are elements of the state vector, $\widehat{\mathbf{x}}_t^s$. The UKF yields estimates of the state vector for the state space representation from Section 3.1, $\widehat{\mathbf{x}}_t$, and its uncertainty, $\mathbf{Q}_t^s$, for each time period. In linear models, one would normally obtain the smoothed posterior density and draw the values of the state vector recursively. For nonlinear models, there are no optimal filters. Rather than drawing directly from the Kalman posterior, the output of the UKF can be thought of as a proposal density for an MH algorithm used to draw the latent thresholds. Given $\widehat{\mathbf{x}}_t^s$ and $\mathbf{Q}_t^s$, the candidate $\widehat{z}_t^*$ can be drawn as a subvector of $\mathbf{x_t}$:

$$\widehat{z}_t^* \sim N\left(\widehat{\mathbf{x}}_t^s, \mathbf{Q}_t^s\right), \tag{12}$$

where the subscripted z indicates the truncation of the state vector. The candidate is accepted with probability $a_z = \min\{A_z, 1\}$, where

$$A_z = \frac{\prod_t \phi\left(\mathbf{y}_t | \pi\left(z_{t-1}|\gamma, \widehat{z}_t^*\right), \boldsymbol{\Psi}, \boldsymbol{\Sigma}\right)}{\prod_t \phi\left(\mathbf{y}_t | \pi\left(z_{t-1}|\gamma, z_t^{*[i]}\right), \boldsymbol{\Psi}, \boldsymbol{\Sigma}\right)},$$

and the superscript $[i]$ represents the past accepted value.

4 Application 1: The Natural Rate of Unemployment

The Phillips curve and Okun's Law are typically estimated using the value of the unemployment rate relative to its natural rate, an unobserved quantity. STAR models have the appealing feature that the threshold unemployment rate can determine the evolution of inflation

[e.g., Deschamps, 2008; Skalin and Teräsvirta, 2002].[7] While the constant STAR threshold leaves it uninterpretable, our time-varying threshold might serve as a meaningful estimate of the natural rate.[8] More precisely, the time-varying threshold is the "tipping level" rate of unemployment at which the mean and/or dynamics of the unemployment rate shift.

4.1 Data

Estimation of the model requires a set of observables, $\mathbf{y}_t$, and the threshold variable, z_{t-1}, that determines the regime process. **For this application, the set of observables includes the first difference of the monthly unemployment rate, the first difference of the monthly log CPI, and an interest rate. The unemployment rate is taken from the BLS's payroll employment survey. Our measure of inflation uses the CPI, excluding food and energy. The interest rate is the 3-month T-Bill. The full sample of data begins in January 1968 and ends in December 2019.**[9]

To make the model self-exciting, the threshold variable, z_{t-1}, is chosen to be one of the set of observables—in this case, the first lag of the level of the unemployment rate.[10] We choose one lag for the threshold evolution (i.e., $m = 1$ in eq. (5)) and three lags for the VAR (i.e., $p = 3$).[11]

[7] The literature also uses unobserved components decompositions that interpret the permanent component of unemployment as the natural rate [e.g., Basistha and Nelson, 2007; Clark, 1989; Doménech and Gómez, 2006; Jaeger and Parkinson, 1994; Sinclair, 2009] and Markov-switching models to characterize the shifting relationship between inflation and unemployment [e.g., Chauvet, Juhn, and Potter, 2002; Phelps and Zoega, 1998].

[8] A number of studies find evidence in favor of a time-varying natural rate [see, for example, Ball and Mankiw, 2002; Gordon, 1995, 1997; Summers, 1986]. The time-varying natural rate models can account for changes in labor force participation, discouraged workers, changes in unemployment compensation, and expectations of future policies.

[9] We end the sample prior to the start of the COVID recession. Extending the sample may require structural breaks; we leave this for future research.

[10] In response to a referee, we determined that estimated threshold is sensitive to the variance of the correlation parameter, ρ. Not surprisingly, when the true threshold is constant, the algorithm requires a fairly tight prior for ρ around zero. We discuss the simulation exercise further in the Sensitivity Analysis Section C in the Appendix.

[11] Our prior for the threshold parameter is close to a random walk so one lag seems appropriate. We choose 3 lags for the VAR to cover the number of months in a quarter.

4.2 Results

Figure 1 shows the estimated threshold, along with the unemployment rate and the CBO estimate of the natural rate.[12] The latent threshold exhibits fluctuations (i) similar to the unemployment rate, (ii) smoother than the unemployment rate, but (iii) more volatile than the CBO natural rate.

Figure 2 shows the regime weights—i.e., the values of the transition function for each period. When $z_{t-1} > z^*_{t-1}$, $\pi\left(z_{t-1}; z^*_{t-1}\right) \to 1$, which we characterize as the slack regime. The weight on the slack regime increases around the beginning of the NBER recessions and decreases during the recovery periods following the trough. The slack regime weight falls slower after the 1991, 2000, and 2007-09 recessions than other recessions, reflecting the jobless recoveries [Schreft and Singh, 2003; Engemann and Owyang, 2010].[13]

We identify the shocks using a Cholesky ordering with unemployment ordered first, inflation second, the interest rate third, and the latent threshold last. This implies unemployment and inflation do not contemporaneously respond to interest rates and no variable contemporaneously responds to the latent threshold.[14] We compute both regime-dependent (RDIRFs) and generalized impulse responses [GIRFs, see Koop, Pesaran, and Potter, 1996].

Figure 3 shows the RDIRFs for the three variables and the latent threshold, the shocked variable in the columns and the responding variable in the rows. To compute the RDIRFs, we set either $\pi\left(z_{t-1}\right) = 0$ (expansion; light gray) or $\pi\left(z_{t-1}\right) = 1$ (slack; dark gray) at the time of the shock and hold the regime constant for the life of the response.

Shocks to the macro variables are consistent with the VAR literature, including the price puzzle [Hanson, 2004]. Shocks occurring during expansion exhibit greater persistence than shocks originating during slack periods. For the most part, the qualitative responses to shocks are the same across the threshold; however, the responses are quantitatively larger in the low unemployment regime.

[12] The unscented Kalman filter proposal density yielded acceptance rates that average around 20 percent across time periods.

[13] In these cases, the smooth transition has an advantage over Markov-switching or TAR models. Since those models characterize the business cycle as *discrete* phases, modeling jobless recoveries requires a third regime.

[14] Because the unemployment rate and CPI enter the VAR in differences, the impulse responses are converted into levels. See the Appendix for details on how to parameterize the model for these impulse responses.

The responses to shocks to the threshold are shown in the right-most column. In the expansion regime, a shock that increases the threshold produces a large, persistent increase in the unemployment rate and a decrease in the interest rate. Alternatively, in the slack regime, a shock that increases the threshold tightens labor market conditions and raises interest rate.

For RDIRFs, the response of the unemployment rate does not affect the underlying state of the economy. However, the level of the unemployment rate determines the state by construction. GIRFs, on the other hand, allow the regime to respond to the model variables. We construct the GIRFs under two scenarios: (1) $\pi\left(z_{t-1}; z^*_{t-1}\right) < 0.2$; and (2) $\pi\left(z_{t-1}; z^*_{t-1}\right) > 0.8$, corresponding to periods of labor market tightness versus labor market slack, respectively. In each scenario, a shock (to any variable) could push the unemployment rate across the threshold. Because of the nonlinearity, the responses will also vary depending on the size and direction of the shock, as well as the future shocks to all variables. We consider positive shocks when unemployment starts below the threshold and negative shocks when unemployment starts above the threshold.

Figure 4 plots the GIRFs to the four shocks (in the columns) for the four variables (in the rows). We invert the responses to shocks in the slack scenario (2) to facilitate comparison. The direction of the responses are consistent across scenarios—at least the first three to four months—suggesting that the regimes affect the magnitude of the responses but not their direction.[15] The right-most column shows the effect of a change in the threshold, which is larger in the expansion scenario than in the slack scenario.

Figure 5 plots the responses for the same scenarios, but holding the threshold constant at the mean value of the observed unemployment rates within the sample, 6.1 percent. Imposing a constant threshold produces responses to an unemployment shock that are more similar across regimes. Perhaps more importantly, the impulse responses across the regimes have completely different characteristics than those in the time-varying threshold model. This suggests that, in some cases, incorporating the time variation may be required to maximize differences in the regime dynamics.

[15] The estimated threshold is very close to a random walk. Shocks to other model variables have only an impact effect on the threshold.

5 Application 2: Government Spending Shocks

Beginning with Auerbach and Gorodnichenko (2012), a number of papers have argued for [Fazarri, Morley, and Panovska, 2015; Caggiano et al., 2015; and others] and against [Owyang, Ramey, and Zubairy, 2013; RZ] state-dependence in the effect of government spending shocks. Many of these papers test whether the macroeconomic response to fiscal policy varies when there is slack in the economy and crowding out is less likely. In many of the papers, however, the threshold level of economic activity—however it is measured—is assumed to be constant.

Time-varying slack might be relevant in a study such as RZ, who estimate multipliers over a very long sample fixing the unemployment threshold at 6.5 percent.[16] We construct GIRFs from our model with a time-varying threshold and use these to compute state-dependent multipliers.

5.1 Data

The baseline system contains four variables: (i) the first difference of output relative to potential; (ii) the first difference of government spending relative to potential output; (iii) the first difference of the unemployment rate; and (iv) a news shock. The model differs from the RZ TAR model in that we include the unemployment rate to make the model self-exciting, ordered ahead of output. As in RZ, we compute potential output as a sixth order polynomial time trend, excluding the same period around WWII that RZ exclude. For this application, the data are quarterly, beginning in 1946:I and ending in 2015:IV, when the RZ data end. As before, **we use the first lag of the level of the unemployment rate** to drive the regime process and we use 4 lags in the VAR, consistent with RZ.

We include the contemporaneous value of the news shock constructed and described in RZ, X_t, as an exogenous regressor in the VAR, allowing the coefficient to be state-dependent

[16] In their paper, RZ estimate the multiplier using a state-dependent version of local projections (LP). The LP framework is a parametric, direct-multistep method of computing impulse responses. In a later section, they show that the same results are obtained when using a version of a threshold autoregression (TAR) that is nested by our STAR model.

(similarly to θ_{ip} and Γ_i).[17] Thus, to construct the multipliers, we consider the contemporaneous effect of these exogenous news shocks, ρ_i, and construct impulse responses accordingly based on the GIRF technique. In this particular application, the analogue to eq. (1) is the following:

$$\mathbf{y}_{it} = \alpha_i + \sum_{p=1}^{P_i} \theta_{ip} \mathbf{y}_{t-p} + \Gamma_i \left(z_{t-1} - z^* \right) + \rho_i X_t + \boldsymbol{\varepsilon}_{it}. \tag{13}$$

The remainder of the model follows the structure described in Section 2. We choose one lag for the threshold evolution (i.e., $m = 1$) and include four lags for the VAR (i.e., $P_i = 4$ for all i).

5.2 Results

Figure 6 shows the estimated threshold, along with the unemployment rate and a constant 6.5-percent value for the natural rate (as employed by Ramey and Zubairy (2018)).[18] Interestingly, our estimate of the latent threshold lies below 6.5 percent for the entirety of the sample.[19] The threshold exhibits fluctuations similar to the unemployment rate but is considerably smoother. There are a few periods where both the magnitude and the sign of the difference between the unemployment rate and the estimated threshold differ from what would be inferred with a constant threshold [i.e., recessions around 1954, 1970, and 2001]. However the remaining recessionary periods are identified similarly, at least in terms of the sign of the gap.

Figure 7 shows the estimated regime weights, the values of the transition function for each period. Here, the shaded areas represent the periods of slack computed using the RZ constant threshold. For the time-varying threshold, the weight on the slack regime increases around the beginning of the NBER recessions and decreases during the recovery period. We again

[17] The shock is computed as a percent of GDP. The news shock is designed by RZ to be exogenous by construction.

[18] For this application, the use of the unscented Kalman filter as a proposal density yielded acceptance rates that average around 10 percent across time periods. We estimated a constant-threshold TAR with the same data as our application and compared the BIC of this model to ours. Our model (1268.104) is preferred to both a smooth-transition VAR with constant threshold (1297.602) and a threshold VAR with a constant threshold (1270.867).

[19] Our sample is shorter than the full sample used by RZ. Importantly, our sample does not include the Great Depression.

capture the jobless recoveries following the 1991, 2000, and 2007-09 recessions. Additionally, we see a similar dynamic throughout the 1970's and 1980's where labor market slack seems to have persisted throughout much of both decades. The constant threshold approach has fewer slack periods, most of which occur during the recovery phases after the NBER troughs.

As before, we compute GIRFs to the exogenous news shock, X_t, and compute the government spending multipliers. GIRFs are converted into levels and are computed under the same scenarios: (1) an expansion scenario, $\pi\left(z_{t-1}; z^*_{t-1}\right) < 0.2$; and (2) slack scenario, $\pi\left(z_{t-1}; z^*_{t-1}\right) > 0.8$.

With this setup, we can investigate whether (i) the government spending multiplier is larger during periods of slack and (ii) the total increase in output is larger than the increase in government spending. The former asks whether the multiplier is larger in scenario 2 than in scenario 1. The latter asks whether the multiplier in either case is greater than 1. To produce the state-dependent multipliers, let $\Delta G_j(h)$ be the response of government spending to a shock to variable j at horizon h and let $\Delta Y_j(h)$ be the response of output to a shock to variable j at horizon h. We consider the 2-year and 4-year cumulative multipliers with respect to a shock to X_t as:

$$\left[\sum_{h=0}^{H} \frac{\Delta G_X(h)}{\Delta X}\right]^{-1} \sum_{h=0}^{H} \frac{\Delta Y_X(h)}{\Delta X},$$

for $h = \{8, 16\}$ quarters.

Table 2 reports the posterior mean estimates of the multipliers for the two scenarios for both horizons $h = \{8, 16\}$ quarters. The multipliers during periods of labor-market tightness in the expansion scenario are less than 1 and similar in magnitude at both horizons. However, we find evidence of a much larger multiplier during periods of slack, reaching a level close to 3 after two years and falling to just below 2 after four years.

These results are consistent with those in, for example, Auerbach and Gorodnichenko (2012) and contrast those in RZ. There are a number of reasons for this finding. First, estimating the threshold allows us to maximize the difference in the responsiveness of output to government spending across regimes. Second, allowing the threshold to vary over time further

facilitates maximizing this difference. Third, our scenarios originate the shock in periods in which the regime is entrenched (or well defined). Assuming that there are differences across the regimes, our scenario definitions may amplify the differences. An increase in government spending increases output and decreases the unemployment rate. If this effect is larger in the slack regime, the longer the economy stays in the slack regime, the larger the multiplier will be. Thus, starting the economy in a deeper recession should amplify the effect.

6 Conclusions

The standard smooth transition autoregression is a popular alternative to Markov-switching models. One drawback of the standard STAR model is that it must be estimated with a constant threshold (or constant thresholds in the case of multiple-regime models). We extend the STAR model to include a time-varying threshold. The added complexity of the time-varying threshold requires a nonlinear alternative to the Kalman filter to generate proposals for an MH step. We present two applications of the model: (i) interpreting the threshold as the natural rate of unemployment in the Phillips-curve and (ii) exploiting the model-implied states of the economy to produce government spending multipliers which can vary throughout the business cycle.

In the first application, the regimes in the STAR model depend in part on the sign of the unemployment gap since the threshold can be interpreted as the natural rate of unemployment. Moreover, allowing the threshold to vary lets us analyze the effect of shocks to the natural rate. We find that the effect of these shocks differs across regimes.

In the second application, we find that fiscal policy shocks seem to produce different effects depending on the prevailing economic conditions at the time of the shock. Government spending multipliers during expansions are less than 1 but are much larger, reaching values of at least 2, during economic downturns.

References

[1] Auerbach, A.J. and Gorodnichenko, Y., 2012. Fiscal multipliers in recession and expansion. *Fiscal policy after the financial crisis* (pp. 63-98). University of Chicago Press.

[2] Ball, L., Mankiw, N. G., 2002. The NAIRU in theory and practice. *Journal of Economic Perspectives, 16*(4), pp.115-136. DOI: 10.1257/089533002320951000.

[3] Basistha, A., Nelson, C. R., 2007. New measures of the output gap based on the forward-looking new keynesian phillips curve. *Journal of Monetary Economics, 54*(2), pp.498-511. DOI: 10.1016/j.jmoneco.2005.07.021.

[4] Caggiano, G., Castelnuovo, E., Colombo, V. and Nodari, G., 2015. Estimating fiscal multipliers: News from a non-linear world. *The Economic Journal*, *125*(584), pp.746-776. DOI: 10.1111/ecoj.12263.

[5] Chan, K. S., Tong, H., 1986. On estimating thresholds in autoregressive models. *Journal of Times Series Analysis, 7*(3), pp.179-190.

[6] Chan, J., Jeliazkov, I., 2009. MCMC Estimation of Restricted Covariance Matrices. *Journal of Computational and Graphical Statistics, 18*(2), pp.457-480. DOI: 10.1198/jcgs.2009.08095.

[7] Chauvet, M., Juhn, C., Potter, S., 2002. Markov switching in disaggregate unemployment rates. *Empirical Economics, 27*(2), pp.205-232. DOI: 10.1007/s001810100101.

[8] Chib, S. 1993. Bayes regression with autoregressive errors: a Gibbs sampling approach. *Journal of Econometrics, 58*(3), pp.275-294. DOI: 10.1016/0304-4076(93)90046-8.

[9] Chib, S., Greenberg, E., 1995. Understanding the metropolis-hastings algorithm. *The American Statistician, 49*(4), pp.327-335.

[10] Chib, S., Greenberg, E., 1996. Markov chain monte carlo simulation methods in econometrics. *Econometric Theory, 12*(3), pp.409-431. DOI: 10.1017/S0266466600006794.

[11] Clark, P. K., 1989. Trend reversion in real output and unemployment. *Journal of Econometrics, 40*(1), pp.15-32. DOI: 10.1016/0304-4076(89)90027-4.

[12] Deschamps, P. J., 2008. Comparing smooth transition and Markov switching autoregressive models of US unemployment. *Journal of Applied Econometrics 23*(4), pp.435-462. DOI: 10.1002/jae.1014.

[13] Doménech, R., Gómez, V., 2006. Estimating potential output, core inflation, and the NAIRU as latent variables. *Journal of Business and Economic Statistics, 24*(3), pp.354-365. DOI:10.1198/073500105000000315.

[14] Dueker, M.J., Psaradakis, Z., Sola, M. and Spagnolo, F., 2013. State-dependent threshold smooth transition autoregressive models. *Oxford Bulletin of Economics and Statistics, 75*(6), pp.835-854. DOI: 10.1111/j.1468-0084.2012.00719.x.

[15] Engemann, K., Owyang, M. T., 2010. Whatever Happened to the Business Cycle: A Bayesian Analysis of Jobless Recoveries. *Macroeconomic Dynamics, 14*(5), pp.709-726. DOI: 10.1017/s1365100509990812.

[16] Fazzari, S.M., Morley, J. and Panovska, I., 2015. State-dependent effects of fiscal policy. *Studies in Nonlinear Dynamics & Econometrics, 19*(3), pp.285-315. DOI: 10.1515/snde-2014-0022.

[17] Galvão, A.B. and Owyang, M.T., 2018. Financial stress regimes and the macroeconomy. *Journal of Money, Credit and Banking, 50*(7), pp.1479-1505. DOI: 10.1111/jmcb.12491.

[18] Gordon, R. J., 1995. Is there a trade-off between unemployment and productivity growth? NBER Working Papers 5081. DOI: 10.3386/w5081.

[19] Gordon, R. J., 1997. The time-varying NAIRU and its implications for economic policy. *Journal of Economic Perspectives, 11*(1), pp.11-32. DOI: 10.1257/jep.11.1.11.

[20] Granger, C. W. J., Teräsvirta, T., 1993. Modelling Nonlinear Economic Relationships. *OUP Catalogue*, Oxford University Press, number 9780198773207.

[21] Hamilton, J. D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the econometric society, 57*, pp.357-384. DOI: 10.2307/1912559.

[22] Hanson, M., 2004. The "price puzzle" reconsidered. *Journal of Monetary Economics, 51*(7), pp.1385-1413. DOI: 10.1016/j.jmoneco.2003.12.006.

[23] Jaeger, A., Parkinson, M., 1994. Some evidence on hysteresis in unemployment rates. *European Economic Review, 38*(2), pp.329-342. DOI: 10.1016/0014-2921(94)90061-2.

[24] Julier, S.J. and Uhlmann, J.K., 1997, July. New extension of the Kalman filter to nonlinear systems. *Signal processing, sensor fusion, and target recognition VI* (Vol. 3068, pp. 182-193). Spie. DOI: 10.1117/12.280797.

[25] Koop, G., Pesaran, M. H. and Potter, S. M., 1996. Impulse Response Analysis in Nonlinear Multivariate Models. *Journal of Econometrics, 74*(1), pp.119-147. DOI: 10.1016/0304-4076(95)01753-4.

[26] Lopes, H.F., Salazar, E., 2006. Bayesian model uncertainty in smooth transition autoregressions. *Journal of Time Series Analysis, 27*(1), pp.99-117. DOI: 10.1111/j.1467-9892.2005.00455.x.

[27] Owyang, M.T., Ramey, V.A. and Zubairy, S., 2013. Are government spending multipliers greater during periods of slack? Evidence from twentieth-century historical data. *American Economic Review, 103*(3), pp.129-34. DOI: 10.1257/aer.103.3.129.

[28] Phelps, E.S., Zoega, G., 1998. Natural-rate theory and OECD unemployment. *The Economic Journal, 108*(448), pp.782-801. DOI: 10.1111/1468-0297.00315.

[29] Ramey, V.A. and Zubairy, S., 2018. Government spending multipliers in good times and in bad: evidence from US historical data. *Journal of political economy, 126*(2), pp.850-901. DOI: 10.1086/696277.

[30] Rothman, P., van Dijk, D., Franses, P. H., 2001. Multivariate STAR analysis of money-output relationship. *Macroeconomic Dynamics, 5*(4), pp.506-532. DOI: 10.1017.S1365100500000444.

[31] Särkkä, S,. 2008. Unscented Rauch-Tung-Striebel smoother." *IEEE transactions on automatic control, 53*(3), pp.845-849. DOI: 10.1109/TAC.2008.919531.

[32] Schreft, S. L., Singh, A., 2003. A Closer Look at Jobless Recoveries. *Economic Review-Federal Reserve Bank of Kansas City, 88*(2), pp.45-73.

[33] Sinclair, T. M., 2009. The relationships between permanent and transitory movements in US output and the unemployment rate. *Journal of Money, Credit, and Banking, 41*(2-3), pp.529-542. DOI: 10.1111/j.1538-4616.2009.00220.x.

[34] Skalin, J., Teräsvirta, T. 2002. Modeling asymmetries and moving equilibria in unemployment rates. *Macroeconomic Dynamics, 6*(2), pp.202-241. DOI: 10.1017/S1365100502031024.

[35] Summers, L. H., 1986. Estimating the long-run relationship between interest rates and inflation: a response to McCallum. *Journal of Monetary Economics, 18*(1), pp.77-86. DOI: 10.1016/0304-3932(86)90055-3.

[36] van Dijk, D., Teräsvirta, T., Franses, P. H., 2002. Smooth transition autoregressive models—a survey of recent developments. *Econometric reviews, 21*(1), pp.1-47. DOI: 10.1081/ETC-120008723.

[37] Wan, E.A. and Van Der Merwe, R., 2001. The unscented Kalman filter. *Kalman filtering and neural networks*, pp.221-280. DOI: 10.1002/0471221546.ch7.

A Bayesian Estimation Methodology

A.1 Drawing $\boldsymbol{\Psi}$,λ conditional on $\gamma, \boldsymbol{\Sigma}$, and $\{z_t^*\}_{t=1}^T$

Conditional on $\pi(z_{t-1})$, drawing from the posterior distributions for the parameters of (2) is a straightforward application of Chib (1993) and Chib and Greenberg (1996). We can rewrite (2) in the following form:

$$\mathbf{y}_{0t} = \alpha_0 + \sum_{p=1}^{P_0} \theta_{0p}\mathbf{y}_{t-p} + \Gamma_0 (z_{t-1} - z^*) + \boldsymbol{\varepsilon}_{0t}$$

and

$$\mathbf{y}_{1t} = \alpha_1 + \sum_{p=1}^{P_1} \theta_{1p}\mathbf{y}_{t-p} + \Gamma_1 (z_{t-1} - z^*) + \boldsymbol{\varepsilon}_{1t},$$

$$\mathbf{y}_t = \alpha_0 + \sum_{p=1}^{P_0} \theta_{0p}\mathbf{y}_{t-p} + \Gamma_0 (z_{t-1} - z^*) + \pi(z_{t-1}) \left[d_1^\alpha + \sum_{p=1}^{P_1} d_{1p}^\theta \mathbf{y}_{t-p} + d_1^\Gamma (z_{t-1} - z^*) \right] + \varepsilon_t,$$

where $\boldsymbol{\varepsilon}_t \sim N(0, \Omega)$, $d_1^\alpha = \alpha_1 - \alpha_0$, $d_{1p}^\theta = \theta_{1p} - \theta_{0p}$, and $d_1^\Gamma = \Gamma_1 - \Gamma_0$. We use this specification to measure the change in coefficients across regimes. This allows for the sampler to be effective even if γ is small while imposing the identifying restriction that $\gamma > 0$. Then, we can stack $\mathbf{y}_t$ with the latent threshold, z_t^*. Thus, we define

$$\widetilde{\mathbf{y}}_t = \boldsymbol{\Sigma}^{-1/2} \begin{bmatrix} \mathbf{y}_t \\ z_t^* \end{bmatrix}$$

and

$$\widetilde{\mathbf{x}}_t = \boldsymbol{\Sigma}^{-1/2} \begin{bmatrix} \mathbf{x}_t & \mathbf{0}_{1\times 2} \\ \mathbf{0}_{(4+2P)\times 1} & \begin{bmatrix} 1 & z_{t-1}^* \end{bmatrix} \end{bmatrix}.$$

Given the prior $N(m_0, M_0)$, we define $\mathbf{X}$ and $\mathbf{Y}$ as the time-stacked vectors of $\widetilde{\mathbf{x}}_t$ and $\widetilde{\mathbf{y}}_t$, respectively. Then, the joint parameter vector can be drawn from

$$\begin{bmatrix} \boldsymbol{\Psi}' \\ \lambda \end{bmatrix} \sim N(\mathbf{m}, \mathbf{M}), \tag{14}$$

where

$$\mathbf{M} = \left(M_0^{-1} + \mathbf{X}'\mathbf{X}\right)^{-1} \tag{15}$$

and

$$\mathbf{m} = \mathbf{M}\left(M_0^{-1} m_0 + \mathbf{X}'\mathbf{Y}\right). \tag{16}$$

A.2 Drawing γ conditional on $\mathbf{\Psi}$,λ, $\mathbf{\Sigma}$, and $\{z_t^*\}_{t=1}^T$

Given a gamma prior, the posterior for γ is not a standard form; γ, however, can be drawn from an MH step. Then, a candidate, $\widehat{\gamma}$, can be drawn from a gamma proposal density:

$$\widehat{\gamma} \sim \Gamma\left(g_0, G_0\right),$$

and accepted with probability $a_g = \min\{A_g, 1\}$, where

$$A_g = \frac{\prod_t \phi\left(\mathbf{y}_t | \pi\left(z_{t-1} | \widehat{\gamma}, z_t^*\right), \mathbf{\Psi}, \mathbf{\Sigma}\right)}{\prod_t \phi\left(\mathbf{y}_t | \pi\left(z_{t-1} | \gamma^{[i]}, z_t^*\right), \mathbf{\Psi}, \mathbf{\Sigma}\right)} \times \frac{d\Gamma\left(\widehat{\gamma} | g_0, G_0\right)}{d\Gamma\left(\gamma^{[i]} | g_0, G_0\right)},$$

$\gamma^{[i]}$ represents the last accepted value of γ, $\phi(.)$ is the normal pdf, and $d\Gamma$ is the gamma pdf.

A.3 Drawing elements of $\mathbf{\Sigma}$ conditional on $\mathbf{\Psi}$,λ, γ, and $\{z_t^*\}_{t=1}^T$

To draw the components of the variance-covariance matrix Σ, we need to impose the restriction that the diagonal element corresponding to the variance of innovations to the threshold is small. We employ the method developed in Chan and Jeliazkov (2009) for a restricted Inverse-Wishart specification. Due to the positive-definite nature of the variance-covariance matrix, there exist unique matrices L and D such that $L\Sigma L' = D$ and $\Sigma^{-1} = L'D^{-1}L$. L is a lower-triangular matrix with $1's$ on the diagonal and a_{kj}, $1 \leq j < k \leq n+1$, as the free elements. D is a diagonal matrix with $\lambda_k > 0$ along the diagonal:

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{n+1} \end{bmatrix}, \; L = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{21} & 1 & 0 & \cdots & 0 \\ a_{31} & a_{32} & 1 & \cdots & \vdots \\ \vdots & \vdots & & \ddots & \\ a_{(n+1)1} & a_{(n+1)2} & \cdots & & 1 \end{bmatrix}. \tag{17}$$

We follow Chan and Jeliazkov (2009) and consider the following priors for λ_k and a where $a_k = [a_{k1}, ..., a_{k,k-1}]'$ and $a = \left[a_2', ..., a_{n+1}'\right]'$:

$$\begin{aligned} \lambda_k &\sim IG\left(v_{k0}/2, \delta_{k0}/2\right), \\ a \;\; &| \;\; \lambda \sim N\left(\mathbf{a}_0, \mathbf{A}_0\right). \end{aligned}$$

We implement the estimation procedure outlined in Chan and Jeliazkov (2009) and impose that $\lambda_{n+1} = 0.01$.[20]

B Adapting the State-Space to Construct Impulse Responses in Levels

In the first application for the natural rate of unemployment, we need to address how the observed unemployment rate enters into the model in both differences (as an element of

[20] See Chan and Jeliazkov (2009) for a full discussion of the econometric technique.

the VAR) and in levels (via the error-correction term). Therefore, when computing impulse responses, both of these dynamics must be captured as the shock propagates through the system. We adjust the coefficient matrices in the state-space representation of $\mathbf{y}_t$ in order to express how the levels of all variables in the VAR evolve following a shock. Recall that the VAR consists of monthly changes in the unemployment rate and CPI and the level of the 3-month T-Bill rate. The error-correction term and the threshold variable, z_{t-1}, use the level of the unemployment rate. We can write the the state-dependent processes as the following, suppressing the i notation to designate the state for clarity but acknowledge that these dynamics would apply in either $i = \{0, 1\}$:

$$\begin{aligned}\begin{bmatrix} y_{1t} - y_{t,t-1} \\ y_{2t} - y_{2,t-1} \\ y_{3t} \end{bmatrix} &= \begin{bmatrix} \theta_{11,1} & \theta_{12,1} & \theta_{13,1} \\ \theta_{21,1} & \theta_{22,1} & \theta_{23,1} \\ \theta_{31,1} & \theta_{32,1} & \theta_{33,1} \end{bmatrix} \begin{bmatrix} y_{1,t-1} - y_{t,t-2} \\ y_{2,t-1} - y_{2,t-2} \\ y_{3,t-1} \end{bmatrix} + \begin{bmatrix} \theta_{11,2} & \theta_{12,2} & \theta_{13,2} \\ \theta_{21,2} & \theta_{22,2} & \theta_{23,2} \\ \theta_{31,2} & \theta_{32,2} & \theta_{33,2} \end{bmatrix} \begin{bmatrix} y_{1,t-2} - y_{t,t-3} \\ y_{2,t-2} - y_{2,t-3} \\ y_{3,t-2} \end{bmatrix} \\ &\quad + ... + \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \\ \Gamma_3 \end{bmatrix} \left[y_{1,t-1} - y^*_{t-1} \right] + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix},\end{aligned}$$

where y_1 is the unemployment rate, y_2 is the CPI, y_3 is the interest rate, and y^*_{t-1} is the estimate of the time-varying latent threshold. Rewriting this in levels produces the following:

$$\begin{aligned}\begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{bmatrix} &= \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ 0 \end{bmatrix} + \theta_1 \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} - \theta_1 \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \\ 0 \end{bmatrix} + \theta_2 \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \\ y_{3,t-2} \end{bmatrix} - \theta_2 \begin{bmatrix} y_{1,t-3} \\ y_{2,t-3} \\ 0 \end{bmatrix} \\ &\quad + ... + \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \\ \Gamma_3 \end{bmatrix} \left[y_{1,t-1} - y^*_{t-1} \right] + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \\ &= \left[\tilde{I} + \theta_1 \right] \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \begin{bmatrix} \theta_{11,2} - \theta_{11,1} & \theta_{12,2} - \theta_{12,1} & \theta_{13,2} \\ \theta_{21,2} - \theta_{21,1} & \theta_{22,2} - \theta_{22,1} & \theta_{23,2} \\ \theta_{31,2} - \theta_{31,1} & \theta_{32,2} - \theta_{31,1} & \theta_{33,2} \end{bmatrix} \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \\ y_{3,t-2} \end{bmatrix} \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \phi_2 - \widetilde{\phi}_1 \\ &\quad + ... + \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \\ \Gamma_3 \end{bmatrix} \left[y_{1,t-1} - y^*_{t-1} \right] + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix},\end{aligned}$$

where $\tilde{I} = diag([1, 1, 0])$.

Ultimately, the state space for the level responses takes the form:

$$\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p} \end{bmatrix} = \begin{bmatrix} \tilde{I} + \Gamma + \theta_1 & \theta_2 - \widetilde{\theta}_1 & \cdots & -\widetilde{\theta}_p \\ I_3 & \mathbf{0}_{3\times3} & & 0 \\ \mathbf{0}_{3\times3} & I_3 & & 0 \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \mathbf{0}_{3\times1} \\ \vdots \\ \mathbf{0}_{3\times1} \end{bmatrix}. \tag{18}$$

In the second application, all variables enter the VAR in differences so the adjustments shown above to the last element in the VAR are not necessary.

C Sensitivity Analysis

We conducted some limited simulation experiments to explore the efficacy of our methodology when the true data-generating process has a fixed threshold. Our results suggest that the time-varying threshold STAR can produce an estimate of the constant threshold if the prior on the correlation between the threshold innovation and the VAR shocks, ρ, is rather tight. We simulate from a DGP using the (mean) model estimates from Application #1 but instead with a constant threshold. We then estimated the time-varying threshold model. In the empirical specification, the threshold innovation may be correlated with the shocks to the observable data,

$$Cov(\boldsymbol{\varepsilon}_t, u_t) = \rho,$$

where ρ is a vector of parameters that governs this cross-correlation. In our baseline application, the prior for ρ is $N(f_0, F_0)$ where $f_0 = 0$, and $F_0 = 1$. This works well when $\rho \neq 0$. When working with the simulated dataset with a constant threshold and $\rho = 0$, we found that a tighter prior—shrinking ρ closer to zero—was necessary. We experimented with two alternatives:

To assess the degree to which our baseline results may be sensitive to the prior on ρ, we re-estimated Application 1 using two alternative specifications for the prior: $F_0 = \{0.1, 0.01\}$. Figure 7 plots the unemployment rate (black), the posterior mean estimate of our baseline theshold ($F_0 = 1$, red), and the posterior mean estimates of the time-varying thresholds for $F_0 = 0.1$ (blue) and $F_0 = 0.01$ (green). In the first half of the sample, the thresholds produced from all three priors are similar. After the mid-1990s, we start to see some variation where the magnitude of the unemployment gap would be sensitive to the prior. As a gauge of relative fit of the threshold model itself, we considered the sign of the gap between z and the threshold for all periods, $[z_t - z_t^*]_{t=1}^T$. The sign of the unemployment gaps are the same for our baseline and the alternatives 76% ($F_0 = 0.1$) and 87% ($F_0 = 0.01$) of the time.

Table 1: Priors for Estimation

Parameter	Prior Distribution	Hyperparameters
$\{z_t^*\}_{t=1}^T$	MH draw via UKF proposal	n/a
$\alpha, \theta, \Gamma, \lambda$	$N(\mathbf{m}_0, \mathbf{M}_0)$	$m_0 = 0_{\kappa \times 1}$; $M_0 = I_\kappa^*$
γ	$\Gamma(g_0, G_0)$	$g_0 = 46$; $G_0 = 0.1$
ρ	$N(f_0, F_0)$	$f_0 = 0$; $F_0 = 1$
$\Sigma = \begin{bmatrix} \Omega & \rho \\ \rho & 0.01 \end{bmatrix}$	$\lambda_k \sim IG(v_{k0}/2, \delta_{k0}/2)$	$v_{k0} = 1$; $\delta_{k0} = 1$
$= L^{-1} D \left(L^{-1}\right)'$	$D = diag([\lambda_1, ..., \lambda_{N+1}]$	
$L = \begin{bmatrix} \mathbf{I}_N & \mathbf{0}_{N \times 1} \\ a & 1 \end{bmatrix}$	$a \mid \lambda \sim N(\mathbf{a}_0, \mathbf{A}_0)$	$a_0 = 0_{N \times 1}$; $A_0 = I_N$

Table 1: Parameterization of the Priors for Bayesian Estimation. For the normal prior on the VAR, error correction, and threshold coefficients: $\kappa = n(4 + P_0 + P_1) + m + 1$. *The diagonal elements of the variance term, $\mathbf{M}_0$, corresponding to the change in the VAR coefficients for regime 1 are set equal to 0.1 to contrain the amount of variation.

Table 2: State-Dependent Multipliers	
	Multiplier
2-year	
$\pi\left(z_{t-1}; z_{t-1}^*\right) < 0.2$	0.22
$\pi\left(z_{t-1}; z_{t-1}^*\right) > 0.8$	2.97
4-year	
$\pi\left(z_{t-1}; z_{t-1}^*\right) < 0.2$	0.28
$\pi\left(z_{t-1}; z_{t-1}^*\right) > 0.8$	1.90

Table 2: State-dependent government spending multipliers computed relative to an exogenous news shock. We condition the generalized impulse responses on the weight placed upon the slack regime being low (tight labor markets) versus high (slack labor markets). The table displays the 2-year and 4-year cumulative multipliers based upon the behavior of output relative to government spending in response to the news shock. Details regarding the calculation of the multiplier are provided in Section 5.2.

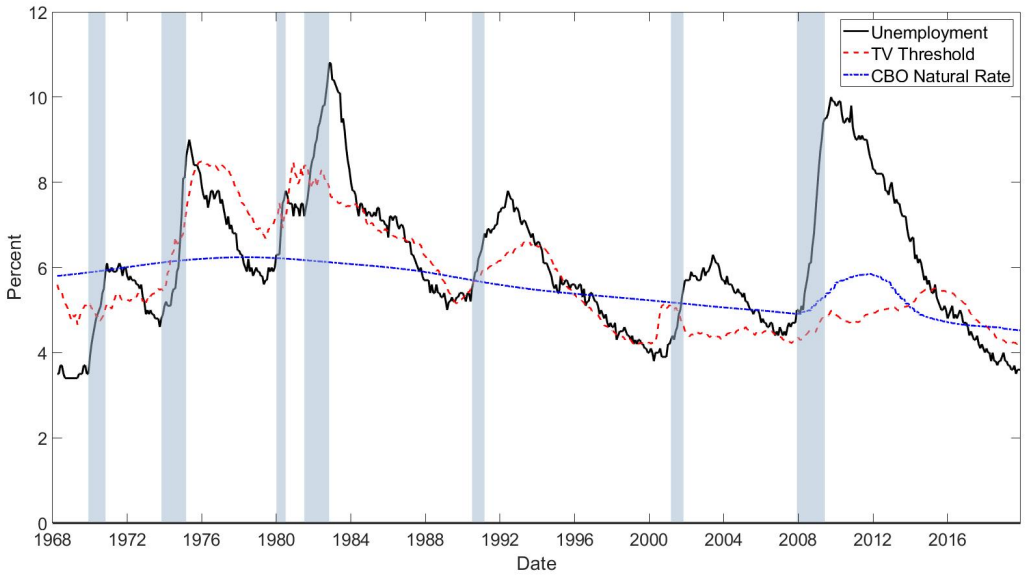

Figure 1: Latent threshold estimated in Application 1: Observed data on the unemployment rate, the CBO estimate of the natural rate, and the posterior mean estimate of the latent threshold.

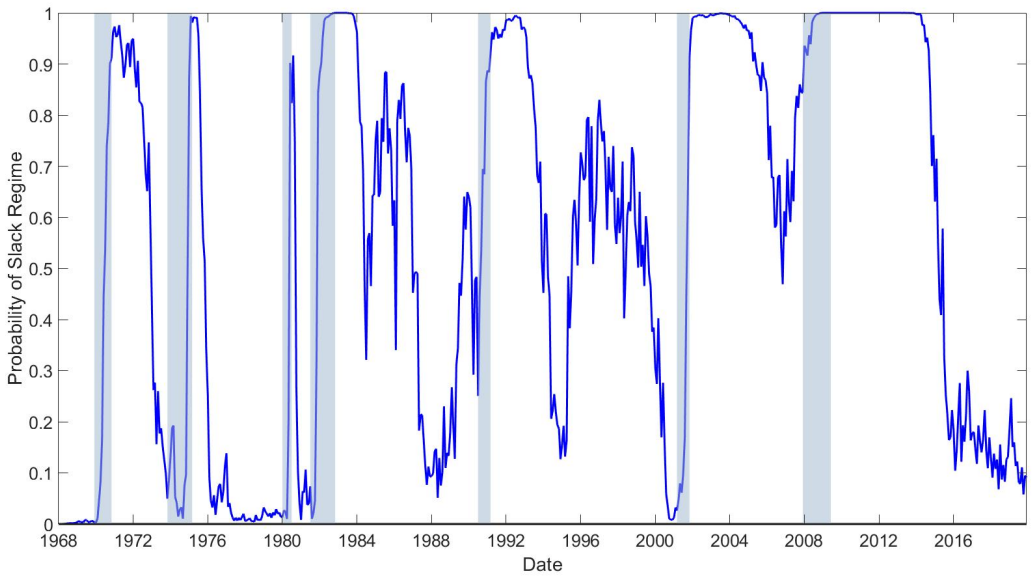

Figure 2: Posterior Mean Estimates of the Regime Probabilities for Time-Varying Threshold Model in Application 1. The time series plots the model-implied estimates for the values of the transition function in each period.

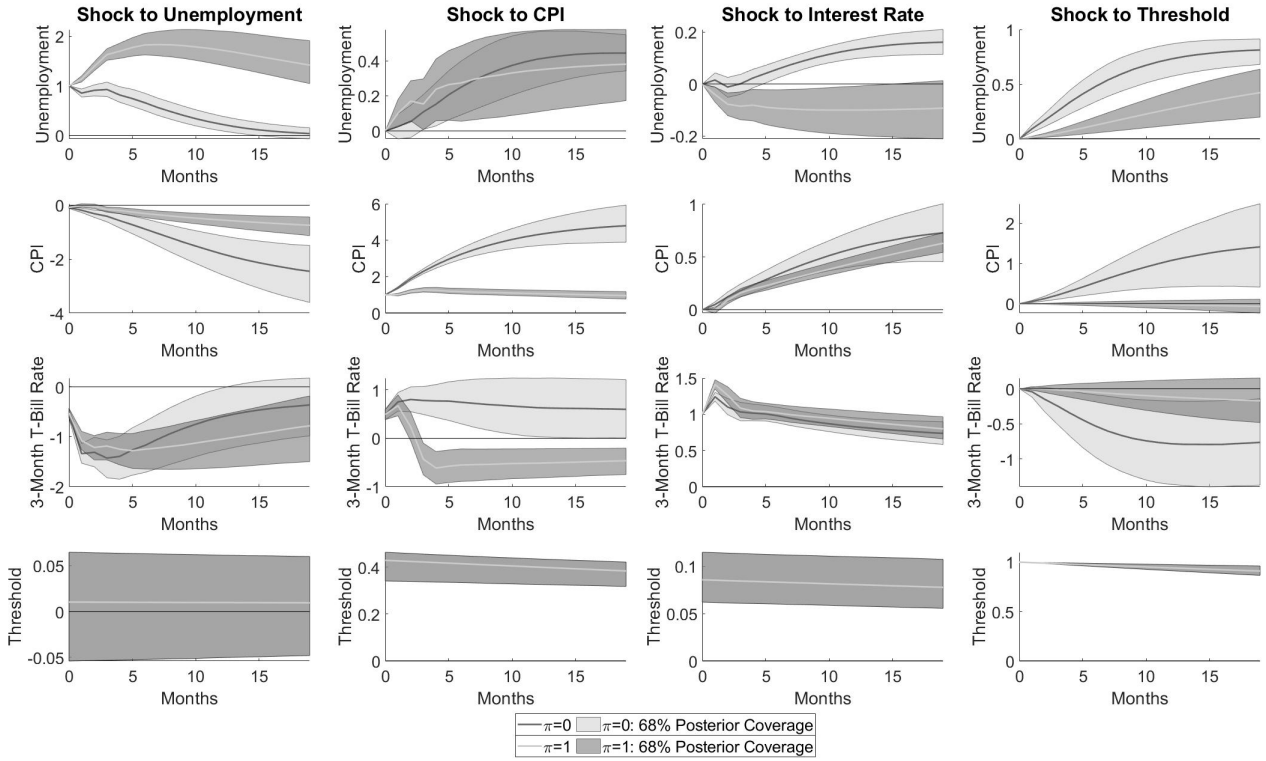

Figure 3: Regime-Dependent Impulse Responses for Application 1. This figure shows the responses to shocks for all variables within the model, conditional on being completely in one regime versus the other (i.e. unemployment being above or below the threshold). The regime at the time of the shock is held constant for the life of the response.

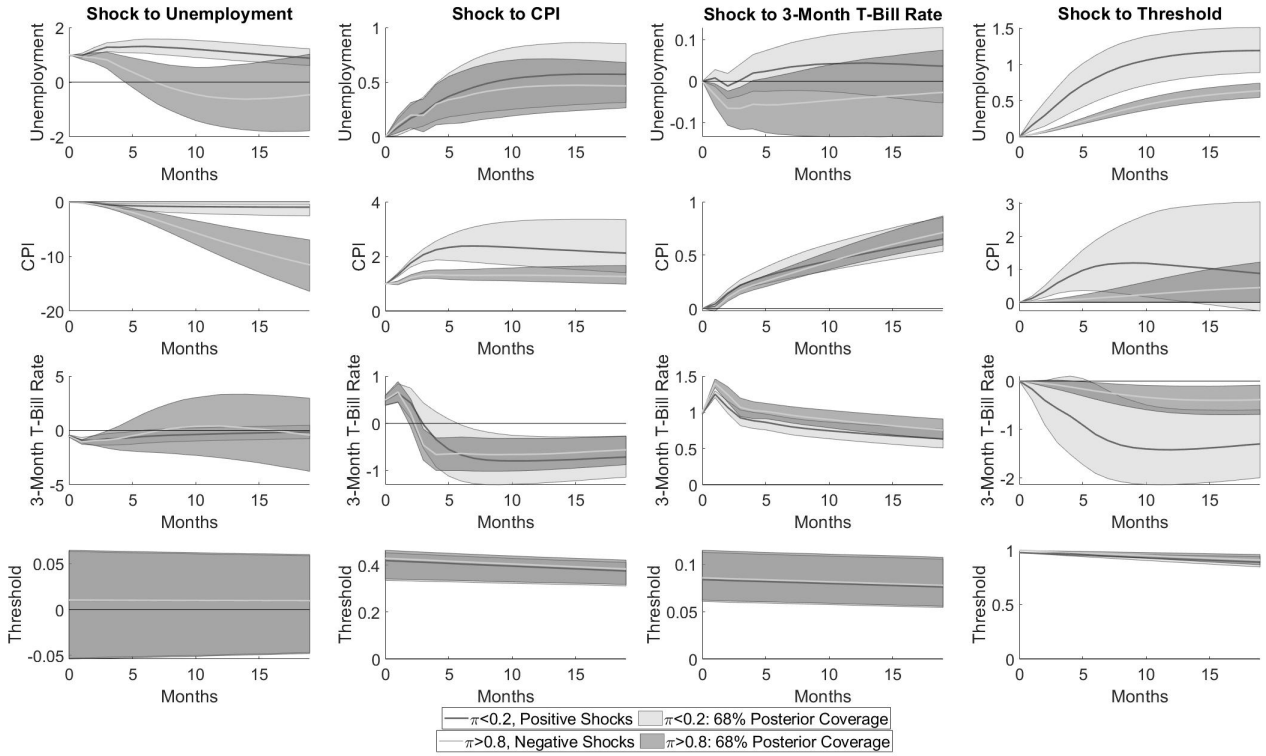

Figure 4: Generalized Impulse Reponses for Application 1. The GIRF conditions on the weight placed upon the slack regime being low (tight labor markets) versus high (slack labor markets). We consider the response of macro variables within the model, accounting for the chance that unemployment may cross the threshold and the probability of each regime will be time-dependent. For initial conditions when the probability of regime 1 is less than 0.2, we consider positive shocks to all variables. When conditioning on the probability of regime 1 being greater than 0.8, we consider negative shocks and invert the responses for comparison.

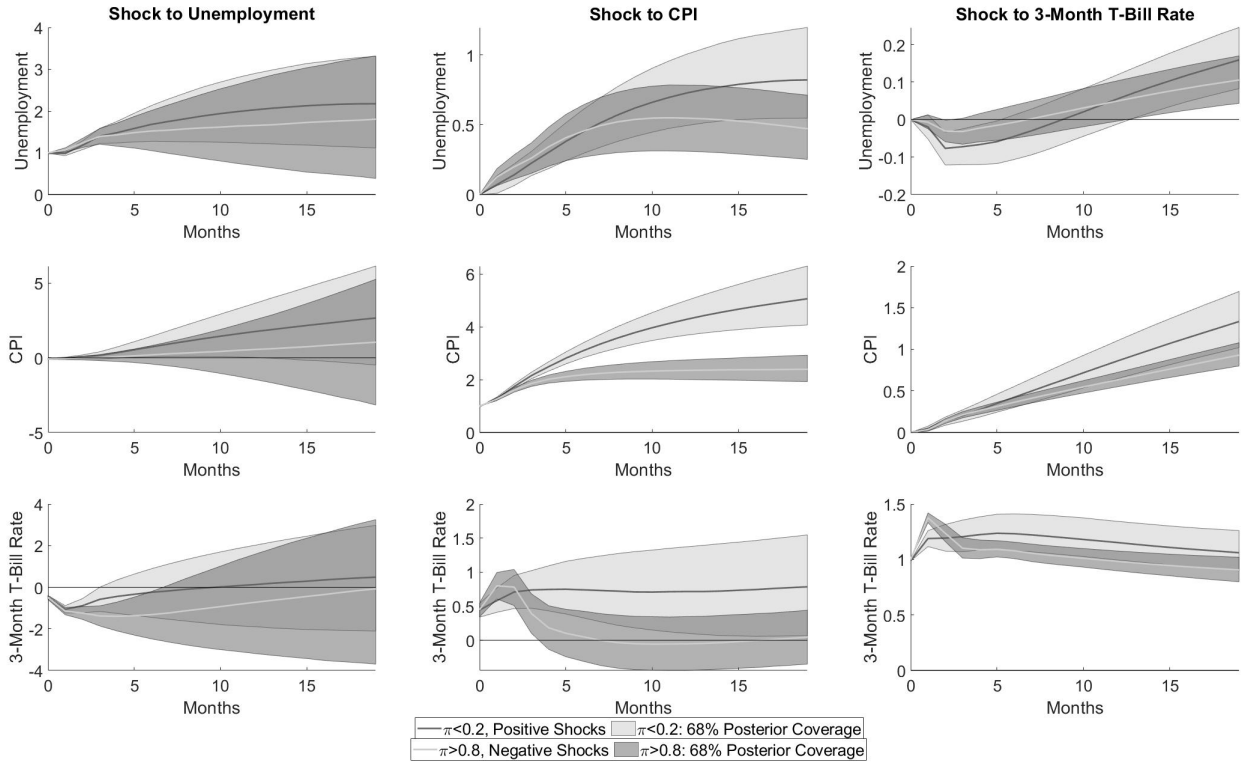

Figure 5: Generalized Impulse Reponses for Application 1 with a Constant Threshold. This alternative fixes the threshold at a constant equal to the mean value within the sample, 6.1 percent. The GIRF conditions on the weight placed upon the slack regime being low (tight labor markets) versus high (slack labor markets). We consider the response of macro variables within the model, accounting for the chance that unemployment may cross the threshold and the probability of each regime will be time-dependent. For initial conditions when the probability of regime 1 is less than 0.2, we consider positive shocks to all variables. When conditioning on the probability of regime 1 being greater than 0.8, we consider negative shocks and invert the responses for comparison.

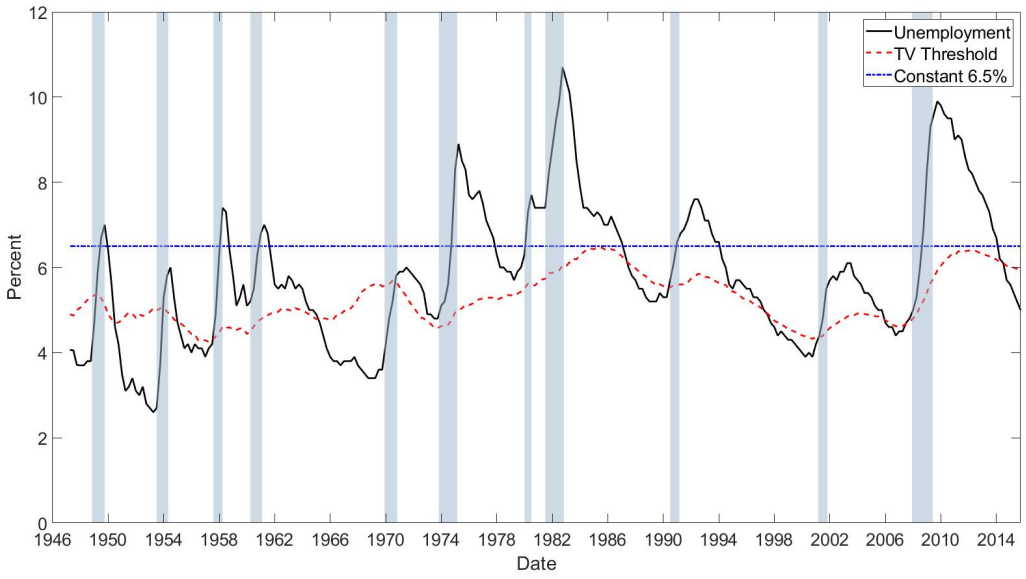

Figure 6: Latent threshold produced in Application 2: Observed data on the unemployment rate, the posterior mean estimate of the latent threshold, and a constant 6.5-percent level.

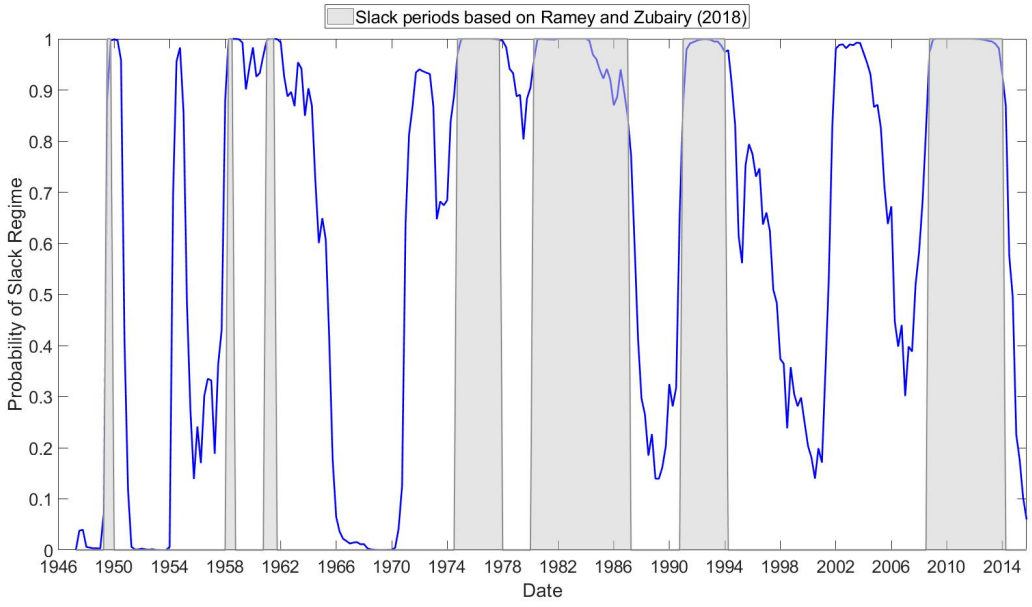

Figure 7: Posterior Mean Estimates of the Regime Probabilities for Time-Varying Threshold Model in Application 2. The time series plots the model-implied estimates for the values of the transition function in each period. The shaded areas represent the periods of slack computed using the RZ constant threshold of 6.5 percent.

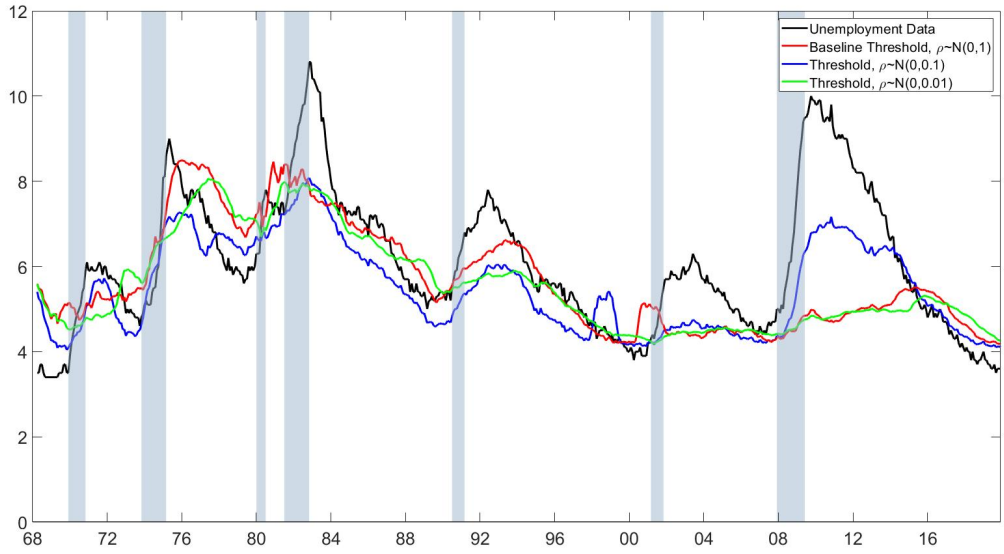

Figure 8: Latent threshold estimated with data from Application 1 in the paper utilizing the same alternative priors as we employed for the simulation exercise.