Research Division
Federal Reserve Bank of St. Louis
*Working Paper Series*

# A Time-Varying Threshold STAR Model of Unemployment and the Natural Rate

**Michael Dueker**
**Michael T. Owyang**
**and**
**Martin Sola**

Working Paper 2010-029A
http://research.stlouisfed.org/wp/2010/2010-029.pdf

September 2010

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

---

# A Time-Varying Threshold STAR Model of Unemployment and the Natural Rate*

Michael Dueker†
Russell Investments

Michael T. Owyang‡
Research Division
Federal Reserve Bank of St. Louis

Martin Sola
Department of Economics
Universidad Torcuato di Tella and Birkbeck College, University of London

First Draft: February 2010
This Draft: September 20, 2010

**Abstract**

Smooth-transition autoregressive (STAR) models have proven to be worthy competitors of Markov-switching models of regime shifts, but the assumption of a time-invariant threshold level does not seem realistic and it holds back this class of models from reaching their potential usefulness. Indeed, an estimate of a time-varying threshold level of unemployment, for example, might serve as a meaningful estimate of the natural rate of unemployment. More precisely, within a STAR framework, one might call the time-varying threshold the "tipping level" rate of unemployment, at which the mean and dynamics of the unemployment rate shift. In addition, once the threshold level is allowed to be time-varying, one can add an error-correction term—between the lagged level of unemployment and the lagged threshold level—to the autoregressive terms in the STAR model. In this way, the time-varying latent threshold level serves dual roles: as a demarcation between regimes and as part of an error-correction term.

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

*The authors benefitted from conversations with Jim Hamilton, Tatevik Sekhposyan, and Tara Sinclair as well as comments from participants at the 2009 Conference on Business Cycle Models at the University of California, Riverside and the 17th Annual Meeting of the Society of Nonlinear Dynamics and Econometrics. Kristie M. Engemann provided research assistance. The view expressed herein are those of the authors alone and do not reflect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

†mdueker@russell.com
‡owyang@stls.frb.org

# 1 Introduction

The unemployment rate is a key macroeconomic indicator and is the foundation of a number of theoretical or empirical relationships such as the Phillips curve and Okun's Law. An important element of these relationships is the so-called natural rate of unemployment. The natural rate is often defined as the non-accelerating inflation rate of unemployment—that is, the rate of unemployment which applies no upward or downward pressure on inflation. The natural rate, however, is an unobserved quantity, typically estimated in multivariate models as the trend or the permanent component of unemployment.

Despite (or perhaps because of) its potential significance for policy discussions, recent studies have questioned the validity of the natural rate and, in particular, the Phillips curve relationship [e.g., Galí, Gertler, and López-Salido (2005); Ball and Mankiw (2002)]. The Phillips curve is an apparent empirical relationship between inflation and the unemployment rate in which, at the natural rate, inflation is stable. Thus, in order to validate the Phillips curve relationship, one would need to accurately estimate the natural rate, requiring in turn an empirical model of the unemployment rate.

Many classes of empirical models have been estimated with the unemployment rate, often with the goal of obtaining the natural rate. First, smooth-transition autoregressions (STAR) have been used extensively to study unemployment [e.g., Skalin and Teräsvirta (2002); Deschamps (2008)]. The smooth transition models have the appealing feature that the threshold rate of unemployment may determine the evolution of inflation. Second, trend-cycle decompositions (and unobserved components) can separate the permanent and temporary components of unemployment [e.g., Clark (1989); Jaeger and Parkinson (1994); Sinclair (2009); Basistha and Nelson (2007); Doménech and Gómez (2006)]. Standard unobserved components models often identify the permanent component as that which evolves as a unit root. The transitory component, then, is the stationary portion of the decomposition often interpreted as the cyclical component. Finally, Markov-switching models can model a shifting relationship between inflation and unemployment [e.g., Phelps and Zoega (1998); Chauvet, Juhn, and Potter (2001); Owyang (2001)].

One of the benefits of the latter two models is that the natural rate can move across time as the state of the economy changes. A number of studies have found evidence in favor of a time-varying natural rate [see, for example, Summers (1986); Gordon (1997, 1998); Ball and Mankiw (2002); many others]. 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.

In this paper, we propose an extension of the STAR model [Granger and Teräsvirta (1993)] which allows for potential time-variation of the natural rate. In particular, we construct a STAR model in which the latent threshold is both endogenous and time-varying. In addition, we incorporate an error-correction term which depends on the deviation of unemployment from the latent threshold. An estimate of a time-varying threshold level of unemployment, for example, might serve as a meaningful estimate of the natural rate of unemployment. More precisely, within a STAR framework, one might call the time-varying threshold the "tipping level" rate of unemployment, at which the mean and dynamics of the unemployment rate shift. In addition, once the threshold level is allowed to be time-varying, one can add an error-correction term—between the lagged level of unemployment and the lagged threshold level—to the autoregressive terms in the STAR model. In this way, the time-varying latent threshold level serves dual roles: as a demarcation between regimes and as part of an error-correction term.

In the literature, there are models with multiple levels for the threshold in multiple-regime models, but the thresholds themselves are constant across time. In terms of a time-varying threshold, only special cases 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. This approach is akin to older models of heteroscedasticity, where the variance was posited to be a function of an observable variable.

Because the type of time-varying threshold we consider makes maximum likelihood estimation impossible, we estimate the model in a Bayesian environment. Lopes and Salazar (2006) proposed using MCMC to estimate smooth transition models. They use a joint 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 because of both the time-variation in the threshold and because the threshold appears as the latent attractor in the error-correction term. In addition, the nonlinearity in the transition function requires a more complicated proposal density; in this case, we use the unscented Kalman filter—a nonlinear version of the classic Kalman filter—to obtain the hyperparameters of a normal proposal.

We find our time-varying threshold tracks unemployment in a manner similar to an HP-filtered trend. However, our threshold is typically smoother than the filtered trend. In addition, we find a negative relationship between the inflation rate and the cyclical component of unemployment, defined as the difference between the actual unemployment rate and the threshold. This finding is consistent with a Phillips curve-type relationship.

The rest of the paper develops as follows: Section 2 presents the STAR model with the endogenous, 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. Section 4 presents the empirical results. Section 5 discusses the results in the context of the time varying Phillips curve models. Section 6 concludes.

## 2 The Time-varying Threshold Model

Let $\mathbf{y}_t$ be an $(n \times 1)$ vector of variables of interest. Define two separate $p$-lag processes, $\mathbf{y}_{0t}$ and $\mathbf{y}_{1t}$, as

$$\mathbf{y}_{0t} = \alpha_0 + \sum_{i=1}^{p_0} \theta_{0i} \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_{0t}$$

and

$$\mathbf{y}_{1t} = \alpha_1 + \sum_{i=1}^{p_1} \theta_{1i} \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_{1t},$$

where $\boldsymbol{\varepsilon}_{0t} \sim N(\mathbf{0}_n, \Omega_0)$ and $\boldsymbol{\varepsilon}_{1t} \sim N(\mathbf{0}_n, \Omega_1)$. Then, the standard smooth-transition autoregression model [Granger and Teräsvirta (1993)], STAR($p$), can be thought of as a combination of these two parameterized processes:

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

where $\pi(z_{t-1})$ is a transition function, $p = \max\{p_0, p_1\}$. The transition function determines the weights of each autoregression on the path of $\mathbf{y}_t$ in terms of the two latent regimes defined by $\pi(z_{t-1}) = 0$ and $\pi(z_{t-1}) = 1$. The path of the economy is determined by $\pi(z_{t-1})$, which is bounded by zero and one and a function of past values of the driving variable, $z_{t-1}$. When $\pi(z_{t-1}) = 0$, the steady-state $\mathbf{y}_t$ is $\alpha_0 \left[1 - \sum_{i=1}^{p} \theta_{0i}\right]^{-1}$. Conversely, when $\pi(z_{t-1}) = 1$, the steady-state $\mathbf{y}_t$ is $\alpha_1 \left[1 - \sum_{i=1}^{p} \theta_{1i}\right]^{-1}$.

Figure 1 (transition function) about here

The transition function depends on the threshold variable, $z_{t-1}$, which influences the state of the economy. While the transition function can take a number of forms, we assume the following representation[1]:

$$\pi(z_{t-1}) = \Phi\left(\gamma\left(z_{t-1} - z^*\right)\right), \tag{2}$$

where $\gamma$ is a slope parameter, $z^*$ is a (fixed) threshold, and $\Phi(.)$ is the normal cdf. To normalize the regimes, we impose $\gamma < 0$. In (2), the regime process is determined by the sign of the deviation of $z_{t-1}$ from the threshold $z^*$. If $z_t$ is less than $z^*$, the transition function, $\pi(z_{t-1})$, moves further toward the first regime (in this case, normalized to $\pi(z_{t-1}) = 1$). Figure 1 plots a number of parameterizations of the transition function for fixed $z^*$. Increasing the magnitude of $\gamma$ increases the transition speed between regimes.

Figure 2 (3D time-varying transition function) about here

We are interested in the case in which the threshold parameter, $z^*$, varies over time. This might occur in a number of cases. For example, in a model of unemployment, we may believe the natural rate is time varying. Alternatively, we could be modeling a time-varying inflation target in a Taylor rule. We can rewrite (2) to account for this potential time-variation:

$$\pi\left(z_{t-1}; z_{t-1}^*\right) = \Phi\left(\gamma\left(z_{t-1} - z_{t-1}^*\right)\right), \tag{3}$$

where $\Phi(.)$ is the normal cdf. In addition, we need to specify a process for the evolution of the time-varying threshold. Figure 2 plots the transition function with time-varying $z_t^*$. Suppose that the time-varying threshold is an autoregressive process:

[1] 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.

$$z_t^* = \lambda_0 + \sum_{j=1}^{m} \lambda_j z_{t-j}^* + u_t, \tag{4}$$

where $u_t$ is normalized to have unit variance.[2] The threshold innovation may be correlated with the shocks to the observable data,

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

where $\rho$ 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

$$v_t \sim N\left(0, \begin{bmatrix} \Omega & \rho \\ \rho & 1 \end{bmatrix}\right).$$

Because we allow for a time-varying threshold, we can introduce an error-correction term into the threshold autoregression:

$$\begin{aligned} \mathbf{y_t} \quad = \quad & \pi\left(z_{t-1}; z_{t-1}^*\right)\left[\alpha_0 + \sum_{i=10}^{p} \theta_{0i}\mathbf{y}_{t-i} + \Gamma_0\left(z_{t-1} - z_{t-1}^*\right)\right] \\ & + \left[1 - \pi\left(z_{t-1}; z_{t-1}^*\right)\right]\left[\alpha_1 + \sum_{i=1}^{p} \theta_{1i}\mathbf{y}_{t-i} + \Gamma_1\left(z_{t-1} - z_{t-1}^*\right)\right] + \varepsilon_t, \end{aligned} \tag{5}$$

where $\Gamma_0$ and $\Gamma_1$ are the error-correction coefficients. Given the specification of the transition function, we have a regime change when $z_t$ rises above or falls below from $z_t^*$. If the first variable in $\mathbf{y}$ is the first difference of the threshold variable $y_{1t} = \Delta z_t$, then there is a natural interpretation to the error correction term. Essentially, the error correction term pushes $z_t$ back toward the threshold.

## 3 Estimation

Because the time-varying latent threshold enters the regime weights, $\pi(z_{t-1})$ and $1 - \pi(z_{t-1})$, the model cannot be estimated by maximum likelihood. It is also not possible to derive exact

[2] The threshold could also be a function of exogenous shifters as in Dueker et al. (2010).

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 key to the MH step is to find a tractable proposal density as close as possible to the target distribution for the latent threshold series, $\{z_t^*\}_{t=1}^T$.

The algorithm partitions the set of model parameters into four 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) $\lambda$, the coefficients in the latent variable equation; (3) $\mathbf{\Sigma} = [\Omega, \rho]$, the elements in the variance-covariance matrix for joint system; (4) $\gamma$, 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. 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 following subsections discuss the methods used to generate draws from each parameter block's conditional distribution.[3] The sampler is then executed for 10,000 iterations, discarding the first 5,000 before forming the joint posterior.

### 3.1 Drawing $\mathbf{\Psi}$,$\lambda$ conditional on $\gamma, \mathbf{\Sigma}$, and $\{z_t^*\}_{t=1}^T$

Conditional on $\pi(z_{t-1})$, drawing from the posterior distributions for the parameters of (6) is a straightforward application of Chib (1993) and Chib and Greenberg (1996). We can rewrite (5) in the following form:

$$\mathbf{y}_t = \mathbf{\Psi}\mathbf{x}_t + \varepsilon_t, \tag{6}$$

where

$$\mathbf{x}_t = \left[ \begin{array}{cc} \pi(z_{t-1})\,\mathbf{y}_t^\dagger & (1 - \pi(z_{t-1}))\,\mathbf{y}_t^\dagger \end{array} \right],$$

and

$$\mathbf{y}_t^\dagger = \left[1, \mathbf{y}_{t-1}, ..., \mathbf{y}_{t-p}, \left(z_{t-1} - z_{t-1}^*\right)\right].$$

[3] Appendix 1 contains the parameterization for the prior.

Then, we can stack these with the latent threshold, $z_t^*$. Then, define

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

and

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

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

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

where

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

and

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

### 3.2 Drawing $\gamma$ conditional on $\mathbf{\Psi}$,$\lambda$, $\mathbf{\Sigma}$, and $\{z_t^*\}_{t=1}^T$

The draw of the transition function parameters is slightly more problematic. Given a uniform prior, the posterior for $\gamma$ is intractable. However, $\gamma$ can be drawn from an MH step. Then, a candidate, $\widehat{\gamma}$, can be drawn from the proposal density:

$$\widehat{\gamma} \sim U(\kappa_1, \kappa_2),$$

$\kappa_1 < \kappa_2 < 0$, and accepted with probability $a_g = \min\{A_g, 1\}$, where

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

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

### 3.3 Drawing $\mathbf{\Omega}, \boldsymbol{\rho}$ conditional on $\mathbf{\Psi}$,$\lambda$, $\gamma$, and $\{z_t^*\}_{t=1}^T$

To draw the variance-covariance matrix of the observables equation, $\mathbf{\Omega}$, suppose $\mathbf{\Omega}^{-1}$ has the prior $W\left(\frac{\varpi_0}{2}, \frac{D_0}{2}\right)$. Then, the posterior is

$$\mathbf{\Omega}^{-1} \sim W\left(\frac{\varpi_0 + T}{2}, \frac{D_0 + \varepsilon_t \varepsilon_t'}{2}\right),$$

where $W(.,.)$ represents the Wishart distribution and $\varepsilon_t = \mathbf{y}_t - \pi(z_{t-1})\,\mathbf{y}_{0t} - [1 - \pi(z_{t-1})]\,\mathbf{y}_{1t}$. Once we obtain a draw for the observables variance-covariance matrix, the correlations between the observables errors and the threshold errors, $\rho_i = Cov(\varepsilon_{it}, u_t)$, can be drawn. Define the vector $\varphi$ such that

$$u_t = \varphi' \boldsymbol{\varepsilon}_t + \nu_t,$$

where $\nu_t \sim N(0, 1)$. Then, if $\varphi$ has prior $N(f_0, F_0)$, it can be drawn from $N(\mathbf{f}, \mathbf{F})$, where

$$\mathbf{F} = \left(F_0^{-1} + \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon}\right)^{-1}$$

and

$$\mathbf{f} = \mathbf{F}\left(F_0^{-1} f_0 + \boldsymbol{\varepsilon}' \mathbf{u}\right).$$

### 3.4 Drawing $\{z_t^*\}_{t=1}^T$ conditional on $\mathbf{\Psi}$,$\gamma$, $\lambda$, and $\mathbf{\Sigma}$

The last step in the sampler is to draw the series of latent thresholds. This draw, as indicated above, is complicated by the fact that the posterior distribution is intractable and by the fact that the transition function is nonlinear. We can solve the first problem with an MH step; however, the choice of the proposal density in the MH step is made more difficult by the nonlinearity. In some nonlinear models [e.g., Hamilton's (1989) Markov-switching model], the posterior distribution for the latent variable can be computed analytically via the Kalman filter. We take a similar approach here, accounting for the nonlinearity by employing a variation of the Kalman filter using the unscented transformation.

We can rewrite the model as the state equation in a state-space representation:

$$\mathbf{x_t} = \mathbf{g}(\mathbf{x_{t-1}}) + \epsilon_t, \tag{7}$$

where the nonlinear function $\mathbf{g}(.)$ contains, among others, the transition function. Given this state-space representation, inferred values for the state variable $\mathbf{x}_t$ — which here includes the data of interest, $\mathbf{y}_t$; the driving data, $z_t$; and the latent threshold $z_t^*$ — can be obtained from the unscented Kalman filter (UKF). The UKF is a nonlinear filter that serves as an alternative to the extended Kalman filter, which uses first-order Taylor-series approximations to any nonlinear functions in the measurement and transition equations. Instead, the UKF tracks the state variable by computing its distribution across a set of deterministic points called sigma points.[4]

The unscented transformation applies the nonlinear function $\mathbf{g}(.)$ on each of the sigma points. The corresponding prediction or update in the Kalman filter is taken as the weighted sum of the transformed sigma points. Let

$$\boldsymbol{\chi}^p = \begin{cases} \mathbf{x} & , \; for \quad p = 0 \\ \mathbf{x} + \left(\sqrt{(L+\zeta)\,P}\right)_p & , \; for \quad p = 1, ..., L \\ \mathbf{x} - \left(\sqrt{(L+\zeta)\,P}\right)_{p-L} & , \; for \quad p = L+1, ..., 2L \end{cases} \tag{8}$$

define the sigma points for any stage of the filter, where $\zeta = a^2(L+\kappa) - L$. Here, $a$ and $\kappa$ 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 $i$th column of the lower triangular Cholesky factorization of the square matrix $X$. The matrix $P$ is the uncertainty surrounding the state vector $\mathbf{x}$. Given the set of sigma points, we propagate $\{\boldsymbol{\chi}^p\}$ through the function $\mathbf{g}(.)$ to recover a prediction for each of the sigma points:

$$\widetilde{\boldsymbol{\chi}}^p = \mathbf{g}(\boldsymbol{\chi}^p), \quad \text{for} \;\; p = 0, ..., 2L.$$

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

The predicted state vector, $\widetilde{\mathbf{x}}$, is the weighted sum of the $\widetilde{\boldsymbol{\chi}}^p$'s. The updated state vector is obtained from propagated predicted state vector. Details of the algorithm are provided in the appendix.

The series of latent thresholds, $\{z_t^*\}_{t=1}^T$, are elements of the state vector, $\mathbf{x}_t$. The UKF yields smoothed estimates of the state vector for (7), $\widehat{\mathbf{x}}_t^s$, and its uncertainty, $P_t^s$, for each time period. This output of the UKF can be thought of as the hyperparameters of a proposal density for the MH algorithm used to draw the latent thresholds. Given $\widehat{\mathbf{x}}_t^s$ and $P_t^s$, the candidate $\widehat{z}_t^*$ can be drawn as a subvector of $\mathbf{x_t}$:

$$\widehat{z}_t^* \sim N\left(\widehat{\mathbf{x}}_{zt}^s, P_{zt}^s\right),$$

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

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

where, as above, the superscript $[i]$ represents the past accepted value.

# 4 An Application to the U.S. Unemployment Rate

We estimate the system, (3), (4), and (5), using the Bayesian algorithm described above. For our application, we employ the unemployment rate, inflation, and a short term interest rate. In addition, we choose one lag for the threshold evolution (i.e., $m = 1$) and three lags for the VAR (i.e., $p = 3$).

## 4.1 Data

Figure 3 (data) about here

Estimation of the model requires two sets of possibly overlapping data. First, we require a set of observables for $\mathbf{y}_t$. In this case, we use the change in the monthly unemployment rate, inflation, and an interest rate. The unemployment rate is the annualized monthly, seasonally-adjusted measure taken from the BLS's payroll employment survey. Inflation is the annualized seasonally-adjusted monthly rate of change of the CPI prices, excluding food and energy. Finally, the interest rate is the 30-day T-Bill. The full sample of data begins in January 1968 and ends in February 2010. Next, we must choose the variable, $z_{t-1}$, which will determine the regime process. In the literature on STAR models, it is standard to use one or more of the set of observables—the self-exciting threshold models. In this case, we use the first lag of the unemployment rate. Figure 3 plots the data for reference.

## 4.2 Results

Figure 4 (threshold, u rate include HP filtered trend) about here

Figure 4 shows the estimated threshold, along with the unemployment rate and its HP-filtered trend.[5] Both the latent threshold and the Hodrick-Prescott-filtered trend exhibit fluctuations similar to those from the unemployment rate. Obviously, both the threshold and the HP trend are smoother than the unemployment rate. Additionally, though, the threshold appears to lead the HP trend unemployment rate. The threshold level is smoother than the HP trend, however, such that deviations of the threshold from the unemployment rate larger than those for the HP trend.

Figure 5 (posterior regimes) about here

Figure 5 shows the estimated regime probabilities, the values of the transition function for each period. For the most part, the posterior probability of being in regime 1 increases around NBER recessions and decreases during expansions. For the period following the 1990 and 2001 recessions, however, the posterior regime probability remains high, reflecting the jobless recoveries. Obviously, given (3), regime 1 should roughly correspond to periods in which unemployment is greater than the threshold.

Table 2 (posterior distributions) about here

Table 2 presents the means of the posterior distributions for the parameters from the estimation of the model with the unemployment rate, inflation, and a short-term interest rate. We note a

[5] For our application, the use of the unscented Kalman filter as a proposal density yielded acceptance rates of about 60 percent.

few of the features for some of the estimated parameters of interest. First, the effect of regime switching in this model appears to manifest mostly in the interest rate response to the gap between unemployment and the threshold. Monetary policy (or the expectation of future monetary policy) is more sensitive to the gap in regime 1, the periods consistent with NBER recessions. Thus, during "ordinary times", monetary policy does not respond to the difference between unemployment and the threshold. Second, the threshold appears to be closely approximated by a random walk with drift – at the very least, the threshold is governed by a highly persistent process. Third, the threshold is negatively correlated with other shocks. While the correlation with unemployment and inflation is weak, an increase interest rate yields decrease in the threshold. This feature is especially obvious when considering the regime-dependent impulse responses.

Figures 6 and 7 (impulse responses) about here

From the state-space model described above, we can compute regime-dependent impulse responses, i.e., VAR responses computed as if the system were regime-invariant. In order to identify the shocks and compute their associated impulse responses, we must impose some identifying restrictions. In this case, we impose a Cholesky ordering with the latent threshold ordered last. Unemployment is ordered first, inflation is second, and interest rates are third. This implies unemployment and inflation do not contemporaneously respond to interest rates and no variable can respond contemporaneously to the latent threshold.

Figures 6 and 7 depict the regime-dependent responses of the three variables and the latent threshold to the various model shocks. The first two columns present the responses to unemployment and inflation shocks. The two shocks have opposite but expected effects on interest rates (i.e., monetary policy). When unemployment rises, interest rates fall; on the other hand, when inflation rises, interest rates rise. The last panel in the third column in each of the figures depicts the response of the latent threshold to a contractionary monetary shock. The model produces inflation and unemployment responses consistent with the VAR literature, including the price puzzle. In both regimes, the contractionary monetary shock produces a relatively persistent decline in the threshold. Suppose that unemployment is below the threshold. Then, increasing interest rates (perhaps through expectations) appears to shrink the wedge between unemployment and the threshold through both a change in the unemployment rate (small) and a decrease in the threshold.

## 5 Implications for the Phillips Curve and the Natural Rate

A simple expectations-augmented Phillips curve relates inflation, $\pi_t$, and unemployment, $U_t$, as

$$\pi_t - \pi_t^e = -b\,(U_t - U_t^*) + v_t,$$

where $\pi_t^e$ are inflation expectations and $U_t^*$ is the natural rate. If one begins with such a model, the natural rate can be identified by imposing statistical restrictions [see Staiger, Stock, and Watson (1997)]. In particular, the cyclical component of the unemployment rate should be mean zero and the should be inversely related to the gap between inflation and inflation expectations. Instead of imposing such a relationship, we ask whether the time-varying threshold identified in the previous section has properties consistent with those implied by the Phillips curve for the natural rate.

Figure 8 (threshold, cyclical component, inflation) about here

Figure 8 plots the inflation rate alongside the cyclical unemployment rate defined by $z_t - z_t^*$. Almost by definition, the cyclical unemployment rate is mean zero, consistent with the first restriction identifying the natural rate. Because the regimes roughly coincide with the periods in which the unemployment rates is above or below the threshold, it is illuminating to return to the regime-dependent impulse responses. This time, we consider the last columns in Figures 6 and 7—the responses to a shock to the threshold. In particular, we highlight the variation across the regimes for inflation and the interest rate. When unemployment is below the threshold (regime 0), a positive shock to $z^*$ widens the gap and leads to an increase in both inflation and a contractionary rise in interest rates. Conversely, when unemployment is above the threshold, a positive shock to $z^*$ closes the gap and produces the opposite responses to both inflation and unemployment. Thus, in these ways, the threshold does appear to have features consistent with a Phillips-curve notion of the natural rate.

# 6 Conclusions

The standard smooth transition autoregression is a popular alternative to regime 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 that we estimate the model with Bayesian methods, including a nonlinear alternative to the Kalman filter to generate proposals for our MH step. We applied the model to the change in the unemployment rate, inflation, and a short-term interest rate, assuming that the unemployment rate governed the regime process.

We found that the estimated time-varying threshold has properties which may allow it to be interpreted as the natural rate of unemployment. If we interpret the time-varying threshold as the natural rate, the regimes in the STAR model depend in part on the sign of the employment gap. Moreover, allowing the threshold to vary lets us analyze the effect of shocks to the natural rate. We found that the effect of these shocks differs across regimes.

## References


[1] Ball, Laurence and Mankiw, N. Gregory. "The NAIRU in Theory and Practice." *Journal of Economic Perspectives*, Fall 2002, 16(4), pp. 115-136.

[2] Basistha, Arabinda and Nelson, Charles R. "New Measures of the Output Gap Based on the Forward-looking New Keynesian Phillips Curve." *Journal of Monetary Economics*, March 2007, 54(2), pp. 498-511.

[3] Chauvet, Marcelle; Juhn, Chinhui; and Potter, Simon. "Markov Switching in Disaggregate Unemployment Rates." *Empirical Economics*, 2002, 27(2), pp. 205-232.

[4] Chib, Siddhartha. "Bayes Regression with Autoregressive Errors: A Gibbs Sampling Approach." *Journal of Econometrics*, August 1993, 58(3), pp. 275-294.

[5] Chib, Siddhartha and Greenberg, Edward. "Understanding the Metropolis-Hastings Algorithm." *American Statistician*, November 1995, 49(4), pp. 327-335.

[6] Chib, Siddhartha and Greenberg, Edward. "Markov Chain Monte Carlo Simulation Methods in Econometrics." *Econometric Theory*, August 1996, 12(3), pp. 409-431.

[7] Clark, Peter K. "Trend Reversion in Real Output and Unemployment." *Journal of Econometrics*, January 1989, 40(1), pp. 15-32.

[8] Deschamps, Philippe J. "Comparing Smooth Transition and Markov Switching Autoregressive Models of US Unemployment." *Journal of Applied Econometrics*, June/July 2008, 23(4), pp. 435-462.

[9] Doménech, Rafael and Gómez, Victor. "Estimating Potential Output, Core Inflation, and the NAIRU as Latent Variables." *Journal of Business and Economic Statistics*, July 2006, 24(3), pp. 354-365.

[10] Dueker, Michael; Psaradakis, Zacharias; Sola, Martin; and Spagnolo, Fabio. "State-Dependent Threshold STAR Models." working paper, March 2010.

[11] Galí, Jordi; Gertler, Mark; and López-Salido, J. David. "Robustness of the Estimates of the Hybrid New Keynesian Phillips Curve." *Journal of Monetary Economics*, September 2005, 52(6), pp. 1107-1118.

[12] Granger, Clive W.J. and Teräsvirta, Timo. *Modelling Nonlinear Economic Relationships*. New York: Oxford University Press, 1993.

[13] Hamilton, James D. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle." *Econometrica*, March 1989, 57(2), pp. 357-384.

[14] Jaeger, Albert and Parkinson, Martin. "Some Evidence on Hysteresis in Unemployment Rates." *European Economic Review*, February 1994, 38(2), pp. 329-342.

[15] Julier, Simon J. and Uhlmann, Jeffrey K. "New Extension of the Kalman Filter to Nonlinear Systems," in Proceedings of SPIE, Vol. 3068: Signal Processing, Sensor Fusion, and Target Recognition VI, July 1997, pp. 182-193.

[16] Lopes, Hedibert F. and Salazar, Esther. "Bayesian Model Uncertainty in Smooth Transition Autoregressions." *Journal of Time Series Analysis*, January 2006, 27(1), pp. 99-117.

[17] Owyang, Michael T. "Persistence, Excess Volatility, and Volatility Clusters in Inflation." Federal Reserve Bank of St. Louis *Review*, November/December 2001, 83(6), pp. 41-51.

[18] Phelps, Edmund S. and Zoega, Gylfi. "Natural-Rate Theory and OECD Unemployment." *Economic Journal*, May 1998, 108(448), pp. 782-801.

[19] Särkkä, Simo. "Unscented Rauch-Tung-Striebel Smoother." *IEEE Transactions on Automatic Control*, April 2008, 53(3), pp. 845-849.

[20] Sinclair, Tara M. "The Relationships between Permanent and Transitory Movements in U.S. Output and the Unemployment Rate." *Journal of Money, Credit, and Banking*, March-April 2009, 41(2-3), pp. 529-542.

[21] Skalin, Joakim and Teräsvirta, Timo. "Modeling Asymmetries and Moving Equilibria in Unemployment Rates." *Macroeconomic Dynamics*, April 2002, 6(2), pp. 202-241.

[22] van Dijk, Dick; Teräsvirta, Timo; and Franses, Philip Hans. "Smooth Transition Autoregressive Models—A Survey of Recent Developments." *Econometric Reviews*, 2002, 21(1), pp. 1-47.

[23] Wan, Eric A. and van der Merwe, Rudolph. "The Unscented Kalman Filter," in Simon Haykin, ed., *Kalman Filtering and Neural Networks*. New York: Wiley, September 2001, pp. 221-282.

# A Details of the UKF

## A.1 Filtering

We begin with a set of initial values. We then augment the mean with the expectation of the transition noise

$$\mathbf{x}^a_{t-1|t-1} = \begin{bmatrix} \mathbf{x}_{t-1|t-1} \\ E\left[\epsilon_t\right] \end{bmatrix}$$

and augment the state covariance

$$P^a_{t-1|t-1} = \begin{bmatrix} P_{t-1|t-1} & \mathbf{0}_{N,N} \\ \mathbf{0}_{N,N} & \mathbf{I}_N \end{bmatrix},$$

where $\mathbf{x}_{t-1|t-1}$ and $P_{t-1|t-1}$ are the estimates of the state and its covariance matrix at time $t-1$. Our task is to construct a set of $2L+1$ sigma points, where $L$ is the dimension of $\mathbf{x}^a_{t-1|t-1}$. Let

$$\boldsymbol{\chi}^p_{t-1|t-1} = \begin{cases} \mathbf{x}^a_{t-1|t-1} & , \; for \quad p=0 \\ \mathbf{x}^a_{t-1|t-1} + \left(\sqrt{\left(L+\zeta\right)P^a_{t-1|t-1}}\right)_p & , \; for \quad p=1,...,L \\ \mathbf{x}^a_{t-1|t-1} - \left(\sqrt{\left(L+\zeta\right)P^a_{t-1|t-1}}\right)_{p-L} & , \; for \quad p=L+1,...,2L \end{cases}$$

define the initial sigma points, where $\zeta = a^2\left(L+\kappa\right) - L$. Here, $a$ and $\kappa$ 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 $i$th column of the lower triangular Cholesky factorization of the square matrix $X$. Given the set of initial sigma points, we can then propagate $\left\{\boldsymbol{\chi}^p_{t-1|t-1}\right\}$ through the function $\mathbf{g}\left(.\right)$ to recover a prediction for each of the initial sigma points:

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

The predicted states and covariances can then be extrapolated from a weighted sum of the propagated sigma points:

$$\widehat{\mathbf{x}}_{t|t-1} = \sum_{p=0}^{2L} w^p_s \boldsymbol{\chi}^p_{t|t-1}$$

and

$$\mathbf{P}_{t|t-1} = \sum_{p=0}^{2L} w^p_c \left[\boldsymbol{\chi}^p_{t|t-1} - \widehat{\mathbf{x}}_{t|t-1}\right]\left[\boldsymbol{\chi}^p_{t|t-1} - \widehat{\mathbf{x}}_{t|t-1}\right]',$$

where $w^0_s = L\left(L+\zeta\right)^{-1}$, $w^0_c = L\left(L+\zeta\right)^{-1} + \left(1-\alpha^2+\beta\right)$, and $w^p_s = w^p_c = \frac{1}{2}\left(L+\zeta\right)^{-1}$ for all other $p$. We form the updated state similar to the standard Kalman filter

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

where $K_t$ is the Kalman gain defined by

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

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

$$\mathbf{P}_{yz} = \sum_{p=0}^{2L} w_c^p \left[\boldsymbol{\chi}_{t|t-1}^p - \widehat{\mathbf{x}}_{t|t-1}\right] \left[\boldsymbol{\gamma}_{t|t}^p - \widehat{\mathbf{y}}_t\right]'$$

and $\mathbf{P}_{yy}$ is the predicted covariance

$$\mathbf{P}_{yy} = \sum_{p=0}^{2L} w_c^p \left[\boldsymbol{\gamma}_{t|t}^p - \widehat{\mathbf{y}}_t\right] \left[\boldsymbol{\gamma}_{t|t}^p - \widehat{\mathbf{y}}_t\right]'.$$

The updated covariance is defined by

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

## A.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 impossible to implement in all cases. The Rauch-Tung-Striebel (RTS) filter is alternative, forward smoother for the Kalman filter. Särkkä (2008) constructed a smoother using principles similar to RTS. The unscented RTS smoother begins by augmenting the unscented Kalman filter with a step that recomputes the state estimate:

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

with covariance matrix

$$P_t^s = P_t + D_t\left[P_{t+1}^s - P_{t+1|t}\right]D_t'.$$

The smoother gain, $D_t$, defined by

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

where

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

and

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

| Table 1: Priors for Estimation | | |
|---|---|---|
| Parameter | Prior Distribution | Hyperparameters |
| $\{z_t^*\}_{t=1}^T$ | latent threshold $\sim$ MH draw unscented Kalman filter proposal | $n/a$ |
| $\alpha, \theta, \Gamma, \lambda$ | $N(\mathbf{m}_0, \mathbf{M}_0)$ | $\mathbf{m}_0 = \mathbf{0}_{\kappa \times 1}$ ; $\mathbf{M}_0 = \mathbf{I}_\kappa$ |
| $\gamma$ | $U(\kappa_1, \kappa_2)$ | $\kappa_1 = -10$ ; $\kappa_2 = 0$ |
| $\rho$ | $N(f_0, F_0)$ | $f_0 = 1$ ; $F_0 = 1$ |
| $\Omega^{-1}$ | $W\left(\frac{\varpi_0}{2}, \frac{D_0}{2}\right)$ | $\varpi_0 = 1$ ; $D_0 = \mathbf{I}_n$ |

Table 2: VECM Posterior Distributions

| Parameter | Mean | Parameter | Mean |
|---|---|---|---|
| $\alpha_0$ | $\begin{bmatrix} 0.00002 \\ 0.0008 \\ 0.0139 \end{bmatrix}$ | $\alpha_1$ | $\begin{bmatrix} 0.00004 \\ 0.0008 \\ 0.0156 \end{bmatrix}$ |
| $\theta_{01}$ | $\begin{bmatrix} 0.0562 & 0.0916 & -0.0055 \\ 0.0028 & 0.3236 & 0.0206 \\ -0.5129 & 0.1841 & 0.7011 \end{bmatrix}$ | $\theta_{11}$ | $\begin{bmatrix} 0.0530 & 0.0876 & -0.0043 \\ 0.0212 & 0.2786 & 0.0195 \\ -0.4960 & 0.1972 & 0.6473 \end{bmatrix}$ |
| $\theta_{02}$ | $\begin{bmatrix} 0.2030 & 0.0666 & -0.0024 \\ -0.0167 & 0.2883 & -0.0115 \\ -0.1768 & 0.2405 & 0.0501 \end{bmatrix}$ | $\theta_{12}$ | $\begin{bmatrix} 0.1759 & 0.0584 & -0.0030 \\ -0.0149 & 0.2499 & -0.0099 \\ -0.2292 & 0.2743 & 0.0411 \end{bmatrix}$ |
| $\theta_{03}$ | $\begin{bmatrix} 0.1717 & 0.0070 & 0.0004 \\ -0.0674 & 0.1106 & -0.0046 \\ 0.0140 & -0.0258 & 0.0683 \end{bmatrix}$ | $\theta_{13}$ | $\begin{bmatrix} 0.1420 & 0.0042 & 0.0004 \\ -0.0690 & 0.0807 & -0.0026 \\ 0.0344 & 0.0326 & 0.0441 \end{bmatrix}$ |
| $\Gamma_0$ | $\begin{bmatrix} -0.0007 \\ -0.0004 \\ -0.0013 \end{bmatrix}$ | $\Gamma_1$ | $\begin{bmatrix} -0.0007 \\ 0.0001 \\ 0.0123 \end{bmatrix}$ |
| $\Omega$ | $\begin{bmatrix} 0.0292 & -0.0022 & -0.0067 & -0.0045 \\ -0.0022 & 0.0342 & 0.0888 & -0.0111 \\ -0.0067 & 0.0888 & 1.6827 & -0.2595 \\ -0.0045 & -0.0111 & -0.2595 & 1.0000 \end{bmatrix}$ | $\gamma$ | 1.5183 |
| $\lambda_0$ | 0.2166 | $\lambda_1$ | 0.9633 |

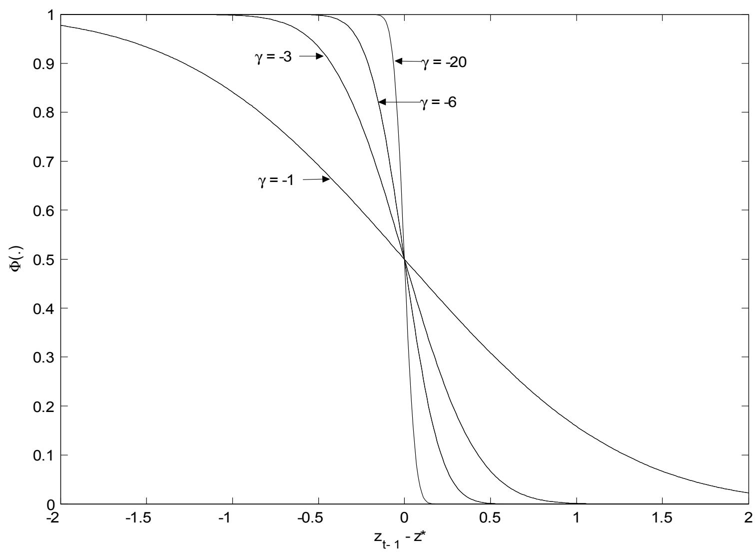


Figure 1: Transition Function

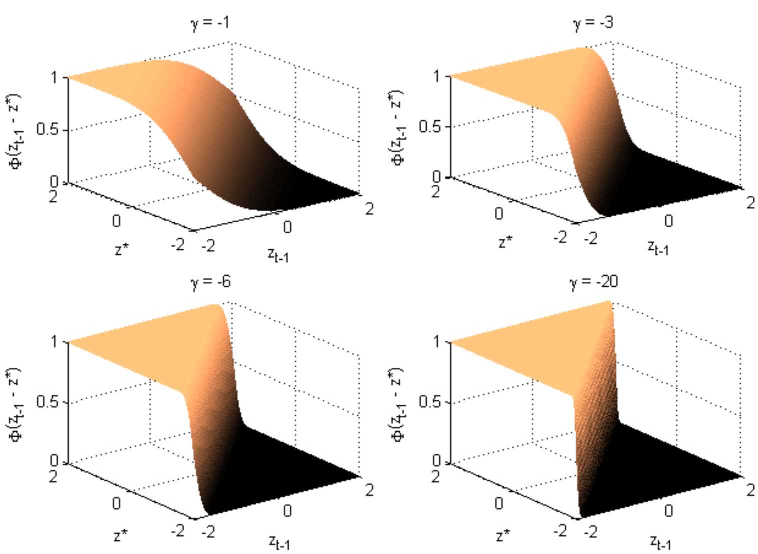


Figure 2: 3-D Time-Varying Transition Function

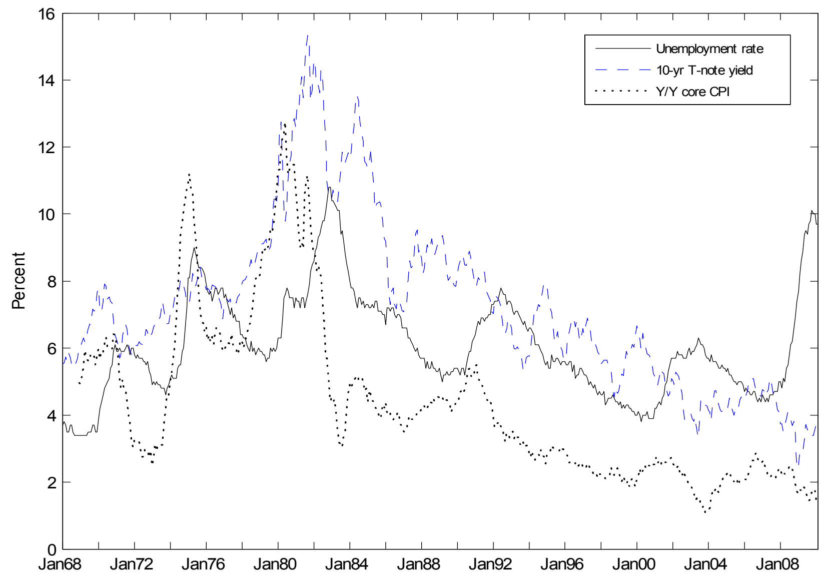


Figure 3: The Data

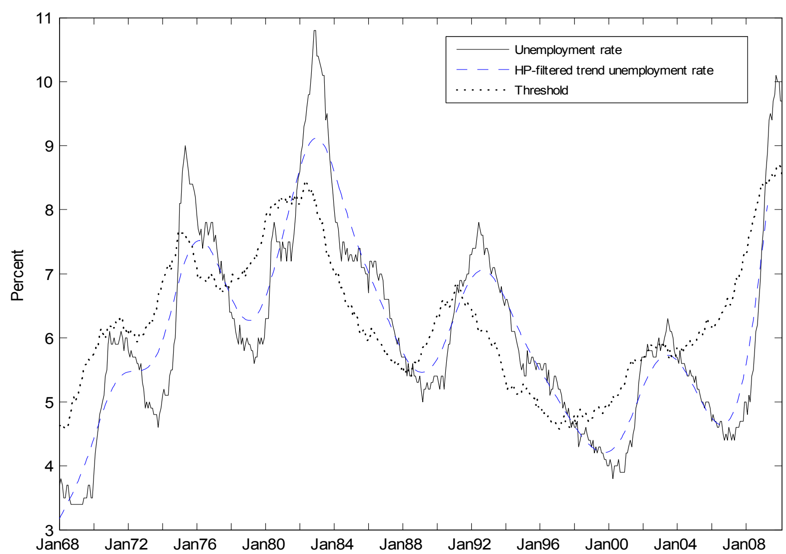


Figure 4: Unemployment Rate and Latent Threshold Level

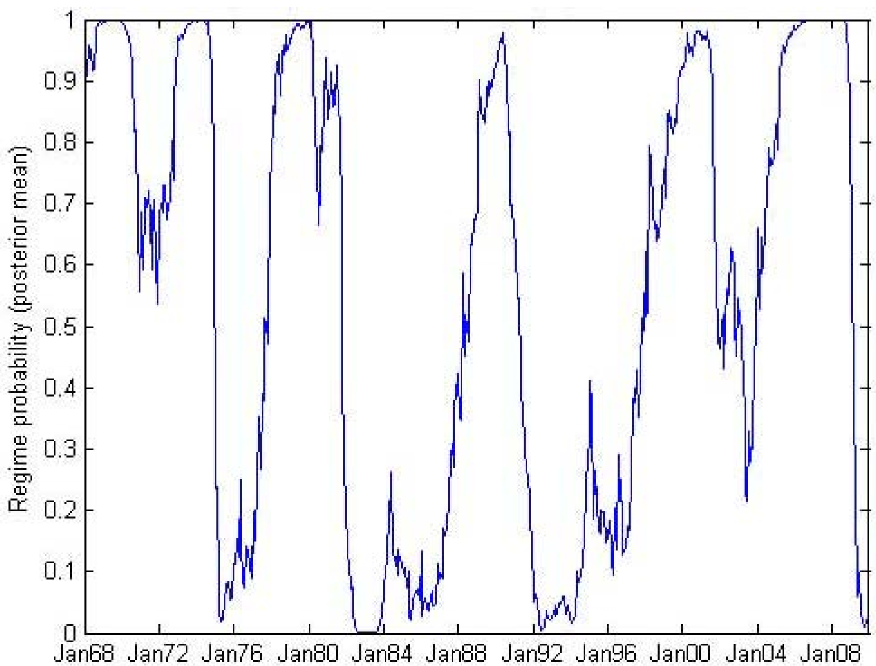


Figure 5: Regime Probilities for Time-Varying Threshold Model

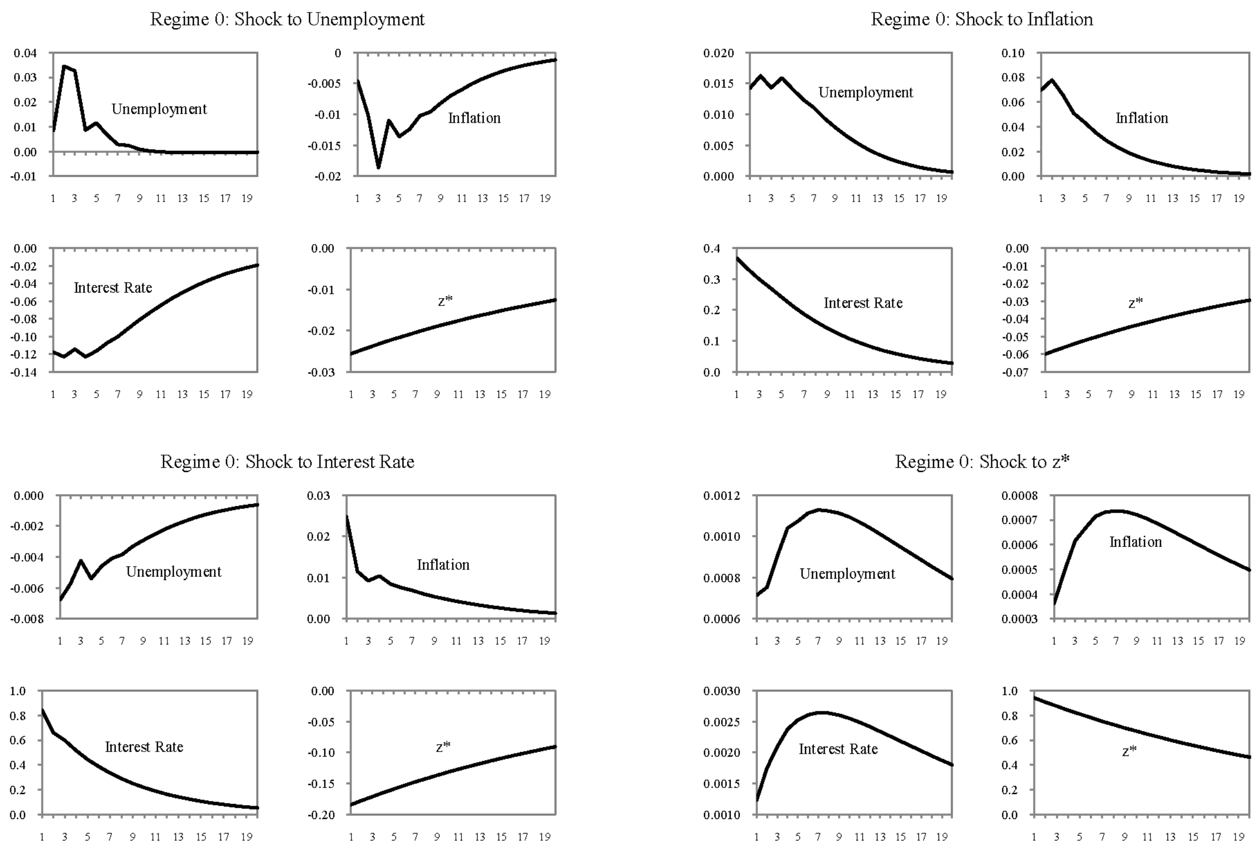


Figure 6: Impulse Responses in Regime 0

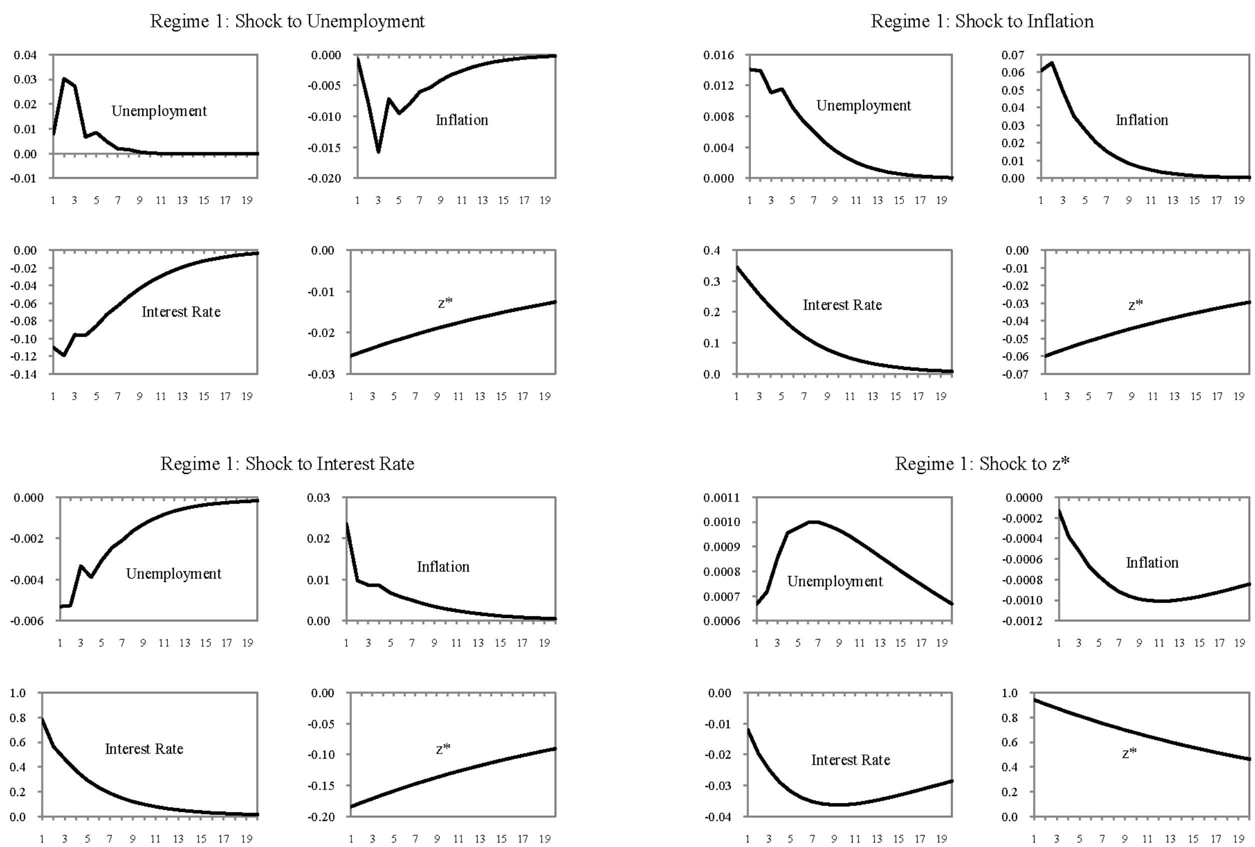


Figure 7: Impulse Responses in Regime 1

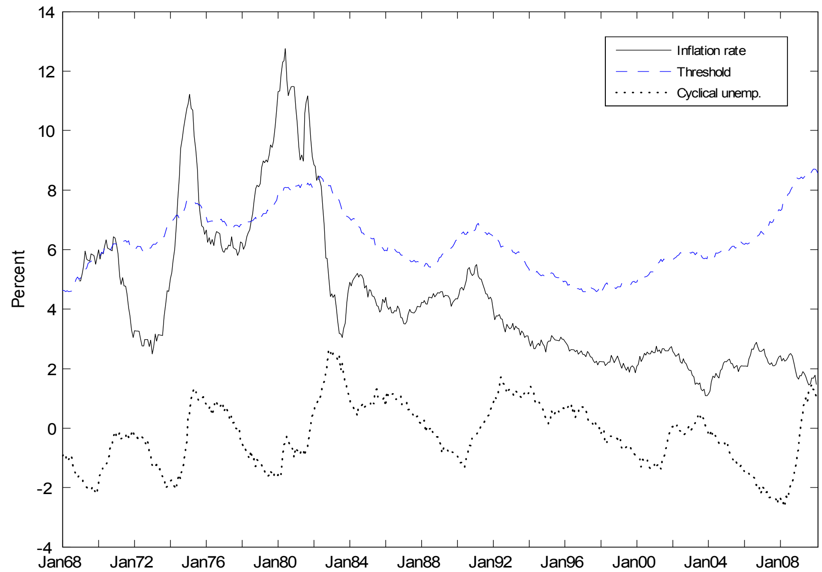


Figure 8: Inflation Rate and Latent Threshold Level