A Time-Varying Threshold STAR Model of Unemployment

MICHAEL DUEKER[a]
MICHAEL OWYANG[b]
MARTIN SOLA[c,d]
[a]Russell Investments
[b]Federal Reserve Bank of St. Louis
[c]Departamento de Economia, Universidad Torcuato di Tella
[d]School of Economics, Mathematics and Statistics, Birkbeck College, University of London

January 2009

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.

1 Introduction

Our starting point in modeling the rate of unemployment is the contemporaneous smooth-transition autoregressive model from Dueker, Sola and Spagnolo (2007), which weights the two regimes by the ex ante probability that the unemployment rate will be below/above the contemporaneous value of the threshold:

$$P(y_t < y_t^* \mid I_{t-1}, \Theta_0)$$

and

$$P(y_t \geq y_t^* \mid I_{t-1}, \Theta_1),$$

where Θ_0 and Θ_1 are the parameters that pertain to the two regimes and the respective regime weights are

$$P_{0t} = \frac{P(y_t < y_t^* \mid I_{t-1}, \Theta_0)}{P(y_t < y_t^* \mid I_{t-1}, \Theta_0) + P(y_t \geq y_t^* \mid I_{t-1}, \Theta_1)} \tag{1}$$

$$P_{1t} = \frac{P(y_t \geq y_t^* \mid I_{t-1}, \Theta_1)}{P(y_t < y_t^* \mid I_{t-1}, \Theta_0) + P(y_t \geq y_t^* \mid I_{t-1}, \Theta_1)} \tag{2}$$

Note that $P_{0t} + P_{1t} = 1$ this way.

With the time-varying threshold comes the possibility of adding an error-correction term to a STAR model:

$$\begin{aligned} \mathbf{y_t} &= P_{0t}[\mu_0 + \sum_{i=1}^{p} \Phi_0^{(i)} \mathbf{y_{t-i}} + \Gamma_0(y_{t-1} - y_{t-1}^*)] + \\ &\quad P_1 t[\mu_1 + \sum_{i=1}^{p} \Phi_1^{(i)} \mathbf{y_{t-i}} + \Gamma_1(y_{t-1} - y_{t-1}^*)] + \epsilon_t \end{aligned} \tag{4}$$

We model the time-varying threshold level of the unemployment rate as an autoregressive process, although the coefficient θ could be close to one to accommodate a high level of persistence in the threshold level:

$$y_t^* = \lambda + \theta y_{t-1}^* + w_t \tag{5}$$

The threshold is endogenous in that its innovation is correlated with the shocks to the observable data:

$$Cov(\epsilon_t, w_t) \neq 0.$$

and the regime-specific covariance matrices are denoted Ω_0 and Ω_1.

Estimation algorithm for the endogenous, time-varying threshold model

Because the vector of latent thresholds enters the regime weights, P_{0t} and P_{1t}, the model cannot be estimated by maximum likelihood. Nor is it possible to derive exact conditional distributions for Gibbs sampling. Instead, a Metropolis-Hastings algorithm is needed for Bayesian estimation of this model. The key is to have a good, yet tractable, proposal draw for the latent threshold series, $\{y_t^*\}_{t=1}^T$.

The parameter groupings for Markov Chain Monte Carlo estimation of the model are:

$\{y_t^*\}_{t=1}^T$	latent threshold series $\sim$ M-H draw, unscented Kalman filter proposal
$\phi, \rho, \mu, \gamma, \lambda, \theta$	regression coeffs. $\sim$ M-H draw Normal proposal
Ω_0, Ω_1	cov. matrices $\sim$ inverted Wishart

Estimation results for U.S. unemployment

We applied the model to the U.S. unemployment rate since 1968. Figure 1 shows the relationship we found between the unemployment rate and the estimated time-varying threshold level. We also include out-of-sample forecasts of both the unemployment rate and the threshold level. Starting from November 2008, when the unemployment rate was 6.7 percent, the STAR model predicts that the unemployment rate would reach 7.5 percent at its peak in November 2009. This projected peak seems low, but it is important to bear in mind that this forecast is coming from a univariate model of the unemployment rate. With additional information about the current state of the business cycle, the model should be able to provide accurate forecasts of unemployment and keep the interesting self-referential unemployment dynamics captured by the model.

Appendix: analysis using the unscented Kalman filter

We can rewrite the model in the preceding section in a state space representation

$$\begin{aligned} \mathbf{y}_t &= H\mathbf{z}_t + \varepsilon_t, \\ \mathbf{z_t} &= \alpha + \mathbf{g}(\mathbf{z_{t-1}}) + \mathbf{u}_t \end{aligned} \tag{6}$$

where the nonlinear function $\mathbf{g}(.)$ contains the cumulative density weighting function P_{at}.

Given this state-space representation, inferred values for the latent variable $\mathbf{z}_t$ 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. The UKF tracks the state variable by computing its distribution across a set of deterministic points called sigma points.

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

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

and augment the state covariance

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

where

$\mathbf{s}_{t-1|t-1}$ 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 $\mathbf{s}^a_{t-1|t-1}$. Let

$$\chi^p_{t-1|t-1} = \begin{cases} \mathbf{z}^a_{t-1|t-1} & , \quad for \quad p = 0 \\ \mathbf{z}^a_{t-1|t-1} + \left(\sqrt{(L+\lambda)\, P^a_{t-1|t-1}}\right)_i & , \quad for \quad p = 1, ..., L \\ \mathbf{z}^a_{t-1|t-1} - \left(\sqrt{(L+\lambda)\, P^a_{t-1|t-1}}\right)_{i-L} & , \quad for \quad p = L+1, ..., 2L \end{cases}$$

define the initial sigma points, where $\lambda = a^2 (L + \kappa) - L$, and a and κ are user-chosen parameters that govern the spread and scale, repsectively. Here, $\left(\sqrt{X}\right)_i$ is the ith column of the lower triangular Cholesky factorization of the square marix X. Given the set of initial sigma points, we can then propagate $\left\{\chi^p_{t-1|t-1}\right\}$ through the transition function $\mathbf{g}(.)$ to recover

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

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

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

and

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

The weights are defined as

$$w^p_s = \begin{cases} \frac{\lambda}{L+\lambda} & , \quad for \quad p = 0 \\ \frac{1}{2(L+\lambda)} & , \quad for \quad p = 1, ..., 2L \end{cases}$$

and

$$w^p_c = \begin{cases} \frac{\lambda}{L+\lambda} + (1 - a^2 + b) & , \quad for \quad p = 0 \\ \frac{1}{2(L+\lambda)} & , \quad for \quad p = 1, ..., 2L \end{cases},$$

where a is defined as above and b is a parameter that incorporates prior information of the distribution.

Filter recursions

The updating step proceeds in a similar manner. We can augment the predicted state and covariance with the expectation and covariance of the measurement noise

$$\mathbf{z}_{t|t-1}^{a} = \begin{bmatrix} \widehat{\mathbf{z}}_{t|t-1} \\ \mathbf{v}_t \end{bmatrix}$$

and

$$P_{t|t-1}^{a} = \begin{bmatrix} P_{t|t-1} & 0_{n,n} \\ 0_{n,n} & \mathbf{\Omega} \end{bmatrix},$$

where $\mathbf{\Omega}$ is the covariance of the measurement noise. We again compute $2L+1$ sigma points from

$$\widehat{\chi}_{t|t-1}^{p} = \begin{cases} \mathbf{z}_{t|t-1}^{a} & , \ for \quad p = 0 \\ \mathbf{z}_{t|t-1}^{a} + \left(\sqrt{(L+\lambda)\, P_{t|t-1}^{a}}\right)_i & , \ for \quad p = 1, ..., L \\ \mathbf{z}_{t|t-1}^{a} - \left(\sqrt{(L+\lambda)\, P_{t|t-1}^{a}}\right)_{i-L} & , \ for \quad p = L+1, ..., 2L \end{cases}.$$

We then propagate the update sigma points through the measurement equation:

$$\gamma_{t|t}^{p} = \mathbf{H}\widehat{\chi}_{t|t-1}^{p}, \quad p = 0, ..., 2L.$$

The updated propagation points, $\gamma_{t|t}^{p}$, are used to construct the predicted measurement error

$$\widehat{\mathbf{y}}_t = \sum_{p=0}^{2L} w_s^p \gamma_{t|t}^{p},$$

where the weights are defined as before. We then form the updated state similar to the standard Kalman filter

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

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

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

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

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

The updated covariance is defined by

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

Smoothing

Multi-move sampling of the latent attractor, $\{s^*\}_{t=1}^T$, requires backwards sampling from one-period smoothed inferences of the state vector.

Here the unscented Rauch-Tung-Striebel smoother comes into play. The URTS smoother (Sarkka, 2007) begins by augmenting the unscented Kalman filter with a smoothing step that recomputes the state estimate. The smoothed state is

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

and

with covariance matrix

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

where D_t is the smoother gain defined by

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

where

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

with

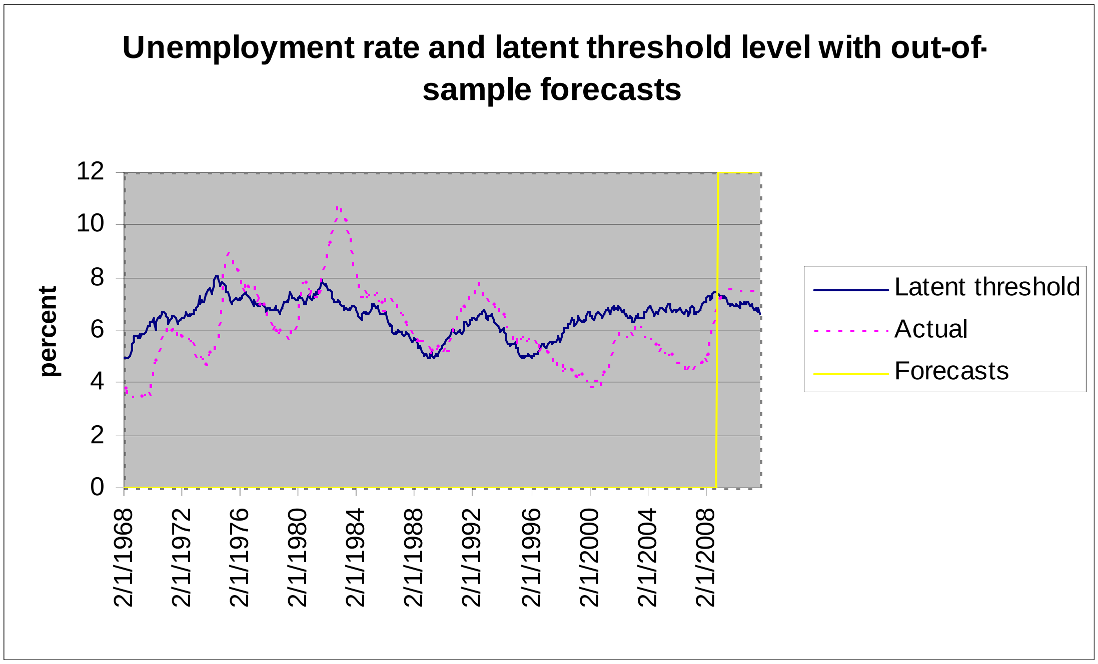
$$\chi_{t|t} = \chi_{t|t-1} + K_t(y_t - \hat{y}_{t|t-1}).$$

References

[1] Julier, Simon J. and Jeffery K. Uhlmann. "A New Extension of the Kalman Filter to nonlinear Systems," in The Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls, Multi Sensor Fusion, Tracking and Resource Management II, SPIE, 1997.

[2] Särkkä, Simo. "Unscented Rauch-Tung-Striebel Smoothing," forthcoming IEEE, 2007.

[3] Wan, Eric and van der Merwe, Rudolph. "The Unscented Kalman Filter," in S. Hakin, *Kalman Filtering and Neural Networks*, Wiley, September 2001.


Unemployment rate and latent threshold level with out-of-sample forecasts
percent
12
10
8
6
4
2
0
2/1/1968
2/1/1972
2/1/1976
2/1/1980
2/1/1984
2/1/1988
2/1/1992
2/1/1996
2/1/2000
2/1/2004
2/1/2008
Latent threshold
Actual
Forecasts