Regime-Switching Models

May 18, 2005

James D. Hamilton

Department of Economics, 0508

University of California, San Diego

La Jolla, CA 92093-0508

jhamilton@ucsd.edu

Prepared for: Palgrave Dictionary of Economics

Many economic time series occasionally exhibit dramatic breaks in their behavior, associated with events such as financial crises (Jeanne and Masson, 2000; Cerra, 2005; Hamilton, 2005) or abrupt changes in government policy (Hamilton, 1988; Sims and Zha, 2004, Davig, 2004). Of particular interest to economists is the apparent tendency of many economic variables to behave quite differently during economic downturns, when underutilization of factors of production rather than their long-run tendency to grow governs economic dynamics (Hamilton, 1989, Chauvet and Hamilton, 2005). Abrupt changes are also a prevalent feature of financial data, and the approach described below is quite amenable to theoretical calculations for how such abrupt changes in fundamentals should show up in asset prices (Ang and Bekaert, 2003; Garcia, Luger, and Renault, 2003; Dai, Singleton, and Wei, 2003).

Consider how we might describe the consequences of a dramatic change in the behavior of a single variable y_t. Suppose that the typical historical behavior could be described with a first-order autoregression,

$$y_t = c_1 + \phi y_{t-1} + \varepsilon_t, \tag{1}$$

with $\varepsilon_t \sim N(0, \sigma^2)$, which seemed to adequately describe the observed data for $t = 1, 2, ..., t_0$. Suppose that at date t_0 there was a significant change in the average level of the series, so that we would instead wish to describe the data according to

$$y_t = c_2 + \phi y_{t-1} + \varepsilon_t \tag{2}$$

for $t = t_0 + 1, t_0 + 2, ...$ This fix of changing the value of the intercept from c_1 to c_2 might help the model to get back on track with better forecasts, but it is rather unsatisfactory as a probability law that could have generated the data. We surely would not want to maintain that the change from c_1 to c_2 at date t_0 was a deterministic event that anyone would have been able to predict with certainty looking ahead from date $t = 1$. Instead there must have been some imperfectly predictable forces that produced the change. Hence, rather than claim that expression (1) governed the data up to date t_0 and (2) after that date, what we must have in mind is that there is some larger model encompassing them both,

$$y_t = c_{s_t} + \phi y_{t-1} + \varepsilon_t, \tag{3}$$

where s_t is a random variable that, as a result of institutional changes, happened in our sample to assume the value $s_t = 1$ for $t = 1, 2, ..., t_0$ and $s_t = 2$ for $t = t_0 + 1, t_0 + 2, ...$ A complete description of the probability law governing the observed data would then require a probabilistic model of what caused the change from $s_t = 1$ to $s_t = 2$. The simplest such specification is that s_t is the realization of a two-state Markov chain with

$$\Pr(s_t = j | s_{t-1} = i, s_{t-2} = k, ..., y_{t-1}, y_{t-2}, ...) = \Pr(s_t = j | s_{t-1} = i) = p_{ij}. \tag{4}$$

Assuming that we do not observe s_t directly, but only infer its operation through the observed behavior of y_t, the parameters necessary to fully describe the probability law governing y_t are then the variance of the Gaussian innovation σ^2, the autoregressive coefficient ϕ, the two intercepts c_1 and c_2, and the two state transition probabilities, p_{11} and p_{22}.

The specification in (4) assumes that the probability of a change in regime depends on the past only through the value of the most recent regime, though, as noted below, nothing in the approach described below precludes looking at more general probabilistic specifications. But the simple time-invariant Markov chain (4) seems the natural starting point and is clearly preferable to acting as if the shift from c_1 to c_2 was a deterministic event. Permanence of the shift would be represented by $p_{22} = 1$, though the Markov formulation invites the more general possibility that $p_{22} < 1$. Certainly in the case of business cycles or financial crises, we know that the situation, though dramatic, is not permanent. Furthermore, if the regime change reflects a fundamental change in monetary or fiscal policy, the prudent assumption would seem to be to allow the possibility for it to change back again, suggesting that $p_{22} < 1$ is often a more natural formulation for thinking about changes in regime than $p_{22} = 1$.

A model of the form of (3)-(4) with no autoregressive elements ($\phi = 0$) appears to have been first analyzed by Lindgren (1978) and Baum, et. al. (1980). Specifications that incorporate autoregressive elements date back in the speech recognition literature to Poritz (1982), Juang and Rabiner (1985), and Rabiner (1989), who described such processes as "hidden Markov models". Markov-switching regressions were introduced in econometrics by Goldfeld and Quandt (1973), the likelihood function for which was first correctly calculated by Cosslett and Lee (1985). The formulation of the problem described here, in which all objects of interest are calculated as a by-product of an iterative algorithm similar in spirit to a Kalman filter, is due to Hamilton (1989, 1994). General characterizations of moment and stationarity conditions for such processes can be found in Tjøstheim (1986), Yang (2000), Timmermann (2000), and Francq and Zakoïan (2001).

Suppose that the econometrician observes y_t directly but can only make an inference about the value of s_t based on what we see happening with y_t. This inference will take the form of two probabilities

$$\xi_{jt} = \Pr(s_t = j|\Omega_t; \boldsymbol{\theta}) \tag{5}$$

for $j = 1, 2$, where these two probabilities sum to unity by construction. Here $\Omega_t = \{y_t, y_{t-1}, ..., y_1, y_0\}$ denotes the set of observations obtained as of date t, and $\boldsymbol{\theta}$ is a vector of population parameters, which for the above example would be $\boldsymbol{\theta} = (\sigma, \phi, c_1, c_2, p_{11}, p_{22})'$, and which for now we presume to be known with certainty. The inference is performed iteratively for $t = 1, 2, ..., T$, with step t accepting as input the values

$$\xi_{i,t-1} = \Pr(s_{t-1} = i|\Omega_{t-1}; \boldsymbol{\theta}) \tag{6}$$

for $i = 1, 2$ and producing as output (5). The key magnitudes one needs in order to perform this iteration are the densities under the two regimes,

$$\eta_{jt} = f(y_t|s_t = j, \Omega_{t-1}; \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y_t - c_j - \phi y_{t-1})^2}{2\sigma^2}\right], \tag{7}$$

for $j = 1, 2$. Specifically, given the input (6) we can calculate the conditional density of the tth observation from

$$f(y_t|\Omega_{t-1}; \boldsymbol{\theta}) = \sum_{i=1}^{2} \sum_{j=1}^{2} p_{ij}\xi_{i,t-1}\eta_{jt} \tag{8}$$

and the desired output is then

$$\xi_{jt} = \frac{\sum_{i=1}^{2} p_{ij}\xi_{i,t-1}\eta_{jt}}{f(y_t|\Omega_{t-1}; \boldsymbol{\theta})}. \tag{9}$$

As a result of executing this iteration, we will have succeeded in evaluating the sample conditional log likelihood of the observed data

$$\log f(y_1, y_2, ..., y_T|y_0; \boldsymbol{\theta}) = \sum_{t=1}^{T} \log f(y_t|\Omega_{t-1}; \boldsymbol{\theta}) \tag{10}$$

for the specified value of $\boldsymbol{\theta}$. An estimate of the value of $\boldsymbol{\theta}$ can then be obtained by maximizing (10) by numerical optimization.

Several options are available for the value ξ_{i0} to use to start these iterations. If the Markov chain is presumed to be ergodic, one can use the unconditional probabilities

$$\xi_{i0} = \Pr(s_0 = i) = \frac{1 - p_{jj}}{2 - p_{ii} - p_{jj}}.$$

Other alternatives are simply to set $\xi_{i0} = 1/2$ or estimate ξ_{i0} itself by maximum likelihood.

The calculations do not increase in complexity if we consider an $(r \times 1)$ vector of observations $\mathbf{y}_t$ whose density depends on N separate regimes. Let $\Omega_t = \{\mathbf{y}_t, \mathbf{y}_{t-1}, ..., \mathbf{y}_1\}$ be the observations through date t, $\mathbf{P}$ be an $(N \times N)$ matrix whose row j, column i element is the transition probability p_{ij}, $\boldsymbol{\eta}_t$ be an $(N \times 1)$ vector whose jth element $f(\mathbf{y}_t|s_t = j, \Omega_{t-1}; \boldsymbol{\theta})$ is the density in regime j, and $\hat{\boldsymbol{\xi}}_{t|t}$ an $(N \times 1)$ vector whose jth element is $\Pr(s_t = j|\Omega_t, \boldsymbol{\theta})$. Then (8) and (9) generalize to

$$f(\mathbf{y}_t|\Omega_{t-1}; \boldsymbol{\theta}) = \mathbf{1}'(\mathbf{P}\hat{\boldsymbol{\xi}}_{t-t|t-1} \odot \boldsymbol{\eta}_t) \tag{11}$$

$$\hat{\boldsymbol{\xi}}_{t|t} = \frac{\mathbf{P}\hat{\boldsymbol{\xi}}_{t-t|t-1} \odot \boldsymbol{\eta}_t}{f(\mathbf{y}_t|\Omega_{t-1}; \boldsymbol{\theta})} \tag{12}$$

where $\mathbf{1}$ denotes an $(N \times 1)$ vector all of whose elements are unity and $\odot$ denotes element-by-element multiplication. Markov-switching vector autoregressions are discussed in detail in Krolzig (1997). Vector applications include describing the comovements between stock prices and economic output (Hamilton and Lin, 1996) and the tendency for some series to move into recession before others (Hamilton and Perez-Quiros, 1996). There further is no requirement that the elements of $\boldsymbol{\eta}_t$ be Gaussian densities or even from the same family of densities. For example, Dueker (1997) studied a model in which the degrees of freedom of a Student t distribution change depending on the economic regime.

One is also often interested in forming an inference about what regime the economy was in at date t based on observations obtained through a later date T, denoted $\hat{\boldsymbol{\xi}}_{t|T}$. These are referred to as "smoothed" probabilities, an efficient algorithm for whose calculation was developed by Kim (1994).

The calculations in (11) and (12) remain valid when the probabilities in $\mathbf{P}$ depend on lagged values of $\mathbf{y}_t$ or strictly exogenous explanatory variables, as in Diebold, Lee and Weinbach (1994), Filardo (1994), and Peria (2002). However, often there are relatively few transitions among regimes, making it difficult to estimate such parameters accurately, and most applications have assumed a time-invariant Markov chain. For the same reason, most applications assume only $N = 2$ or 3 different regimes, though there is considerable promise in models with a much larger number of regimes, either by tightly parameterizing the relation between the regimes (Calvet and Fisher, 2004), or with prior Bayesian information (Sims and Zha, 2004).

In the Bayesian approach, both the parameters $\boldsymbol{\theta}$ and the values of the states $\mathbf{s} = (s_1, s_2, ..., s_T)'$ are viewed as random variables. Bayesian inference turns out to be greatly facilitated by Monte Carlo Markov chain methods, specifically, the Gibbs sampler. This is achieved by sequentially (for $k = 1, 2, ...$) generating a realization $\boldsymbol{\theta}^{(k)}$ from the distribution of $\boldsymbol{\theta}|\mathbf{s}^{(k-1)}, \Omega_T$ followed by a realization of $\mathbf{s}^{(k)}$ from the distribution of $\mathbf{s}|\boldsymbol{\theta}^{(k)}, \Omega_T$. The first distribution, $\boldsymbol{\theta}|\mathbf{s}^{(k-1)}, \Omega_T$, treats the historical regimes generated at the previous iteration,

$s_1^{(k-1)}, s_2^{(k-1)}, ..., s_T^{(k-1)}$, as if fixed known numbers. Often this conditional distribution takes the form of a standard Bayesian inference problem whose solution is known analytically using natural conjugate priors. For example, the posterior distribution of ϕ given other parameters is a known function of easily calculated OLS coefficients. An algorithm for generating a draw from the second distribution, $\mathbf{s}|\boldsymbol{\theta}^{(k)}, \Omega_T$, was developed by Albert and Chib (1993). The Gibbs sampler turns out also to be a natural device for handling transition probabilities that are functions of observable variables, as in Filardo and Gordon (1998).

It is natural to want to test the null hypothesis that there are N regimes against the alternative of $N + 1$, for example, when $N = 1$, to test whether there are any changes in regime at all. Unfortunately, the likelihood ratio test of this hypothesis fails to satisfy the usual regularity conditions, because under the null hypothesis, some of the parameters of the model would be unidentified. For example, if there is really only one regime, the maximum likelihood estimate $\hat{p}_{11}$ does not converge to a well-defined population magnitude, meaning that the likelihood ratio test does not have the usual χ^2 limiting distribution. To interpret a likelihood ratio statistic one instead needs to appeal to the methods of Hansen (1992) or Garcia (1998). An alternative is to rely on generic tests of the hypothesis that an N-regime model accurately describes the data (Hamilton, 1996), though these tests are not designed for optimal power against the specific alternative hypothesis of $N + 1$ regimes. A test recently proposed by Carrasco, Hu, and Ploberger (2004) that is easy to compute but not based on the likelihood ratio statistic seems particularly promising. Other alternatives are to use Bayesian methods to calculate the value of N implying the largest value for the marginal likelihood (Chib, 1998) or the highest Bayes factor (Koop and Potter, 1999), or to compare models on the basis of their ability to forecast (Hamilton and Susmel, 1994).

A specification where the density depends on a finite number of previous regimes, $f(\mathbf{y}_t|s_t, s_{t-1}, ..., s_{t-m}, \Omega_{t-1}; \boldsymbol{\theta})$ can be recast in the above form by a suitable redefinition of regime. For example, if s_t follows a 2-state Markov chain with transition probabilities $\Pr(s_t = j|s_{t-1} = i)$ and $m = 1$, one can define a new regime variable s_t^* such that $f(\mathbf{y}_t|s_t^*, \Omega_{t-1}; \boldsymbol{\theta}) = f(\mathbf{y}_t|s_t, s_{t-1}, ..., s_{t-m}, \Omega_{t-1}; \boldsymbol{\theta})$ as follows:

$$s_t^* = \begin{cases} 1 & \text{when } s_t = 1 \text{ and } s_{t-1} = 1 \\ 2 & \text{when } s_t = 2 \text{ and } s_{t-1} = 1 \\ 3 & \text{when } s_t = 1 \text{ and } s_{t-1} = 2 \\ 4 & \text{when } s_t = 2 \text{ and } s_{t-1} = 2 \end{cases}.$$

Then s_t^* itself follows a 4-state Markov chain with transition matrix

$$\mathbf{P}^* = \begin{bmatrix} p_{11} & 0 & p_{11} & 0 \\ p_{12} & 0 & p_{12} & 0 \\ 0 & p_{21} & 0 & p_{21} \\ 0 & p_{22} & 0 & p_{22} \end{bmatrix}.$$

More problematic are cases in which the order of dependence m grows with the date of the observation t. Such a situation often arises in models whose recursive structure causes the density of y_t given Ω_{t-1} to depend on the entire history $y_{t-1}, y_{t-2}, ..., y_1$ as is the case in ARMA, GARCH, or state-space models. Consider for illustration a GARCH(1,1) specification in which the coefficients are subject to changes in regime, $y_t = h_t v_t$, where $v_t \sim N(0, 1)$

and

$$h_t^2 = \gamma_{s_t} + \alpha_{s_t} y_{t-1}^2 + \beta_{s_t} h_{t-1}^2. \tag{13}$$

Solving (13) recursively reveals that the conditional standard deviation h_t depends on the full history $\{y_{t-1}, y_{t-2}, ..., y_0, s_t, s_{t-1}, ..., s_1\}$. One way to avoid this problem was proposed by Gray (1996), who postulated that instead of being generated by (13), the conditional variance is characterized by

$$h_t^2 = \gamma_{s_t} + \alpha_{s_t} y_{t-1}^2 + \beta_{s_t} \tilde{h}_{t-1}^2 \tag{14}$$

where

$$\tilde{h}_{t-1}^2 = \sum_{i=1}^{N} \hat{\xi}_{i,t-1|t-2} \left(\gamma_i + \alpha_i y_{t-2}^2 + \beta_i \tilde{h}_{t-2}^2 \right).$$

In Gray's model, h_t in (14) depends only on s_t since $\tilde{h}_{t-1}^2$ is a function of data Ω_{t-1} only. An alternative solution, due to Haas, Mittnik, and Paolella (2004), is to hypothesize N separate GARCH processes whose values h_{it} all exist as latent variables at date t,

$$h_{it}^2 = \gamma_i + \alpha_i y_{t-1}^2 + \beta_i h_{i,t-1}^2 \tag{15}$$

and then simply pose the model as $y_t = h_{s_t} v_t$. Again the feature that makes this work is the fact that h_{it} in (15) is a function solely of the data Ω_{t-1} rather than the states $\{s_{t-1}, s_{t-2}, ..., s_1\}$.

A related problem arises in Markov-switching state-space models, which posit an unobserved state vector $\mathbf{z}_t$ characterized by

$$\mathbf{z}_t = \mathbf{F}_{s_t} \mathbf{z}_{t-1} + \mathbf{Q}_{s_t} \mathbf{v}_t$$

with $\mathbf{v}_t \sim N(\mathbf{0}, \mathbf{I}_n)$, with observed vectors $\mathbf{y}_t$ and $\mathbf{x}_t$ governed by

$$\mathbf{y}_t = \mathbf{H}_{s_t}^{'} \mathbf{z}_t + \mathbf{A}_{s_t}^{'} \mathbf{x}_t + \mathbf{R}_{s_t} \mathbf{w}_t$$

for $\mathbf{w}_t \sim N(\mathbf{0}, \mathbf{I}_r)$. Again the model as formulated implies that the density of $\mathbf{y}_t$ depends on the full history $\{s_t, s_{t-1}, ..., s_1\}$. Kim (1994) proposed a modification of the Kalman filter equations similar in spirit to the modification in (14) that can be used to approximate the log likelihood. A more common practice recently has been to estimate such models with numerical Bayesian methods, as in Kim and Nelson (1999).

References

Albert, James, and Siddhartha Chib (1993), "Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts," *Journal of Business and Economic Statistics* 11, 1-15.

Ang, Andrew, and Geert Bekaert (2002), "International Asset Allocation with Regime Shifts", *Review of Financial Studies* 15, 1137-1187.

Ang, A., and Bekaert, G. (2001), "Regime Switches in Interest Rates," *Journal of Business and Economic Statistics*, forthcoming.

Baum, Leonard E., Ted Petrie, George Soules, and Norman Weiss (1980), "A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains," *Annals of Mathematical Statistics* 41, 164-171.

Calvet, Laurent, and Adlai Fisher (2004), "How to Forecast Long-Run Volatility: Regime-Switching and the Estimation of Multifractal Processes," *Journal of Financial Econometrics* 2, 49-83.

Carrasco, Marine, Liang Hu, and Werner Ploberger (2004) "Optimal Test for Markov Switching," Working paper, University of Rochester.

Cerra, Valerie, and Sweta Chaman Saxena (2005), "Did Output Recover from the Asian Crisis?" *IMF Staff Papers* 52, 1-23.

Chauvet, Marcelle, and James D. Hamilton (2005), "Dating Business Cycle Turning Points," in *Nonlinear Analysis of Business Cycles,* edited by Costas Milas, Philip Rothman, and Dick van Dijk, Elsevier, forthcoming.

Chib, Siddhartha (1998), "Estimation and Comparison of Multiple Change-Point Models," *Journal of Econometrics* 86, 221-241.

Cosslett, Stephen R., and Lung-Fei Lee (1985), "Serial Correlation in Discrete Variable Models," *Journal of Econometrics* 27, 79-97.

Dai, Qiang, Kenneth J. Singleton, and Wei Yang (2003), "Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bonds," working paper, Stanford University.

Davig, Troy (2004), "Regime-Switching Debt and Taxation," *Journal of Monetary Economics* 51, 837-859.

Diebold, Francis X., Joon-Haeng Lee, and Gretchen C. Weinbach (1994), "Regime Switching with Time-Varying Transition Probabilities," in C. Hargreaves, ed., *Nonstationary Time Series Analysis and Cointegration*, Oxford: Oxford University Press.

Dueker, Michael (1997), "Markov Switching in GARCH Processes and Mean-Reverting Stock-Market Volatility," *Journal of Business and Economic Statistics* 15, 26-34.

Filardo, Andrew J. (1994), "Business Cycle Phases and Their Transitional Dynamics," *Journal of Business and Economic Statistics* 12, 299-308.

Filardo, Andrew J., and Stephen F. Gordon (1998), "Business Cycle Durations," *Journal of Econometrics* 85, 99-123.

Francq, C., and J.-M. Zakoïan (2001), "Stationarity of Multivariate Markov-Switching ARMA Models," *Journal of Econometrics* 102, 339-364.

Garcia, Rene (1998), "Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models," *International Economic Review* 39, 763-788.

Garcia, Rene, Richard Luger, and Eric Renault (2003), "Empirical Assessment of an Intertemporal Option Pricing Model with Latent Variables," *Journal of Econometrics* 116, 49-83.

Goldfeld, Stephen M., and Richard E. Quandt (1973), "A Markov Model for Switching Regressions," *Journal of Econometrics* 1, 3-16.

Gray, Stephen F. (1996), "Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process," *Journal of Financial Economics* 42, 27-62.

Haas, Markus, Stefan Mittnik, and Marc Paolella (2004), "A New Approach to Markov-Switching GARCH Models," *Journal of Financial Econometrics* 2, 493-530.

Hamilton, James D. (1988), "Rational-Expectations Econometric Analysis of Changes in Regime: An Investigation of the Term Structure of Interest Rates," *Journal of Economic Dynamics and Control* 12, 385-423.

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

Hamilton, James D. (1994), *Time Series Analysis*, Princeton, NJ: Princeton University Press.

Hamilton, James D. (1996), "Specification Testing in Markov-Switching Time-Series Models," *Journal of Econometrics* 70, 127-157.

Hamilton, James D. (2005), "What's Real About the Business Cycle?", *Federal Reserve Bank of St. Louis Review,* forthcoming.

Hamilton, James D., and Gang Lin (1996), "Stock Market Volatility and the Business

Cycle," *Journal of Applied Econometrics*, 11, 573-593.

Hamilton, James D., and Gabriel Perez-Quiros (1996), "What Do the Leading Indicators Lead?", *Journal of Business* 69, 27-49.

Hamilton, James D., and Raul Susmel (1994), "Autoregressive Conditional Heteroskedasticity and Changes in Regime," *Journal of Econometrics* 64, 307-333.

Hansen, Bruce E. (1992), "The Likelihood Ratio Test under Non-Standard Conditions," *Journal of Applied Econometrics 7, S61-82.* Erratum, 1996, 11, 195-198.

Jeanne, Olivier and Paul Masson (2000), "Currency Crises, Sunspots, and Markov-Switching Regimes," *Journal of International Economics* 50, 327-350.

Juang, Biing-Hwang, and Lawrence R. Rabiner (1985), "Mixture Autoregressive Hidden Markov Models for Speech Signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP-30, 1404-1413.*

Kim, Chang Jin (1994), "Dynamic Linear Models with Markov-Switching," *Journal of Econometrics* 60, 1-22.

Kim, Chang Jin, and Charles R. Nelson (1999), *State-Space Models with Regime Switching*, Cambridge, Massachusetts: MIT Press.

Koop, Gary, and Simon N. Potter (1999), "Bayes Factors and Nonlinearity: Evidence from Economic Time Series," *Journal of Econometrics* 88, 251-281.

Krolzig, Hans-Martin (1997), *Markov-Switching Vector Autoregressions : Modelling, Statistical Inference, and Application to Business Cycle Analysis,* Berlin: Springer.

Lindgren, G. (1978), "Markov Regime Models for Mixed Distributions and Switching

Regressions," *Scandinavian Journal of Statistics* 5, 81-91.

Rabiner, Lawrence R. (1989), "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceedings of the IEEE* 77, 257-286.

Peria, Maria Soledad Martinez (2002), "A Regime-Switching Approach to the Study of Speculative Attacks: A Focus on EMS Crises," in James D. Hamilton and Baldev Raj, editors, *Advances in Markov-Switching Models,* Heidelberg: Physica-Verlag.

Poritz, Alan B. (1982), "Linear Predictive Hidden Markov Models and the Speech Signal," *Acoustics, Speech and Signal Processing, IEEE Conference on ICASSP '82,* vol. 7, 1291-1294.

Sims, Christopher, and Tao Zha (2004), "Were There Switches in U.S. Monetary Policy?", working paper, Princeton University.

Timmermann, Allan (200), "Moments of Markov Switching Models," *Journal of Econometrics* 96, 75-111.

Tjøstheim, Dag (1986), "Some Doubly Stochastic Time Series Models," *Journal of Time Series Analysis* 7, 51-72.

Yang, Min Xian (2000), "Some Properties of Vector Autoregressive Processes with Markov-Switching Coefficients," *Econometric Theory* 16, 23-43.