# Estimation of Markov Regime-Switching Regression Models with Endogenous Switching

| | |
|---|---|
| **Authors** | Chang-Jin Kim, Jeremy M. Piger, and Richard Startz |
| **Working Paper Number** | 2003-015C |
| **Revision Date** | June 2003 |
| **Citable Link** | https://doi.org/10.20955/wp.2003.015 |
| **Suggested Citation** | Kim, C.-J., Piger, J.M., Startz, R., 2003; Estimation of Markov Regime-Switching Regression Models with Endogenous Switching, Federal Reserve Bank of St. Louis Working Paper 2003-015. URL https://doi.org/10.20955/wp.2003.015 |

# Estimation of Markov Regime-Switching Regression Models with Endogenous Switching

**Chang-Jin Kim***
*Korea University and University of Washington*

**Jeremy Piger**
*Federal Reserve Bank of St. Louis*

**Richard Startz**
*University of Washington*

First Draft: March, 2003
This Draft: November, 2005

## Abstract

Following Hamilton (1989), estimation of Markov regime-switching regressions typically relies on the assumption that the latent state variable controlling regime change is exogenous. We relax this assumption and develop a parsimonious model of endogenous Markov regime-switching. Inference via maximum likelihood estimation is possible with relatively minor modifications to existing recursive filters. The model nests the exogenous switching model, yielding straightforward tests for endogeneity. In Monte Carlo experiments, maximum likelihood estimates of the endogenous switching model parameters were quite accurate, even in the presence of certain model misspecifications. As an application, we extend the volatility feedback model of equity returns given in Turner, Startz and Nelson (1989) to allow for endogenous switching.

Keywords: Endogeneity, Regime-Switching
JEL Classification: C13, C22, G12

---

* **Kim:** Dept. of Economics, Korea University, Anam-Dong, Seongbuk-ku, Seoul, 136-701, Korea, (cjkim@korea.ac.kr); **Piger:** Research Department, Federal Reserve Bank of St. Louis, 411 Locust St., St. Louis, MO 63166, USA, (piger@stls.frb.org); **Startz**: Dept. of Economics, University of Washington, Box 353330, Seattle, WA 98195, USA (startz@u.washington.edu). Startz acknowledges financial support from the Cecil and Jane Castor Professorship at the University of Washington. We thank the editor, two anonymous referees, Barry Arnold, Robert Beaver, Michael Dueker, James Morley, Charles Nelson and seminar participants at USC and the 2003 NBER Summer Institute for helpful comments. The views expressed in this paper are solely the responsibility of the authors, and should not be interpreted as reflecting the views of the Federal Reserve Bank of St. Louis or of any other person associated with the Federal Reserve System.

Recent decades have seen extensive interest in time-varying parameter models of macroeconomic and financial time series. One notable set of models are regime-switching regressions, which date to at least Quandt (1958). Goldfeld and Quandt (1973) introduced a particularly useful version of these models, referred to in the following as a Markov-switching model, in which the latent state variable controlling regime shifts follows a Markov-chain, and is thus serially dependent. In an influential article, Hamilton (1989) extended Markov-switching models to the case of dependent data, specifically an autoregression.

The vast literature generated by Hamilton (1989) typically assumes that the regime shifts are exogenous with respect to all realizations of the regression disturbance. In this paper we work with Markov-switching regressions of the type considered by Hamilton (1989) and various extensions, but relax the exogenous switching assumption. We develop a model of endogenous Markov regime-switching that is based on a probit specification for the realization of the latent state. The model is quite parsimonious, and admits a test for endogenous switching as a simple parameter restriction. The model parameters can be estimated via maximum likelihood with relatively minor modifications to the recursive filter in Hamilton (1989).

Why are we motivated to investigate Markov-switching regressions with endogenous switching? Many of the model's applications are in macroeconomics or finance in situations where it is natural to assume the state is endogenous. As an example, it is often the case that the estimated state variable has a strong business cycle correlation. This can be seen in recent applications of the regime-switching model to identified monetary VARs, such as Sims and Zha (2002) and Owyang (2002). It is not hard to imagine that the shocks to the regression, such as the macroeconomic shocks to the VAR, would be correlated with the business cycle. As another example, some applications of the model contain parameters that represent the reaction of agents to realization of the state (see for example Turner, Startz and Nelson, 1989). However, it is likely that agents do not observe the state, but instead draw inference based on some information set, the contents of which are unknown to the econometrician. Use of the actual state to proxy for this inference leads to a regression with measurement error in the explanatory variables, and thus endogeneity.

In order to evaluate the performance of maximum likelihood estimates of the endogenous switching model parameters, as well as tests for endogenous switching, we conduct a battery of Monte Carlo experiments. These experiments suggest that: 1) When the true Markov-switching process is endogenous, maximum likelihood estimation assuming exogenous switching yields biased parameter estimates, 2) Maximum likelihood estimates of the endogenous switching model were close to their true values, as were quasi-maximum likelihood estimates obtained from data generated by a non-Gaussian endogenous switching model, and 3) The likelihood ratio test for endogenous switching was close to having correct size.

As an application, we extend the volatility feedback model of equity returns given in Turner, Startz and Nelson (1989) to allow for endogenous switching. As discussed above, this model provides a setting in which we might reasonably expect the Markov-switching state variable to be endogenous. We find strong statistical evidence for endogenous switching in the model and that allowing for endogeneity has substantial effects on parameter estimates.

It should be noted that the model of endogenous switching developed in this paper has much in common with an earlier literature using switching regressions. This earlier literature, such as Maddala and Nelson (1975), was often concerned with endogenous switching, as the primary applications were in limited dependent variable contexts such as self-selection and market disequilibrium settings. The model we have presented here can be interpreted as an extension of the Maddala and Nelson (1975) approach, which was a model of independent switching, to the Hamilton (1989) regime-switching model, in which the state process is serially dependent.

In the next section we lay out a two-regime Markov-switching regression model with endogenous switching and discuss maximum likelihood estimation. Section 3 generalizes this model to the *N*-regime case. Section 4 gives the results of Monte Carlo experiments evaluating the performance of parameter inference and tests for endogenous switching. In Section 5 we present an empirical example based on a model of volatility feedback in equity markets taken from Turner, Startz and Nelson (1989). Section 6 concludes.

## 2. A Two-Regime Endogenous Switching Model

### *2.1 Model Specification*

Consider the following Gaussian regime-switching model for the sample path of a time series, $\{y_t\}_{t=1}^{T}$:

$$y_t = x_t' \beta_{S_t} + \sigma_{S_t} \varepsilon_t , \qquad (2.1)$$

$$\varepsilon_t \sim i.i.d. N(0,1) ,$$

where $y_t$ is scalar, $x_t$ is a ($k \times 1$) vector of observed exogenous or predetermined explanatory variables, which may include lagged values of $y_t$, and $S_t = i$ is the state variable. Denote the number of regimes by *N,* so that $i$ = 1, 2, …., *N*. We begin with the case where $N = 2$. In addition to aiding intuition, the two-regime case is a popular specification in applied work.[1]

The state variable is unobserved and is assumed to evolve according to a Markov chain with transition probabilities:

$$P(S_t = i \mid S_{t-1} = j, z_t) = P_{ij}(z_t)\,. \tag{2.2}$$

In (2.2), the transition probabilities are influenced by a ($q \times 1$) vector of observed exogenous or predetermined variables $z_t$, where $z_t$ may include elements of $x_t$. The Markov chain is assumed to evolve independently of all observations of those elements of $x_t$ not included in $z_t$. We assume the Markov process is stationary, with unconditional probabilities $P(S_t = i \mid \bar{z}) = P(S = i \mid \bar{z})$, $\forall\, t$.[2]

The transition probabilities (2.2) are constrained to be in [0,1] using a probit specification for $S_t$:

$$S_t = \begin{cases} 1 & \text{if} \quad \eta_t < \left(a_{S_{t-1}} + z_t^{'} b_{S_{t-1}}\right) \\ 2 & \text{if} \quad \eta_t \geq \left(a_{S_{t-1}} + z_t^{'} b_{S_{t-1}}\right) \end{cases}, \tag{2.3}$$

[1] As the regime ordering is arbitrary, we assume that the model in 2.1 is appropriately normalized to achieve identification. See Hamilton, Waggoner and Zha (2004) for detailed discussion of this issue.

[2] Several special cases of (2.2) are worth mentioning. The unrestricted model is the time-varying transition probability Markov-switching model of Goldfeld and Quandt (1973), Diebold, Lee and Weinbach (1994) and Filardo (1994). When the transition probabilities are not influenced by $S_{t-1}$, we have the time-varying transition probability independent switching model of Goldfeld and Quandt (1972). When the transition probabilities are not influenced by $z_t$, we have the fixed transition probability Markov-switching model of Goldfeld and Quandt (1973) and Hamilton (1989). When the transition probabilities are influenced by neither $z_t$ or $S_{t-1}$, we have the fixed transition probability independent switching model of Quandt (1972).

$\eta_t \sim i.i.d.N(0,1)$.

The transition probabilities are then:

$$p_{1j}(z_t) = P(\eta_t < (a_j + z_t' b_j)) = \Phi(a_j + z_t' b_j), \tag{2.4}$$

$$p_{2j}(z_t) = P(\eta_t \geq (a_j + z_t' b_j)) = 1 - \Phi(a_j + z_t' b_j),$$

where $\Phi$ is the standard normal cumulative distribution function.

To model endogenous switching, assume that the joint density function of $\varepsilon_t$ and $\eta_t$ is bivariate normal:

$$\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim N(0, \Sigma), \quad \Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \tag{2.5}$$

where $\varepsilon_t$ and $\eta_{t-h}$ are uncorrelated $\forall\, h \neq 0$. Regime-switching models found in time-series applications nearly always make the assumption that $\varepsilon_t$ is independent of $S_{t-h}$, $\forall$ *h,* which corresponds to the restriction that $\rho = 0$ in the model presented here.[3]

[3] In recent work, Chib and Dueker (2004) develop a non-Markov regime switching model in which observable variables are related to the sign of a Gaussian autoregressive latent state variable, the innovations to which are allowed to be correlated with the model residual through a bivariate normal specification as in (2.5). The authors develop Bayesian procedures to estimate this model.

*2.2 Maximum Likelihood Estimation*

Let $\Omega_t = (x_t^{'}, x_{t-1}^{'}, ..., x_1^{'}, z_t^{'}, z_{t-1}^{'}, ..., z_1^{'})^{'}$ and $\xi_t = (y_t, y_{t-1}, ..., y_1)^{'}$ be vectors containing observations observed through date *t*, and $\theta = (\beta_1, \sigma_1, a_1, b_1, \beta_2, \sigma_2, a_2, b_2, \rho)^{'}$ be the vector of model parameters. The conditional likelihood function for the observed data $\zeta_t$ is constructed as $L(\theta) = \prod_{t=1}^{T} f(y_t \mid \Omega_t, \xi_{t-1}; \theta)$, where:

$$f(y_t \mid \Omega_t, \xi_{t-1}; \theta) \tag{2.6}$$

$$= \sum_i \sum_j f(y_t \mid S_t = i, S_{t-1} = j, \Omega_t, \xi_{t-1}; \theta) \Pr(S_t = i, S_{t-1} = j \mid \Omega_t, \xi_{t-1}; \theta).$$

The weighting probability in (2.6) is computed recursively by applying Bayes' Rule given the initial unconditional probabilities $P(S_0 = i \mid \bar{z}) = P(S = i \mid \bar{z})$:

$$\Pr(S_t = i, S_{t-1} = j \mid \Omega_t, \xi_{t-1}; \theta) = P_{ij}(z_t) \Pr(S_{t-1} = j \mid \Omega_t, \xi_{t-1}; \theta), \tag{2.7}$$

$$\Pr(S_t = i \mid \Omega_{t+1}, \xi_t; \theta) = \Pr(S_t = i \mid \Omega_t, \xi_t; \theta)$$

$$= \frac{1}{f(y_t \mid \Omega_t, \xi_{t-1}; \theta)} \sum_j f(y_t \mid S_t = i, S_{t-1} = j, \Omega_t, \xi_{t-1}; \theta) \Pr(S_t = i, S_{t-1} = j \mid \Omega_t, \xi_{t-1}; \theta).$$

To complete the recursion in (2.6)-(2.7), we require the regime-dependent conditional density function, $f(y_t \mid S_t = i, S_{t-1} = j, \Omega_t, \xi_{t-1}; \theta)$. For the exogenous switching case, when $\rho = 0$, this density function is Gaussian:

$$f\left(y_t \mid S_t = i, S_{t-1} = j, \Omega_t, \xi_{t-1}; \theta\right) = \frac{1}{\sigma_i}\phi\left(\frac{y_t - x_t'\beta_i}{\sigma_i}\right), \tag{2.8}$$

where $\phi$ is the standard normal probability density function. However, for more general values of $\rho \in (-1,1)$, $f\left(y_t \mid S_t = i, S_{t-1} = j, \Omega_t, \xi_{t-1}; \theta\right)$ is:[4]

$$f\left(y_t \mid S_t = 1, S_{t-1} = j, \Omega_t, \xi_{t-1}; \theta\right) = \frac{\phi\left(\frac{y_t - x_t'\beta_1}{\sigma_1}\right)\Phi\left(\frac{\left(a_j + z_t'b_j\right) - \rho\left(\frac{y_t - x_t'\beta_1}{\sigma_1}\right)}{\sqrt{1-\rho^2}}\right)}{\sigma_1 p_{1j}(z_t)}, \tag{2.9}$$

$$f\left(y_t \mid S_t = 2, S_{t-1} = j, \Omega_t, \xi_{t-1}; \theta\right) = \frac{\phi\left(\frac{y_t - x_t'\beta_2}{\sigma_2}\right)\Phi\left(\frac{-\left(a_j + z_t'b_j\right) + \rho\left(\frac{y_t - x_t'\beta_2}{\sigma_2}\right)}{\sqrt{1-\rho^2}}\right)}{\sigma_2 p_{2j}(z_t)}.$$

The appendix provides a derivation of (2.9).

When $S_t$ is endogenous, maximum likelihood estimation assuming $S_t$ is exogenous, and thus based on the distribution in (2.8), is inconsistent in general. To see this, note that:

$$E\left(\varepsilon_t \mid S_t = 1, S_{t-1} = j; \theta\right) = E\left(\varepsilon_t \mid \eta_t < \left(a_j + z_t'b_j\right)\right) = -\rho\frac{\phi\left(a_j + z_t'b_j\right)}{\Phi\left(a_j + z_t'b_j\right)}, \tag{2.10}$$

$$E\left(\varepsilon_t \mid S_t = 2, S_{t-1} = j; \theta\right) = E\left(\varepsilon_t \mid \eta_t \geq \left(a_j + z_t'b_j\right)\right) = \rho\frac{\phi\left(a_j + z_t'b_j\right)}{1 - \Phi\left(a_j + z_t'b_j\right)}.$$

[4] The density (2.9) belongs to the "skew-normal" family of density functions, which are commonly credited to Azzalini (1985). See Arnold and Beaver (2002) for a survey of this literature.

Thus, when $\rho \neq 0$, the regime-dependent conditional mean of $\varepsilon_t$ is non-zero, implying that maximum likelihood estimates based on (2.8) suffer from the ordinary problem of omitted variables. Another, less obvious, source of inconsistency arises because $f(y_t \mid S_t = i, S_{t-1} = j, \Omega_t, \xi_{t-1}; \theta)$ is non-Gaussian when $\rho \neq 0$, as is clear from (2.9). In this case maximum likelihood estimation based on (2.8) is Quasi-maximum likelihood estimation, which, as pointed out in Campbell (2002), is inconsistent for regime-switching models in general.

*2.3 Testing for Endogeneity*

In the model of endogenous switching presented above, the null hypothesis that $S_t$ is exogenous is equivalent to the scalar restriction $\rho = 0$. Thus, a test for exogeneity can be carried out by any suitable test of the restriction $\rho = 0$, such as a Wald or likelihood ratio test.

## 3. An N-Regime Endogenous Switching Model

In this section we generalize the two-regime Gaussian endogenous-switching model presented in Section 2 to *N* regimes. We begin by modifying the probit specification of the transition probabilities given in (2.3). Suppose the realization of $S_t$ is now determined by the outcome of $\eta_t \sim i.i.d.N(0,1)$ as follows:

$$S_t = \left\{ \begin{array}{lllcl} 1 & \text{if} & -\infty & \leq \eta_t < & \left(a_{1,j} + z_t^{'} b_{1,j}\right) \\ 2 & \text{if} & \left(a_{1,j} + z_t^{'} b_{1,j}\right) & \leq \eta_t < & \left(a_{2,j} + z_t^{'} b_{2,j}\right) \\ . & & & & \\ . & & & & \\ N-1 & \text{if} & \left(a_{N-2,j} + z_t^{'} b_{N-2,j}\right) & \leq \eta_t < & \left(a_{N-1,j} + z_t^{'} b_{N-1,j}\right) \\ N & \text{if} & \left(a_{N-1,j} + z_t^{'} b_{N-1,j}\right) & \leq \eta_t < & \infty \end{array} \right\}. \quad (3.1)$$

The transition probabilities, $p_{ij}(z_t)$, are then given as follows:

$$p_{ij}\left(z_t\right) = \Phi\left(c_{i,j,t}\right) - \Phi\left(c_{i-1,j,t}\right), \quad (3.2)$$

where $c_{0,j,t} = -\infty$, $c_{N,j,t} = \infty$, and $c_{i,j,t} = \left(a_{i,j} + z_t^{'} b_{i,j}\right)$ for $0 < i < N$.

Again, to model endogenous switching, assume that the joint density of $\varepsilon_t$ and $\eta_t$ is bivariate normal as in (2.5):

$$\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim N(0, \Sigma), \quad \Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \quad (2.5)$$

where $\varepsilon_t$ and $\eta_{t-h}$ are uncorrelated $\forall\, h \neq 0$. Let the vector of model parameters be $\theta = \left(\theta_1^{'}, \theta_2^{'}, ..., \theta_N^{'}, \rho\right)^{'}$, where $\theta_i = \left(\beta_i, \sigma_i, a_i, b_i\right)^{'}$. Given $f\left(y_t \mid S_t = i, S_{t-1} = j, \Omega_t, \xi_{t-1}; \theta\right)$, the likelihood function, $L(\theta)$, can again be constructed using the recursion in (2.6)-(2.7). It is shown in the appendix that:

$$f\left(y_t \mid S_t = i, S_{t-1} = j, \Omega_t, \xi_{t-1}; \theta\right)$$

$$= \frac{\phi\left(\frac{y_t - x_t'\beta_i}{\sigma_i}\right)\left(\Phi\left(\frac{c_{i,j,t} - \rho\left(\frac{y_t - x_t'\beta_i}{\sigma_i}\right)}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{c_{i-1,j,t} - \rho\left(\frac{y_t - x_t'\beta_i}{\sigma_i}\right)}{\sqrt{1-\rho^2}}\right)\right)}{\sigma_i p_{ij}(z_t)}. \tag{3.3}$$

Finally, as with the two regime endogenous switching model, a test of the null hypothesis that $S_t$ is exogenous is equivalent to a test of the restriction $\rho = 0$.

## 4. Monte Carlo Analysis

In this section we provide Monte Carlo evidence regarding maximum likelihood estimation of the endogenous switching model and associated tests for endogeneity. Given its prominence in the applied literature, we focus on the two-regime model with fixed, Markov-switching transition probabilities, so that $b_1 = b_2 = 0$. We begin by evaluating the performance of maximum likelihood estimation when the true model is the endogenous switching model presented in Section 2 with varying levels of $\rho$. We then investigate the sensitivity of maximum likelihood estimation based on the joint normality assumption in (2.5) to departures from this Gaussian assumption in the data generating process. Such a departure renders the estimator based on (2.5) a Quasi-maximum likelihood (QML) estimator, which is inconsistent for Markov-switching models in general. Our Monte Carlo experiments then provide some limited evidence of how badly the QML estimator performs in practice.

For each Monte Carlo experiment, 1000 simulated series are generated from the model given in (2.1)-(2.3). We consider two sample sizes for the simulated series, $T = 200$ and $T = 500$.

For each simulation, we generate the vector of exogenous explanatory variables as $x_t = \begin{bmatrix} 1 & x_t^* \end{bmatrix}$, where $x_t^* \sim i.i.d.N(0,2)$, and fix the vector of regime switching parameters to $\beta_1 = (\beta_{0,1}, \beta_{1,1})'$ $= (1.0, 1.0)'$, $\beta_2 = (\beta_{0,2}, \beta_{1,2})' = (-1.0, -1.0)'$, $\sigma_1 = 0.33$, $\sigma_2 = 0.67$. We consider three different sets of transition probabilities corresponding to moderate persistence ( $p_{11} = 0.7$, $p_{22} = 0.7$ ), high persistence ( $p_{11} = 0.9$, $p_{22} = 0.9$ ), and differential persistence ( $p_{11} = 0.7$, $p_{22} = 0.9$ ). We also consider three different values for $\rho$, corresponding to high correlation $\rho = 0.9$, moderate correlation $\rho = 0.5$, and zero correlation $\rho = 0$.

We consider two different joint density functions for $\varepsilon_t$ and $\eta_t$, labeled DGP1 and DGP2. DGP1 is the bivariate normal distribution in (2.5). DGP2 relaxes this joint normality assumption. Instead, $\varepsilon_t$ is generated as a standard normal random variable, while $\eta_t$ is generated as a weighted sum of $\varepsilon_t$ and a *t*-distributed random variable with four degrees of freedom. The weighting is calibrated so that $(\varepsilon_t, \eta_t)'$ has covariance matrix:

$$\Sigma = \begin{bmatrix} 1 & \rho\gamma_4 \\ \rho\gamma_4 & \gamma_4^2 \end{bmatrix},$$

where $\gamma_4^2 = 2$ is the variance of a *t*-distributed random variable with four degrees of freedom.

For each simulated time series, two maximum likelihood estimates and associated standard errors are computed.[5] The first, which we label the "exogenous" estimator, assumes that $\rho = 0$, and is thus based on the recursion in (2.6)-(2.7), using (2.8) to measure

[5] All computations were performed in GAUSS 6.0 using the OPTMUM numerical optimization package. Standard errors were based on second derivatives of the log-likelihood function.

$f(y_t \mid S_t = i, S_{t-1} = j, \Omega_t, \xi_{t-1}; \theta)$. The second, which we label the "endogenous" estimator, allows for $\rho \neq 0$, and is thus based on the recursion in (2.6)-(2.7), using (2.9) to measure $f(y_t \mid S_t = i, S_{t-1} = j, \Omega_t, \xi_{t-1}; \theta)$. Finally, we also record the outcome of 5% nominal size Wald and likelihood ratio tests of the null hypothesis $\rho = 0$. For those cases where $\rho = 0$ in the data generating process, these tests document the empirical size of the 5% nominal size tests. For those cases where $\rho \neq 0$, we use size-adjusted critical values, taken from the Monte Carlo simulations generated with $\rho = 0$, to measure the power of the tests.

Tables 1-5 show the results of the Monte Carlo experiments investigating maximum likelihood estimation of the endogenous switching model. Each table shows the mean of the 1000 maximum likelihood point estimates of $\beta_1$, $\beta_2$, $\sigma_1$, $\sigma_2$, the mean of the standard errors for these parameter estimates, and the rejection rate of the tests of $\rho = 0$.[6] Table 1 gives results when the data generating process has exogenous switching, that is $\rho = 0$. In this case, both the exogenous and endogenous estimator are exact maximum likelihood estimators, but the endogenous estimator is inefficient, as it does not restrict $\rho = 0$. As is clear from the table, for both sample sizes and all values of the transition probabilities, the exogenous and endogenous estimators produce estimates of the model parameters that are very close to their true values. As would be expected, the endogenous estimator is less efficient than the exogenous estimator, with the average standard error of the estimates consistently higher for the endogenous estimator.

Tables 2 and 3 give results when the true data generating process includes endogenous switching of the form in DGP1. The tables demonstrate the estimation bias that occurs when the endogenous state variable is treated as exogenous in estimation. When the exogenous estimator is used, the mean estimates of $\beta_{0,1}$ and $\beta_{0,2}$ are far from their true values, with the bias larger for higher values of $\rho$. The mean estimates of $\sigma_1$ and $\sigma_2$ are also biased downward. Note that the mean estimates are nearly identical in the $T = 200$ and $T = 500$ cases, suggesting the bias is not a small sample phenomenon. Also note that the estimates of $\beta_{1,1}$ and $\beta_{1,2}$ are close to their true values. The accuracy of these parameter estimates can be traced to the model assumption, maintained in the Monte Carlo samples, that $x_t^*$ is independent of the endogenous state variable $S_t$. Finally, Tables 2 and 3 also demonstrate that the endogenous estimator produces very accurate estimates of the endogenous switching model. Indeed, for both sample sizes and all values of the transition probabilities and $\rho$ considered, the mean parameter estimates are nearly identical to their true values.

Tables 4 and 5 present results for DGP2, that is when the joint density between $\varepsilon_t$ and $\eta_t$ is non-normal. For the particular joint density function considered, the approximation provided by the normality assumption is quite good. Again, for both sample sizes and all values of the transition probabilities and $\rho$ considered, the mean parameter estimates from the endogenous estimator are very close to their true values. While this result may not generalize to non-normal distributions more generally, it is suggestive that the quality of the endogenous estimator procedure is not hyper-sensitive to the joint-normality assumption. Not surprisingly, the exogenous estimator, which ignores the potential for the state variable to be endogenous all together, continues to produce biased parameter estimates under DGP2.

[6] Model estimation also produces estimates of the transition probabilities, and, in the case of the endogenous estimator, the correlation parameter $\rho$. Although not reported, results for these parameter estimates are qualitatively similar to those for the conditional mean and variance parameters of the regression model.

Tables 6 and 7 report the empirical size and size-adjusted power of the Wald and likelihood ratio test of $\rho = 0$. From Table 6, the 5% Wald test is somewhat oversized, with rejection rates as high as 12% when $T = 200$. However, the 5% likelihood ratio test has roughly correct size for both sample sizes considered. Table 7 demonstrates that the likelihood ratio and Wald test have similar size-adjusted power against the alternatives considered.

Overall, the Monte Carlo experiments confirm that maximum likelihood estimates using the endogenous estimator are quite good for the examples considered, while the exogenous estimator produces substantially biased parameter estimates when the true process has endogenous switching. Also, the likelihood ratio test appears to be a fairly reliable test for endogenous switching. In the next section we turn to an empirical application of the endogenous switching model.

## 5. Application: Measurement Error and Estimation of the Volatility Feedback Effect

A stylized fact of U.S. equity return data is that the volatility of realized returns is time-varying and predictable. Given this, classic portfolio theory would imply that the equity risk premium should also be time-varying and respond positively to the expectation of future volatility. However, the data suggest that realized returns and realized volatility, as measured by squared returns, are negatively correlated.[7]

One explanation for the observed data is that while investors do require an increase in expected return in exchange for expected future volatility, they are often surprised by news about realized volatility. This "volatility feedback effect" creates a reduction in prices in the period in which the increase in volatility is realized. If the volatility feedback effect is strong enough, it

[7] For a recent discussion of this result, see Brandt and Kang (2004).

may create a negative contemporaneous correlation between realized returns and volatility in the data. The volatility feedback effect has been investigated extensively in the literature, see for example French, Schwert and Stambaugh (1987), Turner, Startz and Nelson (1989), Campbell and Hentschell (1992), Bekaert and Wu (2000) and Kim, Morley and Nelson (2002).

Turner, Startz and Nelson (1989), hereafter TSN, model the volatility feedback effect with a Markov-switching model:

$$r_t = \theta_1 E\left(\sigma_{S_t}^2 \mid \Psi_{t-1}\right) + \theta_2 \left(E\left(\sigma_{S_t}^2 \mid \Psi_t^*\right) - E\left(\sigma_{S_t}^2 \mid \Psi_{t-1}\right)\right) + \sigma_{S_t} \varepsilon_t , \qquad (5.1)$$

$$\varepsilon_t \sim i.i.d. N(0,1),$$

where $S_t$ is a discrete Markov-switching variable taking on values 1 or 2, with transition probabilities $p_{ij}$ parameterized as in equation (2.4). For normalization we assume $\sigma_2^2 > \sigma_1^2$, so that $S_t = 2$ is the high volatility state.

The model in (5.1) is motivated as follows. At the beginning of period *t*, the risk premium, $\theta_1 E\left(\sigma_{S_t}^2 \mid \Psi_{t-1}\right)$, is determined based on the expectation of period *t* volatility formed with information available at the end of period *t*-1. During period *t* additional information regarding volatility is observed. By the end of period *t*, this information is collected in the information set $\Psi_t^*$. When $E\left(\sigma_{S_t}^2 \mid \Psi_t^*\right) \neq E\left(\sigma_{S_t}^2 \mid \Psi_{t-1}\right)$, information about volatility revealed during the period has surprised agents. If $\theta_2 < 0$, surprises that reveal greater probability of the high-variance state are viewed negatively by investors, and thus reduce the contemporaneous return.

One estimation difficulty with the model in (5.1) is that there exists a discrepancy between the investors' and the econometrician's data set. In particular, while $\Psi_{t-1}$ may be summarized by all data up to $t$-1, that is $\Psi_t = \{r_{t-1}, r_{t-2}, \ldots\}$, the information set $\Psi_t^*$ includes information that is not summarized in the researcher's data set on observed returns. This is because, while the researcher's data set is collected discretely at the beginning of each period, the market participants continuously observe trades that occur during the period.

To handle this estimation difficulty, TSN use the actual volatility, $\sigma_{S_t}^2$, as a proxy for $E(\sigma_{S_t}^2 \mid \Psi_t^*)$. That is, they estimate:

$$r_t = \theta_1 E(\sigma_{S_t}^2 \mid \Psi_{t-1}) + \theta_2 (\sigma_{S_t}^2 - E(\sigma_{S_t}^2 \mid \Psi_{t-1})) + \sigma_{S_t} u_t \qquad (5.2)$$

$$u_t \sim N(0,1)$$

In essence, this approximation replaces the estimated probability of the state, $P(S_t = i \mid \Psi_t^*)$, with one if $S_t = i$ and zero otherwise. Assuming these differ, this introduces classical measurement error into the state variable in the estimated equation, thus rendering it endogenous.

The existing literature estimates (5.2) assuming the state variable is exogenous. However, the techniques developed in Section 2 can be used to estimate the volatility feedback model allowing for endogeneity, as well as to test for endogeneity. Here we estimate (5.2) using monthly returns for a value-weighted portfolio of all NYSE-listed stocks in excess of the one-month Treasury Bill rate over the sample period 1952-1999, the same data as used in Kim, Morley and Nelson (2002). Table 8 summarizes the results.

The first panel of Table 8 shows estimates when endogeneity is ignored. These estimates, which are similar to those in TSN, are consistent with both a positive relationship between the risk premium and expected future volatility ($\theta_1 > 0$) and a substantial volatility feedback effect ($\theta_2 << 0$). The estimates also suggest a dominant volatility feedback effect, that is $\theta_1$ is very small relative to $\theta_2$. The second panel shows the estimates when endogeneity is allowed, so that the correlation parameter $\rho$ is estimated. The estimate of $\rho$ is substantial, equaling -0.40. Both the Wald and likelihood ratio test reject the null hypothesis that $\rho = 0$ at the 10% level (the p-values are 0.026 and 0.081 respectively). The primary difference in the parameter estimates is for the volatility feedback coefficient $\theta_2$, which is estimated to be about one-third smaller when endogeneity is allowed than when it is ignored. Thus, while there is still evidence of a strong volatility feedback effect, it is substantially smaller than that implied by the model with no allowance for endogeneity.

**6. Conclusion**

We have developed a model of Markov-switching in which the latent state variable controlling the regime shifts is endogenously determined. The model is quite parsimonious, and admits a test for endogenous switching as a simple parameter restriction. The model parameters can be estimated via maximum likelihood with relatively small modifications to the recursive filter in Hamilton (1989). Monte Carlo experiments suggest that maximum likelihood estimation of the endogenous switching model and the likelihood ratio test for endogeneity performed quite well for the data generating processes considered. We apply the model to test for endogenous switching in the volatility feedback model of equity returns given in Turner, Startz and Nelson (1989).

## Appendix

*Derivation of (2.9) and (3.3):*

We proceed by generalizing the derivation of the univariate skew-normal density function given in Arnold and Beaver (2002). The random variables described in (2.5) can be written as:

$$\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} = A \begin{bmatrix} \varepsilon_t \\ \omega_t \end{bmatrix} , \tag{A.1}$$

where $\omega_t \sim i.i.d.N(0,1)$, and $A = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix}$ is the Cholesky decomposition of $\Sigma$, so that $AA^{'} = \Sigma$. From (A.1):

$$\eta_t = \rho\varepsilon_t + \sqrt{1-\rho^2}\,\omega_t . \tag{A.2}$$

We can then write, suppressing $\Omega_t, \xi_{t-1}$, and $\theta$ from the conditioning set for convenience:

$$\begin{aligned} & f(y_t \mid S_t = i, S_{t-1} = j) \\ & \quad = f\left(y_t \mid \left(c_{i-1,j,t}\right) \le \eta_t < \left(c_{i,j,t}\right)\right) \\ & \quad = f\left( y_t \mid \frac{\left(c_{i-1,j,t}\right) - \rho\varepsilon_t}{\sqrt{1-\rho^2}} \le \omega_t < \frac{\left(c_{i,j,t}\right) - \rho\varepsilon_t}{\sqrt{1-\rho^2}} \right), \end{aligned} \tag{A.3}$$

where $c_{i-1,j,t}$ and $c_{i,j,t}$ are defined in Section 3. Consider the cumulative probability distribution function:

$$\Pr\left( y_t < g \mid \frac{(c_{i-1,j,t}) - \rho\varepsilon_t}{\sqrt{1-\rho^2}} \le \omega_t < \frac{(c_{i,j,t}) - \rho\varepsilon_t}{\sqrt{1-\rho^2}} \right) \tag{A.4}$$

$$= \frac{\Pr\left( y_t < g, \frac{(c_{i-1,j,t}) - \rho\varepsilon_t}{\sqrt{1-\rho^2}} \le \omega_t < \frac{(c_{i,j,t}) - \rho\varepsilon_t}{\sqrt{1-\rho^2}} \right)}{\Pr\left( \frac{(c_{i-1,j,t}) - \rho\varepsilon_t}{\sqrt{1-\rho^2}} \le \omega_t < \frac{(c_{i,j,t}) - \rho\varepsilon_t}{\sqrt{1-\rho^2}} \right)}$$

The denominator of (A.4) is:

$$\Pr\left( \frac{(c_{i-1,j,t}) - \rho\varepsilon_t}{\sqrt{1-\rho^2}} \le \omega_t < \frac{(c_{i,j,t}) - \rho\varepsilon_t}{\sqrt{1-\rho^2}} \right) = p_{ij}(z_t) = \Phi(c_{i,j,t}) - \Phi(c_{i-1,j,t}). \tag{A.5}$$

The numerator of (A.4) is:

$$\Pr\left( y_t < g, \frac{(c_{i-1,j,t}) - \rho\varepsilon_t}{\sqrt{1-\rho^2}} \le \omega_t < \frac{(c_{i,j,t}) - \rho\varepsilon_t}{\sqrt{1-\rho^2}} \right)$$

$$= \int_{-\infty}^{g} \int_{((c_{i-1,j,t}) - \rho\varepsilon_t)/\sqrt{1-\rho^2}}^{((c_{i,j,t}) - \rho\varepsilon_t)/\sqrt{1-\rho^2}} f(y_t, \omega_t) d\omega d\,\varepsilon$$

$$= \int_{-\infty}^{g} \int_{((c_{i-1,j,t}) - \rho\varepsilon_t)/\sqrt{1-\rho^2}}^{((c_{i,j,t}) - \rho\varepsilon_t)/\sqrt{1-\rho^2}} \frac{1}{\sigma_i} \phi\left( \frac{y_t - x_t^{'}\beta_i}{\sigma_i} \right) f(\omega_t) d\omega d\varepsilon$$

$$= \int_{-\infty}^{g} \frac{1}{\sigma_i} \phi\left(\frac{y_t - x_t^{'}\beta_i}{\sigma_i}\right) \left( \Phi\left( \frac{c_{i,j,t} - \rho\left(\frac{y_t - x_t^{'}\beta_i}{\sigma_i}\right)}{\sqrt{1-\rho^2}} \right) - \Phi\left( \frac{c_{i-1,j,t} - \rho\left(\frac{y_t - x_t^{'}\beta_i}{\sigma_i}\right)}{\sqrt{1-\rho^2}} \right) \right) d\varepsilon \,. \qquad \text{(A.6)}$$

Combining (A.5)-(A.6) and differentiating with respect to $g$ yields:

$$f(y_t \mid S_t = i, S_{t-1} = j) \qquad \text{(3.3)}$$

$$= \frac{\phi\left(\frac{y_t - x_t^{'}\beta_i}{\sigma_i}\right) \left( \Phi\left( \frac{c_{i,j,t} - \rho\left(\frac{y_t - x_t^{'}\beta_i}{\sigma_i}\right)}{\sqrt{1-\rho^2}} \right) - \Phi\left( \frac{c_{i-1,j,t} - \rho\left(\frac{y_t - x_t^{'}\beta_i}{\sigma_i}\right)}{\sqrt{1-\rho^2}} \right) \right)}{\sigma_i p_{ij}(z_t)},$$

which is the density function (3.3). In the case where $N = 2$, we have:

$$f\left(y_t \mid S_t = 1, S_{t-1} = j\right) = \frac{\phi\left(\frac{y_t - x_t^{'}\beta_1}{\sigma_1}\right) \Phi\left( \frac{\left(a_{1,j} + z_t^{'}b_{1,j}\right) - \rho\left(\frac{y_t - x_t^{'}\beta_1}{\sigma_1}\right)}{\sqrt{1-\rho^2}} \right)}{\sigma_1 p_{1j}(z_t)}$$

$$f\left(y_t \mid S_t = 2, S_{t-1} = j\right) = \frac{\phi\left(\frac{y_t - x_t^{'}\beta_2}{\sigma_2}\right) \Phi\left( \frac{-\left(a_{1,j} + z_t^{'}b_{1,j}\right) + \rho\left(\frac{y_t - x_t^{'}\beta_2}{\sigma_2}\right)}{\sqrt{1-\rho^2}} \right)}{\sigma_2 p_{2j}(z_t)},$$

which, upon renaming $a_{1,j} = a_j$ and $b_{1,j} = b_j$, is the density function in (2.9).

## References


Arnold, B.C. and R.J. Beaver, 2002, Skewed Multivariate Models Related to Hidden Truncation and / or Selective Reporting, *Sociedad de Estadistica e Investigacion Operativa,* 11, 7-54.

Azzalini, A., 1985, A Class of Distributions which Includes the Normal Ones, *Scandinavian Journal of Statistics,* 12, 171-178.

Bekaert, G. and G. Wu, 2000, Asymmetric Volatility and Risk in Equity Markets, *Review of Financial* Studies, 13, 1-42.

Brandt, M. and Q. Kang, 2004, On the Relationship between the Conditional Mean and Volatility of Stock Returns: A Latent VAR Approach, 72, 217-257.

Campbell, J. Y. and L. Hentschel, 1992, No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns, *Journal of Financial Economics*, 31, 281-318.

Campbell, S.D., 2002, Specification Testing and Semiparametric Estimation of Regime Switching Models: An Examination of the US Short Term Interest Rate, Brown University Department of Economics Working Paper #2002-26.

Chib, S. and M. Dueker, 2004, Non-Markovian Regime Switching with Endogenous States and Time Varying State Lengths, mimeo.

Diebold, F., Lee, J-H. and G. Weinbach, 1994, Regime Switching with Time-Varying Transition Probabilities, in *Non-stationary Time Series Analysis and Cointegration,* ed. C. Hargreaves, Oxford University Press, Oxford, U.K.

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

French, K.R., Schwert, G. and R. F. Stambaugh, 1987, Expected Stock Returns and Volatility, *Journal of Financial Economics,* 19, 3-29.

Goldfeld, S.M. and R.E. Quandt, 1972, in *Nonlinear Methods in Econometrics*, North Holland, Amsterdam.

Goldfeld, S.M. and R.E. Quandt, 1973, A Markov Model for Switching Regressions, *Journal of Econometrics,* 1, 3-16.

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

Hamilton, J.D. Waggoner, D.F., and T. Zha, 2004, Normalization in Econometrics, Federal Reserve Bank of Atlanta Working Paper # 2004-13.

Kim, C.J., Morley, J.C. and C.R. Nelson, 2002, Is There a Positive Relationship Between Stock Market Volatility and the Equity Premium?, *Journal of Money, Credit and Banking*, forthcoming.

Maddala, G.S. and F. Nelson, 1975, Switching Regression Models with Exogenous and Endogenous Switching, *Proceedings of the American Statistical Association,* 423-426.

Owyang, M., 2002, Modeling Volcker as a Non-Absorbing State: Agnostic Identification of a Markov-Switching VAR, working paper, Federal Reserve Bank of St. Louis.

Quandt, R.E., 1958, The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes, *Journal of the American Statistical Association,* 53, 873-880.

Quandt, R.E., 1972, A New Approach to Estimating Switching Regressions, *Journal of the American Statistical Association*, 67, 306-310.

Sims, C. and T. Zha, 2002, Macroeconomic Switching, mimeo.

Turner, C. M., Startz, R. and C.R. Nelson, 1989, A Markov Model of Heteroskedasticity, Risk, and Learning in the Stock Market, *Journal of Financial Economics*, 25, 3-22.

**Table 1**
Monte Carlo Results: Estimation
DGP 1, $\rho = 0$ (exogenous switching)

| **$T$ = 200** | $\beta_{0,1} = 1.0$ | $\beta_{0,2} = -1.0$ | $\beta_{1,1} = 1.0$ | $\beta_{1,2} = -1.0$ | $\sigma_1 = 0.33$ | $\sigma_2 = 0.67$ |
|---|---|---|---|---|---|---|
| $p_{11} = 0.7$ $p_{22} = 0.7$ | | | | | | |
| **Exog. Estimator** | 1.00 (0.04) | -1.00 (0.07) | 1.00 (0.02) | -1.00 (0.03) | 0.33 (0.03) | 0.66 (0.05) |
| **Endog. Estimator** | 0.99 (0.08) | -0.99 (0.16) | 1.00 (0.02) | -1.00 (0.03) | 0.34 (0.03) | 0.68 (0.06) |
| $p_{11} = 0.7$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.00 (0.05) | -1.00 (0.06) | 1.00 (0.03) | -1.00 (0.03) | 0.32 (0.04) | 0.66 (0.04) |
| **Endog. Estimator** | 1.00 (0.09) | -1.00 (0.07) | 1.00 (0.03) | -1.00 (0.03) | 0.32 (0.04) | 0.67 (0.04) |
| $p_{11} = 0.9$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.00 (0.04) | -1.00 (0.07) | 1.00 (0.02) | -1.00 (0.03) | 0.33 (0.03) | 0.66 (0.05) |
| **Endog. Estimator** | 1.00 (0.04) | -1.00 (0.08) | 1.00 (0.02) | -1.00 (0.03) | 0.33 (0.03) | 0.66 (0.05) |
| **$T$ = 500** | $\beta_{0,1} = 1.0$ | $\beta_{0,2} = -1.0$ | $\beta_{1,1} = 1.0$ | $\beta_{1,2} = -1.0$ | $\sigma_1 = 0.33$ | $\sigma_2 = 0.67$ |
| $p_{11} = 0.7$ $p_{22} = 0.7$ | | | | | | |
| **Exog. Estimator** | 1.00 (0.02) | -1.00 (0.04) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.02) | 0.67 (0.03) |
| **Endog. Estimator** | 1.00 (0.05) | -1.00 (0.10) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.02) | 0.67 (0.03) |
| $p_{11} = 0.7$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.00 (0.03) | -1.00 (0.04) | 1.00 (0.02) | -1.00 (0.02) | 0.33 (0.02) | 0.67 (0.03) |
| **Endog. Estimator** | 1.00 (0.06) | -1.00 (0.05) | 1.00 (0.02) | -1.00 (0.02) | 0.33 (0.02) | 0.67 (0.03) |
| $p_{11} = 0.9$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.00 (0.02) | -1.00 (0.04) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.02) | 0.67 (0.03) |
| **Endog. Estimator** | 1.00 (0.03) | -1.00 (0.05) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.02) | 0.67 (0.03) |

Notes: Each cell contains the mean of the maximum likelihood point estimates and the mean of the standard errors of these estimates from the Monte Carlo experiment. Exog. estimator refers to the maximum likelihood estimator assuming the state process is exogenous, so that $\rho = 0$. Endog. estimator refers to the maximum likelihood estimator allowing the state process to be endogenous, so that $\rho \in (-1, 1)$.

**Table 2**
Monte Carlo Results: Estimation
DGP 1, $\rho = 0.5$ (endogenous switching)

| ***T* = 200** | $\beta_{0,1} = 1.0$ | $\beta_{0,2} = -1.0$ | $\beta_{1,1} = 1.0$ | $\beta_{1,2} = -1.0$ | $\sigma_1 = 0.33$ | $\sigma_2 = 0.67$ |
|---|---|---|---|---|---|---|
| $p_{11} = 0.7$ $p_{22} = 0.7$ | | | | | | |
| **Exog. Estimator** | 1.11 (0.03) | -1.23 (0.07) | 1.00 (0.02) | -1.00 (0.03) | 0.30 (0.02) | 0.62 (0.05) |
| **Endog. Estimator** | 1.00 (0.07) | -1.00 (0.13) | 1.00 (0.02) | -1.00 (0.03) | 0.33 (0.03) | 0.67 (0.07) |
| $p_{11} = 0.7$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.15 (0.05) | -1.10 (0.05) | 1.00 (0.03) | -1.00 (0.03) | 0.31 (0.04) | 0.65 (0.04) |
| **Endog. Estimator** | 1.00 (0.08) | -1.00 (0.07) | 1.00 (0.02) | -1.00 (0.03) | 0.33 (0.03) | 0.67 (0.05) |
| $p_{11} = 0.9$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.06 (0.03) | -1.11 (0.07) | 1.00 (0.02) | -1.00 (0.03) | 0.32 (0.02) | 0.65 (0.05) |
| **Endog. Estimator** | 1.00 (0.04) | -1.00 (0.08) | 1.00 (0.02) | -1.00 (0.03) | 0.33 (0.04) | 0.66 (0.05) |
| ***T* = 500** | $\beta_{0,1} = 1.0$ | $\beta_{0,2} = -1.0$ | $\beta_{1,1} = 1.0$ | $\beta_{1,2} = -1.0$ | $\sigma_1 = 0.33$ | $\sigma_2 = 0.67$ |
| $p_{11} = 0.7$ $p_{22} = 0.7$ | | | | | | |
| **Exog. Estimator** | 1.11 (0.02) | -1.23 (0.04) | 1.00 (0.01) | -1.00 (0.02) | 0.31 (0.02) | 0.63 (0.03) |
| **Endog. Estimator** | 1.00 (0.04) | -1.00 (0.09) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.02) | 0.67 (0.04) |
| $p_{11} = 0.7$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.14 (0.03) | -1.09 (0.03) | 1.00 (0.02) | -1.00 (0.02) | 0.31 (0.02) | 0.65 (0.02) |
| **Endog. Estimator** | 1.00 (0.05) | -1.00 (0.04) | 1.00 (0.02) | -1.00 (0.02) | 0.33 (0.03) | 0.67 (0.03) |
| $p_{11} = 0.9$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.05 (0.02) | -1.11 (0.04) | 1.00 (0.01) | -1.00 (0.02) | 0.32 (0.02) | 0.66 (0.03) |
| **Endog. Estimator** | 1.00 (0.02) | -1.00 (0.05) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.02) | 0.67 (0.03) |

Notes: Each cell contains the mean of the maximum likelihood point estimates and the mean of the standard errors of these estimates from the Monte Carlo experiment. Exog. estimator refers to the maximum likelihood estimator assuming the state process is exogenous, so that $\rho = 0$. Endog. estimator refers to the maximum likelihood estimator allowing the state process to be endogenous, so that $\rho \in (-1,1)$.

**Table 3**
Monte Carlo Results: Estimation
DGP 1, $\rho = 0.9$ (endogenous switching)

| ***T* = 200** | $\beta_{0,1} = 1.0$ | $\beta_{0,2} = -1.0$ | $\beta_{1,1} = 1.0$ | $\beta_{1,2} = -1.0$ | $\sigma_1 = 0.33$ | $\sigma_2 = 0.67$ |
|---|---|---|---|---|---|---|
| $p_{11} = 0.7$  $p_{22} = 0.7$ | | | | | | |
| **Exog. Estimator** | 1.21 (0.03) | -1.42 (0.05) | 1.00 (0.01) | -1.00 (0.03) | 0.25 (0.02) | 0.52 (0.04) |
| **Endog. Estimator** | 1.00 (0.04) | -1.00 (0.08) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.03) | 0.67 (0.06) |
| $p_{11} = 0.7$  $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.26 (0.05) | -1.17 (0.05) | 1.00 (0.02) | -1.00 (0.03) | 0.29 (0.03) | 0.60 (0.04) |
| **Endog. Estimator** | 1.00 (0.06) | -1.00 (0.06) | 1.00 (0.02) | -1.00 (0.02) | 0.33 (0.04) | 0.67 (0.04) |
| $p_{11} = 0.9$  $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.10 (0.03) | -1.20 (0.07) | 1.00 (0.02) | -1.00 (0.03) | 0.31 (0.02) | 0.63 (0.05) |
| **Endog. Estimator** | 0.99 (0.04) | -1.00 (0.07) | 1.00 (0.01) | -1.00 (0.03) | 0.33 (0.02) | 0.67 (0.05) |
| ***T* = 500** | $\beta_{0,1} = 1.0$ | $\beta_{0,2} = -1.0$ | $\beta_{1,1} = 1.0$ | $\beta_{1,2} = -1.0$ | $\sigma_1 = 0.33$ | $\sigma_2 = 0.67$ |
| $p_{11} = 0.7$  $p_{22} = 0.7$ | | | | | | |
| **Exog. Estimator** | 1.21 (0.02) | -1.42 (0.03) | 1.00 (0.01) | -1.00 (0.02) | 0.25 (0.01) | 0.52 (0.02) |
| **Endog. Estimator** | 1.00 (0.03) | -1.00 (0.05) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.02) | 0.67 (0.04) |
| $p_{11} = 0.7$  $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.26 (0.03) | -1.17 (0.03) | 1.00 (0.01) | -1.00 (0.02) | 0.29 (0.02) | 0.60 (0.02) |
| **Endog. Estimator** | 1.00 (0.04) | -1.00 (0.04) | 1.00 (0.01) | -1.00 (0.01) | 0.33 (0.02) | 0.67 (0.03) |
| $p_{11} = 0.9$  $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.10 (0.02) | -1.20 (0.04) | 1.00 (0.01) | -1.00 (0.02) | 0.31 (0.02) | 0.64 (0.03) |
| **Endog. Estimator** | 1.00 (0.02) | -1.00 (0.04) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.03) | 0.67 (0.03) |

Notes: Each cell contains the mean of the maximum likelihood point estimates and the mean of the standard errors of these estimates from the Monte Carlo experiment. Exog. estimator refers to the maximum likelihood estimator assuming the state process is exogenous, so that $\rho = 0$. Endog. estimator refers to the maximum likelihood estimator allowing the state process to be endogenous, so that $\rho \in (-1, 1)$.

**Table 4**
Monte Carlo Results: Estimation
DGP 2, $\rho = 0.5$ (endogenous switching)

| ***T* = 200** | $\beta_{0,1} = 1.0$ | $\beta_{0,2} = -1.0$ | $\beta_{1,1} = 1.0$ | $\beta_{1,2} = -1.0$ | $\sigma_1 = 0.33$ | $\sigma_2 = 0.67$ |
|---|---|---|---|---|---|---|
| $p_{11} = 0.7$ $p_{22} = 0.7$ | | | | | | |
| **Exog. Estimator** | 1.13 (0.03) | -1.26 (0.06) | 1.00 (0.02) | -1.00 (0.03) | 0.30 (0.02) | 0.61 (0.05) |
| **Endog. Estimator** | 1.00 (0.06) | -1.00 (0.13) | 1.00 (0.02) | -1.00 (0.03) | 0.33 (0.03) | 0.67 (0.07) |
| $p_{11} = 0.7$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.15 (0.05) | -1.10 (0.05) | 1.00 (0.02) | -1.00 (0.03) | 0.31 (0.03) | 0.64 (0.04) |
| **Endog. Estimator** | 1.00 (0.08) | -1.00 (0.07) | 1.00 (0.02) | -1.00 (0.03) | 0.32 (0.04) | 0.67 (0.04) |
| $p_{11} = 0.9$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.06 (0.03) | -1.11 (0.07) | 1.00 (0.02) | -1.00 (0.03) | 0.32 (0.02) | 0.65 (0.05) |
| **Endog. Estimator** | 1.00 (0.04) | -1.00 (0.08) | 1.00 (0.02) | -1.00 (0.03) | 0.33 (0.03) | 0.66 (0.05) |
| ***T* = 500** | $\beta_{0,1} = 1.0$ | $\beta_{0,2} = -1.0$ | $\beta_{1,1} = 1.0$ | $\beta_{1,2} = -1.0$ | $\sigma_1 = 0.33$ | $\sigma_2 = 0.67$ |
| $p_{11} = 0.7$ $p_{22} = 0.7$ | | | | | | |
| **Exog. Estimator** | 1.13 (0.02) | -1.26 (0.04) | 1.00 (0.01) | -1.00 (0.02) | 0.30 (0.02) | 0.61 (0.03) |
| **Endog. Estimator** | 1.00 (0.04) | -1.00 (0.08) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.02) | 0.67 (0.04) |
| $p_{11} = 0.7$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.15 (0.03) | -1.10 (0.03) | 1.00 (0.02) | -1.00 (0.02) | 0.31 (0.02) | 0.65 (0.02) |
| **Endog. Estimator** | 1.00 (0.05) | -1.00 (0.04) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.03) | 0.67 (0.03) |
| $p_{11} = 0.9$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.06 (0.02) | -1.11 (0.04) | 1.00 (0.01) | -1.00 (0.02) | 0.32 (0.02) | 0.66 (0.03) |
| **Endog. Estimator** | 1.00 (0.02) | -1.00 (0.05) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.02) | 0.67 (0.03) |

Notes: Each cell contains the mean of the maximum likelihood point estimates and the mean of the standard errors of these estimates from the Monte Carlo experiment. Exog. estimator refers to the maximum likelihood estimator assuming the state process is exogenous, so that $\rho = 0$. Endog. estimator refers to the maximum likelihood estimator allowing the state process to be endogenous, so that $\rho \in (-1,1)$.

**Table 5**
Monte Carlo Results: Estimation
DGP 2, $\rho = 0.9$ (endogenous switching)

| **T = 200** | $\beta_{0,1} = 1.0$ | $\beta_{0,2} = -1.0$ | $\beta_{1,1} = 1.0$ | $\beta_{1,2} = -1.0$ | $\sigma_1 = 0.33$ | $\sigma_2 = 0.67$ |
|---|---|---|---|---|---|---|
| $p_{11} = 0.7$ $p_{22} = 0.7$ | | | | | | |
| **Exog. Estimator** | 1.21 (0.03) | -1.42 (0.05) | 1.00 (0.01) | -1.00 (0.03) | 0.25 (0.02) | 0.52 (0.04) |
| **Endog. Estimator** | 1.01 (0.04) | -1.01 (0.08) | 1.00 (0.01) | -1.00 (0.02) | 0.32 (0.03) | 0.66 (0.06) |
| $p_{11} = 0.7$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.26 (0.04) | -1.17 (0.05) | 1.00 (0.02) | -1.00 (0.03) | 0.28 (0.03) | 0.59 (0.04) |
| **Endog. Estimator** | 1.00 (0.06) | -1.00 (0.06) | 1.00 (0.02) | -1.00 (0.02) | 0.33 (0.04) | 0.66 (0.04) |
| $p_{11} = 0.9$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.10 (0.03) | -1.20 (0.07) | 1.00 (0.02) | -1.00 (0.03) | 0.31 (0.02) | 0.63 (0.05) |
| **Endog. Estimator** | 1.00 (0.04) | -1.00 (0.07) | 1.00 (0.01) | -1.00 (0.03) | 0.33 (0.02) | 0.67 (0.05) |
| **T = 500** | $\beta_{0,1} = 1.0$ | $\beta_{0,2} = -1.0$ | $\beta_{1,1} = 1.0$ | $\beta_{1,2} = -1.0$ | $\sigma_1 = 0.33$ | $\sigma_2 = 0.67$ |
| $p_{11} = 0.7$ $p_{22} = 0.7$ | | | | | | |
| **Exog. Estimator** | 1.20 (0.02) | -1.43 (0.03) | 1.00 (0.01) | -1.00 (0.02) | 0.25 (0.01) | 0.52 (0.02) |
| **Endog. Estimator** | 1.01 (0.02) | -1.01 (0.05) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.02) | 0.66 (0.04) |
| $p_{11} = 0.7$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.25 (0.03) | -1.17 (0.03) | 1.00 (0.01) | -1.00 (0.02) | 0.29 (0.02) | 0.60 (0.02) |
| **Endog. Estimator** | 1.00 (0.04) | -1.00 (0.04) | 1.00 (0.01) | -1.00 (0.01) | 0.33 (0.03) | 0.67 (0.03) |
| $p_{11} = 0.9$ $p_{22} = 0.9$ | | | | | | |
| **Exog. Estimator** | 1.10 (0.02) | -1.20 (0.04) | 1.00 (0.01) | -1.00 (0.02) | 0.31 (0.02) | 0.63 (0.03) |
| **Endog. Estimator** | 1.00 (0.02) | -1.00 (0.04) | 1.00 (0.01) | -1.00 (0.02) | 0.33 (0.02) | 0.67 (0.03) |

Notes: Each cell contains the mean of the maximum likelihood point estimates and the mean of the standard errors of these estimates from the Monte Carlo experiment. Exog. estimator refers to the maximum likelihood estimator assuming the state process is exogenous, so that $\rho = 0$. Endog. estimator refers to the maximum likelihood estimator allowing the state process to be endogenous, so that $\rho \in (-1,1)$.

**Table 6**
Monte Carlo Results:
Empirical Size of 5% Nominal Size Test of $\rho = 0$

| **$T$ = 200** | | **Empirical Size: Wald Test** | **Empirical Size: LR Test** |
|---|---|---|---|
| $p_{11} = 0.7$ | $p_{22} = 0.7$ | 11.8% | 6.8% |
| $p_{11} = 0.7$ | $p_{22} = 0.9$ | 9.6% | 5.5% |
| $p_{11} = 0.9$ | $p_{22} = 0.9$ | 7.4% | 5.5% |
| **$T$ = 500** | | | |
| $p_{11} = 0.7$ | $p_{22} = 0.7$ | 6.4% | 4.6% |
| $p_{11} = 0.7$ | $p_{22} = 0.9$ | 7.3% | 6.7% |
| $p_{11} = 0.9$ | $p_{22} = 0.9$ | 5.2% | 4.9% |

**Table 7**

Monte Carlo Results: Empirical Size-Adjusted Power of 5% Nominal Size Test of $\rho = 0$

| **$T$ = 200** | | **Empirical Power: Wald Test** | **Empirical Power: LR Test** |
|---|---|---|---|
| DGP 1, $\rho = 0.5$ | | | |
| $p_{11} = 0.7$ | $p_{22} = 0.7$ | 39.9% | 46.3% |
| $p_{11} = 0.7$ | $p_{22} = 0.9$ | 65.4% | 67.0% |
| $p_{11} = 0.9$ | $p_{22} = 0.9$ | 86.3% | 85.8% |
| DGP 1, $\rho = 0.9$ | | | |
| $p_{11} = 0.7$ | $p_{22} = 0.7$ | 99.8% | 100.0% |
| $p_{11} = 0.7$ | $p_{22} = 0.9$ | 100.0% | 99.9% |
| $p_{11} = 0.9$ | $p_{22} = 0.9$ | 100.0% | 100.0% |
| DGP 2, $\rho = 0.5$ | | | |
| $p_{11} = 0.7$ | $p_{22} = 0.7$ | 65.9% | 68.1% |
| $p_{11} = 0.7$ | $p_{22} = 0.9$ | 67.0% | 69.2% |
| $p_{11} = 0.9$ | $p_{22} = 0.9$ | 79.4% | 79.7% |
| DGP 2, $\rho = 0.9$ | | | |
| $p_{11} = 0.7$ | $p_{22} = 0.7$ | 100.0% | 100.0% |
| $p_{11} = 0.7$ | $p_{22} = 0.9$ | 99.8% | 99.8% |
| $p_{11} = 0.9$ | $p_{22} = 0.9$ | 99.9% | 99.9% |

**Table 7 (cont.)**

Monte Carlo Results: Empirical Size-Adjusted Power of 5% Nominal Size Test for $\rho = 0$

| ***T* = 500** | | **Empirical Power: Wald Test** | **Empirical Power: LR Test** |
|---|---|---|---|
| DGP 1, $\rho = 0.5$ | | | |
| $p_{11} = 0.7$ | $p_{22} = 0.7$ | 87.9% | 88.4% |
| $p_{11} = 0.7$ | $p_{22} = 0.9$ | 96.4% | 96.6% |
| $p_{11} = 0.9$ | $p_{22} = 0.9$ | 99.9% | 99.9% |
| DGP 1, $\rho = 0.9$ | | | |
| $p_{11} = 0.7$ | $p_{22} = 0.7$ | 100.0% | 100.0% |
| $p_{11} = 0.7$ | $p_{22} = 0.9$ | 100.0% | 100.0% |
| $p_{11} = 0.9$ | $p_{22} = 0.9$ | 100.0% | 100.0% |
| DGP 2, $\rho = 0.5$ | | | |
| $p_{11} = 0.7$ | $p_{22} = 0.7$ | 96.5% | 96.9% |
| $p_{11} = 0.7$ | $p_{22} = 0.9$ | 99.0% | 99.0% |
| $p_{11} = 0.9$ | $p_{22} = 0.9$ | 99.9% | 99.9% |
| DGP 2, $\rho = 0.9$ | | | |
| $p_{11} = 0.7$ | $p_{22} = 0.7$ | 100% | 100% |
| $p_{11} = 0.7$ | $p_{22} = 0.9$ | 100% | 100% |
| $p_{11} = 0.9$ | $p_{22} = 0.9$ | 100% | 100% |

**Table 8**
Maximum Likelihood Estimates of the Turner, Startz, and Nelson (1989) Volatility-Feedback Model

| Parameter | Ignoring Endogeneity | Accounting for Endogeneity |
|---|---|---|
| $\theta_1$ | 0.31<br>(0.10) | 0.36<br>(0.10) |
| $\theta_2$ | -1.55<br>(0.45) | -1.07<br>(0.45) |
| $\sigma_1$ | 0.40<br>(0.02) | 0.40<br>(0.02) |
| $\sigma_2$ | 0.75<br>(0.07) | 0.74<br>(0.07) |
| $a_1$ | 2.05<br>(0.20) | 2.05<br>(0.17) |
| $a_2$ | -1.09<br>(0.21) | -1.16<br>(0.22) |
| $\rho$ | --- | -0.40<br>(0.18) |
| Log Likelihood | -372.41 | -370.89 |

Notes: Maximum likelihood estimates computed in GAUSS 6.0 using the OPTMUM numerical optimization package. Standard errors, reported in parentheses, were based on second derivatives of the log-likelihood function.