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

June 2003

Abstract

Following Hamilton (1989), estimation of Markov regime-switching regressions nearly always relies on the assumption that the latent state variable controlling the regime change is exogenous. We relax this assumption and develop two techniques to estimate Markov-switching models with endogenous switching. The first extends the endogenous switching models in Maddala and Nelson (1975) to the Hamilton (1989) Markov regime-switching regression. The second is based on the interpretation of the endogenous switching regression as a regression model with endogenous unobserved dummy variables. In this case the regression can be estimated using instrumental variables techniques. For both techniques, identification is achieved when the transition probabilities of the regime-switching process are influenced by observed exogenous variables. However, even in the fixed transition probability case, identification is achieved if the state process is serially dependent and lagged values of the state process are exogenous. This is true even though the lagged state is unobserved. The bias correction techniques also admit straightforward tests for endogeneity. Monte Carlo experiments confirm that the proposed bias correction techniques perform quite well in practice. We apply the procedure to the volatility feedback model of equity returns given in Turner, Startz and Nelson (1989).

Keywords: Endogeneity, Regime-Switching, Instrumental Variables
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). We thank Michael Dueker 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. Corresponding author: Jeremy Piger.

There is mounting evidence that empirical models of many economic time series, particularly macroeconomic and financial series, are characterized by parameter instability. This has sparked an explosion of interest in time-varying parameter models. One notable set of models are switching regressions with unknown sample separation, in which parameters move discretely between a fixed number of regimes, with the switching controlled by an unobserved state variable. Switching regressions have a rich history in econometrics, dating back 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 the 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 spawned by Hamilton (1989) has typically assumed that the regime shifts are exogenous of all realizations of the regression error. However, the earlier literature on switching regressions, such as Maddala and Nelson (1975), was often concerned with endogenous switching, as the primary applications of switching regression models in this literature were in limited dependent variable contexts such as self-selection and market disequilibrium problems.

In this paper we work with switching regressions of the type considered by Hamilton (1989) and various extensions, but relax the exogenous switching assumption. We show that failure of the exogeneity assumption leads to significant bias in the coefficients of a regime-switching regression when estimation methods based on the exogeneity assumption are used. We then consider two estimation techniques for correcting this bias. In the first, we extend the endogenous switching models in Maddala and Nelson (1975) to the Hamilton (1989) regime switching regression, in which the switching process is serially dependent, specifically Markov-switching. This approach relies on a joint normality assumption for the relationship between the regression error terms and the innovations to the state realization equation. If one is unwilling to make this assumption, we develop an alternative technique, based on the interpretation of the endogenous switching regression as a regression model with endogenous unobserved dummy variables. In this case the regression can be estimated using instrumental variables techniques. Both bias correction techniques also admit a straightforward test for endogeneity.

For both of these estimation techniques, we show that for serially dependent state processes, such as a Markov-switching state process, the lagged state can provide information necessary for identification, providing it is uncorrelated with the current regression error. This is true even though the lagged state is unobserved. Additional information is obtained when the transition probabilities of the switching process are influenced by exogenous variables, as in the so called "time-varying transition probability" case. We use Monte Carlo experiments to confirm the efficacy of the proposed bias-correction procedures and find that their performance is quite good for empirically relevant data generating processes.

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 would be natural to assume that the state is endogenous. As an example, in many models the estimated state variable has a strong business cycle correlation, often corresponding with recessions. 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 recessions. 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 the next section we lay out a canonical Markov-switching regression and document the biases that arise when the latent state variable is correlated with the regression error. Section 3 develops the two bias correction strategies, discusses identification, and presents tests for endogeneity. Section 4 gives the results of a Monte Carlo experiment documenting the performance of the proposed bias correction procedures. 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).

2. Endogenous Regime-Switching and Estimation Bias

The model we consider is a linear regression with regime-switching coefficients and residual variance. Assume that the dependent variable, y_t, is generated by one of two true regression equations:[1]

$$\begin{aligned} y_t &= \alpha_0 + x_t'\beta_0 + \varepsilon_{0t} \qquad \text{iff } S_t = 0 \\ y_t &= \alpha_1 + x_t'\beta_1 + \varepsilon_{1t} \qquad \text{iff } S_t = 1 \end{aligned} \tag{1}$$

where $S_t = \{0,1\}$ is a state variable that indicates which of the two regression equations generates the observation y_t and x_t is a k x 1 vector of exogenous or predetermined explanatory variables measured at time t, which may include lagged values of y_t. The residuals, ε_{it}, have zero mean and variances σ_i^2. We will focus on Gaussian maximum likelihood estimation throughout the paper and thus assume that ε_{it} is Gaussian.

We are interested in the case of unobserved S_t and thus require a probability law governing S_t for estimation purposes. Here we consider a regime-switching process for S_t that encompasses several popular specifications in the literature. In particular we assume that the probability that $S_t = i$ depends on S_{t-1} and on a vector of exogenous or predetermined variables z_t, where z_t may include elements of x_t. Formally:

$$
\begin{aligned}
&P(S_t = 1 \mid S_{t-1} = 1, \ldots\ldots, S_{t-j} = i, z_t) = P(S_t = 1 \mid S_{t-1} = 1, z_t) = p(z_t) \\
&P(S_t = 0 \mid S_{t-1} = 1, z_t) = 1 - p(z_t) \\
&P(S_t = 0 \mid S_{t-1} = 0, \ldots\ldots, S_{t-j} = i, z_t) = P(S_t = 0 \mid S_{t-1} = 0, z_t) = q(z_t) \\
&P(S_t = 1 \mid S_{t-1} = 0, z_t) = 1 - q(z_t)
\end{aligned}
$$

We parameterize $p(z_t)$ and $q(z_t)$ using the following probit specification:

$$
\begin{aligned}
&S_t = \begin{Bmatrix} 0 & \text{if} & S_t^* < 0 \\ 1 & \text{if} & S_t^* \geq 0 \end{Bmatrix} \\
&S_t^* = a_0 + a_1 S_{t-1} + a_2 z_t + a_3 z_t S_{t-1} + \eta_t \qquad (2) \\
&\eta_t \sim N(0,1)
\end{aligned}
$$

So that:

$$
\begin{aligned}
&p(z_t) = P(\eta_t > (-(a_0 + a_1) - (a_2 + a_3) z_t)) = 1 - \Phi(-(a_0 + a_1) - (a_2 + a_3) z_t) \\
&q(z_t) = P(\eta_t \leq (-a_0 - a_2 z_t) = \Phi(-a_0 - a_2 z_t)
\end{aligned}
$$

where Φ is the standard normal cumulative distribution function.

[1] We focus here on a two-regime model, however, all of the results are easily generalized to the n regime case.

Several special cases of (2) are worth mentioning. The unrestricted model represents the time-varying transition probability Markov-switching model (TVP-MS) of Goldfeld and Quandt (1973), Diebold, Lee and Weinbach (1994) and Filardo (1994). When $a_2 = a_3 = 0$, we have the fixed transition probability Markov-switching model (FTP-MS) of Goldfeld and Quandt (1973) and Hamilton (1989), that is $p(z_t) = p$, $q(z_t) = q$. If $a_1 = a_2 = a_3 = 0$ we have the fixed transition probability independent switching model (FTP-IS) of Quandt (1972), so that $q = 1 - p$. Finally, if $a_1 = a_3 = 0$, we have the time-varying transition probability independent switching model (TVP-IS) of Goldfeld and Quandt (1972), $q(z_t) = 1 - p(z_t)$.

The three residual terms in the model, $\varepsilon_{0t}, \varepsilon_{1t}$ and η_t have the following covariance structure:

$$\operatorname{cov}(\varepsilon_{0t}, \varepsilon_{1t}, \eta_t) = \Sigma = \begin{bmatrix} \sigma_0^2 & \sigma_{\varepsilon_0,\varepsilon_1} & \rho\sigma_0 \\ \sigma_{\varepsilon_0,\varepsilon_1} & \sigma_1^2 & \rho\sigma_1 \\ \rho\sigma_0 & \rho\sigma_1 & 1 \end{bmatrix} \tag{3}$$

Maximum likelihood estimation of (2) via the recursive filter given in Hamilton (1989) or the expectation maximization (EM) algorithm given in Hamilton (1990) will yield consistent parameter estimates under the assumption that S_t is uncorrelated with ε_{0t} and ε_{1t}, that is $\rho = 0$.[2] Relaxation of this assumption will lead to biased estimation of the parameters α_i and β_i. To see this, note that equation (1) can be rewritten as follows:

$$\begin{aligned} y_t &= \alpha_0 + E(\varepsilon_{0t} \mid \mathrm{S_t} = 0) + x_t'\beta_0 + u_{0t} \qquad \text{iff } S_t = 0 \\ y_t &= \alpha_1 + E(\varepsilon_{1t} \mid \mathrm{S_t} = 1) + x_t'\beta_1 + u_{1t} \qquad \text{iff } S_t = 1 \end{aligned} \tag{4}$$

[2] Note that $\sigma_{\varepsilon_0,\varepsilon_1}$ does not appear in the likelihood function and is thus not estimable. See Maddala (1984) for details.

where $u_{it} = \varepsilon_{it} - E(\varepsilon_{it} \mid S_t = i)$. From equation (4), we can see that S_t introduces predictability into ε_{it} that varies perfectly with the regime switching intercept. Thus, there is no information to separately identify α_i from $E(\varepsilon_{it} \mid S_t = i)$. In general, maximum likelihood estimation will yield biased estimates of both α_i and β_i. However, in the special case where x_t and S_t are independent, the bias will be contained to α_i only. In the next section we present techniques to correct the bias introduced by the endogenous state variable.

3. Bias Correction in the Endogenous Switching Model

3.1 Bias Correction Assuming Joint Normality

Suppose that the joint distribution between $\varepsilon_{0t}, \varepsilon_{1t}$ and η_t is multivariate Gaussian:

$$\begin{bmatrix} \varepsilon_{0t} \\ \varepsilon_{1t} \\ \eta_t \end{bmatrix} \sim N(0, \Sigma), \quad \Sigma = \begin{bmatrix} \sigma_0^2 & \sigma_{\varepsilon_0, \varepsilon_1} & \rho\sigma_0 \\ \sigma_{\varepsilon_0, \varepsilon_1} & \sigma_1^2 & \rho\sigma_1 \\ \rho\sigma_0 & \rho\sigma_1 & 1 \end{bmatrix} \tag{5}$$

In this case, the model in (1), (2) and (5) is closely related to the switching regression with endogenous switching considered by Maddala and Nelson (1975). The main addition we consider here is to allow the unobserved state process to be serially dependent, specifically Markov-switching, as in Goldfeld and Quandt (1973) and Hamilton (1989). In the following, we outline a maximum likelihood estimation procedure for this Markov-switching regression with endogenous switching.

Using the joint normal assumption given in (5), one can derive explicit characterizations of the bias terms from equation (4), $E(\varepsilon_{it} \mid S_t = i)$. In particular, note from (5) that:

$$
\begin{aligned}
E(\varepsilon_{0t} \mid S_t = 0) &= E(\varepsilon_{0t} \mid \eta_t < -a_0 - a_1 S_{t-1} - a_2 z_t - a_3 z_t S_{t-1}) \\
&= \rho\sigma_0 \frac{-\phi(-a_0 - a_1 S_{t-1} - a_2 z_t - a_3 z_t S_{t-1})}{\Phi(-a_0 - a_1 S_{t-1} - a_2 z_t - a_3 z_t S_{t-1})} = \rho\sigma_0 M_0 \\
E(\varepsilon_{1t} \mid S_t = 1) &= E(\varepsilon_{1t} \mid \eta_t \geq -a_0 - a_1 S_{t-1} - a_2 z_t - a_3 z_t S_{t-1}) \\
&= \rho\sigma_1 \frac{\phi(-a_0 - a_1 S_{t-1} - a_2 z_t - a_3 z_t S_{t-1})}{(1 - \Phi(-a_0 - a_1 S_{t-1} - a_2 z_t - a_3 z_t S_{t-1}))} = \rho\sigma_1 M_1
\end{aligned}
\tag{6}
$$

where ϕ is the standard normal density function. Using (6), we can correct for bias introduced by the endogenous state variable by rewriting (4) as:

$$
\begin{aligned}
y_t &= \alpha_0 + \rho\sigma_0 M_0 + x_t' \beta_0 + u_{0t} \qquad \textit{iff } S_t = 0 \\
y_t &= \alpha_1 + \rho\sigma_1 M_1 + x_t' \beta_1 + u_{1t} \qquad \textit{iff } S_t = 1
\end{aligned}
\tag{7}
$$

The parameters of (7) are only identified if M_0 and M_1 vary, otherwise there is no information to identify α_i from $\rho\sigma_i M_i$. From (6), we can see this is achieved in two ways. First, if the state process depends on exogenous or predetermined observable variables z_t, as in the TVP-MS and TVP-IS models, the model is identified. Second, even when $a_2 = a_3 = 0$, as in the FTP-MS model, the model is still identified so long as $\alpha_1 \neq 0$. That is, if S_t is serially correlated and S_{t-1} is exogenous, S_{t-1} serves to identify the model.

Assuming identification, consistent estimates of the parameters of (1) are achieved by maximum likelihood estimation of (7), which can be performed computationally using the recursive filter given in Hamilton (1989) or the expectation maximization (EM) algorithm discussed in Hamilton (1990). As these techniques are well known, we will not repeat them in detail here. Note though that computation of the unconditional likelihood requires that the likelihood for (7) be computed conditional on each possible combination of S_t and S_{t-1}. This requires that we characterize the variances of the forecast errors in, u_{0t} and u_{1t}. These are given by:

$$\begin{aligned} &\operatorname{var}(u_{0t} \mid S_t = 0, M_0, x_t) \\ &\qquad = \rho^2 \sigma_0^2 (1 - M_0 (M_0 + a_0 + a_1 S_{t-1} + a_2 z_t + a_3 z_t S_{t-1})) + \sigma_0^2 (1 - \rho^2) \\ &\operatorname{var}(u_{1t} \mid S_t = 1, M_1, x_t) \\ &\qquad = \rho^2 \sigma_1^2 (1 - M_1 (M_1 + a_0 + a_1 S_{t-1} + a_2 z_t + a_3 z_t S_{t-1})) + \sigma_1^2 (1 - \rho^2) \end{aligned} \tag{8}$$

3.2 Estimation Based on Instrumental Variables

The bias correction technique presented in Section 3.1 relied on a specific distributional assumption describing the relationship between ε_{it} and η_t. In this section we present a bias correction strategy, based on instrumental variables estimation, which makes no such distributional assumption.

To motivate this approach, consider the alternative representation of the switching regression model in (1):

$$\begin{aligned} &y_t = \alpha_0 (1 - S_t) + \alpha_1 S_t + x_t' \beta_0 (1 - S_t) + x_t' \beta_1 S_t + \varepsilon_{0t} (1 - S_t) + \varepsilon_{1t} S_t \\ &\varepsilon_{it} \sim N(0, \sigma_i^2) \end{aligned} \tag{9}$$

where S_t is defined in (2). Equation (9) is a regression equation involving an unobserved dummy variable, S_t. If $\rho \neq 0$ in (3), this dummy variable is endogenous, and instrumental variables estimation seems a natural way to proceed. The role of an instrument is to separate an endogenous variable into endogenous and exogenous components. The standard two-stage least squares approach proceeds by replacing the endogenous variables in the regression, in this case S_t and $1-S_t$, with their exogenous components obtained using the instrument. An alternative procedure is to add the endogenous components as additional explanatory variables to the model, thereby controlling for all endogenous variation in the endogenous variables of interest. Here we follow this second approach.

To begin, consider a useful alternative characterization of S_t as an autoregressive process of order one:

$$\begin{aligned} S_t &= c + \lambda S_{t-1} + v_t \\ c &= 1 - q(z_t) \\ \lambda &= -1 + p(z_t) + q(z_t) \end{aligned} \tag{10}$$

where conditional on $S_{t-1} = 0$:

$v_t = -(1-q(z_t))$ with probability $q(z_t)$

$v_t = q(z_t)$ with probability $1-q(z_t)$

and conditional on $S_{t-1} = 1$:

$v_t = (1-p(z_t))$ with probability $p(z_t)$

$v_t = -p(z_t)$ with probability $1-p(z_t)$

As discussed in Hamilton (1989), the residual v_t is mean zero and uncorrelated with S_{t-1}.

Assume that while S_t may be endogenous, $S_{t-j}, j > 0$ is uncorrelated with ε_{0t} and ε_{1t}. Again, z_t is assumed to be predetermined or exogenous. Under these assumptions, the exogenous component of S_t can be constructed as $c + \lambda S_{t-1}$ while the endogenous component can be constructed as $v_t(S_t, S_{t-1}, z_t)$. To correct for bias, the regression in (9) is reformulated as:

$$y_t = \alpha_0(1-S_t) + \alpha_1 S_t + x_t^{'}\beta_0(1-S_t) + x_t^{'}\beta_1 S_t + \kappa_0 v_t(S_t, S_{t-1}, z_t)(1-S_t) + \kappa_1 v_t(S_t, S_{t-1}, z_t)S_t + \omega_{0t}(1-S_t) + \omega_{1t}S_t \quad (11)$$

where $\omega_{it} = \varepsilon_{it} - \kappa_i v_t(S_t = i, S_{t-1}, z_t)$ and is Gaussian as ε_t is Gaussian and the regression is conditioned on v_t.

The term $v_t(S_t, S_{t-1}, z_t)$ serves as a control variable for the endogenous component of S_t in (11), allowing the coefficients α_0 and α_1 to capture the effect of the exogenous component of S_t. Identification requires that we can identify $S_t - v_t(S_t, S_{t-1}, z_t)$ as the exogenous component of S_t. This is achieved in two ways. First, if an exogenous or predetermined z_t exists, as in the TVP-MS and TVP-IS models, the model is identified. Second, even when $a_2 = a_3 = 0$, as in the FTP-MS model, the model is still identified so long as $\alpha_1 \neq 0$. That is, if S_t is serially correlated and S_{t-1} is exogenous, S_{t-1} serves as an instrument for S_t. This is true even though S_{t-1} is unobserved.

Again, maximum likelihood estimation of (11) requires that we specify the variance of the forecast errors of (11), ω_{0t} and ω_{1t}, denoted $\text{var}(\omega_{it} \mid S_t = i, v_t, x_t)$. Recalling that $\omega_{it} = \varepsilon_{it} - \kappa_i v_t(S_t = i, S_{t-1}, z_t)$, we then have:

$$\text{var}(\omega_{it} \mid S_t = i, v_t, x_t) = \text{var}(\varepsilon_{it} \mid S_t = i, v_t, x_t) \quad (12)$$

as $\kappa_i v_t(S_t = i, S_{t-1}, z_t)$ is constant given the conditioning variables. Note that $\text{var}(\varepsilon_{it} \mid S_t = i, v_t, x_t) \neq \text{var}(\varepsilon_{it} \mid S_t) = \sigma_i^2$, as v_t provides information regarding η_t beyond that contained in S_t. Instead, $\text{var}(\varepsilon_{it} \mid S_t = i, v_t, x_t)$ depends on the unknown joint distribution between ε_{it} and η_t.

In the FTP-MS case, in which $a_2 = a_3 = 0$, $\text{var}(\varepsilon_{it} \mid S_t = i, v_t, x_t)$ takes on one of only four different values, corresponding to the four different combinations of $S_t = i, S_{t-1} = j,\ i, j = 0,1$. In this case, maximum likelihood estimation can be implemented through the introduction of four new parameters. However, in those cases where the transition probabilities of the switching process are time varying, and thus depend on z_t, $\text{var}(\varepsilon_{it} \mid S_t = i, v_t, x_t)$ will be an unknown function of z_t and estimation of (11) via maximum likelihood will require some approximation of this unknown function. Thus, while consistent estimates of (11) can be obtained in the time-varying transition probability case, standard errors calculations involve an approximation. Once the forecast error variances have been specified, maximum likelihood estimation of (11) can again be performed computationally using the recursive filter given in Hamilton (1989) or the expectation maximization (EM) algorithm discussed in Hamilton (1990).

3.3 Testing for an Endogenous State Variable

Estimation of the models in (7) and (11) provide a straightforward test for endogeneity in S_t. Given valid instruments, a test of the null hypothesis of no endogeneity can be performed as a test of the null hypothesis that $\rho = 0$ in (7) or $\kappa_0 = 0$ and $\kappa_1 = 0$ in (11). Note that this can be interpreted as an application of the Wu (1973) test for endogenous regressors, where in this case the regressor is an unobserved latent state variable.

4. Monte Carlo Assessment of the Proposed Bias Correction Procedure

To evaluate the proposed bias correction procedures we perform a series of Monte Carlo experiments. In each Monte Carlo experiment, 1000 data sets are generated from the model given in equations (1)-(3). To model correlation between S_t and ε_t we assume that η_t in equation (2) and ε_t have the joint normal distribution given in (5). Note that this assumption is the same as that used to develop the bias correction procedure based on (7) described in Section 3.1, meaning this bias correction procedure will be optimal.

The Monte Carlo experiments are calibrated as follows: Each experiment is performed for three values of the correlation parameter $\rho = 0.3$, 0.6 and 0.9 and two sample sizes, $T = 200$ and 500. The vector z_t is assumed to be scalar and is generated as an independent standard normal random variable. The vector x_t is also assumed to be scalar and is generated to be exogenous, but correlated with S_t. We achieve this by generating x_t as a standard normal random variable added to the exogenous part of S_t, which from equation (2) is equal to $a_0 + a_1 S_{t-1} + a_2 z_t + a_3 z_t S_{t-1}$. This yields correlations between x_t and S_t ranging from 0.15 to 0.4, depending on the parameterization of equation (3) used in the Monte Carlo experiment. Equation (2) is calibrated with $\alpha_0 = 2, \alpha_1 = -2, \beta_0 = 1$, $\beta_1 = -1$, $\sigma_0 = 0.5$ and $\sigma_1 = 1$. We consider two different sets of parameters for the state process (2). These are:

DGP1: TVP-MS: $a_0 = -0.5, a_1 = 1, a_2 = 0.5, a_3 = 1$

DGP2: FTP-MS: $a_0 = -0.5, a_1 = 1, a_2 = 0, a_3 = 0$

Based on this calibration we perform three sets of Monte Carlo experiments. In the first, data is generated from DGP1 and the model is estimated without bias correction. This experiment will verify the bias in the regression parameters discussed in section 2. In the second and third experiments, data is generated from DGP1 and DGP2, and both of the bias correction procedures discussed in section 3, that based on equation (7) and that based on equation (11), are implemented. In each case, the likelihood function for the appropriate model is constructed using the filter given in Hamilton (1989) and maximized computationally. As was discussed in section 3.2, in the time-varying transition probability case, estimation of equation (11) requires an approximation for the forecast error variance to make maximum likelihood estimation feasible. Here we use the approximation that the forecast error variances are independent of z_t. The results of the Monte Carlo experiments are contained in Tables 1-4. Each table shows the mean and standard deviation of the 1000 maximum likelihood point estimates of the parameters of (1) and (2), which will henceforth be referred to as the "mean" and "standard deviation."

Table 1 presents the results when data is generated according to DGP1, and no bias correction procedure is employed. The table demonstrates that there is significant bias introduced into the regime-switching coefficients α_i and β_i by the endogenous S_t. For example, when $T = 200$ and $\rho = 0.6$, the mean estimates of α_0 and α_1 are over 3 standard deviations from the true values. The bias is somewhat smaller in the estimates of β_0 and β_1, the mean estimates of which are around one standard deviation from the true values. As would be expected, the bias in both sets of parameters is larger when $\rho = 0.9$ and smaller when $\rho = 0.3$. The bias is not mitigated as the sample size grows – when $T = 500$, the mean estimates of α_i and β_i are nearly the same as for the $T = 200$ case.

In Table 2, we again generate data according to DGP1, but now estimate the model employing the bias correction procedures from equation (7) (Table 2a) and equation (11) (Table 2b). DGP1 supplies us with information to identify the model parameters both from the time varying transition probabilities, z_t, and the serial dependence of the state, S_{t-1}. The Monte Carlo results suggest that both bias correction procedures are very effective. In particular, the mean estimates of α_i and β_i are now close to the true values for both sample sizes considered.

Table 3 holds the results when the bias correction technique is applied to DGP2, in which there are no observed variables with which to identify the model parameters. Instead, only the unobserved lagged state, S_{t-1}, is available for identification. Table 3 shows that for both bias correction procedures, the estimates of α_i and β_i are very close to the true values. The standard deviations of the estimates in Tables 3 are in general somewhat larger than those in Table 2. This is not surprising, as in Table 2 (based on DGP1) there is more information to identify the parameters, in the form of z_t, than in Tables 3.

Overall, the Monte Carlo experiments confirm that 1) endogenous regime switching can introduce significant bias into the maximum likelihood estimates of a Markov regime-switching model and 2) the bias correction procedures proposed in Section 3 are quite effective in eliminating this bias. In the next section we turn to an empirical application of the bias correction procedures.

5. An 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, the expected return of the market portfolio above the risk-free rate, 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.

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 also 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 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).

One approach to modeling time-varying volatility is through the use of a regime-switching model. For example, Turner, Startz and Nelson (1989), henceforth TSN, model the excess of equity returns over the risk free rate, r_t, as arising from a normal distribution with time dependent expectation μ_t and variance:

$$\begin{aligned} &\sigma_{r,t}^2 = \sigma_{r,0}^2(1-S_t) + \sigma_{r,1}^2 S_t, \quad \sigma_{r,1}^2 > \sigma_{r,0}^2 \\ &S_t = (0,1) \end{aligned} \tag{12}$$

S_t follows a FTP-MS process, that is the process in (3) with $a_2 = a_3 = 0$. The transition probabilities are then given by:

$$\begin{aligned} &P(S_t = 0 \mid S_{t-1} = 0) = q = P(\eta_t \le (-a_0) = \Phi(-a_0) \\ &P(S_t = 1 \mid S_{t-1} = 1) = p = P(\eta_t > (-(a_0 + a_1))) = 1 - \Phi(-(a_0 + a_1)) \end{aligned}$$

In this model, excess returns are drawn from either a high variance distribution ($S_t = 1$) or a low variance distribution ($S_t = 0$). Also, since the state variable S_t is persistent, deviations of the return variance above its unconditional mean are predictable. We would thus expect a risk premium that varies with the predictable component of the variance. This can be modeled as follows:

$$\mu_t = \theta_0 + \theta_1 P(S_t \mid \Psi_{t-1}) \tag{13}$$

Here, $P(S_t \mid \Psi_{t-1})$ measures the probability of the high variance state at time *t* given an information set dated time *t*-1. TSN assume that Ψ_{t-1} includes all past returns, $r_{t-1}......r_1$ and $P(S_t \mid \Psi_{t-1})$ is thus equivalent to the filtered probability of the state. In (13), θ_0 is equal to the risk premium when there is no probability placed on the occurrence of the high-variance state. Assuming that θ_1 is positive, the risk premium is increasing in the probability of the high-variance state.

The model above does not incorporate volatility feedback in determining the actual excess return. TSN incorporate volatility feedback by modeling r_t as follows:

$$\begin{aligned} r_t &= \mu_t + \theta_2 (P(S_t \mid \Psi_t^*) - P(S_t \mid \Psi_{t-1})) + \zeta_t \\ \zeta_t &\sim N(0, \sigma_{\varsigma,t}^2) \\ \sigma_{\varsigma,t}^2 &= \sigma_{\varsigma,0}^2 (1 - S_t) + \sigma_{\varsigma,1}^2 S_t \end{aligned} \tag{14}$$

The model in (14) can be motivated as follows. At the beginning of period *t*, the risk premium μ_t is determined based on the expectation of the high-variance state in period *t* based on *t*-1 information. During period t additional information regarding the incidence of the high-variance state is observed. By the end of period t, this information can be collected in the information set Ψ_t^*. When $P(S_t \mid \Psi_t^*) \neq P(S_t \mid \Psi_{t-1})$, information revealed during the period has surprised agents regarding the occurrence of the high-variance state. If $\theta_2 < 0$, surprises that reveal higher probability of the high-variance state are viewed negatively by investors, and thus reduce the contemporaneous return.

One difficulty in estimating the above model of volatility feedback is that there exists a discrepancy between the investors' information set and the econometrician's data set. In particular, while Ψ_{t-1} may be summarized by all data up to t-1, that is $\Psi_t = \{r_{t-1}, r_{t-2}, ...\}$, the information set Ψ_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. This is a particular problem when the period is relatively long, as is the case for the monthly data set used by TSN.

To handle this estimation difficulty, TSN use the actual state, S_t, as a proxy for $P(S_t \mid \Psi_t^*)$. That is, they estimate:

$$
\begin{aligned}
& r_t = \mu_t + \theta_2 (S_t - P(S_t \mid \Psi_{t-1})) + \zeta_t^* \\
& \zeta_t^* \sim N(0, \sigma^2_{\varsigma^*,t}) \\
& \sigma^2_{\varsigma^*,t} = \sigma^2_{\varsigma^*,0}(1 - S_t) + \sigma^2_{\varsigma^*,1} S_t
\end{aligned}
\tag{15}
$$

This approach introduces measurement error into the explanatory variables of the estimated equation, as S_t is a noisy measure of $P(S_t \mid \Psi_t^*)$. In particular, note that the residual in (15) is

$\zeta_t^* = \zeta_t + \theta_2(P(S_t \mid \Psi_t^*) - S_t)$ so that S_t will be correlated with the residual ζ_t^* in equation (15). The results from section 2 suggest that maximum likelihood estimates of the parameters of equation (15) will then be biased.

The bias correction techniques introduced in Section 3 can be used to correct for this bias. We show this here using the technique of Section 3.2, in which a joint normality assumption is employed. As there are no time-varying transition probabilities in the TSN model, that is $z_t = 0$, there is a single instrument in the model, S_{t-1}. Thus identification requires that S_t be Markov-switching, rather than independent switching. Formally, equation (7) for the model in (15) becomes:

$$\begin{aligned} r_t &= \mu_t + \theta_2(S_t - P(S_t \mid \psi_{t-1})) + \rho\sigma_{\varsigma^*,0}M_0(1-S_t) + \rho\sigma_{\varsigma^*,1}M_1S_t + \omega_t \\ \omega_t &= \varsigma_t^* - \rho\sigma_{\varsigma^*,0}M_0(1-S_t) - \rho\sigma_{\varsigma^*,1}M_1S_t \end{aligned} \tag{16}$$

We estimate the model in (16) using monthly returns for a value-weighted portfolio of all NYSE-listed stocks in excess of the one-month Treasury Bill rate, our measure of the risk free rate, over the sample period 1952-2000. This is the same data as used in Kim, Morley and Nelson (2002). The first panel of Table 4 shows these estimates when no bias correction is employed, that is $\rho = 0$. These estimates, which are similar to those in TSN, are consistent with both a positive relationship between the risk premium and expected future volatility ($\theta_0 > 0$ and $\theta_1 > 0$) and a substantial volatility feedback effect ($\theta_2 << 0$). The estimates are also consistent with a dominant volatility feedback effect, that is θ_1 is very small relevant to θ_2. The second panel shows the estimates when bias correction is employed, so that ρ is estimated. There is substantial evidence of endogeneity, as ρ is estimated to be large in absolute value and statistically significant. Although the parameter estimates are very imprecise, correcting for this endogeneity does yield substantial changes in the point estimates of the model parameters. First, θ_1 is estimated to be nearly ten times the estimate for the model with no bias correction, suggesting a risk premium that is much more responsive to forecasted future volatility. Second, θ_2 is estimated to be half the estimate for the model with no bias correction, suggesting a strong volatility feedback effect, but less so than for the model with no bias correction.

6. Conclusion

We have explored the implications of relaxing the assumption of exogeneity for the latent state variable in a Markov regime-switching regression. We show that endogeneity of the state variable can lead to significant estimation bias and have considered two techniques to correct for this estimation bias. The first extends the endogenous switching models in Maddala and Nelson (1975) to the Hamilton (1989) regime switching regression, in which the switching process is serially dependent. This approach relies on a joint normality assumption for the relationship between the regression error terms and the innovations to the state realization equation. The second technique is based on the interpretation of the endogenous switching regression as a regression model with endogenous unobserved dummy variables. In this case the regression can be estimated using instrumental variables techniques.

For both techniques, identification can be achieved using information from external, exogenous variables that affect the transition probabilities of the regime-switching process. However, even in the fixed transition probability case, we have shown that identification is achieved if the state process is serially dependent, and lagged values of the state are exogenous. This is true even though the lagged state is unobserved. Monte Carlo experiments confirm that the proposed bias-correction procedure is quite good for empirically relevant data generating processes. We apply the technique to the volatility feedback model of equity returns given in Turner, Startz and Nelson (1989).

References

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

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.

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., 1990, Analysis of Time Series Subject to Changes in Regime, *Journal of Econometrics,* 45, 39-70.

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.

Maddala, G.S., 1984, Disequilibrium, Self-Selection and Switching Models, in *Handbook of Econometrics,* eds. Z. Griliches and M.D. Intrilligator, North Holland, Amsterdam.

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 Switchign 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.

Wu, D., 1973, Alternative Tests of Independence Between Stochastic Regressors and Disturbances, *Econometrica,* 41, 733-750.

Table 1
Monte Carlo 1: Maximum Likelihood Estimation of Time-Varying Transition Probability Markov-Switching Regression Model (TVP-MS) – No Bias Correction

	$\rho = 0.3$	$\rho = 0.6$	$\rho = 0.9$
Parameter	**Parameter Estimates: Mean / Std. Dev.**	**Parameter Estimates: Mean / Std. Dev.**	**Parameter Estimates: Mean / Std. Dev.**
T=200			
$\alpha_0 = 2$	1.91 / 0.05	1.83 / 0.05	1.74 / 0.05
$\alpha_1 = -2$	-1.77 / 0.13	-1.55 / 0.12	-1.34 / 0.10
$\beta_0 = 1$	0.97 / 0.04	0.95 / 0.04	0.93 / 0.04
$\beta_1 = -1$	-1.06 / 0.07	-1.12 / 0.07	-1.19 / 0.06
$a_0 = -0.5$	-0.51 / 0.16	-0.52 / 0.16	-0.53 / 0.16
$a_1 = 1$	1.03 / 0.28	1.06 / 0.28	1.07 / 0.29
$a_2 = 0.5$	0.54 / 0.19	0.56 / 0.18	0.56 / 0.19
$a_3 = 1$	1.13 / 0.50	1.13 / 0.53	1.11 / 0.51
T=500			
$\alpha_0 = 2$	1.91 / 0.04	1.82 / 0.03	1.74 / 0.03
$\alpha_1 = -2$	-1.78 / 0.08	-1.56 / 0.07	-1.34 / 0.06
$\beta_0 = 1$	0.98 / 0.02	0.95 / 0.02	0.93 / 0.02
$\beta_1 = -1$	-1.06 / 0.04	-1.12 / 0.04	-1.18 / 0.04
$a_0 = -0.5$	-0.51 / 0.10	-0.51 / 0.10	-0.52 / 0.10
$a_1 = 1$	1.02 / 0.17	1.03 / 0.17	1.04 / 0.17
$a_2 = 0.5$	0.52 / 0.11	0.53 / 0.11	0.54 / 0.11
$a_3 = 1$	1.03 / 0.25	1.04 / 0.27	1.05 / 0.29

Notes: Each column contains the mean and standard deviation of the 1000 maximum likelihood point estimates from the Monte Carlo experiment.

Table 2a
Monte Carlo 2: Maximum Likelihood Estimation of Time-Varying Transition Probability Markov-Switching Regression Model (TVP-MS) – Bias Correction from Equation (7)

	$\rho = 0.3$	$\rho = 0.6$	$\rho = 0.9$
Parameter	**Parameter Estimates: Mean / Std. Dev.**	**Parameter Estimates: Mean / Std. Dev.**	**Parameter Estimates: Mean / Std. Dev.**
T=200			
$\alpha_0 = 2$	2.00 / 0.09	2.00 / 0.09	2.00 / 0.07
$\alpha_1 = -2$	-1.98 / 0.23	-1.97 / 0.21	-1.97 / 0.16
$\beta_0 = 1$	1.00 / 0.04	0.99 / 0.04	0.99 / 0.04
$\beta_1 = -1$	-1.01 / 0.09	-1.02 / 0.08	-1.02 / 0.07
$a_0 = -0.5$	-0.52 / 0.16	-0.51 / 0.16	-0.51 / 0.14
$a_1 = 1$	1.02 / 0.29	1.01 / 0.27	1.03 / 0.22
$a_2 = 0.5$	0.53 / 0.18	0.52 / 0.17	0.51 / 0.14
$a_3 = 1$	1.12 / 0.48	1.10 / 0.44	1.09 / 0.37
T=500			
$\alpha_0 = 2$	1.99 / 0.06	1.99 / 0.05	2.00 / 0.04
$\alpha_1 = -2$	-1.98 / 0.14	-1.97 / 0.12	-1.98 / 0.1
$\beta_0 = 1$	1.00 / 0.03	1.00 / 0.02	1.00 / 0.02
$\beta_1 = -1$	-1.01 / 0.05	-1.02 / 0.05	-1.02 / 0.04
$a_0 = -0.5$	-0.50 / 0.09	-0.50 / 0.09	-0.49 / 0.08
$a_1 = 1$	1.00 / 0.17	1.01 / 0.16	1.00 / 0.13
$a_2 = 0.5$	0.50 / 0.10	0.51 / 0.10	0.51 / 0.08
$a_3 = 1$	1.04 / 0.25	1.03 / 0.24	1.02 / 0.19

Notes: Each column contains the mean and standard deviation of the 1000 maximum likelihood point estimates from the Monte Carlo experiment.

Table 2b
Monte Carlo 2: Maximum Likelihood Estimation of Time-Varying Transition Probability Markov-Switching Regression Model (TVP-MS) – Bias Correction from Equation (11)

	$\rho = 0.3$	$\rho = 0.6$	$\rho = 0.9$
Parameter	**Parameter Estimates: Mean / Std. Dev.**	**Parameter Estimates: Mean / Std. Dev.**	**Parameter Estimates: Mean / Std. Dev.**
T=200			
$\alpha_0 = 2$	2.00 / 0.14	2.01 / 0.13	2.01 / 0.13
$\alpha_1 = -2$	-1.97 / 0.29	-1.96 / 0.26	-1.96 / 0.26
$\beta_0 = 1$	1.00 / 0.05	1.00 / 0.05	1.00 / 0.04
$\beta_1 = -1$	-1.01 / 0.10	-1.02 / 0.09	-1.02 / 0.09
$a_0 = -0.5$	-0.51 / 0.16	-0.51 / 0.16	-0.51 / 0.15
$a_1 = 1$	1.02 / 0.28	1.03 / 0.29	1.02 / 0.25
$a_2 = 0.5$	0.50 / 0.19	0.51 / 0.18	0.50 / 0.16
$a_3 = 1$	1.13 / 0.51	1.11 / 0.50	1.09 / 0.43
T=500			
$\alpha_0 = 2$	2.00 / 0.08	2 / 0.08	2.01 / 0.07
$\alpha_1 = -2$	-1.98 / 0.17	-1.96 / 0.17	-1.97 / 0.15
$\beta_0 = 1$	1.00 / 0.03	1.00 / 0.03	1.00 / 0.03
$\beta_1 = -1$	-1.01 / 0.06	-1.02 / 0.06	-1.02 / 0.05
$a_0 = -0.5$	-0.50 / 0.10	-0.50 / 0.10	-0.49 / 0.09
$a_1 = 1$	1.01 / 0.17	1.01 / 0.16	0.99 / 0.13
$a_2 = 0.5$	0.50 / 0.11	0.51 / 0.11	0.50 / 0.09
$a_3 = 1$	1.05 / 0.27	1.03 / 0.25	1.00 / 0.21

Notes: Each column contains the mean and standard deviation of the 1000 maximum likelihood point estimates from the Monte Carlo experiment.

Table 3a
Monte Carlo 3: Maximum Likelihood Estimation of Fixed Transition Probability Markov-Switching Regression Model (FTP-MS) – Bias Correction from Equation (7)

	$\rho = 0.3$	$\rho = 0.6$	$\rho = 0.9$
Parameter	**Parameter Estimates: Mean / Std. Dev.**	**Parameter Estimates: Mean / Std. Dev.**	**Parameter Estimates: Mean / Std. Dev.**
T=200			
$\alpha_0 = 2$	2.01 / 0.14	2.01 / 0.13	2.00 / 0.09
$\alpha_1 = -2$	-2.01 / 0.28	-2.01 / 0.25	-1.99 / 0.17
$\beta_0 = 1$	0.99 / 0.05	1.00 / 0.05	0.99 / 0.04
$\beta_1 = -1$	-1.01 / 0.11	-1.01 / 0.09	-1.02 / 0.07
$a_0 = -0.5$	-0.49 / 0.15	-0.50 / 0.14	-0.50 / 0.14
$a_1 = 1$	0.98 / 0.22	0.99 / 0.21	1.00 / 0.20
$a_2 = 0$	---	---	---
$a_3 = 0$	---	---	---
T=500			
$\alpha_0 = 2$	2.00 / 0.08	2.00 / 0.08	2.00 / 0.06
$\alpha_1 = -2$	-1.99 / 0.16	-1.99 / 0.14	-1.99 / 0.12
$\beta_0 = 1$	1.00 / 0.03	0.99 / 0.03	0.99 / 0.02
$\beta_1 = -1$	-1.00 / 0.06	-1.01 / 0.06	-1.01 / 0.05
$a_0 = -0.5$	-0.50 / 0.09	-0.49 / 0.09	-0.49 / 0.09
$a_1 = 1$	1.00 / 0.14	0.99 / 0.14	0.99 / 0.13
$a_2 = 0$	---	---	---
$a_3 = 0$	---	---	---

Notes: Each column contains the mean and standard deviation of the 1000 maximum likelihood point estimates from the Monte Carlo experiment.

Table 3b
Monte Carlo 3: Maximum Likelihood Estimation of Fixed Transition Probability Markov-Switching Regression Model (FTP-MS) – Bias Correction from Equation (11)

	$\rho = 0.3$	$\rho = 0.6$	$\rho = 0.9$
Parameter	**Parameter Estimates: Mean / Std. Dev.**	**Parameter Estimates: Mean / Std. Dev.**	**Parameter Estimates: Mean / Std. Dev.**
T=200			
$\alpha_0 = 2$	2.01 / 0.20	2.01 / 0.19	2.03 / 0.19
$\alpha_1 = -2$	-2.02 / 0.39	-2.02 / 0.37	-2.01 / 0.34
$\beta_0 = 1$	1.00 / 0.06	0.99 / 0.05	0.99 / 0.04
$\beta_1 = -1$	-1.01 / 0.11	-1.02 / 0.10	-1.02 / 0.08
$a_0 = -0.5$	-0.49 / 0.15	-0.50 / 0.15	-0.50 / 0.15
$a_1 = 1$	0.97 / 0.21	0.99 / 0.23	0.99 / 0.21
$a_2 = 0$	---	---	---
$a_3 = 0$	---	---	---
T=500			
$\alpha_0 = 2$	2.01 / 0.11	2.00 / 0.11	2.02 / 0.10
$\alpha_1 = -2$	-2.00 / 0.23	-2.00 / 0.22	-2.00 / 0.21
$\beta_0 = 1$	1.00 / 0.03	0.99 / 0.03	0.99 / 0.02
$\beta_1 = -1$	-1.01 / 0.06	-1.01 / 0.06	-1.01 / 0.05
$a_0 = -0.5$	-0.49 / 0.09	-0.50 / 0.10	-0.49 / 0.09
$a_1 = 1$	0.98 / 0.13	0.99 / 0.14	0.98 / 0.13
$a_2 = 0$	---	---	---
$a_3 = 0$	---	---	---

Notes: Each column contains the mean and standard deviation of the 1000 maximum likelihood point estimates from the Monte Carlo experiment.

Table 4
Maximum Likelihood Estimates of Volatility-Feedback Model from Turner, Startz and Nelson (1989)
(standard errors in parenthesis)

Parameter	Without Bias Correction	With Bias Correction
θ_0	0.06 (0.06)	0.04 (0.07)
θ_1	0.020 (0.30)	0.19 (0.34)
θ_2	-2.11 (0.52)	-1.12 (0.49)
$\sigma_{\varsigma^*,0}$	0.38 (0.01)	0.40 (0.02)
$\sigma_{\varsigma^*,1}$	0.67 (0.04)	0.74 (0.06)
a_0	-2.11 (0.52)	-2.06 (0.16)
a_1	3.05 (0.27)	3.26 (0.30)
ρ	---	-0.42 (0.19)