THE OBJECTIVE FUNCTION OF SIMULATION ESTIMATORS NEAR THE BOUNDARY OF THE UNSTABLE REGION OF THE PARAMETER SPACE

George Tauchen*

Abstract—The paper examines the role of stability constraints in estimation by dynamic simulation. In particular, it analyzes the behavior of the objective function on either side of the boundary of the stability region of the parameter space. The main finding is that stability constraints may be ignored because the simulation-based objective function contains a built-in penalty to enforce stability. A key caveat, however, is that the dynamic stability of the auxiliary model that defines the moment conditions must be checked and enforced. An attempt to fit via simulation to moments defined by a dynamically unstable auxiliary model can be expected to lead to an ill-behaved objective function.

I. Introduction

THIS PAPER examines the behavior of the objective function of dynamic simulation estimators on either side of the boundary of the unstable region of the parameter space. The estimation context is the dynamic simulation framework of Laroque and Selanie (1989), Smith (1990, 1993), Ingram and Lee (1991), Duffie and Singleton (1993), Gourieroux et al. (1993), Gourieroux and Monfort (1996), Bansal et al. (1993, 1995), Gallant and Tauchen (1996a), and elsewhere. In this context, the underlying economic model, which here is termed the *maintained model*, defines a data generating process from which it is straightforward to generate artificial time series. The data generating process maps a realization of the underlying forcing process and a value of a parameter vector to an outcome process, of which some or all components are observed by the econometrician.

The central question addressed in this paper is whether one needs to impose stability constraints on the maintained model, that is, whether one needs to restrict the parameter space to that region where the maintained model generates stationary and ergodic data. The issue is of practical importance because the econometric theory presupposes stability, while deducing the actual stability constraints for a particular dynamic model can be quite formidable.

The question is investigated for a class of simulation estimators termed minimum simulated deviation (MSD). An MSD estimator mimics in the simulation the condition that an auxiliary function, called the deviation function, vanishes when averaged over the observed data. The MSD class is sufficiently broad to include many computationally tractable dynamic simulation-based estimators. The investigation starts with the case where the auxiliary deviation function is a general polynomial function of the data. It then proceeds to cases involving simulation estimators for the first and second moments of a time series. The analysis of first moment dynamics utilizes recent results from the unit-roots literature to get precise information on the behavior of the simulated objective function around the boundary of the unstable region. The analysis of the second moment dynamics does not attain quite such sharp results, as there are additional subtle aspects to the behavior of simulated second moments. Nonetheless, the analytical results and an illustrative example indicate close parallels between the first and second moment cases.

The upshot is that, subject to a major caveat, one need not worry about imposing dynamic stability on the maintained model itself. Dynamic stability is self-enforcing. If the optimizer wanders into the region of the parameter space where the maintained model is unstable, then the data simulator generates a wildly explosive simulated realization. The time-series properties of this explosive realization are very much unlike the time-series properties of the observed data set over which the auxiliary deviation function is averaged. Hence, the objective function attains an exceedingly high value. Lack of stability thus entails an automatic and very stringent penalty. Practical experience is consistent with this. Typically, the optimizer is seen to wander once or twice into the unstable region, but it quickly figures out that this is a very unpromising region of the parameter space and stays away from then on.

The caveat is that automatic stability is ensured if the auxiliary model that defines the moment conditions is itself dynamically stable. As seen in section III, the use of a dynamically unstable auxiliary model can be expected to define an MSD objective function with very poor numerical properties in both the stable and the unstable regions of the parameter space. Thus, stability must be checked and enforced on the auxiliary model.

II. Setup and Notation

A. Expectations under the Maintained Model

The maintained model is described as follows. Assume the conditional density h of a vector of outcome variables v_t is given by

$$v_t \sim h(v_t | v_{t-1}, w_t, \psi) \tag{1}$$

where w_t is a vector forcing process, and ψ is an $l_\psi \times 1$ parameter vector. The forcing process w_t is assumed to be generated as

$$w_t \sim q(w_t | w_{t-1}, \gamma) \tag{2}$$

where $q(w_t | w_{t-1}, \gamma)$ is the conditional density of w_t given w_{t-1}, and γ is an $l_\gamma \times 1$ parameter vector. The restriction to only one lag of v_{t-1} appearing in equation (1) and one lag of w_t appearing in equation (2) is without loss of generality.

Received for publication December 11, 1995. Revision accepted for publication June 10, 1997.

* Duke University.

The author wishes to thank three referees and the editor for comments that substantially improved this paper. The research is supported by the National Science Foundation.

Any system with longer (but finite) lag lengths can be put into a first-order scheme as above by suitably redefining the variables. Define the state variable $s_t = (v_t' \quad w_t')'$ and the combined parameter vector $\rho = (\psi' \quad \gamma')' \in R \subset \Re^{l_\rho}$, where ρ is of length $l_\rho = l_\psi + l_\gamma$ and R is the parameter space. Not all elements of the state vector s_t may be observed, and we take

$$y_t = \phi(s_t)$$

as the observation equation, where y_t represents the observed vector and ϕ is a smooth function.

The maintained model thus defines a stochastic data generator for the state process $\{s_t\}$ and hence for the process $\{y_t\}$. It is assumed to generate stationary and ergodic data for each $\rho \in R$.

The econometrician observes the data set $\{\tilde{y}_t\}_{t=1}^n$. The maintained model is assumed to be correctly specified in the sense that the observed data set has been generated through it for some true value $\rho_0 \in R$. The econometrician's task is to estimate ρ_0. The optimal strategy is, of course, dynamic maximum likelihood. Use

$$\hat{\rho}_{\mathrm{MLE}} = \underset{\rho \in R}{\operatorname{argmax}} \sum_{t=L+1}^{n} \log\left[p(\tilde{y}_t | \tilde{y}_{t-L}, \ldots, \tilde{y}_{t-1}, \rho)\right]$$

where $p(y_t | y_{t-L}, \ldots, y_{t-1}, \rho)$ is the transition density of y_t given its past. But for a wide class of interesting models, the transition density is not available in simple closed form, rendering $\hat{\rho}_{\mathrm{MLE}}$ infeasible. The infeasibility of dynamic maximum likelihood motivates the use of simulated moment techniques.

For a candidate value $\rho \in R$, the econometrician generates a long simulated realization $\{\hat{y}_\tau(\rho)\}_{\tau=1}^N$. This is done by priming the lags, simulating the processes v_t and w_t by taking random draws iteratively through equations (1) and (2), letting the system run for a long time to let the effects of transients die out, and then retaining the last N values of simulated $\hat{y}_\tau(\rho)$. The econometrician computes via Monte Carlo integration the expectation of nonlinear functions under the maintained model. If $g(y_{t-L}, \ldots, y_t)$ is a function of a stretch of length $L+1$, then

$$E_\rho[g(y_{t-L}, \ldots, y_t)] = \int \cdots \int g(y_{t-L}, \ldots, y_t) \times p(y_{t-L}, \ldots, y_t | \rho) \prod_{j=0}^{L} dy_{t-j}$$

means the expectation of $g(y_{t-L}, \ldots, y_t)$ under the maintained model given $\rho \in R$, where $p(y_{t-L}, \ldots, y_t | \rho)$ is the implied joint density of $y_{t-L}, \ldots, y_t$ under the maintained model. It is presumed that g is sufficiently regular that the expectation exists. This expectation is computed via Monte Carlo integration as

$$E_\rho[g(y_{t-L}, \ldots, y_t)] \approx \frac{1}{N} \sum_{\tau=L+1}^{N} g[\hat{y}_{\tau-L}(\rho), \ldots, \hat{y}_\tau(\rho)].$$

For models that generate identically and independently distributed (iid) data, N might not need to be too large for the approximation to be quite accurate. However, for models that display high persistence in mean or variance, experience suggests that N needs to be on the order of 50,000–100,000 or more to get accurate answers (Tauchen (1997), Gallant and Tauchen (1997)), and Monte Carlo acceleration schemes can be quite helpful (Andersen and Lund (1997)).

B. *Simulation Estimators*

Duffie and Singleton (1993) analyze a dynamic *simulated method of moments* (SMM) estimator based on the vector

$$\Delta_N(\rho) = \frac{1}{N} \sum_{\tau=L+1}^{N} \nu[\hat{y}_{t-L}(\rho), \ldots, \hat{y}_t(\rho)] - \frac{1}{n} \sum_{t=L+1}^{n} \nu(\tilde{y}_{t-L}, \ldots, \tilde{y}_t) \tag{3}$$

where $\nu(y_{t-L}, \ldots, y_t)$ is a vector-valued function of a stretch of length $L+1$. The objective function for SMM is

$$Q_N^{\mathrm{SMM}}(\rho) = \Delta_N(\rho)' W \Delta_N(\rho) \tag{4}$$

where W is an appropriate weighting matrix. SMM minimizes the length of the difference between the expectation of $\nu(y_{t-L}, \ldots, y_t)$ computed from the simulation and computed from the data.

Here we consider a somewhat more general simulation estimator, termed *minimum simulated deviation* (MSD), which is based on an auxiliary deviation function $\eta(y_{t-L}, \ldots, y_t, \theta)$. The deviation function η maps $y_{t-L}, \ldots, y_t$ and an auxiliary parameter vector $\theta \in \Re^{l_\theta}$ to $\Re^{l_\theta}$. Let $\tilde{\theta}$ be defined by

$$\frac{1}{n} \sum_{t=L+1}^{n} \eta(\tilde{y}_{t-L}, \ldots, \tilde{y}_t, \tilde{\theta}) = 0. \tag{5}$$

The preceding comprises a system of l_θ equations in l_θ unknowns, which are presumed to admit a unique solution $\tilde{\theta}$. The MSD estimator mimics condition (5) in the simulation. Define

$$m_N(\rho, \tilde{\theta}) = \frac{1}{N} \sum_{\tau=L+1}^{N} \eta[\hat{y}_{t-L}(\rho), \ldots, \hat{y}_t(\rho), \tilde{\theta}] \tag{6}$$

which is the analogue of equation (5), but computed by averaging with respect to the maintained model instead of the empirical cumulative distribution functions (cdf) of the data. Define the objective function

$$Q_N(\rho) = m_N(\rho, \tilde{\theta})' W m_N(\rho, \tilde{\theta}) \tag{7}$$

where W is a weighting matrix. The MSD estimator is

$$\hat{\rho} = \underset{\rho}{\operatorname{argmin}}\, Q_N(\rho). \tag{8}$$

This estimator minimizes the expected length (computed via simulation) of the deviation function η. Evidently, the SMM estimator is an MSD estimator, as its objective function (4) takes the form of equation (7) with $\eta(\cdot, \theta) = \nu(\cdot) - \theta$.

A specific example of an MSD estimator is given as follows. Let y_t be scalar, $L = 2$, and $\theta \in \Re^4$. Set

$$\eta_1(y_{t-2}, y_{t-1}, y_t, \theta) = y_t - \theta_1 \tag{9}$$

$$\eta_2(y_{t-2}, y_{t-1}, y_t, \theta) = (y_t - \theta_1)^2 - \theta_2 \tag{10}$$

$$\eta_3(y_{t-2}, y_{t-1}, y_t, \theta) = \theta_2^{-1}(y_t - \theta_1)(y_{t-1} - \theta_1) - \theta_3 \tag{11}$$

$$\begin{aligned}\eta_4(y_{t-2}, y_{t-1}, y_t, \theta) = {} & [(y_t - \theta_1) - \theta_3(y_{t-1} - \theta_1)]^2 \\ & \times [(y_{t-1} - \theta_1) \\ & - \theta_3(y_{t-2} - \theta_1)]^2 - \theta_4.\end{aligned} \tag{12}$$

Then $(1/n)\sum_{t=3}^{n} \eta(\tilde{y}_{t-2}, \tilde{y}_{t-1}, \tilde{y}_t, \tilde{\theta}) = 0$ gives $\tilde{\theta}_1$ as the unconditional mean, $\tilde{\theta}_2$ as the unconditional variance, $\tilde{\theta}_3$ as the first-order autocorrelation coefficient, and $\tilde{\theta}_4$ as the first-order serial covariance of the squared linear innovation, which, in general, is nonzero for time-varying volatility models. Implementing MSD amounts to matching these moments. For this particular choice of η, the weighting matrix of the MSD estimator should be a weighted covariance estimator (Newey and West (1987), Andrews (1991)) in order to take account of serial correlation in the moment functions. Note, for subsequent reference, that each of the functions comprising η is a polynomial function of y_t and its lags. This will be important in the stability analysis later.

The *efficient method of moments* (EMM) estimator of Bansal et al. (1993, 1995) and Gallant and Tauchen (1996a) is an MSD estimator. For EMM, the deviation function η is defined by the scores of an auxiliary time-series model, that is, a conditional density $f(y_t | y_{t-L}, \ldots, y_{t-1}, \theta)$, $\theta \in \Theta \subset \Re^{l_\theta}$ for $\{y_t\}$, which is called the *score generator*. The score generator is a statistical model for $\{y_t\}$ that need not have anything directly to do with the maintained model. Let $\tilde{\theta}$ denote the quasi-maximum-likelihood (QML) estimator

$$\tilde{\theta} = \underset{\theta\in\Theta}{\operatorname{argmax}} \frac{1}{n} \sum_{t=L+1}^{n} \log [f(\tilde{y}_t | \tilde{y}_{t-L}, \ldots, \tilde{y}_{t-1}, \theta)]. \tag{13}$$

The first-order conditions are

$$\frac{1}{n} \sum_{t=L+1}^{n} s_f(\tilde{y}_{t-L}, \ldots, \tilde{y}_t, \tilde{\theta}) = 0 \tag{14}$$

where

$$\begin{aligned} & s_f(y_t | y_{t-L}, \ldots, y_{t-1}, \theta) \\ & \quad = \frac{\partial}{\partial\theta} \log [f(y_t | y_{t-L}, \ldots, y_{t-1}, \theta)]. \end{aligned} \tag{15}$$

Taking $\eta = s_f$, then the first-order condition (14) is a special case of equation (5). The objective function for EMM is

$$Q_N^{\text{EMM}}(\rho) = \hat{m}_N(\rho, \tilde{\theta})' \tilde{I}^{-1} \hat{m}_N(\rho, \tilde{\theta}) \tag{16}$$

where

$$\tilde{I} = \frac{1}{n} \sum_{t=L+1}^{n} s_f(\tilde{y}_{t-L}, \ldots, \tilde{y}_t, \tilde{\theta}) s_f'(\tilde{y}_{t-L}, \ldots, \tilde{y}_t, \tilde{\theta})$$

is the estimate of the information matrix. If the auxiliary model fails to provide an adequate statistical description of the data, then a weighted covariance estimator would be used for $\tilde{I}$, as discussed in Gallant and Tauchen (1996a).

Using arguments parallel to those of Duffie and Singleton (1993), Gallant and Tauchen (1996a), and Tauchen (1997), one can show that so long as the deviation function $\eta(\cdot, \theta)$ in equation (5) is sufficiently regular that there exists a pseudotrue value $\bar{\theta}$ such that $\sqrt{n}(\tilde{\theta} - \bar{\theta}) \to N(0, V_\theta)$, then the MSD estimator has a well-behaved asymptotic distribution. Doing so formally is beyond the scope of this paper.

Smith (1990, 1993) introduces a different simulation estimator, which is extended by Gourieroux et al. (1993) and termed *indirect inference*. This estimator matches directly to the QML estimator $\tilde{\theta}$ defined in equation (13). Set

$$\begin{aligned} \bar{\theta}_N(\rho) = {} & \underset{\theta}{\operatorname{argmax}} \frac{1}{N} \sum_{\tau=L+1}^{N} \\ & \log [f(\hat{y}_\tau(\rho) | \hat{y}_{\tau-L}(\rho), \ldots, \hat{y}_{\tau-1}(\rho), \theta)] \end{aligned} \tag{17}$$

which is the implied estimate of θ, that is, the simulated *binding function*, when QML estimation is applied to the simulated data set $\{y_\tau(\rho)\}_{\tau=1}^{N}$. The objective function for indirect inference is

$$Q_N^{\text{II}}(\rho) = [\bar{\theta}_N(\rho) - \tilde{\theta}]' [\tilde{H}^{-1} \tilde{I} \tilde{H}^{-1}]^{-1} [\bar{\theta}_N(\rho) - \tilde{\theta}]$$

where $\tilde{H}$ and $\tilde{I}$ are estimates of the Hessian and information matrices, respectively. Indirect inference involves a Hessian estimation, and, in order to evaluate $Q_N^{\text{II}}(\rho)$, it entails evaluating the binding function (17) at each simulated realization $\{\hat{y}_\tau(\rho)\}_{\tau=1}^N$, which can be numerically very intensive.

Because of the optimization in equation (17), the indirect inference estimator is not an MSD estimator as defined in equation (8), which only involves evaluation of simulation means as in equations (3) and (6). Interestingly, as noted by a referee, and as will be seen below, the behavior of the indirect inference objective function on the unstable region turns out to be different from that of an MSD estimator. The behavior is difficult to analyze in more complex models like that of Section IIIC because of the additional optimization associated with evaluating the binding function. In what follows we concentrate on the properties of the objective functions of the numerically tractable MSD estimators. Where possible, we point out the different behavior of indirect inference.

III. Main Results

The central topic of interest for this paper is the domain over which ρ is allowed to vary in performing the minimization of $Q_N(\rho)$ for the MSD estimator as defined in equation (8). It is simple enough to write down

$$\operatorname*{argmin}_{\rho \in \mathcal{R}} Q_N(\rho). \tag{18}$$

However, this presumes the econometrician actually has figured out the exact region $\mathcal{R}$ over which the maintained model generates stationary and ergodic data. Typically, the parameter space $\mathcal{R}$ is a subset of a larger and more natural space $\mathcal{R}^* = \mathcal{R} \cup \mathcal{R}_b \cup \mathcal{R}_e \subset \Re^{l_\rho}$. The regions $\mathcal{R}$, $\mathcal{R}_b$, and $\mathcal{R}_e$ are mutually disjoint, and for $\rho \in \mathcal{R}$ the maintained model generates strictly stationary and ergodic data; for $\rho \in \mathcal{R}_b$ it generates borderline (unit-root type) nonstationary data; for $\rho \in \mathcal{R}_e$ it generates explosive data. The presumption that the true data-generation process (DGP) is stationary and ergodic means $\rho_0 \in \mathcal{R}$. In formulating the model, however, it is often more natural to work in terms of $\mathcal{R}^*$ than $\mathcal{R}$.

In actual practice, computing equation (18) means that the stability constraint has to be imposed during the estimation. For anything more complicated than a univariate AR(1), characterizing the stability set $\mathcal{R}$ within $\mathcal{R}^*$ can be burdensome. Even for an AR(2), the stability set depends on the parameters in a complicated manner, as suggested by the famous diagram for a second-order deterministic autoregression (Baumol (1970, p. 221)). Monahan (1984) provides a strategy for imposing stability on a univariate AR(L). More generally, though, for any kind of complicated nonlinear problem, the characterization of $\mathcal{R}$ represents a formidable challenge. The natural thing to do is simply to ignore the stability constraints and attempt to compute

$$\operatorname*{argmin}_{\rho \in \mathcal{R}^*} Q_N(\rho)$$

without worrying too much about problems when the optimizer tries to evaluate $Q_N(\rho)$ for $\rho \in \mathcal{R}_b$ or for $\rho \in \mathcal{R}_e$. We investigate the consequences of doing so.

A. *Polynomial Deviation Functions*

Consider the example defined by equations (9)–(12), and note in particular that the η functions are polynomial functions of y_t and its lags. Using polynomial functions to define moment conditions for SMM estimation is typical in practice, for example, Foster and Viswanathan (1995). Likewise, polynomial functions typically appear in the score function of reasonable auxiliary models for EMM estimation. For example, if $f(y_t | y_{t-L}, \ldots, y_{t-1}, \theta)$ is based on the exponential family, then polynomials will appear in the score function.

Suppose, for simplicity, that y_t is a scalar process and that $\eta(h, \theta) = \sum_{k=0}^K \alpha_k(\theta) y^k$ is a scalar-valued function polynomial whose coefficients $\alpha_k(\theta)$ depend on a scalar θ. The components of η in equations (9)–(12) are polynomials like this. Then

$$\sum_{k=0}^{K} \alpha_k(\tilde{\theta}) \left(\frac{1}{n} \sum_{t=1}^{n} \tilde{y}_t^k \right) = 0 \tag{19}$$

determines $\tilde{\theta}$. Throughout, we condition on the observed data set $\{\tilde{y}_t\}_{t=1}^n$, and hence $\tilde{\theta}$ in equation (19) is fixed. Specializing equation (6) to this case, which is the mimicking of equation (19), gives

$$m_N(\rho, \tilde{\theta}) = \sum_{k=0}^{K} \alpha_k(\tilde{\theta}) \left(\frac{1}{N} \sum_{\tau=1}^{N} \hat{y}_\tau^k(\rho) \right). \tag{20}$$

Now let $\rho \in \mathcal{R}^* - \mathcal{R}$, which is the nonergodic region on which one expects $|\hat{y}_N(\rho)| \to \infty$ as $N \to \infty$. Consider

$$\frac{|\hat{y}_N(\rho)|}{N} \to \infty. \tag{21}$$

By this it is meant that for each $c > 0$ there exist N_c such that $|\hat{y}_N(\rho)| \geq N_c$ for all $N \geq N_c$. This is simple linear divergence of the simulated realization, which would be easily satisfied by, say, a realization from an autoregression with a coefficient on or outside the unit circle. The Cesaro sum of any power of $\hat{y}_\tau(\rho)$ must likewise diverge:

Proposition 1: Suppose $\rho \in \mathcal{R}^* - \mathcal{R}$ and equation (21) holds. Then, for each $k \geq 1$,

$$\left| \frac{1}{N} \sum_{\tau=1}^{N} \hat{y}_\tau(\rho)^k \right| \to \infty. \tag{22}$$

Proof: For sufficiently large N, $|\hat{y}_N(\rho)| \geq 1$ and so $|\hat{y}_N(\rho)|^k \geq |\hat{y}_N(\rho)|$; hence, $|\hat{y}_N(\rho)|^k/N \to \infty$. But

$$\frac{|\hat{y}_N(\rho)^k|}{N} \leq \left|\frac{1}{N}\sum_{\tau=1}^{N} \hat{y}_\tau(\rho)^k\right| + \frac{N-1}{N}\left|\frac{1}{N-1}\sum_{\tau=1}^{N-1} \hat{y}_\tau(\rho)^k\right|$$

because any sequence $\{y_\tau\}_{\tau=1}^N$ satisfies $|y_N| \leq |\Sigma_{\tau=1}^N y_\tau| + |\Sigma_{\tau=1}^{N-1} y_\tau|$. Therefore

$$\infty = \lim_{N\to\infty} \frac{|\hat{y}_N(\rho)^k|}{N} \leq 2 \lim_{N\to\infty} \left|\frac{1}{N}\sum_{\tau=1}^{N} \hat{y}_\tau(\rho)^k\right|$$

which establishes equation (22). □

From proposition 1 we expect the magnitude of each of the individual terms comprising the outer sum on the right-hand side of equation (20) to diverge. Hence we expect, in general, for $\rho \in \mathcal{R}^* - \mathcal{R}$, that $|m_N(\rho, \tilde{\theta})| \to \infty$ as $N \to \infty$, and thus we expect the MSD objective function (7) to satisfy $Q_N(\rho) \to \infty$ as $N \to \infty$ when $\rho \in \mathcal{R}^* - \mathcal{R}$. In other words, for sufficiently large N the MSD objective function incorporates an automatic penalty on parameters in the unstable region of the parameter space.

There is a caveat in this reasoning. Although each of the terms in equation (20) diverges, there could potentially be values of ρ that generate simulated data that exactly (or nearly) match roots of the polyomial. In other words, there potentially can be isolated values $\rho_N^\# \in \mathcal{R}^* - \mathcal{R}$ such that

$$m_N(\rho_N^\#) \approx 0 \tag{23}$$

even though each of the individual terms of the polynomial defining $m_N(\rho, \tilde{\theta})$ is divergent on $\rho \in \mathcal{R}^* - \mathcal{R}$. This can happen if there are unstable dynamics embedded in the deviation function $\eta(y_{t-L}, \ldots, y_t, \tilde{\theta})$ and there exists a value $\rho_N^\# \in \mathcal{R}^* - \mathcal{R}$ such that the realization $\{\hat{y}_\tau(\rho_N^\#\}_{\tau=1}^N$ reproduces those unstable dynamics.

Further insight is available in subsequent analysis.

B. First Moment Analysis

We consider the case where one fits an AR(1) via MSD. One would never actually do this in practice, but the analysis provides some crucial insights. More can be said in this case, because the unit-root literature provides specific information on the rates at which simulated data diverge in the nonergodic region of the parameter space.

Suppose the true (and maintained) DGP is

$$y_t = \rho y_{t-1} + \epsilon_t, \qquad \epsilon_t \text{ iid } N(0, 1) \tag{24}$$

and $\rho_0 \in (-1, 1)$. The econometrician knows the DGP, observes $\{\tilde{y}_t\}_{t=1}^n$, and must estimate ρ. Evidently $\hat{\rho}_{\text{MLE}} = \Sigma_{t=2}^n \tilde{y}_{t-1}\tilde{y}_t/\Sigma_{t=2}^n \tilde{y}_{t-1}^2$ is readily available, and there is no actual need to undertake estimation by simulation. Nonetheless, for the sake of examining the role of the stability constraints, we consider MSD estimation for this special case.

Take the deviation function as

$$\eta(y_{t-1}, y_t, \theta) = y_{t-1}(y_t - \theta y_{t-1}). \tag{25}$$

This deviation function entails matching to the first-order autocorrelation coefficient, or, as would amount to the same thing, it is EMM using

$$y_t = \theta y_{t-1} + \epsilon_t, \qquad \epsilon_t \text{ iid } N(0, 1)$$

as the score generator to define $f(y_t|\, y_{t-1}, \theta)$. Specializing equation (5) to this situation gives

$$\tilde{\theta} = \frac{\sum_{t=2}^{n} \tilde{y}_{t-1}\tilde{y}_t}{\sum_{t=2}^{n} \tilde{y}_{t-1}^2}.$$

The natural estimate of the weight matrix (which is actually scalar) is $W = \tilde{\mathcal{I}}^{-1}$, where

$$\tilde{\mathcal{I}} = \frac{1}{n}\sum_{t=2}^{n} \tilde{y}_{t-1}^2(\tilde{y}_t - \tilde{\theta}\tilde{y}_{t-1})^2.$$

Everything from now on is conditional on the fixed data set $\{\tilde{y}_t\}_{t=1}^n$. Hence, $\tilde{\theta}$ and $\tilde{\mathcal{I}}$ are fixed throughout. We are interested in the behavior of the objective function $Q_N(\rho)$ for a given $\tilde{\theta}$ as ρ varies and $N \to \infty$.

Simulated data $\{\hat{y}(\rho)_\tau\}_{\tau=1}^N$ are generated by drawing $\{\hat{\epsilon}_\tau\}_{\tau=1}^N$ iid $N(0, 1)$, $\hat{y}_0$ given, and forming $\hat{y}_\tau(\rho) = \rho\hat{y}_{\tau-1}(\rho) + \hat{\epsilon}_\tau$, $\tau = 1, 2, \ldots, N$. All limits will be taken as simulated sample size $N \to \infty$, with the econometric sample size fixed at n. All randomness is with respect to the random number generator that delivers $\{\hat{\epsilon}_\tau\}_{\tau=1}^N$. The intrinsic randomness of the underlying model (24) is irrelevant here because we are conditioning on the particular data set $\{\tilde{y}_t\}_{t=1}^n$.

Specializing the expression (6) gives

$$m_N(\rho, \tilde{\theta}) = \frac{1}{N}\sum_{\tau=2}^{N} \hat{y}_{\tau-1}(\rho)[\hat{y}_\tau(\rho) - \tilde{\theta}\hat{y}_{\tau-1}(\rho)]$$

for the simulated moment conditions. Thus equation (7) specializes to

$$Q_N(\rho) = \frac{[m_N(\rho, \tilde{\theta})]^2}{\tilde{\mathcal{I}}}. \tag{26}$$

Write

$$m_N(\rho, \tilde{\theta}) = A_N(\rho) + B_N(\rho)\,(\rho - \tilde{\theta})$$

where

$$A_N(\rho) = \frac{1}{N}\sum_{\tau=2}^{N} \hat{y}_{\tau-1}(\rho)\,[\hat{y}_\tau(\rho) - \rho\hat{y}_{\tau-1}(\rho)]$$

$$= \frac{1}{N}\sum_{\tau=2}^{N} \hat{y}_{\tau-1}(\rho)\hat{\epsilon}_\tau$$

and

$$B_N(\rho) = \frac{1}{N}\sum_{\tau=2}^{N} [\hat{y}_{\tau-1}(\rho)]^2.$$

Observe that

$$Q_N(\rho) = \frac{[A_N(\rho) + B_N(\rho)\,(\rho - \tilde{\theta})]^2}{\tilde{I}}.$$

The limiting behavior of $A_N(\rho)$ and $B_N(\rho)$, and thus $Q_N(\rho)$, depends on ρ. We consider only $-1 < \rho < 1$ and $\rho \geq 1$, as the situation $\rho < -1$ is not very relevant from a practical perspective and can be handled in a similar manner as $\rho > 1$ anyway. By standard results,

$$\text{for } -1 < \rho < 1, \qquad A_N(\rho) \xrightarrow{\text{as}} 0 \tag{27}$$

and

$$\text{for } -1 < \rho < 1, \qquad B_N(\rho) \xrightarrow{\text{as}} \frac{1}{1 - \rho^2}. \tag{28}$$

On the other hand, for $\rho \geq 1$, there exist (Abadir (1993, theorem 3.1)) sequences $\{\alpha_N(\rho)\}$, $\{\beta_N(\rho)\}$, such that $\alpha_N(\rho) \to \infty$, $\beta_N(\rho) \to \infty$, $\alpha_N(\rho)/\beta_N(\rho) \to 0$, as $N \to \infty$, and

$$\frac{A_N(\rho)}{\alpha_N(\rho)} \xrightarrow{D} \mathcal{A}(\rho)$$

$$\frac{B_N(\rho)}{\beta_N(\rho)} \xrightarrow{D} \mathcal{B}(\rho)$$

where $\mathcal{A}(\rho)$ and $\mathcal{B}(\rho)$ are nondegenerate random variables (with respect to the random number generator that generates $\{\hat{\epsilon}_\tau\}_{\tau=1}^{N}$) whose realizations depend also on the value of ρ. For $\rho > 1$, the sequence $\beta_N(\rho) = \text{const} \cdot \rho^{2N}$. Anything that is $O[\beta_N(\rho)]$ diverges rapidly indeed.

The behavior of $Q_N(\rho)$ in equation (26) depends critically on whether the estimate $\tilde{\theta}$ lies in the stable region or not.

We start with the case $\tilde{\theta} \in (-1, 1)$. In this case, for $-1 < \rho < 1$ and N large, by equations (27) and (28) we have that

$$Q_N(\rho) \approx \frac{(\rho - \tilde{\theta})^2}{\tilde{I}\,(1 - \rho^2)^2}, \qquad \text{for } -1 < \rho < 1. \tag{29}$$

For $\rho \geq 1$, $[\beta_N(\rho)]^{-2}Q_N(\rho) \xrightarrow{D} (\rho - \tilde{\theta})^2\mathcal{B}^2(\rho)/\tilde{I}$. Hence for N large,

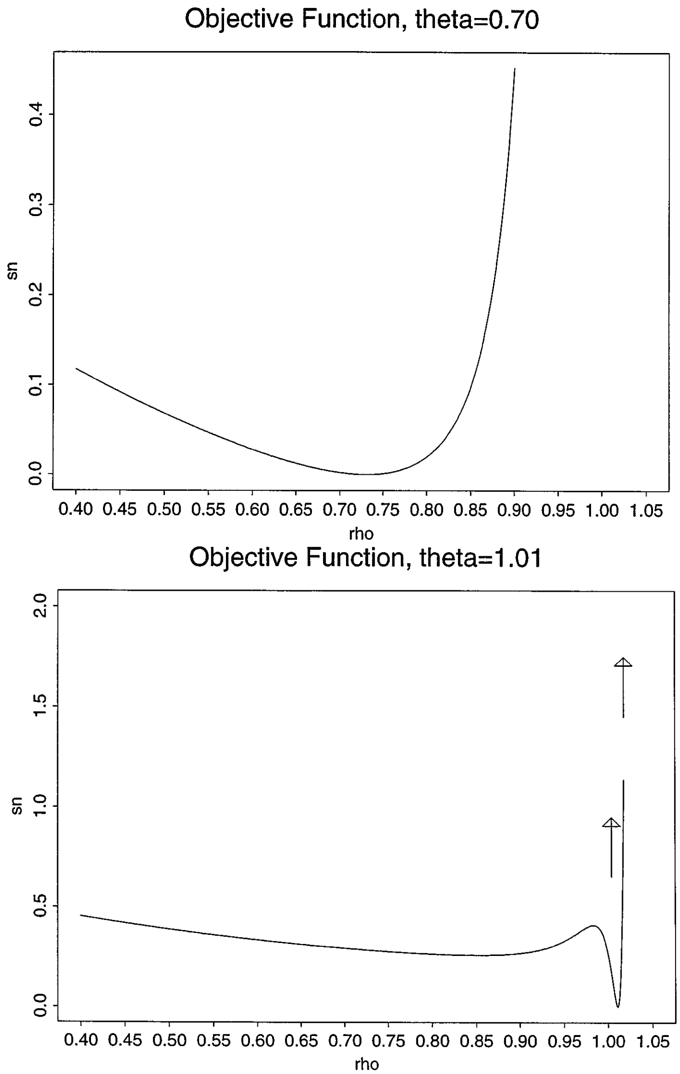
$$Q_N(\rho) \approx \beta_N^2(\rho)\,\frac{(\rho - \tilde{\theta})^2}{\tilde{I}}\,\mathcal{B}^2(\rho), \qquad \text{for } \rho \geq 1. \tag{30}$$

Considering equations (29) and (30), as $N \to \infty$, the objective function $Q_N(\rho)$ is asymptotically minimized at $\rho = \tilde{\theta}$ while $Q_N(\rho)$ diverges (rapidly) to infinity for $\rho \geq 1$. The objective function places a huge penalty on ρ entering the unstable region and, in effect, incorporates a built-in penalty function to enforce stability on the fit.

The behavior of $Q_N(\rho)$ is different if $\tilde{\theta} \geq 1$. For $-1 < \rho < 1$, $Q_N(\rho)$ still behaves as in equation (29). Also, if $\rho \geq 1$ and $\rho \neq \tilde{\theta}$, then $[\beta_N(\rho)]^{-2}Q_N(\rho) \xrightarrow{D} (\rho - \tilde{\theta})^2\mathcal{B}^2(\rho)/\tilde{I}$ and $Q_N(\rho) \to \infty$. However, at $\rho = \tilde{\theta}$, $[\beta_N(\tilde{\theta})]^{-2}Q_N(\tilde{\theta}) = [A_N(\tilde{\theta})/\beta_N(\tilde{\theta})]^2/\tilde{I} \xrightarrow{\text{as}} 0$. Thus $Q_N(\rho)$ diverges (rapidly) to infinity everywhere in the unstable region except at $\rho = \tilde{\theta}$, where $Q_N(\tilde{\theta}) \to 0$. The objective function can thus be expected to admit a sharp local minimum at ρ near $\tilde{\theta}$ in the unstable region. However, if $\tilde{\theta} > 1$ then, since $Q_N(\rho)$ diverges for $\rho \in [1, \tilde{\theta}) \cup (\tilde{\theta}, \infty)$, but converges for $\rho \in (-1, 1)$, by continuity we can expect $Q_N(\rho)$ to admit another local minimum in the stable region.

Figure 1 illustrates these analytical results. The top panel shows a realization of the objective function for $N = 100$ for a case where the estimator computed from observed data is $\tilde{\theta} = 0.70$. The objective function is nicely behaved and admits a local minimum near 0.70, as expected. The bottom panel shows a realization (with the same random numbers) and $N = 100$ but $\tilde{\theta} = 1.01$. As expected, there is a sharp global minimum for $\rho \approx 1.01$ and a second false local minimum in the stable region. The optimization software might have a great deal of difficulty handling this objective function and might easily stop at the false minimum in the stable region.

FIGURE 1.—OBJECTIVE FUNCTION $Q_N(\rho)$



The top panel shows a simulated realization of the objective function for $N = 100$ and $\tilde{\theta} = 0.70$. The bottom panel shows a simulated realization for $N = 100$ and $\tilde{\theta} = 1.01$. Note the local minimum at about $\rho = 0.85$ in the lower panel.

The deviation function (25) is a polynomial function of the data, and the behavior in figure 1 is consistent with the analysis of section IIA. Almost everywhere in the unstable region $\rho \in R^* - R$ the simulated mean $m_N(\rho, \tilde{\theta})$ rapidly diverges. The exception occurs when $\tilde{\theta}$ itself lies in the unstable region, and there is thus an isolated point in $R^* - R$ where the maintained model can replicate the unstable dynamics.

The message is that for MSD estimation, there is no reason to impose stability constraints in the second-round optimization in computing argmin $[Q_N(\rho)]$, as the constraints will be self-imposed via the nature of the objective function. It is important to check that the deviation function itself contains no embedded instabilities as they can lead to a numerical nightmare.

A referee pointed out that the objective function of indirect inference behaves differently, though the conclusions are somewhat the same. For this example,

$$Q_N^{\mathrm{II}}(\rho) = [\bar{\theta}_N(\rho) - \tilde{\theta}]^2 \tag{31}$$

where

$$\bar{\theta}_N(\rho) = \frac{\sum_{\tau=2}^{N} \hat{y}_{\tau-1}(\rho)\hat{y}_\tau(\rho)}{\sum_{\tau=2}^{N} \hat{y}_{\tau-1}(\rho)^2}. \tag{32}$$

By well-known results $\bar{\theta}_N(\rho) \to \rho$ at rate $N^{1/2}$ for $\rho \in (-1, 1)$ and at rate N for $\rho = 1$, while Abadir's (1993) results indicate that $\bar{\theta}_N(\rho) \to \rho$ exponentially for $\rho > 1$. In effect, for large N the indirect inference objective function is

$$Q_N^{\mathrm{II}}(\rho) \approx (\rho - \tilde{\theta})^2$$

a simple quadratic function of ρ. Thus, if $-1 < \tilde{\theta} < 1$, then indirect inference will not, for large N, generate an estimate of ρ in the unstable region. This is analogous to MSD but the mechanism is different, and the objective function does not diverge as $N \to \infty$ for ρ in the unstable region. (Actually, in practice with machine arithmetic, $Q_N^{\mathrm{II}}(\rho)$ could be quite ill-behaved on $R^* - R$ for large N because both the numerator and the denominator of equation (32) diverge rapidly and the arithmetic becomes numerically unstable.) On the other hand, if $\tilde{\theta} \geq 1$, then indirect inference can likewise be expected to generate an unstable estimate, though there will not be the false local minimum as in the bottom panel of figure 1.

C. Second Moment Analysis

To assess second moment properties, we investigate estimation of the stochastic volatility model for security returns,

$$y_t = \exp(h_t/2)z_{1t} \tag{33}$$

$$h_t = \rho_1 + \rho_2 h_{t-1} + \rho_3 z_{2t}. \tag{34}$$

Here z_{1t} and z_{2t} are unobserved iid $N(0, 1)$ processes and y_t is the observed return. In the above, h_t is log volatility, which is generated via the autonomous autoregression of the second equation. The stochastic volatility model, which is reviewed at great length in Ghysels et al. (1995), is a widely used model for describing random fluctuations of financial volatility. As is well understood, the likelihood function of stochastic volatility models can be difficult to evaluate, but expectations and simulations are easy, which motivates use of moment matching estimation as in Andersen and Sorensen (1996) and Gallant et al. (1997).

The key parameter vis-à-vis stability is ρ_2. Consider a simulated realization $\{\hat{y}_\tau\}$ generated with $\rho_2 > 1$, which implies an explosive trajectory $|\hat{h}_\tau| \rightarrow \infty$ for volatility. In the remainder, we suppress the dependence of simulated data on ρ to help simplify formulas. Observe that

$$E_{\tau-1}\left(\frac{\hat{y}_\tau^2}{\hat{y}_{\tau-1}^2}\right) = \frac{1}{z_{1,\tau-1}^2} \exp\left[(\rho_2 - 1)\hat{h}_{\tau-1}\right] \exp\left(\rho_1 + \frac{\rho_3^2}{2}\right)$$

where $E_{\tau-1}(\cdot)$ means $E(\cdot|\{z_{1s}, z_{2s}\}_{s=1}^{\tau-1}, h_0)$. Let B denote an arbitrary large positive constant. Along the explosive trajectory, $\{\hat{h}_\tau\}$ will visit the region $[h \in \Re\colon \quad h < -B]$ infinitely often, in which times $\exp[(\rho_2 - 1)\hat{h}_{\tau-1}] \approx 1$ and $\hat{y}_\tau^2/\hat{y}_{\tau-1}^2$ is roughly constant on average; but it will also visit the region $[h \in \Re\colon \quad h > B]$ infinitely often, in which case $\exp[(\rho_2 - 1)\hat{h}_{\tau-1}]$ is huge and $\hat{y}_\tau^2/\hat{y}_{\tau-1}^2$ grows very rapidly on average. Explosive growth of the form

$$\frac{\hat{y}_\tau^2}{\hat{y}_{\tau-1}^2} \rightarrow \infty \tag{35}$$

turns out to be important below. Such growth is faster than exponential growth, which entails $\hat{y}_\tau^2/\hat{y}_{\tau-1}^2 \rightarrow$ constant >1.

Suppose, for simplicity, that $\rho_1 = 0$, $\rho_3 = 1$ are known and the only task is to estimate ρ_2. Consider MSD estimation of ρ_2 by using ARCH(1) (Engle, 1982) as the score generator for EMM,

$$f(y_t|y_{t-1}, \theta_1, \theta_2) = \frac{1}{(2\pi\sigma_t^2)^{1/2}} \exp\left(-\frac{1}{2\sigma_t^2} y_t^2\right)$$

where

$$\sigma_t^2 = \theta_1 + \theta_2 y_{t-1}^2.$$

The deviation function is

$$\eta(y_{t-1}, y_t, \theta_1, \theta_2) = \frac{\partial}{\partial\theta} \log[f(y_t|y_{t-1}, \theta_1, \theta_2)].$$

Also for simplicity, we concentrate on the behavior of the second component of the score function $\eta_2(y_{t-1}, y_t, \theta_1, \theta_2) = (\partial/\partial\theta_2)\log[f(y_t|y_{t-1}, \theta_1, \theta_2)]$ near the unstable region for ρ_2. Let $\tilde{\theta}_1$ and $\tilde{\theta}_2$ denote the QML estimates obtained by fitting the ARCH(1) model to the observed realization $\{\tilde{y}_t\}_{t=1}^n$. Let $\{\hat{y}_\tau\}_{\tau=1}^N$ denote a simulation given a candidate value of $\rho = (0 \quad \rho_2 \quad 1)$. Basic calculations give

$$\begin{aligned}\frac{1}{N}\sum_{\tau=2}^{N} \eta_2(\hat{y}_{\tau-1}, \hat{y}_\tau, \tilde{\theta}_1, \tilde{\theta}_2) \\ = -\frac{1}{2N}\sum_{\tau=2}^{N}\left(\frac{\hat{y}_{\tau-1}^2}{\hat{\sigma}_\tau^2}\right)\left(1 - \frac{\hat{y}_\tau^2}{\hat{y}_{\tau-1}^2}\frac{\hat{y}_{\tau-1}^2}{\hat{\sigma}_\tau^2}\right)\end{aligned} \tag{36}$$

where $\hat{\sigma}_\tau^2 = \tilde{\theta}_1 + \tilde{\theta}_2 \hat{y}_{\tau-1}^2$. Some care is needed in assessing this deviation function, which is not a polynomial. Along an explosive trajectory where $|\hat{y}_\tau| \rightarrow \infty$, the simulated variance $\hat{\sigma}_\tau^2 = \tilde{\theta}_1 + \tilde{\theta}^2\hat{y}_{\tau-1}^2$ diverges as well, and in fact

$$\frac{\hat{y}_{\tau-1}^2}{\hat{\sigma}_\tau^2} = \frac{1}{(\tilde{\theta}_1/\hat{y}_{\tau-1}^2) + \tilde{\theta}_2} \rightarrow \frac{1}{\tilde{\theta}_2}$$

on an explosive trajectory. Thus, in equation (36) the terms involving $\hat{y}_{\tau-1}^2/\hat{\sigma}_\tau^2$ converge to $1/\tilde{\theta}_2$, so that equation (36) becomes

$$\begin{aligned}\frac{1}{N}\sum_{\tau=2}^{N} \eta_2(\hat{y}_{t-1}, \hat{y}_t, \tilde{\theta}_1, \tilde{\theta}_2) \approx -\frac{1}{2N} \\ \times \sum_{\tau=2}^{N}\left(\frac{1}{\theta_2}\right)\left(1 - \frac{\hat{y}_\tau^2}{\hat{y}_{\tau-1}^2}\frac{1}{\theta_2}\right).\end{aligned} \tag{37}$$

However, from equation (35) $\hat{y}_\tau^2/\hat{y}_{\tau-1}^2$ can be expected to grow explosively along extended segments of the realization, which in turn can be expected to drive the right-hand side to infinity, and thereby penalize very heavily values of ρ_2 in the unstable region.

The preceding suggests that the objective function for MSD estimation of stochastic volatility models should also diverge on the unstable region of the parameter space and thereby enforce stability. Table 1 illustrates this numerically. The table shows the values of the concentrated objective function

$$Q_N^*(\rho_2) = \min_{\rho_1,\rho_3} Q_N(\rho_1, \rho_2, \rho_3) \tag{38}$$

when the stochastic volatility model (33)–(34) is estimated via EMM on daily S&P stock returns, 1977–1992. The auxiliary model is an ARCH-type model with 11 lags in the volatility and a modified Hermite polynomial of degree 4 for the error density. (This auxiliary model is one of the sample seminonparametric models estimated in Gallant and Tauchen (1996c) as illustration.) The objective function in table 1 has

TABLE 1.—CONCENTRATED OBJECTIVE FUNCTION OF STOCHASTIC VOLATILITY MODEL

ρ_2	$Q_N(\rho_2)$
0.900	50.036
0.910	47.325
0.920	44.135
0.930	40.499
0.940	36.557
0.950	32.599
0.960	29.069
0.970	26.490
0.980	25.282
0.990	25.461
1.000	2.8×10^3
1.005	6.4×10^{17}

Note: ρ_2 is the autoregressive parameter of the volatility autoregression (34) and Q_N is the concentrated objective function with ρ_1 and ρ_3 concentrated out.

a minimum very close to unity ($\hat{\rho} = 0.9833$), and, much like the top panel of figure 1, it is seen to be extremely asymmetric due to the heavy penalty placed on values of ρ_2 in the unstable region.

IV. Conclusion

There are two models involved in estimation by dynamic simulation. One is the maintained model, which is used to generate simulated data and whose parameters are of intrinsic economic interest. The other is the auxiliary model, which is used to define the moment conditions. Formal use of an auxiliary model to define moment conditions is introduced by Smith (1990, 1993) and followed upon by Gourieroux et al. (1993) for indirect inference and by Bansal et al. (1993, 1995) for EMM. Although dynamic SMM estimation in the style of Ingram and Lee (1991) and Duffie and Singleton (1993) does not directly require an auxiliary model, researchers typically have one in mind when formulating the moment conditions.

This paper considers the role of dynamic stability conditions for each of these two models. The focus of the investigation is on a class of estimators termed minimum simulated deviation (MSD), which encompasses SMM and EMM, but not the more computationally demanding indirect inference estimator. The main finding is that the stability conditions should be enforced on the auxiliary model. If that is done, then the stability conditions can be ignored for estimation of the maintained model. Instability can be expected to be heavily penalized by the objective function itself. If stability is not enforced on the auxiliary model, and if the auxiliary model should happen to contain embedded unstable dynamics, then the objective function of MSD can be expected to be ill-behaved as the simulations attempt to replicate the unstable dynamics. A more limited analysis of indirect inference suggests that somewhat the same conclusions hold, though the penalty on the unstable region of the parameter space is not as severe as that of MSD.

The issue then becomes that of actually enforcing stability on the estimated auxiliary model. Given a specification $f(y_t | y_{t-L}, \ldots, y_{t-1}, \theta)$ for the conditional density of the data, there is no assurance that the unrestricted QML estimator $\tilde{\theta}$ will land in the stable region. Current practice is not to impose stability directly on the optimization for estimation of the auxiliary model, but rather to check via simulation that the fit is stable. An example is Andersen and Lund (1997), who use semiparametric GARCH models as auxiliary models for EMM estimation of continuous-time stochastic volatility models of the short interest rate. They estimate a family of semiparametric GARCH and E-GARCH (exponential GARCH) models and then carefully check stability. They find that fitted semiparametric GARCH models tend to be unstable, and thus unusable, while fitted semiparametric E-GARCH models are stable. A similar strategy is adopted in the applications of Tauchen (1997) and Gallant and Tauchen (1996b, 1997).

The difficulty in enforcing stability on an auxiliary model $f(y_t | y_{t-L}, \ldots, y_{t-1}, \theta)$ appears not to be conceptual but rather computational. Gonzalez-Farias and Dickey (1994) note that for a Gaussian AR(1), if one includes the contribution to the likelihood of the initial datum y_1 and thereby does full maximum likelihood instead of conditional maximum likelihood, then the maximum-likelihood estimator cannot land outside the unit circle. The Gonzalez-Farias and Dickey idea is likely to generalize. Inclusion of the contribution of the unconditional loglikelihood of the initial observations, that is, $\log [f(y_1, \ldots, y_L, \theta)]$, where $f(y_1, \ldots, y_L | \theta)$ is the implied joint density of the initial L observations, should enforce stability. The difficulty is that for a general, flexible, conditional density $f(y_t | y_{t-L}, \ldots, y_{t-1}, \theta)$, it is infeasible to deduce a tractable closed form for the unconditional density $f(y_{t-L}, \ldots, y_t | \theta)$. As a work-around, Gallant and Tauchen (1996c) implement flexible conditional densities of the form $f[y_t | \mathcal{T}(y_{t-L}), \ldots, \mathcal{T}(y_{t-1}), \theta]$, where the nonlinear transformation $\mathcal{T}(\cdot)$ on the conditioning variables imposes nonlinear mean reversion on the conditional density and thereby rules out explosions a priori. Some ongoing research suggests that this strategy is potentially effective for enforcing stability on the auxiliary model.

REFERENCES

Abadir, Karim, "On the Asymptotic Power of Unit Root Tests," *Econometric Theory* 9 (1993), 189–221.

Andersen, Torben G., and Jesper Lund, "Estimating Continuous-Time Stochastic Volatility Models of the Short Term Interest Rate," *Journal of Econometrics* 77 (1997), 343–378.

Andersen, Torben G., and Bent Sorensen, "GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study," *Journal of Business and Economic Statistics* 14 (1996), 328–352.

Andrews, Donald W. K., "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica* 59 (1991), 307–346.

Bansal, Ravi, A. Ronald Gallant, Robert Hussey, and George Tauchen, "Computational Aspects of Nonparametric Simulation Estimation," in David A. Belsley (ed.), *Computational Techniques for Econometrics and Economic Analysis* (Boston: Kluwer, 1993).

——— "Nonparametric Estimation of Structural Models for High-Frequency Currency Market Data," *Journal of Econometrics* 66 (1995), 251–287.

Baumol, William J., *Economic Dynamics,* 3rd ed. (Toronto, Ont: Macmillan, 1970).

Duffie, Darrell, and Kenneth J. Singleton, "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica* 61 (1993), 929–952.

Engle, Robert F., "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica* 50 (1982), 987–1007.

Foster, Douglas, and S. Viswanathan, "Can Speculative Trading Explain the Volume–Volatility Relation?" *Journal of Business and Economic Statistics* 13 (1995), 379–398.

Gallant, A. Ronald, David A. Hsieh, and George Tauchen, "Estimation of Stochastic Volatility Models with Diagnostics," *Journal of Econometrics* 81 (1997), 159–192.

Gallant, A. Ronald, and George Tauchen, "Which Moments to Match?" *Econometric Theory* 12 (1996a), 657–681.

——— "Specification Analysis of Continuous Time Models in Finance," in P. E. Rossi (ed.), *Modeling Stock Market Volatility: Bridging the Gap to Continuous Time* (New York: Academic Press, 1996b).

——— ‘‘SNP: A Program for Nonparametric Time Series Analysis, a User’s Guide,’’ manuscript, University of North Carolina at Chapel Hill (1996c). (Available along with code via anonymous ftp at ftp.econ.duke.edu/home/arg/snp.)

——— ‘‘Estimation of Continuous Time Models for Stock Returns and Interest Rates,’’ *Macroeconomic Dynamics* 1 (1997), 135–168.

Ghysels, Eric, Andrew Harvey, and Eric Renault, ‘‘Stochastic Volatility,’’ in G. S. Maddala (ed.), *Handbook of Statistics,* vol. 14: *Statistical Methods in Finance* (Amsterdam: North-Holland, 1995).

Gonzalez-Farias, Gloria, and David A. Dickey, ‘‘An Unconditional Maximum Likelihood Test for a Unit Root,’’ manuscript, North Carolina State University (1994).

Gourieroux, C., and A. Monfort, *Simulation-Based Econometric Methods* (New York: Oxford University Press, 1996).

Gourieroux, C., A. Monfort, and E. Renault, ‘‘Indirect Inference,’’ *Journal of Applied Econometrics* 8 (1993), S85–S118.

Ingram, Beth F., and B. S. Lee, ‘‘Simulation Estimation of Time Series Models,’’ *Journal of Econometrics* 47 (1991), 197–205.

Laroque, Guy, and Bernard Selanie, ‘‘Estimation of Multi-Market Fix-Price Models: An Application of Pseudo Maximum Likelihood Models,’’ *Econometrica* 57 (1989), 831–860.

Monahan, John, ‘‘A Note on Enforcing Stationarity in Autoregressive-Moving Average Models,’’ *Biometrika* 71 (1984), 403–404.

Newey, Whitney K., and Kenneth D. West, ‘‘A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator,’’ *Econometrica* 55 (1987), 703–708.

Smith, Anthony A., ‘‘Three Essays on the Solution and Estimation of Dynamic Macroeconomic Models,’’ Ph.D. Dissertation, Duke University (1990).

——— ‘‘Estimating Nonlinear Time Series Models Using Simulated Vector Autoregressions,’’ *Journal of Applied Econometrics* S8 (1993), S63–S84.

Tauchen, George, ‘‘New Minimum Chi-Square Methods in Empirical Finance,’’ in D. Kreps and K. Wallis (eds.), *Advances in Econometrics, 7th World Congress* (Cambridge, UK: Cambridge University Press, 1997).