stants follows immediately from the expression for the moment-generating function.

Therefore we have an expression for $f(\mathbf{s})$ as

$$f(\mathbf{s}) = \sum_{i=0}^{\infty} \beta^{i+1} \exp(\mathbf{a}_i'\mathbf{s} + c_i),$$

where

$$\begin{aligned}\mathbf{a}_{i+1} &= \mathbf{F}'\mathbf{a}_i + \mathbf{F}'\mathbf{a} \\ &= \mathbf{F}'(1 - \mathbf{F}')^{-1}\mathbf{a} - \mathbf{F}'^{i+2}(1 - \mathbf{F}')^{-1}\mathbf{a}\end{aligned}$$

and

$$c_{i+1} = c_i + (\mathbf{a}_i + \mathbf{a})'\mathbf{b} + (\mathbf{a}_i + \mathbf{a})'\mathbf{\Sigma}(\mathbf{a}_i + \mathbf{a})/2.$$

For large i,

$$\mathbf{a}_i = \boldsymbol{\lambda} + o(1/i)$$

and

$$c_i = (i + 1)\delta + \psi + o(1/i),$$

where

$$\boldsymbol{\lambda} = \mathbf{F}'(1 - \mathbf{F}')^{-1}\mathbf{a},$$

$$\delta = \mathbf{a}'(1 - \mathbf{F})^{-1}(\mathbf{b} + \mathbf{\Sigma}(1 - \mathbf{F}')^{-1}\mathbf{a}/2),$$

and

$$\begin{aligned}\psi = &-\mathbf{a}'\mathbf{F}(1 - \mathbf{F})^{-2}(\mathbf{b} + \mathbf{\Sigma}(1 - \mathbf{F}')^{-1}\mathbf{a}/2) \\ &- \mathbf{a}'(1 - \mathbf{F})^{-1}\mathbf{\Sigma}\mathbf{F}'(1 - \mathbf{F}')^{-2}\mathbf{a}/2 \\ &+ \sum_{i=1}^{\infty} \mathbf{a}'(1 - \mathbf{F})^{-1}\mathbf{F}^{i}\mathbf{\Sigma}\mathbf{F}'^{i}(1 - \mathbf{F}')^{-1}\mathbf{a}/2.\end{aligned}$$

The last sum in the equation for ψ can be related to the stationary variance of the AR(1) process. It is worth noting that all autoregressive moving average (ARMA) processes can be written as vector AR(1) processes, although with a possibly higher-dimensional state vector and a possibly singular variance matrix.

For the series to converge it is necessary that $|\beta \exp(\delta)| < 1$. Even though this condition is satisfied for the parameter values used in the simulations, however, convergence is very slow since $\beta \exp(\delta)$ is nearly 1. Convergence can be improved by adding and subtracting terms from the geometric series that for large i, approximate the series for $f(\mathbf{s})$. The approximating terms are $\beta^{i+1} \exp((i + 1)\delta + \boldsymbol{\lambda}'\mathbf{s} + \psi)$. Hence a usable form for $f(\mathbf{s})$ is

$$\begin{aligned}f(\mathbf{s}) = &\exp(\boldsymbol{\lambda}'\mathbf{s} + \psi)\frac{\beta \exp(\delta)}{1 - \beta \exp(\delta)} \\ &+ \sum_{i=0}^{\infty} \beta^{i+1}(\exp(\mathbf{a}_i'\mathbf{s} + c_i) \\ &- \exp((i + 1)\delta + \boldsymbol{\lambda}'\mathbf{s} + \psi)).\end{aligned}$$

An alternative method of doing the simulations consists of simulation of the AR(1) process directly together with a numerical solution of the pricing equation to obtain v_{it}. Written in terms of the $\log(w, x)'$ coordinates, a Gauss–Hermite quadrature seems natural. For some of the sets of parameter values used in the simulations, $12 \otimes 12$ point quadrature was used to numerically solve the preceding equation. Agreement of the numerical solution with the analytic solution was extremely good. The method used by Tauchen to solve the pricing equation for his simulations was similar to this numerical method, but the spacing of the grid points was not chosen to be those of the Gauss–Hermite quadrature rule.

Because a numerical solution of the previously mentioned type soon becomes infeasible as the dimension of the stochastic process increases, a simulation investigation of more complicated processes rapidly becomes very difficult. The infinite series expansion remains tractable for higher-dimensional processes, however, and may be useful when larger systems are studied.

Comment

Lars Peter Hansen
Department of Economics, University of Chicago, Chicago, IL 60637

1. INTRODUCTION

Tauchen's article is a very nice study of the small sample properties of generalized method of moments (GMM) estimators for a dynamic stochastic nonlinear econometric model. The parameter estimates for such a model must be calculated using numerical methods and the small sample properties deduced via Monte Carlo methods. Although the results reported in this article appear to be quite limited in scope, they are important because they are the first set of results about the small sample properties of GMM estimators for dynamic nonlinear models. Prior to this study, the only set of results available relied exclusively on asymptotic distribution theory.

My comments are divided into three segments. First, I make some general remarks about the GMM methodology for models such as the one studied here. Next, I make some specific comments about some of the results reported in Tauchen's article. Finally, I comment on approximation issues and nonparametric maximum likelihood estimators.

2. GMM ESTIMATION AND INFERENCE

Nonlinear stochastic dynamic models have proved to be formidable to solve analytically or even numerically. This led many researchers to use limited information methods for estimating preference parameters and testing overidentifying moment restrictions. The limited information estimators used have instrumental variables estimators as antecedents and can be viewed as special cases of GMM estimators. When applied to the study of nonlinear models, such estimators are tractable to implement in practice as long as the number of underlying parameters of interest is small. The estimation method has the advantage of not requiring an explicit solution for the model or requiring a specification of a parametric law of motion for the stochastic forcing variables in the economic model. On the other hand, for these methods to be applicable, some rather severe restrictions on how unobservable (to the econometrician) forcing variables enter the model must be satisfied (see Garber and King 1984; Hansen and Singleton 1982).

The GMM estimation method applied in this context does not require an explicit solution for the law of motion for the state vector (containing both endogenous and forcing variables), but this advantage is not achieved without cost. Validation of the model via predictive performance or comparisons with less restricted estimated laws of motion is not a by-product of this estimation method. Furthermore, the parameter estimates may not be particularly useful without simultaneously considering the impact of these parameters on hypothetical modifications in the underlying economic environment. The only way of assessing this impact is to calculate implied laws of motion for the endogenous state variables.

In spite of these limitations, the GMM estimation method as applied here may still be a useful device for discrimination among members of a class of dynamic nonlinear econometric models. Even though the implied law of motion for the state variables is not analyzed explicitly, unconditional moment restrictions are tested. With this perspective, the focus of the statistical analysis shifts away from estimation and toward testing the overidentifying restrictions. An extreme version of this view would treat the underlying parameters as nuisance parameters that must be estimated to test the implied unconditional moment restrictions.

Pursuing this line further, consider the representative consumer model estimated by Singleton and me (Hansen and Singleton 1982) and studied by Tauchen. One formal rationalization for the representative consumer model allows for heterogeneity in preferences and endowments among individual consumers by assuming the existence of a complete set of security markets. In this case, the implied preferences for the representative consumer are mongrel preferences, because they depend on the initial distribution of endowments across consumers. This approach to aggregation limits the interest in the estimated preference parameters, because these parameters will not remain invariant with respect to hypothetical changes in the underlying economic environment that alter wealth distributions. Tests of the moment restrictions remain valid even though the preference parameters cease to have a structural interpretation. With this perspective, the small sample evidence reported by Tauchen on the behavior of the tests of the overidentifying restrictions is much more interesting than the evidence reported on the behavior of the parameters estimator.

Viewing the preference parameters as nuisance parameters is probably too extreme for two reasons. First, if the mongrel aggregation approach is used, then a more nonparametric approach to estimating the preference parameters would seem to be appropriate. Second, there is an alternative approach to aggregation that also rationalizes a representative consumer. This form of aggregation, suggested by Gorman (1953), restricts substantially the heterogeneity of preferences across consumers but only places mild restrictions on the heterogeneity of endowments (e.g., see Eichenbaum, Hansen, and Richard 1985). Under Gorman aggregation, the preference parameters of the representative consumer no longer depend on the endowment distribution. Thus the estimated parameters may still be useful in simulation experiments. These simulation exercises, although requiring a solution of the dynamic model, do not necessarily require evaluation of this solution for a large number of hypothetical parameter values as would be required for maximum likelihood estimation. The parameterization of preferences adopted by Singleton and me (Hansen and Singleton 1982) and by Tauchen are consistent with Gorman aggregation.

3. GMM ESTIMATORS THAT HAVE THE SMALLEST ASYMPTOTIC COVARIANCE MATRICES

There is an infinite dimensional class of GMM estimators that can be used in estimating the parameters of models such as the one Tauchen studies. In prior work, I derived greatest lower bounds on the asymptotic covariance matrices of this class of estimators (see Hansen 1985a,b). These calculations can be viewed as time series counterparts to calculations performed by Amemiya (1977), Chamberlain (1983), and Jorgenson and Laffont (1974). Tauchen uses one of these bounds in his analysis and studies the behavior of the idealized GMM estimator that attains this bound. One of the puzzling aspects of his results is that the idealized estimator performs poorly in small samples vis-à-vis the other GMM estimators that he constructs using ad hoc choices of instrumental variables.

To better understand the nature of this finding, it would be helpful if some subsequent research were pursued. Except for the idealized estimator, all of the GMM estimators considered by Tauchen solve minimization problems with a criterion function that satisfies the minimum chi-squared property. For instance, when these criterion functions are evaluated at the true parameter value, they are distributed asymptotically as chi squares with degrees of freedom equal to the number of orthogonality conditions. When evaluated at the parameter estimators, these criterion functions are distributed asymptotically as chi squares with degrees of freedom equal to the number of overidentifying restrictions. The criterion function for the idealized estimator also could have been chosen to have this chi-squared property; however, a different choice of criterion function was used by Tauchen. I would find it interesting to see comparable results for the idealized estimator obtained by minimizing a criterion function that possesses the chi-squared property just described. By modifying the criterion function in this way, one could use the asymptotic distribution theory to help assess the behavior of the criterion function evaluated at hypothetically true values of the parameter. To put the idealized estimator on a more equal footing with the other estimators, it would also be of interest to include some additional orthogonality conditions so that the idealized criterion function embedded overidentifying restrictions. These suggestions may not resolve Tauchen's puzzling findings, but they could provide some additional insights.

A related set of results that I would find interesting is a systematic report of the behavior of the criterion functions evaluated at the true parameter value. In addition, the behavior of idealized objective functions that use the inverse of the true asymptotic covariance matrix of sample counterparts to the population orthogonality conditions could be simulated. These idealized objective functions could be evaluated at the true parameter vector. Such calculations would help in understanding divergence between the asymptotic results and the small sample results. For instance, in cases in which the divergence is nontrivial, one could assess whether the problem is the failure of the central limit approximation, the use of an estimator of the true asymptotic covariance matrix, or the nonquadratic character of the criterion function used for estimation.

4. APPROXIMATION VIA MARKOV CHAINS AND ASYMPTOTIC EFFICIENCY

Tauchen shows how to approximate stationary lognormal processes by Markov chains. Solving for the equilibrium return process implied by the economic model is considerably easier when the joint consumption growth and dividend growth rates are components of a stationary Markov chain than when these processes are lognormally distributed. The computational advantages to solving Markov chain models carries over to a variety of other nonlinear rational expectations models (e.g., see Lucas and Stokey 1985). An alternative approach is to adopt a convenient joint process for consumption growth rates and equilibrium returns and to work backward and construct a corresponding dividend process. This second approach was adopted, at least implicitly, by Hansen and Singleton (1983). A more general approach along this same vein was suggested by Sims (1985). Sims's approach was to impose a convenient time series law of motion for some subset of endogenous variables and/or multipliers for constraints and to solve backward to deduce the implied properties of the unobserved (to the econometrician) forcing variables. This alternative approach makes model solution and hence estimation considerably easier at the cost of sacrificing direct control over the behavior of the exogenous forcing processes.

The Markov chain approximation, however, is of interest for two other reasons. Chamberlain (1983) showed that for models with independent and identically distributed observation vectors, GMM estimators can approximate arbitrarily well a nonparametric maximum likelihood bound on the asymptotic covariance matrix. Chamberlain's result provides a nice link between GMM and maximum likelihood estimation. Chamberlain's strategy for proving this result is to approximate members of a general class of distributions that satisfy conditional moment restrictions by a corresponding class of multinominal distributions that satisfy the same conditional moment restrictions. Markov chain processes can be viewed as the time series counterpart to the multinominal distributed random vectors used by Chamberlain for approximation. It would be of interest to investigate the extent to which Chamberlain's conclusions carry over to GMM estimation of time series models. An approach to this problem might entail extending the approximation from parametric families of processes (e.g., lognormal processes) to more general Markov processes.

The following is a second reason. I have already argued that a disadvantage of GMM estimation as applied here is that it does not result in an estimated time series law of motion, whereas a nonparametric maximum likelihood approach would result in an approximate law of motion. Rather than concentrating future research on bootstrapping GMM estimators, it might be useful to also investigate nonparametric maximum likelihood estimation of nonlinear time series models subject to conditional moment restrictions.

ADDITIONAL REFERENCES

Amemiya, T. (1977), "The Maximum Likelihood and the Nonlinear Three-Stage Least Squares Estimator in the General Nonlinear Simultaneous Equation Model," *Econometrica,* 45, 955–968.

Eichenbaum, M. S., Hansen, L. P., and Richard, S. (1985), "The Dynamic Equilibrium Pricing of Durable Consumption Goods," unpublished manuscript.

Garber, P. M., and King, R. G. (1984), "Deep Structural Excavation? Identification Problems in Euler Equation Methods," unpublished manuscript.

Gorman, W. M. (1953), "Community Preference Fields," *Econometrica,* 21, 63–80.

Hansen, L. P. (1985b), "Using Martingale Difference Approximations to Obtain Covariance Matrix Bounds for Generalized Methods of Moments Estimators," National Opinion Research Center Discussion Paper 85-16, University of Chicago, Department of Economics.

Jorgenson, D. W., and Laffont, J. (1974), "Efficient Estimation of

Non-linear Simultaneous Equations With Additive Disturbances," *Annals of Economic and Social Measurement,* 3, 615–640.
Lucas, R. E., Jr., and Stokey, N. L. (1985), "Money and Interest in a Cash-in-Advance Economy," Center for Mathematical Studies in Economics and Management Science Discussion Paper 628, Northwestern University, Department of Managerial Economics and Decision Sciences.
Sims, C. (1985), "Solving Nonlinear Stochastic Equilibrium Models 'Backwards'," Center for Economic Research Discussion Paper 206, University of Minnesota, Dept. of Economics.

Comment

Peter E. Rossi
Graduate School of Business, University of Chicago, Chicago, IL 60637

The key contribution of Tauchen's article is a method of simulating equilibrium asset prices in a simple exchange economy. The equilibrium distribution of asset prices can be determined by solving an integral functional equation. Tauchen uses a discrete approximation method to approximate solutions to the integral equation. His simulation techniques can be applied both to the study of the properties of asset pricing models and to the analysis of the sampling distribution of various proposed estimators of model parameters.

In particular, Tauchen's simulation procedure makes possible a Monte Carlo analysis of the finite sample properties of Hansen and Singleton's (1982) generalized method of moments (GMM) estimators. The GMM procedure has gained popularity because of the relative ease of implementation [as contrasted with maximum likelihood estimators (MLE's)] and because the GMM procedure does not require specification of the exact probability model for the possibly unobserved exogenous or "driving" processes. Although Hansen (1982, 1985a) provided asymptotic approximations to the distribution of the GMM estimators, Tauchen is the first to provide evidence on their small sample properties. Furthermore, his procedure for simulating artificial economies appears to have wide applicability to other dynamic competitive equilibrium problems that are much more complicated than the simple consumption-based asset pricing models.

In general, the Monte Carlo results have the same flavor as the exact finite sample results available for instrumental variables methods in the simultaneous equations model [see, e.g., Anderson and Sawa (1979) for the case of two-stage least squares]. The estimator of the risk aversion parameter exhibits substantial small sample bias, and there appears to be a bias-variance trade-off. As the number of lagged values of dividend and consumption growth instruments increases, the bias increases with a more than commensurate decrease in variance so that mean squared error (MSE) and mean absolute deviation (MAD) decline. Perhaps the most striking result is that the test of the overidentifying restrictions has a rejection rate less than the nominal significance level. This suggests that the resounding rejection of these simple time-additive models based on the asymptotic distribution of these test statistics is not due to the inadequacy of the asymptotic approximation.

As Tauchen points out, the use of discrete approximations to solve functional equations has a long history in the physical science literature. Mehra and Prescott (1985) were the first to use a discrete approximation approach to these dynamic competitive equilibria in economics. They assumed that the consumption and dividend processes are discrete and solved the asset pricing problem, whereas Tauchen regards the discrete process as an approximation to an underlying continuous process. It is not necessary to bring these discrete methods to bear on the simple asset pricing model, however.

A direct attack on the problem of simulating equilibrium prices exploits the fact that the first-order conditions in this model imply that, at each time period, asset prices can be expressed as the expected discounted dividend stream.

$$P_t = E_t\left[\sum_{i=0}^{\infty} d_{t+i}\beta^i(u'(c_{t+i+1})/u'(c_{t+i}))\right] \tag{1}$$

can easily be evaluated using Monte Carlo integration. Merely draw sample paths from the process that generates $\{(c_t, d_t)\}$. The computations involved in producing a realization from a simple linear time series model are trivial, even with large numbers of assets. One suspects very fast convergence could be obtained in computing these integrals because of the discount factor and the usual martingale convergence results. Thus, it might be possible to truncate the sum at a small number of terms.

Another advantage of this approach is the built-in check derived by taking the expectation of (1) with respect to the stationary distributions:

$$E[P_t] = \sum_{i=0}^{\infty} E\left[d_{t+i}\beta^i(u'(c_{t+i+1})/u'(c_{t+i}))\right]. \tag{2}$$

We can check the accuracy of the simulated price series by comparing the sample mean of $\{P_t\}$ with $E[P_t]$. In addition, error bounds are available and easily calculated for Monte Carlo integration. These bounds are independent of the dimension of the integral. Monte