Journal of Econometrics 47 (1991) 197–205. North-Holland

Simulation estimation of time-series models*

Bong-Soo Lee
University of Minnesota, Minneapolis, MN 55455, USA

Beth Fisher Ingram
University of Iowa, Iowa City, IA 52242, USA

Received January 1986, final version received July 1989

A formal econometric treatment of the estimation of the parameters of a fully specified stochastic equilibrium model is proposed. The method, estimation by simulation, yields an estimator which is shown to have an asymptotic normal distribution. A goodness-of-fit test based on a chi-square statistic is also derived.

1. Introduction

In recent years, economists have tried to explain various empirical regularities by solving stochastic equilibrium models which explicitly use the assumption that economic agents dynamically optimize under uncertainty. In general, closed-form expressions for the endogenous variables in these models cannot be found.[1] Recently, several methods have been suggested for finding simulated solution paths for stochastic, dynamic general equilibrium models.[2] Using any of these methods, the simulated solution paths which the researcher derives will depend on a set of model parameters to which values

*The authors are grateful to Christopher Sims, Yasuo Amemiya, Wayne Fuller, Charles Whiteman, two anonymous referees, and an Associate Editor for helpful suggestions and comments.

[1]The exception is when the problem is formulated in such a way that the economic agents are assumed to maximize a quadratic objective function subject to a linear constraint set. See Hansen and Sargent (1980).

[2]Fair and Taylor (1983), Labadie (1984), Marcet (1988), Sims (1985), Tauchen (1986). Taylor (1988) summarizes these methods.

must be assigned before simulations can be generated. Ideally, this would be accomplished through estimation of the parameters using observed data. In many instances, the estimation procedure has involved a naive grid search on a parameter space which is restricted on either theoretical or empirical grounds. In fact, several authors have cited the need for a formal econometric treatment of this issue [Kydland and Prescott (1982), Labadie (1984)]. To recapitulate, we would like the model we develop to mimic observed empirical regularities quantitatively, and we want to treat goodness of fit of the model in a more rigorous, systematic fashion.

The purpose of this paper is to propose an econometric estimation strategy which uses the complete representation of a stochastic equilibrium model and allows us to test whether the underlying theory is quantitatively consistent with regularities observed in some data set. We give conditions under which the simulation estimator is asymptotically normal, and derive a goodness-of-fit test based on the chi-square distribution.

Recently, McFadden (1986) and Pakes and Pollard (1986) have proposed similar simulation estimators, mainly for use in the discrete-response problem. These authors do not allow serial correlation in the data set; the simulation estimator proposed here implicitly permits the disturbance terms to be serially correlated, a condition which is likely to be relevant for many economic applications, especially when the researcher is using time-series data. The criterion functions in the McFadden and Pakes–Pollard papers are, however, more general in that each allows for a certain degree of discontinuity. We require that the derivative of the function be continuous in the mean (or supercontinuous) for the proof of the asymptotic normality of the estimator. This assumption is stronger than the assumptions made in either McFadden or Pakes–Pollard. In certain situations, we also require knowledge of the ergodic distribution of the variables being simulated.

It is well known that the generalized method-of-moments procedure proposed by Hansen (1982) can be used to estimate nonlinear rational-expectations models without explicitly solving for the full stochastic equilibrium [see Hansen and Singleton (1982)]. A typical procedure has been to make use of the moment conditions implied by the stochastic Euler equations derived from an optimization problem. Like the Euler equation approach, the estimation procedure advanced here can be formulated as a special case of the generic GMM estimator proposed by Hansen. The simulation estimation methodology, however, can be used on some models in which the stochastic Euler equation approach fails.

The rest of the paper is organized as follows. Section 2 introduces estimation by simulation, discusses the consistency and the asymptotic normality of the estimator and outlines a chi-square test. Section 3 concludes the paper.

2. Simulation estimation

Suppose we have a fully specified, stochastic general equilibrium model which generates an $m \times 1$ vector stochastic process, $\{y_j(\beta),\ j \geq 1\}$, denoted $\{y_j(\beta)\}$. The $l \times 1$ parameter vector β contains the underlying parameters of the model defining tastes, technology, etc., and may also contain parameters of auxiliary equations added to the model or parameters used in obtaining the simulated solution path. Under the null hypothesis that the model is a true description of some aspect of the economy when evaluated at β_0, the true β, $\{y_j(\beta_0)\}$ will have a counterpart in an observed data set, $\{x_t,\ t \geq 1\}$, denoted $\{x_t\}$.[3] That is, the simulated data $\{y_j(\beta_0)\}$ is assumed to be drawn from the same distribution as the observed data $\{x_t\}$, which is also an $m \times 1$ vector stochastic process. In practice, we would generate $\{y_j(\beta),\ j = 1, \ldots, N\}$, a finite realization of $\{y_j(\beta)\}$, and observe $\{x_t,\ t = 1, \ldots, T\}$, a finite realization of $\{x_t\}$.

Heuristically, the proposed simulation estimator of β is obtained by equating the sample counterparts of moments of the simulated process, $\{y_j(\beta)\}$, to sample counterparts of moments of the observed data process, $\{x_t\}$. Let

$$H_T(x) = \frac{1}{T} \sum_{t=1}^{T} h(x_t) \quad \text{and} \quad H_N(y(\beta)) = \frac{1}{N} \sum_{j=1}^{N} h(y_j(\beta)).$$

That is, $H_T(x)$ is an $s \times 1$ vector of statistics calculated as a time average of some function of the observed data, and $H_N(y(\beta))$ is a corresponding vector of statistics calculated from the economic model using simulated data.

Assuming the processes $\{x_t\}$ and $\{y_j(\beta)\}$ are ergodic,

$$H_T(x) \xrightarrow{\text{a.s.}} \mathrm{E}[h(x_t)] \quad \text{as} \quad T \to \infty,$$

and, for each β,

$$H_N(y(\beta)) \xrightarrow{\text{a.s.}} \mathrm{E}\big[h(y_j(\beta))\big] \quad \text{as} \quad N \to \infty.$$

Furthermore, under the null hypothesis that the economic model is correct at β_0,

$$\mathrm{E}[h(x_t)] = \mathrm{E}\big[h(y_j(\beta_0))\big].$$

[3]Note that we have not excluded the possibility that the general equilibrium model generates series which are unobservable in the real world. $\{y_j(\beta)\}$ is assumed to contain only observable series.

What is shown in this paper is that if an estimate of β, $\hat{\beta}_{TN}$, is chosen so that the weighted, squared difference between $H_N(y(\beta))$ and $H_T(x)$ is minimized, the resulting estimator is consistent and asymptotically normal.

The simulation estimator, $\hat{\beta}_{TN}$, can be defined as follows:

Definition. Given a random, symmetric, $s \times s$ weighting matrix W_T of rank at least l, the simulation estimator, $\hat{\beta}_{TN}$, is the solution to

$$\min_{\beta} \left[H_T(x) - H_N(y(\beta)) \right]' W_T \left[H_T(x) - H_N(y(\beta)) \right].$$

We will fix an integer $n = N/T > 1$ and define functions $g_T(\beta)$ and $f_t(\beta)$:[4]

$$g_T(\beta) = \frac{1}{T} \sum_{t=1}^{T} f_t(\beta) = \frac{1}{T} \sum_{t=1}^{T} \left[h(x_t) - \frac{1}{n} \sum_{k=1}^{n} h(y_{k,t}(\beta)) \right].$$

Here we have indexed simulations by the pair (k,t), $k = 1,\dots,n$ and $t = 1,\dots,T$. For example, $y_{k,t}(\beta) = y_{n(t-1)+k}(\beta)$. By assuming $n > 1$, we are requiring that the length of the simulated series be longer than the length of the observed data series, and that as $T \to \infty$, N also increases so that N/T stays fixed at n. As will be shown below, the larger that n is chosen, the smaller is the asymptotic covariance matrix for the estimator.[5]

The formulation of the estimator as a solution to a minimization problem makes it a special case of the generic estimator found in Hansen (1982). The consistency and asymptotic normality of our estimator can be established by verifying that the assumptions of Hansen's theorems hold here. The mapping of our framework into his regularity conditions warrants the following comments. In particular, consistency[6] of the estimator requires that $\{y_j(\beta), 1 \le j \le N\}$ and $\{x_t, 1 \le t \le T\}$ be independent of each other and that $\{y_j(\beta)\}$ and $\{x_t\}$ be stationary and ergodic.[7] Further, $\mathrm{E}[f_t(\beta)]$ must have a unique

[4]This notation corresponds to the notation used in Hansen (1982).

[5]Relaxing the assumption that n is an integer simply makes the notation more complicated. Allowing $n = N/T$ to be a real number greater than one, we redefine $f_t(\beta)$ such that

$$\begin{aligned} g_T(\beta) &= \frac{1}{T} \sum_{t=1}^{T} f_t(\beta) = \frac{1}{T} \sum_{t=1}^{T} h(x_t) - \frac{1}{Tn} \sum_{t=1}^{Tn} h(y_j(\beta)) \\ &= \frac{1}{T} \sum_{t=1}^{T} \left[h(x_t) - \frac{1}{n} \sum_{j=n_0}^{n_1} h(y_j(\beta)) \right], \end{aligned}$$

where $n_0 = [1 + (t-1)n]$, $n_1 = [nt]$, and $[m]$ represents the smallest integer less than or equal to m.

[6]Wooldridge and White (1985) give very general conditions under which an optimization estimator such as that examined here is consistent. However, their conditions are difficult to verify in practice.

[7]This assumption is verifiable in specific cases for $\{y_j(\beta)\}$ and, of course, must be assumed for $\{x_t\}$.

zero at β_0.[8] Also, $h(y_j(\beta))$ must be continuous in the mean. Formally, if S represents the parameter space, this condition requires

$$\lim_{\delta\to 0} \mathrm{E}\Big[\sup\Big\{\big|h(y_j(\beta_0)) - h(y_j(\beta_1))\big| : \beta_0, \beta_1 \in S, |\beta_0 - \beta_1| < \delta\Big\}\Big] = 0.$$

This assumption differentiates the analysis from that of Pakes–Pollard and McFadden who each allow for some degree of discontinuity in their criterion functions. Under the conditions outlined above, the simulation estimator $\hat{\beta}_{TN}$ exists and converges almost surely to β_0 as $T, N \to \infty$.

To derive the asymptotic distribution of $\hat{\beta}_{TN}$, some additional conditions must be verified. First, $\partial h(y_j(\beta))/\partial\beta$ must be continuous in the mean at β_0, and the matrix $B \equiv \mathrm{E}[\partial h(y_j(\beta))/\partial\beta]$ must exist, be finite, and have full rank. We require that the weighting matrix, W_T, used in the criterion function converge in probability to a matrix W. Hansen makes use of a central limit theorem applicable to martingale difference sequences, and this requires the following assumptions. Define $w_t = f_t(\beta_0)$ and

$$v_i = \mathrm{E}[w_t | w_{t-i}, w_{t-i-1}, \dots] - \mathrm{E}[w_t | w_{t-i-1}, w_{t-i-2}, \dots], \qquad i \geq 0.$$[9]

We assume $\mathrm{E}[w_t w_t']$ exists and is finite, $\mathrm{E}[w_t | w_{t-i}, w_{t-i-1}, \dots]$ converges in mean square to zero, and

$$\sum_{i=0}^{\infty} \mathrm{E}[v_i v_i']^{1/2} \quad \text{is finite.}$$

Define

$$R_x(i) = \mathrm{E}\{[h(x_t) - \mathrm{E}(h(x_t))][h(x_{t-i}) - \mathrm{E}(h(x_{t-i}))]'\},$$

$$R_y(i) = \mathrm{E}\{[h(y_t(\beta_0)) - \mathrm{E}\{h(y_t(\beta_0))\}] \times [h(y_{t-i}(\beta_0)) - \mathrm{E}\{h(y_{t-i}(\beta_0))\}]'\},$$

and

$$\Omega = \sum_{i=-\infty}^{\infty} R_x(i) \qquad \left(= \sum_{i=-\infty}^{\infty} R_y(i) \quad \text{under the null hypothesis}\right).$$

[8]This means that the model must be identified.

[9]$\{v_i: i \geq 0\}$ is a martingale difference sequence.

Given our assumptions,

$$\sqrt{T}\left[H_T(x) - \mathrm{E}(h(x_t))\right] \xrightarrow{\mathrm{D}} \mathrm{N}(0,\Omega),$$

$$\sqrt{N}\left[H_N(y(\beta_0)) - \mathrm{E}(h(y(\beta_0)))\right] \xrightarrow{\mathrm{D}} \mathrm{N}(0,\Omega).$$

Hence, given the independence of $\{x_t\}$ and $\{y_j(\beta_0)\}$,

$$\operatorname{cov}\{H_T(x) - H_N(y(\beta_0))\} = (1+1/n)\Omega.$$

Under the above conditions and the assumption that $N/T \to n$ as $T, N \to \infty$,

$$\sqrt{T}\left(\hat{\beta}_{TN} - \beta_0\right) \xrightarrow{\mathrm{D}} \mathrm{N}\left(0, (B'WB)^{-1}B'W(1+1/n)\Omega WB(B'WB)^{-1'}\right) \quad \text{as} \quad T, N \to \infty.$$

The asymptotic covariance matrix for the estimator depends on the choice of weighting matrix in the criterion function. Hansen shows that the optimal choice for W (the choice which yields the smallest asymptotic covariance matrix for the estimator) is $W = [(1+1/n)\Omega]^{-1}$. In that case,

$$\sqrt{T}\left(\hat{\beta}_{TN} - \beta_0\right) \xrightarrow{\mathrm{D}} \mathrm{N}\left(0, \left[B'(1+1/n)^{-1}\Omega^{-1}B\right]^{-1}\right) \quad \text{as} \quad T, N \to \infty.$$

To implement the strategy recommended here, we need a consistent estimator for Ω. One possibility is

$$\hat{\Omega} = \left[\sum_{i=-p+1}^{p-1} \frac{1}{T} \sum_{t=1+i}^{T} u_{t+p}u'_{t+p-i}\right],$$

where $u_{t+p} = h(x_{t+p}) - (1/T)\sum h(x_{t+p})$ and p is the number of population autocovariances determined by the order of nonzero autocorrelations of $h(x_t)$. It has been pointed out, however, that $\hat{\Omega}$ need not be positive definite in any finite sample when p is not zero. Following Newey and West (1987), a consistent, positive semi-definite estimate of Ω can be computed according to

$$\tilde{\Omega} = \frac{1}{T}\sum_{t=1}^{T} u_{t+p}u'_{t+p} + \sum_{i=1}^{p-1}\left[w(i,p-1)\left\{\frac{1}{T}\sum_{t=i+1}^{T} u_{t+p}u'_{t+p-i} + \frac{1}{T}\sum_{t=i+1}^{T} u_{t+p-i}u'_{t+p}\right\}\right],$$

where

$$w(i, p-1) = 1 - i/p.$$

An examination of the covariance matrix of the estimator displays the advantage of choosing N to be larger than T. The randomness in the estimator is derived from two sources: the randomness in the simulation and the randomness in the real data. As $n = N/T$ gets large, the importance of the randomness in the simulation to the covariance matrix of the estimator declines. That is, choosing N to be much larger than T allows for a reduction in the variance of the estimator.[10]

An underlying assumption of the derivation of the properties of $\hat{\beta}_{TN}$ is that we have generated a realization of a stationary stochastic process, $\{y_j(\beta)\}$, given values for the elements of the parameter vector β. In practice, this requires that the initial values we use, $y_0(\beta)$, be picked from the stationary distribution for the vector stochastic process $\{y_j(\beta)\}$. This distribution will depend on the unknown parameter vector β. To illustrate the issue, suppose:

$$y_j = \alpha y_{j-1} + e_j \quad \text{and} \quad e_j = \rho e_{j-1} + \omega_j,$$

where $\omega_j \sim \mathrm{NID}(0,1)$, $|\alpha| < 1$, and $|\rho| < 1$. Assume that y_j is a scalar. β is the vector $[\alpha \;\; \rho]$. To simulate the process $\{y_j(\beta)\}$, we need to choose values for β, y_0, and e_0. Given these values and a realization for $\{\omega_j\}$, we can calculate $\{e_j,\ j = 1, \ldots, N\}$ and $\{y_j(\beta),\ j = 1, \ldots, N\}$. If e_0 is chosen from the stationary distribution for $\{e_j\}$, then $\{e_j,\ j = 1, \ldots, N\}$ and, thus, $\{y_j(\beta),\ j = 1, \ldots, N\}$, will represent finite realizations drawn from their respective stationary stochastic processes. In our example, the unconditional distribution of e_j is $\mathrm{N}(0, 1/(1-\rho^2))$ so that e_0 must be drawn from the $\mathrm{N}(0, 1/(1-\rho^2))$ distribution.

As this example illustrates, in the linear case it is easy to calculate the unconditional distribution of e_0 (and, thus, y_0). However, if y_j depends on past y_{j-s} in a nonlinear fashion, it may be very difficult to ensure that y_0 has been drawn from the stationary distribution for $\{y_j(\beta)\}$. A practical suggestion in this circumstance would be to use an arbitrary y_0 and generate a realization from the stochastic process $\{y_j(\beta)\}$ of length $2N$, throw away the first N data points, and use only the remaining N data points. If N is large enough, the effect of the initial y_0 on the generated series should be negligible, and $\{y_j,\ j = N+1, \ldots, 2N\}$ would represent a realization from the stationary stochastic process $\{y_j(\beta)\}$.[11]

[10] Pakes–Pollard and McFadden make a similar point.

[11] It is the belief of an Associate Editor that the requirement that the stationary ergodic distribution of the variables being simulated be known could rule out many interesting dynamic models. WTF?

One final issue which needs to be mentioned is whether the random errors for the simulated series should be redrawn at every iteration in the minimization of the criterion function. In other words, the algorithm which is implemented to solve the minimization problem will search over various values for β, calculating $\{y_j(\beta)\}$ and, thus, $H_N(y(\beta))$ for each value of β. We assume that the random errors used to calculate $\{y_j(\beta)\}$ are held fixed throughout this routine. Hence, the estimator is the value of β which minimizes the criterion function for a given random draw. Allowing for new draws at every iteration in the minimization algorithm would violate our assumptions concerning continuity of the objective function.[12] More specifically, a new random draw means a new value for the criterion function even when the value of β is held fixed. The algorithm would have difficulty sorting out a change in the criterion function due to a variation in the random draw from a change due to a modification in β.

The estimation procedure outlined above sets l linear combinations of the s statistics $[H_T(x) - H_N(y(\beta))]$ equal to zero asymptotically by using the random weighting matrices $B'W_T$. The vector of statistics can be used to obtain a test statistic for the goodness of fit of the model under consideration since T times the minimized value of the criterion function:

$$[H_T(x) - H_N(y(\beta))]'W_T[H_T(x) - H_N(y(\beta))]$$

(given an optimal choice of W_T) can be shown to be asymptotically distributed as a chi-square random variable with $s-l$ degrees of freedom. In fact, as shown in Hansen (1982),

$$T\left[H_T(x) - H_N\left(y(\hat{\beta}_{TN})\right)\right]'W_T\left[H_T(x) - H_N\left(y(\hat{\beta}_{TN})\right)\right]$$

$$\xrightarrow{\mathrm{D}} \chi^2(s-l),$$

where W_T is a consistent estimator of W.

3. Concluding remarks

In recent years, various methods have been proposed and implemented to provide complete solution paths to a fully specified stochastic equilibrium model, the aim of which is to explain certain empirical regularities observed in a real data set. In many cases, the comparison of the economic model to the real data has been fairly casual, and the estimation of the parameters of the model has been inexact. This paper is an attempt to address this issue.

[12]Note that allowing the random errors to be redrawn would also violate the weaker continuity assumptions found in McFadden and Pakes–Pollard.

As new methods are developed for quick and cheap simulation of economic models, it will become more straightforward to perform the additional step of estimating the parameters of the model using the simulated series.[13] In this paper, the simulation estimator was shown to have an asymptotically normal distribution under fairly general conditions; a goodness-of-fit test was also derived.

References

Fair, R.C. and John B. Taylor, 1983, Solution and maximum likelihood estimation of dynamic nonlinear rational expectations models, Econometrica 51, 1169–1185.

Hansen, Lars Peter, 1982, Large sample properties of generalized method of moments estimators, Econometrica 50, 1929–1954.

Hansen, Lars Peter and Thomas J. Sargent, 1980, Formulating and estimating dynamic linear rational expectations models, Journal of Economic Dynamics and Control 2, 7–46.

Hansen, Lars Peter and Kenneth J. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, Econometrica 50, 1269–1286.

Ingram, Beth Fisher, 1988, General equilibrium modelling of asset prices, rationality versus rules of thumb, Working paper no. 88-22, May (University of Iowa, Iowa City, IA).

Kydland, Finn and Edward C. Prescott, 1982, Time to build and aggregate fluctuations, Econometrica 50, 1345–1370.

Labadie, Pamela, 1984, A test of risk premia behavior in an overlapping generations model, Manuscript (Graduate School of Business, Columbia University, New York, NY).

Lee, Bong-Soo, 1986, Inflation, stock returns and interest rates, Ph.D. dissertation, Aug. (University of Minnesota, Minneapolis, MN).

Marcet, Albert, 1988, Solution of nonlinear models by parametrizing expectations, Manuscript (Carnegie-Mellon University, Pittsburgh, PA).

McFadden, Daniel, 1986, A method of simulated moments for estimation of multinomial probits without numerical integration, Manuscript, July (Massachusetts Institute of Technology, Cambridge, MA), forthcoming in Econometrica.

Newey, Whitney K. and Kenneth D. West, 1987, A simple, positive definite, heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703–708.

Pakes, Ariel and David Pollard, 1986, The asymptotics of simulation estimators, Working paper no. 8628, Sept. (Social Systems Research Institute, University of Wisconsin, Madison, WI).

Sims, Christopher A., 1985, Solving nonlinear rational expectations models 'backwards', Working paper no. 206 (Economic Research Center, University of Minnesota, Minneapolis, MN).

Tauchen, George, 1986, Quadrature based methods for obtaining approximate solutions to the integral equations of nonlinear rational expectations models, Manuscript (Duke University, Durham, NC).

Taylor, John, 1988, Solving nonlinear stochastic growth models: A comparison of alternative methods, Unpublished manuscript (Stanford University, Stanford, CA).

Wooldridge, Jeffrey and Halbert White, 1985, Consistency of optimization estimators, Manuscript, May (University of California, San Diego, CA).

[13]This method of estimation has been successfully applied in Ingram (1988) and Lee (1986).