

ESTIMATION BY SIMULATION

Bong-Soo Lee  
Iowa state University

Beth Fisher Ingram  
Cornell University

Staff Paper No. 179  
August 1987

#179

ESTIMATION BY SIMULATION\*

Abstract

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

Bong-Soo Lee  
Iowa State University  
Ames, IA 50011

Beth Fisher Ingram  
480 Uris Hall  
Cornell University  
Ithaca, NY 14853

January 1986

Revised: August 1987

\* The authors are grateful to Chris Sims, an anonymous referee and an Associate Editor for helpful suggestions and comments.

## I. 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. When the problem is formulated in such a way that economic agents are assumed to maximize a quadratic objective function subject to linear constraints, linear optimal decision rules can be derived and a complete characterization of the equilibrium solution path (in the form of linear difference equations) is possible. In addition, the problem can be separated into a forecasting and a non-stochastic optimization problem by the certainty equivalence principle (Sargent (1979), Hansen and Sargent (1982)).

For models which are not linear-quadratic, the certainty equivalence principle doesn't apply and a computationally practical method for finding closed form solution paths is rarely available. One way to handle this problem is the quadratic approximation method proposed by Kydland and Prescott (1982). They suggest approximation of the non-quadratic objective function by a quadratic function in the neighborhood of the model's steady state. The equilibrium decision rules and solution paths for the approximate economy can then be derived and the certainty equivalence principle can be used, but at the cost of approximation errors. A second approach, suggested by Labadie (1984) and Sargent (1984), uses successive approximations of the stochastic Euler equations based on a contraction mapping theorem. A final procedure, called the "backwards mapping method," has been proposed by Novales (1983) and Sims (1985). They suggest taking a symmetric view of controlled variables and

forcing variables so that an arbitrary stochastic process for controlled variables implies a corresponding stochastic process for the uncontrolled variables. This method generates mutually consistent processes for the exogenous disturbances and choice variables without any need for approximation (for examples, see Ingram (1986) and Lee (1986)).

Using one of these methods of solving stochastic equilibrium models, we can generate complete solution paths to the model through appropriate simulations. However, these simulated solution paths to the model depend on a set of parameters from the model whose value is unknown. To generate a solution path, values must be assigned to the parameters in some way. Ideally, this would be accomplished through estimation of the parameters using observed data. In many instances, the estimation procedure has been reduced to a naive grid search on the parameter space which is restricted on either theoretical or empirical grounds. In fact, the treatment of the estimation of the parameters of the complete stochastic equilibrium model has not been systematic, and the need for a formal econometric treatment has been mentioned often (Kydland and Prescott (1982), Labadie (1984)).

It is well known that the generalized method of moments procedure can be used to estimate nonlinear rational expectations models without explicitly solving for the full stochastic equilibrium (Hansen (1982), Hansen and Singleton (1982)). The GMM procedure is convenient in that it requires neither a complete, explicit representation of the economic environment nor a complete system of equations generating the data.

However, the GMM procedure may not be available in some cases.<sup>1</sup> In addition, we may want to use the complete structure of the model to make forecasts or to organize a broad range of dynamics of the model in a way which allows us to compare the fit of our model to another model. Unlike GMM, the approach suggested in this paper also allows the researcher to choose the aspects of the model which he would like to emphasize in the estimation. In other words, we would like the model we develop to mimic 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. Recently, McFadden (1986) and Pakes and Pollard (1986) have proposed similar simulation estimators, mainly for use in the discrete response problem. While these authors do not allow serial correlation in the data set, the estimator by simulation in this paper implicitly permits the disturbance terms to be serially correlated. The criterion functions in these 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 and Pollard.

---

<sup>1</sup> Labadie (1984) found that the GMM procedure could not be used since it was difficult to construct a disturbance term that was convenient for econometric estimation. Instead, she used a variation of the method of minimum distance.

The rest of the paper is organized as follows. Section 2 introduces estimation by simulation and establishes the asymptotic normality of the estimator. Section 3 outlines a chi-square test of the model and Section 4 concludes the paper.

## Section 2. Consistency and Asymptotic Normality

Suppose we have a fully specified, stochastic general equilibrium model which generates an  $m \times 1$  vector of simulated series,  $y_j(\beta)$ ,  $j=1, \dots, N$  from an  $l \times 1$  parameter vector  $\beta$ .  $\beta$  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, the model is a true description of some economic phenomenon, so that  $y_j(\beta)$  is assumed to have a counterpart in an observed data set, denoted  $x_t$ ,  $t=1, \dots, T$ .<sup>2</sup>  $x_t$  is also an  $m \times 1$  vector. Heuristically, the proposed simulation estimator of  $\beta$  is obtained by equating moments of the simulated series,  $y_j(\beta)$ , to moments of the observed data series,  $x_t$ .

Let

$$H_T(x) = \frac{1}{T} \sum_{t=1}^T h(x_t)$$

$$H_N(y(\beta)) = \frac{1}{N} \sum_{j=1}^N h(y_j(\beta))$$

That is,  $H_T(x)$  is an  $s \times 1$  vector of statistics calculated, using the real data, as a time average of some function of the observations, and

---

<sup>2</sup> 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.

$H_N(y(\beta))$  is a corresponding vector of statistics calculated, using the simulated data, from the model. The simulation estimator,  $\hat{\beta}_{TN}$ , can then be defined as follows.

Definition. Given a random, symmetric,  $s \times s$  weighting matrix  $W_T$  of rank at least 1,  $\hat{\beta}_{TN}$  is chosen to solve the following:

$$\text{Min}_{\beta} [H_T(x) - H_N(y(\beta))] ' W_T [H_T(x) - H_N(y(\beta))]$$

We will fix a real number  $n=N/T > 1$  and define  $g_t(\beta)$  such that<sup>3</sup>:

$$G_T(\beta) = \frac{1}{T} \sum_{t=1}^T g_t(\beta) = \frac{1}{T} \sum_{t=1}^T h(x_t) - \frac{1}{T \cdot n} \sum_{j=1}^{T \cdot n} h(y_j(\beta))$$

$$= \frac{1}{T} \sum_{t=1}^T \left[ h(x_t) - \frac{1}{n} \sum_{j=n_0}^{n_1} h(y_j(\beta)) \right]$$

where  $n_0 = [1 + (t-1)n]$ ,  $n_1 = [nt]$  and  $[m]$  represents the smallest integer less than or equal to  $m$ . Essentially, we are requiring that the length of the simulated series be longer than the length of the observed data series, and that as  $T \rightarrow \infty$ ,  $N$  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.

The consistency of the simulation estimator can be shown in several ways. Wooldridge and White (1985) give very general conditions under which an optimization estimator such as that examined here is consistent. However, in practice, their conditions are difficult to verify. Hansen (1982) provides a set of assumptions similar to those

---

<sup>3</sup> If  $n$  is an integer, then  $g_t(\beta) = h_t(x) - (1/n) \sum h(y_{k,t}(\beta))$ , where simulations may be indexed by the pair  $(k,t)$ ,  $k = 1, \dots, n$  and  $t=1, \dots, T$ .

used to show the consistency of nonlinear instrumental variables estimators. We repeat Hansen's theorem 2.1 here.

Theorem Suppose the following:

C1.  $x_t$  and  $y_j(\beta)$  are independent, and are stationary, ergodic processes.

C2.  $(S, \sigma)$  is a separable metric space, where  $\beta_0 \in S$  and  $S$  is a compact subset of  $R^1$ .

C3.  $h(y_j(\beta))$  is Borel measurable for each  $\beta \in S$  and continuous on  $S$  for each  $y$  in  $R^m$ .  $h(x_t)$  is continuous for each  $x$  in  $R^m$ .

C4.  $h(y_j(\beta))$  is continuous in the mean at  $\beta_0$ . That is,

$$\lim_{\delta \rightarrow 0} E \left[ \sup \{ |h(y_j(\beta_0)) - h(y_j(\beta))| : \beta_0, \beta_1 \in S, |\beta_0 - \beta_1| < \delta \} \right] = 0$$

C5.  $E(g_t(\beta))$  exists and is finite for all  $\beta \in S$  and  $E(g_t(\beta_0)) = 0$ .

C6.  $W_T \rightarrow W$  almost surely.

Then if  $g_t(\beta)$  has a unique zero at  $\beta_0$ , the simulation estimator  $\hat{\beta}_{TN}$  exists and converges almost surely to  $\beta_0$ .

The following assumptions are sufficient to guarantee the asymptotic normality of the simulation estimator:

N1.  $x_t$  and  $y_j(\beta)$  are independent, and are stationary, ergodic processes.

N2.  $S$  is an open subset of  $R^1$  that contains  $\beta_0$ .

N3.  $h(x_t)$ ,  $h(y_j(\beta))$  and  $dh(y_j(\beta))/d\beta$  are Borel measurable for each  $\beta \in S$  and  $dh(y_j(\beta))/d\beta$  is continuous on  $S$  for each  $y_j(\beta) \in R^m$ .

N4.  $dh(y_j(\beta))/d\beta$  is continuous in the mean at  $\beta_0$ . That is,

$$\lim_{\delta \rightarrow 0} E \left[ \sup \left\{ \left| \frac{dh(y_j(\beta_0))}{d\beta} - \frac{dh(y_j(\beta))}{d\beta} \right| : \beta_0, \beta_1 \in S, |\beta_0 - \beta_1| < \delta \right\} \right] = 0$$

N5.  $E[dh(y_j(\beta_0))/d\beta] = B$  exists, is finite and has full rank.

N6.  $E[g_t g_t']$  exists and is finite,

$E[g_t | g_{t-1}, g_{t-1-1}, \dots]$  converges in mean square to zero, and

$\sum_{i=0}^T E[v_i v_i']^*$  is finite,  
 $i=0$

where  $v_i = E[g_t | g_{t-1}, g_{t-1-1}, \dots] - E[g_t | g_{t-1-1}, g_{t-1-2}, \dots]$ <sup>4</sup>

N7.  $\hat{\beta}_{TN} \rightarrow \beta_0$  in probability as  $T \rightarrow \infty$ .

N8.  $W_T \rightarrow W$  in probability as  $T \rightarrow \infty$ .

N1 ensures that time averages of functions of these series will converge to the corresponding population moments. For example,  $H_T(x)$  will converge almost surely (as  $T$  increases to infinity) to  $E[h(x)]$  and  $H_N(y(\beta))$  will converge almost surely (as  $N$  increases to infinity) to  $E[h(y(\beta_0))]$ , where  $\beta_0$  is the 'true' parameter vector. Under the null hypothesis that the model we are simulating is a true description of the real world, the population moments of  $h(x)$  and  $h(y(\beta_0))$  will be equal so that  $E[h(x)] = E[h(y(\beta_0))]$ . In addition, N1 guarantees that the influence of the initial values assumed for the simulated series will decrease as  $N$  gets large.<sup>5</sup>

N6 entails conditions which are sufficient to allow use of a central limit theorem for stationary, ergodic martingale processes.

In N7 we are assuming either that C1 - C6 are satisfied or that consistency of the estimator has been shown in some other fashion.

Following methods outlined in Hansen (1982) we can prove:

Proposition 1. Given N1 - N8,

<sup>4</sup>  $\{v_i: i \geq 0\}$  is a martingale difference sequence.

<sup>5</sup> This is important because, as pointed out by an Associate Editor, the simulation of a macroeconomic model involves choosing initial values for endogenous state variables. Under N1, the distribution of the estimator will not be affected by this choice.

$$\sqrt{T} (\hat{\beta}_{TN} - \beta_0) \xrightarrow{D} N(0, (aB)^{-1} a' S_G a' (aB)^{-1'}) \text{ as } T \rightarrow \infty$$

where B is defined as above,  $a=B'W$  and  $S_G = \text{Cov}\{H_T(x) - H_N(y(\beta))\}$ .

Proof (Sketch): After substitution of  $G_T(\beta)$ , the simulation estimator is chosen to solve:

$$\text{Min}_{\beta} [G_T(\beta)]' W_T' [G_T(\beta)]$$

Define the  $l \times s$  matrix  $a_T = DG_T(\beta)' W_T$  where  $DG_T(\beta)$  is the Jacobian matrix of  $G_T$ . The first order condition for a solution to this minimization problem may be written:

$$a_T G_T(\hat{\beta}_{TN}) = 0$$

Under N2 - N4, we may expand this quantity in a Taylor's series:

$$a_T G_T(\hat{\beta}_{TN}) = 0 = a_T G_T(\beta_0) + a_T B_N(\bar{\beta})(\hat{\beta}_{TN} - \beta_0)$$

or,

$$\sqrt{T} (\hat{\beta}_{TN} - \beta_0) = -\sqrt{T} (a_T B_N(\bar{\beta}))^{-1} a_T G_T(\beta_0)$$

where the  $i^{\text{th}}$  row of  $B_N(\bar{\beta})$  is  $dh^i(y_j(\bar{\beta}_i))/d\beta$  and  $\bar{\beta}_i$  lies on the line between  $\beta_0$  and  $\hat{\beta}_{TN}$  for  $i = 1, \dots, s$ . By N7,  $\hat{\beta}_{TN} \xrightarrow{a.s.} \beta_0$  which implies that  $\bar{\beta}_i \xrightarrow{a.s.} \beta_0$ ,  $i=1, \dots, s$  as  $T \rightarrow \infty$ . Hence, under N1, N4 and N7,  $B_N(\bar{\beta}) \xrightarrow{a.s.} B$ , where B is defined as in N5.

Define:

$$R_x(i) = E\{[h(x_t) - E(h(x_t))][h(x_{t-1}) - E(h(x_{t-1}))]'\}$$

$$R_y(i) = E\{[h(y_t(\beta_0)) - E(h(y_t(\beta_0)))] [h(y_{t-1}(\beta_0)) - E(h(y_{t-1}(\beta_0)))]'\}$$

$$S = \sum_{i=-\infty}^{\infty} R_x(i) \quad (= \sum_{i=-\infty}^{\infty} R_y(i) \text{ under the null hypothesis})$$

Given N1, N5, N6, N7 and N8,

$$\sqrt{T} [H_T(x) - E(h(x_t))] \xrightarrow{L} N(0, S)$$

$$\sqrt{N} [H_N(y(\beta_0)) - E(h(y(\beta_0)))] \xrightarrow{L} N(0, S)$$

Hence,

$$\sqrt{T} (\hat{\beta}_{TN} - \beta_0) \xrightarrow{L} N(0, (aB)^{-1} a(1 + (1/n)) S a' (aB)^{-1}). \quad \text{Q.E.D.}$$

Given the definition of  $a_T$ , N8 requires that  $W_T$  be chosen so that it converges in probability to an  $s \times s$  matrix,  $W$ , of rank at least 1. Hansen (1982) shows (in theorem 3.2) that the optimal choice for  $W$  (the choice which yields the smallest asymptotic covariance matrix for the estimator) is  $W = [(1 + 1/n)S]^{-1}$ . In that case,

$$\sqrt{T} (\hat{\beta}_{TN} - \beta_0) \xrightarrow{L} N(0, [B(1 + 1/n)^{-1} S^{-1} (B')^{-1}])$$

An examination of the covariance matrix of the estimator reveals why  $N$  is chosen 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. From a practical standpoint, one may be able to neglect the randomness in the simulation and to use only the real data in deriving a consistent estimator of  $S_g$ .

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  $\beta$ , calculating  $y_j(\beta)$  each time. 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  $\beta$  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.<sup>6</sup>

### Section 3. A Chi-Square Test

The estimation procedure in the previous section 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  $a_T$ . These 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  $s-1$  degrees of freedom.

Proposition 2. Given  $N1 - N7$  and setting  $a_T = B' W_T$ ,

$$T [H_T(x) - H_N(y(\hat{\beta}_{TN}))]' W_T [H_T(x) - H_N(y(\hat{\beta}_{TN}))] \xrightarrow{D} \chi^2(s-1)$$

where  $W_T$  is a consistent estimator of  $W$ .

Proof.

Recall from the previous proof:

$$G_T(\hat{\beta}_{TN}) = G_T(\beta_0) + B_N(\bar{\beta})(\hat{\beta}_{TN} - \beta_0).$$

Substitute  $\hat{\beta}_{TN} - \beta_0 = -(a_T B_N(\bar{\beta}))^{-1} a_T G_T(\beta_0)$  and multiply by  $\sqrt{T}$ :

$$\sqrt{T} G_T(\hat{\beta}_{TN}) = [I - B_N(\bar{\beta})(a_T B_N(\bar{\beta}))^{-1} a_T] \sqrt{T} G_T(\beta_0)$$

Given our assumptions, we know that:

$$B_N(\bar{\beta}) \xrightarrow{P} B$$

$$a_T \xrightarrow{P} a$$

$$\sqrt{T} G_T(\beta_0) \xrightarrow{D} N(0, W_T^{-1})$$

$$\sqrt{T} G_T(\hat{\beta}_{TN}) \xrightarrow{D} N(0, [I - B(aB)^{-1}a] W_T^{-1} [I - B(aB)^{-1}a]')$$

<sup>6</sup> Note that allowing the random errors to be redrawn would also violate the weaker continuity assumptions found in McFadden and Pakes and Pollard.

It can be shown that this latter covariance matrix is singular with rank  $s-1$ .

Now choose  $C_T' W_T C_T = I$  (i.e.,  $W_T^{-1} = C_T C_T'$ ) such that  $C_T \xrightarrow{P} C$  and  $C$  is nonsingular. Then

$$\sqrt{T} C_T^{-1} G_T(\beta_{TN}) \xrightarrow{D} N(0, I - C^{-1} B(B'WB)^{-1} B' C^{-1'})$$

Since this covariance is idempotent and has rank  $s-1$ ,

$$T G_T(\beta_{TN})' W_T G_T(\beta_{TN}) \xrightarrow{D} X^2(s-1). \quad \text{Q.E.D.}$$

#### Section 4. 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. However, the literature has been lacking a formal econometric treatment of the estimation of the parameters of these models. This paper is an attempt to address this need.

The simulation estimator was shown to have an asymptotically normal distribution under fairly general conditions; a goodness of fit test was also derived. Given the inexpensive and powerful nature of estimation by simulation as a solution procedure, the fact that it may also be used as an estimation procedure widens its applicability to economic problems.

## References

- Hansen, Lars Peter. 1982. "Large sample properties of generalized method of moments estimators," *Econometrica*, 50:1929-54.
- Hansen, Lars Peters 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*, 50:1269-86.
- Ingram, Beth Fisher. 1986. "General equilibrium modelling of asset prices: rationality versus rules of thumb," Ph.D. dissertation, University of Minnesota, August.
- Kydland, Finn and Edward C. Prescott. 1982. "Time to build and aggregate fluctuations," *Econometrica*, 50:1345-70.
- Labadie, Pamela. 1984. "A test of risk premia behavior in an overlapping generations model," manuscript, Graduate School of Business, Columbia University.
- Lee, Bong-Soo. 1986. "Inflation, stock returns and interest rates," Ph.D. dissertation, University of Minnesota, August.
- McFadden, Daniel. 1986. "A method of simulated moments for estimation of multinomial probits without numerical integration," manuscript, Massachusetts Institute of Technology, July. (Forthcoming in *Econometrica*).
- Novales, Alphonso. 1983. "Some stochastic equilibrium models of the interest rate," Ph.D. dissertation, University of Minnesota, July.
- Pakes, Ariel and David Pollard. 1986. "The asymptotics of simulation

estimators," Social Systems Research Institute (University of Wisconsin) working paper #8628, September.

Sargent, Thomas J. 1979. Macroeconomic Theory. Academic Press, New York.

Sargent, Thomas J. 1984. "Notes on solving rational expectations models," December.

Sims, Christopher A. 1985. "Solving nonlinear rational expectations models 'backwards'," Economic Research Center Working paper #206, University of Minnesota.

Wooldridge, Jeffrey and Halbert White. 1985. "Consistency of optimization estimators," University of California, San Diego manuscript, May.