POLICY DETERMINATION THROUGH SIMULATION OF NON-LINEAR ECONOMETRIC MODELS

By

Robert James Whitacre

WORKING PAPER NO. 213

INSTITUTE FOR DEVELOPMENT STUDIES
UNIVERSITY OF NAIROBI
P.O. BOX 30197
N A I R O B I

April, 1975

SYNOPSIS

Simulation of stochastic non-linear econometric models is known to have desirable analytic content only when the error terms affecting the structural equations are incorporated within the simulation procedure. This paper demonstrates that stochastic simulation, repeated solution of a model with error terms explicitly incorporated in quantified form, results in an empirical distribution function for the endogenous variables which converges uniformly to the true distribution function. This result allows the construction of confidence intervals on the paths of the endogenous variables of the model. Furthermore the Bayes' Principle is extended to cover optimal policy determination for a finite set of available policies for stochastic non-linear econometric models; resulting in a practical procedure available for development planning.

I. Introduction and Summary

The technique of simulation, repeated solution of a model for its endogenous variables given alternative specifications of the exogenous variables, is typically regarded as a tool of "last resort" to be used only when analytical techniques are not available for obtaining solutions. In the case of nonlinear econometric models explicit analytical solutions are extremely difficult to obtain even when the number of equations is small. When the number of equations is large exact solution is, given the present state of knowledge, impossible to obtain.

When considering the solution of a stochastic non-linear econometric model for planning purposes in the face of objective criteria there are additional difficulties stemming directly from the stochastic error terms which are contained within the individual structural equations of the model.[1] In particular the application of nonstochastic simulation procedures to econometric models that contain nonlinearities in the structural equations, e.g. simulation which does not take account explicitly of the error terms attached to the structural equations in quantified form, yields results that are not consistent with the properties of the reduced form of the model. Highly undesirable implications of such nonstochastic procedural techniques are threefold: (1) Any derived policy designed to satisfy specified objective criteria will be unlikely to be optimal (or even close to optimal) for the stochastic econometric model being investigated. Since a stochastic econometric model presumably is a close approximation of economic phenomena within an economy this implies that the derived macroeconomic policy is unlikely to fulfill policy objectives in actual application. (2) Predictions of the path of endogenous variables over time will be biased, the extent of such bias not being ascertainable from any nonstochastic framework of analysis short of actual solution or approximation through analytic expansion of the reduced form equations by means of

---

1.) This is true for the World Bank model of the Kenyan economy.

polynomials.[2] And (3) since the technique results in unique nonstochastic estimators of parameters for which the error distribution is unknown it is impossible to determine confidence intervals and significance tests. This precludes evaluation of predictive performance and reliability both in terms of individual parameter estimates and the model as a whole in terms of probability statements.[3]

Stochastic simulation of nonlinear systems, repeated runs of the model with error terms explicitly included in the structural equations of the model in quantified form, can yield estimates of the time path of the endogenous variables which are consistent with the estimators that could be obtained from the reduced form equations. In addition stochastic simulation results in a posterior error distribution attached to the various parameters of the model which converges asymtotically to the associated multivariate distribution function. Thus, given sufficient stochastic simulation, the multivariate distribution function of the endogenous variables can be asymtotically approximated from the posterior empirical frequency distribution; this convergence being uniform. This implies that meaningful confidence intervals and significance tests can be used to analyze the predictive performance and reliability of individual parameter estimates and the model as a whole; this can be done by utlizing the empirical frequency distribution.

Since results consistent with the reduced form equations of a nonlinear model must be obtained from stochastic simulation of the model through a large number of runs utilizing explicit error terms obtained from the frequency distributions attached to the structural

---

2) Approximations of this latter sort however are unlikely to lead to valid comparisons of diverse policy proposals as the derived polynomial expansions are done with respect to the means of the endogenous and exogenous variables. The resulting accuracy of such polynomial expansion becomes increasingly suspect the further a policy specification in terms of the values given to the exogenous variables lies from the means of the exogenous variables in the data base. This error compounds those which are normally expected for policies differing from the existing data base.

3) When considering the first set of policy recommendations and projections of the World Bank model which were derived from a nonstochastic simulation of what is a stochastic non-linear simultaneous equations model the above conclusions seem quite relevant.

equations there is a very definite problem in obtaining a macroeconomic policy package which adequately fulfills predetermined objective planning criteriae. In the case of nonstochastic simulation it is possible to specify an objective function defined over the specific set of criteria and approximate the optimal policy (that which best satisfies plan objectives) to any degree of desired accuracy through a number of approximation techniques. Among the approximation techniques available are the sequential simplex technique, Hookes-Jeeves method, Gauss-Seidel method, Newton's method, the Newton-Raphson method, the modified Newton-Raphsan method, the relaxation method, the modified relaxation method, inner linearization, outer linearization and Box's Complex method.[4,5] Under stochastic simulation it seems possible to adopt each and every one of these techniques so that derived policies approximate optimal policies in an expectational sense. These extensions are now being conducted by the author. Under the necessary modifications the computational requirements for an adequately close approximation increase considerably, however today's high speed computers should prove capable to the task in empirical applications.

Section II, The Reduced Form, Nonstochastic and Stochastic Simulation, presents the argument involved in demonstrating why nonstochastic simulation yields estimates of the values of endogenous variables which will be systematically biased. Then it is shown why stochastic simulation yields results that are consistent with the original specification of the model.[6]

4.) If it were desirable to find a macroeconomic policy package through nonstochastic simulation specification first of an objective function defined over policy objectives, followed by subsequent approximation by one of these techniques, would appear to be the proper course of action.

5.) For any purely deterministic model, such as the dynamic linear programming Kensim model of the Kenyan economy, simulation according to one of these techniques, in conjunction with appropriately specified objective criteria, would seem entirely appropriate.

6. The results in this section rely heavily upon an analysis by E.R. Howrey and H.H. Kelejean contained in /̄6/.

Section III goes on to establish that confidence intervals for endogenous variables can be established by considering the empirical frequency distribution of the values of the endogenous variables generated by repeated trials of stochastic simulation procedures. This yields a technique capable of evaluating the model in terms of actual historical values. Finally the problem of choosing an objective function for policy determination in terms of endogenous random variables is discussed. If complete specification of the gains (or losses) accruing to the various possible states that may arise is made then the decision theory approach of choosing that policy which minimizes the implied risk function is found to be the appropriate procedure. Finally an empirical technique is specified which is capable of choosing optimal policy when only a finite set of policy alternatives is specified. The problem of finding an optimum policy when the policy set is unconstrained is currently under investigation by the author.

II. Reduced Form, Nonstochastic and Stochastic Simulation

Consideration is of a model non-linear in the endogenous variables but linear in parameters. The conclusions concerning nonstochastic simulation obvious extends to models possibly non-linear in both endogenous variables and parameters, as the analysis here concerns a special case of this latter category.

In general form this model can be represented as

$$(1) \qquad y_t - x_t H_1 + F(y_t, y_{t-1}, x_t) H_2 + R(y_{t-1}, x_t) H_3 + u_t$$

where y_t is the vector of endogenous variables at time t, x_t is the vector of endogenous variables at time t (additional lagged values of the endogenous and exogenous variables could be introduced without affecting the analysis), H_1, H_2, and H_3 are respectively, G X K, M_1 X K, and M_2 X K matrices of parameters; F is a $1 \times M_1$ vector of observations at time t on M_1 functions f_{it} each of which is different and dependent upon at least one endogenous variable. Also at least one it is nonlinear. In a similar manner R is a 1 X M_2 vector of observations on M_2 functions

r_{it}. Assumption is made that the underlying process is a stationary stochastic one and that the probability limits of the sample moments are equal to their corresponding expectations. In order to preserve linearity in the parameters requirement is made that if f_{it} contains the jth element of y_t then the i, jth element of H_2 is zero.

Consider f_{it}. Each f_{it} can be considered a random variable, then

(2) $$E(f_{it}/x_t, y_{t-1}) = s_{it}, \quad i=1,\ldots\ldots\ldots M_1$$

where $s_{it} = s_i(x_t, x_{t-1})$, assuming of course these expectations exist. Now we can express f_{it} as

(3) $$f_{it} = s_{it} + w_{it}, \qquad i = 1, \ldots\ldots\ldots, M_1$$

where w_{it} is a stochastic element so that $E(w_{it}/x_t, y_{t-1}) = 0$.

Thus

(4) $$F(y_t, y_{t-1}, x_t) = S(y_{t-1}, x_t) + W_t,$$

where S and W are 1 X M_1 vectors whose ith elements, respectively, are s_{it} and w_{it}, Substituting (4) into (1)

(5) $$y_t = x_t H_1 + S(y_{t-1}, x_t) H_2 + R(y_{t-1}, x_t) H_3 = e_t = J(y_{t-1}, x_t) = e_t,$$

where $e_t = u_t + W_t H_2$. Now $E(e_t/x_t, y_{t-1}) = 0$ and, thus

$E(y_t/x_t, y_{t-1}) = J(y_{t-1}, x_t)$. The set of equations defined by (5) may be termed the reduced form equations of our model. These equations however are not a solution for the endogenous variables in terms of x_t, y_{t-1}, and linear combinations of the structural disturbances in u_t. This is clear since to arrive at these equations a nonlinear transformation of additive error terms was necessary to obtain the reduced form.

Consider the following example.

(6) $y_{1t} = b_1 x_t + u_{1t}$.

(7) $y_{2t} = b_2 y_{1,t-1} + b_3 \exp(y_{1t}) + u_{2t}$,

where the disturbances are normally distributed with means zero, variances α_1^2; and covariance $\alpha_{12} \neq 0$. Assume that each u_{it} is not autocorrelated and, furthermore, is independent of x_t.
The solution for y_{2t} is then

(8) $y_{2t} = b_2 y_{1t-1} + b_3 \exp(b_1 x_t) \exp(u_{1t}) + u_{2t}$.

Since $E(\exp(u_{1t})/x_t, y_{1t}) = \exp(\alpha_1^2/2)$ we can define u_{3t} with expectation zero such that

(9) $\exp(u_{1t}) = \exp(\alpha_1^2/2) + u_{3t}$.

Substituting (9) into (8)

(10) $y_{2t} = b_2 y_{1t-1} + b_4 \exp(b_1 X_t) + z_t$,

where $b_4 = b_3 \exp(\alpha_1^2/_2)$, and the reduced - form disturbance

$z_t = u_{2t} + b_3 \exp(b, X_t) u_{3t}$. Here z has expectation zero.

In comparing (10) with (7) setting $z_t = 0$ cannot be obtained by merely setting u_{1t} and u_{2t} equal to zero, thus (7) is not a solution. Also z_t is not a linear function of the disturbances, it is heteroskedastic while the structural disturbances are homoskedastic. Also nonlinear transformations of time series data not autocorrelated in the error terms of the structural equations will have autocorrelation in the reduced-form disturbances.

Suppose that the parameter matrices H_1, H_2, and H_3 are known; that is we eschew a discussion of estimation. Suppose we also have generated solutions for the system y_t^*, t= 1, ..., through some algorithmic approximation technique ignoring error terms. Then let

(11) $$y_t^* = T_1(y_{t-1}^*, x_t).$$

Assume that the solution of the original structural system including error terms is

(12) $$y_t = T_2(y_{t-1}, x_t, u_t).$$

(13) $$y_t^* = T_2(y_{t-1}^*, x_t, 0) = T_1(y_{t-1}^*, x_t).$$

Now since T_1 is not equal to J, multiplier analysis based on nonstochastic simulation will not apply to multipliers deriving from the full model including structural errors. Also since the structural disturbances will generally be combined in a nonadditive fashion with reduced-form parameters the elements of y_t^* will diverge systematically from the elements y_t. That is given a stochastic process

(14) $$E(.y_t - y_t^*)/y_0, x_t) = A(y_o, x_t) \neq 0.$$

Thus, ideally, validation should be carried out in terms of the multivariate distribution theory, corresponding to the estimates of the structural parameters and the various tests for randomness that concern the structural disturbances.

The outcome then is clear. In the case of nonlinear simultaneous equation models nonstochastic simulation will yield biased results, the bias not disappearing in the asymtotic limit.

Given that nonstochastic simulation is an exercise of questionaable value the problem remains that of finding estimates and associated confidence intervals concerning both the endogenous variables and the reduced form parameters. Let v_t be a vector of disturbances at time t generated from the structural disturbances u_t. Then let y_t^{**} be the solution of the resulting equations, e.g.,

(15) $$y_t^{**} = T_2(y_{t-1}^{**}, x_t, v_t), \; t = 1, \ldots.$$

Then if the x_t are held fixed in repeated trials for v_t taken from the structural disturbances.

$$\text{(16)} \qquad \text{plim}\; N^{-1} \sum_{i=1}^{T} T_2[y^{**}_{t-1}(i),\; x_t(i)] = J(y^{**}_{t-1},\; x_t).$$

Thus the properties of the reduced-form equations may be studied in terms of the simulation results corresponding to different time paths of x_t.[7] The problem remains of determining proper policy choice criterion given this state of affairs.

III. Decision Making Under Stochastic Simulation

Ideally when considering the various policies available it would be most desirable to have a 'best' decision rule for finding an optimal policy course. In the context of a stochastic environment such a rule would generate the smallest risk no matter what the true state of nature. Unfortunately, situations in which a best decision rule exists are rare, and almost exclusively limited to deterministic models. For each fixed state of nature there may be a best action to take. However, this best action will differ for different states of nature, so that no one policy will be best overall.

Phrasing this argument another way, as a practical procedure it is quite possible to specify an objective function defined over the endogenous variables of a simultaneous equations model and approximate an optimal policy over a period of time through a number of techniques, most of these having been mentioned in Section I, the approximation being achievable to any desired degree of accuracy. Under stochastic simulation, for each set of specified values of the error terms affecting the structural equations, it would indeed be possible to derive an optimal policy under one of the aforementioned approximation techniques. Such policy however would be conditional upon the specified errors and would not, as a general rule, be the optimal policy under an alternative specification of the structural disturbances. The problem thus reduces to one of choice, that of finding a "reasonable" rule, "reasonable" here meaning a rule that is better than "just guessing", e.g. a rule independent of the observed value of experiments.

7. The result stated in equation (16) is also stated in Howrey and Kelejian (6,p.310) but with no rigorous demonstration of its truth. (No proof is presented). The validity of equation (16) may be taken as a corrollary to Theorem 1 in Section III.

It is suggested here that given the situation the most desirable principle for ordering the various decision rules is the Bayes' criterion. Choice of policy should be that which minimizes Bayes risk, or in economic terms, maximizes expected welfare. Specifically given a decision rule δ (or in economic terms a specified policy quantitatively defined on the controllable parameters and exogenous variables) in conjunction with a parameter space B with prior distribution τ then the Bayes risk of δ is

(17) $$r(\tau,\delta) = E[R(T,\delta)]$$

where T is a random variable over B with distribution τ
Then the minimum Bayes' risk is given as:

Definition 1: A decision rule δ_o is said to be Bayes with respect to a prior distribution τ B if

(18) $$r(\tau_1\delta_o) = \inf_{\delta \Sigma D^*} r(\tau,\delta)$$

Here D is the set of all admissible decision rules.

In the context of a simultaneous equations model applying the Bayes' principle in terms of an actual empirical exercise seems quite feasible if the distribution τ were known, the number of policies were finite and losses or gains were specified for each possible state that could arise. In this situation the distribution over the statespace would be calculated for each δ, weighted by losses for each state and then the expectation taken. δ_o would then be the minimum of this finite set of expectations.

The problem in applying this procedure to stochastic simulotion of a simultaneous equation model is that the relevant prior distribution τ is not known, as this is the error distribution of the solution to the structural equations. If it were known then application of stochastic simulation would be unnecessary as the expectations required by the Bayes' principle could be calculated directly. What is clearly desired of our stochastic simulation procedure is a means of approximating t through a specified procedure of stochastic simulation.

In order to establish the desired result it is necessary to extend the famous Glivenko-Cantelli Theorem[8] (also known as the Central Statistical Theorem) to the multivariate case.

<u>Theorem 1</u>: If F is the multivariate distribution function of a vector of random variables X^i, $1 \leq i \leq n$, where $X = \begin{bmatrix} X^1 \\ \cdot \\ \cdot \\ \cdot \\ X^n \end{bmatrix}$ and

F_n is the empirical distribution function of X in n independent and identical trials, then

C = 1.

with probability 1.

In other words 1, $F_n(X)$ F(X) uniformly in X.

Proof: See Appendix 1

Theorem 1 in conjunction with the Bayes' Principle then yields a practical technique for evaluating alternative policies for a simultaneous equations macroeconomic model. Specifically a set of alternative policy courses is initially specified and then a large number of simulations for each of these policies is undertaken (presumably upon the computer) specifying different stochastic disturbance terms for all of the structural equations. The resulting posterior distribution of the values of the endogeneous random variables, since we have a large sample and theorem 1 applies, will converge uniformly to the unknown distribution function of these same endogeneous random variables. The mean of this empirical distribution function for a given policy, weighted at each frequency by its' respective welfare - value yields the expected welfare gain of this policy. The best policy within the resulting set of expected welfare gains, according to the Bayes' Principle, is that with the greatest expected welfare gain.

---

8) See [9], p. 20.

It is noted here that the resulting optimal policy from this procedure is not the optimum optimorum. Restriction has been made to a finite set of policies. Within the context of macroeconomic planning for an actual economy this set would presumably reflect diversity both as to focus and implementation over time. As examples one policy might be a focus upon agricultural development whereas another might be an emphasis on industry. However since the desirable properties of this simulation technique are all asymtotic the set of policies to be compared might necessarily have to be rather small. Consider for example estimation of the posterior error distribution for a nonlinear simultaneous equation model over several time periods. Each and every simulation normally entails specification of the disturbance terms and approximate solution of the resulting set of nonlinear equations for each and every time period; presumably by one of the aforementioned algerithmic techniques. This is neither a costless nor timeless operation. Furthermore this process has to be repeated a "large number" of time for each alternative policy. A "large number" typically depending upon the number of endogenous variables since the resulting error distribution on the endogeneous variables will typically be autocorrelated and heteroskedastic. A "large number" should probably not be less than one hundred simulations for each policy specification for even a small model. All, however, is not lost as solutions in simulation after the first one has been completed can be utilized as starting values in the solution of successive simulations; thereby reducing the computationa burden considerably.

While the preceding paragraphs give a practical technique for evaluating alternative policies under the existing computer capabilities, the question remains as to how to proceed to an approximation to the optimum optimorum; that is, how to find a policy which is indeed close to a Bayes' solution when the policy set is unrestricted; unrestricted in the sense that all feasible values of the exogeneous variables are allowed within the set of available policies. Examination of this problem will be investogated in a subsequent paper as inclusion here would make the present investigation unduly long.

Appendix 1

Proof of Theorem 1

Denote the i^{th} component of $F(X)$ by $F^i(X^i)$ and that of $F_n(X)$ by $F_n^i(X^i)$.

Set $F(\bar{X}-0) = F(\bar{X})P[X < \bar{X}]$, $F(\bar{X}+0) = P[X \leq \bar{X}]$

so that $P[\bar{X} = X] = F(\bar{X}+0) - F(\bar{X})$. The function F so define determines the probability distribution of X, e.g. the distribution function of X. If all X_j, $j \neq i$ are held fixed and, as above we denote the i^{th} component of $F(X)$ as $F^i(X^i)$ then $F^i(X^i)$ defines the conditional probability distribution of F with respect to the random variable X^i.

Let X^i_{jk} be the smallest value X^i such that $F(X^i) \geq \frac{j}{k} \leq F(X^i+0)$.

Since the conditional frequency of the event $[\bar{X}^i \leq X^i_{jk}]$ is $F^i_n(X^i_{jk})$ and its conditional probability is $F^i(X^i_{jk})$, it follows by Borel's result[*] that $P\,A^i_{jk} = 1$ where $A^i_{jk} = [F_n(X^i_{jk}) \to F(X^i_{jk})]$. Similarly, $P\,A''_{jk} = 1$ where $A''_{jk} = [F_n(X^i_{jk} + 0) \to F(X^i_{jt} + 0)]$ Let $A_{jk} = A'_{jk}A''_{jk}$

and let $\theta = \pm 0$, then

$$A_k = \bigcap_{j=1}^{k} A_{jk} = \left\{ \sup_{1 \leq j \leq k} \left| F_n(X^i_{jk} + \theta) - F(X^i_{jk} + \theta) \right| \to 0 \right\}.$$

---

*) Borel's result is that if $\forall\, \Sigma$, $\sum_{n=1}^{\infty} P[X_n \geq \frac{1}{\Sigma}] < \infty$, then $P[X_n \neq 0] = 0$. This is also known as the strong law of large numbers

Now $P \bigcup_j A_j = P\,A_1 + P\,A_1^c \cdot A_2 + P\,A_1^c\,A_2^c\,A_3 + \cdots \leq \sum_j P\,A_j$.

Therefore $P(A_k^c) = P \bigcup_{j=1}^{k} A^c_{jk} \geq \sum_{j=1}^{k} P\,A^c_{jk} = 0$

and, hence, $P\ A_k = 1$. Upon setting $A = \bigcap_{k=1}^{\infty} A_k$, it follows similarly that $P\ A = 1$.

On the other hand, for every X^i between X^i_{jk} and $X^i_{j+1,k}$

$$F(X^i_{jk}+0) \leq F(X^i) \leq F(X^i_{j+1,k}),$$

$$F_n(X^i_{jk}+0) \leq F_n(X^i) \leq F_n(X^i_{j+1,k})$$

While for every X^i_{jk}

$$0 \leq F_n(X^i_{j+1,k}) - F(X^i_{jk}+0) \leq \frac{1}{k}.$$

Therefore

$$F_n(X^i) - F(X^i) \leq F_\nu(X^i_{j+1,k}) - F(X^i_{jk}+0) \leq F_n(X^i_{j+1,k}) - F(X^i_{j+1,k}) + \frac{1}{k}$$

and

$$F_n(X^i) - F(X^i) \geq F_n(X^i_{jk}+0) - F(X^i_{j+1,k}) \geq F_n(X^i_{jk}+0) - F(X^i_{jk}+0) - \frac{1}{k}.$$

Thus for all X and k

$$| F_n(X^i) - F(X^i) | \leq \sup_{1\leq j\leq k} |F_n(X^i_{jk}+0) - F(X^i_{jk}+0)| + \frac{1}{k}$$

Or

$$\Delta_n = \sup_{-\infty<X^i<\infty} |F_n(X^i) - F(X^i)| \leq \sup_{1\leq j\leq k} | F(X^i_{jk}+\theta - F(X^i_{jk}+\theta) | + \frac{1}{k}.$$

Hence $P\ | \Delta_n \to 0 | \geq PA = 1$. Thus every empirical conditional probability distribution for any random variables $X^i, 1\leq i\leq n$, regardless of the conditional values, converges uniformly to the corresponding conditional probability function. Now since this convergence is independent of the conditional values the same uniform convergence will hold for the empirical marginal probability distributions and the marginal distribution functions for all X_i, $1\leq i\leq$. Denote the marginal distribution function for the random variable X_i by $\overline{F}(X^i)$. Finally the multivariate distribution function $F(X)$ can be regarded as the product space of all $\overline{F}(X^i)$, $1 \leq i \leq n$. This establishes the desired results.

REFERENCES

1. Beveridge, G.S.G. and Schechter, R.S., Optimization: Theory and Practice. Mcgraw-Hill, New York, 1970.

2. Ferguson, T.S., Mathematical Statistics, Academic Press, New York, 1967.

3. Fisk, P.R., Stechastically Dependent Equations. Harper New York, 1967.

4. Hood, W.C. and Koopmans, T.C. Studies in Econometric Method, Wiley, New York, 1953.

5. Loeve, M.M., Probability Theory, Van Nostrand, 3rd edition, Princeton.

6. Computer Simulation Experiments with Models of Economic Systems, Naylor, T. (ed.) Wiley, New York, 1971.

7. The Design of Computer Simulation Experiments, Naylor, T. (ed.), Duke University Press, Durham, N.C., 1969.

8. Wald, A., Sequential Analysis, Wiley, 1947, New York.

9. Wald, A., Statistical Decision Functions. Wiley, 1950, New York.

10. International Bank for Reconstraction and Development, The Second Decade ; A Basic Economic Report on Kenya, Vol. I. V.

11. Slater C., and Walsham, G. "A General systems simulation of the Kenyan economy. "July, 1974 I.D.S. Working Paper 174.