NONLINEAR SIMULTANEOUS EQUATIONS: ESTIMATION AND PREDICTION

Stephen M. Goldfeld and Richard E. Quandt

Econometric Research Program
Research Memorandum No. 79
October 21, 1965

The research described in this paper was supported by National Science Foundation Grant NSF G 24462 and ONR Contract No. Nonr 1858(16). The computer facilities used are supported by National Science Foundation Grant NSF-GP579.

Princeton University
Econometric Research Program
92-A Nassau Street
Princeton, N. J.

ABSTRACT

The purpose of this paper is to begin an investigation of the small sample properties of certain estimators of the coefficients of systems of simultaneous nonlinear equations. Sampling experiments are used in connection with two specific nonlinear models. The estimating methods investigated comprise direct least squares, various forms of two-stage least squares and full-information maximum likelihood. The relative performances of the various methods are evaluated on the basis of informal comparisons of their respective mean absolute errors and root mean square errors and also by more formal tests of significance. Direct least squares is found to be, as expected, the worst estimating method. The other two methods are rather more comparable with full-information maximum likelihood holding the edge for both theoretical and experimental reasons.

NONLINEAR SIMULTANEOUS EQUATIONS: ESTIMATION AND PREDICTION

1. Introduction

The realization that economic variables are jointly dependent has led to the simultaneous-equation approach in econometrics. This approach recognizes that in a complete system of simultaneous equations (1) the conditions of the Markov theorem of least squares are not satisfied and (2) that the application of ordinary least squares to each equation does not yield consistent estimates.

Considerable attention has been devoted to devising various alternative estimating methods which all share the property of consistency. Among these methods one may note full-information maximum likelihood, limited-information maximum likelihood, two-stage least squares, three-stage least squares, unbiased -k, indirect least squares, and others.[1]

Typically these methods have been devised for and applied to linear structural models of the type

$$By + \Gamma z = u$$

where y and z are vectors of endogenous and exogenous variables respectively, u is a vector of unobservable error terms and B and Γ are matrices of coefficients to be estimated.

A rational choice among alternative methods requires knowledge of the small sample distributions of the various types of estimators. These

---

1. See [6], [9], [10], [11].

have been examined in various ways. For a few isolated cases, Basmann has derived the exact finite sample distribution.[2] Others have engaged in more or less extensive sampling experiments.[3] Although these sampling experiments have yielded some results of broad usefulness, they have not been fully satisfactory, if for no other reason than the pervasive suspicion that the results are peculiar to the specification of the models and structures used in the experiments.

The purpose of this paper is to analyze the small-sample properties of the parameter estimates of nonlinear systems. As such it can be regarded as a natural extension of previous studies. Our study was further motivated by the following questions: (1) To what extent do previous experiments with linear models have relevance for nonlinear models, i.e., to what extent do the substantive conclusions derived from linear cases hold for nonlinear models? (2) Is there a variety of roughly equally sensible estimating methods for nonlinear models? (3) Can an algorithm which has acceptable convergence properties (i.e., an algorithm which converges rapidly to the true maximum) be devised for the calculation of full-information maximum-likelihood estimates?

We intend this paper to represent a modest beginning to the study of nonlinear models. Our method of investigation is that of sampling experiments. We are aware that all the criticisms that have been raised against sampling experiments as a research tool in this area can be repeated here with equal justification. Although the results are limited, we nevertheless feel that they will prove of some usefulness in estimating

---

2. See [1], [2], [3].

3. See [11], [12], [13], [14].

the coefficients of nonlinear systems.

The next section is devoted to a discussion of the models selected for study, the estimating methods examined, and some interesting characteristics of the models themselves that arise out of their nonlinearity. Section 3 describes the design of the experiments performed. Section 4 is devoted to an analysis of the results of the sampling experiments and Section 5 contains some conclusions.

2. Models, Methods and Identification

Models. Two basic sets of structural equations have been employed in the study. They are as follows:

Model I.

$$b_{11}\log y_{1t} + b_{12}\log y_{2t} + b_{13}z_t + b_{14} = u_{1t}$$

$$b_{21}y_{1t} + b_{22}y_{2t} + b_{23}z_t = u_{2t}$$

with covariance matrix Σ_I, and

Model II.

$$b_{11}y_{1t} + b_{12}y_{2t}^2 + b_{13}z_t + b_{14} = u_{1t}$$

$$b_{21}y_{1t}z_t + b_{22}y_{2t} + b_{24} = u_{2t}$$

with covariance matrix Σ_{II}, where, in both models, z is an exogenous variable, (u_1, u_2) are jointly normally distributed with mean = (0,0) and the indicated covariance matrix and where $E(u_t u_{t-\theta}) = 0$ for all $\theta \neq 0$.

It should be mentioned that the models analyzed in this paper were not chosen to resemble any particular economic model.[4] Sampling experiments

4. It may be noted, however, that Model I is not unlike some recent specifications of models of the demand and supply for money. See [8] and

were performed with both models. The experiments, the detailed characteristics of which are described in the next section, consisted basically of generating normally distributed error terms u_i, solving the structural equations for the values of the endogenous variables and using the resulting data series to estimate the coefficients of the equations in several ways. The estimates were then used to generate predictions for the endogenous variables.[5]

Estimating Methods. For purposes of estimation we normalized the equations in both models. In Model I the first equation was normalized on $\log y_1$ and the second equation was normalized on y_2 (i.e., $b_{11} = b_{22} = 1$). In Model II the first equation was normalized on y_1 and the second on y_2 (i.e., $b_{11} = b_{22} = 1$).[6]

It was not our intention to achieve completeness in the variety of

---

(Footnote 4 continued)
references therein. Model I is actually a variant of an example used by Fisher in his discussion of identification in nonlinear models. See [5].

5. Since in general Model I cannot be solved for the endogenous variables in closed form, we obtained solution values by solving an equivalent problem which is to minimize
$$[(b_{11}\log y_{1t} + b_{12}\log y_{2t} + b_{13}z_t + b_{14} - u_{1t})^2 + (b_{21}y_{1t} + b_{22}y_{2t} + b_{23}z_t - u_{2t})^2].$$
It may be noted that, for the particular structural parameters utilized in this paper, the problem of solving for the endogenous variables reduces to determining the real roots of a certain quintic polynomial (which can be shown to have only two distinct real roots). As a result of the minimization above there is a minor element of approximation in the data. However, since the criterion for the acceptance of a sample point was that the equation had to be satisfied to within 10^{-3}, the effect will be negligible. Because this procedure is time consuming, no predictions were generated for Model I. The second model can be solved directly and posed no problem of this kind.

6. As has been noted before, [13], normalization, in the context of direct least squares and two-stage least squares, introduces an element of arbitrariness. However, since we have normalized in the same way for both methods they are, at least, comparable.