Issued Under Office of Naval Research
Contract No. Nonr 1858(16)

SMALL-SAMPLE PROPERTIES OF VARIOUS SIMULTANEOUS-EQUATION ESTIMATORS: THE RESULTS OF SOME MONTE-CARLO EXPERIMENTS

J. G. Cragg

Econometric Research Program
Research Memorandum No. 68
October, 1964

The research described in this paper was supported by the Office of Naval Research.

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

ABSTRACT

This study reports the results of 172 Monte Carlo experiments on the small-sample properties of some simultaneous equation estimators. The estimators studied were Direct Least-Squares, Two-Stage Least-Squares, Nagar's Unbiased k-Class Estimator, Limited-Information Maximum-Likelihood, Three-Stage Least-Squares, Full-Information Maximum-Likelihood and Least-Squares used to estimate the reduced form. The estimates of the structural equations, their standard errors, the estimates of the variance-covariance matrix of disturbances, the estimates of the reduced-form coefficients and the predictive ability of the methods were examined.

The study investigated the small-sample properties of the estimators in a variety of situations, concentrating especially on the performances of the estimators relative to each other. First was examined a series of cases falling under the standard, statistical, simultaneous-equation model. Among the features studied were the effects of having different true structures generate the endogenous variables and of having different sets, amounts and types of exogenous data. It was found that the differences between the estimators were not usually great. The frequencies with which one estimator came closer to the population parameters of a structure than another one differed for different parameters and were sensitive to the true values of the parameters and to the sets of data used to estimate the structure. In many experiments it was possible to establish a ranking of the methods into three groups: full-model methods; consistent single-

equation methods and the straightforward applications of least-squares, but this ranking of the estimators was not found in several experiments.

A second part of the experiments investigated the effects of violating the stochastic assumptions of the standard model. Errors in the exogenous variables, stochastic coefficients, auto-correlated disturbances and heteroskedastic disturbances were used in these experiments. Only the use of stochastic coefficients had much effect on the estimators. It altered both their relative and absolute performances. Surprisingly, errors in the exogenous variables did not seem to lead to pronounced biases of the estimates and auto-correlated and heteroskedastic disturbances did not affect the reliability of the standard errors.

The third subject studied was misspecification of the form of the structural equations. This involved the specification of zeroes for coefficients whose true values were non-zero, the estimation of coefficients whose true values were zero and the omission of an equation. It was found that minor misspecifications had little effect on the estimators, but that misspecifications which affected important coefficients or the identifiability of the structure led to quite serious biases of the estimates, to larger dispersions for their estimates, to unreliable standard errors, to changed rankings of the methods and, in some cases, to completely ridiculous full-information maximum-likelihood estimates.

The closeness of the estimators to each other and their sensitivity to the exact sets of data and structures used made it impossible to judge any one method the best for use in econometric models. The estimators usually gave adequate indications of the orders of magnitude of the parameters of the structures, though they did not provide great precision in estimation. The standard errors frequently led to reliable confidence intervals for the estimates, but this was not always the case, especially when the structure was misspecified.

PREFACE

This study arose from dissatisfaction with the lack of knowledge about the small-sample properties of the estimators of systems of simultaneous equations. Simultaneous-equation models seem appropriate for many econometric studies. The advent of modern computers has greatly simplified the task of estimating such systems. However, the usefulness of the estimated models must be open to doubt if the properties of the estimates are unknown. This study was designed to shed more light on the small-sample properties of the simultaneous-equation estimators.

The bulk of the study reports the findings of over 150 Monte Carlo experiments on the small-sample properties of the estimators of simultaneous-equation systems. These experiments investigated the performance of the estimators in a variety of different situations. They threw a good deal of light on the small-sample properties of the estimators, especially on their performances relative to each other, though they far from succeed in answering all the questions of interest about the estimators.

The presentation of the results of Monte Carlo experiments is at best a tedious undertaking since the conclusions rest on the exact types of data investigated and arise from the consideration of specific and limited sets of estimates. To facilitate the task of readers in finding the main points of the study, short, indented summaries are presented at the ends of the sections presenting the results of the experiments.

One point should be made about the presentation of the results. The study tells how the estimators performed; it does not tell why they performed as they did. The reason for this is simple. If I knew why the results occurred, I would already know enough about the small-sample properties of the estimators to render the Monte Carlo experiments redundant. Nevertheless, heuristic explanations of the results might have been attempted and the findings presented in terms of these explanations. This was done but rarely. All too often the results of the experiments did not conform to what had been expected on heuristic grounds. It seemed better not to develop plausible explanations for these findings. Such explanations, while they would not have been at variance with the results, would otherwise have rested on no sounder grounds than the conjectures which had been overthrown. Although many heuristic beliefs about the properties of the estimators were confirmed by the study, it seemed wise not to present the results in terms of such suppositions. I can see no reason to presume that these are really the correct explanations when equally plausible conjectures were so often not supported by the results. The findings of this study rest on what happened in the Monte Carlo experiments, not on possible explanations of the results of these experiments.

My chief debt in this study is to Professor Richard E. Quandt. His support, instruction and encouragement, and his invariable willingness to discuss problems which arose in carrying out the project benefited the study immeasurably. Professors Michio Hatanaka and Stephen M. Goldfeld also made innumerable valuable and helpful comments. My thanks go to Messrs. G. J. Mitchell, Robert Cave and E. T. Irons, all of The Institute for Defense Analyses, for help in understanding and programming the 1604 computer and in using the operating system under which the programs for this study were run. The study was conducted while I was a member of the Econometric Research Program of Princeton University. I am indebted to Professor Oskar Morgenstern, its director, for his constant support and interest in the study. Finally, I am indebted to my wife, not only for the drawing of the graphs of Appendix E, but also for patiently and cheerfully putting up with being a "computer-widow" while the Monte Carlo experiments were being carried out. Needless to say, all errors and shortcomings in the design and presentation of this study are entirely my own.

TABLE OF CONTENTS

seek to identify the forces at work in a given situation and to describe the ways in which they are connected. The systems may express equilibrium positions of an economic process, be static cross-sections of it or be formulations of causal relations which determine the path of a process through time. They may relate to small parts of the economy or attempt to encompass all of it. Within the limits of their subject matter, they may try to be complete descriptions of the major forces at work or be only partial explanations of them. They may contain as few relations as the elementary supply and demand curves of the partial-equilibrium analysis in price theory or have vast numbers of equations as does general-equilibrium theory. What is common to these different systems is the formulation of models in terms of several relationships which operate simultaneously to determine economic phenomena.

Many useful insights and results can be achieved by economic theory with only a few, fairly general assumptions about the characteristics of the relationships such as the signs of their partial derivatives. Nevertheless, very frequently intuition and *a priori* reasoning alone, however valuable they may be in describing the processes at work, can give little if any indication even of the direction of change resulting from specified changes in the causal factors. Forces pull in opposite directions so that their combined effects are unknown without quantitative knowledge of their significance. Nor can much be said of the implications for the economy of the occurrence of particular sets of conditions. Frequently the theorist can say nothing of total effects or at best can only present a taxonomy of possible outcomes. These seek to identify the forces at work in a given situation and to describe the ways in which they are connected. The systems may express equilibrium positions of an economic process, be static cross-sections of it or be formulations of causal relations which determine the path of a process through time. They may relate to small parts of the economy or attempt to encompass all of it. Within the limits of their subject matter, they may try to be complete descriptions of the major forces at work or be only partial explanations of them. They may contain as few relations as the elementary supply and demand curves of the partial-equilibrium analysis in price theory or have vast numbers of equations as does general-equilibrium theory. What is common to these different systems is the formulation of models in terms of several relationships which operate simultaneously to determine economic phenomena.

Many useful insights and results can be achieved by economic theory with only a few, fairly general assumptions about the characteristics of the relationships such as the signs of their partial derivatives. Nevertheless, very frequently intuition and a priori reasoning alone, however valuable they may be in describing the processes at work, can give little if any indication even of the direction of change resulting from specified changes in the causal factors. Forces pull in opposite directions so that their combined effects are unknown without quantitative knowledge of their significance. Nor can much be said of the implications for the economy of the occurrence of particular sets of conditions. Frequently the theorist can say nothing of total effects or at best can only present a taxonomy of possible outcomes. These difficulties become more and more pronounced as the size of models increases. Estimation of the relationships becomes necessary for finding the implications of the theory. If the use of theory for prediction is desired, some estimation of the system becomes crucial, though it must be emphasized that direct estimation of the parameters of the system is far from being the only way of going about prediction or the testing of economic theory.[2]

A large step needs to be taken in proceeding from the formulation of most theoretical economic models to the estimation of their parameters. Economic theory seldom specifies the mathematical form of its equations or the stochastic nature of the system. It rarely states how, if at all, the variables are to be measured or how the theory might need to be adapted so that it would serve to describe possible statistical observations. To estimate a system of equations, these difficulties have to be overcome.[3] Finally a method of estimation has to be selected.

---

[2]Discussion of the need for and the use of estimates of the coefficients of simultaneous economic systems occurs in many references in the Bibliography to this study, particularly in J. Marschak, "Statistical Inference in Economics: An Introduction," in T. C. Koopmans (ed.) Statistical Inference in Dynamic Economic Models, (New York: Wiley , 1950), and "Economic Measurements for Policy and Prediction," in William C. Hood and T. C. Koopmans (eds.), Studies in Econometric Method, (New York: Wiley, 1953).

[3]Economic theories presented in forms which do not present most of the difficulties outlined, but instead are more nearly ready for statistical implementation, are apt to be developed in connection with what are essentially econometric studies. See, for example, Edwin Kuh, Capital Stock Growth: A Micro-Econometric Approach, (Amsterdam: North-Holland, 1963). Examples of how these problems are encountered and dealt with in the course of econometric model-building are found in many studies in which models are constructed. See, for example. L. R. Klein and A. S. Goldberger, An Econometric Model of the United States 1929-1952, (Amsterdam: North-Holland, 1955).

The situation can be summed up. Economic theory postulates a set of variable quantities in the economy, Y, whose formation is to be explained or predicted. There exists a finite number of variables, Z, and a set of functions, F, such that $Y = F(Z)$. Among Z may be a set of variables representing factors not explicitly considered by the theory. Neither Y nor Z may be capable of exact measurement or of measurement at all. The problem facing the econometrician in building models is to give explicit form to F, to choose the observations to correspond to a set of Y and Z and to estimate the model--that is, to calculate from the observations of Y and Z a set of numbers, $\hat{\alpha}$, to approximate the parameters, α, of F.

The most widely discussed statistical formulation for systems of simultaneous equations in econometrics is the Cowles-Commission, general, linear model:[4]

$$BY_t + \Gamma Z_t = U_t \tag{1.1}$$

$$B = (B^{-1})^{-1} \tag{1.2}$$

$$E(U_t) = 0 \qquad \text{(all t)} \tag{1.3}$$

[4] See especially J. Marschak, "Statistical Inference in Economics: An Introduction," and T. C. Koopmans, H. Rubin and R. B. Leipnik, "Measuring the Equation Systems of Dynamic Economics," both in T. C. Koopmans (ed.), op. cit. By a model we shall mean a set of mathematical equations involving random variables, mathematical variables and parameters. We shall adopt the restricted meaning for a linear model of a model linear in the parameters, the random variables and the mathematical variables. Cf. Franklin A. Graybill, An Introduction to Linear Statistical Models, (New York: McGraw-Hill, 1961), pp. 96-97.

$$(1.4) \qquad E(U_t U_m') = \Sigma = (\Sigma^{-1})^{-1} \qquad t = m$$
$$= 0 \qquad t \neq m$$

$$(1.5) \qquad E(U_t Z_m') = 0 \qquad t \geqq m$$

where Y_t is a vector of values for the g endogenous variables in the observation numbered t,[5] Z_t the t^{th} observation of k predetermined variables, U_t a vector of g unobservable stochastic disturbances at t and B and Γ are matrices of coefficients. (The operator E denotes taking the mathematical expectation.) Sometimes it is further specified that U_t, in addition to being independent of values of itself at other observations and of values of the predetermined variables and having finite moments at least up to the second, has the multivariate normal distribution:

$$(1.6) \qquad P(U_t) = (2\Pi)^{-1/2g} \, |\Sigma|^{-1/2} \exp - \frac{1}{2} U_t' \Sigma^{-1} U_t \quad .$$

Before one can estimate this model, further restrictions have to be placed on B and Γ to assure that it is identified.[6]

---

[5] t need not have any sequential meaning, though it may have, for example, if the model is supposed to apply to observations in different points of time. The form of (1.5) as being for $t \geqq m$ rather than as being for all t and m makes sense, of course, only if the numbering sequence for the observations is not arbitrary.

[6] A structure of a model is identified if from the *a priori* knowledge embodied in the model (including normalization rules) and a knowledge of the parameters of the likelihood function of the observations one can deduce the parameters of the structure of the model uniquely. See, for example, Tjalling C. Koopmans, "Identification Problems in Economic Model Construction," in William C. Hood and T. C. Koopmans (eds.), *op. cit.*, Chapter II.

Even if all the steps needed to go from an economic theory to an econometric model are taken, a number of problems remain. First, the specification of the model (1.1)-(1.6) is fairly restrictive and may not be met by the processes generating economic data. Second, the properties of the estimators of structures of (1.1)-(1.6) are largely unknown. Not only does this make the choice of one estimator from the several competing possibilities difficult, it also leaves the nature and value of estimated econometric models uncertain.

2. Limitations of the Standard Model

Most of the restrictions implied by the model (1.1)-(1.6) are of dubious validity in economic studies, with the possible exception of (1.2) and (1.3). (Non-zero means for the disturbance terms can be taken care of by the constants of the equations included in the Γ-matrices.) Insofar as the assumptions are not met, the considerations which led to the development of techniques for estimating structures from this model may not be valid and use of such techniques becomes a rather dubious undertaking.

In many ways the most questionable of the assumptions is that the underlying model is sufficiently well known to justify the explicit form of the equations (1.1) both in the variables to be included and in the restrictions on the B- and Γ-matrices that have to be imposed to achieve identification. These are certainly a good deal more precise than economic theory can usually justify. Furthermore, when the explanatory variables of a model are correlated with each other, it is quite likely that misspecification of the form of the equations, and especially of the variables to be included in them, will not be detected in the course of estimating the model. Unless there are strong theoretical reasons for believing that the forms of the equations are correct, the validity (and usefulness) of any model may be extremely questionable.

Additive disturbances are another questionable assumption. The lack of precision in a relationship may well depend on the sizes of the other variables in it. This might make multiplicative disturbances seem a better specification. Or the behavior patterns themselves may be subject to variations in the sense that the importance given to different variables in decision-making varies from time to time or from individual to individual. If the model describes relationships supposed to hold over a period of time or to describe the aggregate or average behavior of a group of economic agents, one might prefer a specification in which the coefficients B and Γ were stochastic variables of the system.[7]

Assumption (1.4) may often be open to doubt. The assumption of constant variance and the assumption of inter-observation independence have little or no general basis in economic theory. Indeed, there may be good reasons for expecting the residuals to be inter-related in time-

[7]Cf. Leonid Hurwitz, "Systems with Nonadditive Disturbances," and Herman Rubin, "Note on Random Coefficients," being respectively Chapters XVIII and XIX of T. C. Koopmans (ed.), op. cit.

series data and for them not to have the same distributions at all observations both in time-series and in cross-section data. There is nothing in economic theory to make one presume that the sizes of the disturbances are independent of the size or wealth (say) of the entities or events which the observations record or of other values in a temporal process.

Assumption (1.5) is also of very questionable validity in economic models. To begin with, it rules out errors of measurement in the explanatory variables. The assumption that the exogenous data are measured without error is hardly descriptive of economic data.[8] Even if the exogenous data were measured exactly, it is frequently the case that the data do not represent exactly the quantities used in economic theory because of different definitions and classifications used in collecting the data. But the inappropriateness of assumption (1.5) runs deeper than the problems of errors in data.

The assumption expressed in (1.5) is that the influence of the exogenous variables on the endogenous ones occurs only through channels explicitly recognized by the structural equations and that the factors influencing the latent variables have no effect on the endogenous variables. There is nothing in economic theory to make one presume that the distribution of the disturbances should be completely independent of the predetermined variables. At the heart of economic

---

[8] Cf. Oskar Morgenstern, <u>On the Accuracy of Economic Observations</u>, (Princeton: Princeton University Press, 1963). The specification does permit errors of measurement of the endogenous variables provided such errors obey assumptions (1.4) and (1.5).

theory lies a belief in great and widespread interdependence among the variables in the economy. While in developing theories and in building models it is convenient to focus attention on a few important relationships, there may still be other, less important connections among the variables through variables excluded from the simpler models. The exogenous variables of most econometric models can hardly be considered exogenous to the whole economic system. The factors influencing these variables may have a direct, although possibly minor, influence on the endogenous variables. In view of these considerations, it seems likely that some of the unexplained _structural_ influences on the endogenous variables should be related to the predetermined variables. In building simultaneous-equation models, it is supposed that only one of the equivalent ways in which the equations may be written represents the behavior of the economic actors. Other connections between the exogenous and endogenous variables may be present which are _not_ supposed to be described by the structural equations (1.1). A correct specification of (1.1) may entail a violation of (1.5). Specification (1.5), on which the available estimators rest, may be both arbitrary and invalid.[9]

Despite the difficulties involved with using the model (1.1)-(1.5) in econometric work, a fair number of techniques has been proposed for estimating the coefficients of this model and it is virtually the only simultaneous-equation model for which estimators do exist. A good deal of work has been done on the theoretical, asymptotic properties of

[9] See Fritz C. Holtz, _Economic Shock Models_, (Oslo: Norwegian Universities Press, 1962), pp. 10-24, for a presentation of the view that assumption (1.5) is inappropriate.

the estimates arising from the various techniques. A smaller group of studies has started the exploration of their small-sample properties both by mathematical analysis and by Monte Carlo experiments. However, it is fair to say that as yet little is known about which technique is best for estimating economic models or how adequately the techniques perform in various situations. This is both because of the sparseness of our knowledge of their small-sample properties and because it is likely that economic data are generated by processes substantially different from those assumed in the derivation and justification of the techniques and in most of the work done on them. This leads to the dilemma that economic theory suggests the necessity of estimating simultaneous-equation systems, but that such systems, when estimated, are of doubtful use.

There could be several ways of attacking the problems of simultaneous-equation estimation. For the problems arising from the model not being appropriate, one could try to develop estimators for other models, or one can examine how sensitive the estimators are to violations of the assumptions, asking, in effect, whether the use of inappropriate methods matters very much. This latter alternative is important not only because other methods are not at present available, but also, and more importantly, because frequently one would not know the correct specification and because it seems likely that methods based on different specifications would require more specific *a priori* knowledge than do present ones. To attack the problems of which estimator to choose and its small-sample properties, one can try to derive analytically the distributions of the estimators or one can examine their properties using Monte Carlo techniques. Examination of the general characteristics of the estimators may also give indications of their relative merits.

3. Plan of This Study

This study aims to cast more light on the small-sample properties of various estimators of the linear model and to examine the sensitivity of the estimators to violations of the assumptions of the Cowles-Commission model. The main tools used in examining these questions are a number of Monte Carlo experiments. The study proceeds as follows:

Chapter II describes the estimating methods under study. It also discusses the results of previous Monte Carlo studies. Chapter III describes the Monte Carlo experiments performed in this study. Chapters IV, V, and VI present the results of experiments which used the standard model to generate the data. They try to throw light on the small-sample properties of the estimators and to indicate their reliability in cases where the Cowles-Commission model is appropriate. Chapter VII presents the results of experiments where the assumptions are violated. This part of the study seeks to indicate how serious violations of the assumptions are for simultaneous-equation estimation and to indicate which of the methods studied performs best in these situations. Chapter VIII discusses the results of one type of violation of the assumptions not covered in Chapter VII, namely misspecification of the form of the structural equations, particularly in the variables included in the equations.

The experiments conducted far from exhaust the aspects of the properties of the estimators which are of interest. However, it is hoped that they shed some useful light on the small-sample properties of the estimators. The findings of the study are summarized in Chapter IX.

CHAPTER II

The Estimators of Simultaneous Equations

1. Introduction

There is a number of estimators of the model (1.1)-(1.5). In this chapter we review the methods of estimation which will be investigated in this study. The review is short, incomplete and non-rigorous since good, over-all presentations of the estimators are available,[10] as well as many individual articles on their properties. The chapter then outlines the findings about the small-sample properties of the estimators made by the Monte Carlo experiments conducted by others.

Throughout the chapter we shall use the following notation:

T - the number of observations

g - the number of endogenous variables (and equations)

K - the number of predetermined variables

g_m - the number of endogenous variables included in the m^{th} equation $(m = 1,\dots,g)$

k_m - the number of predetermined variables included in the m^{th} equation

Y - the $T\times g$ matrix of observations of the endogenous variables

---

[10] J. Johnston, Econometric Methods, (New York: McGraw-Hill, 1963), Chapter 9; Arthur S. Goldberger, Econometric Theory, (New York: Wiley, 1964), Chapter 7; Gregory C. Chow, A Comparison of Alternative Estimators of Simultaneous Equations, (Yorktown Heights, New York: International Business Machines Corporation, Thomas J. Watson Research Center, Report RC-781, 1962).

Z - the $T \times K$ matrix of observations of the predetermined variables

U - the $T \times g$ matrix of (latent) structural disturbances

y_t, z_t, u_t - the t^{th} rows of Y, Z, U

y_m - the endogenous variable on which the m^{th} equation is normalized

Y_m - the other endogenous variables included in the m^{th} equation

Z_m - the predetermined variables included in the m^{th} equation

$Z_{\upsilon m}$ - the predetermined variables excluded from the m^{th} equation

B, Γ - the coefficient matrices of the system $YB' + Z\Gamma' = U$

u_m - the $T \times 1$ vector of the disturbances of the m^{th} equation

Π - the $g \times K$ matrix of the reduced-form coefficients of the system, $Y = -Z\Gamma'(B^{-1})' + U(B^{-1})'$

$$= Z\Pi' + V$$

$\Sigma = E(u_t'u_t)$

The m^{th} equation will be written:

$$YB_m' + Z\Gamma_m' = u_m$$

or, equivalently,

$$y_m = Y_m\beta_m' + Z_m\gamma_m' + Z_{\upsilon m}\gamma_{\upsilon m}' + u_m$$

or

$$y_m = Y_m\beta_m' + Z_m\gamma_m' + u_m \quad .$$

Parts of the reduced-form will be written:

$$[y_m \quad Y_m] = Z\Pi'_m + [v_m \quad V_m]$$

$$= Z_m \Pi'_{mm} + Z_{\supset m} \Pi'_{m \supset m} + [v_m \quad V_m]$$

and $$Y_m = Z\Pi'_{m_1} + V_m \quad .$$

We shall assume that the models to be estimated obey assumptions (1.1)-(1.5). Each equation, m, of the system of g structural equations,

$$(2.1) \qquad YB' + Z\Gamma' = U \quad ,$$

to be estimated is normalized on one of the endogenous variables.[11] The models discussed will be assumed to be identified. The restrictions placed on them will be the specification that at least $(g-1)$ structural coefficients in each equation be zero.[12]

---

[11] There is an important conceptual difference between estimators which use the normalization rule as a constraint in estimating the equations and those for which it can be imposed after estimation since it affects the estimates only as a scalar transformation of all estimated coefficients. See Gregory C. Chow, *op. cit.*, for a discussion of the importance of normalization rules. The difference is conceptual in that one can calculate the estimates for all methods with the normalization rule specified *ab initio*.

[12] This assures that the order condition for identification will be met. See Tjalling C. Koopmans, "Identification Problems in Economic Model Construction," in William C. Hood and Tjalling C. Koopmans (eds.), *Studies in Econometric Method*, (New York: Wiley, 1953), p. 38. Implicit in the discussion is the assumption that the rank condition is also met.

2. The Estimators

Direct Least Squares. The simplest method for estimating the equations (2.1) is to estimate each equation using the formulae of classical least-squares one equation at a time. In any equation, m, the variable on which the equation is normalized is taken as the dependent variable:

$$y_m = Y_m\beta'_m + Z_m\gamma'_m + u_m \quad . \tag{2.2}$$

β_m and γ_m are then estimated by the standard formula[13] to minimize the sum of squares of the residuals from the estimated equation. This method of estimating the structural equations will be referred to as Direct Least-Squares (DLS). Equations (2.2) violates the assumptions of the classical regression model since the regressors Y_m are not, in general, independent of the disturbances, u_m. The DLS estimates are in general inconsistent. They may also be biased.[14] The usual formulae for the standard errors of the estimates will not, in general, lead to estimates with the properties of the usual least-squares estimates of the variance of the estimates of the coefficients about the true values. They can, of course, still be calculated.

13

$$\begin{bmatrix} \hat{\beta}'_m \\ \hat{\gamma}'_m \end{bmatrix} = \begin{bmatrix} Y'_mY_m & Y'_mZ_m \\ Z'_mY_m & Z'_mZ_m \end{bmatrix}^{-1} [Y_m \quad Z_m]' y_m \quad .$$

14 Note that since the inverse of a random matrix enters into the formulae for the estimates, the first moment of the distribution of the estimates may not exist and so the bias may be undefined. See Maurice G. Kendall, The Advanced Theory of Statistics, (London: Griffin, 1950), vol. 3, pp. 265-269. This may also be true for the other estimates of the structural equations discussed in this section.

Least-Squares Estimation of the Reduced Form. Least-squares can be used appropriately in simultaneous-equation models for estimating the reduced form:

$$(2.3) \qquad Y = -Z\Gamma'(B^{-1})' + U(B^{-1})' = Z\Pi' + V \quad .$$

The disturbances V are independent of the regressors Z (by assumption 1.5). However, the specification that certain structural coefficients be zero will usually imply restrictions on the reduced form. Since

$$\Pi = -B^{-1}\Gamma \quad ,$$

$$B_m\Pi = -B_mB^{-1}\Gamma = -\Gamma_m$$

and in particular,

$$(2.4) \qquad (1 - \beta_m)\Pi_{m\upsilon m} = \gamma_{\upsilon m} = 0 \quad .$$

The dimensions of $\Pi_{m\upsilon m}$ are $(g_m) \times (K - k_m)$ and by assumption, $(K - k_m) \geqq (g_m - 1)$. For a unique solution of (2.4) the rank of $\Pi_{m,\upsilon m}$ must equal $(g_m - 1)$. If $(K - k_m) > (g_m - 1)$, unrestricted estimates of Π, such as those obtained by using least-squares, would not usually lead to the rank of $\Pi_{m,\upsilon m}$ being $g_m - 1$. For this reason, the method can be referred to as least-squares with no restrictions. In this study, however, it will be called least-squares applied to estimate the reduced form directly (LSRF).

Two features of LSRF should be noted. First, LSRF minimizes not only the sum of squares of the residuals of any one equations, it also minimizes the generalized variance:[15]

$$(2.5) \qquad |\hat{V}'\hat{V}| = |(Y - Z\Pi')'(Y - Z\Pi')| \; .$$

That is, the estimate of Π minimizing (2.5) is the LSRF estimate. Second, if V has the multi-variate normal distribution, LSRF estimates of Π are the unrestricted maximum-likelihood estimates if Z is exogenous.[16]

Two-Stage Least-Squares. The easiest consistent estimator of (1.1)-(1.5) to compute can be derived from a number of principles of varying degrees of sophistication which have led to its being given a variety of names. Here we refer to it as Two-Stage Least-Squares (2SLS).

2SLS attempts to estimate the system (2.1) by estimating each equation separately as (2.2). DLS breaks down as an estimator because Y_m and u_m are not independent. If we rewrite (2.2) as

$$(2.6) \qquad \begin{aligned} y_m &= (Y_m - V_m)\beta'_m + Z_m\gamma'_m + u_m + V_m\beta'_m \\ &= Z\Pi'_{m_1}\beta'_m + Z_m\gamma'_m + u_m + V_m\beta'_m \end{aligned}$$

[15] See Arthur S. Goldberger, op. cit., p. 320.

[16] See Tjalling C. Koopmans and William C. Hood, "The Estimation of Simultaneous Linear Economic Relationships," in Hood and Koopmans (eds.), op. cit., pp. 151-155.

we would have an equation suitable for least-squares estimation since Z is assumed uncorrelated with u_m and V_m. Unfortunately V_m is unknown. The essence of 2SLS is to replace V_m in (2.6) by their LSRF estimates, $\hat{V}_m$. 2SLS estimates

$$(2.7) \qquad y_m = (Y_m - \hat{V}_m)\beta'_m + Z_m\gamma'_m + q_m$$

where

$$(2.8) \qquad q_m = u_m + \hat{V}_m\beta'_m$$

by minimizing $q'_m q_m$. The 2SLS estimator[17] is:

$$(2.9) \qquad \begin{bmatrix} \hat{\beta}'_m \\ \hat{\gamma}'_m \end{bmatrix} = [(Y_m - \hat{V}_m \quad Z_m)'(Y_m - \hat{V}_m \quad Z_m)]^{-1}[Y_m - \hat{V}_m \quad Z_m]'y_m$$

$$= \begin{bmatrix} Y'_mY_m - \hat{V}_m\hat{V}_m & Y'_mZ_m \\ Z'_mY_m & Z'_mZ_m \end{bmatrix}^{-1} [Y_m - \hat{V}_m \quad Z_m]'y_m \quad .$$

---

[17] 2SLS can be derived in a number of other ways. In particular, it can be considered an instrumental variable estimator with Z_m serving as their own instruments and $(Y_m - \hat{V}_m) = Z(Z'Z)^{-1}Z'Y_m$ as the instruments of Y_m. See Goldberger, op. cit., pp. 331 ff. 2SLS can also be derived as the estimator minimizing the difference of the sum of squares of the residuals of the regression of the vector $[y_m \quad Y_m] \begin{bmatrix} 1 \\ -\beta'_m \end{bmatrix}$ on Z_m from the sum of squares of the regression of the same vector on Z. See Johnston, op. cit., pp. 261-262. Finally, 2SLS can be derived as an (approximate) application of Aitken's generalization of least squares. See H. Theil, Economic Forecasts and Policy, (Amsterdam: North-Holland, 1961), pp. 337-338. 2SLS was derived by H. Theil and independently by R. L. Basmann. For discussion see H. Theil, op. cit., pp. 225-232, 334-355; and R. L. Basmann, "A Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation," Econometrica, vol. 25 (1957), pp. 77-83.

2SLS can be shown to be consistent[18] and the asymptotic variance of the 2SLS estimates is

(2.10)
$$\sigma_m \begin{bmatrix} Y_m'Y_m' - \hat{V}_m'\hat{V}_m & Y_m'Z_m \\ Z_m'Y_m & Z_m'Z_m \end{bmatrix}^{-1}$$

where σ_m can be consistently estimated by

(2.11)
$$\hat{\sigma}_m = \frac{1}{T}(y_m - Y_m\hat{\beta}_m' - Z_m\hat{\gamma}_m')(y_m - Y_m\hat{\beta}_m' - Z_m\hat{\gamma}_m') \quad .$$

(2.10) and (2.11) provide a formula which may be used to try to calculate the standard errors, even in finite samples--though then they may not have the usual properties of standard errors. One difficulty is to be noted: σ_m may be estimated consistently by $\frac{T\hat{\sigma}_m}{T-g_m-k_m+1}$ or with other adjustments. By analogy to the pure single-equation case, where $\hat{\sigma}_m$ is a biased estimator of σ_m, it might be thought appropriate to adjust (2.11) for "degrees of freedom" before inserting the estimate of σ_m in (2.10) to calculate the standard errors. Such an adjustment is <u>not</u> derived from the finite sample distribution of the 2SLS estimates--but it may make the standard errors more useful, as is indicated in Chapter IV.

---

[18] $(Y_m - \hat{V}_m)$ is independent of u_m only asymptotically. 2SLS may be biased. As with DLS, the moments of the finite-sample distribution of the 2SLS estimates may not exist and so the bias would be undefined. R. L. Basmann, "A Note on the Exact Finite Sample Frequency Functions of the Generalized Classical Linear Estimators in Two Leading Over-Identified Cases," <u>Journal of the American Statistical Association</u>, vol. 56 (1961), pp. 619-636, has proved that the 2SLS estimates do not have finite moments in certain cases.

The k-Class Estimators. To obtain the next group of single-equation estimators, we start with (2.9) and insert a parameter, k, into the estimating formula:

$$(2.12)\quad \begin{bmatrix} \hat{\beta}'_m \\ \hat{\gamma}'_m \end{bmatrix}_k = \begin{bmatrix} Y'_mY_m - k\hat{V}'_m\hat{V}_m & Y'_mZ_m \\ Z'_mY_m & Z'_mZ_m \end{bmatrix}^{-1} [Y_m - kV_m \quad Z_m]'\, y_m \,.$$

This defines the k-class estimators.

It can be shown that if $P\lim(k-1) = 0$ then $(\hat{\beta}_m, \hat{\gamma}_m)_k$ are consistent estimators of (β_m, γ_m). If $P\lim T^{1/2}(k-1) = 0$ then the asymptotic covariance matrix of $(\hat{\beta}_m, \hat{\gamma}_m)_k$ is the same as the 2SLS asymptotic variance,[19] (2.10). This suggests a formula, based on the asymptotic distribution, for estimating the standard errors of the k-class estimates based on

$$(2.13)\quad (\hat{\sigma}_m)_k \begin{bmatrix} Y'_mY_m - k\hat{V}'_m\hat{V}_m & Y'_mZ_m \\ Z'_mY_m & Z'_mZ_m \end{bmatrix}^{-1}$$

where

$$(\hat{\sigma}_m)_k = \frac{1}{T}\left[y_m - [Y_m \quad Z_m]\begin{bmatrix} \hat{\beta}'_m \\ \hat{\gamma}'_m \end{bmatrix}_k\right]' \left[y_m - [Y_m \quad Z_m]\begin{bmatrix} \hat{\beta}'_m \\ \hat{\gamma}'_m \end{bmatrix}_k\right] .$$

[19] This follows from theorems of A. L. Nagar found in Anirudh Lal Nagar, Statistical Estimation of Simultaneous Economic Relationships, (Rotterdam: private, 1959), Chapters III and IV. Note that DLS is a k-class estimator with $k = 0$.

Important work on the approximate small-sample bias and variance of the k-class estimators has been done by A. L. Nagar. Under the assumption that k and Z are nonstochastic, Nagar[20] derived the bias of k-class estimators to the order T^{-1} and their second-moment matrix to the order T^{-2}.

Define $k = 1 + \frac{p}{T}$. Nagar showed that the bias of the k-class estimators (to the order T^{-1}) is:

$$E\left[\begin{bmatrix} \hat{\beta}'_m \\ \hat{\gamma}'_m \end{bmatrix}_k - \begin{bmatrix} \beta'_m \\ \gamma'_m \end{bmatrix}\right] = \left[\frac{p-(K-g_m-k_m)}{T}\right]\begin{bmatrix} Y'_mY_m - V'_mV_m & Y'_mZ_m \\ Z'_mY_m & Z'_mZ_m \end{bmatrix}^{-1}\begin{bmatrix} E(Y'_mu_m) \\ E(Z'_mu_m) \end{bmatrix}.$$

It follows that if $k = 1 + \frac{K-g_m-k_m}{T}$, the k-class estimator is unbiased to the order T^{-1}. This estimator will be referred to as unbiased-k (UBK). If $(K-g_m-k_m) = 0$, that is, if there is one over-identifying restriction on equation m, 2SLS is the UBK estimator.

In the course of deriving the approximate second moment of the k-class estimates (to the order T^{-2}), Nagar derived an estimator for which the determinant of the second moment was minimized--the minimum-second moment estimator to the order T^{-2} (MSM). The value of k for this estimator depends on (unknown) population parameters and would have to be estimated.

[20] Nagar, op. cit. To obtain the approximate bias and covariance Nagar had to impose conditions to avoid the problem of infinite moments so that his approximations would go through. The conditions are (essentially) that the stochastic part of the model be relatively unimportant. Whether the conditions would be met by the data of econometric models is, of course, a moot point.

Limited-Information Maximum-Likelihood. One other member of the k-class estimators is of particular interest. This estimator has k equal to the smallest root of

$$(2.14) \qquad |V_r'V_r - \ell V_L'V_L| = 0$$

where

$$V_r = [y_m \quad Y_m] - Z_m(Z_m'Z_m)^{-1}Z_m'[y_m \quad Y_m]$$

and

$$V_L = [y_m \quad Y_m] - Z(Z'Z)^{-1}Z'[y_m \quad Y_m] \quad .$$

If U is normally distributed, this is the maximum likelihood estimator of the single equation (2.2). This estimator[21] will be referred to as Limited-Information Maximum-Likelihood (LIML).

It can be shown that for LIML $P \lim T^{1/2}(k-1) = 0$ so that LIML is a consistent estimator with (asymptotic) variance-covariance matrix (2.10).

[21] See T. W. Anderson and H. Rubin, "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," Annals of Mathematical Statistics, vol. 20 (1949), pp. 46-63; and T. W. Anderson and H. Rubin, "The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations," Annals of Mathematical Statistics, vol. 21 (1950), pp. 570-583. See also Koopmans and Hood, *op. cit.*; and Goldberger, *op. cit.*, pp. 338-345. LIML is also the estimator minimizing $|(v_m \quad V_m)'(v_m \quad V_m)|$ subject to (2.4), and the estimator minimizing the ratio of the sum of squares of the residuals of the regression of $[y_m \quad Y_m]\upsilon_m$ on Z_m to the sum of squares of the residuals of $[y_m \quad Y_m]\upsilon_m$ regressed on Z. (υ_m is a vector of coefficients. The normalization rule leading to the coefficients $(1 - \beta_m)$ is not an intrinsic part of LIML.)

Full-Information Maximum-Likelihood. Under assumptions (1.1)-(1.6), the (logarithmic) likelihood function of the sample Y,Z is:

$$(2.15)\quad L(B,\Gamma,\Sigma) = Q + T\log|B| - \frac{T}{2}\log|\Sigma| - \sum_{t=1}^{T}(y_tB' + z_t\Gamma')\,\Sigma^{-1}(y_tB' + z_t\Gamma')$$

where Q is a constant. Maximization of (2.15) with respect to B, Γ and Σ leads to the estimator of (2.1) known as Full-Information Maximum-Likelihood (FIML).[22]

Maximimzation of (2.15) requires the solution of nonlinear equations in B and Γ and cannot be done directly. Iterative methods are available, however, so that the estimates can be calculated.[23]

It can be shown that FIML is consistent, efficient and asymptotically normally distributed. Asymptotically, the variance-

[22] See T. C. Koopmans, H. Rubin and R. B. Leipnik, "Measuring the Equation Systems of Dynamic Economics," in Tjalling C. Koopmans (ed.), Statistical Inference in Dynamic Economic Models, (New York: Wiley, 1950), pp. 53-237. Other useful presentations are in Goldberger, op. cit., pp. 352-356; Chow, op. cit., pp. 14-22; and T. J. Rothenberg and C. T. Leenders, "Efficient Estimation of Simultaneous Equation Systems," Econometrica, vol. 32 (1964), pp. 57-76. FIML can also be derived as the estimator minimizing the generalized variance $|V'V| = |[Y+Z\Gamma'(B^{-1})']'[Y+Z\Gamma'(B^{-1})']|$. See Goldberger, loc. cit. It can also be derived as the estimator minimizing Σ subject to a suitable normalization rule. See Chow, loc. cit.

[23] Computation of FIML estimates is discussed in Appendix A. See also Koopmans, Rubin and Leipnik, op. cit.; Herman Chernoff and Nathan Divinsky, "The Computation of Maximum-Likelihood Estimates of Linear Structural Equations," in Hood and Koopmans (eds.) op. cit., Chapter X; and T. M. Brown, "Simplified Maximum Likelihood and Comparative Structural Estimates," Econometrica, vol. 27 (1959), pp. 638-653.

covariance matrix of the FIML estimates of B, Γ can be shown to be $-L_F^2$, where L_F^2 is the matrix of second partial derivatives of the likelihood function about the FIML estimates. This suggests obtaining standard errors for FIML in finite samples from the second derivatives of the likelihood function about the FIML estimates.

Three-Stage Least-Squares. The final method to be outlined here seems to rest on much shakier grounds than FIML. It is, however, a full-model method and somewhat easier to compute. It is known as Three-Stage Least-Squares (3SLS).[24]

Premultiplying (2.1) by Z', the equations can be written:

$$\begin{aligned} Z'y_1 &= Z'Y_1\beta_1' + Z'Z_1\gamma_1' + Z'u_1 \\ &\vdots \\ Z'y_g &= Z'Y_g\beta_g' + Z'Z_g\gamma_g' + Z'u_g \end{aligned}$$

or

$$(2.16)\quad \begin{bmatrix} Z'y_1 \\ Z'y_2 \\ \vdots \\ Z'y_g \end{bmatrix} = \begin{bmatrix} Z'X_1 & 0 & \cdots & \\ 0 & Z'X_2 & \cdots & \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & Z'X_g \end{bmatrix} \begin{bmatrix} \delta_1' \\ \delta_2' \\ \vdots \\ \delta_g' \end{bmatrix} + \begin{bmatrix} Z'u_1 \\ Z'u_2 \\ \vdots \\ Z'u_3 \end{bmatrix}$$

[24] Three-Stage Least Squares is presented in Arnold Zellner and H. Theil, "Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations," Econometrica, vol. 30 (1962), pp. 54-78. See also Albert Madansky, "On the Efficiency of Three-Stage Least-Squares Estimation," Econometrica, vol. 32 (1964), pp. 51-57; and J. D. Sargan, "Three-Stage Least-Squares and Full Maximum Likelihood Estimates," Econometrica, vol. 32 (1964), pp. 77-81.

The disturbances of (2.16) are heteroskedastic. This might suggest the application of Aitken's generalized least-squares to the complete system of equations (2.16). Unfortunately, the application would require knowledge of the parameters Σ. 3SLS replaces Σ by its 2SLS estimates.[25] 3SLS can be shown to be consistent and asymptotically to have the same distribution as the FIML estimates.

3SLS has the weakness that it is necessary to invoke the 2SLS estimates of Σ at one point, rather than estimating Σ simultaneously with the other parameters. It does however, provide for the simultaneous estimation of all the structural coefficients and is easier to compute than FIML.

3. The Choice of Estimator

Choice of the most suitable estimator for econometric models could be based on a comparison of the small-sample properties of the estimators--if their small-sample distributions were known. Unfortunately the exact small-sample distributions have been derived in only a few special cases and then it is the distribution of only one estimator which is known. To judge the estimators we can fall back on three ways of comparing them: comparison of their asymptotic distributions, comparison of the principles from which they are derived and comparison of their performances in samples of estimates taken from the small-sample distributions.

[25] 3SLS and its computation are discussed in more detail in Appendix A.

<u>Asymptotic Properties</u>. Comparison of the asymptotic properties leads to a fairly clear-cut ranking of the methods. All methods except DLS are asymptotically unbiased and consistent. The full-model methods are more efficient than the consistent single-equation methods. DLS does retain its minimum-variance property, but this is now about its biased asymptotic mean. Otherwise, the asymptotic properties help little in the choice of an estimator. All the consistent k-class estimators discussed in section 2 converge to the same asymptotic distribution. All the full-model methods have the same limit. Thus it is impossible to choose among 2SLS, UBK, MSM and LIML or between FIML and 3SLS on the basis of their asymptotic distributions.

<u>Principles of Derivation</u>. Comparisons of the principles from which the estimators can be derived may help to guide the choice of an estimator. FIML is the maximum-likelihood estimator, and takes full account of the simultaneity of the model and can be regarded as the correct generalization of least-squares.[26] 3SLS recognizes the simultaneity but is a less happy generalization of least-squares, <u>especially</u> since it substitutes 2SLS estimates of Σ at a crucial stage when it might better use estimates of Σ to be derived simultaneously with those of the structural coefficients. The single-equation methods ignore the simultaneity of the model except insofar as they use all the predetermined variables. LIML is a maximum-likelihood estimator. Both LIML and 2SLS

[26] This is the thesis of Chow, <u>op. cit.</u>, which is an extended discussion of the estimators based on this principle of derivation.

are generalizations of least-squares, but in different ways.[27] DLS ignores the simultaneity of the model entirely and may be considered a poor application of least-squares.

In terms of simplicity of principle or of computation, the ranking of the methods is quite different. 2SLS, LIML, 3SLS and FIML (in that order) are more subtle methods than DLS in the sense that neither is the direction of minimization so evident nor are the computations as easy. It may be doubted if this subtlety is warranted. First, the computations are more precarious.[28] Second, if the data of econometric models do not conform fully to the model (1.1)-(1.5), it may not be of much advantage to take account of simultaneity in sophisticated ways while ignoring the other features of the data. Finally, if little confidence can be placed in the theoretical specification of an econometric model, it may not be particularly appropriate to estimate it by a complicated method. If the object is really to explore whether there are apparent relationships between economic series or how different combinations of observations seem to fit together, a simple method such as DLS, or the simple principle of DLS, may lead to the most appropriate descriptive statistics. Without firmer theoretical and empirical knowledge

[27] See Chow, op. cit., for the view that LIML is a superior generalization. See R. L. Basmann, "An Expository Note on Estimation of Simultaneous Structural Equations," Biometrics, vol. 16 (1960), for the view that 2SLS is superior.

[28] LIML may be more subject to computational error than 2SLS for two reasons. First, to calculate it we need to compute the characteristic root of (2.12). Second, as a k-class estimator, the matrix inversion in (2.9) may be more difficult since there are indications that the matrix passes through singularity at values of k-class to the LIML value. FIML and 3SLS both require the inversion of larger matrices than the k-class estimators.

of the nature and generating processes of economic data, it may be doubted whether use of subtle generalizations of regression are really worthwhile.

Sampling Techniques. The third way of trying to evaluate the methods, inference about their properties from samples of estimates, has been taken by a number of studies. The results of these Monte Carlo studies are sketched in the next section. Here we discuss briefly the techniques.

The basic techniques of a Monte Carlo study are easily described.[29] A sample of data from a specific statistical model is first obtained. From this sample, the parameters of the model are estimated. This process is repeated using new samples of data to obtain further estimates until a sample of estimates has been obtained. (These repetitions of the process are referred to as the replications of the study.) Inferences about the distributions of the estimates of the parameters of the model are made from the sample of estimates.

Unlike most problems where the techniques of statistical inference are used, the problems investigated by Monte Carlo studies are, at least in principle, open to deductive, mathematical analysis. Such

[29] Although there is merit in Robert Summers' suggestion that the technique be called distribution sampling, limiting the term "Monte Carlo study" to more refined investigations (Robert Summers, "A Capital Intensive Approach to the Small Sample Properties of Various Simultaneous Estimators," typescript, p. 12, and H. F. Trotter and J. W. Tukey, "Conditional Monte Carlo for Normal Samples," in H. A. Meyer (ed.), Symposium on Monte Carlo Methods (New York: Wiley, 1956), p. 64), we use the term in the wider sense since this is how such studies investigating the simultaneous-equation estimators have usually been described.

analysis, by providing exact, certain knowledge of the distributions of the estimates, would render the results of Monte Carlo studies redundant. In addition, the inferences of a Monte Carlo study can never be held with certainty but only hold up to (prescribed) levels of confidence. However, the technique can provide answers to questions which seem intractible to mathematical analysis, may provide such answers in an easier and quicker fashion than would analysis, and may suggest which problems most cry for analysis and give an indication of the ways analysis should proceed.[30] In studying simultaneous-equation estimators, Monte Carlo studies have gone some way in overcoming the drawback suffered by not knowing the small-sample distributions and supplementing the considerations on which to base the choice of an estimator for econometric models.

One other point about Monte Carlo studies should be stressed. As in any statistical investigation, the results rest on validity of assumptions about the distributions being investigated which are required for tests used in the investigation. In particular, since there is a possibility that the lower-order moments of the distributions of the estimates of systems of simultaneous equations do not exist, considerable skepticism about results which assume the existence of these moments may be appropriate. However, it is not necessary to make such assumptions to be able to conduct useful Monte Carlo studies since there are statistical tests which do not rest on the assumption that finite moments exist.

[30] Cf. R. L. Basmann, On the Exact Finite Sample Distribution of Generalized Classical Linear Structural Estimators, (Santa Barbara: General Electric Company, 1960).

4. Results of Previous Monte Carlo Studies

There have been several Monte Carlo studies on the small-sample properties of the simultaneous-equation estimators.[31] There are, of course, many differences between the separate studies. They investigated different sets of methods, used different models and employed different types of data for their experiments.

Each of the studies had a number of interesting specific findings.[32] In addition, they suggested some more general features of the small-sample performances of the estimators. DLS usually had larger biases than any one of the other estimators. On the other hand, the DLS

---

[31] Harvey M. Wagner, "A Monte Carlo Study of Estimates of Simultaneous Linear Equations," *Econometrica*, vol. 26 (1959), pp. 117-133; A. L. Nagar, "A Monte Carlo Study of Alternative Simultaneous-Equation Estimators," *Econometrica*, vol. 28 (1960), pp. 573-590; R. L. Basmann, *On the Exact Finite Sample Distribution of Generalized Classical Linear Structural Estimators*, (Santa Barbara: Technical Military Planning Operation, General Electric Company, Report SP-91, 1960); R. E. Quandt, *Some Small Sample Properties of Certain Structural Equation Estimators*, (Princeton: Econometric Research Program, Princeton University, Research Memorandum 46, 1962); R. E. Quandt, *On Certain Small Sample Properties of k-class Estimators*, (Princeton: mimeographed, 1963); Robert Summers, *A Capital Intensive Approach to the Small Sample Properties of Various Simultaneous Equation Estimators*, (typescript); G. W. Ladd, "Effects of Shocks and Errors in Estimation: An Empirical Comparison," *Journal of Farm Economics*, vol. 38 (1956), pp. 485-495; W. A. Neiswanger and T. A. Yancy, "Parameter Estimates and Autonomous Growth," *Journal of the American Statistical Association*, vol. 54 (1959), pp. 389-402.

[32] These are well summarized in Johnston, *op. cit.*, Chapter 10, though Johnston does not cover Quandt's experiments.

estimates usually had smaller dispersions about their central tendencies.[33] The consistent estimators all seemed to be fairly close to each other in performance in the experiments in which more than one of them were investigated. Several of the studies indicated that the standard errors of the consistent techniques seemed to give fairly reliable indications of the dispersions of the estimates about the true values when the structural equations were correctly specified. On the other hand, the DLS standard errors were not reliable except as indicators of the dispersions of the estimators about their (biased) central tendencies.

The Monte Carlo studies which have been conducted shed a good deal of light on the small-sample properties of the simultaneous-equation estimators. There is, however, a large number of questions which they have left unanswered. Partly this is because of limitations in the design of the experiments, partly because of the small number of studies which have been conducted, partly because the results of different experiments were not entirely in agreement with each other.

The prime drawback in the experiment designs has been the limited number of methods which have usually been compared with each other. With the exception of the Wagner-Nagar studies and the experiments conducted by Summers, only one or two consistent estimators were used in addition to DLS. These studies were necessarily silent on the

[33]These findings were indicated by the studies of Wagner, Nagar, Quandt, Basmann and Summers. It is to be noted that the smaller dispersions of DLS than other methods were not found invariably, especially when FIML was included in the experiments. Nor was the larger bias invariably found.

properties of methods not included. Only Summers investigated the performance of a full-model method.[34]

The small number of experiments which has been conducted means that many aspects which might be of interest have not been explored. The independence of the different studies makes their results difficult to compare and to use in investigating the problem which has been most extensively explored: the performances of the methods when different models, structures and sets of data are used. Other problems have received much less attention. For example, only two pairs of experiments conducted by Summers throw light on the effects of having different numbers of observations from which to estimate the structure. As for the sensitivity of the performances of the methods to incorrect specifications of the forms of the structural equations, only a few experiments conducted by Summers and Neiswanger and Yancy have investigated this wide and important area at all. How sensitive the performances of the estimators are to the validity of the stochastic assumptions under which the estimating techniques were derived has had equally limited exploration.

The results of the Monte Carlo experiments have not been wholly in agreement with each other. For example, while most experiments found DLS to have been the poorest method, this was not true for some of Quandt's

[34]These remarks are not intended as criticism of the studies. Nor will we criticize the ways in which the studies were conducted though one might well wish that other analyses of the data had been made.

experiments. Summers found that different amounts of collinearity did not substantially affect the relative performances of methods in correctly specified models; Quandt found that this did affect his results. Further, it was frequently found that the performances of the methods were different for different coefficients of the models.

In view of these short-comings to our knowledge, it would seem desirable to have further studies conducted to see if ambiguities in previous results can be cleared up, to see if the findings of other studies are specific to the structures used or would also hold when other models, structures and sets of data are used, and to investigate whether the results are sensitive to the assumptions of the Cowles-Commission model.

CHAPTER III

Design of the Monte Carlo Experiments

1. General Aspects of the Studies

Succeeding parts of this study report the results of a number of Monte Carlo experiments conducted on some of the techniques described in Chapter II. This chapter describes the experiments conducted.

Seven estimating methods were investigated. Four are single-equation estimators: Direct Least Squares (DLS), Two-Stage Least-Squares (2SLS), Nagar's Unbiased k-Class Estimator (UBK) and Limited-Information Maximum-Likelihood applied to single equations (LIML). Two are full-model methods: Three-Stage Least-Squares (3SLS) and Full-Information Maximum-Likelihood with unrestricted variance-covariance matrix of disturbances (FIML). The seventh method is Least-Squares applied to estimate the reduced form directly (LSRF). These seven methods were chosen because they seemed to be the most attractive ones and those of most interest to econometricians on the basis of their theoretical properties or the ease of their computation. Other methods were excluded primarily in order to limit the amount of computing time required to complete the study.

For the six methods which deal with the structural equations, estimates of the coefficients of the equations and the "standard errors" of these estimates, as suggested by the asymptotic properties of each method, were studied. In addition, the variance-covariance matrices of the residuals of the structural equations as estimated by the six methods were investigated. For all seven methods, the estimates of the reduced forms were calculated and predictions of the endogenous variables were made for three observations not included in the sample used to compute the estimates.

All these aspects were studied because it cannot be presumed that the performance of the methods would have the same characteristics for each of them and the different aspects would tend to be of interest to different sorts of applications. Estimates of the coefficients of the structural equations are important to studies of the structure of economic models and the parameters of economic systems. The calculated standard errors are supposed to give an indication of the reliability of the estimates of the coefficients and are of especial importance to testing hypotheses about them. The reliability of these standard errors is therefore of great importance. One might well prefer a method whose performance was inferior in other ways to another one if its standard errors were more reliable in the sense that they have a smaller tendency to underestimate or overestimate the variability of the estimates of the coefficients or themselves showed less variation.

The variance-covariance matrix of residuals are estimates of the parameters Σ of the model (1.1)-(1.5). They are of interest for this reason. In addition, the estimates of Σ can be interpreted as showing the variance contributed by features in the process generating the endogenous variables which are not explained by the econometric model and the size of the interactions between these unexplained features. The variance-covariance matrix of residuals can be interpreted as estimating the imprecision of the equations of the theoretical model or the extent to which the theoretical model fails to explain the observations. It seems worthwhile to investigate whether the estimates of Σ are reliable or whether they give an over--or under--optimistic picture of the adequacy of the structural equations in explaining the endogenous variables.

The reduced form is of interest when problems concerning the effect of changes in an exogenous variable on the endogenous variables are studied. When models contain lagged-endogenous variables among the predetermined variables, one-period effects--the impact multipliers--only are studied and not estimates of total, multi-period effects. The performance of the estimators in prediction is important when the main reason for building a model is prediction of future values of the endogenous variables.

The Monte Carlo studies were concerned with the performance of the methods in two types of situations. First were investigated the estimates of structures where the data were generated in conformance with the Cowles-Commission linear model (1.1)-(1.6). This is an investigation of how well the methods perform in situations for which they were designed (with the exception of DLS whose underlying assumptions are violated by this model). We were concerned in this part of the study with how the various estimators perform the tasks for which they were designed and especially with their performances relative to each other.

The second set of experiments was on models in which the assumptions of the model (1.1)-(1.6) were violated in some way. By investigating the estimates of models which do not conform to the assumptions (1.1)-(1.6), we may gain some insight into how seriously the violations of the assumptions affect the quality of the estimates, whether use of the methods in such cases tends to produce seriously misleading estimates and whether the performances of the methods relative to each other are much affected by violations of the assumptions.

The experiments were conducted with the use of a Control Data Corporation 1604 computer. Various aspects of the computer and the programs used, which are relevant to this study, are discussed in Appendix A. While the 1604 is a computer of large capacity and high speed and while Princeton University made very generous allocations of machine time to this study, both the size of the computer and time considerations restricted the experiments which could be conducted. These restrictions limited the study to the performances of the estimators on small structures with fairly modest numbers of observations. Computation considerations also account for many of the limitations of the design of the series of experiments. Both time and space considerations limited the number of replications in each experiment to fifty.

2. The Generation of Data and the Computation of Estimates

The Role of the Program-Parameters. To understand the differences between the various experiments it will be necessary to describe the ways in which the generation of the data was varied for the different experiments.

The computer programs used in this study had seventeen parameters which defined the type of experiment to be conducted in any particular run. The program-parameters controlled the exact type of data generated for an experiment by "telling" the programs which of the possible combinations of the alternatives built into the program should be used in the experiment. A particular set of values for the program-parameters defined exactly an experiment to be conducted.

The program-parameters selected the structure to be studied. They specified the number of observations to be used to estimate the structure. They defined the particular types of data to be used and also selected the sets of random numbers to be used. The effect of each program-parameter will be described when we discuss the performances of the estimators on the types of data each produced.

The Exogenous Data. One of the program-parameters specified the number of observations to be used to estimate the structure, T. The first exogenous variable had a value of unity at all observations, being a dummy variable to which the constants of the equations applied. The other exogenous data were produced by using the random-number generator described in Appendix A. The basic sets of exogenous data used in this study were rectangularly distributed, pseudo-random numbers lying in the range 0-100. The basic observations were generated so that each value of an exogenous variable was independent of all other values of the exogenous data.[35] Enough of these numbers were produced to form a $(T+3) \times K$ matrix of observations of the K exogenous variables. The three observations generated in addition to those needed for estimation were produced so that predictions could be made. Once a set of exogenous variables had been chosen, their values remained the same for all the replications of an experiment. The ways in which these basic exogenous variables were altered to make them show features of data which it was desired to study are described in Chapters V, VI and VII.

<u>The Structural Disturbances</u>. The structural disturbances used in the experiments were formed from normally and independently distrubuted pseudo-random numbers. For each of the T observations, a vector of g structural disturbances, U_t $(t=1,\ldots,T)$, was formed by linear combinations of g such independently distributed variates so that $E(U_t) = 0$ and $E(U_t^{\prime}U_t) = \Sigma$. ($g$ is the number of structural equations.) The linear combinations to be used, and so the value of Σ, were specified by the program-parameter which specified the form of the structural

[35]In small samples, correlations would sometimes appear between the exogenous variables even though the expected correlations were zero. In some instances these correlations were quite high: in one case the coefficient of determination for 20 observations on two exogenous variables was $R^2 = .37$. While this value was high enough that some problems of multicollinearity might be present, it was decided not to eliminate samples with such high correlations from the sets of exogenous variables used. It seems unlikely that economic data would not have at least this much correlation. This acceptance was also based on the fact that the high correlation was only noted after several runs using these data had been made. Tests of the random-number generator indicated that this high value of R^2 was only the result of bad luck and did not reveal a deficiency in the generator.

equations and the values of the structural coefficients, B and Γ. [Cf. (1.1).] The structural disturbances of different observations as thus generated were independent. The ways this was altered for some experiments will be discussed in Chapter VII.

Generation of the Endogenous Variables. Having obtained the exogenous variables, Z, and the $T \times g$ matrix of structural disturbances, U , the endogenous variables Y were generated using the coefficients specified to the experiment:

$$Y = -Z\Gamma'(B^{-1})' + U(B^{-1})' \quad .$$

The endogenous and exogenous data thus formed were then used to obtain estimates of the structure. The structure was estimated by each of the methods from the same set of data so that the estimates of one method are not independent of those of another.

The Replications of the Experiments. Having obtained estimates of the structure, another set of structural disturbances and of endogenous data was generated from which another set of estimates was formed. The disturbances of the separate replications of this process were independent from those of other replications. When other stochastic features were introduced into the generation of the data, such as measurement error of the exogenous variables, they too were independent in the various replications. All parameters of the models were the same in all replications. The replications of the process of generating the data and making estimates were continued until estimates had been obtained from fifty sets of data. In the rest of the study a Monte Carlo experiment will mean the obtaining of a sample of estimates by this process from a population, all of whose parameters were unchanged, and the analysis of these estimates.

3. Problems of Estimation

Appendix A presents the formulae used in computing the estimates discussed in Chapter II. Here we only discuss some of the difficulties faced in these computations and the ways in which they were overcome.

General Computational Difficulties. The prime difficulty was that some singular or near-singular matrices were encountered which could not be inverted. Whenever a matrix could not be inverted, the sample in which it occurred was abandoned and an additional one added to the samples used. The same procedure was adopted when it was impossible to find the characteristic root needed for LIML. Otherwise all results were accepted even if they might seem dubious. In particular, no check was made to see whether certain matrices were positive definitive when they should have been. Indeed, estimates were accepted even when they had negative "squared standard errors". Although at first sight such estimates might seem to be recognizably ridiculous, they were kept in the sample of estimates for two reasons. First, it cannot be presumed that such a blatent failure of the "standard errors" based on asymptotic properties to give useful information about the dispersion of the estimates necessarily signals that the estimates themselves will be ridiculous. Second, if this is a signal that the estimates are poor, dropping these estimates from the sample may lead to mistaken conclusions that one method is a good deal better than another simply because the evidence unfavorable to the first method has been deleted from the study. However, while these estimates were kept in the sample, the occurrence of negative "squared standard errors" was noted so that one could see if the estimates for which this occurred were ridiculous.[36] This was not a wholly satisfactory procedure. These cases arose either through failure to invert a positive-definite matrix correctly or failure to achieve the maximum of a function. They usually arose through an inability to perform the computations without much rounding error. It is possible that rounding error built up sufficiently to lead to meaningless results. However, since in the matrix inversion program a check was made for indications that rounding error might be serious, It was felt that all results which could be computed should be accepted as being valid applications of the estimating techniques.

<u>The Problems of FIML</u>. FIML is the method which presents the greatest computational problems. The problems involved can be described as (1) the choice of the direction in which to ascend the likelihood function from a given point, (2) the choice of the size of step to be taken in that direction before recalculating the direction of ascent and

[36] Our procedure here is different from that of R. Summers who dropped cases where negative standard variances occurred. See Robert Summers, "A Capital-Intensive Approach to the Small Sample Properties of Various Simultaneous Equation Estimators," (typescript). Other investigators do not seem to have encountered this problem. It arose most frequently with FIML estimates and Summers' is the only other Monte Carlo study which investigated this method.

(3) the choice of a criterion for deciding that the iterations have led to a point sufficiently close to the maximum to warrant accepting the estimates as the FIML estimates. The exact ways in which FIML was computed are described in Appendix A. The ways in which these problems were met were all somewhat arbitrary and there are different possible ways of arriving at the FIML estimates. It is possible that there will be differences between the estimates obtained using these different ways. In trying out the program and various schemes for obtaining FIML estimates, it was found that such differences tended to be very small.

A more serious problem is that there may be more than one relative maximum in the likelihood function.[37] There is no guarantee that the method used for computing FIML will converge to the highest of these points. In view of these difficulties, for the discussion of the results of the Monte Carlo studies we shall define Full-Information Maximum-Likelihood estimates to be those obtained by using the methods outlined in Appendix A and shall not worry about possible differences in the estimates to be obtained by using other methods.

4. Investigations of Cases where the Assumptions of the Standard Model are Fulfilled

Many of the Monte Carlo experiments in this study were concerned with the performance of the estimators in situations falling within the

---

[37] See T. C. Koopmans, H. Rubin and R. B. Leipnik, "Measuring the Equation Systems of Dynamic Economics," in Tjalling C. Koopmans (ed.), Statistical Inference in Dynamic Economic Models (New York: Wiley, 1950), pp. 235-236.

specifications of the model (1.1)-(1.6) and with the stability of the results in different experiments. We were concerned here with the properties of the estimating techniques and with the suitability of studying these properties by Monte Carlo studies.

Strictly speaking, a Monte Carlo experiment on the techniques for estimating simultaneous equations, which uses sets of data generated by one structure and with one set of exogenous variables, provides information only about the performance of the techniques in estimating that structure with those values for the exogenous variables. This information might be of interest if one already had an empirical structure and wished to see how reliable estimates of it would be if the estimates of the structure were used as population parameters of a Monte Carlo study. The aim of most Monte Carlo studies, however, and certainly of the ones reported here is to provide information which can be generalized at least to the estimation of similar structures and, hopefully, even to the estimation of fairly different ones. This would seem to suggest that in studying any aspect of the techniques one should use a large variety of models and data sets to show that the results are not particular to specific structures. Such a procedure would greatly limit the number of aspects that could be studied. But the subject of the generality of the results is of sufficient importance that at least some attention must be paid to it.

It was decided to concentrate on studying a great many different aspects of the generation of the data rather than to study each separate aspect with many structures and sets of data. Most of the experiments were conducted on only one structure with but one set of exogenous variables and with what was, in essence, the same set of disturbances. It is hoped that the findings of these experiments will carry over to cases where other aspects are also varied and to other structures. A check on how limited the results of the experiments might be was made by having for each aspect studied some experiments which used different structures or sets of data. We began the study with an investigation of the hypothesis that the particular structures and data sets used do not greatly affect the results.

Three types of experiments were conducted on the robustness[38] of the methods with respect to the model and data sets used in the experiments. First, different structures of the same model were analyzed using the same exogenous variables.[39] Second, other models were studied. Third, experiments were conducted using other sets of exogenous data and of structural disturbances to see if the particular sets chosen had much effect on the experiments.

A second group of experiments was intended to throw light on the generality of the results, but was also designed to investigate how the methods perform with different amounts and types of data. The effects of having different numbers of observations and the effects of having disturbances of different sizes, but with the same correlations between

---

[38] By "robustness" we mean that the performance of the estimators is insensitive to alterations of assumptions (1.1)-(1.6) or the particular values of the parameters of such a system and not that the estimates tend to a normal distribution.

[39] The structures used are described in sections 4 and 5 of Chapter V.

them, were investigated. The effect of having exogenous data of different types was next studied. Multicollinearity in the exogenous variables--correlations among them--was introduced in several ways to examine the effect of this common feature of economic data. Next, trends were introduced into the exogenous data. Then runs were conducted in which the exogenous variables were auto-correlated--that is, there were correlations between the exogenous variables in one period and those of previous periods. Finally, a number of experiments was conducted where lagged-endogenous variables served as predetermined variables to see if the estimation of "dynamic" models showed different characteristics from estimation of models containing no lagged variables. A number of experiments combined two or more of the aspects to see if the performances of the methods were different when several of the features were present from what they were when the features were present singly.[40] The results of these experiments and the ways in which the generation of the data was adapted to incorporate the various features are described in Chapters V and VI.

5. Investigations of Cases where the Assumptions of the Standard Model are Violated

Some of the Monte Carlo experiments for this study examined the effects on the various estimation techniques of violating the assumptions

[40] It should be emphasized that these experiments were mostly conducted with only one structure and with data derived from one set of random numbers. They do not meet the problem that the results are limited by the small number of independent cases which were studied. Indeed, the desire to conduct them led to the limitation on the number of independent cases studied.

of the model (1.1)-(1.6) in generating the data or in estimating the structure.

The first set of these experiments examined the importance of errors of observations in the exogenous variables. Economists seldom have data of great accuracy;[41] instead, economic data typically are subject to measurement errors of not insignificant amounts which arise from a variety of sources. Such errors violate assumption (1.5). Errors of various sizes were added to the exogenous variables in some of the runs. The results of estimating the structural equations using these data were compared with estimates of the same structure using error-free data to see how serious a problem measurement error presents for the building of econometric models. Like the other experiments described in this section, these experiments also aimed at trying to give an indication of how urgent might be the development of techniques based on assumptions which conform more fully to conditions typical of economic data.

Another subject for investigation was the effect of disturbance terms which do not conform to the rather rigid assumptions of the model (1.1)-(1.6). There are several ways of violating these assumptions. Assumption (1.4) may be violated by making the expected variance-covariance matrix of disturbances vary from one observation point to another; that is, by the disturbances being heteroskedastic. Heteroskedasticity may creep into economic data because (a) trends in the economy, or in the model, affect the influences unaccounted for by the

[41] Cf. Oskar Morgenstern, On the Accuracy of Economic Observations (Princeton: Princeton University Press, 1963).

model; (b) the nature of parts of the economy not explicitly included in the structure being examined, although connected to it through the residuals, might be changing in ways which make the variance of the residuals change; or (c) the variance of the disturbances may be related to the level at which the economic segment being studied is operating. Such considerations led to conducting some experiments with heteroskedastic disturbances.

The assumption that the disturbances of observations numbered t are independent of those at other values of t is another one whose validity may frequently be questioned in economic models. For example, if a dynamic process, where values of the variables at one period depend on previous values of themselves, impinges on the model through the disturbances, serial correlation of the residuals is likely to arise. Some of the experiments were conducted with auto-correlated disturbances.

The coefficients of an economic model may themselves be stochastic. At least two lines of argument lead to such a belief. First, it may be felt that the behavioral equations are not immutable laws but represent only tendencies whose strength and character may vary from time to time or between different economic units. The factors making the relationships stochastic are then not only additive disturbances. The disturbances may apply directly to the coefficients of the model. To express it another way, the importance given to certain factors in making decisions may not be constant, but may vary about a set of central values. The parameters of models representing these processes would then be stochastic. Another way to justify the introduction of stochastic coefficients comes from considering the aggregative nature of most economic models. By representing an aggregate or average of a variety of individual behavior patterns, the coefficients may vary because of changes in the importance of parts of the aggregates without the parameters of the behavior of the individual parts themselves changing. This consideration would also make one expect the parameters (coefficients) of econometric models to be stochastic. In view of these considerations, a few runs were conducted with the values of the coefficients generating the data varying stochastically from observation to observation in each sample.

The final type of violation of the assumptions concerned the misspecification of the form of the equations, in the sense of incorrectly specifying the variables which do or do not enter into one or more of the equations. Misspecification in this sense is a serious danger in building econometric models where there may often be doubt about whether a variable should be in a particular equation of a model or in the model at all. A particular danger is that one may feel that a variable should enter the model somewhere but not be sure in which equation it should be placed. Since economic theory very seldom gives firm guidance on the exact form of a model and is itself usually of a tentative nature, the danger of misspecification is almost always present. It is important to try to discover how seriously such misspecification affects the results of estimating a model. Misspecification is of further importance to this study since it is sometimes argued that single-equation methods have considerable advantages over full-model methods in that the estimates of correctly specified equations are not affected by misspecification elsewhere. Because of the importance of the danger of misspecification, a fairly large number of experiments was conducted to investigate it. The details of the misspecifications involved and the results of introducing them are reported in Chapter VIII.

Some experiments were conducted with more than one violation of the assumptions. It was impossible, of course, to make experiments with all the possible combinations of violations. However, since there is little reason to presume that the effects of several violations act like simple combinations of the effects of the violations taken separately,[42] these experiments are of considerable importance. They may well be those of greatest interest in the study since it may well happen that econometric models violate the assumptions of the Cowles-Commission model in several ways. The details of experiments involving violations of the assumptions other than misspecification of the form of the equations are found in Chapter VII.

The experiments conducted far from exhaust the cases of possible interest. For instance, the effects of linearizing a nonlinear model and then estimating the linearized model were not examined. It is believed that the experiments do throw light on some of the important problems of estimating systems of simultaneous equations under conditions other than those of (1.1)-(1.6).

[42] Cf. J. Johnston, Econometric Methods (New York: McGraw-Hill, 1963), pp. 215-216, for an example in the single equation case where the effect of two violations seems to be quite different from either of them separately.

6. Problems Encountered in Analyzing the Results

The analysis of the results of the experiments presented a number of very serious difficulties. In large part these arose from the facts that the problem of which estimator is best and of how good a particular estimator is are not at all well defined.

There is no single criterion by which to judge the various methods of estimating systems of simultaneous equations. The goodness of an estimator must depend essentially on the purposes for which it is calculated and the effects or penalties of its estimates not being equal to the population parameters. A measure of the goodness of an estimator depends on and should be derived from the utility functions of those for whom a model is built. However, if one adopts this approach, one should also argue that the estimator to be used should itself be derived so that the utility function is maximized over the estimated parameters. The derivation of such estimators is probably unfeasible for at least two reasons. First, the amount of work needed to derive such estimators whenever it is desired to estimate a model may far outweigh the disadvantages arising from using one of the available methods.[43] Second, models are frequently built not for a specific purpose but for dissemination among a large number of people who are interested in the area with which the model deals. These users might not possess the same utility functions and their utility functions, being unknown, could not be used to

[43] Such estimators might well turn out to be hideously difficult to compute, which is another reason for doubting the feasibility of this approach.

derive an appropriate estimator.[44] But the lack of a utility function on which to base judgement of "general-purpose" estimators does not mean that the problem can be ignored entirely. It effectively prevents us from adopting any single measure as a criterion on the goodness of an estimator. Instead a variety of measures has to be examined and the over-all ranking of methods becomes rather arbitrary and ambiguous whenever all these measures do not indicate the same rankings for the various estimators.

A particular difficulty in analyzing the results is that some of the more attractive parametric statistics may not be estimates of population parameters since these parameters do not exist for the distributions being sampled. The problem that the samll-sample distributions of the estimates may not possess finite lower-order moments has already been mentioned in this study. While we did examine both the averages and the variances of the estimates, one should not put too much reliance on these statistics.

Even if they did exist, basing the analysis simply on a few parametric statistics such as the second moment around the true value would not necessarily be justified. Other measures of dispersion might better describe the part which the dispersion of an estimate plays in the users' utility function. For example, if the utility function were

[44] These subjects have also been discussed in Richard E. Quandt, Some Small Sample Properties of Certain Structural Equation Estimators, (Princeton: Econometric Research Program, Princeton University, Research Memorandum No. 48, 1963), pp. 8-12, and in T. C. Koopmans and William C. Hood, "The Estimation of Simultaneous Linear Economic Relations," in William C. Hood and T. C. Koopmans (eds.) Studies in Econometric Method (New York:Wiley, 1963).

linear in the dispersion of the estimates around the true value, the mean absolute error of estimates $\hat{\theta}_i$ of θ, $\sum_i |\hat{\theta}_i - \theta|$, might be a more appropriate measure than the second moment about the true value, $\sum_i (\hat{\theta}_i - \theta)^2$.

It does seem reasonable to base judgement of the various methods on how closely, in some sense, their estimates tend to come to the true values of the parameters being estimated. Three aspects are of particular importance in examining the estimates the methods in the light of this criterion. First is the difference between the central tendency of the distribution of the estimates and their true, population values--that is, the amount by which the estimates miss the population values on the average. Second is the dispersion of the estimates about their central tendencies and, finally, the dispersion of the estimates about the true values--that is, how widely the estimates tend to be scattered away from the population values being estimated. In analyzing the data, a variety of measures of each of these features was examined. This was done because we do not have a unique criterion by which to judge each of the aspects. Each of the features was examined separately since we do not have a satisfactory way of weighting the aspects to obtain a single, over-all measure by which to judge the estimators.

The need to analyze a variety of measures for each of the types of estimates complicated one of the most serious, practical difficulties which the study encountered. This difficulty arises from the need to compress the huge amount of results produced by the Monte Carlo experiments into useful form. Even if we ignore the basic, raw estimates of the structures and concentrate on summary statistics, the quantity of such statistics is overwhelming.[45] To try to present even a couple of them for each coefficient or other quantity estimated would turn the study into nothing but a vast series of tables whose sheer bulk would prevent any use ever being made of them. It was necessary to compress, aggregate and summarize the results ruthlessly. This process necessarily means that we ignore and, indeed, lose some aspects of estimation which might be of interest.

A few tables of summary data are presented in Appendix D.[46] It is not expected that anyone will study all entries of these tables; Chapters IV through VIII give the gist of what they show. The tables are included so that anyone wishing to pursue a subject farther may do so and to show, in part, on what the findings are based. It is hoped that these tables will help to indicate the absolute performances of the estimators and throw light on the nature of the results while keeping the number of tables down to a reasonable amount.

Another difficulty to be faced in analyzing the data was the possibility that one or more of the estimates might not be the ones intended. The problems of FIML estimation have been mentioned in an earlier section. It is also possible that a number analyzed as being

---

[45] The paper on which the "raw estimates" were printed occupies over 6000 cubic inches. All the summary tables dealing with the separate estimates occupy about as much space. This chapter, if written on paper of similar thickness, would occupy less than 20 cubic inches.

[46] Other tables are available on request, at the cost of Xerox reproduction. These are described in Appendix B. Data from the separate experiments are also available on the same terms.

a particular type of estimate of some parameter of a model might have been a mistaken one, quite different from the true estimate, because of failure of the computing equipment. It is highly unlikely that the computer itself made mistakes affecting the estimates.[47] However, a good deal of transmission of data between the computer and magnetic tape had to occur before the final results were obtained because the amount of data produced by an experiment exceeded the internal storage capacity of the machine. In addition, it was sometimes necessary to copy the results from one tape to another. Input-output operations are more subject to mechanical failure than are the computation operations. Guards against and tests for this type of error were built into the programs, but some such errors may well have occurred. When a number was detected which "looked" wrong or which was inconsistent with other results in the experiment, the replication in which this occurred was dropped from the sample.[48] However, some erroneous numbers may have slipped through and were analyzed at face value. We could do nothing further about this difficulty, but must warn of its presence.

---

[47]It should be emphasized that this is certainly not impossible--though unlikely--and there might also have been undetected "bugs" still in the programs. Rounding error did, of course, occur.

[48]For example, both the estimates of the structural equations and of the reduced form were stored on magnetic tape. If on analyzing the data the structural equations as read from the tape did not lead to exactly the same reduced form as the one read from the tape, the estimates of the doubtful replication could be dropped (or, of course, the experiment could be redone). This happened only once in the course of the experiments. The doubtful experiment was repeated to obtain correct results since more than one replication was doubtful.

7. Descriptive Statistics of the Estimates

The basic output of each Monte Carlo experiment was fifty sets of estimates of the parameters of the structure being studied made by each of the methods investigated. It should be noted that the estimates of different parameters by any one method are not independent of each other since in each replication they were all estimated from the same data. For the same reason, the estimates of any particular parameter by each of the different methods were not independent of each other. These facts were used when comparisons of the estimates of different methods and of different parameters were made. Each method made fifty independent estimates of any one parameter. These estimates are a random sample drawn from the finite-sample distribution of the estimator of that parameter. This section outlines the statistics which were used to describe the distributions of the estimates of each of the parameters made by each of the methods.

Two measures of central tendency were used, the averages and the medians of the estimates of each parameter made by each method. It should be emphasized again that although we did calculate averages and variances and do present results based on them, not much importance should be given them in the light of the problem that the distributions of the estimates may not have any finite moments.

A variety of measures of dispersion about the central values was used. The standard deviations were computed. The inter-quartile range, the range, and the range of the central eighty percent of the estimates--that is, the difference between the forty-fifth and the sixth smallest observations--were examined. The use of only these three inter-quartile ranges is arbitrary. There can be no reason to suppose that only these three are of interest or that use of other quantiles would necessarily give the same sorts of results as these.

An even larger variety of measures of dispersion about the true values was examined and most weight was put on them in evaluating the performances of the estimators. The parametric statistic used was the square root of the (estimated) second moment about the true value, the root-mean-square error. Quantile statistics of the distribution of the absolute values of the differences between the estimates and the true values were also calculated. The median, the third quartile, the ninth decile and the highest values were used. Again, as was mentioned in connection with the inter-quartile ranges, the particular quantiles chosen were arbitrary. It may be noted that the inherent criterion for the choice of method involved in examining the largest deviations from the true values is that one should avoid methods whose worst deviations are larger than those of the others and accept those whose worst deviations are the smallest. It thus implies a rather extreme criterion by concentrating only on the worst estimates. The other quantile statistics are consistent estimates of the intervals in which the corresponding lower fractiles of the distributions fall.

Another type of statistic of dispersion about the true values which was used is the number of times, n_i, the estimates for each coefficient made by each method fell within $.i$ of the true values. n_i/n is a consistent estimator of the probability:

$$P_i(\theta - i\theta \leqq \hat{\theta} \leqq \theta + i\theta) = \int_{\theta - i\theta}^{\theta + i\theta} f(\hat{\theta})d\hat{\theta}$$

where n is the number in the sample (fifty throughout this study) and $f(\hat{\theta})$ is the frequency function of $\hat{\theta}$, an estimator of θ. Again the values of i chosen are necessarily arbitrary. n_i for $i = 2$ and for $i = 4$ were used in analyzing the estimates.

The number of times the estimate had the wrong sign was also examined. Often the sign of certain coefficients is of as much interest in an econometric model as the values of the coefficients and economic theory frequently suggests what the sign will be. Thus for an estimate to have the same sign as the population parameter may be considered to be an important quality. Estimators rarely having the wrong sign may be preferred to others which have the wrong sign more often even if the others are superior on different criteria. One would expect that this measure would be dependent on the true value of the parameter being estimated and especially on its closeness, in some sense, to zero.

To keep down the bulk of the tables presented in Appendix D and in the text and to simplify the discussion of the results, it was desirable to aggregate the summary statistics of the estimates of different parameters to get single, over-all numbers for the performance of estimators as judged by each of the measures examined. This was done in two ways. The first one was the summation of the results. For counting measures (e.g., the number of times an estimate had the wrong sign), the summary statistics are the total number of times a particular condition occurred. That is, if the statistic for the i^{th} coefficient by the j^{th} method was s_{ij}, the summary statistic for that method in an experiment, S_j, was $S_j = \sum_i s_{ij}$. Where the magnitude of a statistic was related to the size of the true value being estimated, the procedure was to divide each summary statistic for a parameter by the absolute value of the population parameter and then to take the mean of these statistics over all parameters. That is, the summary statistics were $S_j = \frac{1}{n} \sum_{i=1}^{n} \frac{s_{ij}}{|\theta_i|}$ where $|\theta_i|$ is the absolute value of the i^{th} parameter and n the number of parameters. These descriptive statistics are referred to as the <u>typical descriptive statistics</u>.

The typical descriptive statistics are <u>not</u> regarded as being estimates of the parameters of the distributions of the estimates. It is not presumed that all estimates have the same distribution nor that the normalization performed would give them the same distributions. The only function of the typical descriptive statistics is to describe briefly the <u>orders of magnitude</u> which were encountered in the study. It is hoped that they may help to indicate the sizes of the features of the distributions which were typically found in the course of the study when the presentation of these features for the individual parameters would have taken up too much space. Comparison of the estimators was not based on these statistics but on the statistics of the individual parameters, except when the comparisons for all the parameters led to the same results. Even then the typical descriptive statistics are used only as a convenient short-cut to the presentation of conclusions based on other material.

8. The Comparisons of the Estimators

One of the chief aims of the study was to compare the estimators and determine how well they performed relative to each other. This was done in a number of ways for each of the experiments conducted.

The basic way in which comparisons between the estimators were made was to count the numbers of times one method did better than each of the other methods. The criterion by which the estimates of each parameter were judged was the absolute size of their deviations from the true value of the parameter. The estimates compared with each other were, of course, those made in the same replication of the experiment. Since several methods were to be compared, there was a fairly large number of these pairwise comparisons. It was found convenient and useful in most cases to use ranking statistics instead of pairwise comparisons for summarizing and present the results.

The ranking was done as follows. The estimates for each coefficient in each replication were ranked by their absolute deviations from the true values, unity being assigned to the estimate closest to the true value, two to the second closest, and so on. The ranks obtained by the estimates of each method for any one parameter in the fifty sets of observations (replications of the experiment) were then summed to provide an over-all ranking of the methods for each coefficient. The method which had the lowest sum of ranks was judged to be the best at estimating that coefficient, on the basis of the ranking statistics.[49]

[49]The connection between the ranking scheme and the pairwise comparisons is worth noting. Suppose that for the m methods to be compared an...

The over-all ranking of the methods based on the sum of the ranks of the individual replications is invariant up to a linear transformation of the ranking scheme. However, other ranking schemes might give quite different results.[50] Another drawback is that it is not insensitive to what methods are included in the study. The standings of two methods could be reversed by the deletion of a third method from the ranking.[51] Finally, the ranking scheme gives as much weight to a replication in which the methods made almost identical estimates of a parameter

---

[49]...(m×m) table, $A = \{a_{ij}\}$, is drawn up so that a_{ij} shows the number of times the i^{th} method came closer to the true values than the j^{th} method in a set of n estimates of a parameter. ($a_{ii} = 0$) Then

$$R_j = n + \sum_{i=1}^{m} a_{ij}$$

is the total of the ranks of the estimates of the j^{th} method summed over the n sets of estimates. The ranking scheme based on the totals of the ranks of the n sets of estimates has an interesting property. Suppose that the rows and columns of the table are rearranged so that the first row and column of the table refers to the method with the highest rank-total, the second row and column to the method with the second highest rank-total, and so on. That is, rearrange the table so that $i > j$ if $R_i \leq R_j$. Then this rearrangement minimizes the sum of the elements lying above the main diagonal of the table; that is, it minimizes $\sum_{i=1}^{m-1} \sum_{j>i} a_{ij}$. A perfect ranking would have $\sum_{i=1}^{m-1} \sum_{j>i} a_{ij} = 0$ --that is, with a perfect over-all ranking of the methods, a method which placed better than another in the ranking would always do better than the other one. See J. Durbin, "Incomplete Blocks in Ranking Experiments," The British Journal of Sociology (Statistical Section), 1951, pp. 85-90.

[50]For example, one might use the squares of the ranks used here. This could lead to different results.

[51]For example, suppose that there are three methods and that, in the notation of footnote 49, $a_{23} > a_{32}$. If $a_{12} - a_{13} > a_{23} - a_{32}$, method 3 would beat method 2 in the over-all rankings even though method 2 would beat method 3 if method 1 were ignored. This is, of course, closely related to the problem of finding a social preference function. See K. J. Arrow, Social Choice and Individual Values, (New York: Wiley, 1951).

as to replications where there are pronounced differences between the estimates of different methods. One might be unhappy with a ranking scheme which indicated that one method was better than another if the first method only was better than the other when there were slight differences between the estimates while it did very much worse whenever there were large differences. This is possible with the ranking scheme used. Heavy reliance is placed on various types of totals of ranks for the methods in discussing the results of the experiments. The short-comings of the procedure should, in consequence, be kept in mind in evaluating the conclusions.

The totals of the ranks of the estimates of a particular parameter were used to judge the relative performance of the estimators of that parameter. It is desirable to indicate not only what the over-all ranking of the methods is but also to measure how strongly the over-all ranking is supported by the ranks in the individual replications. That is, it is desirable to measure the strength or consistency of the ranking. Kendall's coefficient of concordance, W, was used for this purpose. Kendall's W measures the degree of association in the performances of the estimates relative to each other, as judged by the ranks of the estimates in the separate replications, inbetween the replications of an experiment.[52]

[52] Other uses of Kendall's W were also made in the study and will be described presently. Denote the sum of the ranks of the j^{th} method in estimating a parameter R_j (A as in footnote 49). Kendall's W is defined:

$$W = \sum_{j=1}^{m} \left(R_j - \frac{\sum_{j=1}^{m} R_j}{m}\right) \Big/ \frac{1}{12} n^2(m^3 - m) \quad , \qquad \ldots$$

The performances of the methods relative to each other were investigated not only for the estimates of each parameter but also for a whole group of parameters. This was done in two ways. First, the sums of the ranks of the estimates over all the replications and over all the parameters were calculated. Second, the sums over the replications of the ranks of the estimates of each parameter were ranked. The sums over the parameters of the ranks thus formed were then taken.[53]

---

[52] ...where m is the number of methods and n the number of items for which the methods are to be compared--for example, the number of replications in an experiment. Kendall's W is discussed in Sidney Siegel, Nonparametric Statistics for the Behavioral Sciences, (New York: McGraw-Hill, 1956), pp. 229-238. See also M. G. Kendall, Rank Correlation Methods, (London: Griffin, 1948), Chapter VI. These references discuss the handling of ties, which is not discussed here. In the individual estimates ties were very unlikely and, in fact, only occurred when two methods were calculated from exactly the same formulae. (This happens when there are no over-identifying restrictions or only one in which cases LIML and UBK respectively are the same as 2SLS.)

[53] Let r_{ijk} be the rank of the estimate of the j^{th} parameter made in the i^{th} replication by the k^{th} method. Define $S_{jk} = \sum_i r_{ijk}$, the rank-total for the k^{th} method in estimating the j^{th} parameter. For each parameter the S_{jk} were ranked, unity being assigned to the smallest, two to the second smallest and so on. Let R_{jk} be these ranks. The first types of rank-totals used to get an over-all ranking of the methods for an experiment were $T_k^1 = \sum_j S_{jk}$. The second type was $T_k^2 = \sum_j R_{jk}$. An example may help to clarify the rankings. Let three estimators have made two estimates of two parameters. The estimates and the various rank-totals were:

	Parameter 1 = .2			Parameter 2 = .3		
Replication	Method 1	Method 2	Method 3	Method 1	Method 2	Method 3
			Estimates			
1	.1	0	1.5	.5	-.1	.6
2	.21	.8	.2	.4	.5	.11
			Rankings			
	1	2	3	1	3	2
	2	3	1	1	3	2
	$S_{11}=3$	$S_{12}=5$	$S_{13}=4$	$S_{21}=2$	$S_{22}=6$	$S_{23}=4$
	$R_{11}=1$	$R_{12}=3$	$R_{13}=2$	$R_{21}=1$	$R_{22}=3$	$R_{23}=2$

For both types of rank-totals Kendall's W was calculated. The first type of ranking gives an over-all measure of the performances of the methods relative to each other as suggested by the individual estimates. The second gives an over-all ranking of the methods based on judgement of the performances of the estimators in estimating each of the parameters. It is of interest to see to what extent a ranking of the methods based on the rank-totals of the estimates of each parameter would be repeated in the rankings for the different parameters. The second type of ranking indicates this.

Another use of ranking was made in evaluating the relative performances of the methods according to the descriptive statistics. For each of the descriptive statistics of each parameter the estimators were ranked by the size of their descriptive statistics. The sums of these ranks over the parameters were taken and Kendall's W calculated to give statistics of the relative performances of the methods as judged by the descriptive statistics of their small-sample distributions. Pairwise comparisons of the number of times each method had a smaller descriptive statistic than another were also computed.

One other type of measure of the relative performance of the methods was used. The number of samples for which each method produced the best or the worst estimates of each parameter was calculated. These

---

53...Then

$$T_1^1 = 5 \qquad T_2^1 = 11 \qquad T_3^1 = 8$$

and

$$T_1^2 = 2 \qquad T_2^2 = 6 \qquad T_3^2 = 4.$$

are all-or-nothing measures and depend heavily on which other methods are included in the study. They give no weight to rankings, but they do have interest. Underlying interest in these measures is the criterion that one should choose the estimate $\hat{\theta}_i$ to maximize the probability,

$$P_r(|\hat{\theta}_i-\theta|) \leqq \min_{j\neq i}(|\hat{\theta}_j-\theta|)$$

where j runs over all other estimators of θ taken into consideration, or the different criterion of choosing $\hat{\theta}_i$ to minimize

$$P_r(|\hat{\theta}_i-\theta|) \geqq \max_{j\neq i}(\hat{\theta}_j-\theta|) \quad .$$

These criteria, as we shall see, often give ambiguous results, but they also were found to throw light on the properties of the estimators which would otherwise have been ignored. Between them they can highlight the dilemma to be faced in choosing an estimator if one method is frequently the best estimator but is also quite often the worst while another method is rarely either best or worst.

9. Tests for the Significance of Differences found in each Experiment

It is of interest not only to describe the differences between the estimators but also to test whether these differences were significant or could have arisen by chance.

A number of tests investigated the significance of differences between the methods found in making pairwise comparisons of the methods.

First, the significance of the differences in the dispersions of the estimators about the true values of the parameters being estimated were examined. The number of times one estimate was closer to the true parameter value than another, besides being a useful descriptive statistic, is a test statistic of the hypothesis that the probability, $P_r[|\hat{\theta}_i-\theta| < |\hat{\theta}_j-\theta|] = 1/2$. The binomial test was used for this with the .05 level of significance chosen. The number of parameters for which one method was significantly closer to the true values than each of the others was then counted. It may be remarked that here and elsewhere in the pairwise comparisons, two-tailed tests were used since we did not wish to predict which methods would be better than other ones.

A more powerful pairwise test used to compare the dispersions of two methods was the Wilcoxon matched-pairs signed-ranks test.[54] This is a non-parametric test for differences in central tendency between the distributions of two related samples. In this study it was used to test for different central tendencies in the distributions of the absolute differences between the estimates of the various methods and the population parameters.

In both these tests there was the possibility that one method would perform significantly better than another one for some parameters while it would be significantly poorer in estimating other ones. If this occurred, the results of the experiments as to which methods were better than other ones have a basic ambiguity for judgement of which

---

[54] This test is described in Siegel, op. cit., pp. 75-83.

method to choose. Even if for some parameters one method did not do significantly better than another estimator while it was significantly worse for others, the relative performances of the methods in estimating the parameters might vary significantly over the parameters estimated. For example, although one method might be better than the other for all parameters, it might have surpassed the other method significantly more frequently in estimating some parameters than in estimating the other ones. The Cochrane Q test could be used to test whether the differences in relative performance were significant.[55] This was not done in most experiments for two reasons. First, with as many as twenty-four parameters to be estimated by as many as seven methods, the number of possible comparisons was very large and time-consuming. Second, in most experiments it could be inferred from the statistics of the number of times one method beat another whether the test would find many significant differences. This alone seemed to suffice to answer whether the performances of the estimators were similar for different parameters.

The sign test and the Wilcoxon matched-pairs signed-ranks test were also used to test for different central tendencies of the distributions of the estimates which were made by the various methods. However, it is probably of more interest to test the significance of differences in the central tendencies of each estimator from the population values rather than to test for significant differences between the estimators. This was done with two tests.

[55]This test is described in Siegel, *op. cit.*, pp. 161-166.

The bias of the averages of the distributions was tested by the t-test with the statistic $\frac{\bar{\theta}_i - \theta_i}{S_{\hat{\theta}_i}}$ where $\bar{\theta}_i$ is the mean of the estimates $\hat{\theta}_i$ of θ_i and $S_{\hat{\theta}_i}$ is their standard deviation. The doubts about the usefulness of parametric statistics apply, of course, to this test and little reliance was placed on it. The significance of the bias of the medians was investigated by a binomial test of the hypothesis that $P_r(\hat{\theta}_i > \theta) = P_r(\hat{\theta}_u > \theta) = .5$. (This test essentially counts the number of estimates lying between the median and the true value and sees if it could have arisen by chance. Under the null-hypothesis of no bias the expected number of these estimates is zero.)

The significance of Kendall's W can be tested by $\chi_r^2 = n(m-1)W$ where m is the number of methods and n the number of replications. This statistic is approximately distributed as χ^2 with $(m-1)$ degrees of freedom under the null-hypothesis that there is no over-all agreement among the rankings. This is the statistic used in Friedman's two-way analysis of variance, which is a test of the hypothesis that the members of k samples come from the same population, or more explicitly, that the m different columns of ranks in a table are all from the same population.[56] The test was used to investigate whether a systematic ranking of the methods was suggested by the sums of ranks over the individual estimates or over the different parameters estimated, or

[56] The test is described in Siegel, op. cit., pp. 166-172. See also M. Friedman, "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance," The Journal of the American Statistical Association, vol. 37 (1937), pp. 675-701.

whether differences between the methods in the totals of their ranks could have arisen by chance if there were not differences in the relative performances of the methods. The significance of χ_r^2 for the totals of the ranks of the estimates of each separate parameter was tested at the .05 and .01 levels. The test was also used for the sum over the parameters of the ranks assigned to the descriptive statistics for each parameter. It was used, in addition, for the sums over the parameters of the ranks of the rank-totals of the individual estimates of each parameter.[57] It is not appropriate for the totals of the ranks of all the separate estimates of all the parameters[58] since in each replication the estimates of one parameter are not independent of those of others while the test assumes that the separate sets of ranks used are independent. The same objection might be made for using the test on the rankings of the descriptive statistics since the descriptive statistics of one parameter are not independent of those of another. However, it seems appropriate to test the null-hypothesis that the rankings of the descriptive statistics act as if they were random. Rejection of the null-hypothesis here would simply indicate that over the particular sample used for estimation the methods tended to perform, relative to each other, in a systematic way and not that their relative performances in making independent estimates of different parameters showed a non-random ranking of the estimators.

---

[57]That is, for the T_i^2 of footnote 53.

[58]That is, for the T_i^1 of footnote 53.

One other test was used to examine whether the relative performances of the estimators were significantly different in estimating some parameters from those in estimating others. When counting statistics (such as the number of estimates falling within twenty percent of the true values) were used to describe the distributions of the estimates of the individual parameters, the χ^2 test was used to test whether the performances of the methods in estimating some parameters were significantly different from the performances for other parameters. The contingency coefficients for the tables of the number of times each method had the characteristic for each parameter were also computed although little use was made of them in analyzing the results.[59]

10. The Analysis of the Standard Errors

The statistics described in the foregoing sections were used to analyze the estimates of the structural coefficients, the reduced-form coefficients and the elements of the variance-covariance matrices of the structural residuals. They were also used for the predictions calculated in each replication of an experiment. Different procedures were used to evaluate the standard errors computed for each of the structural-equation estimators in each experiment.

The standard errors may be regarded as attempts to estimate the root-mean-square errors of the distributions of the estimates of

---

[59] The contingency coefficient is described in Siegel op. cit., pp. 196-202. The same reference discusses the χ^2 test on pp. 175-179.

the structural coefficients or as estimates of their standard deviations. It seems of interest to see how well they did this by calculating for each parameter the quantile of the distribution of the standard errors where the standard deviation and the root-mean-square error of the estimates of the parameter fell. If they occurred at high quantiles, one might judge that the standard errors tended to underestimate the variability of the estimates of the coefficients. This is, of course, a fairly rough-and-ready measure and rests on the presumption that the standard deviations and root-mean-square errors are appropriate and meaningful measures of the dispersions of the estimates. Not much emphasis was placed on these statistics even as purely descriptive statistics.

Most of the analysis of the standard errors was done in quite another way. It is often supposed that the deviations of the estimates of a coefficient from the population value of the coefficient divided by their standard errors is distributed as Student's t. This belief is an extension from the single-equation case where it is true. It is not clear, however, what the degrees of freedom of the t-distribution should be. One might use the number of observations minus the number of coefficients to be estimated in the equation in which the coefficient studied falls. One might also use the number of observations minus the total number of coefficients to be estimated in the structure. Intuitively, the first might seem more appropriate than the second for the single-equation methods, but the second might seem to be more plausible for the full-model methods. However, there seems to be no very good reason to suppose that some other number of degrees of freedom would not be the appropriate one--even supposing that the hypothesis that t is the appropriate distribution is correct. We did not test the hypothesis that $\frac{(\hat{\theta}_i-\theta)}{S_{\hat{\theta}}}$ where $\hat{\theta}$ is an estimate of θ and $S_{\hat{\theta}}$ is the standard error of $\hat{\theta}$, is distributed as t because it was felt that in this form it is not a very interesting hypothesis. Instead, we tested the hypothesis that if one used values of the t-distribution to obtain the ninety-five percent confidence limits around the estimates, the true value of the coefficient being estimated would fall within those limits ninety-five percent of the time. This was tested for values of t based on the two different numbers of degrees of freedom suggested. In addition, the hypothesis was tested that if one would make correct inferences about ninety-five percent of the time and if one used the rough approximation that the true values of the coefficients fall within two standard errors of the estimate. The test statistics used were the number of times $\frac{|\hat{\theta}_i-\theta|}{S_{\hat{\theta}_i}} > k$ where k took on the three values mentioned. A one-tailed binomial test on the resulting numbers was then performed to test whether $P_r[(\frac{|\hat{\theta}_i-\theta|}{S_{\hat{\theta}_i}}) > k] > .05$. The choice of the ninety-five percent confidence level as a criterion by which to judge the standard errors was arbitrary, but it was felt that this level is probably the one most commonly used in economics. In a small selection of cases, the test was also applied using the ninety-percent and the ninety-nine-percent confidence intervals. The quality of the conclusions about the adequacy of the standard errors was not altered by using these different levels. We may note that the ninety-five percent confidence interval based on the t-distribution with one number of degrees of freedom is also a confidence interval for the t-distribution with other degrees of freedom--though of course the confidence percentage is not ninety-five. This fact was used in testing the standard errors so that for both applications of the t-distribution two confidence intervals were examined.

11. The Comparison of the Results of Different Experiments

In comparing the results of different experiments, we must bear in mind that the data of one experiment were not usually independent of the data of others. They would differ in only one aspect of their generation and would otherwise be the same. For example, when multicollinearity was studied, a linear combination of the exogenous data which would otherwise have been used served as the exogenous variables. The disturbances were the same as in experiments without multicollinearity.

Comparisons of the relative standings of the methods in different experiments were based on the rank-totals calculated in analyzing the results of the separate experiments. The sums over the parameters of the ranks of the descriptive statistics in separate experiments were compared in two ways. First, the totals of the ranks in the separate experiments were summed over a set of experiments to be compared. Second, the totals in each experiment were ranked. These ranks were then summed over the set of experiments.

The rankings based on the individual estimates in different experiments were compared in several ways. First, the totals of the ranks of the estimates over all replications and over all parameters of each experiment were summed over a set of experiments. Second, the totals of the ranks of all the estimates in an experiment were ranked. These ranks were then summed over a set of experiments. Third, the totals over the parameters of the ranks of the sums over the replications of the ranks of the estimates of each parameter were summed over a set of experiments. Finally, the totals over the parameters of the ranks of the sums over the replications of the ranks of the estimates of each parameter were ranked and these ranks were summed over the set of experiments to be compared.[60]

Kendall's W was calculated for each of the types of rank-totals investigated. The significance of W was tested only for the totals of the ranks given to the over-all rank-totals in the separate experiments.[61] It will be noted that when the ranks of inter-dependent experiments are summed to get an over-all ranking of the methods for a group of experiments, we are examining whether the relative performances of the methods in estimating from essentially the same sets of data are similar, not whether the methods perform similarly in estimating the structures from different sets of data.

---

[60] Let T^1_{kq} and T^2_{kq} be the values of T^1_k and T^2_k in the q^{th} experiment where T^1_k and T^2_k are defined in footnote 53. Let R^1_{kq} and R^2_{kq} be the ranks of T^1_{kq} and T^2_{kq}. Then the four types of statistics examined were (a) $\sum_q T^1_{kq}$, (b) $\sum_q R^1_{kq}$, (c) $\sum_q T^2_{kq}$ and (d) $\sum_q R^2_{kq}$.

[61] That is, the significance of W was tested only for statistics (b) and (d) in the previous footnote. The test cannot distinguish between agreement (or lack of it) coming from similar performances of the methods in different experiments and similar performances for different parameters in one experiment, which is why the test was not used for (a) and (c).

A number of tests were performed to examine whether the differences in the relative performances of the methods in different experiments were statistically significant. Which tests were used depended on whether the stochastic elements of one experiment were independent of those of another. The close relationships between some of the experiments made it possible to attribute with greater confidence the differences between the experiments to differences in the ways in which the data were generated and prompted the use of tests which took this feature of the generation of the data into account. These tests examined whether the frequency with which one method came closer to the true values than another one was significantly different in one experiment from the frequency in another one.

When the data of two experiments were related, the McNemar test[62] was used to investigate the null-hypothesis that for each parameter the number of times one method came closer to the population value than another was the same in two experiments. The data of one replication in the first experiment was related to the data of a replication of the second. The McNemar test compared the estimates made in each replication of one experiment with the estimates made in the corresponding replication of the other experiment. The test proceeded by counting the number of times method i made a better estimate than method j in the first experiment and made a poorer estimate than method j in the second experiment. It also counted the number of times method i was inferior to method j in the first experiment and superior in the

[62] See Siegel, op. cit., pp. 63-67, for a description of this test.

second. Under the null-hypothesis that the frequency with which one method would beat the other method was the same in both experiments, the expected number of times that one method changed for the better relative to the other method was the same as the number of times it changed for the worse. This is the hypothesis tested by the McNemar test. It will be noted that replications in which the relative performances of the two methods are the same in both experiments are ignored by the test. When estimates could not be obtained for a replication in one of the experiments being compared, the estimates for this replication in the other experiment were dropped from consideration in the test. The number of matched replications in two experiments was, in consequence, often somewhat less than the total number of replications in each experiment. The McNemar test was also used to test whether the number of parameters for which one method was judged better than another by the descriptive statistics was significantly different from the number in another experiment.

When two experiments used the same model, but independent sets of data, the χ^2 test was used to determine whether the frequency with which one method did better than another in estimating a parameter of the model differed significantly in a pair of experiments. The χ^2 test was also used to compare the performances of the methods in two experiments as judged by the descriptive statistics.

Tests for significant differences between experiments were performed for only a limited selection of the experiments conducted. The limitation on the number of tests performed was based on two reasons.

First, the comparison of experiments replication by replication using the McNemar test required not insignificant amounts of computer time and the storage of results on magnetic tape while waiting for other experiments to be conducted did tie up magnetic tapes in an inconvenient fashion. Second, if significant differences between one experiment and another occurred which could not reasonably be ascribed to Type I error, this was sufficient to answer the question whether the types of differences in the generation of data for the two experiments led to significant changes in the relative performance of the methods. Other experiments which exhibited the same sorts of differences in generating the data or more extreme versions of these differences were useful in examining the effects of these different ways of generating the data. Such further experiments helped in examining the question of how much and in what ways the performances of the methods were affected by different ways of generating the data. It was felt, however, that further tests of the significance of these differences would not add much to the study, especially when significant differences had once been found and further experiments on the same aspect were conducted or when extreme versions of different ways of generating data led to significant differences while less pronounced versions led to results which could be considered to be intermediate to those of the extreme experiments. Since failure to reject the null-hypothesis is not equivalent to accepting it, and may only be the result of having a small sample, it was decided to allow the inconvenience of performing a large number of tests to limit the number of experiments compared.

In examining the effects of using different types of data, the sizes of the descriptive statistics were used to give an indication of the importance of changes in performance which occurred in the different experiments. Such comparisons are necessarily rather rough and are meaningful only in conjunction with the exact ways in which the various sets of data were generated. However, it is hoped that they give an indication of how reliable are the estimates of structures made by the different techniques under various conditions and so can serve to some extent to indicate their usefulness for economic research. It is to be noted that tests were not performed to examine the significance of differences in the dispersions of the estimates of any one method in estimating the structures of two different experiments. The experimental controls built into the programs were not strong enough to make this an interesting test. The dispersions in different experiments could be expected to be different since the stochastic elements were usually of different importance in explaining the endogenous variables.

CHAPTER IV

The Basic Experiment

1. Introduction

This chapter reports in some detail the results of the first Monte Carlo experiment conducted for this study. This experiment, referred to as the basic experiment, will be used throughout the study as a standard of comparison for other experiments. Having dwelt on the results of this experiment at length, we can discuss more briefly the findings of other experiments in terms of how they differ from or are similar to those of the basic experiment.

The data for the basic experiment were generated in conformance to the statistical model (1.1)-(1.6). The experiment used the structure shown in Table IV-1, which is called MOD37A2.[63] Estimates were made from twenty observations in each replication of the experiment.[64] The data were of the sort described in Chapter III, Section 2. The pairwise correlations between the exogenous variables used are shown in Table VI-1. The disturbances accounted for about ten percent of the variances of the endogenous variables.

---

[63] Program-parameter 1 specified the structure to be used in each experiment. Thus in Table 1 of Appendix C, which shows the program-parameters for the first set of experiments, MOD37A2 is shown as the entry for program-parameter 1 in line A, which refers to the basic experiment. The role of the program-parameters is discussed in Chapter III, Section 2.

[64] Program-parameter 2 specified the number of observations used.

TABLE IV-1

MOD37A2

$$B = \begin{bmatrix} 1 & -.89 & -.16 \\ -.54 & 1 & 0 \\ 0 & -.29 & 1 \end{bmatrix}$$

$$-\Gamma = \begin{bmatrix} 44 & .74 & 0 & 0 & .13 & 0 & 0 \\ 62 & 0 & .70 & 0 & .96 & 0 & .06 \\ 40 & 0 & .53 & .11 & 0 & .56 & 0 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 35.24 & 34.48 & 31.12 \\ & 36.68 & 29.84 \\ & & 40.64 \end{bmatrix}$$

Note: B, $-\Gamma$, and Σ are defined in Chapter I, equations (1.1)-(1.5). The size of Σ shown is that for the standard run, with program-parameter 13 = 2.

2. Estimates of the Structural Coefficients

Bias. Table IV-2 shows the biases of the medians and the averages of the estimates of the structural coefficients made in the basic experiment.[65] Several points stand out in the table. First, the biases of the estimates of each of the consistent methods were fairly small. The biases of DLS were larger than those of the consistent methods. Second, differences in the biases of the consistent methods were usually slight and no method always had the smallest bias.[66] Third, the median biases of the consistent methods were rarely significantly different from zero at the .05 level. It is noteworthy that the biases of LIML and FIML were never significant. Very few of the significant biases would have been found significant if the .01 level had been used. With only one exception, the medians of the DLS estimates were significantly biased at the .05 level. Many of these biases were also significant at the .01 level.

Dispersion. The dispersions of the estimates of the structural coefficients about their central tendencies were quite large. Table IV-3 shows the inter-quartile ranges of the estimates. The inter-quartile ranges of 3SLS and FIML averaged roughly thirty percent of the true values of the coefficients being estimated. The inter-quartile ranges of DLS were about forty percent of the true values while those of the other k-class estimates were approximately forty-five percent of the true values.

---

[65] Bias here is taken to mean the differences between the median or the average of the estimates and the true value of a coefficient.

[66] The significance of differences in the central tendencies of the estimates of the different methods is discussed below when we examine their performances relative to each other.

The ranges of the estimates were, of course, substantially larger. The ranges of the full-model estimates, the DLS estimates and the estimates of the consistent k-class methods averaged about 100, 120 and 150 percent of the true values respectively.

The dispersions of the estimates of the consistent methods about the true values were similar to their dispersions about the central tendencies--as might have been expected from the small biases that were found. The DLS estimates had larger dispersions about the true values than about their central tendencies. To indicate further the sizes of the dispersions encountered in the basic experiment, the second part of Table IV-3 shows the ninth deciles of the distributions of the absolute deviations of the estimates from the true values.

<u>Relative Performances--Rankings of the Estimators</u>. The estimates of each structural coefficient were ranked by their absolute deviation from the true value of that coefficient. Table IV-4 records the totals of these ranks for the estimates of each coefficient made in the fifty replications of the basic experiment. It also shows the totals of the ranks for the estimates of all the coefficients. The most striking point to notice is that no method seems to stand out as being either a great deal better or worse than the others. This is brought out by the fairly low value of .112 for Kendall's W for the sums of the ranks of all the estimates. In the rankings for the individual coefficients, Kendall's W

TABLE IV-2

The Basic Experiment - Biases of Estimates of the Structural Coefficients

Biases of the Medians

True Value	DLS	2 SLS	UBK	LIML	3 SLS	FIML
.89	.058*	- .004	- .010	- .016	- .004	- .005
.16	.055*	.005	.007	.023	.002	.008
44.00	-9.476*	-1.976	-1.380	- .939	-1.423	- .199
.74	- .042*	.018	.023	.028	.012	.007
.13	- .107*	.001	.007	.014	.006	- .005
.54	.062*	.018*	.014*	.009	.009*	- .001
62.00	-21.825*	-4.573*	-3.536*	-1.106	-2.884*	.749
.70	- .094*	- .031*	- .028*	- .020	- .020*	- .014
.96	- .108*	- .006	.004	.010	- .018	- .011
.06	.002	.002	.002	.002	.002	.002
.29	.034*	.004	.002	.000	- .001	- .007
40.00	-13.689*	-2.948	-2.342	-1.047	-1.387	.476
.53	- .056*	- .012	- .010	- .004	- .002	- .001
.11	.016*	.006	.006	.004	.009	.003
.56	.034*	.027*	.027*	.024	.017	.010

Biases of the Averages

True Value	DLS	2 SLS	UBK	LIML	3 SLS	FIML
.89	.069	- .011	- .016	- .023	- .001	- .008
.16	- .059	.013	.018	.024	- .000	.004
44.00	-8.026	.138	.626	1.450	- .334	1.598
.74	- .047	.007	.010	.015	.006	.006
.13	- .113	.016	.024	.036	.008	.013
.54	.062*	.005	.001	- .008	.006	- .001
62.00	-20.985*	-1.053	.378	3.277	-1.956	.977
.70	- .090*	- .014	- .008	.003	- .012	- .007
.96	- .107	- .008	- .001	.015	- .008	.000
.06	- .003	- .005	- .005	- .005	- .004	- .002
.29	.034	.005	.003	- .001	.003	- .002
40.00	-14.934	-2.841	-2.140	- .383	-2.268	.012
.53	- .054	- .013	- .010	- .004	- .005	- .001
.11	.023	.008	.007	.005	.005	.003
.56	.031	.018	.018	.016	.012	.010

*Significant at the .05 level

TABLE IV-3

The Basic Experiment - Dispersions of the Estimates of the Structural Coefficients

Inter-Quartile Ranges

True Values	DLS	2SLS	UBK	LIML	3SLS	FIML
.89	.096	.103	.106	.106	.071	.071
.16	.157	.174	.166	.156	.070	.058
44.00	9.626	14.849	14.778	13.572	14.969	15.185
.74	.087	.106	.108	.103	.059	.058
.13	.204	.190	.199	.210	.136	.120
.54	.027	.037	.037	.036	.035	.031
62.00	9.217	16.441	17.994	17.791	17.386	18.757
.70	.064	.072	.075	.073	.055	.063
.96	.099	.142	.141	.170	.128	.135
.06	.055	.068	.069	.071	.030	.031
.29	.043	.053	.054	.060	.055	.061
40.00	20.255	19.481	19.254	22.676	23.574	23.718
.53	.064	.082	.083	.082	.080	.078
.11	.064	.081	.083	.082	.034	.039
.56	.076	.082	.082	.082	.069	.063

Ninth-Decile Absolute Deviations

True Values	DLS	2SLS	UBK	LIML	3SLS	FIML
.89	.155	.145	.155	.156	.075	.077
.16	.189	.192	.196	.187	.082	.076
44.00	18.207	20.798	21.527	21.932	16.392	18.329
.74	.137	.133	.134	.136	.087	.069
.13	.285	.252	.270	.278	.174	.188
.54	.088	.065	.069	.073	.039	.033
62.00	30.505	27.751	28.996	33.216	17.710	18.312
.70	.148	.124	.131	.124	.087	.093
.96	.198	.169	.173	.187	.136	.110
.06	.060	.071	.072	.074	.047	.030
.29	.067	.056	.057	.059	.059	.056
40.00	29.581	26.261	25.846	26.125	24.964	22.322
.53	.113	.110	.112	.116	.099	.108
.11	.083	.088	.089	.091	.059	.058
.56	.112	.117	.117	.118	.086	.072

TABLE IV-4

The Basic Experiment - Rankings of the Estimates of the Structural Coefficients by Absolute Deviation from the True Values

Coefficient	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W
β_{12}	217	179	187	197	134	136	.128**
β_{13}	203	190	202	206	130	119	.180**
γ_{11}	226	170	168	172	161	153	.077**
γ_{12}	208	178	192	206	128	138	.135**
γ_{15}	205	195	195	204	121	130	.171**
β_{21}	273	191	162	165	136	123	.328**
γ_{21}	261	194	175	151	141	128	.267**
γ_{23}	272	196	162	164	126	130	.333**
γ_{25}	230	177	171	189	137	146	.126**
γ_{24}	139	183	237	256	122	113	.421**
β_{32}	220	161	153	169	167	180	.065**
γ_{31}	217	174	171	171	149	168	.058*
γ_{33}	227	176	162	174	154	157	.083**
γ_{34}	206	191	187	185	130	151	.093**
γ_{36}	230	196	174	162	142	146	.127**
Totals	3334	2751	2698	2771	2078	2118	.112
Rankings of table†	86.0	61.5	52.0	65.5	22.0	28.0	.737**

**Significant at the .01 level
* Significant at the .05 level

† This line records the totals of the ranks given to the methods in ranking each line of the table.

TABLE IV-5

The Basic Experiment - Pairwise Comparisons of the Estimators for the Structural Coefficients

Number of times closer to the true values*

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	256	261	272	181	196
2SLS	494	-	358	367	249	281
UBK	489	392	-	392	253	274
LIML	478	383	356	-	250	262
3SLS	569	501	497	500	-	355
FIML	554	469	476	488	395	-

Significant differences - sign test**

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	1	1	1	0	0
2SLS	10	-	1	1	0	0
UBK	9	4	-	1	0	0
LIML	9	4	0	-	0	0
3SLS	15	10	11	10	-	0
FIML	13	8	7	7	0	-

Significant Differences - Wilcoxon test***

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	1	1	1	0	0
2SLS	7	-	3	3	0	0
UBK	7	0	-	2	0	0
LIML	4	0	0	-	0	0
3SLS	15	11	11	11	-	0
FIML	14	10	10	10	0	-

*Entries in each row record the number of times the method came closer to the true values than each of the other methods in estimating all the separate coefficients.

**Entries in each row record the number of coefficients for which the sign test found the dispersions of the method from the true values to be significantly smaller than those of each of the other methods at the .05 level.

***Each row records the number of coefficients for which the Wilcoxon signed-ranks matched-pairs test found the central tendency of the distribution of the absolute deviations from the true values of one method to have been significantly smaller than those of each of the other methods at the .05 level.

TABLE IV-6

The Basic Experiment - Rankings of the Descriptive Statistics of the Dispersions of the Estimates of the Structural Coefficients*

Statistic	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W
Standard Deviations	29	56	74	89	32	35	.784**
Inter-quartile Ranges	36	60	72	67	37	43	.317**
80% Ranges	31	62	76	80	31	35	.668**
Ranges	31	60	77	83	33	31	.734**
Root Mean Square Errors	83	43	60	78	24	27	.810**
Median Absolute Deviations	76	60	60	60	29	30	.452**
Third Quartile Absolute Deviations	81	56	59	64	27	28	.571**
Ninth Decile Absolute Deviations	72	54	64	77	26	22	.698**
Largest Absolute Deviations	60	60	65	72	26	32	.450**

* Table entries are the totals of the ranks of the descriptive statistics of the estimates of the individual structural coefficients.

**Significant at the .01 level.

varied substantially[67] from a high of .421 to a low of .058. All these values of W were significant at the .05 level.

The sum over all replications and coefficients of the ranks of the estimates suggests a (weak) ranking of the various methods. This ranking puts 3SLS and FIML in the first places. Next come the three consistent k-class estimators, UBK, 2SLS, and LIML. In the last place is found DLS. The totals of the ranks of the consistent k-class estimates were very close together. The differences in the rank-totals between these methods was considerably smaller than the differences between the total for any one of them and the totals for the full-model methods or DLS. Similarly, the two full-model methods were very close together.

For each coefficient, the sums of the ranks of the estimates of that coefficient were themselves ranked. The sums of these ranks over all the coefficients are shown in the last line of Table IV-4. These ranks showed fairly high concordance as is indicated by the value of .737 for Kendall's W. The lack of complete agreement between the coefficients as to the over-all ranking of the methods can be attributed largely, but not entirely, to lack of agreement over whether FIML or 3SLS was best and over the performances of the consistent k-class estimators relative to each other.

Relative Performances--Pairwise Comparisons. Pairwise comparisons between the various estimating techniques bring out the extent of the

[67] We are hampered here by not having a test for the significance of differences in the values of W. We discuss below the significance in the differences of the performances of one method relative to another one in estimating one coefficient from those in estimating other coefficients.

differences in the dispersions of the estimates from the true values. The first part of Table IV-5 shows the number of times each method was closer to the true values of the structural coefficients than each of the other methods. In some instances, a method came closer to the true values more frequently than a method which had slightly surpassed it in the totals of the ranks of all the estimates shown in Table IV-4. For example. FIML came closer to the true values than 3SLS more than half the time. These cases did not upset the suggested ranking of the estimators into three groups: they occurred only between estimators found in the same group, not between estimators in different groups.

The second and third parts of Table IV-5 show the number of coefficients for which the sign test and the Wilcoxon matched-pairs signed-ranks test found significant differences in the central tendencies of the distributions of the absolute deviations of estimates of the different estimators from the true values. Several things about these tables are worth noting. First, it was not usually the case that significant differences were found for all coefficients. Usually there were significant differences between a pair of methods for less than half the coefficients. Second, there were few significant differences between methods which had been placed close together in the rankings of the estimates by their absolute deviations from the true values. In particular, no significant differences were found between 3SLS and FIML. Third, there were instances where a test would find one method significantly better than another for some coefficients and significantly poorer for one other coefficient. The two tests did not always agree about the significance of differences between estimators for a particular coefficient. Further, there were cases where one test found more coefficients for which one method was significantly better than another than coefficients for which it was worse while the other test found it poorer more frequently. This happened in comparing the consistent k-class estimators with each other.[68]

The Cochrane Q test was used to determine whether the frequency with which one method came closer to the true values than another differed significantly among the coefficients. It was found that there were differences among the coefficients in the comparisons of every pair of methods which were significant well beyond the .01 level. This finding indicates that differences found in the concordances of the rank-totals of the estimates of the coefficients may be significant. It also makes precarious any attempt to evaluate the gain one could expect from choosing one method rather than another one. The probability that one method will surpass another is different for different coefficients of a structure. The differences seemed to follow no simple pattern.[69] The gain would seem to depend in an unknown (and possibly, since it may depend on the true values of the coefficients, unknowable) way on the coefficients in which one was particularly interested.

Before leaving the pairwise comparisons of the estimators, it is worth noting that the sign test and the Wilcoxon signed-ranks matched-pairs test found many coefficients for which there were significant differences

[68] The differences in the findings of the tests are due to differences in the definition of the central tendencies for which the tests investigate the significance of differences.

[69] For example, the differences did not seem to depend on whether the coefficients were in the beta- or the gamma-matrices.

in the central tendencies of the distributions of the estimates of different methods. Indeed, such significant differences were found more frequently than when the tests were used to compare the distributions of the absolute deviations of the estimates from the true values. The t-test also indicated that the averages of any one method were frequently significantly different from those of other methods. There did not, however, seem to be any pattern in whether a method was found to have a higher or lower central tendency than another one when the differences were significant. The tests do indicate that the differences between the methods shown in Table IV-2 were frequently significant even though they were small.

Relative Performances--Descriptive Statistics. The descriptive statistics of the dispersions of the estimates of each coefficient were ranked according to the sizes of the descriptive statistics. Table IV-6 presents the sums of the ranks over all the coefficients. The statistics for the dispersions of the estimates about the true values indicated the same ranking of the methods as was found in the individual estimates of the structural coefficients. There was not, however, complete agreement among the coefficients as to the ranking of the methods. Much of this was accounted for by disagreements over the standings of the full-model methods relative to each other and of the consistent k-class estimators relative to one another. There was, however, incomplete agreement among the coefficients over the ranking of the methods into groups. It is to be noted that in the rank-totals for the ninth decile and the largest absolute deviations of the estimates, DLS was placed with the consistent k-class estimators.

DLS was judged to be at least as good as the full-model methods and surpassed the consistent k-class estimators in the rankings of the descriptive statistics of dispersion about the true values. There was not full agreement among the coefficients as to the rankings of the estimators. LIML, however, was usually ranked as the weakest of the methods.

No method had estimates with a sign different from that of the population value of a coefficient very often. The frequency of wrong signs among the various methods agreed with the rankings of the methods by dispersion from the true values. The estimates which had the wrong sign were usually estimates of the smaller coefficients.

The number of times the estimates came within twenty percent of the true values gave the rankings further support. The frequency with which a method fell within twenty percent of the true values varied considerably between the separate coefficients estimated. The frequency with which a method came within twenty percent of the true value of a coefficient <u>relative</u> to the frequency for another method did not differ significantly between the coefficients. The frequencies with which the methods fell within forty percent of the true values showed the same pattern.

Each method produced the estimate which came closest to the true values of the structural coefficients in a fair number of replications. Most methods also had the worst estimates quite often. 2SLS was remarkable in <u>never</u> making the worst estimate, but it also seldom made the best estimate. FIML and DLS both produced the best estimates and the worst estimates more frequently than 2SLS and UBK. FIML was the method which was most frequently the best while DLS was most frequently the worst.

Summary. Investigation of the estimates of the structural coefficients of the basic experiment led to a number of findings. Bias was not an important problem for the consistent methods though it was for DLS. Only DLS biases were frequently found to be significantly different from zero. The dispersions of the estimates were quite large. The differences in the relative performances of the methods were not very great and sometimes differed for different coefficients. No method was greatly better or worse than the other. It was possible to find a ranking of the methods with the full-model methods first, the consistent k-class estimators second and DLS last. The ranking was found both for the deviations of the individual estimates from the true values and for the descriptive statistics of the dispersions of the estimates about the true values. It did not seem to be possible to choose between the estimators in any one group. We may note that the ranking of the methods into three groups is the one suggested by their asymptotic properties. Significant differences in the frequencies with which one method came closer to the true values than another were not found very often, but there were significant differences in the frequencies with which one method surpassed another in estimating some coefficients from those in estimating others. The descriptive statistics of dispersion about the central tendencies of the estimates ranked DLS with the full-model methods ahead of the consistent k-class estimators.

3. Standard Errors

Standard errors for all estimates of the structural coefficients were calculated. They were computed either according to the formulae for the asymptotic standard deviations of the estimates or, in the case of DLS, from the formulae for the standard errors in the single-equation model. The usefulness of these standard errors for making inferences about the variability of the estimates of the coefficients is examined in this section.

The standard errors tended to under-estimate the variability of the estimates of the structural coefficients. The root-mean-square errors of the estimates of the structural coefficients fell in the upper half of the distributions of the standard errors for all the coefficients of the basic model. There was some variability in the exact place the root-mean-square errors fell, but for over two-thirds of the coefficients, they fell in the highest quartile of the distributions of the standard errors of the consistent methods. For DLS, the root-mean-square errors for eight of the fifteen structural coefficients were greater than the largest standard error of the estimates of the coefficients. All but one of the other root-mean-square errors fell in the fourth quartiles of the distribution of the DLS standard errors. For the standard deviations of the estimates of the structural coefficients, the performances of DLS were quite different. Two of the standard deviations fell in the second quartiles; seven fell in the third quartiles and only six lay in the fourth quartiles. The positions of the standard deviations of the other estimators were much the same as the places of the root-mean-square errors, as one might expect from the lack of pronounced biases in the estimates of the coefficients.[70]

We turn now to the usefulness of the standard errors for forming confidence intervals for the estimates of the structural coefficients or for testing hypotheses about the true values of the structural coefficients. The statistics examined here are the deviations of the estimates of the structural coefficients from the true values divided by their standard errors, called in this study the t-ratios.

The hypothesis that ninety-five percent of the t-ratios would fall within the 95 percent confidence interval of the t-distribution was tested using two different degrees of freedom for the distribution. The first was the number of observations minus the number of the coefficients to be estimated in each equation. The second was the number of observations minus the total number of coefficients to be estimated in the structure. In the basic experiment these were fifteen and five degrees of freedom respectively.

The number of times the t-ratios fell outside the two intervals is recorded in Table IV-7. About ten percent of the ratios for the consistent methods fell outside the first interval and four to seven percent of the ratios fell outside the second. In neither case could one reject

[70] The distributions of the standard errors were usually skewed to the right. We may note that the averages of the standard errors were usually smaller than the root-mean-square errors or the standard deviations of the estimates of the structural coefficients. This would seem to indicate a tendency for the standard errors to understate the variability of the estimates of the coefficients.

TABLE IV-7

The Basic Experiment - Number of Times the t-Ratios Fell Outside the 95% Confidence Intervals of the t-Distribution

Degrees of Freedom - 15***

Coefficient	DLS	2SLS	UBK	LIML	3SLS	FIML
β_{12}	10*	4	4	4	3	5
β_{13}	8*	6*	6*	4	2	4
γ_{11}	11*	4	3	4	5	4
γ_{12}	7*	3	3	4	4	3
γ_{15}	11*	3	3	4	5	7*
β_{21}	36*	5	5	5	5	1
γ_{21}	34*	5	6*	4	7*	6*
γ_{23}	29*	6*	5	6*	5	3
γ_{25}	20*	8*	8*	9*	8*	7*
γ_{27}	3	4	4	4	4	8*
β_{32}	14*	6*	6*	5	5	5
γ_{31}	14*	8*	8*	5	9*	6*
γ_{33}	7*	6*	5	6*	5	5
γ_{34}	5	4	4	4	5	4
γ_{36}	6*	6*	6*	4	5	6*
Totals	215	78	76	72	77	74
Percent of t-ratios	29	10	10	10	10	10

TABLE IV-7 (Cont.)

The Basic Experiment - Number of Times the t-Ratios Fell Outside the 95% Confidence Intervals of the t-Distribution

Degrees of Freedom - 5**

Coefficient	DLS	2SLS	UBK	LIML	3SLS	FIML
β_{12}	6*	2	2	1	2	2
β_{13}	3	3	3	3	2	3
γ_{11}	7*	3	2	2	2	0
γ_{12}	3	3	3	3	3	1
γ_{15}	7*	3	3	3	3	3
β_{21}	31*	3	3	2	1	1
γ_{21}	24*	3	3	2	2	0
γ_{23}	18*	3	3	3	3	2
γ_{25}	16*	6*	6*	5	4	2
γ_{27}	2	2	2	2	4	4
β_{32}	11*	4	4	3	5	2
γ_{21}	9*	5	4	4	4	1
γ_{33}	7*	4	4	4	4	3
γ_{34}	2	2	2	2	0	3
γ_{36}	3	3	3	3	1	2
Totals	149	49	47	42	40	29
Percent of t-ratios	20	7	6	6	5	4

* Significantly higher than 5% of the t-ratios at the .05 level.

** This is the number of observations minus the total number of coefficients to be estimated in the basic experiment.

*** This is the number of observations minus the number of coefficients to be estimated in each structural equation of the basic experiment.

for many of the coefficients the hypothesis that not more than five percent of the ratios would fall outside the interval. However, the hypothesis could be rejected more frequently when fifteen degrees of freedom were used than when five were used. More of the DLS ratios fell outside the intervals and for both intervals the hypothesis was frequently rejected. The hypothesis that 95 percent of the t-ratios would be less than two in absolute value fared poorly and could frequently be rejected for the consistent methods as well as for DLS.[71]

If the t-distribution with five degrees of freedom were the distribution of the t-ratios, about 92 percent of the t-ratios would fall within the 95 percent confidence interval of the t-distribution with fifteen degrees of freedom. About 89 percent of the ratios would be less than two. While the percentages of the ratios falling in the three ranges tended to be slightly smaller than suggested by the hypothesis, one could have rejected it for very few of the structural coefficients as estimated by the consistent methods using any one of the three ranges.

Before concluding that the t-ratios follow the t-distribution with five degrees of freedom, it should be remembered that the standard errors were not adjusted for degrees of freedom.[72] We might have adjusted them for the number of coefficients to be estimated in each equation. If the t-ratios thus adjusted had followed the t-distribution with fifteen

[71] In testing hypotheses about the number of t-ratios falling in given intervals, the Poisson approximation to the binomial distribution was used and judgement was based on the .05 significance level. One-tailed tests were used.

[72] Cf. Chapter II, Section 2.

degrees of freedom, about six, eleven and fifteen percent of the ratios actually used would have been expected to fall outside the three ranges examined. These percentages are slightly higher than those found for most coefficients as estimated by the consistent methods. The hypothesis that the adjusted ratios followed the suggested distribution would have been rejected seldom. Such adjustments would not have led to the DLS standard errors being reliable.

Summary. The standard errors of each of the methods tended to understate the dispersions of the estimates about the true values. There were indications that the DLS standard errors gave fairly adequate measures of the dispersions of the estimates about the (biased) central tendencies. The standard errors of the consistent methods would have been reliable for making inferences about the dispersions of the estimates if adjustment for degrees of freedom had first been made or if inferences were based on the hypothesis that the t-ratios followed the t-distribution with degrees of freedom equal to the number of observations minus the total number of coefficients to be estimated in the structure. There did not seem to be any important differences between the consistent methods in the reliability of their standard errors.

4. Estimates of the Variance-Covariance Matrix of Disturbances

The variance-covariance matrices of the structural residuals calculated for all the estimators in the replications of the basic experiment showed several interesting features. It may be noted that in analyzing the residuals, both the variances and the covariances were given equal treatment as estimates of parameters of the structure. It should also be noted that the variance-covariance matrices were not adjusted for degrees of freedom and that their elements are analyzed as being estimates of the (expected) variance-covariance matrix of the structural disturbances, Σ.

The first part of Table IV-8 shows the biases of the medians of the estimates of the elements of Σ. All were significantly biased towards zero at the .05 level. The averages of the estimates were also biased towards zero. The sign test and the Wilcoxon matched-pairs signed-ranks test both found that the central tendencies of the distributions of the estimates of any one method for most of the elements of Σ were significantly different from those for the other methods.

The rankings of the estimates by their absolute deviations from the true values followed the ranking of the estimates to be found in the biases. The second part of Table IV-8 records the totals over all replications of the ranks for the estimates of the individual elements of Σ. The same over-all ranking was found for most of the descriptive statistics of the dispersions of the estimates about the true values. In pairwise comparisons of the methods, the differences in the number of times one method came closer to the true values of Σ than another method were frequently found to be significant. In no case was a method which placed better than another in the totals of the ranks for all the estimates of the elements of Σ found to be significantly poorer than the other one in the pairwise comparisons. The frequencies with which one method beat

TABLE IV-8

The Basic Experiment - Estimates of the Elements of Σ

Biases of the Medians*

Element	Value	DLS	2SLS	UBK	LIML	3SLS	FIML
σ_{11}	35.24	-15.36	-13.27	-12.92	-12.84	-10.20	-9.05
σ_{12}	34.48	-20.26	-13.38	-12.89	-12.84	-10.71	-7.14
σ_{13}	31.12	-15.76	-12.04	-11.74	-11.50	- 8.78	-5.75
σ_{22}	36.68	-17.70	-13.16	-12.56	-10.85	-11.23	-8.12
σ_{23}	29.84	-16.84	-10.67	-10.26	- 9.11	- 6.92	-5.78
σ_{37}	40.64	-14.28	-12.47	-12.26	-11.45	-10.62	-9.05
Rankings		36	30	24	17	13	6

Ranking - Totals of Estimates by Deviations from the True Values

Element	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W**
σ_{11}	255	204	158	156	154	123	.252
σ_{12}	297	211	171	141	135	95	.579
σ_{13}	272	215	177	159	123	103	.435
σ_{22}	279	213	169	132	144	113	.433
σ_{23}	273	221	182	160	114	100	.488
σ_{33}	255	205	171	148	149	122	.264
Totals	1631	1269	1028	896	819	657	.392
Rankings	36	30	24	16	14	6	.975

Inter-quartile Ranges

Element	Value	DLS	2SLS	UBK	LIML	3SLS	FIML
σ_{11}	35.24	10.77	14.68	15.14	14.71	12.87	13.85
σ_{12}	34.48	5.53	12.97	13.70	16.96	12.92	14.18
σ_{13}	31.12	9.27	12.29	12.48	11.38	12.64	14.91
σ_{22}	36.68	9.55	15.57	16.47	18.99	12.60	15.59
σ_{23}	29.84	8.24	13.63	14.20	14.06	14.10	15.14
σ_{33}	40.64	14.13	16.67	16.86	16.97	17.70	18.85
Rankings		6	17	27	26	20	30

* All median biases are significantly different from zero at the .05 level.

** All values of W are significantly different from zero at the .01 level.

another did vary significantly among the different elements estimated. The ranking of the estimators suggested by the deviations from the true values was:

1) FIML
2) 3SLS
3) LIML
4) UBK
5) 2SLS
6) DLS

There was not as pronounced a tendency for methods to group together in the rank-totals as was found for the structural coefficients and reversals of order in the rankings of the individual estimates from the over-all ranking occurred less frequently. The ranking seemed to be slightly stronger for the covariances than for the variances in that it was found more often in the rankings of the individual estimates.

The descriptive statistics of dispersion about the central tendencies of the distributions of the estimates showed a ranking which was somewhat different from that for the dispersions of the estimates about the true values. The third part of Table IV-8 shows the inter-quartile ranges of the estimates and the totals of the ranks of the inter-quartile ranges of the estimates of each element of Σ. For the other statistics also DLS ranked as the best method and FIML tended to do most poorly.

Had we adjusted the variance-covariance matrices for degrees of freedom before analyzing them, the biases of the estimates would have been smaller but would usually have been negative except for the full-model methods which would have been almost unbiased. The adjustment would have upset the rankings by the deviations from the true values at least to the extent that the full-model methods would have lost much of their superiority over the single-equation methods.

Summary. A clear-cut ranking of the methods in estimating the elements of Σ was established. This ranking was based more on the biases of the methods than on the variability of their estimates. It could have been upset by adjusting the estimates for "lost degrees of freedom", an adjustment which would usually have improved the estimates. There were not pronounced differences in the estimates of the variances from those of the covariances of the residuals, but there were significant differences among the elements for the frequency with which one method came closer to the true values than another one.

5. Estimates of the Reduced-Form Coefficients

Another method, LSRF, was introduced into the experiments for estimating the reduced form. LSRF is the best linear unbiased estimator of the reduced-form coefficients when no account is taken of the *a priori* restrictions placed on the reduced form by the form of the structural equations. It is also the maximum-likelihood estimator if these restrictions are ignored. Comparisons of the other estimators with LSRF will indicate whether one can gain by taking the *a priori* information into account by using the other estimators even though their use is less firmly justified in other respects.

Bias. Table IV-9 shows the biases of the medians of the estimates of the reduced-form coefficients. By and large the biases are small. The biases of LSRF were usually larger than those of the other methods. Only for DLS were the biases found frequently to be significantly different from zero. The biases of the averages of the estimates were quite similar to those of the medians. Quite frequently the differences in the central tendencies of two methods were found to be significant. However, this did not occur very often for the comparisons of LSRF with the other consistent methods so one cannot safely conclude that the biases of other methods were usually significantly smaller than those of the unbiased LSRF.

Dispersions. Tables IV-10 and IV-11 show the inter-quartile ranges and the ninth-decile absolute deviations from the true values of the estimates of the reduced-form coefficients. The dispersions can hardly be judged to be small, but the LSRF dispersions were usually larger than those of other methods.

Relative Performances--Rankings of the Estimators. Table IV-12 records the totals of the ranks for the absolute deviations from the true values of the estimates of each reduced-form coefficient. As with the structural coefficients, the rank-totals for the estimators were fairly close to each other. Kendall's W was .147 for the sums of the ranks of all the estimates. It varied from .018 to .417 for the rank-totals of the estimates of the separate coefficients. For five coefficients Kendall's W was not significantly different from zero at the .05 level.

TABLE IV-9

The Basic Experiment - Biases of the Medians of the Estimates of the Reduced-Form Coefficients

True Values	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
219.40	-24.372*	-4.470	-3.980	-2.396	-3.778	- .466	-3.407
1.50	.280*	.089	.078	.064	.040*	.001	.053
1.50	.132*	- .018	- .025	.028	.002	.011	.017
.04	.000	.005	.006	.005	.000	.003	.095
2.08	.062	.008	.004	.004	- .017	- .017	.015
.18	- .035	.022	.021	.026	- .005	.008	.067
.11	.048	.007	.005	.004	.002	- .004	- .065
180.47	-23.790*	-2.781	-2.091	-1.291	-2.244	1.472	-2.005
.81	.264*	.075	.060	.043	.040*	.002	.033
1.51	.089*	- .013	- .017	- .019	- .004	- .014	- .014
.02	.001	.003	.004	.003	.000	.002	.077
2.08	.064	.011	.008	.006	- .014	- .013	.011
.10	- .010	.010	.009	.007	.000	.002	.048
.12	.039	.009	.007	.006	.004	- .005	- .050
92.34	-15.344*	-3.562	-2.881	-2.319	-1.957	-1.813	-3.159
.23	.116*	.014	.008	- .004	.009	- .002	.036
.97	.025*	- .012	- .013	- .014	- .001	- .003	- .006
.12	.016*	.008	.007	.005	.008	.003	.036
.60	.093*	.013	.008	- .005	.001	- .004	.002
.59	.036*	.029	.028	.023	.016	.019	.044
.04	.016*	.001	.000	.000	- .001	- .002	- .039
Rankings	120	100	85	70	52	49	104

* Significantly different from zero at the .05 level.

TABLE IV-10

The Basic Experiment - Inter-Quartile Ranges of the Estimates of the Reduced-Form Coefficients

True Values	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
219.395	49.969	52.019	52.430	51.753	46.563	49.026	77.575
1.497	.228	.246	.251	.264	.190	.190	.307
1.497	.193	.220	.222	.256	.239	.302	.232
.036	.056	.050	.050	.045	.023	.021	.312
2.081	.508	.504	.503	.495	.489	.476	.457
.181	.244	.199	.193	.194	.103	.073	.398
.114	.122	.133	.134	.136	.051	.052	.335
180.474	36.808	37.336	37.379	37.294	33.827	37.737	60.170
.808	.173	.184	.187	.197	.152	.147	.244
1.509	.162	.180	.181	.213	.192	.227	.186
.019	.034	.031	.030	.029	.013	.012	.267
2.084	.399	.379	.378	.371	.380	.365	.343
.098	.152	.123	.126	.122	.059	.045	.289
.121	.121	.144	.146	.148	.054	.057	.268
92.337	22.265	23.739	23.686	25.015	23.122	21.462	35.915
.234	.079	.091	.092	.094	.065	.058	.168
.968	.102	.099	.098	.099	.105	.143	.107
.116	.072	.084	.085	.086	.033	.039	.136
.604	.188	.196	.198	.200	.194	.204	.227
.588	.099	.107	.109	.107	.072	.067	.167
.035	.035	.045	.045	.045	.014	.011	.155
Totals of the ranks for each coefficient	76	87	92	98	52	55	128

TABLE IV-11

The Basic Experiment - Ninth-Decile Absolute Deviations of the Estimates of the Reduced-Form Coefficients

True Values	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
219.395	67.077	55.118	54.422	56.757	44.362	39.059	80.731
1.497	.472	.426	.419	.401	.255	.179	.532
1.497	.358	.309	.314	.318	.286	.306	.333
.036	.064	.053	.053	.052	.024	.024	.439
2.081	.563	.598	.601	.607	.524	.548	.579
.181	.245	.265	.270	.259	.101	.094	.520
.114	.146	.137	.140	.142	.088	.069	.395
180.474	53.617	42.141	42.526	45.779	34.958	32.176	64.541
.808	.441	.345	.337	.322	.204	.155	.410
1.509	.265	.229	.234	.238	.219	.236	.255
.019	.036	.031	.031	.031	.014	.013	.335
2.084	.427	.457	.457	.457	.421	.436	.465
.098	.144	.156	.158	.147	.060	.051	.419
.121	.149	.145	.147	.150	.092	.071	.311
92.337	36.139	29.614	29.201	27.105	25.443	23.481	37.147
.234	.174	.116	.119	.122	.088	.077	.230
.968	.145	.128	.129	.128	.120	.134	.140
.116	.088	.098	.098	.100	.064	.062	.198
.604	.281	.244	.243	.244	.231	.222	.279
.588	.159	.142	.141	.138	.096	.082	.271
.035	.049	.049	.049	.049	.030	.026	.185
Totals of the ranks for each coefficient	111	86	91	92	36	33	139

TABLE IV-12

The Basic Experiment - Rankings of the Estimates of the Reduced-Form Coefficients by Absolute Deviations from the True Values

Coefficient	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
π_{11}	228	200	197	196	149	166	264	.124**
π_{12}	293	216	195	194	134	130	238	.281**
π_{13}	246	184	176	190	183	211	210	.051*
π_{14}	207	204	205	193	132	123	336	.417**
π_{15}	217	207	200	205	177	182	212	.019
π_{16}	228	215	217	197	132	118	293	.304**
π_{17}	239	207	203	204	117	119	311	.391**
π_{21}	243	195	193	194	153	161	261	.134**
π_{22}	303	215	194	193	137	129	229	.297**
π_{23}	229	179	181	196	187	216	212	.032
π_{24}	217	205	201	190	126	119	342	.466**
π_{25}	221	206	204	204	175	176	214	.027
π_{26}	235	211	209	189	131	122	303	.329**
π_{27}	236	215	209	208	126	116	290	.319**
π_{31}	236	205	200	209	153	160	237	.094**
π_{32}	303	191	173	176	134	134	289	.409**
π_{33}	218	184	187	200	186	214	211	.018
π_{34}	216	197	194	197	139	165	292	.196**
π_{35}	235	189	188	204	184	194	206	.026
π_{36}	227	218	208	191	150	146	260	.146**
π_{37}	243	208	201	207	122	110	309	.400**
Totals	5020	4251	4135	4137	3127	3211	5519	.147
Totals of ranks of each line of table	135	89.5	70.5	77	34.5	46.5	135	.755**

* Significantly different from zero at the .05 level.

** Significantly different from zero at the .01 level.

TABLE IV-13

The Basic Experiment - Pairwise Comparisons of the Estimators for the Reduced-Form Coefficients

Significant Differences - Sign Test*

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	0	0	0	0	0	9
2SLS	7	-	0	0	0	0	12
UBK	7	0	-	0	0	0	13
LIML	5	0	0	-	0	0	14
3SLS	16	14	15	14	-	0	15
FIML	13	11	12	12	1	-	14
LSRF	3	0	0	0	0	0	-

Significant Differences - Wilcoxon Test**

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	0	0	0	0	0	10
2SLS	4	-	0	0	0	1	15
UBK	4	0	-	0	0	1	15
LIML	4	2	2	-	0	0	16
3SLS	16	15	15	15	-	1	15
FIML	15	15	15	15	2	-	15
LSRF	2	0	0	0	0	0	-

Number of Times Closer to the True Values***

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	---	382	386	393	263	287	619
2SLS	668	---	495	472	333	378	753
UBK	664	555	---	532	329	373	762
LIML	657	578	518	---	335	363	762
3SLS	787	717	721	715	---	493	790
FIML	763	672	677	687	557	---	783
LSRF	431	297	288	288	260	267	---

* Each row records the number of coefficients for which the sign test found the method to have significantly smaller dispersions about the true values than each of the other methods at the .05 level.

** Table entries as in the first part of the table except that the Wilcoxon signed-ranks matched pairs test was used.

*** Each row records the number of times the method came closer to the true values than each other method in all the estimates of all the coefficients.

TABLE IV-14

The Basic Experiment - Rankings of the Descriptive Statistics of Dispersions of the Estimates of the Reduced-Form Coefficients*

Statistic	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
Standard Deviations	71	87	103	104	40	36	147	.721**
Inter-quartile Ranges	76	87	92	98	52	55	128	.335**
80% Ranges	78	91	103	97	38	37	144	.692**
Ranges	75	93	107	93	40	33	147	.751**
Root Mean Square Error	125	74	92	86	38	31	142	.821**
Median Absolute Deviations	114	83	79	88	47	54	123	.383**
3rd Quartile Absolute Deviations	125	83	68	78	50	47	137	.592**
9th Decile Absolute Deviations	111	86	91	92	36	33	139	.711**
Largest Absolute Deviations	103	80	93	92	32	41	147	.732**

* Table entries are the totals of the ranks of the descriptive statistics of the individual reduced-form coefficients.

** Significant at the .01 level.

The ranks of the rank-totals for the separate coefficients showed quite high agreement. Kendall's W was .76 for the sums of these ranks over all the coefficients.

An over-all ranking of the methods into three groups was suggested by the rank-totals of the individual estimates. In the first group were 3SLS and FIML with 3SLS slightly the better. In the second group fell UBK, LIML and 2SLS. The last group held DLS and LSRF with DLS seeming to be slightly better than LSRF.

The descriptive statistics of dispersion of the estimates about the true values showed the same ranking of the estimators into three groups as did the separate estimates. Table IV-14 presents the totals of the ranks for the descriptive statistics of the reduced-form coefficients. Reversals of order in the rankings for the individual coefficients from that suggested by the totals for all coefficients were not entirely confined to methods which were grouped together in the rankings. FIML and 3SLS were not always the best; DLS and LSRF were not always the poorest methods. It is interesting that LSRF had the largest absolute deviations from the true values of all the methods for each of the coefficients. The numbers of times the estimates had the wrong signs and the numbers of times they came within twenty or forty percent of the true values showed the same rankings of the methods into groups as had the other descriptive statistics.

The rankings of the descriptive statistics of dispersion about the central tendencies of the estimates placed DLS between the full-model methods and the other k-class estimators. Otherwise, the rankings of the methods were the same as found for the dispersions about the true values.

For the reduced-form coefficients as for the structural coefficients, FIML was the method which made the best estimate most frequently. LSRF made the worst estimates most often with DLS making them the second most frequently. FIML, DLS and LSRF amde both the best and the worst estimates more frequently than did the consistent k-class estimators. 2SLS made the worst estimate only once.

Relative Performances--Pairwise Comparisons. Table IV-13 summarizes the findings of the pairwise comparisons of the estimators by their deviations from the true values of the reduced-form coefficients. Only in the comparisons of the full-model methods with the k-class estimators and of LSRF with the other methods were significant differences between the estimators found for many coefficients. As in the structural coefficients, it was found that the frequency with which any one method came closer to the true values than another varied significantly among the coefficients estimated.

The third part of Table IV-13 shows the number of times one method came closer to the true values than each other method in all the estimates of the reduced-form coefficients. FIML was closer to the true values more often than 3SLS even though 3SLS slightly surpassed FIML in the rank-totals of the estimates. Otherwise the table indicates the same relative standings of the methods as did the rank-totals for all the separate estimates. In particular, DLS was slightly better than LSRF.

Summary. Bias was not a serious problem for the consistent methods in estimating the reduced-form coefficients. A ranking of the methods into three groups was suggested which followed the ranking of the methods based on their asymptotic properties. LSRF was apt to be ranked as the poorest method. There seemed to be a definite gain from taking into account the restrictions on the reduced-form implied by the form of the structural equations. This was true, though not very clearly, even for DLS.

6. Predictions of the Endogenous Variables

Three predictions of each endogenous variable were made for observations not used in estimating the structure in each replication. These predictions were compared with the values of the endogenous variables for the observations before the disturbances were added to the endogenous variables.

Descriptive Statistics. Table IV-15 shows some of the descriptive statistics of the distributions of the predictions. The biases were very small and were very rarely significant. The dispersions also were small. The inter-quartile ranges of the predictions averaged less than five percent of the predicted values for each method.

Relative Performances. Table IV-16 records the rank-totals of the predicted values from the true ones for each of the values predicted and the rank-totals for the descriptive statistics of the distributions of the predictions. As was found for the reduced-form coefficients, the rank-totals of the estimates of the methods were all fairly close together. For the totals of the ranks for all predictions over all replications of

TABLE IV-15

The Basic Experiment - Descriptive Statistics for the Predictions**

Biases of the Medians

Variable	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
$y_{1,T+1}$	-.006	-.005	-.005	-.004	-.004	.000	-.004
$y_{2,T+1}$	-.009	-.005	-.003	-.002	-.002	.000	-.002
$y_{3,T+1}$	-.021*	-.006	-.005	-.002	-.002	.001	-.003
$y_{1,t+2}$	.013*	-.003	-.004	-.006	.005	.006	-.009
$y_{2,T+2}$	.018*	-.002	-.003	-.006	.006	.006	-.008
$y_{3,T+2}$	-.005	-.003	-.002	-.002	.000	.003	-.011
$y_{1,T+3}$	-.003	.001	.001	.001	-.002	-.001	.011
$y_{2,T+3}$	-.001	.000	.000	.000	-.002	.000	.010*
$y_{3,T+3}$	.006	.005	.005	.004	.000	.001	.010*

Inter-Quartile Ranges

Variable	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
$y_{1,T+1}$	.029	.029	.029	.030	.032	.031	.040
$y_{2,T+1}$	.029	.028	.028	.029	.029	.029	.039
$y_{3,T+1}$	.026	.028	.028	.028	.034	.030	.041
$y_{1,T+2}$	.064	.053	.053	.059	.049	.051	.052
$y_{2,T+2}$	.065	.065	.064	.063	.050	.052	.052
$y_{3,T+2}$	.055	.051	.051	.050	.049	.049	.070
$y_{1,T+3}$	.018	.020	.020	.020	.015	.015	.045
$y_{2,T+3}$	.016	.017	.018	.018	.012	.013	.043
$y_{3,T+3}$	.018	.019	.019	.018	.016	.019	.046

Ninth-Decile Absolute Deviations

Variable	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
$y_{1,T+1}$	.053	.052	.052	.052	.045	.043	.061
$y_{2,T+1}$	.051	.048	.048	.046	.041	.040	.056
$y_{3,T+1}$	.057	.053	.052	.051	.047	.047	.061
$y_{1,T+2}$	.081	.072	.072	.072	.064	.054	.083
$y_{2,T+2}$	.084	.072	.071	.071	.062	.056	.083
$y_{3,T+2}$	.065	.066	.066	.066	.066	.062	.098
$y_{1,T+3}$	.025	.024	.024	.023	.021	.034	.046
$y_{2,T+3}$	.020	.023	.023	.023	.021	.023	.044
$y_{3,T+3}$	.023	.024	.024	.024	.018	.020	.047

* Significantly different from zero at the .05 level.

** Table entries are the descriptive statistics of the distributions of the predicted values divided by the true values of the variables predicted.

TABLE IV-16

The Basic Experiment - Rankings of the Estimators for the Predictions

Ranking - Totals of the Individual Estimates by Deviations from the True Values

Variable	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
$y_{1,T+1}$	195	206	209	206	180	163	241	.052*
$y_{2,T+1}$	217	208	204	203	170	163	235	.055*
$y_{3,T+1}$	256	189	176	188	189	173	229	.081**
$y_{1,T+2}$	222	201	203	215	175	169	215	.036
$y_{2,T+2}$	227	198	196	216	177	174	212	.034
$y_{3,T+2}$	197	184	201	230	170	163	255	.092**
$y_{1,T+3}$	194	192	191	186	167	168	302	.184**
$y_{2,T+3}$	184	195	197	198	170	170	296	.164**
$y_{3,T+3}$	207	205	205	183	162	153	285	.161**
Totals	1899	1768	1782	1825	1560	1496	2270	.067
Rank - totals for ranks of each line	49	37.5	37.5	40	19	10.5	58.5	.721**

Rank - Totals for the Descriptive Statistics

Statistic	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
Standard Deviation	33	37	43	46	15	15	63	.780**
Inter-quartile Ranges	39	35	35	37	22	30	54	.250*
80% Ranges	38	34	43	46	16	12	63	.821**
Ranges	30	37	43	50	12	17	63	.859**
Root mean square Error	39	38	41	43	15	13	63	.787**
Median Absolute Deviation	42	41	38	33	17	20	61	.580**
3rd Quartile Absolute Deviation	33	40	35	42	20	20	62	.551**
9th Decile Absolute Deviation	43	41	38	31	29	17	62	.616**
Largest Absolute Deviation	40	35	38	47	12	17	63	.797**

* Significantly different from zero at the .05 level.
** Significantly different from zero at the .01 level.

the experiment, Kendall's W was only .067. The value of W for several of the values predicted was not significantly different from zero at the .01 level.

The ranking of the methods into three groups which was established for the reduced-form coefficients was found for the predictions with only one major exception. DLS tended to be ranked only slightly worse than the consistent k-class methods and might better be grouped with them than with LSRF. There was not complete agreement over the rankings either among the individual estimates, or among the ranks of the rank-totals for the separate values predicted, or among the ranks of the descriptive statistics for the individual values. In the rankings of the methods by the descriptive statistics of dispersion about the central tendencies of the predictions, DLS was apt to be placed ahead of the other k-class estimators.

Significant differences between the estimators occurred fairly infrequently in the pairwise comparisons of the predictions by their deviations from the true values. Only in the comparisons of LSRF with the other methods were significant differences found for a majority of the predicted values by either the sign test or the Wilcoxon test.

Summary. The performances of the estimators in making predictions were similar to their performances in estimating the reduced-form coefficients. However, DLS seemed to be slightly better at prediction than at estimation of the reduced form. The dispersions of the estimates were small and all methods seemed to be fairly accurate predictors.

7. The Findings of the Basic Experiment

Several important conclusions are suggested by the results of the basic experiment:

1. Bias is not a serious problem for the consistent estimators.
2. The differences in the performances of the methods are not very pronounced.
3. The exact frequencies with which one method surpasses another differs among the coefficients of a structure.
4. The results suggest an over-all ranking of the methods by their dispersions from the true values. This ranking places the methods in three groups: the full-model methods, the consistent k-class estimators and the other methods. This follows the asymptotic ranking of the methods.
5. DLS had almost as small dispersions about its central tendencies as had the full-model methods about theirs and surpassed the other methods according to this criterion.
6. The standard errors of the consistent methods were reliable for making inferences if they were adjusted for degrees of freedom or if they were supposed to follow the t-distribution with a small number of degrees of freedom.
7. The elements of the variance-covariance matrices of residuals tended to underestimate the elements of the variance-covariance matrices of the structural disturbances.

These findings are based on the results of only one experiment. In the next chapter we investigate whether they will be found in the results of other experiments which also use data generated in conformance to the model (1.1)-(1.6).

CHAPTER V

The Estimators in the Standard Model

1. Introduction

A fairly large proportion of the Monte Carlo experiments reported in this study used data generated in conformance with the Cowles-Commission model (1.1)-(1.6). These experiments investigated the performances of the methods of estimating systems of linear, simultaneous equations for a variety of specific models which met the assumptions of the model. The findings of these experiments are presented in this chapter and the next one.

The discussion of this chapter relies on the results of the basic experiment in that the results of other experiments are compared with those of the basic experiment. The discussion of each experiment is much shorter than that of the basic experiment. This is not because the experiments were analyzed less fully. Each experiment was analyzed separately in the same way as the basic experiment, but the results are presented only if they differ substantially from those of the basic experiment or are of especial interest.

2. Different Disturbance Sets

The second and third experiments used two sets of disturbances not employed in the basic experiment. Different sets of disturbances were obtained by specifying different numbers with which to begin the sequence of pseudo-random numbers used in obtaining the disturbances.[73] The experiments using the different sets of disturbances were conducted to test whether the results of the experiments seemed to be dependent on the sets of disturbances used. In particular, they investigated whether the disturbances of the basic experiment were such that the results of that experiment were a misleading representation of the performances of the methods in estimating the structure. These experiments also served to triple the size of the sample of estimates of the basic structure, allowing more sharpness in the investigations of the distributions of the estimates.[74]

There was substantial agreement among the results of the experiments about the ranking of the methods by their dispersions about the true values of the parameters being estimated. This was true for

[73]The generation of the disturbances and the pseudo-random-number generator are described in Appendix A. Program-parameter 14 specified the set of disturbances to be used in an experiment in the sense of the starting point for the sequence of random numbers used. Different sets were obtained by setting the program-parameter equal to 0, 1, or 2.

[74]The three experiments were analyzed separately in full. The combined estimates of the three experiments were given only partial analysis. This was done for three reasons. First, a full, combined analysis would have required substantial amounts of additional programming. Second, many of the experiments reported on later in this study used data which were closely related to those of the basic experiment and it was desired to check that the results of the basic experiment with which these other experiments were to be compared were not atypical. This was done by comparing them with the results of experiments using independent data. It would have been impossible if the results of the three experiments had been combined before they were analyzed. Third, since other experiments are related closely to the basic experiment, it was important to have an analysis for the fifty sets of estimates in that experiment alone. Having obtained this, it was easier to analyze the other experiments separately and then compare the results.

each of the types of parameters examined. With only a very few exceptions, the reversals in the order of the methods in one experiment from that suggested by the totals of the ranks for all the experiments were between methods which had been grouped together in the results of the basic experiment and not between methods of different groups. This was true for the over-all rankings of the methods suggested in each experiment by the ranks of the absolute deviations of the estimates from the true values. It also held for the rankings based on the descriptive statistics of the distributions of the estimates. There was also close agreement among the experiments for the rankings of the methods into groups by their dispersions about the central tendencies of the estimates. This was especially true if DLS was grouped with the full-model methods.

Rankings of the methods for several of the statistics used to analyze the experiments are shown in Table V-1. The sums of the ranks over the parameters for the statistics in each experiment were totalled over the three experiments. The rankings of the totals appear in the table. The values of Kendall's W for the totals of the ranks for the statistics of the individual parameters in the three experiments are also presented in the table. The totals over the parameters of the ranks in each experiment were themselves ranked. Kendall's W for the totals of these ranks over the three experiments is shown in the last column of the table. The first three lines of the A-set of tables in Appendix C show more fully the results of the experiments.

The hypothesis that the distributions of the estimates in the three experiments were the same was tested in two ways. First, the distributions of the estimates of each parameter of each method in the three experiments were compared. The Komolgorov-Smirnov[75] test could very rarely reject the hypothesis that the distributions were the same in the different experiments. Second, the frequencies with which each method came closer to the true values than each of the other methods were examined. The frequencies for each pair of methods in estimating each parameter in the experiments were compared. The χ^2 test could reject very infrequently the hypothesis that the frequencies did not differ among the experiments. On the basis of the results of these tests, and the substantial agreement among the experiments as to the rankings of the methods, we conclude that the estimates of the basic experiment were not an atypical sample of estimates of the basic structure.

The figures in Appendix E present the cumulative distributions of the estimates made in the three experiments. The cumulative distributions provided confirmation of our conclusion that the methods fall into three distinct groups whose performances are substantially different while differences between members of any one group are much less pronounced. So close were the consistent k-class estimators to each other that it was usually impossible to plot the three distributions in such a way that they could be distinguished from each other. The same was true for the full-model methods.

[75] This test is described in Sidney Siegel, Nonparametric Statistics for the Behavioral Sciences, (New York: McGraw-Hill, 1956), pp. 127-136.

Comparisons of the cumulative distributions tend to over-emphasize the differences between the groups of methods. In forming the cumulative distributions, the relationships between the estimates of the methods in the separate replications were broken. The estimates falling at any given percentile of the cumulative distributions of different methods were not usually made in the same replication. The differences in the performances of the methods as suggested by the number of times each method was closer to the true values than each of the others were much smaller than might be suggested by the seemingly clear-cut differences shown by many of the cumulative distributions.

One conclusion drawn from the results of the basic experiment has to be altered slightly. This is the conclusion that a clear-cut ranking for the estimates of the elements of Σ could be made. In the other two experiments there was a slightly higher occurrence of reversals of order of the methods from the over-all ranking for some of the statistics for the elements of Σ. These reversals of order occurred largely among methods whose performances could be grouped together in estimating other parameters. The cumulative distributions of the estimates also indicated that the estimates of Σ by the different methods could be grouped as were the estimates of the other parameters, though the grouping together of the methods was not as pronounced as for the estimates of other parameters.

The combined estimates of the three experiments allow us to study the performances of the standard errors more closely. The number of times the t-ratios[76] fell outside the three intervals examined is shown in Table V-2. The hypothesis that confidence intervals based on the t-distribution with five degrees of freedom are reliable fared better than the hypothesis that fifteen is the appropriate number of degrees of freedom for the t-distribution. The former hypothesis could be rejected for none of the t-ratios of the consistent methods at the .01 level. The latter hypothesis was rejected for a few coefficients. The percentages of the t-ratios falling outside the intervals were closer to those to be expected under the hypothesis that the appropriate degrees of freedom is the number of observations minus the total number of coefficients to be estimated than to those expected under the other hypothesis. The t-ratios of the consistent methods fell outside the intervals less frequently than would be expected if the t-ratios, adjusted for degrees of freedom, followed the t-distribution with degrees of freedom equal to the number of observations minus the number of coefficients to be estimated in each equation. The numbers of times they fell outside the intervals were not, however, significantly smaller. The DLS t-ratios usually fell outside the intervals significantly more often than would be expected under any one of the hypotheses.

<u>Summary</u>. The experiments which used other disturbance sets had the same results as the basic experiment. Comparisons of the three experiments indicated that the sample of estimates of the basic experiment was not an unrepresentative one.

[76] The t-ratios are the deviations of the estimates from the true values divided by their standard errors.

TABLE V-1

Ranking of Methods, First Three Experiments

Aspect	Measure	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W for Estimates of Separate Coefficients	Kendall's W for Totals of Separate Experiments
Structural Coefficients	Standard Deviation	1	4	5	6	2	3		.80	.90
	Interquartile Range	1	4	6	5	2	3		.46	.87
	Root Mean Square Error	6	3	4	5	1	2		.81	.98
	Median Absolute Deviation	6	4	5	3	1	2		.60	.87
	Largest Deviation	3	4	5	6	1	2		.52	.93
	Number in 20% of True Value	6	3	4	5	2	1			.95
	Rankings of all Estimates	6	4	3	5	1	2		.13	.91
	Ranking of rank-totals of each Coefficient Estimated	6	5	3	4	1	2		.78	.86
Residual Variance-Covariance	Standard Deviation	1	2	3	4	5	6		.78	1.00
	Interquartile Range	1	2	5	4	3	6		.59	.86
	Root Mean Square Error	6	5	4	3	1	2		.65	.92
	Median Absolute Deviation	6	5	4	3	2	1		.80	1.00
	Largest Deviation	5	3	1	4	2	6		.12	.38
	Number in 20% of True Value	6	4	5	3	2	1			.90
	Rankings of all Estimates	6	5	4	3	2	1		.35	1.00
	Rankings of rank-totals of each Variance-Covariance Element	6	5	4	3	2	1		.97	1.00

TABLE V-1 (Cont.)

Aspect	Measure	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W for Estimates of Separate Coefficients	Kendall's W for Totals of Separate Experiments
Reduced Form Coefficients	Standard Deviation	3	4	5	6	2	1	7	.63	.95
	Interquartile Range	4	3	5	6	1	2	7	.39	.84
	Root Mean Square Error	6	3	4	5	2	1	7	.73	.98
	Median Absolute Deviation	6	5	4	3	2	1	7	.50	.94
	Largest Absolute Deviation	6	3	5	4	1	2	7	.52	.94
	Number in 20% of True Value	6	4.5	4.5	3	2	1	7	.14	.97
	Rankings of all Estimates	6	5	3	4	1	2	7	.66	.96
Predictions	Standard Deviation	3	4	5	6	2	1	7	.85	.97
	Interquartile Range	4	3	6	5	2	1	7	.40	.84
	Root Mean Square Error	4	3	6	5	2	1	7	.81	.97
	Median Absolute Deviation	6	5	4	3	2	1	7	.50	.97
	Largest Absolute Deviation	3	4	5	6	2	1	7	.78	.87
	Coefficient of Inequality	4	3	6	5	2	1	7	.81	.96
	Rankings of all Predictions	6	3	5	4	2	1	7	.07	.98
	Rankings of Rank-totals of each Quantity Predicted	6	3	4	5	2	1	7	.69	.94

TABLE V-2

Number of t-Ratios Falling outside the Three Intervals in the First Three Experiments

Outside Interval -2 to +2

Coefficient	DLS	2SLS	UBK	LIML	3SLS	FIML
β_{12}	45*	14*	13	14*	15*	12
β_{13}	27*	18*	17*	15*	10	21*
γ_{11}	45*	17*	17*	18*	20*	19*
γ_{12}	24*	15*	15*	15*	16*	16*
γ_{15}	40*	18*	18*	17*	16*	15*
β_{21}	110*	17*	13	12	15*	8
γ_{21}	99*	15*	13	12	18*	14*
γ_{23}	78*	14*	14*	12	14*	10
γ_{25}	62*	16*	18*	16*	24*	23*
γ_{27}	9	14*	14*	14*	9	19*
β_{32}	51*	19*	18*	18*	17*	14*
γ_{31}	48*	19*	20*	18*	18*	14*
γ_{33}	30*	19*	18*	18*	14*	17*
γ_{34}	15*	14*	14*	13	9	21*
γ_{36}	18*	17*	16*	16*	13	15*
Totals	701	247	239	229	228	238
Percentages	31.2	11.0	10.6	10.2	10.1	10.6

TABLE V-2 (Cont.)

Outside 95% Confidence Interval of t-Distribution with Fifteen Degrees of Freedom

Coefficient	DLS	2SLS	UBK	LIML	3SLS	FIML
β_{12}	45*	12	12	9	10	11
β_{13}	26*	12	13	11	7	15*
γ_{11}	28*	15*	14*	15*	18*	12
γ_{12}	22*	13	12	12	13	11
γ_{15}	37*	14*	13	14*	14*	12
β_{21}	102*	12	12	10	11	7
γ_{21}	95*	11	13	10	14*	12
γ_{23}	71*	11	10	11	11	4
γ_{25}	57*	17*	17*	13	17*	16*
γ_{27}	7	12	12	12	8	15*
β_{32}	44*	14*	14*	12	12	13
γ_{31}	45*	18*	17*	13	16*	13
γ_{33}	21*	14*	13	15*	12	10
γ_{34}	14*	11	11	11	8	13
γ_{36}	15*	15*	15*	13	11	12
Totals	626	201	198	181	182	176
Percentages	28.3	8.9	8.8	8.0	8.1	7.8

TABLE V-2 (Cont.)

Outside 95% Confidence Interval of t-Distribution with Five Degrees of Freedom

Coefficient	DLS	2SLS	UBK	LIML	3SLS	FIML
β_{12}	26*	6	6	5	8	4
β_{13}	13	6	6	6	3	8
γ_{11}	24*	9	8	8	10	5
γ_{12}	15*	9	9	9	10	4
γ_{15}	25*	8	8	8	10	5
β_{21}	83*	6	6	5	3	5
γ_{21}	71*	7	7	6	6	3
γ_{23}	43*	8	7	6	5	2
γ_{25}	43*	11	11	7	8	5
γ_{27}	22	4	4	4	5	9
β_{32}	27*	10	10	9	9	5
γ_{31}	27*	10	9	9	7	4
γ_{33}	13	8	8	7	7	5
γ_{34}	5	6	6	6	3	8
γ_{36}	6	6	6	6	4	7
Totals	421	114	111	101	98	77
Percentages	18.7	5.0	4.9	4.5	4.4	3.4

*Significantly different from five percent at the .01 level.

3. Different Exogenous Data Sets

The distributions of the estimates studied in the basic experiment were based on a fixed set of exogenous data which did not vary in the replications. To check whether the results obtained depended on the set of exogenous data used, two experiments were conducted using different sets of exogenous data. The different sets were obtained by starting the sequence of random numbers used to generate the exogenous variables with different numbers.[77] The fourth and fifth experiments of the first part of the experiment design used different sets of exogenous data.[78]

Fairly high agreement among the results of the experiments using different sets of exogenous data was found. The concordances between the rankings, as measured by Kendall's W, were as high or higher than those shown in Table V-1. Almost without exception the reversals of order in the rankings of the experiments occurred between methods whose behavior was found to be similar to each other in the basic experiment and not between members of groups whose performances had been found to be different.

Pairwise comparisons of the methods revealed a few interesting differences between the experiments. Tables V-3, V-4, and V-5 show the

---

[77] Program-parameter 15 specified the set of exogenous data to be used. The generation of the data is discussed in Chapter III, Section 2 and Appendix A. The sets of data are referred to as data sets 0, 1, and 2.

[78] These are experiments D and E of the first set of experiments, Table 1 of Appendix C. Summary statistics for them are found in the A-set of tables in Appendix D.

TABLE V-3

Different Sets of Exogenous Data - Significant Differences Between Experiments using Data Sets 0 and 1*

Structural Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	6	6	6	6	6
2SLS	2	-	1	1	0	1
UBK	2	1	-	1	0	0
LIML	2	1	0	-	0	0
3SLS	1	0	0	0	-	0
FIML	1	0	0	0	0	-

Elements of Σ

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	0	0	0	0	1
2SLS	0	-	0	0	1	0
UBK	0	0	-	0	0	0
LIML	0	0	0	-	1	0
2SLS	0	0	0	0	-	0
FIML	0	0	0	0	0	-

Reduced-Form Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	6	6	6	1	0	2
2SLS	0	-	0	0	0	2	0
UBK	0	0	-	0	0	2	0
LIML	0	1	0	-	0	0	0
3SLS	0	0	1	0	-	0	0
FIML	0	0	0	0	0	-	0
LSRF	0	0	0	0	0	0	-

Predictions

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	0	0	0	0	0	0
2SLS	3	-	0	0	0	0	0
UBK	3	3	-	0	0	0	0
LIML	2	0	0	-	0	0	0
2SLS	0	0	0	0	-	0	0
FIML	0	0	0	0	0	-	0
LSRF	1	0	0	0	0	0	-

*See notes following Table V-5.

TABLE V-4

Different Sets of Exogenous Data - Significant Differences Between Experiments using Data Sets 0 and 2*

Structural Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	3	3	3	1	1
2SLS	2	-	4	4	0	0
UBK	2	3	-	0	0	0
LIML	2	3	1	-	0	0
3SLS	0	0	0	0	-	0
FIML	0	0	0	0	0	-

Elements of Σ

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	0	0	0	0	0
2SLS	0	-	0	0	1	0
UBK	0	0	-	0	1	-
LIML	0	0	0	-	1	1
3SLS	0	0	0	0	-	0
FIML	0	0	0	0	0	-

Reduced-Form Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	0	0	0	0	0	2
2SLS	2	-	0	0	0	0	0
UBK	3	2	-	0	0	0	0
LIML	3	1	0	-	0	0	1
3SLS	0	0	0	1	-	0	0
FIML	0	0	0	0	0	-	0
LSRF	0	0	0	0	0	0	-

Predictions

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	3	3	2	2	1	3
2SLS	0	-	0	1	0	1	0
UBK	0	0	-	0	0	0	0
LIML	0	0	0	-	0	0	0
2SLS	0	0	0	0	-	0	0
FIML	0	0	0	0	0	-	0
LSRF	0	0	0	0	0	0	-

*See notes following Table V-5.

TABLE V-5

Different Sets of Exogenous Data – Significant Differences Between Experiments using Data Sets 1 and 2*

Structural Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	0	0	0	0	0
2SLS	4	-	0	0	0	0
UBK	3	2	-	0	0	0
LIML	3	2	0	-	0	0
3SLS	2	0	0	0	-	0
FIML	3	0	0	0	0	-

Elements of Σ

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	0	0	0	0	0
2SLS	0	-	0	0	0	0
UBK	0	0	-	0	0	0
LIML	0	0	0	-	0	0
3SLS	0	0	0	0	-	0
FIML	0	0	0	0	0	-

Reduced-Form Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	0	0	0	1	0	3
2SLS	6	-	0	0	0	0	2
UBK	6	1	-	0	0	0	1
LIML	5	1	0	-	0	0	2
3SLS	1	0	0	0	-	0	0
FIML	1	0	0	0	0	-	0
LSRF	4	0	1	1	0	0	-

Predictions

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	5	5	4	4	3	5
2SLS	0	-	4	1	0	1	0
UBK	0	0	-	1	1	0	0
LIML	0	0	0	-	0	0	0
3SLS	0	0	0	0	-	0	0
FIML	0	0	0	0	0	-	0
LSRF	0	0	0	0	0	0	-

*Each row of Tables V-3 to V-5 shows the number of parameters estimated for which one method was found to have been closer to the true values than each other method significantly more often in the experiment first mentioned than in the other one at the .05 level. The McNemar test was used for these tests. There were fifteen structural coefficients, six elements of Σ, twenty-one reduced-form coefficients and nine predictions.

number of parameters for which the McNemar test found significant differences in the frequencies with which one method came closer to the true values than another in the experiments.[79] Significant differences occurred most frequently for the performances of the k-class estimators relative to each other, especially in the performances of DLS relative to the other k-class methods. While significant differences did not usually occur very frequently, there were cases where they were found for one-half of the parameters of a particular type which were estimated. There did seem to be enough significant differences to indicate that the performances of the k-class estimators relative to each other are sensitive to the sets of exogenous data.

<u>Summary</u>. While the results of the experiments using other sets of exogenous data supported the rankings of the methods established in the basic experiment, there were strong indications that the frequencies with which one method surpassed another depended on the exogenous data sets used.

4. Different Structures--Structural Coefficients

Four experiments were conducted using the same model as in the basic experiment but with different values for the structural

[79]The .05 significance level was used for these tests. Note that this is the significance level for the individual tests, not for the number of parameters for which differences occurred.

coefficients.[80] The variance-covariance matrices of the structural disturbances were the same as in the basic experiment. Indeed, the sets of disturbances and the set of exogenous variables used in these experiments were identical to those of the basic experiment. The structural coefficients of these structures are shown in Table V-6 together with the values of the determinants of the B matrices of the structures. The structural coefficients of MOD37A2, the structure of the basic experiment, are also found in the table. The experiments using the different structures investigated whether the values of the structural coefficients affected the results of the study.

On the whole, the results of the experiments which used different structural coefficients were in substantial agreement with each other. Table V-7 shows the over-all rankings of the methods in the five experiments according to several of the measures studied. However, the concordances between the ranks for the separate experiments were not very great. This is shown by the values of Kendall's W for the sums of the ranks assigned to the rank-totals over the individual parameters of the statistics in the separate experiments. While much of the differences in the values of W from unity can be attributed to reversals of order between the methods which had been grouped together in the results of the basic experiment, important differences in the relative standings

[80] Experiments F, G, H. and I of the first set of experiments used the different structures. (The experiments conducted are shown in Appendix C.)

of the methods in the different groups occurred. It may be noted that the values of Kendall's W in the last column of Table V-7 are usually lower than those in Table V-1 as a result of these reversals of order among the methods.

Table V-8 shows the totals of the ranks for all the estimates of the parameters in each of the experiments summed over all parameters and replications. The ranking of the estimates was by their deviations from the true values of each parameter. Table V-9 records the sums over the parameters of the ranks of the totals of the ranks of the estimates of the separate parameters in each experiment. Several points arise from an examination of these tables. In all the experiments, the rank-totals for the estimates of the different methods were close together and did not suggest very pronounced differences in the relative performances of the methods. The over-all ranking of the methods suggested by the rank-totals was not the same in the different experiments. The differences in the orders of the methods were not entirely confined to methods which were grouped together in the basic experiment. For example, UBK ranked ahead of 3SLS and FIML in MOD37A5 for the ranks of the rank-totals of the individual structural coefficients. The amount of agreement among the ranks of the rank-totals of the separate parameters differed substantially among the experiments for the structural coefficients and the reduced-form coefficients. This was not true for the elements of Σ nor the predictions where in each experiment the agreement was high.

The variations in the amounts of agreement among the rankings of the descriptive statistics of the distributions of the individual parameters were also considerable. The McNemar test found for many of the descriptive statistics that the number of coefficients for which one method was better than another differed significantly among the experiments.

Table V-10 shows the findings of the McNemar test in comparing the estimates of the methods in three of the experiments.[81] Although the table shows the results only for the structural coefficients and the reduced-form coefficients, the estimates of Σ and the predictions were also compared. Significant differences in the relative performances of the methods in the experiments for these parameters also occurred, though not quite as frequently.

The McNemar test found a large number of significant differences between the experiments. This was particularly so in the comparisons of the estimates of the reduced-forms in MOD37A4 and MOD37A5 where significant differences were found for all but one of the coefficients in some of the comparisons. One can conclude from the results of these tests that the frequencies with which one method came closer to the true values of the parameters than others did depend on the true values of the structural coefficients.

It is worth noting what conclusions might have been drawn if the experiment using MOD37A5 were the only one conducted. 3SLS would have been grouped with the consistent k-class methods for the estimates

[81] Only three of the experiments were compared with each other by the test. For the reasons for not conducting all the possible tests, see Chapter III, Section 11.

TABLE V-6

Different Structural Coefficients

Coefficient	MOD37A2	MOD37A1	MOD37A3	MOD37A4	MOD37A5
β_{12}	.89	.64	.76	.38	.69
β_{13}	.16	.22	.04	.44	.21
β_{21}	.74	.72	.46	.27	.33
β_{32}	.29	.17	.58	.49	.77
γ_{11}	44.00	32.00	75.00	61.00	88.00
γ_{12}	.74	.65	.87	.72	.08
γ_{15}	.13	.52	.27	.24	.82
γ_{21}	62.00	18.00	48.00	55.00	73.00
γ_{23}	.70	.68	.76	.59	.32
γ_{25}	.96	.67	.89	.26	.60
γ_{27}	.06	.34	.30	.12	.08
γ_{31}	40.00	37.00	83.00	39.00	43.00
γ_{33}	.53	.95	.38	.19	.41
γ_{34}	.11	.39	.44	.95	.71
γ_{36}	.56	.18	.49	.10	.31
det B	.31	.51	.64	.84	.72

TABLE V-7

Ranking of Measures of Performance of the Methods in Experiments with Different Structural Coefficients

Aspect	Measure	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	W for rank-totals of separate estimates	W for experiment rank-totals
Structural Coefficients	Standard Deviation	1	3.5	5	6	2	3.5		.59	.78
	Interquartile Range	1	3	5	6	2	4		.34	.76
	Root Mean Square Error	5	2	4	6	1	3		.42	.70
	Median Absolute Deviation	6	5	4	3	2	1		.29	.76
	Third Quartile Absolute Deviation	6	3	4	5	1	2		.42	.86
	Largest Deviation	2	4	5	6	1	3		.36	.68
	No. of Wrong Signs	6	3	4.5	4.5	2	1			.69
	No. in 20% of True Value	6	5	3	4	2	1			.85
	Rankings of all Estimates	6	4	3	5	2	1		.08	.90
	Rankings of rank-totals for each Coefficient Estimated	6	4	3	5	2	1		.44	.85
Residual Variance-Covariance	Standard Deviation	1	2	4	5	3	6		.84	.96
	Interquartile Range	1	2	4	5	3	6		.63	.95
	Root Mean Square Error	6	3.5	5	3.5	1	2		.41	.63
	Median Absolute Deviation	6	5	4	3	2	1		.82	.98
	Largest Deviation	4	2	3	5	1	6		.19	.40
	No. in 20% of True Values	6	4	5	3	2	1			.89
	Ranking of all Estimates	6	5	4	3	2	1		.37	1.00
	Ranking of rank-totals for each element Estimated	6	5	4	3	2	1		.96	1.00

TABLE V-7 (Cont.)

Aspect	Measure	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	W for rank-totals of separate estimates	W for experiment rank-totals
Reduced-Form Coefficients	Standard Deviation	2	4	6	5	3	1	7	.44	.58
	Interquartile Range	3	4	5	6	1	2	7	.24	.83
	Root Mean Square Error	6	3	5	4	2	1	7	.36	.43
	Median Absolute Deviation	6	5	4	3	2	1	7	.29	.82
	Third Quartile Absolute Deviation	6	3	4	5	2	1	7	.63	.93
	Largest Deviation	6	3	5	4	2	1	7	.25	.38
	No. in 20% of True Value	7	4	3	5	2	1	6		.90
	Ranking of all Estimates	7	5	4	3	2	1	6	.11	.86
	Ranking of rank-totals for each Coefficient Estimated	7	5	4	3	2	1	6	.52	.82
Prediction	Standard Deviation	2	4	6	5	1	3	7	.35	.53
	Interquartile Range	3	4	5	6	1	2	7	.44	.79
	Root Mean Square Error	6	3	5	4	2	1	7	.33	.44
	Median Absolute Deviation	6	5	4	3	2	1	7	.55	.83
	Third Quartile Absolute Deviation	6	3	4	5	2	1	7	.41	.61
	Largest Deviation	6	4	3	5	1	2	7	.26	.45
	Coefficient of Inequality	6	3	5	4	2	1	7	.33	.44
	Ranking of all Predictions	6	5	3	5	2	1	7	.06	.83
	Ranking of rank-totals of each Quantity Predicted	6	4	3	5	2	1	7	.70	.83

TABLE V-8

Different Structural Coefficients - Rank-totals for All Estimates in Each Experiment*

Structural Coefficients

Structure	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W
MOD37A2**	3334	2751	2698	2771	2078	2118	.112
MOD37A1	3538	2731	2519	2601	2258	2103	.128
MOD37A3	3074	2768	2771	2847	2103	2187	.077
MOD37A4	3111	2701	2759	2793	2223	2163	.067
MOD37A5	3315	2618	2479	2554	2443	2341	.063
Totals	16372	13569	13226	13566	11105	10912	.081

Elements of Σ

Structure	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W
MOD37A2	1631	1269	1028	896	819	657	.392
MOD37A1	1616	1273	1021	918	801	671	.377
MOD37A3	1574	1275	1059	938	788	666	.352
MOD37A4	1645	1358	1031	889	822	655	.401
MOD37A5	1595	1272	1027	846	801	759	.340
Totals	8061	6347	5166	4487	4031	3408	.370

Reduced-Form Coefficients

Structure	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
MOD37A2	5020	4251	4135	4137	3127	3211	5519	.147
MOD37A1	5230	4293	3921	3841	3683	3338	5194	.106
MOD37A3	4700	4171	4224	4279	3188	3436	5402	.107
MOD37A4	5514	4203	4007	3932	3273	3273	5198	.147
MOD37A5	5499	3960	3908	3769	3950	3395	4919	.105
Totals	25963	20778	20195	19958	17221	16653	26232	.113

Predictions

Structure	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
MOD37A2	1899	1768	1782	1825	1560	1496	2270	.067
MOD37A1	2090	1754	1703	1706	1610	1532	2205	.066
MOD37A3	1899	1771	1784	1798	1508	1587	2253	.061
MOD37A4	2129	1783	1734	1854	1440	1459	2301	.108
MOD37A5	1965	1749	1738	1659	1771	1568	2150	.041
Totals	9982	8825	8741	8742	7889	7642	11179	.063

*Entries are the totals of the ranks of all the estimates in an experiment by their deviations from the true values.

**This is the basic experiment.

TABLE V-9

Different Structural Coefficients - Rank-totals for the Ranks of the Rank-totals of the Estimates of the Individual Parameters in each Experiment

Structural Coefficients

Structure	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W
MOD37A2	86	61.5	52	65.5	22	28	.737
MOD37A1	82	63	49	58	39	24	.512
MOD37A3	78	56	58.5	70.5	21.5	30.5	.627
MOD37A4	68.5	65.5	60.5	64	32	24.5	.464
MOD35A5	77	54.5	42	51.5	46	44	.211
Totals	391.5	300.5	262	309.5	160.5	151	.438

Elements of Σ

	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W
MOD37A2	36	30	24	16	14	6	.975
MOD37A1	36	30	24	17	13	6	.984
MOD37A3	36	30	24	18	12	6	1.000
MOD37A4	36	30	24	16	14	6	.975
MOD37A5	36	30	24	14	13	9	.908
Totals	180	150	120	81	66	33	.962

Reduced-Form Coefficients

Structure	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
MOD37A2	135	89.5	70.5	77	34.5	46.5	135	.755
MOD37A1	134	97	71	60.5	63.5	40	122	.582
MOD37A3	111	70.5	86	101	34	57	128.5	.519
MOD37A4	141	94	77	70.5	47	42	116.5	.629
MOD37A5	133	80	74	59.5	81	45.5	115	.451
Totals	654	431	378.5	368.5	260	231	617	.516

Predictions

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
MOD37A2	49	37.5	37.5	40	19	10.5	58.5	.721
MOD37A1	56	39	30.5	32.5	20	15	59	.740
MOD37A3	48	36.5	38.5	41	9	18	61	.817
MOD37A4	57	38	33.5	36.5	11.5	15.5	60	.903
MOD37A5	53	32	33	25.5	35	14.5	59	.625

TABLE V-10

Different Structural Coefficients - Significant Differences Between the Experiments*

MOD37A2 - MOD37A4

Structural Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	3	3	0	3	2
2SLS	4	-	0	0	0	0
UBK	4	4	-	0	0	0
LIML	4	3	0	-	0	0
3SLS	3	1	0	1	-	0
FIML	2	1	0	0	0	-

Reduced-Form Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	5	4	3	1	1	6
2SLS	0	-	0	0	0	0	2
UBK	0	0	-	0	0	0	2
LIML	0	0	0	-	0	0	4
3SLS	0	0	2	3	-	0	3
FIML	0	0	0	0	1	-	2
LSRF	1	1	1	1	0	0	-

MOD37A2 - MOD37A5

Structural Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	4	4	3	1	0
2SLS	2	-	3	2	0	0
UBK	2	2	-	0	0	0
LIML	2	2	1	-	0	0
3SLS	3	4	4	5	-	1
FIML	3	4	3	4	0	-

Reduced-Form Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	11	10	9	5	4	10
2SLS	2	-	5	0	0	0	3
UBK	2	3	-	1	0	0	3
LIML	3	3	0	-	0	0	7
3SLS	3	7	7	9	-	2	6
FIML	2	2	2	3	0	-	2
LSRF	3	1	0	0	0	0	-

MOD37A4 - MOD37A5

Structural Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	7	7	8	7	8
2SLS	2	-	9	9	2	3
UBK	2	3	-	2	2	1
LIML	2	3	3	-	1	0
3SLS	1	1	2	4	-	8
FIML	1	1	1	0	1	-

Reduced-Form Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	17	17	17	16	17	11
2SLS	2	-	17	17	1	7	1
UBK	3	3	-	17	5	6	1
LIML	3	3	1	-	2	4	0
3SLS	2	3	4	11	-	17	3
FIML	3	1	0	1	3	-	0
LSRF	3	8	7	6	6	8	-

*Each row records the number of coefficients for which the relative performance of the method was significantly better in the first structure than in the second at the .05 level.

of the structural and reduced-form coefficients on the basis of the rankings of the estimates of the methods and the rankings of many of the descriptive statistics of their distributions. The results of pairwise comparisons between the methods would also have led to such a conclusion. The sign test and the Wilcoxon test both found 3SLS to have been significantly poorer than each of the consistent k-class methods for several structural coefficients, though in most of the comparisons with these methods the tests found 3SLS to have been significantly better more frequently than it was significantly poorer than the other methods.[82] While one could hardly conclude that 3SLS was poorer than the consistent k-class estimators in MOD37A5, the closeness of the rankings and the results of the pairwise comparisons of the methods in that experiment do indicate that the ranking of the methods into three groups is more uncertain than it would have seemed to be if the results of the experiment were more in concordance with the results of other experiments.

The standard errors of MOD37A5 exhibited some pronounced differences from those of the basic run. A larger proportion of the t-ratios fell outside each of the three intervals examined than in the basic experiment. These differences between the experiments were particularly marked for 3SLS, 2SLS and DLS. The same sort of differences from the basic experiment were found for MOD37A1, although they were less pronounced than in MOD37A5.

[82] We may remark that in the experiments using different structures, as in the basic experiment, there were significant differences in the frequencies with which a method came closer to the true values than other methods for some parameters from those for other parameters.

The numbers of t-ratios falling outside the intervals were far from being the same for all the coefficients in MOD37A5. It was the t-ratios for the coefficients of the second equation which fell outside the intervals more frequently than in the basic experiment. The number of t-ratios falling outside the intervals in the other two equations was much the same as in the basic experiment. This pattern was also found in MOD37A1.

Another peculiarity of MOD37A5 was that in four replications it produced some negative "squared standard errors". All these negative values occurred for FIML and indicate that a maximum of the likelihood function was not attained (or that rounding error had built up to such an extent that one might well doubt the meaningfulness of the estimates). Most of the negative standard errors occurred for estimates of structural coefficients in the second equation. Estimates of the structural coefficients for which negative standard errors were found were frequently very far from the population values of the coefficients. However, in a few cases, the estimates having negative standard errors were not particularly far from the true values. In addition, when negative standard errors occurred in conjunction with very poor FIML estimates, the estimates of other methods, particularly of LIML, were also apt to be very poor, although they were not as far from the true values as the FIML estimates of the coefficients, estimates occurred which were farther from the true values than those which had negative standard errors.

<u>Summary</u>. Although there was fairly high agreement among the experiments which had different values for the structural

coefficients about the over-all rankings of the methods, it was found that the exact frequencies with which one method was closer to the true values than another differed significantly among the experiments. In at least one case the differences were great enough to cause considerable doubt about the over-all ranking of the methods into three groups which was suggested by the basic experiment. The results of the experiments also cast doubt on the adequacy of the standard errors for making inferences about the values of the coefficients. No explanation for these results is suggested. They did not seem to arise from differences in the values of the determinants of the beta-matrices since some of the most pronounced differences occurred between experiments for which the values of the determinant were close together.

5. Different Structures--Different Σ-Matrices

Five experiments investigated the performances of the estimators on structures which had different variance-covariance matrices of disturbances, Σ. Three Σ-matrices were investigated in addition to the one used in the basic experiment. All three were used with the structural coefficients of MOD37A2. These structures were named MOD37AA, MOD37AB, and MOD37AC. In addition, two of the Σ-matrices were used with the structural coefficients of MOD37A4. These structures are referred to as MOD37AD and MOD37AE. The values of the Σ-matrices used are presented in Table V-11. The main differences between the matrices lie in the relative importance of the off-diagonal elements. The disturbances of

MOD37A2 were very highly correlated with each other. Those of MOD37AA were less highly correlated while those of MOD37AB had still smaller correlations. MOD37AC had disturbances whose correlations were all fairly low and were all negative.

A comparison of the results of the experiments which used different Σ-matrices suggests important differences in the relative performances of the estimators when there are different amounts of correlation between the structural disturbances. The main difference is that, with low correlations between the disturbances, the full-model methods showed a tendency not to have better performances than the consistent k-class estimators.

In estimating the structural and the reduced-form coefficients of MOD37AB and MOD37AC, both FIML and 3SLS ranked behind the consistent k-class estimators for most measures according to which they had been superior in estimating the coefficients of MOD37A2. In cases where this reversal of order did not occur, the totals of the ranks for the estimates of the full-model methods were closer to those of the consistent k-class estimators in MOD37AB and MOD37AC than they were in MOD37A2.[83] The position of DLS relative to the full-model methods and, to a lesser extent, relative to the consistent k-class estimators showed an improvement, though not enough to lead to its being ranked ahead of the full-model methods by most measures of dispersion about the true values.

[83] When a reversal of order did not occur, it was usually LIML which continued to rank worse than 3SLS. See the A-tables of Appendix D for details.

TABLE V-11

Variance-Covariance Matrices

Elements of Matrices	MOD37A2 MOD37A4	MOD37AA MOD37AD	MOD37AB MOD37AE	MOD37AC
σ_{11}	35.24	30.80	29.24	38.60
σ_{12}	34.48	21.48	3.32	- 5.92
σ_{13}	31.12	16.76	- 1.24	-14.80
σ_{22}	36.68	37.32	36.20	36.68
σ_{23}	29.84	20.12	4.96	- 2.96
σ_{33}	40.64	46.96	46.60	40.64
Raw Correlation Coefficients				
r_{12}	.96	.63	.11	-.16
r_{13}	.82	.44	-.03	-.37
r_{23}	.77	.48	.12	-.05

TABLE V-12

Different Σ Matrices - Rank-totals of all Estimates*

Structural Coefficients

Structure	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W
MOD37A2	3334	2751	2698	2771	2078	2118	.112
MOD37AA	3108	2656	2600	2645	2356	2385	.037
MOD37AB	2877	2480	2488	2565	2621	2719	.012
MOD37AC	2929	2494	2484	2505	2656	2682	.015

Elements of Σ

Structure	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W
MOD37A2	1631	1269	1028	896	819	657	.392
MOD37AA	1500	1216	1018	894	898	774	.225
MOD37AB	1320	1064	931	877	1070	1038	.075
MOD37AC	1313	1108	989	907	1039	944	.069

Reduced-Form Coefficients

Structure	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
MOD37A2	5020	4251	4135	4137	3127	3211	5519	.147
MOD37AA	4811	4118	4009	4079	3556	3507	5320	.084
MOD37AB	4409	3818	3841	3991	4065	4056	5220	.047
MOD37AC	4426	3941	3939	3959	4031	4202	4902	.025

Predictions

Structure	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
MOD37A2	1899	1768	1782	1825	1560	1496	2270	.067
MOD37AA	1813	1733	1777	1826	1589	1633	2229	.046
MOD37AB	1783	1662	1695	1792	1736	1793	2139	.026
MOD37AC	1940	1734	1744	1724	1740	1738	1980	.013

*Table entries are the sums of the ranks of the absolute deviations from the true values of all the estimates in each experiment.

TABLE V-13

Different Σ Matrices - Significant Differences between MOD37A2 and MOD37AB*

Structural Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	5	5	5	9	7
2SLS	0	-	1	0	7	6
UBK	0	0	-	1	6	6
LIML	0	0	0	-	10	9
3SLS	0	0	0	0	-	0
FIML	0	0	0	0	0	-

Elements of Σ**

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	3	3	3	3	3
2SLS	0	-	3	3	3	3
UBK	0	0	-	2	4	3
LIML	0	0	0	-	4	3
3SLS	0	0	0	0	-	0
FIML	0	0	0	0	0	-

Reduced-Form Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	1	0	1	11	10	2
2SLS	0	-	1	3	13	10	0
UBK	0	0	-	1	12	11	0
LIML	0	0	0	-	11	10	0
3SLS	0	0	0	0	-	0	0
FIML	0	1	0	0	0	-	0
LSRF	0	1	1	2	3	4	-

Predictions

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	2	2	1	1	2	1
2SLS	-	-	1	0	0	0	0
UBK	0	0	-	0	0	1	0
LIML	0	0	0	-	1	1	-
3SLS	0	0	0	0	-	0	0
FIML	0	0	0	0	0	-	0
LSRF	1	0	0	0	0	0	-

*Entries in each row record the number of parameters for which the method came closer to the true values than each other method significantly more often in MOD37A2 than in MOD37AB. The McNemar test and the .05 significance level were used.

**Most of the significant differences occurred for the off-diagonal elements of Σ.

The extent of the changes in the positions of the methods is well indicated by the differences in the totals of the ranks of all estimates presented in Table V-12. For all types of parameter estimated, the changes in the relative standings of the methods occurred. When the changes were not enough to upset the rankings of the methods, the rank-totals of the methods were nevertheless closer to each other when there were low correlations among the disturbances. Usually these changes occurred in the rank-totals for all the separate parameters estimated. It is worth noting that in MOD37AB and MOD37AC, the rank-totals of the full-model methods for the off-diagonal elements of Σ were higher than those for the diagonal elements. In the basic experiment the reverse was true.

FIML and 3SLS were not closer to the true values more often than the consistent k-class estimators for a majority of the structural or reduced-form coefficients in MOD37AB and MOD37AC. In the basic experiment they had been. In MOD37AB and MOD37AC the full-model methods were never found to have significantly smaller dispersions for their estimates of the structural coefficients about the true values than the consistent k-class methods. In the basic experiment, they had significantly smaller dispersion for about two-thirds of the coefficients. However, in MOD37AB and MOD37AC, the sign test and the Wilcoxon test found at most one structural coefficient for which FIML or 3SLS had significantly larger dispersions than any of the consistent k-class methods. The results of the pairwise comparisons of the estimators for the reduced-form coefficients were much the same as for the structural coefficients. One may conclude from the results of MOD37AB and MOD37AC that the full-model methods were not better than the consistent k-class estimators, but not that they were worse.

Table V-13 summarizes the results of the McNemar test in comparing the estimates of MOD37AB with those of the basic experiment. Significant differences between the experiments occurred frequently in the performances of the full-model methods relative to the others in estimating the structural and the reduced-form coefficients. There were few significant differences between the experiments in the performances of the consistent k-class methods relative to each other or in the comparisons of the full-model methods with each other. Few significant differences were found among the predictions. Quite a few significant differences were found between the experiments in the frequencies with which any one method came closer to the true values of Σ than any of the others. These significant differences occurred almost entirely for the off-diagonal elements of Σ and not for the diagonal elements.

The summary statistics of the distributions of the estimates of each method help to indicate what the differences between the experiments were. Most of the descriptive statistics showed the dispersions of the k-class estimators to have been smaller in MOD37AB or MOD37AC than in the basic experiment. The dispersions of the full-model methods were frequently larger. When the differences in the values of the descriptive statistics were in the same direction for both classes of methods, the changes in the descriptive statistics of the k-class estimators were larger than those of the full-model methods if the statistics were smaller in MOD37AB or MOD37AC than in the basic experiment. If the descriptive statistics had become larger, the differences were greater for the full-model methods than for the k-class methods.

The results of the experiment using MOD37AA, which had correlations between the disturbances of sizes intermediate to those of MOD37A2 and MOD37AB, were intermediate to those of the experiments using MOD37AB and the basic experiment. This was true of the rankings of the methods which became closer together than in the basic experiment though not as much so as in MOD37AB. It was also true for the changes in the descriptive statistics.

Comparisons of the results of the experiments which used MOD37AD and MOD37AE with the one using MOD37A4 tended to confirm the conclusions which could be drawn from the experiments using the structural coefficients of MOD37A2 with different Σ-matrices. Once again, the lower the correlations between the structural disturbances, the better tended to be the performances of the consistent k-class estimators relative to the full-model methods and the better the performances of DLS relative to each of the other methods. As the correlations between the disturbances decreased, there was a deterioration of the rankings of the descriptive statistics of the full-model methods. Their descriptive statistics increased in size while those of the single-equation methods either decreased or did not increase as much.

The results of this section may not be very surprising. The full-model methods can be looked at as minimizing the generalized variance of the residuals of the structural equations. The single-equation methods minimize the variances of the residuals of each equation.[84] If the off-diagonal elements of Σ are small, one might expect there to be little gain from concentrating on the generalized variance and that, indeed, one might lose from using one of the more sophisticated methods. It is interesting to find that the results of the experiments bear out this supposition. They suggest that the full-model methods have much advantage over the consistent single-equation methods only if the correlations between the disturbances are high.

The reason why DLS is not an appropriate estimator for simultaneous equations is that the disturbances in each equation are correlated with some of the "explanatory" variables. With lower correlations among the disturbances, this problem may not be as serious and the advantage to be gained from using one of the consistent methods might be expected to be less.[85] The results of the experiments suggest that this is true.

<u>Summary</u>. The performances of the estimators relative to each other were sensitive to the true values of Σ. In particular,

[84] See Chapter II, Section 2 for a discussion of the estimators and what is minimized by each. It should be noted that the single-equation methods minimize the variances of the equations expressed in different variables (or the directions of minimization are not the same for the methods if the equations are expressed in the original variables). 3SLS does not, in fact, minimize the generalized variance, but it can be expressed as an approximation to this.

[85] In MOD37AB, MOD37AA and MOD37AC, the correlations between the structural disturbances and the endogenous variables on which the equations were not normalized were in fact lower than in the basic experiment. The hypothesis that the performance of DLS relative to 2SLS will be better when Σ is diagonal than when it is not, is discussed in R. E. Quandt, <u>On Certain Small Sample Properties of k-class Estimators</u>, Princeton University, mimeographed, 1963, p. 4.

with small correlations between the structural disturbances, the differences in performance between the full-model methods and the consistent k-class estimators were negligible. The performances of DLS were not a great deal worse than those of the consistent methods when there were small correlations. This suggests that at best the rankings of the methods established in the basic experiment hold only if there are large correlations between the structural disturbances.

6. Different Structural Forms

The next set of experiments investigated the performances of the estimators when the endogenous variables were generated by structures of forms different from the one used so far in the experiments. The nine other strcutures used are shown in Table V-14. All the experiments discussed in this section used twenty observations of the variables to estimate the structures.

There were a number of differences between the structures. MODEL37B, MODEL37C, MODEL37D and MODEL37E each had the same number of equations and of exogenous variables as MOD37A2, the structure of the basic experiment. Indeed, the experiments studying these structures used the same exogenous variables and sets of disturbances as did the basic experiment. The forms of these structures differed from that of the basic experiment in the number of identifying restrictions placed on the equations or in the number of zeroes in the B-matrix.[86] The other structure had different numbers of equations or exogenous variables from the basic structure. Each of their equations had from one to four over-identifying restrictions. The data used with these structures were not the same as those of the basic experiment.[87]

The results of the experiments which used different structures were in quite close agreement over the rankings of the methods. The rank-totals for the descriptive statistics in each experiment were themselves ranked. The concordances of these ranks in the experiments were high. Kendall's W was higher than .8 for all measures of dispersion of the estimates of the structural coefficients except for those which concentrated

---

[86] The results of the experiments which used different values for the coefficients of the basic structure were not available when the experiments using MODEL37B, MODEL37C, MODEL37D and MODEL37E were conducted. It had been supposed that the use of different values for the coefficients of a particular structure would not lead to great differences in the results but that having different structural forms might substantially affect the results. It was further supposed that such differences in the results might be ascribed to greater sparseness of the B-matrix or to having different numbers of over-identifying restrictions on the equations. The relations between the structures and data of MODEL37B-MODEL37E and the basic experiment were designed to provide experimental controls to facilitate exploring these suppostions. When it turned out that using different values for the structural coefficients did substantially affect the results, it was decided not to explore these suppositions in detail. The program which was to take advantage of the close relationships between the structures was never completed since this did not seem a very profitable use of computer time. However, since these experiments were completed, their results are presented here at least in a superficial manner.

[87] In analyzing the estimates of the reduced-forms of MODEL48, MODEL56 and MODEL58, only the first twenty-four coefficients were studied. Thus only the estimates of the first three reduced-form equations of MODEL48 and MODEL58 and those of the first four equations of MODEL56 were considered. These limitations were imposed to simplify programming and to save computer time. It was not felt that ways of analyzing the results to allow the study of all the reduced-form coefficients would be justified since the extra information was not felt likely to compensate the very considerable amount of computer time needed to attain the extra results.

on the more extreme quantiles of the distributions of the estimates. Many of the same levels of concordance occurred for the measures of performance in estimating the reduced-form coefficients. The levels of Kendall's W for the predictions tended to be somewhat lower, but values of W lower than .6 were not observed and for most measures W was above .7. High concordances between the experiments occurred for the measures of performance in estimating the elements of Σ which did not give weight to the more extreme quantiles of the observed distributions. For all these aspects, reversals of order in the rankings of the methods in the separate experiments from the over-all ranking for the set of experiments occurred mainly between methods whose performances were very much alike in the basic experiment.

Two features of the estimates of the different structural forms were as noticeable as the high concordances among the experiments for the ranks of the rank-totals for the descriptive measures in each experiment. The first is that in all the experiments, the concordances among the ranks of the deviations of the estimates from the true values of any one structural or reduced-form coefficient in the fifty replications of the experiments were apt to be very low. This was particularly evident when MODEL48 and MODEL58 were used. In those experiments Kendall's W was not significantly different from zero at the .01 level for one-half of the struqtural coefficients. In other experiments, although the values of W were more frequently significantly different from zero, they were apt to be very low.

TABLE V-14

Different Structural Forms

MODEL 5A

$$B = \begin{bmatrix} 1 & -.2 & -.3 \\ -.3 & 1 & 0 \\ 0 & -.2 & 1 \end{bmatrix} \qquad -\Gamma = \begin{bmatrix} 16 & .4 & 0 & .2 & 0 & 0 \\ 32 & 0 & .9 & 0 & .1 & 0 \\ 48 & 0 & .3 & .6 & 0 & .5 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 16.72 & 11.60 & 11.40 \\ & 9.36 & 10.76 \\ & & 15.95 \end{bmatrix}$$

MODEL 37B

$$B = \begin{bmatrix} 1 & 0 & -.70 \\ -.58 & 1 & 0 \\ 0 & -.83 & 1 \end{bmatrix} \qquad -\Gamma = \begin{bmatrix} 71 & .56 & .79 & 0 & .42 & 0 & 0 \\ 51 & 0 & .30 & 0 & .38 & 0 & .19 \\ 20 & 0 & .52 & .81 & 0 & .7 & 0 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 35.24 & 34.48 & 31.12 \\ & 36.68 & 29.84 \\ & & 40.64 \end{bmatrix}$$

TABLE V-14 (Cont.)

MODEL 37C

$$B = \begin{bmatrix} 1 & -.21 & -.27 \\ -.97 & 1 & -.52 \\ -.45 & -.11 & 1 \end{bmatrix} \qquad -\Gamma = \begin{bmatrix} 39 & .89 & 0 & 0 & .43 & 0 & 0 \\ 85 & 0 & .35 & 0 & 0 & 0 & .17 \\ 68 & 0 & 0 & .82 & 0 & .71 & 0 \end{bmatrix}$$

Σ - as in MODEL 37B

MODEL 37D

$$B = \begin{bmatrix} 1 & 0 & -.7 \\ -.58 & 1 & 0 \\ 0 & -.83 & 1 \end{bmatrix} \qquad -\Gamma = \begin{bmatrix} 39 & .89 & 0 & 0 & .43 & 0 & 0 \\ 85 & 0 & .35 & 0 & 0 & 0 & .17 \\ 68 & 0 & 0 & .82 & 0 & .71 & 0 \end{bmatrix}$$

Σ - as in MODEL 37B

MODEL 37E

$$B = \begin{bmatrix} 1 & 0 & -.7 \\ -.58 & 1 & 0 \\ 0 & -.83 & 1 \end{bmatrix} \qquad -\Gamma = \begin{bmatrix} 39 & .89 & 0 & 0 & .43 & 0 & 0 \\ 85 & 0 & .35 & 0 & 0 & 0 & .7 \\ 68 & 0 & 0 & .82 & 0 & .71 & 0 \end{bmatrix}$$

Σ - as in MODEL 37B

TABLE V-14 (Cont.)

MODEL 46

$$B = \begin{bmatrix} 1 & 0 & -.82 & 0 \\ -.37 & 1 & 0 & -.14 \\ -.4 & -.37 & 1 & -.13 \\ 0 & -.51 & 0 & 1 \end{bmatrix} \qquad -\Gamma = \begin{bmatrix} 54 & 0 & 0 & .35 & 0 & .19 \\ 76 & 0 & .98 & 0 & 0 & 0 \\ 88 & 0 & 0 & 0 & 0 & .65 \\ 20 & .27 & 0 & 0 & .43 & 0 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 31.56 & 23.60 & 28.48 & -22.76 \\ & 31.56 & 18.56 & -2.2 \\ & & 39.2 & -26.48 \\ & & & 41.56 \end{bmatrix}$$

MODEL 48

$$B = \begin{bmatrix} 1 & -.64 & -.14 & 0 \\ -.17 & 1 & -.22 & -.68 \\ -.25 & -.76 & 1 & -.85 \\ -.32 & -.09 & -.58 & 1 \end{bmatrix} \qquad -\Gamma = \begin{bmatrix} 65 & .84 & 0 & 0 & 0 & .3 & 0 & 0 \\ 58 & 0 & .27 & 0 & 0 & 0 & .52 & 0 \\ 68 & 0 & 0 & .9 & 0 & 0 & 0 & .25 \\ 42 & 0 & 0 & 0 & .6 & 0 & 0 & 0 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 38.04 & 24.28 & 32.56 & -26.72 \\ & 28.40 & 17.88 & -5.04 \\ & & 42.92 & -31.84 \\ & & & 40.64 \end{bmatrix}$$

TABLE V-14 (Cont.)

MODEL 56

$$B = \begin{bmatrix} 1 & 0 & 0 & -.47 & -.7 \\ 0 & 1 & -.32 & 0 & 0 \\ -.56 & 0 & 1 & -.69 & 0 \\ 0 & -.48 & 0 & 1 & 0 \\ 0 & 0 & -.35 & 0 & 1 \end{bmatrix} \qquad -\Gamma = \begin{bmatrix} 18 & .54 & 0 & 0 & 0 & 0 \\ 6 & 0 & .63 & 0 & 0 & .34 \\ 93 & 0 & 0 & .77 & 0 & 0 \\ 72 & 0 & 0 & 0 & .6 & .19 \\ 41 & 0 & 0 & 0 & .13 & .23 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 29.88 & 17.92 & 21.08 & -21.32 & 6.92 \\ & 30.72 & 19.56 & 1.64 & 4.20 \\ & & 39.04 & -19.32 & 25.36 \\ & & & 40.36 & -14.12 \\ & & & & 39.52 \end{bmatrix}$$

MODEL 58

$$B = \begin{bmatrix} 1 & .78 & 0 & -.84 & .33 \\ -.09 & 1 & .52 & 0 & 0 \\ -.54 & 0 & 1 & -.57 & 0 \\ -.82 & -.37 & 0 & 1 & 0 \\ 0 & 0 & -.77 & 0 & 1 \end{bmatrix} \qquad -\Gamma = \begin{bmatrix} 46 & .47 & 0 & 0 & 0 & 0 & 0 & .6 \\ 72 & 0 & .16 & 0 & 0 & .76 & 0 & .25 \\ 34 & 0 & 0 & .47 & 0 & 0 & .15 & 0 \\ 84 & 0 & 0 & 0 & .22 & 0 & .59 & 0 \\ 38 & 0 & 0 & 0 & 0 & .95 & 0 & .39 \end{bmatrix}$$

Σ - As in MODEL 56

The second noticeable feature was that the ranks of the rank-totals of the estimates of the separate coefficients sometimes showed fairly low concordances for the set of coefficients estimated in an experiment. For the structural coefficients, Kendall's W varied from a high of .737 in the basic experiment to a low of .198 in the experiment using MODEL37E. In estimating the reduced-form coefficients and the elements of Σ and in making predictions similar variations in the concordances of the rankings of the methods were observed. Considerable variations in W between the experiments for the rankings of the descriptive statistics of dispersion also occurred.

The lack of high concordances in any experiment for the estimates of the separate parameters and the low levels observed in some experiments for the ranks of the rank-totals of the separate parameters show well how ambiguous are the results of this study. When different rankings of the methods occurred for the rank-totals of the separate quantities estimated, the number of times which one method surpassed another was frequently not significantly different from one-half of the estimates for many of the parameters involved in the reversals of order. However, there were also several cases where, for some parameters in an experiment, one method would be significantly better than another for some parameters, while for others it would be significantly poorer. The frequencies with which one method was better than another varied significantly among the different coefficients in the experiments. All these findings render an over-all ranking of the methods precarious.

Experiments with sparser B-matrices (MOD37A2, MODEL5A, MODEL37B, MODEL37D, MODEL46 and MODEL56) seemed to have somewhat higher concordances among the ranks of the rank-totals of the separate structural coefficients than those with less sparse B-matrices. However, there was not a systematic tendency for concordance to increase with sparseness and structures with higher sparseness than others sometimes had lower concordances. Other systematic relationships between features of the structural forms, such as the number of over-identifying restrictions, and the levels of concordance were not apparent in the results. Nor did there seem to be many differences in the relative performances of the methods in estimating the elements of B from those of Γ or in estimating one equation and in estimating another. In view of the low values of W for the estimates of any one parameter which were usually found and the significant differences in relative performances found earlier in the experiments using the same structural forms, it is doubtful if any importance should be given to the tendency of the experiments with sparse B-matrices to show higher concordances in the ranks of the rank-totals of the separate structural coefficients than the experiments with less sparse B-matrices.

The performances of the standard errors in the experiments using different structures were similar to those in the basic experiment. In some of the experiments, the confidence limits based on the t-distribution with degrees of freedom equal to the number of observations minus the total number of coefficients to be estimated are meaningless, since there were more coefficients than observations, In these cases, as in the other experiments, the hypothesis that confidence intervals for the *adjusted* t-ratios should be based on the t-distribution with degrees of freedom equal to the number of observations minus the number of coefficients in the equation seemed fairly adequate and could be rejected rarely. The hypothesis that this distribution was appropriate for the unadjusted t-ratios did not fare as well and could quite often be rejected. For DLS none of the hypotheses seemed to fit and all of them could be rejected very often. The standard errors of DLS did, however, seem to give fairly reliable indications of the dispersions of the estimates about their (biased) central tendencies.

Summary. The over-all rankings of the methods in the experiments using other structural forms were quite similar to the rankings established in the basic experiment. The agreement in any one experiment over the ranking among the different parameters was sometimes very low and made the rankings of the methods into groups seem rather precarious.

7. Different Disturbance Sizes

Program-parameter 13 was a scale-factor by which the basic set of structural disturbances in each experiment was multiplied before being used to generate the endogenous variables. In most experiments, program-parameter 13 had a value of 2. In the experiments reported in this section, other values of the scale-factor were used. First, experiments were conducted using MOD37A2, the structure of the basic experiment, with program-parameter 13 set equal to 1, 4, 6, 8 and 10. Then three experiments were conducted using other structures with program-parameter 13 set equal to 8.

The effect of varying the scale-factor was, in the first instance, to change the size of the effect which the latent, stochastic parts of the structure played in the generation of the endogenous variables. Table V-15 shows, in approximate fashion, the percentages of the variation of the endogenous variables accounted for by the disturbances.[88]

One might expect that increasing the size of the disturbances would lead to an increase in the dispersions of the estimates of each of the methods. By and large, the results of the experiments tend to bear this out. In most cases, as the value of the scale-factor increased, so did the descriptive statistics of dispersion for any one of the methods. The increases seemed to be roughly proportional to the increases in the size of the disturbances. However, some exceptions to this pattern occurred. For the structural coefficients there were several descriptive statistics of dispersion which were smaller for the experiment with the scale-factor set at 10 than in the one where it equalled 8. This occurred especially often for the 80% ranges and the third-quartile and ninth-decile absolute deviations of the FIML estimates. For the reduced-form coefficients and for the predictions, instances of descriptive statistics for the experiment with the scale-factor set at 8 being larger than for the one with it set at 10 occurred quite frequently for 3SLS, LIML, FIML and UBK.

An explanation of the seemingly peculiar result that the dispersions of the estimates of some methods seemed to decrease when the size of

[88] The proportionate roles of the disturbances are shown as ranges since the effect of the disturbances was not the same for each variable. The effects shown are for the reduced-form disturbances.

the disturbances increased is to be found in the difficulties encountered in computing the estimates. A number of sets of data could not be used when the scale-factor equalled 10 because of the occurrence of (apparently) singular matrices in computing the FIML estimates. For many of the corresponding sets of data, estimates were obtained when the scale factor was set at 8, but these were often very poor estimates. Furthermore, when a set of data which produced estimates falling in the upper quartile of the distributions of the absolute deviations of the estimates of the experiment with the program-parameter set at 8 could be used when it equalled 10, the resulting estimates in the latter experiment usually fell at higher quantiles of the distributions of absolute deviations than in the former experiment. The abandoning of more of the "difficult" sets of data in one experiment than in the other may thus account for the dispersions in the former experiment sometimes seeming to be smaller than in the latter.

With high values for the scale-factor, a large number of very bad estimates of the structural coefficients and the reduced-form coefficients occurred. While LIML produced some very poor estimates, the method which produced the worst estimates of the structural coefficients was FIML. Some FIML estimates deviated by several hundreds percent from the true values of the structural coefficients. Furthermore, in the experiments with the scale-factor set at 8 or 10, negative "squared standard errors" occurred with some frequency for the FIML estimates. Since FIML and, to a lesser extent, LIML are methods which may pose computational problems and negative "squared standard errors" are _prima facie_ evidence that the FIML estimates have not been computed properly and are not the "true" FIML estimates,[89] one may wonder whether these poor estimates may not just have been the result of rounding error building up to the point that the estimates bore no relation to the data.

Three considerations argue against readily embracing the hypothesis that rounding error or other computational hazards account for the very poor estimates. First, although negative "squared standard errors" occurred for some of the poor FIML estimates, not all the poor estimates had them. Indeed, most of the worst estimates of the coefficients had standard errors of positive sign. At the same time, some of the negative "squared standard errors" occurred for fairly good estimates. Second, when bad FIML estimates occurred in a replication, they were apt to occur for the estimates of only one of the structural equations and not for the other two, even though all FIML estimates are computed simultaneously.[90] Third, and more important, although FIML and LIML often produced very poor estimates of the structural coefficients, their estimates of the reduced-form coefficients and their predictions--both of which are derived from the estimates of the structural coefficients--were not particularly poor. The dispersions of 3SLS for these two aspects were larger than those of FIML. The dispersions of UBK similarly were larger than those of LIML.

[89] See Appendix A for a discussion of the way in which estimates were made and the criterion for accepting estimates as the FIML estimates, even though the matrix of second derivatives might not indicate that a true maximum had been attained.

[90] The second equation was the one which was estimated poorly.

Even for the replications for which LIML or FIML produced very poor estimates, their predictions were often quite good, both in comparison with predictions made by other methods and in comparison with predictions made by these methods in other replications. If computing difficulties had led to estimates of the structural coefficients which were quite unrelated to the sample from which they were estimated, it seems most unlikely that good estimates of the reduced-form coefficients or good predictions would still be made.

Changing the sizes of the disturbances seemed to affect the biases of several of the estimating methods. Not only did the sizes of the biases of the means and the medians of the estimates of the structural and the reduced-form coefficients tend to increase with increases in the sizes of the disturbances, the number of coefficients for which significant median biases occurred also tended to increase for 2SLS, UBK and 3SLS. When the scale-factor was set at 10, these methods had medians which were significantly different from the true values for at least two-thirds of the structural coefficients. For the reduced-form coefficients, 2SLS and 3SLS had significant median biases for ten of the twenty-one coefficients while UBK was significantly biased for six. LIML and FIML, however, had significant biases for at most three structural coefficients and two reduced-form coefficients. Increases in the number of significant median biases did not seem to occur for the predictions of any of the methods as the size of the disturbances increased.

The dispersions of the estimates of the elements of Σ about their true values and about their central tendencies increased in size with the increase in the sizes of the population values of Σ. However, the dispersions as a percentage of the population values showed a uniform increase only for FIML and LIML. For other methods, while there was a tendency for the descriptive statistics to increase, there were enough exceptions to make generalization precarious.

The effects of increasing the value of the scale-factor on the performances of the standard errors were quite pronounced. As the size of the disturbances increased, the number of t-ratios which fell outside each of the three intervals studied also increased. This was true for all methods except LIML. For none of these methods did the standard errors seem to give reliable indications of the extent of the dispersions of the estimates about the true values when the scale-factor was large. Hypotheses which fared well in the basic experiment could be rejected for many of the coefficients. The performances of the standard errors of LIML were a bit more complicated. At first, as the value of the scale-factor was increased, the number of times the t-ratios fell outside the intervals decreased slightly. Then, as the disturbances were made still larger, the numbers increased. In most experiments, the hypotheses that confidence intervals for the t-ratios for LIML should be based on the t-distribution with five degrees of freedom or that the adjusted t-ratios would follow it with fifteen degrees of freedom could seldom be rejected.

The performances of the methods relative to each other did not change in an entirely straightforward way as the sizes of the disturbances increased. In many of the comparisons between pairs of methods, the number of times one method came closer than another to the true values of the structural coefficients would first increase as the size of the disturbances was increased. Then as the size of the disturbances was further increased, the relative performance of the method would start to deteriorate. Usually the turning-point for this behavior was the experiment with the scale-factor set equal to 4. This pattern of change in the relative performances affected all the comparisons of the methods except those of 2SLS with FIML, of UBK with 3SLS and FIML, and of LIML with 3SLS. In these four cases, the relative performances of the k-class estimators improved with all increases in the size of the disturbances. The rank-totals of the deviations of the estimates from the true values, shown in Table V-16, reflect what occurred.[91] The poor positions of FIML for estimating the structural coefficients when the scale-factor was set at 8 or 10 are particularly noticeable.

A comparison of the dispersions of the estimates of the structural coefficients about their central tendencies and of the biases of the estimates may indicate why there were not uniform effects on the pairwise comparisons of the deviations of the methods from the true values from increasing the size of the disturbances. The dispersions of FIML and LIML tended to increase more than those of DLS, 2SLS and 3SLS. DLS, particularly, showed fairly small increases in the dispersions of its

[91] It may be remarked that originally it was planned to conduct only the experiments with the scale-factor set at 4 and 8. The puzzling nature of the changes in relative standing led to our conducting the other experiments. The puzzling changes in the rankings did not occur for the rankings of the descriptive statistics of dispersion about the true values except for the median absolute deviations.

TABLE V-15

Effects of Program-Parameter 13 on the Percentage of Variances of Endogenous Variables Contributed by the Disturbances

Structure	Program-Parameter 13	Variances of Disturbances as Percent of Variances of Endogenous Variables
MOD37A2	1	2.2 - 3.3
MOD37A2	2	8.3 - 12.0
MOD37A2	4	26.6 - 35.4
MOD37A2	6	44.9 - 55.2
MOD37A2	8	59.2 - 68.6
MOD37A2	10	69.4 - 77.4
MOD37AB	2	4.6 - 5.9
MOD37AB	8	43.6 - 50.0
MODEL37D	2	20.4 - 25.9
MODEL37D	8	80.4 - 84.8
MODEL 46	2	9.0 - 18.3
MODEL 46	8	61.3 - 78.2
MOD37A1	2	5.6 - 12.2
MOD37A3	2	6.6 - 10.3
MOD37A4	2	12.0 - 24.1
MOD37A5	2	12.3 - 15.0

TABLE V-16

Different Disturbance Sizes - Totals of the Ranks of all Estimates by Absolute Deviations from the True Values

Structural Coefficients

Scale-Factor	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W
1	3177	2787	2757	2800	2092	2137	.092
2	3334	2751	2698	2771	2098	2118	.112
4	3604	2793	2572	2662	2087	2086	.158
6	3393	2679	2572	2748	2123	2229	.103
8	3168	2561	2490	2659	2146	2788	.058
10	3056	2528	2487	2708	2214	2757	.041

Elements of Σ

Scale-Factor	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W
1	1603	1333	1074	935	757	598	.438
2	1631	1269	1028	896	819	657	.392
4	1649	1215	1002	893	833	708	.366
6	1589	1201	957	868	885	800	.282
8	1534	1188	885	806	870	1017	.282
10	1526	1164	876	801	916	1017	.223

Reduced-Form Coefficients

Scale-Factor	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
1	4650	4302	4236	4191	3145	3288	5586	.132
2	5020	4251	4135	4137	3127	3211	5519	.147
4	5698	4064	3956	3936	3249	3168	5329	.183
6	5987	4055	3767	3777	3593	3143	5078	.189
8	5939	4141	3895	3683	3727	3279	4736	.154
10	5992	4250	3896	3698	3803	3190	4571	.158

Predictions

Scale-Factor	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
1	1846	1811	1828	1846	1511	1477	2281	.075
2	1899	1768	1782	1825	1560	1496	2270	.067
4	2054	1725	1770	1776	1473	1511	2291	.089
6	2106	1681	1743	1753	1566	1562	2189	.066
8	2078	1790	1846	1703	1535	1574	2974	.050
10	2082	1758	1879	1724	1574	1552	2031	.046

TABLE V-17

Different Disturbance Sizes - Significant Differences between Experiments*

Structural Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	3	3	0	3	0
2SLS	2	-	1	0	1	0
UBK	3	2	-	0	0	0
LIML	2	3	0	-	0	0
3SLS	3	2	1	1	-	0
FIML	6	6	6	9	1	-

Elements of Σ

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	0	0	0	0	0
2SLS	0	-	0	0	0	0
UBK	1	0	-	0	0	0
LIML	2	0	0	-	0	0
3SLS	0	0	2	3	-	0
FIML	5	6	6	6	5	-

Reduced-Form Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	10	7	6	7	7	18
2SLS	0	-	1	1	0	0	4
UBK	0	1	-	1	0	0	4
LIML	0	0	0	-	0	0	1
3SLS	2	4	7	10	-	0	5
FIML	0	0	2	4	0	-	7
LSRF	0	0	0	0	0	0	-

Predictions

	DLS	2SLS	UBK	LIML	3SLS	FIML	LFRF
DLS	-	0	0	1	0	0	4
2SLS	0	-	0	0	1	0	0
UBK	0	0	-	2	0	0	1
LIML	0	0	0	-	0	0	0
3SLS	0	0	0	0	-	0	0
FIML	0	0	0	1	1	-	2
LSRF	0	0	0	0	0	0	-

*Each row records the number of parameters estimated for which one method was closer to the true values than another significantly more often in the basic experiment than in the one with the scale-factor equal to 8. The McNemar test was used with the .05 significance level.

estimates about their central tendencies. The biases of DLS, 2SLS and 3SLS showed a tendency to increase much more than those of LIML and FIML. These different ways in which the methods were affected by increases in the size of the disturbances may account for the differences in relative performance. The increases in the biases at first made some methods become relatively poorer, but later the smaller increases in dispersion made them relatively better.

Changes in the relative performances of the methods in estimating the reduced-forms were of a more straightforward pattern than the changes in estimating the structural coefficients. Only in the comparisons of 2SLS and FIML did one method (2SLS) at first become relatively better and then relatively worse as the size of the disturbances was increased. Table V-16 shows the rank-totals for the estimates in the different experiments. The number of times LSRF came closer to the true values than each of the other methods increased with the size of the disturbances. DLS and 3SLS performed relatively more poorly as the scale-factor became larger.

No clear-cut patterns of changes in relative performance of the methods in making predictions emerged from examining the results of the experiments having different sizes of disturbances. Such changes as occurred were small. There did, however, seem to be a slight tendency for the relative performances of LSRF to improve and for those of DLS and FIML to deteriorate. For the estimates of the elements of Σ, the position of FIML relative to the other methods tended to become steadily worse with increases in the sizes of the disturbances. The position of

3SLS relative to the single-equation methods deteriorated but to a less pronounced extent. Other differences in relative performance were slight.

The McNemar test was used to compare the results of the experiments with the scale-factor set at 2, 4 and 8. Table V-17 summarizes the findings for the comparison of the first and last of these experiments. For the structural coefficients, the changes in the positions of FIML, DLS and 2SLS were most frequently significant, but the changes in relative standing were seldom significant for half the coefficients. For the reduced-form coefficients, changes occurred in the standings of DLS, LSRF and 3SLS. For the estimates of Σ the significant changes occurred mainly in the standings of FIML. Not many significant differences were found for the predictions.

In comparing the experiment with the scale-factor set at 2 with the one with it set at 4, significant differences occurred mainly in the relative standings of DLS for the structural and the reduced-form coefficients. There were few changes which were significant for other methods or for other parameters. In comparing the experiment with the scale-factor set at 4 with the one with it set at 8, the pattern of the occurrence of significant differences was much the same as that shown in Table V-17 except that there were fewer significant changes in the performances of DLS relative to the other methods.

The comparisons of the experiments which used other structures with program-parameter 13 set at 8 with the ones using the same structures with the scale-factor equal to 2 showed much the same sorts of differences between the experiments as were found in the experiments using the structure of the basic experiment. In all cases the sizes of the dispersions of the estimates increased with the size of the disturbances. The frequency with which the biases of the medians were found to be significant for methods other than FIML and LIML also increased. In the run using MOD37AB, however, this tendency was not very pronounced. Both FIML and LIML showed significant median biases in one of the other experiments quite frequently. The standard errors, with one exception, became worse as the disturbance size increased. The exception was the standard errors of LIML in MOD37AB which became slightly more reliable with larger disturbances. MOD37AB was a structure with very low correlations between the disturbances.

One might wonder whether the differences in the relative importance of the disturbances in the experiments using different values for the structural coefficients from the basic experiment might explain the differences in the results of those experiments. This seems unlikely for several reasons. The variance of the disturbances relative to the variance of the endogenous variables in these experiments did not vary greatly. (Cf. Table V-15.) The differences between the experiments were larger and more widespread than the differences between the experiments using MOD37A2 with program-parameter 13 set equal to 1 and 4. There was a wider range of variation in the importance of the disturbances between these two experiments than was found between the experiments which used different values for the structural coefficients. Furthermore, significant differences were found most frequently between MOD37A4 and MOD37A5, structures whose relative disturbance-sizes were not very different. Finally, the significant differences found by the McNemar test in the comparisons of the experiments using MOD37A4 and MOD37A5 with the one using MOD37A2 with program-parameter 15 set at 5 were of the same type and of much the same frequency as had been found in comparing these two experiments with the basic experiment.

<u>Summary</u>. Changing the sizes of the disturbances not only changed the dispersions of the methods; it affected significantly their performances relative to each other. The changes in relative standing were not entirely of the same sort for the structural coefficients and the reduced-form coefficients. The medians of several methods were more frequently biased significantly when the disturbances were large. The standard errors were less adequate when there were large disturbances.

8. Conclusion

The results of the experiments reported in this chapter gave some support to the findings of the basic experiment. However, there were several results which were different from those of the basic experiment. The more important conclusions suggested by the experiments are listed below.

1. The differences in the performances of the methods are not very pronounced. In all experiments, the rank-totals of the estimates were fairly close together and the values of Kendall's W were low.

2. Although the ranking of the methods into three groups was supported by many of the experiments, there were cases where the performances of at least one of the full-model methods could hardly be judged superior to, or even different from, those of the consistent k-class estimators. These cases were found when different values of the structural coefficients from those of the basic experiment were used, when there were low correlations between the structural disturbances and when the sizes of the structural disturbances were large. In some experiments using different structures, the ranking was very weak. The inferiority of DLS was often not pronounced.

3. The frequencies with which one method came closer to the true values of the parameters than another were found to depend significantly on the exact set of exogenous data used, the true values of the structural coefficients, the correlations between the disturbances and the sizes of the disturbances. As in the basic experiment, it appeared that the frequency with which one method surpassed another varied significantly among the parameters in any one experiment. One cannot make general statements about the degree of superiority of one method over another on the basis of the results.

4. In some cases the estimates of some methods were found frequently to be biased significantly. This was true when the disturbances were large.

5. In some experiments, the standard errors were not reliable for making inferences about the dispersions of the estimates of the consistent methods.

These conclusions leave the question of the choice of an estimator for econometric models up in the air. In the next three chapters we will investigate whether having special features of the data or the specification of the model would suggest more clear-cut distinctions between the performances of the estimators.

CHAPTER VI

Special Cases of the Standard Model

1. Introduction

Chapter V was largely concerned with differences in the structures to be estimated, which would only be recognizable if the population values of the parameters were known. The true values of the structural coefficients, the sizes of the disturbances and the correlations between the disturbances are unknown for actual econometric models. The different sets of exogenous data used in Chapter V were all from the same population. One could not say that the data of an econometric model resembled one set more than another. While differences in the structural forms could be recognized without knowledge of the values of the parameters, it seems unlikely that most actual econometric models would resemble one of them more closely than another. The present chapter investigates cases of the standard model whose differences from the basic experiment can be recognized without knowing the true values of the parameters of the structure. The subjects studied are the effects of having more observations from which to make estimates, multicollinearity of the exogenous variables, the use of lagged-endogenous variables as predetermined variables, and trends or auto-correlation in the exogenous variables.

Chapter V indicated that the relative performances of the methods are sensitive to the exact sets of data used and the structures which generated the endogenous variables. One might wonder whether this discovery precludes other experiments from leading to useful results since any study of other aspects must entail some changes in the data used. Ascribing differences in the results from those discussed in Chapter V to the special features studied might seem a very dubious undertaking.

The search for significant differences is not the only reason for conducting experiments. Although in Chapter V the performances of the estimators turned out not to differ greatly from each other, this might not be true in other cases. Whether one or more methods are clearly superior to others when different ways of generating the data are used is an interesting question. A second reason for conducting the experiments is to inquire if the weak ranking of the methods found in most experiments of Chapter V is upset, confirmed, or even strengthened by the use of other types of data. Finally, another problem worth investigating is the possibility that other ways of generating the data would strongly affect the absolute performances of the methods, even if the relative performances were not clearly affected.

2. Different Numbers of Observations

The effects of having different numbers of observations were investigated by estimating the basic structure[92] from thirty-five, fifty

[92] This is the structure of the basic experiment, MOD37A2, shown in Table IV-1.

and seventy observations. Then three other structures were investigated using seventy observations in each replication of the experiments.[93]

The increase in the number of observations led to only one significant change in the relative standings of the methods in estimating the structural and reduced-form coefficients and in making predictions. DLS came closer to the true values than each of the other methods less often when seventy observations were used than when there were only twenty observations. These differences between the experiments were statistically significant beyond the .05 level for about half the coefficients and predictions.[94]

There was close agreement among the experiments about the rankings of the estimators.[95] Reversals of order among the consistent k-class estimators and among the full-model methods accounted for almost all the differences from unity in the values of Kendall's W.[96]

---

[93]Program-parameter 2 specified the number of observations. The experiments reported in this section are experiments AF through AK of the first set of experiments, shown in Appendix C, Table 1. The observations used in these experiments were independent from those used in experiments with only twenty observations.

[94]The relative performances of DLS did not necessarily become increasingly worse as the number of observations increased. In estimating the reduced-form and in making perdictions, the relative standings of DLS were often worse in the experiment using thirty-five observations than in the one with fifty observations. Differences between these two experiments were not usually significant at the .05 level.

[95]This is hardly surprising in view of the lack of significant differences in the pairwise comparisons of the methods between the experiments.

[96]This was true for the sums over the experiments of the ranks of the totals over the parameters of the ranks of the descriptive statistics. It also held for the sums over the experiments of the ranks of the sums over all replications and parameters of the ranks of the deviations of the estimates from the true values.

The dispersions of the estimates of the consistent techniques decreased as the number of observations increased. DLS frequently showed the same sorts of improvements as other methods when the number of observations increased, but this did not always apply to the statistics of dispersion about the true values. With more observations, one might expect to have more accurate estimates. This happened for the consistent methods.

The estimates of the elements of Σ showed some interesting changes when the number of observations was increased. With more observations, the consistent methods performed more similarly to each other in the sense that the frequency with which one estimator came closer to the true values than another became more nearly equal to one-half. These changes were significant at the .05 level for most of the individual elements of Σ in the comparisons of the experiments with twenty and seventy observations. Significant differences in the frequencies with which other methods surpassed DLS in the two experiments occurred for only one or two of the elements of Σ. These differences did not upset the ranking of the methods for estimating Σ which was established in the basic experiment, but did make it less clear-cut. The dispersions of the estimates tended to decrease, but not uniformly, as the number of observations increased. Some methods had some descriptive statistics which were smaller in the experiment with thirty-five observations than in the one with fifty observations.

The standard errors showed some changes as the number of observations increased. The standard deviations and the root-mean-square errors of the consistent estimates of the structural coefficients tended to fall at lower quantiles of the distributions of the standard errors. The places of the DLS standard deviations, but not of its root-mean-square errors, tended to decrease as the number of observations increased.

The deviations of the estimates of the structural coefficients from their true values exceeded twice their standard errors less often as the number of observations increased for all the consistent techniques. Fewer t-ratios of the consistent methods exceeded the 95% confidence limits of the t-distribution with degrees of freedom equal to the number of observations minus the number of coefficients in each equation. For neither measure could one reject the hypothesis that only 5% of the estimates would fall outside the specified ranges for any of the consistent estimators in the experiment using seventy observations.[97] The DLS standard errors performed differently from those of other methods: for DLS the frequencies with which the t-ratios fell outside the intervals examined tended to increase with the number of observations.

The effects of using seventy observations in other structures were similar to those found in the basic structure. The typical descriptive statistics of dispersion were smaller in the experiments with seventy

---

[97]For this experiment, the t-value based on the number of coefficients in each equation was less than two. The differences between the t-value based on the total number of coefficients--which had led to reliable inferences in all these experiments--and the t-value based on the number of coefficients in each equation becomes smaller as the number of observations increases. The hypothesis, which did well in the basic experiment, that the <u>adjusted</u> t-ratios follow the t-distribution with degrees of freedom equal to the number of observations minus the number of coefficients in each equation continued to do well in these experiments.

observations than in the ones with twenty and the standard errors were more reliable. The performances of the consistent methods in estimating Σ were closer to each other when seventy observations were used. The performances of DLS with respect to the structural and reduced-form coefficidnts and the predictions, relative to other methods, were significantly worse in the experiments using seventy observations than in the ones using twenty. The relative performances of the other methods did not show significant differences between the experiments.[98]

Summary. Having more observations improved the estimates of all the consistent methods but did not alter the relative performances of the methods. DLS tended to be poorer relative to the other methods when there were more observations. With more observations, DLS had smaller dispersions about its central tendencies which continued to be biased significantly.

3. Multicollinearity

The next subject investigated was the effect of having deliberately introduced correlations between the exogenous variables. Collinearity was

[98]In the structure with low correlations between the structural disturbances (MOD37AB), the relative performances of the full-model methods, especially 3SLS, were better in the experiment with seventy observations than in the one with twenty. These differences, however, were not significant at the .05 level. It is worth noting that the McNemar test found fewer significant differences in the performances of the full-model methods relative to the k-class estimators for the estimates of the structural coefficients in the paired replications of the experiments using MOD37AB and MOD37A2 with seventy observations in each than it had found when comparing the estimates of these structures made from twenty observations.

introduced by multiplying the exogenous data used in the experiments already reported by various matrices so that the exogenous variables were linear combinations of those of earlier experiments. Program-parameter 6 controlled the introduction of multicollinearity. Table VI-1 shows the pairwise correlations between the exogenous data used in the experiments of this section.[99]

Six experiments used the structure of the basic experiment with different values for program-parameter 6. Five other structures were investigated with various values for the program-parameter. Finally, two experiments used the basic structure and multicollinear exogenous data with smaller disturbances in one case and seventy observations, rather than twenty, in the other.[100] We will examine first the results of the experiments using the basic structure with the standard number of observations and set of disturbances.

Bias. The medians of the estimates of the structural and reduced-form coefficients made by 2SLS, UBK and 3SLS were significantly different from the true values more frequently in some experiments than

[99] Values of 2, 3, and 4 for program-parameter 6 introduced correlations between only two or three of the exogenous variables. Values of 5, 6, and 7 for the program-parameter led to all the variables being multi-collinear. The pairwise correlations between most variables increased with the size of program-parameter 6.

[100] The experiments investigating multicollinearity are experiments B through N of the second set of experiments conducted shown in Appendix C, Table 2. In one of the experiments studying different structures, thirty-five observations instead of twenty were used.

in the basic experiment. The finding of significantly biased medians did not increase systematically with the amount of collinearity. Sometimes the medians of the estimates were significantly biased more often in experiments with smaller correlations between the exogenous variables than in ones with smaller correlations.[101]

Dispersions. The dispersions of the estimates both about the true values and about their central tendencies tended to increase as more multicollinearity was introduced. Table VI-2 shows some of the typical descriptive statistics of dispersion for the structural and the reduced-form coefficients.[102] The typical descriptive statistics for the estimates of Σ and for the predictions showed similar increases.

The increases in dispersion had some peculiar features. Often the descriptive statistics for an experiment with lower correlations were quite a bit larger than those for experiments with higher correlations. Occasionally these peculiarities were evident among the typical descriptive statistics as well as among the statistics for the individual coefficients. These peculiar cases were not confined to the estimates

[101] For the estimates of the reduced-form coefficients, significantly biased medians were found most frequently when program-parameter 6 equalled 3, although the correlations were higher when it equalled 4, 5, 6 and 7. Significant biases in the estimates of the structural coefficients occurred most often when the program-parameter was set at 4 and 7. The correlations between the exogenous variables are shown in Table VI-1.

[102] The typical descriptive statistics are the averages of the statistics for the separate coefficients divided by the true values. They are described in Chapter III, Section 7.

of only one experiment.[103] Among the experiments with correlations between all the variables, each experiment frequently had the largest and the smallest descriptive statistics, even though the correlations between the variables were higher in some experiments than in others. The same was true for the experiments with deliberately introduced correlations among only one or two exogenous variables.

These peculiarities in the descriptive statistics could not be explained, as were the peculiarities found when different sizes of disturbances were used, by differences in which sets of data had to be abandoned in the course of estimating the structures. In the experiments with high multicollinearity, singular matrices were found quite often. Examination of the estimates made from the corresponding sets of data in other experiments did not indicate that differences in the sets of data used led to the differences in the descriptive statistics. The estimates made from sets of data which could be used in all the experiments often exhibited the same peculiarities as the descriptive statistics. Experiments with larger correlations between the exogenous variables frequently had estimates with smaller deviations from the true values than ones with lower correlations even when estimates were made from data using identical sets of disturbances.

FIML made a number of very bad estimates when there were large correlations between the exogenous variables. Often these estimates deviated from the true values by several hundreds percent. Negative

[103]The experiment with program-parameter 6 set at six showed the peculiar features most often.

"squared standard errors" also occurred for the FIML estimates. The remarks in Chapter V, Sections 5 and 7, about very bad FIML estimates and negative "squared standard errors" hold here. In particular, the bad estimates of the structural coefficients often led to good estimates of the reduced-form coefficients and to good predictions.[104]

Relative Performances. Table VI-3 records the totals of all the ranks of the deviations of the estimates from the true values in the experiments investigating multicollinearity in the basic structure. The results of most of the experiments suggested the ranking of the estimators into three groups which was established in the basic experiment. However, there were some cases where FIML and 3SLS did not do better than the consistent k-class estimators. The totals of the ranks for the different methods varied considerably among the experiments. These changes did not seem to be closely related to increasing correlations between the exogenous variables. For example, the position of 3SLS and FIML according to the totals of the ranks of all the estimates were poorer when program-parameter 6 equalled 5 or 7 than when it was set at 6. The correlations between the exogenous variables were higher when it was set at 6 than when it equalled 5; they were still higher when program-parameter 6 was set at 7.

[104] In one replication of one experiment, FIML had very bad estimates of the structural coefficients, the reduced-form coefficients and the predicted values. If these estimates had been deleted from the experiment, the findings of the experiment, except about the extreme quantiles of the distributions of the estimates, would not have been changed. We may note that the very bad estimates usually occurred for the estimates of the second structural equation.

Pairwise comparisons of the methods revealed several differences between the experiments. The frequencies with which one method came closer to the true values than another did not seem to vary systematically with changes in the correlations between the exogenous variables. An experiment which had lower correlations between the variables than another would often show larger differences from the basic experiment or other kinds of differences from it than did the other experiment. These peculiar cases happened for all the types of parameters investigated.

The McNemar test investigated the significance of differences between several of the experiments. Table VI-4 summarizes the results for the estimates of the structural coefficients; Table VI-5 shows the findings for the reduced-form coefficients. The test indicated that the puzzling differences between the experiments found in the rankings and the pairwise comparisons of the estimates were due to statistically significant differences between the experiments. This is brought out by the fact that for some structural coefficients, the estimates of the full-model methods came closer to the true values than other methods significantly less often when program-parameter 6 equalled 6 than when it was set at 5 or 7. The comparisons of these three experiments with the basic experiment also indicated that the results of the experiment with program-parameter 6 set at 6 could not be regarded as being mid-way between those of the other two experiments, even though the correlations between its exogenous variables fell between those of the other experiments.

TABLE VI-1

Correlations Between Exogenous Variables

MOD37A2	1	0	0	0	0	0	0
Data Set - 0		1	-.15	-.11	-.18	-.37	-.18
T = 20			1	-.18	-.01	.27	.03
Parameter 6 = 0				1	-.15	-.37	.02
Variances of disturbances as					1	-.09	.05
percentages of variances of						1	.56
endogenous variables: 8.3 - 12.0							1

MOD37A2	1	0	0	0	0	0	0
Data Set - 0		1	.95	-.13	.19	-.12	-.18
T = 20			1	-.16	-.19	-.29	-.17
Parameter 6 = 2				1	-.15	-.37	.02
Variances of disturbances					1	-.09	.05
as percentages of Y-variances:						1	.56
9.3 - 17.7							1

MOD37A2	1	0	0	0	0	0	0
Data Set - 0		1	.62	.52	-.20	-.21	.03
T = 20			1	-.18	-.01	.27	.03
Parameter 6 = 3				1	-.15	-.37	.02
Variances of disturbances					1	-.09	.05
as percentages of Y-variances:						1	.56
7.1 - 10.2							1

MOD37A2	1	0	0	0	0	0	0
Data Set - 0		1	.87	.91	-.10	.06	.06
T = 20			1	.99	-.01	.27	.11
Parameter 6 = 4				1	-.05	.19	.01
Variances of disturbances					1	-.09	.05
as percentages of Y-variances:						1	.56
5.9 - 7.5							1

TABLE VI-1 (Cont.)

MOD37A2	1	0	0	0	0	0	0
Data Set - 0		1	.71	.55	.38	.25	.59
T = 20			1	.58	.52	.57	.87
Parameter 6 = 5				1	.39	.25	.57
Variances of disturbances					1	.37	.57
as percentages of Y-variances:						1	.88
3.3 - 4.3							1

MOD37A2	1	0	0	0	0	0	0
Data Set - 0		1	.87	.73	.59	.30	.29
T = 20			1	.73	.66	.49	.35
Parameter 6 = 6				1	.58	.25	.36
Variances of disturbances					1	.42	.43
as percentages of Y-variances:						1	.72
4.4 - 5.3							1

MOD37A2	1	0	0	0	0	0	0
Data Set - 0		1	.99	.93	.87	.69	.60
T = 20			1	.93	.88	.73	.61
Parameter 6 = 7				1	.87	.67	.64
Variances of disturbances					1	.75	.69
as percentages of Y-variances:						1	.85
3.4 - 4.0							1

MODEL 46	1	0	0	0	0	0
Data Set - 0		1	-.21	.11	.24	-.20
T = 20			1	-.02	-.56	-.03
Parameter 6 = 0				1	-.17	-.12
Variances of disturbances					1	-.10
as percentages of Y-variances:						1
9.0 - 18.3						

TABLE VI-1 (Cont.)

MODEL 46	1	0	0	0	0	0
Data Set - 0		1	.62	.68	-.41	-.18
T = 20			1	.02	-.56	-.03
Parameter 6 = 3				1	-.17	-.12
Variances of disturbances					1	-.10
as percentages of Y-variances:						1
8.6 - 17.7						

MODEL 58	1	0	0	0	0	0	0	0
Data Set - 0		1	.19	.03	-.01	-.07	-.15	-.03
T = 20			1	.08	-.23	-.41	+.16	-.07
Parameter 6 = 0				1	.03	-.17	0	-.24
Variances of disturbances					1	-.01	.03	.23
as percentages of Y-variances:						1	.63	+.02
8.4 - 24.3							1	-.23
								1

MODEL 58	1	0	0	0	0	0	0	0
Data Set - 0		1	.66	.46	.49	0	.66	.30
T = 35			1	.73	.15	-.10	.54	-.09
Parameter 6 = 5				1	.14	.02	.46	-.50
Variances of Disturbances					1	.32	.63	.63
as percentages of Y-variances:						1	.57	.33
1.7 - 4.4							1	.47
								1

MOD37A2	1	0	0	0	0	0	0
Data Set - 0		1	.07	-.26	.07	-.17	0
T = 70			1	-.21	.03	.10	.10
Parameter 6 = 0				1	-.10	-.25	-.14
Variances of disturbances					1	0	-.01
as percentages of Y-variances:						1	.29
5.7 - 8.9							1

TABLE VI-1 (Cont.)

MOD37A2	1	0	0	0	0	0	0
Data Set - 0		1	.80	.51	.55	.33	.70
T = 70			1	.55	.55	.52	.88
Parameter 6 = 5				1	.43	.24	.57
Variances of disturbances					1	.35	.59
as percentages of Y-variances:						1	.84
2.6 - 3.2							1

MOD37A2	1	0	0	0	0	0	0
Data Set - 1		1	.21	-.07	0	-.03	.58
T = 20			1	.35	0	.57	.10
Parameter 6 = 0				1	-.20	.11	.36
Variances of disturbances					1	-.02	.18
as percentages of Y-variances:						1	-.03
5.7 - 7.2							1

MOD37A2	1	0	0	0	0	0	0
Data Set - 2		1	-.16	-.12	-.10	-.11	.14
T = 20			1	.05	.18	.24	.08
Parameter 6 = 0				1	-.01	0	-.21
Variances of disturbances					1	.03	.01
as percentages of Y-variances:						1	.28
5.9 - 8.8							1

TABLE VI-2

Multicollinearity - Typical Descriptive Statistics of Dispersion

Structural Coefficients

Inter-quartile Ranges

Program-Parameter 6	DLS	2SLS	UBK	LIML	3SLS	FIML
0	.386	.437	.444	.453	.293	.288
2	.268	.515	.628	.876	.443	.504
3	.329	.459	.494	.524	.370	.402
4	.598	.837	.869	1.032	.675	.721
5	.341	.493	.507	.551	.500	.528
6	.342	.620	.705	.799	.482	.519
7	.289	.659	.941	1.344	.854	1.514

Third-quartile Absolute Deviations

Program-Parameter 6	DLS	2SLS	UBK	LIML	3SLS	FIML
0	.436	.376	.383	.384	.255	.244
2	.867	.606	.706	.900	.561	.424
3	.624	.446	.454	.441	.382	.348
4	.990	.774	.772	.822	.570	.562
5	.549	.453	.455	.487	.433	.444
6	.856	.570	.614	.678	.435	.449
7	1.151	.971	1.182	1.677	.995	2.666

Reduced-Form Coefficients

Inter-quartile Ranges

Program-Parameter 6	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
0	.597	.595	.597	.590	.313	.297	2.012
2	.828	.677	.745	.834	.484	.483	2.253
3	.710	.674	.722	.723	.501	.487	2.667
4	1.024	1.268	1.280	1.343	.946	.946	12.907
5	.531	.597	.600	.596	.555	.592	11.618
6	.768	.749	.789	.788	.556	.519	2.874
7	2.080	1.244	1.363	1.284	1.172	1.137	5.152

Third-quartile Absolute Deviations

Program-Parameter 6	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
0	.546	.478	.476	.475	.286	.256	1.742
2	.952	.581	.623	.684	.467	.453	1.917
3	.660	.590	.600	.599	.518	.415	2.267
4	1.119	1.229	1.220	1.184	.819	.770	10.764
5	.575	.521	.524	.531	.490	.506	10.162
6	.895	.652	.676	.690	.459	.440	2.382
7	2.189	1.521	1.647	1.288	1.618	.998	4.222

TABLE VI-3

Multicollinearity - Totals of Ranks of Deviations of Estimates from True Values

Structural Coefficients

Program-Parameter 6	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W
0	3334	2751	2698	2771	2078	2118	.112
2	3702	2756	2475	2598	2158	2061	.176
3	3502	2649	2508	2518	2380	2193	.106
4	3703	2730	2478	2589	2125	2125	.172
5	3487	2555	2508	2559	2302	2339	.097
6	3720	2691	2517	2630	2104	2088	.180
7	3483	2620	2371	2443	2365	2468	.094

Elements of Σ

Program-Parameter 6	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W
0	1631	1269	1028	896	819	657	.392
2	1637	1265	992	930	810	666	.390
3	1647	1238	985	842	848	740	.366
4	1686	1259	983	847	818	707	.422
5	1667	1233	973	835	838	754	.380
6	1679	1270	989	862	802	698	.424
7	1499	1166	918	861	920	936	.189

Reduced-Form Coefficients

Program-Parameter 6	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
0	5020	4251	4135	4137	3127	3211	5519	.147
2	5054	3798	4110	4173	3461	3315	5489	.126
3	4301	4087	4103	4048	3942	3313	5606	.093
4	4441	4034	4095	4233	3594	3325	5678	.111
5	4350	3782	3877	3869	3581	3620	6321	.182
6	4800	3921	4121	4148	3377	3327	5706	.135
7	4771	4023	4219	3784	3973	3646	4978	.048

Predictions

Program-Parameter 6	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
0	1899	1768	1782	1825	1560	1496	2270	.067
2	2000	1703	1718	1692	1689	1685	2113	.034
3	1848	1827	1793	1680	1760	1549	2143	.035
4	1896	1737	1781	1790	1598	1589	2209	.047
5	1780	1674	1724	1830	1709	1687	2196	.035
6	1980	1709	1786	1824	1563	1552	2186	.054
7	2099	1785	1821	1596	1782	1659	1858	.027

TABLE VI-4

Multicollinearity - Significant Differences between Experiments, Structural Coefficients

Experiments A-C

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	2	1	2	0	0
2SLS	0	-	2	2	0	0
UBK	0	0	-	0	0	0
LIML	0	2	0	-	0	0
3SLS	1	3	3	3	-	0
FIML	1	2	1	1	0	-

Experiments A-E

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	6	5	5	0	2
2SLS	1	-	1	1	0	0
UBK	1	1	-	0	0	0
LIML	1	1	0	-	0	0
3SLS	2	4	2	2	-	0
FIML	2	3	4	3	1	-

Experiments A-F

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	9	7	5	7	5
2SLS	1	-	5	2	0	0
UBK	1	1	-	0	0	0
LIML	1	1	0	-	0	0
3SLS	0	0	0	0	-	0
FIML	0	0	0	0	0	-

Experiments A-G

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	6	3	3	1	0
2SLS	0	-	5	2	0	0
UBK	0	0	-	0	0	0
LIML	0	0	0	-	0	0
3SLS	1	2	4	3	-	0
FIML	3	4	5	4	0	-

Experiments E-F

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	1	0	0	6	3
2SLS	0	-	4	0	4	2
UBK	0	0	-	0	3	3
LIML	0	0	0	-	2	3
3SLS	0	0	0	0	-	0
FIML	0	0	1	0	0	-

Experiments E-G

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	4	3	2	2	1
2SLS	1	-	4	0	0	0
UBK	1	0	-	0	0	0
LIML	3	0	0	-	1	0
3SLS	3	0	3	2	-	0
FIML	5	0	1	2	0	-

Experiments F-G

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	1	2	1	0	0
2SLS	2	-	1	0	0	0
UBK	6	0	-	0	0	0
LIML	5	0	0	-	0	0
3SLS	6	2	2	3	-	0
FIML	8	2	2	3	0	-

See notes to Table VI-5 for explanation of table entries.

TABLE VI-5

Multicollinearity - Significant Differences between Experiments, Reduced-Form Coefficients

Experiments A-C

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	1	1	2	0	0	2
2SLS	4	-	2	0	0	0	0
UBK	4	0	-	0	0	0	0
LIML	3	0	0	-	0	0	0
3SLS	9	6	7	7	-	1	3
FIML	4	4	3	2	0	-	1
LSRF	6	3	1	1	3	2	-

Experiments A-E

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	1	1	0	0	0	0
2SLS	3	-	0	0	0	0	0
UBK	3	1	-	0	0	0	0
LIML	3	0	0	-	0	0	0
3SLS	9	6	5	7	-	0	0
FIML	8	7	6	6	2	-	0
LSRF	14	9	10	9	5	6	-

Experiments A-F

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	1	0	1	0	0	2
2SLS	2	-	0	0	0	0	0
UBK	4	2	-	0	0	0	0
LIML	4	0	0	-	0	0	0
3SLS	3	7	3	3	-	0	0
FIML	2	1	1	0	0	-	0
LSRF	4	1	0	0	1	0	-

Experiments A-G

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	3	2	3	0	0	5
2SLS	0	-	0	0	0	0	2
UBK	0	0	-	0	0	0	4
LIML	0	0	0	-	0	0	2
3SLS	5	7	5	11	-	0	8
FIML	5	2	1	5	0	-	1
LSRF	1	0	0	0	0	0	-

Experiments E-F

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	4	2	2	6	6	8
2SLS	1	-	0	0	4	5	4
UBK	1	1	-	0	3	6	4
LIML	0	1	0	-	3	4	8
3SLS	1	1	1	0	-	2	6
FIML	1	1	0	0	0	-	5
LSRF	0	0	0	0	0	0	-

Experiments E-G

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	4	4	4	1	4	10
2SLS	1	-	0	0	0	3	12
UBK	2	0	-	1	0	1	11
LIML	2	0	0	-	0	0	8
3SLS	1	3	2	1	-	2	14
FIML	1	2	1	0	0	-	12
LSRF	0	0	0	0	0	0	-

TABLE VI-5 (Cont.)

Experiments F-G

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	0	0	0	0	0	5
2SLS	3	-	1	3	0	0	7
UBK	4	0	-	3	0	0	10
LIML	4	0	0	-	0	0	0
3SLS	4	2	1	6	-	0	6
FIML	2	0	0	0	0	-	0
LSRF	0	0	0	0	0	0	-

Notes:

Entries in each row of the tables in Tables VI-4 and VI-5 record the number of coefficients for which the method came closer to the true values than each other method significantly more often in the experiment mentioned first than in the one mentioned second. The McNemar test and the .05 significance level were used. There were fifteen structural coefficients and twenty-one reduced-form coefficients. The experiments differ in the values assigned to program-parameter 6. The values were:

Experiment	Program-Parameter 6
A	0 (the basic experiment)
C	3
E	5
F	6
G	7

These letters for the experiments are those of the second set of experiments conducted (Appendix C, Table 2). MOD37A2, the structure of the basic experiment, was used in all the experiments.

<u>Standard Errors</u>. In some experiments studying multicollinearity, the standard errors of some of the methods were less reliable than in the basic experiment.[105] The deteriorations were more marked for the coefficients of the second structural equation than for others. In some experiments, one could reject quite frequently the hypothesis that ninety-five percent of the t-ratios fall within the 95% confidence interval of the t-distribution with five degrees of freedom.[106] The hypothesis that the adjusted t-ratios follow the t-distribution with fifteen degrees of freedom also fared poorly.

<u>Findings using the Basic Structure</u>. The results of the experiments investigating multicollinearity in the basic structure had several puzzling features. One might expect multicollinearity to make estimation more difficult for two closely-related reasons. With multicollinearity, the distinctions between the exogenous variables are less clear-cut. While the effects of each variable remain unchanged, the closeness of the variables to each other might make estimation of the part each plays in determining the endogenous variables more difficult. In addition, multicollinearity tends to make the variance-covariance matrix of the exogenous variables more ill-conditioned. This matrix enters in one way or another into the calculations of all the estimates. One might expect the calculations to become more inaccurate with higher correlations

[105] This was particularly true when program-parameter 6 was set at 2 and 7.

[106] This is the number of observations minus the total number of coefficients to be estimated. Fifteen is the number of observations minus the number of coefficients in each equation. The hypothesis mentioned in the text was rejected frequently when program-parameter 6 was set at 2, 5 and 7.

between the exogenous variables. Intuitively, one might expect the confluence of the relationships arising from multicollinearity to affect the more subtle methods to a greater extent than the other ones since their subtlety might lead to their being more confused by the multicollinearity. From the computational side, one would expect the full-model estimators to be more affected than the single-equation methods since they require the inversion of larger matrices. One might also expect DLS to gain relative to the other methods since its calculations are simpler.

The findings of the experiments gave but weak support to these conjectures. While there was a tendency for the dispersions of the estimates to increase with larger correlations between the exogenous variables, the increase was not at all regular.[107] The relative positions of the methods were affected, but again the changes were not closely connected to the changes in the correlations among the variables.

It was found in Chapter V, Section 3, that the performances of the methods were sensitive to the exact sets of exogenous data used.[108] It is impossible to change the correlations between the variables without changing the variables used. The results of the experiments seem to suggest that the effects of changing the sets of data used were themselves

[107] FIML became harder to estimate with increases in the correlations between the exogenous variables in the sense that it tended to require more iterations to reach the FIML estimates.

[108] A comparison of the correlations shown in Table VI-1 with the results of the section indicates that those results cannot be ascribed to different correlations between the variables of the different data sets.

more important than those of changing the levels of multicollinearity. The impossibility of separating the two aspects in the experiments seems to make it impossible to draw many conclusions about the effects of multicollinearity on the estimates.

Other Experiments. The experiments which used other structures usually indicated that multicollinearity led to larger dispersions of the estimates, but this was not always the case. Marked changes in the relative performances of the estimators did arise from multicollinearity, but slight deteriorations in the standings of the full-model methods were apt to occur. There was, however, some variation between the experiments in the apparent effects of having multicollinearity.

The experiments using the basic structure with seventy observations or small disturbances as well as multicollinearity, indicated that multicollinearity increased the dispersions of the estimates and led to slight deteriorations in the relative positions of the full-model methods. These results were similar to those found when twenty observations and the standard size of disturbances were used with the basic structure and the same values for program-parameter 6.

Summary. The dispersions of the estimates tended to increase with larger correlations between the exogenous variables. The relative performances of the full-model methods were apt to deteriorate slightly. However, there were a number of peculiar features in the results and the changes were not very closely related to the changes in correlation between the variables.

It was suggested that this was due to the sensitivity of the estimators to the exact sets of data used. The fact that different sets of data were used, per se, may account for the changes in performance as much as the fact that these differences came from increasing the correlations between the exogenous variables.

4. Lagged-Endogenous Variables

The effects of having lagged-endogenous variables among the predetermined variables were next investigated. Program-parameter 4 controlled the use of lagged-endogenous variables. Letting y_{gt} be the value of the g^{th} endogenous variable at observation t and z_{kt} be the value of predetermined variable k at observation t $(g = 1,\ldots,G;$ $k = 1,\ldots,K;$ $t = 1,\ldots,T)$, the effects of program-parameter 4 were:

Program-Parameter 4	Effect
0	No lagged-endogenous variables
1	$z_{K-1,t} = y_{G-1,t-1}$
2	$z_{K-1,t} = y_{G-1,t-1}$
	$z_{K,t} = y_{G,t-1}$
3	$z_{K-1,t} = y_{G-1,t-1}$
	$z_{K,t} = y_{G,t-1}$
	$z_{K-2,t} = y_{G-2,t-2}$

Other predetermined variables were the same as the exogenous variables used when there were no lagged-endogenous variables. The initial values of the predetermined variables, that is, the values of z_{k1}, were simply the values for the variables when no lagged-endogenous variables were used.

One experiment used the basic structure with program-parameter 4 set at unity. The program-parameter was set at 2 for experiments using the basic structure and two other structures.[109] Experiments were also conducted with two lagged-endogenous variables in the basic structure with deliberately-introduced multicollinearity among the exogenous variables. Finally, three experiments used two different structures with program-parameter 4 set at 3.[110]

The results of the experiments using the basic structure with one and two lagged-endogenous variables did not show many pronounced differences from the results of the basic experiment. The dispersions of the estimates of all methods were smaller with lagged-endogenous variables. This was less pronounced for FIML than for other methods. The dispersions of the estimates were slightly larger in the experiment with two lagged-endogenous variables than in the one with only one.

---

[109]The other structures were MOD37A4 and MODEL37B.

[110]Two of these used MOD37A4 with twenty and seventy observations. The third experiment used MOD37AE. MOD37A4 differed from the basic structure in the values of the structural coefficients. MOD37AE used the same structural coefficients, but had lower correlations between the structural disturbances than MOD37A4. MODEL37B was of different structural form. The experiments reported in this section are experiments O through W of the second set of experiment conducted, Appendix C, Table 2.

The decreases in the dispersions should not, probably, be ascribed to lagged-endogenous variables alone. The use of lagged-endogenous variables reduced the size of the variances of the disturbances relative to the variances of the endogenous variables.[111] The decreases in the relative roles of the disturbances could account for the decreases in the dispersions of the estimates.

There was no tendency for the number of times the medians of the estimates were significantly biased to increase with the use of lagged-endogenous variables. Nor were there many changes in the performances of the standard errors, except that when one lagged-endogenous variable was used, the DLS standard errors were somewhat less misleading than in the basic experiment.

The performances of the methods relative to each other were not greatly affected by the introduction of lagged-endogenous variables into the basic structure. The concordances of the ranks of the rank-totals of the descriptive statistics in the two experiments with lagged-endogenous variables and the basic experiment were usually above .9 for the aspects studied other than the predictions. There were no very pronounced differences in the numbers of times one method came closer to the true values than another between the experiments. There did seem to be a

[111]With program-parameter 4 set equal to unity, the disturbances accounted for from 2% to 10% of the variances of the endogenous variables in the basic structure. With the program-parameter set at 2, the disturbances accounted for from 1.5% to 8% of the variances of the endogenous variables. In the basic experiment, the disturbances accounted for 8% to 12% of the variances of the endogenous variables.

slight tendency for the relative performances of FIML, DLS and 3SLS to be worse when there were two lagged-endogenous variables than when there were none.

The McNemar test investigated the significance of the differences in the experiment with two lagged-endogenous variables from the basic experiment. For the structural coefficients, the estimates of Σ and the reduced-form coefficients there were very few significant differences between the experiments. They were so rare that one would hesitate to say that the performances of the methods relative to each other were affected by the use of lagged-endogenous variables. This is especially true because we have already found that changing the predetermined variables used in an experiment may affect the relative performances of the methods.[112]

More pronounced differences between the experiments were found in the relative performances of the methods for their predictions than for other aspects. The relative positions of FIML, DLS and LSRF seemed to be quite sensitive to the introduction of lagged-endogenous variables according to the rank-totals for all the separate predictions.[113] In comparing the experiment with two lagged-endogenous variables with the basic experiment, the McNemar test found that the performances of FIML

[112]There were significant differences for three of the fifteen structural coefficients in the performances of LIML relative to UBK. In some of the comparisons of the full-model methods with the consistent k-class estimators, significant differences occurred for two coefficients. Otherwise there were significant differences in the performances of the methods for at most one structural coefficient.

[113]These rank-totals are shown in Appendix D, Table B-12.

relative to the consistent k-class estimators were significantly worse when there were two lagged-endogenous variables for a third of the values predicted. DLS did better than the other k-class estimators for two of the values. Other significant differences in the relative performances of pairs of methods occurred for at most one of the values predicted. The values of Kendall's W for the ranks of the rank-totals of the descriptive statistics of the distributions of the predictions in the three experiments using the basic structure were somewhat lower than for the other aspects, but they were usually above .7. Reversals of order from that of the basic experiment occurred quite frequently in the experiments with one and two lagged-endogenous variables between FIML and the consistent k-class estimators and between them and DLS.

The results of using other structures with two lagged-endogenous variables were in general agreement with those using the basic structure. The dispersions of the estimates tended to decrease, but this might be due to the disturbances playing a smaller role when there were lagged-endogenous variables. Except for the predictions, the relative performances of the methods did not seem sensitive to the introduction of lagged-endogenous variables. The standings were somewhat more sensitive in the predictions. However, no pronounced pattern for the changes occurred.[114]

The two experiments using the basic structure with lagged-endogenous variables and deliberately-introduced multicollinearity held

[114] For example, the position of LSRF relative to some other methods improved with the introduction of lagged-endogenous variables in MOD37A2 and MOD37A4, but became worse in MODEL37B.

no great surprises. Lagged-endogenous variables reduced the dispersions of the estimates from the levels in the experiments with multicollinearity without lagged-endogenous variables. This could probably be ascribed both to a reduction of the role of the disturbances and to a lowering of the correlations among the exogenous variables.[115] Striking differences in the relative performances of the methods did not occur. Interestingly enough, such differences as were found showed the experiments with lagged-endogenous variables to have the same sorts of differences among themselves as had been found among the corresponding experiments without lagged-endogenous variables. These differences between the experiments were less pronounced when there were lagged-endogenous variables. The relative positions of the methods were more sensitive to the introduction of lagged-endogenous variables in making predictions than in other aspects, but again no pattern for the effects of lagged-endogenous variables emerged.

The dispersions of the estimates of all parameters were a great deal smaller when three lagged-endogenous variables were used than when there were none. This can probably be attributed to a decrease in the role of the disturbances.[116]

Differences in the relative performances of the methods when there were three lagged-endogenous variables from when there were none

[115] The lagged-endogenous variables were far less highly correlated with the other predetermined variables than were the exogenous variables they replaced.

[116] The structural disturbances accounted for only about one-sixth as many of the variances of the endogenous variables when there were three lagged-endogenous variables as when there were none.

Summary. The introduction of lagged-endogenous variables did not affect the estimators greatly. Changes in the dispersions of the estimates could probably be ascribed to the disturbances playing smaller roles in these experiments than in the basic experiment. Only when three lagged-endogenous variables were used were there many significant differences from the corresponding experiments which used only exogenous variables. Whether these differences should be ascribed to having lagged-endogenous variables rather than just to having different predetermined variables is a moot question.

5. Time-Structures in the Exogenous Variables

The next group of experiments investigated the effects of having systematic relationships in the exogenous variables between the observations of different periods. This was achieved either by the introduction of trends in the variables or by making the exogenous variables autocorrelated. Program-parameter 5 controlled the introduction of time-structures. The effects of different values for this program-parameter are shown in Table VI-7.

The changes in the generation of the data investigated in this section might not be expected, by themselves, to affect the relative performances of the methods. Many of the features studied could have been achieved by re-ordering the observations of the basic experiment to give the appearance that these aspects were present in the data. However, in conjunction with some other features of the generation of

TABLE VI-7

Time-Structures in the Exogenous Variables - Types Investigated

Program-Parameter 5	Effect
0	$z_{kt} = q_{kt}$; all k and t.
1	$z_{kt} = t$; $k = K$, all t;
	$z_{kt} = q_{kt}$; $k \neq K$, all t.
2	$z_{kt} = t$, $k = K$, all t;
	$z_{kt} = (2t/T)q_{kt}$, $k = K-1$, all t;
	$z_{kt} = q_{kt}$, $k = 1,\ldots,K-2$, all t.
3	$z_{kt} = (.5 + t/T)q_{kt}$, $k \neq 1$;
	$z_{kt} = q_{kt} = 1$, $k = 1$.
5	$z_{kt} = q_{kt} - 50 + 100t/T$, $k \neq 1$;
	$z_{kt} = q_{kt} = 1$, $k = 1$.
6	$z_{kt} = .4q_{kt} + .6z_{kt-1}$, $t \neq 1$;
	$z_{kt} = q_{kt}$, $t = 1$.
7	$z_{kt} = .2q_{kt} + .8z_{kt-1}$, $t \neq 1$;
	$z_{kt} = q_{kt}$, $t = 1$.

Symbols: z_{kt} - the k^{th} exogenous variable ($k = 1,\ldots,K$) at the t^{th} observation ($t = 1,\ldots,T$) used when program-parameter 5 had the value indicated.

q_{kt} - the k^{th} exogenous variable at the t^{th} observation in the set of exogenous data used when program-paramater 5 had a velue of zero.

the data studied in the next chapter, they might lead to considerable changes in performance. The experiments discussed in this section were conducted to give bases for comparison for these later experiments.

The need for the experiments of this section was heightened by the sensitivity of the results to the sets of predetermined data used, which was found in Chapter V and in earlier sections of this chapter. One cannot presume that a change in the exogenous data will not affect the relative standings of the methods. The experiments of this section, indeed, provide further explorations of the sensitivity of the results to the sets of data used and many of them were conducted only with this end in mind.

A large variety of experiments investigated the effects of having time-structures both by themselves and together with lagged-endogenous variables or multicollinearity. Several different structures were examined. We shall not discuss the experiments individually but shall only mention special features which stood out in the results of the group of experiments.[118]

<u>Distributions of the Estimates</u>. The dispersions of the estimates showed considerable variation among the experiments where time-structures were introduced into the exogenous variables. Frequently these changes could be attributed to the disturbances accounting for different proportions of the variances of the endogenous variables or to different

---

[118]The experiments discussed in this section are experiments X through AO of the second set of experiments conducted, shown in Appendix C, Table 2.

correlations between the exogenous variables. However, this was not always the case. When one of the exogenous variables was replaced by a trend variable, namely the number of the observation,[119] the dispersions of the estimates were larger than those in the basic experiment, although the disturbances did not play a larger role and the correlations between the exogenous variables were smaller.

In only a few instances were the medians of the estimates significantly biased more frequently than in the basic experiment. The main case occurred for the estimates of 2SLS and 3SLS when the exogenous variables were highly auto-correlated.[120] In that experiment the disturbances accounted for over 50% of the variances of the endogenous variables. Large disturbances were found in Chapter V, Section 7, to lead to the occurrence of significantly biased medians more frequently than in the basic experiment. The number of significant biases was not larger than in the basic experiment when the variables were less highly auto-correlated.[121]

<u>Ranking of the Methods</u>. The ranking of the methods into three groups, as suggested by the results of the basic experiment, was found in most of the experiments which had time-structures in the exogenous variables. When this was not true, the corresponding experiments without time-

---

[119]This is the effect of setting program-parameter 5 equal to unity.

[120]That is, when program-parameter 5 was set at 7.

[121]That is, when program-parameter 5 equalled 6.

structures also had not shown the ranking of the methods into three groups.[122] There was one exception to this pattern of support for the findings of earlier experiments. When MODEL48 was used with highly auto-correlated exogenous variables, FIML was not better than all the consistent k-class estimators for the structural coefficients and 3SLS was not better for reduced-form coefficients. There was some variation between the experiments in the totals of the ranks for the estimates of any one method, but these differences were not usually enough to weaken the rankings greatly. On the other hand, in none of the experiments was the agreement among the ranks of the deviations from the true values of the estimates much closer than in the basic experiment.

<u>Significant Differences between Experiments</u>. The significance of the differences of several of the experiments with time-structured from the basic experiment and from each other were investigated by the McNemar test. In most comparisons, significant differences were found for a few of the parameters estimated. The standings of DLS relative to the other methods and the positions of the full-model methods relative to other methods seemed to be particularly sensitive to the use of time-structures in the exogenous variables. Occasionally the performance of one method relative to another improved significantly for some parameters, while for others it deteriorated significantly. There was nothing in either the changes in the standings or the frequencies with which these

[122] For example, when there were small correlations between the structural disturbances, the full-model methods did not do better than the consistent single-equation methods.

changes were significant to suggest that they were due to anything but to having different sets of exogenous data, regardless of the time-structure introduced, or to having the disturbances play different roles.

Standard Errors. In most of the experiments, the performances of the standard errors were comparable to the performances when no time-structures were introduced. The one exception to this was that the standard errors were not very reliable when highly auto-correlated exogenous variables were used. The results of Chapter V, Section 7, where it was found that very large disturbances led to unreliable standard errors, suggest that the large roles the disturbances played in this experiment may account for the poor performances rather than the introduction of auto-correlated exogenous variables.

Experiments with other Special Features. Time-structures were used in some experiments which also had multicollinearity or lagged-endogenous variables among the predetermined variables. The results of these experiments seemed to be a combination of the results of the experiments which had the features separately. Thus, while multicollinearity tended to make the dispersions of the estimates larger and weakened the ranking of the methods when there were time-structures in the exogenous variables, these effects seemed hardly more pronounced, and were sometimes less so, than when multicollinearity was introduced without time-structures. More remarkable, the combination of lagged-endogenous variables with the time-structures did not greatly affect the results and it did not seem that the two features in conjunction with each other led to results greatly different from those of the basic experiment or from the findings of the experiments which had lagged-endogenous variables only.

Summary. The introduction of time-structures into the exogenous variables did not alter the performances of the methods greatly. Many of the changes which occurred could probably be attributed to there being different correlations between the exogenous variables or to the disturbances playing different roles in the generation of the endogenous variables. Such differences as did occur in the relative standings of the methods seemed only to provide further confirmation of the suggestion that the relative performances of the methods are sensitive to the exact sets of exogenous data used. The use of time-structures with lagged-endogenous variables did not lead to pronounced changes in the performances of the methods.

6. Conclusion

The experiments discussed in this chapter held few surprising results. They did indicate that the results of Monte Carlo experiments are fairly sensitive to the use of different sets of predetermined variables. This sensitivity led to the finding of several surprising results when multicollinearity was investigated. It was probably to blame for many of the differences between the experiments found when investigating lagged-endogenous variables and time-structures in the exogenous variables. Given the difficulties caused by this sensitivity to the actual sets of data used, it seemed, nevertheless, that multicollinearity led to greater dispersions of the estimates, more observations led to smaller dispersions and that the other features, *in themselves*, did not affect the dispersions. In most of the experiments, the ranking of the methods found in earlier chapters still seemed to hold. The weakness of the ranking was again an outstanding feature of the experiments. It should be noted, however, that the experiments of this chapter did not give independent support to earlier findings. The structures used and, in most experiments, the data employed were closely related to those used in the earlier experiments.

CHAPTER VII

Other Statistical, Simultaneous-Equation Models

1. Introduction

A large group of the Monte Carlo experiments reported in this study used data which did not conform entirely to the assumptions of the Cowles-Commission model, (1.1)-(1.6). The findings of experiments in which the statistical assumptions of the standard model, (1.4) and (1.5), were violated in one way or another are discussed in this chapter.

The principal question investigated is the sensitivity of the distributions of the estimates to the assumptions on which the derivations of the estimators were based. The performances of the estimators relative to each other will be examined especially. Because the data are not fully appropriate to the estimators, one might expect the estimates to exhibit different properties in the experiments of this chapter from those of earlier chapters. In many cases, the directions of minimization of the residuals used by the consistent estimators may no longer seem appropriate. Intuitively, the methods which deal with the problem of simultaneity in more subtle ways might be expected to suffer more than the straightforward methods from the violations of the assumptions.[123]

[123] The principles from which the estimators can be derived were discussed briefly in Chapter II. That subtlety does not necessarily make for better estimators was shown in earlier chapters. LIML, which seems to rest on sounder principles than the other k-class estimators and which is, in particular, the maximum-likelihood estimator, did not do better than other consistent single-equation estimators in the earlier experiments except for estimating the elements of Σ.

In particular, the full-model methods might show poorer performances than the single-equation methods and DLS might have better standings relative to the other methods than it had in earlier experiments.

Many of the violations of the assumptions studied in this chapter could be introduced only by altering the exogenous data or the values of the structural coefficients. From the findings of earlier chapters, we can expect the results of the experiments to show some sensitivity to these changes. Small changes in the standings of the methods could be ascribed to this sensitivity rather than to the fact that the change in the data was accompanied by a violation of the assumptions of the standard model. Only pronounced changes in the results can be attributed with much confidence to the violations of the assumptions.

2. Errors in the Exogenous Variables

Measurement error of the exogenous variables was the first violation of the assumptions of the standard model investigated. In these experiments, one set of exogenous data--usually the one used in the basic experiment--generated the endogenous variables. Setting program-parameter 3 equal to unity added stochastic elements to the exogenous data. The structure was then estimated from this new set of data. Program-parameter 16 was a scale-factor controlling the size of these errors.[124] Let

[124] In this section the term "scale-factor" will be used to refer to program-parameter 16 and not to program-parameter 13 which controlled the size of the structural disturbances.

q_{kt} be the exogenous variables used in generating the data $(k = 1,\ldots,K;\ t = 1,\ldots,T)$;

z_{kt} be the exogenous data used for estimating the structure;

v_{kt} be independently and normally distributed pseudo-random numbers of mean zero and unit variance; and

$\underline{h}$ be the value of program-parameter 16.

Setting program-parameter 3 equal to unity led to

$$z_{kt} = q_{kt} + \underline{h}v_{kt}\ , \qquad k \neq 1, \text{ all } t;$$

$$z_{kt} = q_{kt} = 1\ , \qquad k = 1, \text{ all } t.$$

The v_{kt}, like the usual structural disturbances, were not the same in the different replications of an experiment. The q_{kt} remained unchanged throughout the replications, as in other experiments.

Four experiments used the structure and the structural disturbances of the basic experiment[125] with errors added to the exogenous data of that experiment. These experiments had the scale-factor determining the sizes of the errors set at 1, 2, 4, and 8. One experiment had multicollinearity in the exogenous data to which the errors were added.[126]

[125] MOD37A2, shown in Table IV-1, is the structure of the basic experiment.

[126] This experiment also used MOD37A2 and had the scale-factor set at 2. Multicollinearity was introduced by setting program-parameter 6 equal to 5.

Finally, two experiments investigated the effects of measurement errors in other structures.[127]

The results of using the basic structure with measurement errors were more notable for their lack of differences from those of the basic experiment than for any startling effects to be attributed to measurement errors. The dispersions of the estimates of the structural coefficients about their central tendencies and about the true values were larger when there were measurement errors in the exogenous variables. The larger the sizes of the errors, the greater were the dispersions. Indeed, the increases in the dispersions of the estimates seemed to be roughly proportional to the sizes of the scale-factor for the errors added to the exogenous variables. According to most of the descriptive statistics of dispersion, the increases in dispersion were slightly larger for the full-model methods than for others. With the introduction of measurement errors, stochastic elements played a larger role in the structure than when the structural disturbances were the only stochastic parts of the structure. The increases in dispersion might well be due to the larger role of the stochastic elements in the structures, *per se*, regardless of the fact that they were introduced as measurement errors. The changes were quite similar to those found in Chapter V, Section 7, as a result of increasing the sizes of the structural disturbances.

[127] The structures used were MOD37A1, which differed from MOD37A2 in the values of the structural coefficients, and MOD37AB, which differed from MOD37A2 by having smaller correlations between the structural disturbances. These experiments had the scale-factor set at 2. The experiments disussed in this section are experiments B through H of the third set of experiments, Appendix C, Table 3.

The differences of the medians or the averages of the estimates of the structural coefficients from the true values tended to increase in size with the size of the errors in the exogenous variables.[128] The frequencies with which the medians were significantly different from the true values did not increase in any marked fashion with the introduction of errors of measurement. However, in the experiment with the largest errors, several significant median biases were found among the estimates of the consistent methods.[129] Even in this experiment they occurred for at most one-third of the coefficients. The failure to find many significantly biased medians is surprising since errors in the exogenous variables do bias the estimates of the single-equation model.[130] Significant biases may not have been found frequently in the experiments because the dispersions of the estimates were so large in relation to their biases that any biases of the estimators were not statistically significant in samples of fifty estimates rather than because the estimators are actually unbiased.[131] The findings do indicate that errors in the exogenous variables do not lead to serious biases of the estimates of the structural coefficients.

---

[128] This does not appear in most of the typical biases presented in Appendix D since the biases of the separate coefficients largely cancelled out in forming the typical descriptive statistics.

[129] Throughout this chapter we refer to the methods by their properties in the standard model whether or not these properties are true for the models being estimated.

[130] See J. Johnston, Econometric Methods, (New York: McGraw-Hill, 1963), Chapter 6.

[131] This may, of course, also be the explanation for rarely finding significantly biased medians for the estimates of the consistent methods in the experiments of earlier chapters.

Errors in the exogenous variables did not greatly affect the standings of the methods in estimating the structural coefficients. The concordances of the ranks of the rank-totals in the basic experiment and the ones with different sizes of measurement errors were well above .8, as measured by Kendall's W, for most of the measures of dispersion examined. When this was not true,[132] Kendall's W was still above .7 .

Some changes in the relative performances of the methods did occur. FIML and 3SLS came closer to the true values than other methods less often as larger errors in the exogenous variables were introduced. On the other hand, DLS and 2SLS improved somewhat. The deteriorations of the relative positions of FIML were also evident in the rankings of most of the descriptive statistics of dispersion.

The McNemar test found few significant differences between the basic experiment and the one with small measurement errors.[133] These few showed the performances of the full-model methods to have been significantly poorer when there were measurement errors. In comparing the experiment with large measurement errors[134] with the basic experiment, significant differences were found quite often. Table VII-1 records their frequencies.

One other change in the relative performances of the methods bears mentioning. As the errors in the exogenous variables increased, the concordances between the ranks of the deviations of the estimates of all the coefficients from the true values declined. When the scale-

[132]It was not true for the ranks of the median absolute deviations, the interquartile ranges and of the sums over the coefficients of the ranks of the sums over the replications of the ranks of the absolute deviations of the estimates from the true values.

[133]That is, the experiment with the scale-factor set at 2.

[134]That is, the experiment with the scale-factor set at 8.

factor was largest, Kendall's W was only .04 for the totals of these ranks; it was .11 in the basic experiment. There was also less agreement among the separate coefficients about the over-all ranking of the methods suggested by the sums over the replications of the ranks of the deviations from the true values of the estimates of each coefficient. In the basic experiment, Kendall's W was .74 for the sums of the ranks of these totals over the separate coefficients. In the experiment with the largest errors it was only .19. Although this last figure is significantly different from zero at the .05 level, it does indicate that any ranking of the methods in that experiment must be very tenuous.

The standard errors of each consistent method were apt to become poorer as larger errors in the exogenous variables were introduced. However, the deteriorations were slight and were not usually statistically significant if judgement of the standard errors was based on the number of t-ratios falling outside the three intervals examined.[135] One could rarely have rejected the hypothesis that inferences based on the belief that the t-ratios followed the t-distribution with five degrees of freedom or that the adjusted t-ratios followed it with fifteen degrees of freedom would be accurate.

The dispersions of the estimates of Σ made by all methods increased with the size of the errors in the exogenous variables--as

[135] The three intervals were the 95% confidence intervals of the t-distribution with five and fifteen degrees of freedom and the interval -2 to +2. The t-ratios are the deviations of the estimates from the true values divided by their standard errors.

might be expected from the larger roles played by the non-systematic elements. The rankings of the methods for measures of dispersion about the central tendencies of the estimates showed no marked changes, but the rankings for the dispersions about the true values were substantially altered. Any method which had tended to do better than another in the basic experiment became relatively poorer with increasingly large errors in the exogenous variables. The changes in performance between the basic experiment and the one with the largest measurement errors were found to be statistically significant by the McNemar test for all elements of Σ. If the true values of Σ are re-defined to include both the structural disturbances and the measurement errors, the rankings of the methods according to the measures of dispersion from the true values were not changed in any marked way by the introduction of errors in the exogenous variables.

The estimates of the reduced-form coefficients were affected by the introduction of errors in the exogenous variables in much the same ways as were the estimates of the structural coefficients. The dispersions of the estimates of all methods increased with the full-model methods seeming more sensitive than the others. The sizes of the biases of the estimates of the individual coefficients tended to increase, but medians which were significantly different from the true values were found with greatly increased frequency only in the experiment with the largest errors. In that experiment the estimates made by the consistent techniques were significantly biased for from one-third to one-half of the reduced-form coefficients.

TABLE VII-1

Errors in the Exogenous Variables - Significant Differences from the Basic Experiment*

Structural Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	3	2	2	1	0
2SLS	4	-	2	2	0	0
UBK	4	3	-	2	0	0
LIML	4	4	0	-	0	0
3SLS	3	5	4	7	-	0
FIML	4	6	5	5	0	-

Reduced-Form Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	0	0	0	0	0	1
2SLS	6	-	0	0	0	0	0
UBK	7	8	-	1	0	0	1
LIML	7	6	0	-	0	0	0
3SLS	6	6	4	4	-	0	5
FIML	8	10	10	6	6	-	4
LSRF	1	0	0	0	0	0	-

Predictions

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	2	2	2	2	0	2
2SLS	1	-	2	2	0	0	0
UBK	1	2	-	1	0	0	0
LIML	1	2	2	-	0	0	0
3SLS	0	1	1	1	-	0	2
FIML	1	1	1	2	1	-	2
LSRF	0	0	0	0	0	0	-

*Each row of the table records the number of parameters for which the method was closer to the true values than each other method significantly more often in the basic experiment than in the one with the largest errors in the exogenous variables. The McNemar test with the .05 significance level was used.

The relative standings of the methods in estimating the reduced-form coefficients were not very sensitive to the introduction of errors in the exogenous variables. However, the positions of FIML relative to other methods and, to a lesser extent, of 3SLS relative to others deteriorated with the introduction of errors of measurement. The relative positions of LSRF, 2SLS and DLS improved somewhat. For several reduced-form coefficients, the McNemar test found that these changes were significant in comparing the results of the basic experiment with the one with the largest measurement errors. The results of the test are shown in Table VII-1. Significant differences between the experiments for any one pair of methods were not found for more than one-half the reduced-form coefficients. Usually they occurred far less often.

The changes in the predictions were similar to those for the reduced-form coefficients. However, even in the experiment with the largest measurement errors significantly biased medians were found rarely and, as shown in Table VII-1, changes in the relative performances of the methods between the experiment with the largest errors and the basic one were seldom statistically significant.

The experiment with errors in the exogenous variables and deliberately-introduced multicollinearity in the exogenous variables showed no surprising results. Differences between this experiment and the one with measurement error alone were similar to the differences found in Chapter VI, Section 3, between the experiment using the collinear set of data without measurement error and the basic experiment.

Comparisons of the experiments using other structures with errors in the exogenous variables with the experiments using these structures with error-free data were very similar to the comparisons of the experiments using the basic structure with and without measurement errors. The dispersions of the estimates increased somewhat but significantly biased medians were not found more frequently. The rankings of the methods were not substantially affected.[136]

Summary. Errors in the exogenous variables had no pronounced effects of the performances of the methods. Most differences in the results of the experiments could as plausibly have been ascribed to increasing the stochastic parts of the models and to changing the correlations between the stochastic elements of the separate equations as to the fact that these changes arose from adding errors to the exogenous variables.

3. Stochastic Coefficients

The next step in the study was the investigation of the effects of making the structural coefficients which generated the endogenous

[136] It was noteworthy that in the structure with very low correlations between the structural disturbances, the performances of FIML and 3SLS relative to other methods improved with the introduction of measurement errors. It will be remembered that when there were small correlations between the structural disturbances, these methods did not perform better than the consistent k-class estimators. (Cf. Chapter V, Section 5.) The correlations between the errors in the structural equations, both the usual disturbances and the errors added to the exogenous variables, were higher in the structure with small off-diagonal elements of Σ than the correlations between the structural disturbances alone; in the other structures they were lower.

variables stochastic. This happened when program-parameter 8 was set equal to unity. Program-parameter 17 was a scale-factor controlling the size of the stochastic additions to the coefficients. Let $\underline{h}$ be the value of program-parameter 17. Before generating an observation of the endogenous variables, a stochastic addition was made to the structural coefficients. These additions were normally distributed pseudo-random numbers of mean zero and variances equal to $\underline{h}^2$ percent of the true values of the structural coefficients to which each was added. The structure thus altered was used, with the exogenous variables and structural disturbances of the observation, to generate the endogenous variables for the observation. The coefficients on which the equations were normalized were not altered. The additions to the coefficients were different both for the different observations of a replication and for the different replications of an experiment.

Three experiments had stochastic additions to the coefficients of the basic structure with the scale-factor set at 2, 4, and 8. Two other experiments had stochastic additions to the coefficients of other structures. Finally, one experiment had stochastic additions to the coefficients of the basic structure and errors in the exogenous variables.[137]

Analysis of the results of these experiments was complicated by the fairly frequent occurrence of apparently singular matrices. When this difficulty arose, sets of data had to be abandoned. In the experiment with small additions to the basic structure, for example, eleven sets

[137] These are experiments I through N of the third set of experiments conducted, Appendix C, Table 3.

of data had to be dropped before fifty sets of estimates could be obtained. Other experiments left out even more sets of data. Not all the sets omitted from the experiment with small additions were found unusable in other experiments. This meant that the results of the various experiments were not fully comparable. However, despite this difficulty, the comparisons of the experiments revealed several interesting features.

Stochastic coefficients increased the sizes of the dispersions of the estimates of the structural coefficients of the basic structure very considerably. The dispersions were apt to become larger in pronounced fashion as larger stochastic elements were added. Table VII-2 records the sizes of the typical median absolute deviations in the experiments. While the increases were enormous for all methods, those of FIML were substantially greater than those of other methods. DLS was the method least affected. Other descriptive statistics increased in similarly dramatic ways.

The biases of the estimates of the structural coefficients were strongly influenced by the introduction of stochastic coefficients. Not only were the sizes of the differences of the medians of the estimates from the true values larger in the experiments with stochastic coefficients, in most cases the frequencies with which these differences were significantly different from zero also were higher for methods other than DLS. Table VII-3 shows the number of times the medians of the estimates were found to be significantly biased. It will be noted that significant biases were found slightly less frequently in the experiments with the scale-factor set at 4 and at 8 than in the one where it was set at 2.

Having stochastic coefficients sharply affected the relative standings of the methods. DLS did very well in estimating the structural coefficients in the experiments with stochastic coefficients. Its performance relative to other methods was very similar to 3SLS, which retained its good position. 2SLS also had an improved relative performance. On the other hand, FIML, together with LIML, tended to be judged the worst of the methods. These rankings held both for the totals of the ranks of all estimates, which are shown in Table VII-4, and for the ranking of the methods according to the descriptive statistics.

The sign test and the Wilcoxon signed-ranks matched-pairs test found significant differences between the dispersions of pairs of methods in the experiments with stochastic coefficients which were quite different from their findings in the basic experiment. For example, in the experiment with the largest additions to the coefficients, 3SLS had significantly smaller dispersions from the true values than FIML for all the structural coefficients according to the Wilcoxon test. In the basic experiment the test found no significant differences between the two methods. The results of the Wilcoxon test in the experiment with the largest additions to the coefficients and in the basic experiment are shown in Table VII-5.

Table VII-6 records the frequencies with which there were significant differences in the relative performances of the methods in the experiment with the smallest additions to the coefficients of the basic structure from the performances in the basic experiment. The deteriorations of FIML and the improvements of DLS and 2SLS were quite often significant.

Two further points about the effects of stochastic coefficients should be noted. First, although changes in the relative standings of the methods occurred with the introduction of stochastic coefficients, using larger stochastic additions to the coefficients did not further change the standings of the methods. Second, the performances of the methods were all fairly close to each other. Kendall's W for the totals of the ranks of all estimates of the structural coefficients varied from .10 to .12 in the experiments with stochastic coefficients. It was .11 in the basic experiment.

The experiments which used two other structures with stochastic additions to the structural coefficients also showed that their introduction had pronounced effects on the estimates of the structural coefficients. Again there were substantial increases in the dispersions of the estimates, the largest increases being for FIML and smallest for DLS. In one of the experiments, the increases in the frequencies with which the medians were significantly biased were again found. No increase occurred in the other experiment.[138] Once again the relative standings of the methods were influenced. DLS and 2SLS became better while FIML became substantially poorer.[139].

---

[138]The experiment where the increases occurred used MOD37AB which had the same structural coefficients as the basic experiment but had smaller correlations between the structural disturbances. The other experiment used MODEL37E which was of different form from the structure of the basic experiment.

[139]3SLS improved in MOD37AB relative to methods other than 2SLS and DLS. It will be recalled that in the standard run of MOD37AB, 3SLS had not had performances which were superior to the consistent k-class estimators. On the other hand, the relative performances of 3SLS were somewhat poorer in MODEL37E when there were stochastic coefficients.

The standard errors were not greatly affected by the introduction of stochastic coefficients.[140] As judged by the frequencies with which the t-ratios fell outside each of the three intervals examined, 3SLS and FIML had less reliable standard errors when stochastic coefficients were present. As larger stochastic additions to the coefficients were made, the standard errors of these methods became slightly more reliable, though their performances were still poorer than in the basic experiment.

The deteriorations of the performances of the standard errors were largely confined to two coefficients. Both occurred in the B-matrix. One was β_{21} and the other was β_{32}. The summary statistics indicated that the dispersions of the estimates of these coefficients were, if anything, relatively smaller than those of other coefficients. Deteriorations in the performance of the standard errors for other coefficients were very slight. In one of the experiments using a structure different from the structure of the basic experiment, no deterioration in the performances of the standard errors occurred. If anything, they were better when the coefficients were stochastic.[141]

---

[140] FIML and UBK had many negative "squared standard errors" for each of the coefficients estimated in the experiments with stochastic coefficients. In the one with the largest additions to the coefficients, for example, over twenty-five percent of the "squared standard errors" were negative for some coefficients. If these cases were counted among the replications where the standard errors led to misleading conclusions, the standard errors of both methods would have performed a great deal more poorly than in the basic experiment. They were not so counted because negative standard errors are an indication of inaccuracy in computing the estimates and are not otherwise a property of the estimators.

[141] This occurred in MODEL37E.

TABLE VII-2

Stochastic Coefficients - Typical Median Absolute Deviations

Structural Coefficients

Experiment	DLS	2SLS	UBK	LIML	3SLS	FIML
A	.26	.22	.23	.25	.15	.15
I	1.49	1.74	1.96	2.51	1.45	2.23
J	3.14	3.49	4.50	5.27	2.69	4.74
K	5.26	6.09	7.71	8.65	4.81	8.92

Reduced-Form Coefficients

Experiment	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	.36	.28	.28	.28	.15	.14	1.07
I	20.26	10.15	9.69	8.18	12.97	5.97	12.55
J	141.24	72.83	60.99	54.19	72.48	39.28	67.17
K	85.07	46.12	41.83	45.17	49.55	21.49	45.79

Predictions

Experiment	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	.02	.02	.02	.02	.02	.02	.03
I	.59	.37	.30	.29	.40	.22	.27
J	3.75	2.06	1.77	1.42	2.05	1.34	1.43
K	2.00	1.13	1.15	1.10	1.30	.96	1.05

Let θ_{M_i} be the median absolute deviation of the estimates of a method from the value being estimated, θ_i . Table entries record

$$1/Q \sum_{i=1}^{Q} \theta_{M_i}/\theta_i \quad .$$

Experiment A is the basic experiment;
Experiment I had program-parameter 17 set at 2;
Experiment J had program-parameter 17 set at 4;
Experiment K had program-parameter 17 set at 8.

TABLE VII-3

Stochastic Coefficients - Numbers of Significant Median Biases

Structural Coefficients

Experiment	DLS	2SLS	UBK	LIML	3SLS	FIML
A	14	4	4	0	3	0
I	9	8	8	6	9	7
J	7	7	5	5	7	4
K	4	4	3	4	7	2

Reduced-Form Coefficients

Experiment	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	13	0	0	0	2	0	0
I	11	1	0	3	2	9	2
J	0	0	0	1	0	1	0
K	0	3	2	2	0	3	2

Predictions

Experiment	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	3	0	0	0	0	0	2
I	3	0	3	4	0	5	2
J	0	0	0	0	2	0	0
K	1	3	3	2	3	2	3

Table entries record the number of values estimated for which the medians of the estimates of each method were significantly different from the true values at the .05 significance level. Fifteen structural coefficients were estimated. Twenty-one reduced-form coefficients were estimated. Nine values were predicted.

Experiment A was the basic experiment;
Experiment I had program-parameter 17 set at 2;
Experiment J had program-parameter 17 set at 4;
Experiment K had program-parameter 17 set at 8.

TABLE VII-4

Stochastic Coefficients - Rankings of the Estimators*

Structural Coefficients

Experiment	DLS	2SLS	UBK	LIML	3SLS	FIML	Kendall's W
A	3334	2751	2698	2771	2078	2118	.11
I	2174	2271	2763	3111	2248	3183	.11
J	2180	2327	2941	3049	2161	3092	.10
K	2106	2336	2941	3118	2145	3104	.12

Reduced-Form Coefficients

Experiment	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
A	5020	4251	4135	4137	3127	3211	5519	.15
I	5618	4522	3907	3833	4275	3334	3911	.10
J	5659	4644	4139	3737	4271	3121	3829	.12
K	5549	4432	3847	3851	4259	3264	4198	.10

Predictions

Experiment	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	Kendall's W
A	1899	1768	1782	1825	1560	1496	2270	.07
I	2345	1932	1671	1612	1892	1452	1696	.09
J	2360	1911	1790	1612	1826	1457	1644	.09
K	2341	1831	1676	1626	1853	1509	1764	.08

*Table entries are the sums of the ranks for the deviations of the separate estimates of each parameter from the true values of the parameters and the values of Kendall's W for each of the rows of the table.
Experiment A is the basic experiment.
Experiment I had stochastic coefficients with the scale-factor set at 2.
Experiment J had stochastic coefficients with the scale-factor set at 4.
Experiment K had stochastic coefficients with the scale-factor set at 8.

TABLE VII-5

Stochastic Coefficients - Significant Differences between Methods

Structural Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	9 (1)	12 (1)	15 (1)	5 (0)	14 (0)
2SLS	0 (7)	-	11 (3)	15 (3)	0 (0)	14 (0)
UBK	0 (7)	0 (0)	-	0 (2)	0 (0)	3 (0)
LIML	0 (4)	0 (0)	0 (0)	-	0 (0)	1 (0)
3SLS	1 (15)	6 (11)	12 (11)	14 (11)	-	15 (0)
FIML	0 (14)	0 (10)	0 (10)	1 (10)	0 (0)	-

Reduced-Form Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (10)
2SLS	18 (4)	-	0 (0)	0 (0)	0 (0)	0 (1)	3 (15)
UBK	17 (4)	9 (0)	-	0 (0)	3 (0)	0 (1)	2 (15)
LIML	14 (4)	7 (2)	3 (2)	-	6 (0)	0 (0)	1 (16)
3SLS	15 (16)	3 (15)	1 (15)	0 (15)	-	0 (1)	0 (15)
FIML	19 (15)	13 (15)	6 (15)	7 (15)	15 (2)	-	12 (15)
LSRF	15 (2)	10 (0)	1 (0)	6 (0)	5 (0)	0 (0)	-

Predictions

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (6)
2SLS	9 (1)	-	0 (0)	0 (1)	0 (0)	0 (0)	0 (7)
UBK	9 (1)	3 (0)	-	0 (2)	0 (0)	0 (0)	0 (7)
LIML	8 (1)	1 (1)	1 (1)	-	3 (0)	0 (0)	0 (7)
3SLS	9 (4)	0 (0)	0 (3)	0 (3)	-	0 (0)	0 (7)
FIML	9 (4)	2 (0)	0 (0)	0 (3)	6 (0)	-	3 (9)
LSRF	9 (0)	0 (0)	1 (0)	0 (0)	2 (0)	0 (0)	-

Entries in each row record the number of values for which the method was found to have significantly smaller dispersions than each other method in Experiment K. Numbers in parentheses are the frequencies in the basic experiment. The Wilcoxon test was used with the .05 significance level. Experiment K had stochastic coefficients with the scale-factor (program-parameter 17) set at 8.

TABLE VII-6

Stochastic Coefficients - Significant Differences between Experiment I and the Basic Experiment*

Structural Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	1	1	1	0	0
2SLS	10	-	1	1	0	0
UBK	9	8	-	0	1	0
LIML	13	8	1	-	2	0
3SLS	10	5	0	0	-	0
FIML	13	11	7	6	8	-

Reduced-Form Coefficients

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	1	4	2	2	3	17
2SLS	1	-	3	5	0	3	17
UBK	0	0	-	0	0	0	12
LIML	0	0	0	-	0	0	13
3SLS	1	5	10	10	-	3	14
FIML	0	3	4	4	0	-	6
LSRF	0	0	0	0	0	0	-

Predictions

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	2	3	2	0	2	7
2SLS	0	-	3	1	0	0	6
UBK	0	0	-	0	0	0	4
LIML	0	0	1	-	0	0	4
3SLS	0	1	3	3	-	1	5
FIML	0	0	0	0	0	-	1
LSRF	0	0	0	0	0	0	-

*Entries in each row of the table record the number of parameters for which the frequency with which one method came closer to the true values than each of the other methods was significantly smaller in Experiment I of the third set of experiments conducted (Appendix C, Table 3) than in the basic experiment. The McNemar test with the .05 significance level was used for the tests. Experiment I had stochastic coefficients with the scale-factor set at 2.

The t-ratios of DLS were more reliable when stochastic coefficients were used. They seemed to become steadily better as larger stochastic additions were made. They would not, however, have led to reliable inferences about the dispersions of the estimates of the structural coefficients about the true values in any of the experiments.[142]

The introduction of stochastic coefficients had pronounced effects on the estimates of the reduced-form coefficients. These effects were somewhat different from those found for the structural coefficients. First, while the introduction of stochastic coefficients increased the dispersions of the estimates of all methods, larger additions to the coefficients did not necessarily augment the dispersions. In particular, the dispersions of the experiment with the largest additions to the coefficients were often smaller than those of the one with the intermediate-sized coefficients, according to many of the descriptive statistics of dispersion.[143] Second, the biases of the medians of the estimates were not usually found to be significantly different from zero with much increased frequency when stochastic coefficients were used. In the

---

[142] Bias played a less pronounced role, relative to dispersion, in the estimates of DLS when the coefficients were stochastic. In the basic experiment and in the ones with stochastic coefficients, the DLS standard errors gave not inadequate indications of the dispersion of the estimates of the structural coefficients about their central tendencies. The improvement of the DLS standard errors as a representation of the dispersion of the estimates around the true values which occurred when stochastic coefficients were used may be due to the smaller importance of the biases of its estimates. An adequate representation of dispersion about the central tendencies would also be an adequate indication of the dispersion about the true values if the estimates were not seriously biased.

[143] One example where these peculiarities even affected the typical descriptive statistics is shown in Table VII-2.

experiments with the larger stochastic additions to the structural coefficients, the DLS estimates of the reduced-form coefficients had no significantly biased medians. Third, although there were quite pronounced changes in the performances of the methods relative to each other, the patterns of these changes were different from those of the structural coefficients. 3SLS became considerably poorer while the standing of FIML was not affected according to the rank-totals of all the estimates. DLS and 2SLS became poorer relative to other methods. LSRF improved substantially and was apt to surpass DLS, 2SLS and 3SLS. Changes in the frequencies with which one method had significantly smaller dispersions than another supported these impressions of changed performance. This was particularly true for LSRF. (Cf. Table VII-5.)

The findings of the McNemar test in comparing the estimates of the reduced-form coefficients in the basic experiment and the one with the smallest additions to the coefficients of the basic structure are summarized in Table VII-6. The changes found by the test were in line with the changes commented upon. The concordances of the ranks of the deviations of all the separate estimates of the reduced-form coefficients from the true values remained very low in these experiments.

Stochastic coefficients affected the reduced-form estimates of other structures in the same ways as they affected the estimates in the basic structure. Dispersions increased though not as greatly in one of the experiments as in others.[144] The frequency with which

[144] It was the experiment using MODEL37E which had the fairly small increases in the dispersions of the estimates of the reduced-form coefficients as a result of making the coefficients stochastic.

significantly biased medians occurred did not increase. The relative performances of the methods were changed somewhat by the use of stochastic coefficients, but the changes were not entirely the same as those found for the basic structure.[145]

The changes in the predictions in the experiments with stochastic coefficients from those with fixed coefficients were similar to the changes for the reduced-form coefficients. This was true for the absolute performances of the methods as well as for their relative performances.

The changes in the relative performances of the estimates of the reduced-form coefficients and the predictions were not the same as those for the structural coefficients. The standings of the methods for the estimates of the structural coefficients in the experiments with stochastic coefficients supported the conjecture that the more sophisticated methods would be more sensitive to violations of the assumptions of the standard model than the simpler methods.[146] For the reduced-form coefficients and the predictions this was not so. LIML and FIML were

[145] In MODEL37E the performances of FIML as well as of 3SLS became poorer with stochastic coefficients. DLS and 2SLS showed some improvements in standing in that model when stochastic coefficients were used. The changes brought about in MOD37AB by the use of stochastic coefficients were similar to those in the basic structure.

[146] Such an interpretation of the results would seem to indicate that 3SLS should be considered less sophisticated than LIML. It may be wondered whether the comparative robustness of 3SLS was due to its only partially applying the principles from which it can be derived in its use of 2SLS estimates at a crucial point. A fuller application of the principles, estimating Σ simultaneously with the structural coefficients, might not be so comparatively robust as 3SLS, but this was not investigated.

the structural-equation estimators which seemed to be the least sensitive to having stochastic coefficients. The reasons why changes in the results of the experiments occurred with the use of stochastic coefficients remains a mystery.

When both stochastic coefficients and errors in the exogenous variables were used, the results resembled closely those of the experiment with stochastic coefficients alone. The dispersions of the estimates of the reduced-form coefficients, however, were often judged by the descriptive statistics to be a bit smaller when both features were used than when only stochastic coefficients were introduced.

<u>Summary</u>. The introduction of stochastic coefficients had dramatic effects of the dispersions of the estimates. Even with small additions to the coefficients, the estimates of no method came within one-hundred percent of the true values of the structural coefficients half the time. The relative standings of the methods were also affected, but the changes in standing were not the same for the structural and the reduced-form coefficients. For the structural coefficients, the positions of FIML, LIML and UBK deteriorated. For the reduced-form coefficients, these three methods were the least sensitive of the structural-equation estimators to the introduction of stochastic coefficients. It was still true that no method did a great deal better than another.

4. Auto-Correlated Disturbances

Auto-correlated structural disturbances were the next subject investigated. Their use was controlled by program-parameter 9. Different values for the program-parameter introduced moderate auto-correlation in all the disturbances, or led to all the disturbances being highly auto-correlated. Table VII-7 shows the exact effects of the program-parameter.

TABLE VII-7

Types of Auto-Correlated Disturbances Investigated*

Program-parameter 9	Effect
0	$u_{it} = v_{it}$.
1	$u_{it} = v_{it}, \quad i \neq 1, \text{ all } t;$
	$u_{it} = .6v_{it} + .4u_{1,t-1}, \quad i = 1, \ t \neq 1;$
	$u_{it} = v_{it}, \quad i = 1, \ t = 1.$
2	$u_{it} = .5v_{it} + .5u_{i,t-1}, \quad t \neq 1, \text{ all } i;$
	$u_{it} = v_{it}, \quad t = 1, \text{ all } i.$
3	$u_{it} = .2v_{it} + .8u_{i,t-1}, \quad t \neq 1, \text{ all } i;$
	$u_{it} = v_{it}, \quad t = 1, \text{ all } i.$

*Symbols: u_{it} - the value of the i^{th} structural disturbance at the t^{th} observation in the set of data used.

v_{it} - the value of the i^{th} disturbance used when there were no auto-correlated disturbances.

Experiments were conducted with the basic structure using all the types of auto-correlated disturbances provided for by program-parameter 9. Experiments were also conducted investigating the effects of having all the disturbances auto-correlated in other structures.[147]

Introducing auto-correlated disturbances did not affect greatly the results of the experiments using the basic structure. A few differences between the experiments using the basic structure with auto-correlated disturbances and the basic experiment were found. These were not exactly the same when only one of the disturbances was auto-correlated and when they all were.

The dispersions of the estimates of all methods were usually larger when one of the disturbances was auto-correlated than in the basic experiment. The full-model methods deteriorated slightly relative to the single-equation ones, particularly DLS, with respect to the number of times they were closer to the true values of the structural and the reduced-form coefficients. On the whole, these methods also deteriorated in the number of coefficients for which they had smaller descriptive statistics than other methods. LSRF tended to perform better when one of the disturbances was auto-correlated in estimating the reduced-form and in making predictions. The standard errors were slightly better than in the basic experiment, but the changes were not statistically significant.

The dispersions of the estimates were smaller when all the disturbances were auto-correlated than in the basic experiment, especially

[147] These are experiments S through X of the third set of experiments conducted, Appendix C, Table 3.

when the auto-correlation was relatively high. In these and in the basic experiment, the standard errors performed comparably. Nor were the comparative standings of the methods for any of the parameters estimated much different, though DLS tended to come closer to the true values of the structural coefficients than other methods slightly more frequently when there were auto-correlated disturbances. The improvement of DLS also occurred for the reduced-form coefficients, but here it was mainly at the expense of LSRF. The medians of the DLS estimates of the structural coefficients had significantly biased medians less often when the disturbances were auto-correlated.

Auto-correlated disturbances led to only a scattering of significant changes in the results. For the structural coefficients the McNemar test found significant differences most often for the standings of DLS relative to other methods, particularly other single-equation methods. However, the DLS estimates were significantly improved relative to the other methods for not more than one-fifth of the structural coefficients. For the reduced-form coefficients, it was the improvements of DLS and the deteriorations of LSRF which were most often significant. The standings of DLS relative to LSRF were significantly better for eight reduced-form coefficients when there were highly auto-correlated disturbances than in the basic experiment. FIML came closer to the true values than 2SLS and UBK significantly more often when the disturbances were highly auto-correlated for three of the reduced-form coefficients.

The differences between the experiments did not substantially affect the over-all rankings of the methods. The concordances of the ranks of the rank-totals in the separate experiments, both for the ranks of the individual estimates and for the ranks of the descriptive statistics of dispersion, were high for the experiments with auto-correlated disturbances and the basic experiment. Reversals of order from that of the basic experiment occurred almost entirely between methods which were very similar to each other in the basic experiment.

The experiments using structures other than the basic one with auto-correlated disturbances agreed with the experiments using the basic structure about the effects of auto-correlated disturbances. The dispersions of the estimates decreased a bit when all the disturbances were auto-correlated. Pronounced changes in the relative standings of the methods did not occur, but DLS improved slightly in estimating the structural and the reduced-form coefficients. The LSRF estimates of the reduced-form deteriorated slightly in all experiments but one.[148]

Most of the results of using auto-correlated disturbances are not very surprising. None of the methods is designed to take advantage of the auto-correlation, so their relative performances might not be expected to change with the introduction of auto-correlated disturbances. Since auto-correlation does not bias the estimates in the single-equation case, it might not be expected to affect the biases in the simultaneous-equation model. However, the failure of the auto-correlated disturbances to affect the performances of the standard errors does seem surprising. In the single-equation model, auto-correlated disturbances make the usual

[148]The exception was an experiment using MODEL46 with highly auto-correlated disturbances. In that experiment, the improvements of DLS were primarily at the expense of other methods relative to which LSRF also improved.

formulae for the standard errors of the estimates inappropriate.[149] It might be supposed that it would weaken the standard errors in the simultaneous-equation case. This did not happen. The standard errors were as reliable when there were auto-correlated disturbances as when there were not.

> Summary. Introduction of auto-correlated disturbances did not affect the performances of the methods greatly. There were, however, small improvements in the relative standings of DLS in the experiments with auto-correlated disturbances. The standard errors were no less reliable as the result of having auto-correlated disturbances.

5. Heteroskedastic Disturbances

Heteroskedastic disturbances were introduced in two ways. First, the disturbances were made multiplicative factors rather than additive ones so that the variances of the disturbances at each observation depended on the size of the endogenous variables. Second, the variances of the disturbances were made to depend on the sizes of the exogenous variables without direct regard for the sizes of the endogenous variables. Program-parameter 7 introduced the first type of heteroskedasticity; program-parameter 12 controlled the second type. Table VII-8 shows the effects of different values for these program-parameters.

---

[149] See, for example, J. Johnston, Econometric Methods, (New York: McGraw-Hill, 1963), pp. 177ff.

TABLE VII-8

Heteroskedastic Disturbances - Types Investigated*

Program-Parameter	Value	Effect
7	1	$y_{it} = r_{it}\left[1 + \left(\sum_{t=1}^{T} \frac{r_{it}^2}{T}\right)^{-1/2} w_{it}\right]$, all i,t.
7	2	$y_{it} = r_{it}\left[1 + \left(\sum_{t=1}^{T} \frac{r_{it}^2}{T}\right)^{-1/2} w_{it}\right] + q_{it}$, all i,t.
12	1	$u_{it} = v_{it}\left[\frac{T \sum_k z_{kt}}{\sum \sum z_{kt}}\right]$; all t,i = 1.
		$u_{it} = v_{it}$; all t,i $\neq$ 1.
12	2	$u_{it} = v_{it}\left[\frac{T \sum_k z_{kt}}{\sum_k \sum_t z_{kt}}\right]$; all t,i.

*Symbols:

- y_{it} - the endogenous variables actually used with this value of the program-parameter $(i = 1,\ldots,G;\ t = 1,\ldots,T)$.
- r_{it} - the endogenous variables before any disturbances are introduced.
- w_{it} - the reduced-form disturbances usually used in the experiments (i.e., the disturbances if program-parameter 9 were set at zero).
- q_{it} - pseudo-random deviates distributed as w_{it}.
- u_{it} - the structural disturbances actually used.
- v_{it} - the structural disturbances usually used in the experiments (i.e., the disturbances if program-parameter 12 were set at zero).
- z_{kt} - the exogenous variables $(k = 1,\ldots,K)$.

Experiments were conducted with the basic experiment and each of the types of heteroskedastic errors allowed for by the programs. Experiments were also conducted using other structures with heteroskedastic disturbances. One experiment investigated having multicollinear data and heteroskedastic disturbances.[150]

The estimates of the structural and the reduced-form coefficients and of the predictions were not affected by the use of heteroskedastic disturbances. This was true for the biases and the dispersions of the estimates. It was also true for the relative standings of the methods. In comparing some of the experiments with heteroskedastic errors with the basic experiment, the McNemar test found at most one or two parameters for which there were significant differences between the experiments in the number of times one method came closer to the true values than another. The reliability of the t-ratios for making inferences about the dispersions of the estimates about the true values did not seem to be affected by the use of heteroskedastic disturbances.

The results of this section are surprising for the same reasons that those of the section using auto-correlated disturbances might not have been entirely expected. None of the methods takes advantage of possible heteroskedasticity. The appropriateness of the principles from which the methods were derived, relative to each other, do not seem to be changed by the presence of heteroskedastic disturbances, though the principles of all methods seem less suitable. The failure of the

[150]These are experiments O through P and Y through AB of the third set of experiments conducted, Appendix C, Table 3.

performances of the standard errors to be affected by heteroskedastic disturbances is surprising. In the single-equation model, heteroskedastic errors do make the usual standard errors inappropriate.[151]

Summary. The performance of the estimators seemed to be insensitive to the use of heteroskedastic disturbances. This was true for the standard errors as well as for the parameters of the structures which were estimated.

6. Combinations of Features

The effects of generating data with more than one of the special features which have been investigated in this chapter and Chapter VI were investigated next. All but one of the experiments studying several special features together had auto-correlated disturbances. The experiments concentrated on the effects of also having lagged-endogenous variables, trends, auto-correlated exogenous variables or heteroskedastic disturbances.[152]

The results of the experiments in this section will not be discussed individually. Most of them were similar to the results of experiments which have been described earlier. Table VII-9 summarizes the results. Often no entry is made in the table even though the results of an experiment were substantially different from those of the basic

[151] See, for example, Arthur S. Goldberger, Econometric Theory (New York: Wiley, 1964), pp. 231ff.

[152] These are experiments AC through AW of the third set of experiments conducted, Appendix C, Table 3.

experiment. In these cases there was an experiment which had one or more of the program-parameters set at the same values as in the experiment included in the table whose results were similar to those of the experiment in Table VII-9.

Several features of Table VII-9 stand out:

1. The medians of the estimates of the consistent methods were not significantly biased more often than might be expected from previous results. The medians of DLS were significantly biased less frequently in several of the experiments than in earlier ones.

2. The rankings of the methods according to the sums of the ranks of the deviations from the true values of the individual estimates of all the parameters were usually the same as in the basic experiment. The values of Kendall's W were often smaller than in the earlier experiments. When these small values of W occurred, it was not unusual to find that the rankings of the methods according to some of the descriptive statistics of dispersion were different from those of the basic experiment. In these cases the consistent single-equation methods surpassed the full-model estimators or DLS did better than other k-class estimators.

3. In one experiment (Experiment AV which used MODEL 46) the rankings of the methods, even for the totals of the ranks of all estimates, were different from those of the basic

TABLE VII-9

Combinations of Aspects - Summaries of the Results

Notes

Table entries summarize the differences in the results of the experiments from those of other experiments having the special features singly. The terms in the column headings and in the body of the table signify:

Bias - indicates whether the frequency of finding significantly biased medians was different from other experiments.

Rankings - indicates whether the over-all ranking of methods differed from that of the basic experiment.

Standard Errors - shows whether the combinations of factors seemed to affect the usefulness of the t-ratios.*

W - Kendall's W.

R.F. - Reduced-Form Coefficients.

P. - Predictions.

-- No marked differences from the results of experiments having some of the same program-parameters different from zero.

* Described as "much worse" when poorer performances of the t-ratios led to frequent rejection of hypotheses which did well in the basic experiment.

TABLE VII-9 (Noted Continued)

The program-parameters for the experiments were:

Experiment	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
AC	MOD37A2	20	0	0	0	0	1	0	2	0	0	0	2	0	0	2
AD	MOD37A2	20	0	0	2	0	0	0	2	0	0	0	2	0	0	2
AE	MOD37A2	20	0	0	2	0	0	0	0	0	0	2	2	0	0	2
AF	MOD37A2	20	0	0	3	0	0	0	2	0	0	0	2	0	0	2
AG	MOD37A2	20	0	0	3	0	0	0	3	0	0	0	2	0	0	2
AH	MOD37A2	20	0	0	7	0	0	0	3	0	0	0	2	0	0	2
AI	MODEL46	20	0	0	7	0	0	0	3	0	0	0	2	0	0	2
AJ	MOD37A2	20	0	0	7	0	0	0	3	0	0	2	2	0	0	2
AK	MOD37AB	20	0	0	7	0	0	0	3	0	0	2	2	0	0	2
AL	MOD37A2	20	0	2	0	0	0	0	2	0	0	0	2	0	0	2
AM	MOD37A2	20	0	2	3	0	0	0	3	0	0	0	2	0	0	2
AN	MOD37A2	20	0	2	5	0	0	0	3	0	0	0	2	0	0	2
AO	MOD37A2	20	0	2	7	0	0	0	3	0	0	0	2	0	0	2
AP	MOD37A2	20	0	2	7	0	0	0	3	0	0	2	2	0	0	2
AQ	MOD37A4	20	0	3	7	0	0	0	3	0	0	2	2	0	0	2
AR	MOD37AE	20	0	3	7	0	0	0	3	0	0	2	2	0	0	2
AS	MOD37AB	20	0	2	7	0	0	0	3	0	0	2	2	0	0	2
AT	MOD37A2	20	1	2	3	5	0	0	2	0	0	2	2	0	0	2
AU	MOD37A2	20	0	2	3	5	0	0	2	0	0	0	2	0	0	2
AV	MODEL46	20	1	2	3	5	0	0	2	0	0	0	2	0	0	2
AW	MOD37A2	20	0	2	7	6	0	0	3	0	0	2	2	0	0	2

The program-parameters determined the types of data as follows:

Program-parameter	Effect
1	Structure used
2	Number of observations
3	Errors in the exogenous variables
4	Lagged-endogenous variables
5	Trends in the exogenous variables of auto-correlated exogenous variables
6	Multicollinearity in the exogenous variables
7	Heteroskedastic errors--proportional to the endogenous variables
9	Auto-correlated disturbances
12	Heteroskedastic disturbances
13	Scale-factor for the structural disturbances
16	Scale-factor for the errors in exogenous variables

The exact effects of the different values for the program-parameters were discussed when their effects were discussed separately.

TABLE VII-9

Combinations of Aspects - Summaries of the Results

Experiment	Structural Coefficients: Bias	Structural Coefficients: Rankings	Standard Errors	Reduced-form Coefficient and Predictions: Bias	Reduced-form Coefficient and Predictions: Rankings
AC	--	-- W slightly smaller	Worse	DLS less often for R.F.	--
AD	--	--	--	--	-- W slightly larger for P
AE	--	--	--	--	-- W slightly smaller for R.F.
AF	--	--	Much worse	DLS less often for R.F.	W slightly higher
AG	DLS less often	--	Much worse	DLS less often for R.F.	--
AH	--	-- W slightly smaller	Much worse	--	-- W slightly higher
AI	--	--	Much worse	--	-- W smaller for R.F.
AJ	--	-- W smaller	Much worse	--	-- W smaller

TABLE VII-9 (Cont.)

Experiment	Structural Coefficients: Bias	Structural Coefficients: Rankings	Standard Errors	Reduced-form Coefficient and Predictions: Bias	Reduced-form Coefficient and Predictions: Rankings
AK	DLS less often	-- Full-model methods slightly better than in standard run of MOD37AB	--	--	FIML worst methods in P.
AL	--	-- W smaller	A bit worse	DLS less often for R.F.	-- W smaller. FIML poor in P.
AM	DLS less often	-- W smaller	Much worse	DLS less often for R.F.	-- W smaller for R.F.
AN	DLS less often	-- W smaller	Much worse	DLS less often for R.F.	-- W smaller for R.F.
AO	--	-- W smaller	Much worse	--	-- W smaller.
AP	--	--	Much worse	--	--
AQ	--	--	-- (but poor)	--	--
AR	--	-- Full-model methods better than in standard run of MOD37AE	-- (but poor)	--	--

TABLE VII-9 (Cont.)

Experiment	Structural Coefficients		Standard Errors	Reduced-form Coefficient and Predictions	
	Bias	Rankings		Bias	Rankings
AS	DLS less often	-- Full-model method better than in standard run of MOD37AB	-- (but poor)	--	-- (Improvement of full-methods from standard run of MOD37AB)
AT	--	--	--	--	-- W smaller for R.F.
AU	--	--	--	--	--
AV	--	DLS better than other single-equation methods	Much worse except for 3SLS and FIML	--	DLS and LSRF better than single-equation methods. W slightly higher than usual.
AW	-- But frequent biases for 2SLS, UBK and 3SLS	--	-- (but poor)	--	--

experiment. The changes in the rankings also occurred for the descriptive statistics. The results were different from other experiments using the same structure. In an experiment[153] using the basic structure with the values of other program-parameters the same as in Experiment AV, the rankings of the methods were not upset although the values of Kendall's W were smaller.

4. The standard errors of the consistent methods performed more poorly in many of the experiments than in earlier ones. The criterion for this conclusion was the frequencies with which the t-ratios fell outside the three intervals examined. In these experiments, the standard errors would not have led to reliable inferences about the dispersions of the estimates and one could frequently have rejected the hypotheses about the distributions of the t-ratios which did well in the basic experiment.

5. It did not seem possible to find any pronounced systematic pattern for the occurrence of these differences from other experiments. The one exception was that the standard errors seemed to perform very poorly whenever there were lagged-endogenous variables and <u>highly</u> auto-correlated disturbances.

---

[153] This experiment was conducted after all the other experiments in this study. Due to a sudden exclusion from the computer, it was not incorporated in the tables of Appendix D and is not included in Appendix C, Table 3.

Summary. Few pronounced differences from earlier experiments were found among the results of the experiments investigating combinations of aspects. Such differences as were found usually indicated that the performances of the methods became more similar to each other when several of the special features generating the data were used. In many of the experiments the standard errors of the consistent methods gave very poor indications of the variability of the estimates.

7. Conclusion

The violations of the assumptions of the standard simultaneous-equation model did not have dramatic effects on the performances of the estimators. With one major exception, the differences between the experiments of this chapter and the basic experiment were less striking than the differences between the experiments of Chapters V and VI. The exception was the use of stochastic coefficients which led to the estimates having very large dispersions, but it should be remembered that the additional stochastic elements introduced into these experiments were quite large. Stochastic coefficients also led to changes in the relative standings of the methods. LSRF improved its relative performance significantly for estimating the reduced-form coefficients and for making predictions. DLS improved for the structural coefficients.

When small errors were added to the exogenous variables, the medians of the estimates of the consistent methods were not biased much more frequently than in the basic experiment. The results of the experiments with errors in the exogenous variables were similar to the findings of Ladd's study.[154]

In all the experiments of this chapter, the differences between the methods were not very great. No method performed either a great deal better or worse than another. Most of the time the weak ranking found in Chapters IV, V and VI continued to hold.

The performances of the standard errors were similar to those of the basic experiment in most experiments. However, when auto-correlated disturbances were combined with some special features investigated in Chapter VI, the standard errors were found to be a great deal less reliable.

---

[154] G. W. Ladd, "Effects of Shocks and Errors in Estimation: An Empirical Comparison," Journal of Farm Economics, vol. 38 (1956), pp. 485-495.

CHAPTER VIII

Misspecification of the Structural Equations

1. Introduction

The final problem examined by the Monte Carlo experiments of this study was the effect on the performances of the estimators of misspecifying the forms of the structural equations. Four types of misspecification were investigated:

A. Zeroes were specified in the structure to be estimated for coefficients which were non-zero in the structure generating the data.

B. Zeroes were not specified for all the coefficients in the structure to be estimated which were zero in the structure generating the data. This is not, strictly speaking, a misspecification of the structure. It is merely an attempt to estimate the value of a coefficient whose true value happens to be zero. It was, however, convenient to investigate it in this chapter.

C. Zeroes were specified in the structure to be estimated for coefficients which were non-zero in generating the data and coefficients which were zero in generating the data were not so specified in the structure estimated.

D. One equation was omitted from the structure to be estimated.

Program-parameter 10 selected the misspecification to be introduced. The effects of different values of this program-parameter will be discussed when the results of experiments using these values are presented. However, it may help in appreciating the misspecifications to present them in a single table. This is done in Table VIII-1 for the basic structure. The symbol X represents a non-zero coefficient. A number in parentheses above an element of the matrices indicates that with this value for program-parameter 10, the coefficient was not specified to be zero in the structure estimated if it was zero in the basic structure. If the coefficient was not zero in the basic structure, the number above it indicates that it was specified to be zero in the structure estimated.

The experiments conducted have three major short-comings. First, the original experiment-design for this part of the study was too ambitious and could not be completed. The result was an imbalance between the various features studied. Some experiments which might have been of considerable interest were not conducted. On the other hand, experiments of possibly minor interest were completed. Second, only a very limited selection of the possible types of misspecification were studied and for each type not many different cases were investigated. Third, it was discovered after several experiments were finished that there was a mistake in the programs which led to incorrect analysis of the predictions for most of the misspecified runs. The mistake was that the true values of the variables to be predicted, with which the predictions in each replication were compared, were calculated from a misspecified structure rather than from the true one which had generated the data. The mistake affected none of the experiments in which the structural equations were correctly specified, which were reported in earlier chapters. Appendix D has no tables for the predictions of the misspecified experiments since so many of the entries would be misleading. Since the predictions and the correct true values were available, some evaluation of the predictions was possible, but the full analysis, described in Chapter III, was not conducted correctly. It did not seem worthwhile to use computer time to obtain a full and correct analysis of the predictions since this would have required recomputing the experiments instead of conducting other ones of seemingly more interest.

TABLE VIII-1

Misspecifications used with the Basic Structure

$$B - \begin{bmatrix} 1 & \overset{(7,8,9)}{X} & X \\ \overset{(8,9)}{X} & 1 & \overset{(3,6,8,9)}{0} \\ 0 & X & 1 \end{bmatrix}$$

$$\Gamma - \begin{bmatrix} X & X & \overset{(12)}{0} & 0 & X & 0 & 0 & \overset{(2,6)}{0} \\ X & \overset{(4,6)}{0} & X & \overset{(11)}{0} & X & 0 & \overset{(1,9)}{X} & 0 \\ X & 0 & X & X & 0 & X & 0 & \overset{(2,6)}{0} \end{bmatrix}$$

2. Omission of an Exogenous Variable

The first misspecification studied was the omission of an exogenous variable from the structure to be estimated. Setting program-parameter 10 equal to unity deleted the last exogenous variable. This variable was also omitted when the reduced form was estimated by LSRF.

Three experiments investigated the effect of the omission of an exogenous variable in the basic structure and two other structures. In two of the structures, the only non-zero structural coefficient for the omitted variable was very small. The one non-zero coefficient of the omitted variable in the third structure was of moderate size.[155]

Structural Coefficients. Leaving out a variable did not have pronounced effects on the estimates of the structural coefficients. The dispersions of the estimates were of about the same size as when no misspecification was present. Significantly biased medians were found with slightly greater frequency for the full-model methods in the basic structure, for 3SLS in the second structure[156] and for all consistent methods in the third structure.[157] Very little change in the comparative standings of the methods occurred. However, slight deteriorations from

[155] The structures used were MOD37A2, MOD37A4 and MODEL48. For the first and third equations of MOD37A2 and MOD37A4, the misspecification led to UBK being the same as 2SLS. This was also true for the second equation of MODEL48. MOD37A4 differed from it in the values of the coefficients. MODEL48 was of different form; it had more equations and exogenous variables than the basic structure. See Tables IV-1, V-6 and V-14.

[156] MOD37A4.

[157] MODEL 48.

the correctly-specified experiments in the standings of the full-model methods, especially FIML, relative to the consistent k-class estimators were found. In comparing the experiment using the basic structure with the basic experiment, the McNemar test found two or three coefficients for which this change was significant at the .05 level.

Standard Errors. In two of the structures,[158] omitting a variable had little effect on the performances of the standard errors. In the third structure[159] the standard errors were quite a bit more reliable when the misspecification was made.

Estimates of Σ. The dispersions of the estimates of Σ were not much affected by the omission of an exogenous variable. The rankings of the methods were little affected in the basic structure. In the second structure,[160] FIML was poorer when the variable was left out. In the third structure, both FIML and 3SLS performed poorly. The omission of a variable from the structure in effect added another element to the disturbances of the equations in which the variable should have been included. The size of the addition was smallest in the basic structure and largest in the third one. The change in the performances of the methods was as might have been expected from the rankings of the methods in earlier experiments and from the fact that in those experiments FIML and 3SLS were the methods which were least biased towards zero.

[158] MOD37A2 and MOD37A4.

[159] MODEL48.

[160] MOD37A4.

Reduced-Form Coefficients. In one experiment[161] the biases of the medians of the estimates of the reduced-form coefficients were significantly different from zero much more frequently when a variable was omitted, especially in the case of LSRF. In the other experiments, this was not true. The dispersions of the estimates were not affected. The over-all rankings of the methods were little changed. The differences between the experiment using the basic structure with the misspecification and the basic experiment were rarely found to be significant by the McNemar test.

Predictions. The predictions had the most surprising results.[162] Although the dispersions of the estimates were not much affected, the medians of all methods were significantly biased more often in all the structures when an exogenous variable was omitted. LSRF and DLS improved their relative positions, while FIML and 2SLS became poorer. The changes in the positions of DLS and FIML in the experiment using the basic structure from the positions in the basic experiment were found by the McNemar test to be significant for about one-third of the values predicted.

Program-parameter 10 was set at unity for a group of experiments using the basic structure and various features of the generation of the data which were examined in Chapters VI and VII. The first five had

[161] This was the experiment using MODEL48.

[162] The predictions were analyzed correctly when program-parameter 10 was set equal to unity.

time-structures in the predetermined variables, including lagged-endogenous variables. The others had multicollinearity. When lagged-endogenous variables were used, the discarded predetermined variable was one of the lagged-endogenous ones.[163]

Table VIII-2 summarizes the findings of the experiments. Entries in the table record differences from previous experiments having some of the program-parameters set at values used in these experiments. Several features of the table deserve comment.

1. In several of the experiments, especially those with lagged-endogenous variables, the rankings of the methods were changed. For the structural and the reduced-form coefficients, LIML improved greatly. UBK was frequently better than in earlier experiments. For the predictions it was LSRF and DLS which improved.

2. Trends in the exogenous variables did not seem to affect the rankings of the methods even when the same trend was added to the excluded variable as to all the others (experiment AE).

3. In almost all the experiments, the median biases of the consistent methods were signifcant. When lagged-endogenous variables were used and in one of the experiments with multicollinearity the medians of all the predictions were significantly biased.

[163] These are experiments AC-AJ of the fourth set of experiments, Appendix C, Table 4.

Notes to Table VIII-2

Bias - frequency of finding significant biases for methods other than DLS

Misspec. - behaved as in the experiment using MOD37A2 and program-parameter 10 set at unity and no other special features

Rankings - changes in the standings of the methods in the totals of the ranks of all estimates

-- - similar to the results of previous experiments with some of the same program-parameters not set to zero

* This change in the standings is similar to what was found when program-parameter 6 was set at 7 without misspecification (experiment G of the second set of experiments).

The values of the program-parameters in these experiments were:

Experiment	Program-parameter 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
AC	MOD37A2	20	0	2	0	0	0	0	0	1	0	0	2	0	0
AD	MOD37A2	20	0	0	3	0	0	0	0	1	0	0	2	0	0
AE	MOD37A2	20	0	0	5	0	0	0	0	1	0	0	2	0	0
AF	MOD37A2	20	0	2	2	0	0	0	1	1	0	0	2	0	0
AG	MOD37A2	20	0	0	1	0	0	0	0	1	0	2	2	0	0
AH	MOD37A2	20	0	0	0	3	0	0	0	1	0	0	2	0	0
AI	MOD37A2	20	0	0	0	5	0	0	0	1	0	0	2	0	0
AJ	MOD37A2	20	0	0	0	7	0	0	0	1	0	0	2	0	0

It should be noted that the variance of the omitted lagged-endogenous variable was larger than the variance of the exogenous variable it replaced. This may account for the results. Why the ranking of LIML and UBK, but not of 2SLS, improved, remains a mystery. The UBK value of k is less than or equal to that of 2SLS while LIML has a higher k-value. Why the omission of a variable should have affected the results in the experiments with no other special features is also a mystery. Omission of a variable, in effect, puts into the disturbance term one influence which might have been accounted for systematically by the structure. One would not have expected this to change the performances of the methods. Such changes as were found may well be a further demonstration of the sensitivity of the methods to the exact sets of data used rather than a specific effect to be expected when a variable is not included in a structure when it should be.

<u>Summary</u>. The omission of a variable from the structure to be estimated led to several interesting findings. While the dispersions of the methods were but little affected, the frequency of finding significantly biased medians was usually higher when the variable was omitted. For several experiments, the rankings of the methods were affected. This was particularly true when a lagged-endogenous variable was dropped. LSRF became better for the predictions <u>but</u> <u>not</u> the reduced-forms. LIML improved most for the structural and reduced-form coefficients. As in other experiments, however, differences between the methods were not very great.

3. Inclusion of an Extra Exogenous Variable

An extraneous exogenous variable was included in the structure to be estimated when program-parameter 10 was set equal to 2. It was specified that the coefficient for this variable be zero in the second structural equation. In the other two equations, the coefficients were to be estimated. The extra variable, $Z_{K+1,t}$, was formed from random numbers, V_t , distributed as were the other exogenous variables and Z_{Kt}, the last exogenous variable generating the data:

$$Z_{K+1,t} = .5V_t + .5Z_{Kt} \qquad t = 1,\ldots,T.$$

Program-parameter 10 was set at 2 in experiments studying the basic structure and one other structure.[164] A third experiment used the basic structure with deliberately-introduced multicollinearity in the exogenous variables.[165]

By and large, the inclusion of an extra variable had little effect on the estimates. As might be expected from having more coefficients to estimate, the dispersions of the estimates of the structural and reduced-form coefficients were slightly larger.[166] Very little change in the

---

[164] These are experiments E and F of the fourth set of experiments conducted (Appendix C, Table 4). The other structure used was MOD37A1 which differed from MOD37A2, the basic structure, in the values of the structural coefficients.

[165] This is experiment AK of Appendix C, Table 4.

[166] The coefficients whose true values were zero were not included in the formation of the statistics of Appendix D. This was done mainly because these coefficients could not be included in the formation of the typical descriptive statistics since division by the true values was involved in forming these statistics. See Chapter III, Section 7.

rankings of the methods occurred and the McNemar test found almost no coefficients for which the relative performances of the methods were significantly changed by having the extra variable in the basic structure. The rankings of the methods for estimating the elements of Σ were not changed, but the ranking of the methods was stronger than in the basic experiment. However, the McNemar test again found almost no significant differences in the performances of the methods for estimating Σ.

The standard errors, especially those of the coefficients whose true values were zero were a bit poorer as judged by the number of times the t-ratios exceeded two and the 95% confidence limits of the t-distribution with degrees of freedom based on the number of coefficients to be estimated in each equation. When the degrees of freedom were based on the total number of coefficients to be estimated, the deterioration in the performances of the t-ratios was not evident.

In estimating the coefficients whose true values were zero, the ranking of the methods was usually the same as for the non-zero coefficients. However, DLS surpassed 2SLS and UBK in estimating the zero reduced-form coefficients in the basic structure when multicollinearity was introduced.

> Summary. The inclusion of an extraneous variable in the structure estimated had little effect on the performances of the estimators.

4. Failure to Specify All Zero Coefficients

The next group of experiments investigated the effects of failing to specify that coefficients which were zero in generating the data be zero in estimating the structure. That is, an attempt was made to estimate one or more coefficients whose true values were zero. This will be referred to as under-specification of the structure. Four values of program-parameter 10, which are shown with their effects in Table VIII-3, introduced this feature into the experiments. Table VIII-4 presents the experiments conducted with these values for program-parameter 10. Quite a large number of experiments were carried out because some of the results were quite surprising and were not believed at first to be very representative of the effects to be attributed to the failures of specification.

TABLE VIII-3

Failures to Specify Zero Coefficients - Types Studied

Values of Program-Parameter 10	Effect
3	The first zero element of B was to be estimated. In the basic structure, this coefficient was β_{23}.
4	The first zero element of the second row of Γ was estimated. In the basic structure this was γ_{22}.
11	The second zero element of the second row of Γ was to be estimated. In the basic structure this was γ_{24}.
12	The first zero element of the first row of Γ was to be estimated. In the basic structure this was γ_{13}.

TABLE VIII-4

Failure to Specify Zero Coefficients: Experiments Conducted

Experiment*	Structure	Program Parameter 10	Other Special Features
G	MOD37A2	3	None
I	MOD37A2	4	None
W	MOD37A2	11	None
X	MOD37A2	12	None
J	MOD37A2	4	Different exogenous data set
K	MOD37A2	4	Smaller disturbances
L	MOD37A2	4	Seventy observations
H	MOD37A1	3	None
M	MOD37A1	4	None
N	MOD37A3	4	None
O	MOD37AB	4	None
P	MODEL37C	4	None
Q	MODEL37D	4	None

* Experiment letters are those assigned in the fourth set of experiments conducted, Appendix C, Table 4.

TABLE VIII-5

Under-Specified Experiment: Rankings of the Estimates of the Structural Equations of MOD37A2**

Fully-specified Equations

Program-Parameter 10	DLS	2SLS	UBK	LIML	3SLS	FIML
0*	2159	1810	1791	1846	1416	1478
3	2113	1687	1653	1705	1635	1705
4	2007	1641	1655	1688	1804	1705
11	2165	1809	1798	1858	1410	1450
12	2224	1798	1723	1734	1545	1476

Under-Specified Equation

Program-Parameter 10	DLS	2SLS	UBK	LIML	3SLS	FIML
0*	1175	941	907	925	662	640
3	1328	847	847***	816	709	705
4	1182	873	873	739	711	872
11	1156	939	939	932	654	640
12	979	854	854	968	743	852

The Zero Coefficient

Program-Parameter 10	DLS	2SLS	UBK	LIML	3SLS	FIML
3	282	164	164	152	140	148
4	236	169	169	150	152	174
11	213	195	195	203	126	118
12	204	166	166	188	149	177

* The rankings of the first and third equations of the basic experiment are shown with the fully-specified equations of other experiments. The rankings of the second equation appear in the second division. Note that these do not correspond to the fully- and under-specified equations when program-parameter 10 is set at 12.

** Table entries are the sums of the ranks for each of the separate estimates of the coefficients in the equations.

*** In the under-specified equations, 2SLS and UBK were identical.

TABLE VIII-6

Under-Specified Experiments: Significant Differences from Basic Experiment*

Fully-Specified Equations**

Program-parameter 10-3

	DLS	2SLS	UBK	LIML	3SLS	FIML
3SLS	2	5	3	4	-	0
FIML	1	4	3	3	1	-

Program-parameter 11-4

	DLS	2SLS	UBK	LIML	3SLS	FIML
3SLS	3	5	4	5	-	0
FIML	1	4	4	3	0	-

Program-parameter 10-11***

	DLS	2SLS	UBK	LIML	3SLS	FIML
3SLS	0	0	0	0	-	0
FIML	0	0	0	0	0	-

Program-parameter 10-12

	DLS	2SLS	UBK	LIML	3SLS	FIML
3SLS	3	1	1	1	-	1
FIML	1	1	2	1	0	-

Under-Specified Equations

Program-parameter 10-3

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	2	2	1	1	0
2SLS	0	-	α	1	0	0
UBK	0	α	-	1	0	0
LIML	0	0	0	-	0	0
3SLS	0	1	0	0	-	0
FIML	0	1	0	0	0	-

Program-parameter 10-4

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	1	1	1	0	0
2SLS	0	-	α	1	0	0
UBK	0	α	-	2	0	0
LIML	0	0	0	-	0	0
3SLS	0	0	0	0	-	0
FIML	3	2	2	5	0	-

Program-parameter 10-11

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	0	0	0	0	0
2SLS	0	-	α	0	0	0
UBK	0	α	-	0	0	0
LIML	0	0	1	-	0	0
3SLS	0	0	0	0	-	0
FIML	0	0	0	0	0	-

Program-parameter 10-12

	DLS	2SLS	UBK	LIML	3SLS	FIML
DLS	-	0	0	0	0	0
2SLS	0	-	α	1	0	0
UBK	0	α	-	2	0	0
LIML	0	0	0	-	0	0
3SLS	0	0	1	0	-	0
FIML	0	2	3	1	0	-

* Entries in each row record the number of coefficients for which a method did significantly better in the basic experiment than in the ones with program-parameter 10 not equal to zero in pairwise comparisons with each of the other estimators. The McNemar test was used and the .05 significance level.

** With only one exception there were no significant differences in other rows of the table which are therefore omitted. See next note for exception.

*** When program-parameter 10 was set at 11, FIML improved significantly relative to 2SLS and UBK for one coefficient.

α UBK and 2SLS were identical in the under-specified equation.

TABLE VIII-7

Under-Specified Experiments: Very Bad FIML Estimates*

Experiment I - Program-Parameter 10-4

True Values	.54	62.00	0	.70	.96	.06
Experiment						
I	- 41.20	10571.93	54.26	73.21	73.99	6.15
A	.52	63.17	**	.80	.93	.07
I	-114.01	34647.89	155.96	137.71	167.55	4.59
A	.52	63.17	**	.61	.85	.02
I	-748.23	155530.10	1240.90	1269.57	1591.72	147.75
A	.57	50.10	**	.68	.91	.08
I	-3649.07	992165.89	4850.41	5603.61	5407.64	574.88
A	.50	84.02	**	.75	.87	.08
I	- 63.60	14761.35	99.92	102.91	120.45	10.83
A	.54	64.96	**	.71	.91	.07

Experiment X - Program-parameter 10-12

True Values	.89	.16	44.00	.74	0	.13
Experiment						
X	-258.36	57.96	42794.88	211.316	-188.88	500.45
A	.86	.18	46.30	.78	**	.12
X	134.64	-26.79	18741.79	-106.62	- 34.16	-302.33
A	.90	.18	36.00	.67	**	-.06
X	26.49	- 4.41	-3921.28	- 19.33	- 49.75	- 54.26
A	.89	.17	67.72	.78	**	.16
X	38.95	- 5.44	-6167.44	- 26.41	-162.44	- 82.56
A	.86	.14	44.98	.72	**	.20
X	132.54	-28.06	-22638.77	-109.00	-2289.91	-238.50
A	.85	.21	50.48	.72	**	.13

* Table shows estimates of the under-specified equation made by FIML which were particularly poor. Below them are found the estimates made by FIML in the basic experiment from identical sets of data.

** Not estimated in the basic experiment.

The effects of not specifying that one zero coefficient of the basic structure should be zero depended in part on which zero coefficient was not so specified. In experiments I and X, where γ_{22} and γ_{13} were estimated respectively, the medians of the estimates of all the coefficients of the equation in which the failure occurred were significantly biased; in experiments G and W, where β_{23} and γ_{24} were estimated, this was not the case. Although in all experiments the dispersions of all the estimates of the under-specified equation were larger than in the basic experiment, this was much more pronounced in experiments I and X than in experiments G and W. The same was true for the dispersions of the full-model estimates of the fully-specified equations.

The differences between the experiments were also found in the rankings of the estimators. Table VIII-5 presents the totals of the ranks of the separate estimates of the coefficients of the fully specified equations and of the under-specified equation. FIML did quite a bit more poorly in estimating the under-specified equation in experiments X and I than in experiments G and W or than in the basic experiment. For the fully specified equations, 3SLS and FIML showed hardly any deteriorations in relative standings in two of the experiments[167] but were poorer in the other two experiments[168] than in the basic experiment. The experiment with program-parameter 10 set at 11 was remarkable in that the performances of the methods were not much different from the basic experiment

[167]These were experiments W and X.

[168]These were experiments G and I.

for any of the coefficients. Table VIII-6 records the frequencies with which there were significant differences in the relative performances of the methods in the under-specified experiments from those in the basic experiment. Except for the performances of FIML in the under-specified equation of experiment I, significant differences were not found for more than half the coefficients estimated. Most of the changes in the relative performances of the methods were significant for very few of the coefficients. It is noteworthy that in none of the separate experiments were there pronounced differences in the relative performances of the methods, except that DLS was apt to perform poorly.

There was one very remarkable difference in some of the experiments studying under-specification from the basic experiment. In experiments I and X, with program-parameter 10 set at 4 and 12 respectively, FIML made a large number of very wild estimates of the under-specified equation. Table VIII-7 illustrates this phenomenon by presenting five of the poorest estimates of the under-specified equation in each of these experiments. It also shows the estimates made in the corresponding replications of the basic experiment. The estimates in the under-specified experiment are vastly different from those of the basic experiment made from <u>exactly</u> the same data. In each experiment there was a fair number of other very poor FIML estimates which are not shown in the table. In each of these experiments, LIML also made some poor estimates, but they were nothing like as wild as those of FIML. When either of these methods made bad estimates, all the coefficients of the under-specified equation were very poor.

Several remarkable features are to be noted about the very poor FIML estimates. First, the wretched FIML estimates did not lead to very poor estimates of the reduced form. For most of the reduced-form coefficients, indeed, the largest absolute deviations and the ranges of LSRF were larger than those of FIML. (The LSRF estimates in these experiments were the same as in the basic experiment.) The FIML predictions were also very good. Second, terribly wild FIML estimates did not occur for the fully-specified equations. Third, in experiments G and W neither FIML nor LIML produced very bad estimates of any of the structural equations.

The performances of the standard errors in the under-specified experiments were not the same in all the experiments. In experiments I and X, the standard errors of all methods for the coefficients of the under-specified equation were very poor. For example, more than 60% of the 3SLS t-ratios fell outside the 95% confidence interval of the t-distribution with degrees of freedom based on the number of coefficients to be estimated in the under-specified equation. Basing degrees of freedom on the total number of coefficients to be estimated in the structure improved matters very little (nor did adjustment of the t-ratios for "lost" degrees of freedom improve matters much). 2SLS and UBK had similarly poor performances. DLS did even worse. FIML and LIML did badly, but not as poorly as other methods. Only about 40% of the LIML estimates and 50% of the FIML estimates lay outside the range based on the number of coefficients to be estimated in the equation. In both experiments FIML had several negative "squared standard errors". Had these been counted among the misleading cases, FIML would have done as poorly as other methods. Using the standard errors of any one of the methods would have led very frequently to rejecting the hypothesis that the coefficient whose true value was zero was indeed zero.

The standard errors of the under-specified equation of experiment G were slightly poorer than in the basic experiment, but the deterioration in performance was rarely significant. In experiment W there was no deterioration in the performances of the standard errors.

The standard errors of 3SLS and FIML for the estimates of the fully specified equations were slightly poorer in experiments I and X than in the basic experiment. They were, however, quite reasonable, especially in view of the very poor performances in the other equation. There was virtually no deterioration in their performance in the other two experiments.

The dispersions of the estimates of Σ were larger in all experiments than in the basic experiment. When FIML did very poorly in estimating the under-specified equation, its estimates of Σ were also wretched. The rankings of the methods were not much changed, but they did become weaker, especially in experiments I and X.

The dispersions of the estimates of the reduced-form coefficients (except, of course, for LSRF) were larger when the structural equations were under-specified. This was most pronounced when program-parameter 10 equalled 4. The increases were not larger for FIML than for other methods even when judged by the extreme quantiles.

The over-all rankings of the methods in estimating the reduced form were not much altered by under-specifying the experiments. Except in experiment W, FIML and 3SLS had slightly poorer performances relative to the other methods than in the basic experiment. LSRF improved its standings somewhat. Table VIII-8 shows the frequencies with which changes in the relative performances of the methods were significant. In all these experiments, FIML and 3SLS remained the best estimators of the reduced-form coefficients and DLS and LSRF were still the poorest. The performances of the methods in making predictions also showed but small changes from the basic experiment.

The results of some of the under-specified experiments were peculiar enough to warrant further exploration. First, experiments were conducted with the basic structure and program-parameter 10 set equal to 4 to see if the strange results would also arise with a different set of exogenous data, with smaller disturbances, or with more observations in the samples from which the estimates were made. Second, other structures were used.[169]

Attention was focused on several questions in examining the results of these experiments. Would FIML or LIML make wild estimates of the under-specified equation? Would the estimates of the under-specified equation be biased? Would the rankings of the methods be different from those in the corresponding, fully-specified experiments? Would the standard errors of the under-specified equation be poor? Would the

---

[169] Program-parameter 10 was set equal to 3 in one experiment using MOD37A1. It was set at 4 in experiments using MOD37A1, MOD37A3, MOD37AB, MODEL37C and MODEL37D.

TABLE VIII-8

Under-Specified Experiments - Significant Differences in Estimating the Reduced Forms from the Basic Experiment*

Program-Parameter 10-3

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	4	3	3	0	0	6
2SLS	0	-	0	0	0	0	2
UBK	0	0	-	1	0	0	2
LIML	0	0	0	-	0	0	2
3SLS	0	5	5	5	-	0	3
FIML	0	6	6	6	0	-	4
LSRF	0	1	0	0	0	0	-

Program-Parameter 10-4

DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
-	1	0	0	0	0	4
1	-	0	0	0	0	2
1	2	-	0	0	0	2
2	0	0	-	0	0	2
7	8	7	8	-	0	3
3	5	6	6	0	-	3
3	1	0	0	0	0	-

Program-Parameter 10-11

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	0	0	0	0	0	1
2SLS	0	-	3	0	0	1	0
UBK	0	1	-	0	0	1	0
LIML	0	0	0	-	0	1	2
3SLS	0	0	0	0	-	0	0
FIML	0	0	0	0	0	-	0
LSRF	0	0	0	0	0	0	-

Program-Parameter 10-12

DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
-	1	1	0	1	2	6
3	-	3	0	0	0	2
2	0	-	0	0	0	2
5	4	4	-	0	0	3
8	4	3	0	-	0	3
7	2	2	2	0	-	2
2	0	0	0	0	0	-

*Entries in each row of the table record the number of reduced-form coefficients for which the method was significantly better than each of the other methods in the basic experiment than in the under-specified experiment. The McNemar test and the .05 significance level were used.

TABLE VIII-9

Under-Specified Experiments - Summary of Additional Experiments*

	Structural Equation					Standard Errors	
	Under-specified			Fully specified			
Experiment	Bias	Ranks	Bad Estimates	Bias	Ranks	Under-Specified Equation	Fully Specified Equation
J	Yes	Yes	Yes	No	Yes	Yes	Yes
K	Yes except FIML	Yes	Yes	3SLS	Yes	Yes	Yes
L	Yes	Yes LIML very good	Yes	3SLS	Yes	Yes LIML only slightly worse	Yes but very little
H	Yes except LIML and FIML	No	1 bad LIML No bad FIML	No	No	Yes except LIML; FIML slight	Very slight
M	Yes except FIML	Yes	Yes	No	Yes	Yes	Yes
N	Yes except LIML & FIML	Yes	Yes	No	Yes but deterioration small	Yes all very bad	Very slight
O	Yes except FIML	No **	Yes	No	No **	Yes but all only slightly worse, especially FIML & LIML	Very slight
P	No	No	No	No	No	No	No
Q	No	No	No	No	No	Yes but very slight	Very slight

Notes for Table VIII-9

*Bias	- Frequency of finding significantly biased medians. Yes means found often; No, found seldom.
Ranks	- Changes in rank-totals of all estimates. Yes - rank-totals affected as in experiment I. No - rank-totals much as in corresponding fully specified experiment.
Bad Estimates	- Yes - FIML often made very wild estimates and LIML made a few poor ones. No - No wild estimates.
Standard Errors, Under-specified Equations	- Yes - Standard errors of all methods did very poorly. The standard errors of FIML and LIML were not as bad as for other methods. No - Standard errors much the same as in corresponding fully-specified experiment.
Standard Errors, Fully-specified Equations	- Yes - Some deteriorations in performance of 3SLS and FIML standard errors. No - FIML and 3SLS standard errors as in the corresponding fully-specified experiment.
Experiment	- Program-parameters shown in Table VIII-4.

**FIML and 3SLS did poorly in the fully specified experiment using MOD37AB. No deteriorations in their positions here.

rankings of the estimates of 3SLS and FIML of the fully specified equations differ from those in the corresponding, fully specified experiments in the same ways as were found in comparing experiment I with the basic experiment? Table VIII-9 summarizes the answers. In almost all cases the answer was yes. There were, however, some exceptions where the failure of specification did not affect the results very much.[170]

The very poor FIML estimates of the under-specified structural equations, the frequent finding of significantly biased medians for the estimates of that equation, the poor performances of the standard errors and the failure of the poor estimates of the structural equations to lead to poor estimates of the reduced-form may suggest that the difficulties which were encountered in some of the experiments were due to poor identification of the under-specified equation.[171] In all these experiments the equation was identified. However, it was a good deal less well identified in some of them when part of the *a priori* information was not used than when all the information was incorporated in the estimates.

---

[170] One exception was the experiment using MOD37A1 with program-parameter 10 set at 3. There the results were very similar to those obtained when program-parameter 10 was set at 3 in the basic structure. The other exceptions were the experiments using MODEL37C and MODEL37D in which the failures of specification had very little effect. MODEL37C and MODEL37D were of different from from the basic structure. MOD37A1 differed from it in the values of the structural coefficients.

[171] I am indebted to Professor R. E. Quandt for calling this possibility to my attention. It should be pointed out that experiments G, I, J and M were conducted before this possibility was recognized and that it was suggested by the results of those experiments.

Table VIII-10 presents the rank-criterion matrices of the under-identified structural equations.[172] When difficulties were encountered, there was only one non-zero coefficient in one of the other equations which distinguished the form of that equation from the under-specified one. Difficulties were encountered whenever the true value of this structural coefficient was small, especially when it was small in comparison to the identifying coefficient which was dropped because of under-specification. When the value of the remaining coefficient was not very small, no difficulties were encountered.

Summary. Peculiar results, which differed from those of earlier experiments, were obtained in some, but not all, of the experiments in which an attempt was made to estimate coefficients whose true values were zero. The peculiar results consisted of large and significant biases for the estimates of the under-specified equation, very poor performances of the standard errors of the coefficients of the under-specified equation (especially by methods other than the maximum-likelihood estimators), and a large number of very bad estimates of this equation made by FIML. The reason for these results may have been that the equation had become only very weakly identified. If this were true, the very wild FIML estimates may be an asset to users of this estimator. Such estimates were so

---

[172] For a definition and discussion of the rank-criterion matrix see Tjalling C. Koopmans, "Identification Problems in Economic Model Construction," in William C. Hood and Tjalling C. Koopmans (eds.), Studies in Econometric Method (New York: Wiley, 1953), pp. 35-44.

TABLE VIII-10

Under-Specified Experiments - Rank-Criterion Matrices for the Under-Specified Equations*

MOD37A2 - MOD37AB

Second Equation

(3)	(4)**	((11))	
.16	.74	0.	0.
-1.0	0.	.11	.56

First Equation

(12)**			
.70	0.	0.	.06
.53	.11	.56	0.

MOD37A1 - Second Equation

(3)	(4)**		
.22	.65	0.	0.
-1.0	0.	.39	.18

MOD37A3 - Second Equation

	(4)**		
.04	.87	0.	0.
-1.	0.	.44	.49

MODEL37C - Second Equation

(4)			
.89	0.	.43	0.
0.	.82	0.	.71

MODEL37D - Second Equation

	(4)			
.7	.89	0.	.43	0.
-1.0	0.	.82	0.	.71

*Table shows the rank-criterion matrices of the structural equations which in some experiments were not fully specified. An underlined number in parentheses indicates that when program-parameter 10 was set at that value the column was deleted from the rank-criterion matrix by failing to fully specify the equation.

**In the experiments with this structure and this value for program-parameter 10, the peculiar results occurred.

patently absurd that it seems likely that in estimating an empirical structure, one could reject out of hand the possibility that these estimates were an adequate representation of the true structure.

5. Estimating an Unidentified Structure

It was suggested that some surprising results were obtained when not all the possible, valid, *a priori* restrictions were placed on the structure to be estimated because the under-specified equation was only very weakly identified. The next step of the investigation was to estimate a structure which failures of specification had left in reality unidentified. To do this, the failures of specification obtained by setting program-parameter 10 equal to 2, 3 and 4 were all made. This was the result of setting program-parameter 10 equal to 6. Two zero coefficients of the second equation were to be estimated. An extraneous exogenous variable was included in the structure for which coefficients were to be estimated in the first and third equations. As shown in Table VIII-11, the second equation of the basic structure is not, in truth, of different form from the first equation. It *appears* to be different if it is supposed that there is a non-zero coefficient for the extraneous variable in the first equation. Experiment R of the fourth set of experiments (Appendix C, Table 4) used the basic structure with program-parameter 10 set at 6.

The results of experiment R were quite similar to those obtained in the experiment using the basic structure with program-parameter 10 set

TABLE VIII-11

Rank-Criterion Matrix of the Second Structural Equation of MOD37A2 when Program-Parameter 10 = 6

0	0	0*	(first equation)
.11	.56	0*	(third equation)

*Coefficient of the extraneous variable which was to be estimated.

---

at 4.[173] The estimates of the second equation made by all methods were significantly biased. However, as shown in Table VIII-12, the medians of the estimates of coefficients other than the constant were not greatly different from the true values of the structural coefficients. It will be noted that for each coefficient, the medians of the estimates of the different methods were very similar to each other. The dispersions of the estimates were larger than in the basic experiment; they were not larger than in the experiment with program-parameter 10 set at 4 according to most of the descriptive statistics.

FIML made several wild estimates of the coefficients of the unidentified second structural equation. However, these were not as large as the poor estimates of the experiment with program-parameter 10 set at 4, nor did bad estimates occur as frequently as in that experiment. LIML did not produce ridiculous estimates of this equation.

The estimators were all ranked very close to each other for the coefficients of the second equation, as can be seen from Table VIII-13.

---

[173]That is, when the first zero γ-coefficient of the second equation was estimated.

TABLE VIII-12

Median Estimates of the Unidentified Equation, Experiment R*

Coefficient	True Value	DLS	2SLS	LIML	3SLS	FIML
β_{21}	.54	.82	.83	.83	.84	.82
β_{23}	0	-.09	-.08	-.08	-.08	-.08
γ_{21}	62.00	8.84	7.63	7.17	6.73	8.79
γ_{22}	0	-.40	-.40	-.41	-.42	-.41
γ_{23}	.70	.37	.36	.35	.34	.37
γ_{25}	.96	.53	.52	.50	.41	.43
γ_{27}	.06	.03	.03	.03	.03	.03

*N.B. The figures are the medians of the distribution of the estimates of the estimates of the individual coefficients and did not necessarily all occur in the same replication.

TABLE VIII-13

Rankings of Estimates of the Methods, Experiment R*

Equation		DLS	2SLS	UBK	LIML	3SLS	FIML
1,	totals	1088	1010	1040	1108	928	1086
2	β_{21}	155	153	153	173	201	215
	β_{22}	189	155	155	168	177	206
	γ_{21}	152	162	162	175	201	198
	γ_{22}	155	152	152	175	203	213
	γ_{23}	149	162	162	176	197	204
	γ_{25}	158	155	155	181	206	195
	γ_{27}	130	175	175	203	186	181
2,	totals	1088	1114	1114	1251	1371	1412
3,	totals	1248	1048	977	957	932	1099

*Table entries are the totals of the ranks of the separate coefficients in all the replications of the experiment.

DLS was ranked the best by the totals of the ranks of the estimates of all the coefficients in all replications and FIML was poorest. Table VIII-14 shows how often the changes in relative standing from the basic experiment were significant. The changes of FIML, 3SLS and DLS were almost always significant.

The standard errors of the estimates of the unidentified equation performed poorly. The t-ratios of all methods were very frequently greater than two. They fell outside the 95% confidence interval of the t-distribution with degrees of freedom based on the number of coefficients to be estimated in the equation very often. LIML performed slightly better than methods other than FIML, but for most coefficients, it fell outside the range over 60% of the time. FIML did better: it was outside the range only about half the time. If the test interval were based on the number of coefficients in the structure to be estimated, the t-ratios of LIML and FIML fell outside the range no more often than in the basic experiment. The standard errors of other methods were still poor, but not much more than 20% of the time did the t-ratios lie outside the range. Of course, the 95% confidence interval of the t-distribution with only two degrees of freedom is rather large. Adjusting the standard errors for degrees of freedom before using them to make inferences about the dispersions of the estimates of the unidentified equation would not have led to reliable inferences in experiment R. The dispersions of the other two equations of experiment R were larger than in the basic experiment. The rankings of the methods for these equations are shown in Table VIII-12. In both equations FIML did poorly. Table VIII-14 shows how often changes in the standings were significant. The standard errors of FIML and 3SLS were only a little less reliable than in the basic experiment and their performance was similar to that of the experiment using the basic structure with program-parameter 10 set at 2 (that is, the experiment with the extraneous exogenous variable).

The dispersions of the estimates of the reduced-form were a little larger when program-parameter 10 was set at 6 than when it was set at 4. The rankings of the methods were a bit different. DLS and 3SLS did better; FIML did more poorly and was ranked slightly behind 2SLS and UBK. 3SLS, although still first, did more poorly than in the basic experiment. The frequency of significant differences from that experiment are presented in Table VIII-15. For no pair of methods were there significant changes in their relative standings for half the reduced-form coefficients.

TABLE VIII-15

Significant Differences from the Basic Experiment in Estimating the Reduced-Form Coefficients of Experiment R*

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	2	1	1	1	1	3
2SLS	3	-	0	0	0	0	2
UBK	3	0	-	0	0	0	2
LIML	1	3	2	-	0	0	2
3SLS	9	8	8	7	-	0	3
FIML	9	8	10	8	1	-	4
LSRF	4	0	0	0	0	0	-

*Entries in each row are the number of reduced-form coefficients for which the performance of the method relative to each of the others was significantly better in the basic experiment than in experiment R. The McNemar test and the .05 significance level were used.

Since the results of experiment R were surprisingly good, it was decided to use the basic structure with program-parameter 10 set at 6 in several other experiments. Two of these investigated the effect of introducing multicollinearity. One of them used twenty observations and the other one seventy observations.[174] Trends as well as multicollinearity were used in another experiment with program-parameter 10 set at 6.[175] Finally, an experiment with auto-correlated disturbances as well as failure to identify one equation was conducted.[176]

The results of these experiments were quite similar to those of experiment R. The standard errors performed about as poorly. FIML made several wild estimates in each experiment, but the bad estimates did not occur more often than in experiment R nor were they markedly poorer. The medians of the estimates of the structural coefficients of the second equation were similar to those of experiment R. However, in two experiments,[177] the medians of LIML and FIML were not significantly biased. The reason for this was larger dispersions of the estimates rather than medians which were much closer to the true values of the structural coefficients. The changes in the rankings due to setting program-parameter 10 at 6 were usually similar to those found in comparing experiment R with the basic experiment, but they were not usually as pronounced. In estimating the reduced-form coefficients of experiments with multicollinearity

---

[174] These are experiments AL and AM of Appendix C, Table 4.

[175] This is experiment AO of Appendic C, Table 4.

[176] This is experiment AN of Appendix C, Table 4.

[177] These were the experiments with multicollinearity and twenty observations, experiments AL and AO.

in the exogenous variables, DLS was apt to be poorer than in corresponding, fully-specified experiments.

Summary. The results of estimating the unidentified structure were quite surprising. While the performances of the methods could hardly be described as good, they were similar to, but better than, those found in Section 4 when only one of the zero-valued structural coefficients was estimated. If the conjecture that the results of the previous section were due to weak identification was true, it is surprising that failing to identify one equation at all did not lead to even poorer results.

6. An Unidentifiable Structure

All structural coefficients whose values in the structure generating the data were less than .1 were specified to be zero in the structure estimated if program-parameter 10 was set equal to 10. A special structure, MODNOID, was used when program-parameter 10 assumed this value. This structure, shown in Table VIII-16, differed from the structure of the basic experiment, MOD37A2, in two ways. All coefficients which were zero in MOD37A2 had non-zero, but very small, values in MODNOID; and one coefficient, γ_{27}, which was very small in MOD37A2 was larger in MODNOID. The true structure of MODNOID is not identifiable; the structure estimated, which is of the same form as MOD37A2, is identifiable. The specification used for the estimates of MODNOID is almost correct in the sense that the variables deliberately excluded from each equation played a very small

TABLE VIII-16

MODNOID

$$B = \begin{bmatrix} 1 & -.89 & -.16 \\ -.54 & 1. & -.03 \\ -.02 & -.29 & 1. \end{bmatrix}$$

$$-\Gamma = \begin{bmatrix} 44. & .74 & .01 & .02 & .13 & .01 & .02 \\ 62. & .01 & .70 & -.04 & .96 & -.03 & .12 \\ 40. & .04 & .53 & .11 & -.01 & .56 & .01 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 35.24 & 34.48 & 31.13 \\ & 36.68 & 29.84 \\ & & 40.64 \end{bmatrix}$$

role in it. It is of interest to see whether use of this almost correct information would allow good estimates of other structural coefficients and whether the structural-equation estimators would make better estimates of the reduced form than the correctly-specified LSRF.

One experiment used MODNOID with program-parameter 10 set at 10 with the exogenous variables and disturbances of the basic experiment. This structure with the same misspecifications was also used in experiments having multicollinearity, lagged-endogenous variables, auto-correlated disturbances, and heteroskedastic disturbances.[178] The differences of the results of these experiments from the ones having the same features of the data, but no misspecification, were all very similar and can be presented together.

In estimating the structural coefficients, the medians of methods other than DLS and 2SLS were significantly biased more often in the experiments using MODNOID than in experiments without misspecifications. When auto-correlated disturbances were used, FIML and 3SLS were significantly biased even more often. The medians of the estimates of the reduced-form coefficients of all methods except DLS and LSRF were also biased more frequently in the experiments using MODNOID.

FIML and 3SLS were not as clearly better than other methods in estimating the structural coefficients of MODNOID as they were in the corresponding experiments without misspecification. This held both for the standings of the methods on the number of times they came closer to

[178]These are experiments AQ-AT of the fourth set of experiments conducted, Appendix C, Table 4.

the true values than the others and for the standings according to many of the descriptive statistics. The deteriorations of the methods, however, were slight except when auto-correlated disturbances were used. Then FIML did much more poorly. This is surprising in view of the lack of effect from auto-correlated disturbances found in Chapter VI, Section 5, and the absence of auto-correlation from the exogenous variables. In estimating the reduced-form coefficients, LSRF and often DLS tended to be a bit better when MODNOID was used while FIML and 3SLS were poorer. The changes in the standings were small. LSRF, though correctly specified, was not the best method, but instead was usually judged to be the poorest estimator of the reduced form. In the predictions, also, LSRF became slightly better but was still poor.[179]

The standard errors of all methods were slightly less reliable when MODNOID was used than in the basic experiment. Almost no changes were observed in the relative standings of the methods for estimating the residuals.

<u>Summary</u>. The estimates of the misspecified structure differed little from those of the correctly-specified one. Specification that very small coefficients be zero when estimating the structure led not only to the structural-equation estimators making adequate estimates of the major structural coefficients, but also led to these methods making better estimates of the reduced form than the correctly-specified LSRF.

[179] The predictions were correctly analyzed when MODNOID was used.

7. Specification that an Important Non-Zero Coefficient be Zero

When program-parameter 10 was set equal to 7, β_{12} was specified to be zero even though in the structure generating the data it had a non-zero value. One experiment was conducted using the basic structure with program-parameter 10 set at 7.[180] In the basic structure, β_{12} was one of the largest coefficients.

Structural Coefficients. There are two interesting questions in examining the estimates of the structural equations: 1) how did the methods perform in estimating the first structural equation and 2) how were the estimates of other equations by the full-model methods affected by the misspecification.[181]

The estimates of all methods for the coefficients of the mis-specified equation were very poor. Hardly any of the estimates of the coefficients other than the constant came within 40% of the true values and all the medians of the estimates were significantly biased at the .05 level. The rankings of the methods for the estimates of the first equation were peculiar. For two coefficients, DLS was by all odds the best method and LIML did most poorly; LIML did best and DLS worst in estimating the other coefficients. When LIML was worst, FIML was the second poorest and 2SLS the second best. Otherwise no pattern emerged.

[180] This was experiment S of Appendix C, Table 4.

[181] The estimates of the second and third structural equation by the k-class estimators were not affected by the misspecification.

The FIML estimates of the other equations were considerably poorer than in the basic experiment. The McNemar test found that these deteriorations in the positions of FIML were significant for all but one of the structural coefficients. 3SLS did slightly more poorly than in the basic experiment in estimating the coefficients of the second and third structural equations, but the deteriorations in its performances were rarely significant.

Standard Errors. The standard errors of all methods were virtually useless for making inferences about the true values of the coefficients of the misspecified equation. Over 80% of the t-ratios were greater than 2. For the constant, the standard errors were more reliable but were still considerably poorer than in the basic experiment.

The standard errors of FIML for the other equations were a good deal less reliable than in the basic experiment. Those of 3SLS were only slightly less trustworthy.

Residuals. The estimates of the variances of the residuals of the misspecified structural equation were much larger than the variance of the true structural residuals. Usually the estimated residuals were more than seven times as large. The variances of the FIML residuals were almost always larger than those of other methods. FIML also had fairly large estimates for the variances of the residuals of other equations. The medians of the FIML estimates were significantly higher than the true values. It will be recalled that in the basic experiment, FIML tended to underestimate the elements of Σ.

The residuals of 3SLS were not as sensitive to the misspecification. They were similar to those of other methods for the misspecified equation. In the other two equations, the variances of the 3SLS residuals tended to be only slightly higher than in the basic experiment.

Reduced-Form Coefficients. Setting program-parameter 10 equal to 7 led to finding many significantly biased medians for the estimates of the reduced-form coefficients made by methods other than LSRF. Half the FIML estimates were significantly biased; more than 80% of the medians of the other estimators were significantly different from the true values. In addition to increases in the biases, there were also increases in the dispersions of the estimates about their central tendencies.

LIML ranked first according to the sum of the ranks of all the estimates of all the coefficients as judged by their absolute deviations from the true values. LSRF was second with FIML slightly behind it. Kendall's W for these rank-totals was, however, only .06. The agreement between the ranks of the rank-totals for the individual coefficients was also fairly low, Kendall's W being only .19. Many significant changes in the relative standings of the methods from the basic experiment were found by the McNemar test as is shown in Table VIII-17. The improvements of LIML and LSRF stand out. In most cases a method improved significantly for some coefficients while for others there was significant deterioration. Significant differences were not found for all the coefficients.

Summary. Since only one experiment investigated misspecifying an important coefficient to be zero, the conclusions must be tentative. It was not found possible to make a sensible ranking of the methods for estimating the misspecified equation. For this equation, the estimates were poor and the standard errors were worthless. FIML's performance in estimating the other equations deteriorated markedly. 3SLS was scarcely affected. LIML and LSRF did well in estimating the reduced form.

TABLE VIII-17

Misspecification of an Important Coefficient - Significant Differences in the Reduced Form*

	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
DLS	-	12	12	11	6	4	10
2SLS	6	-	12	12	0	4	12
UBK	6	6	-	14	0	4	13
LIML	6	6	6	-	0	4	9
3SLS	5	5	6	13	-	5	12
FIML	7	5	5	8	5	-	13
LSRF	1	0	0	0	0	0	-

*Table entries record the number of times the method in each row was significantly better in the basic experiment than in experiment S of Table VII-2 than each of the other methods. The McNemar test and the .05 significance level were used.

8. Specification of Structures Quite Different from the True Ones

When program-parameter 10 was set at 8 or 9, coefficients whose true values were non-zero were specified to be zero in the structure estimated and one coefficient which was zero in generating the data was estimated. With program-parameter 10 set at 8, β_{12} and β_{21} were both specified to be zero and β_{23} was to be estimated. In addition to these failures of specification, the last exogenous variable was deleted from the structure to be estimated when program-parameter 10 was equal to 9. Neither of the coefficients of the beta matrix which were incorrectly specified to be zero were unimportant in the basic structure.

The structures to be estimated are quite different from the true structures, especially with respect to the inter-dependence of the endogenous variables. The experiments of this section can be interpreted as estimating an econometric model which is based on a mistaken theory of the structural relationships.

Two experiments used the structure and data of the basic experiment with program-parameter 10 set at 8 and at 9. One experiment used multicollinear exogenous data with the basic structure and program-parameter 10 set at 9.[182]

The results of the experiments with program-parameter 10 set at 8 and 9 were very similar to each other. The omission of the last exogenous variable had very little effect on any of the results. The experiments indicated that the effects of specifying that important coefficients be zero were similar to the effects discussed in the previous section.

The medians of the estimates of the coefficients of the first two structural equations were all significantly biased. Some of the medians of the FIML and 3SLS estimates of the coefficients of the third

---

[182] These were experiments T, U and AP of the fourth set of experiments, Appendix C, Table 4.

equation, which was correctly specified, were also significantly biased. Any resemblance of the medians of the estimates of the first two equations to the true values appeared to be largely coincidental. They were, however, not much poorer than the medians of the estimates of the unidentified equation of the experiment with program-parameter 10 set at 6, shown in Table VIII-12, nor were the medians of the estimates of the full-model methods of the first equation greatly worse than in the experiment with program-parameter 10 set equal to 7.

The dispersions of the estimates of the coefficients of the first two structural equations around their central tendencies were a bit larger than in the basic experiment, but not very greatly so. The increases in dispersion were greater for the full-model methods than for the k-class estimators for most coefficients. However, the full-model methods were still ranked fairly well by the dispersions of their estimates from their central tendencies. Rankings of the estimates by their deviations from the true values largely reflected which methods had medians closer to the true values. The rankings were quite different for different coefficients and an over-all ranking of the methods did not seem sensible. It may be noted, however, that the relative positions of DLS were often better than in the basic experiment.

The standard errors of the coefficients of the misspecified equations were not useful for making inferences about the true values of the coefficients. They would have led to reliable inferences about the central tendencies of the distributions of the estimates.

The variances of the residuals of the misspecified structural equations were invariably very large. They clearly indicated that the structure did not fit the data very well. The variances of the FIML residuals of the first equation were usually substantially smaller than those of the experiment where only one coefficient was incorrectly specified to be zero.[183]

All the structural-equation estimators produced significantly biased medians for almost all the reduced-form coefficients. Their dispersions about these central tendencies were smaller than the dispersions of LSRF. LSRF, however, ranked as the best estimator of the reduced-form. FIML and 3SLS were the poorest. For many of the coefficients, these last two methods were significantly poorer than others in the number of times they came closer to the true values. All structural-equation methods ranked as poorer predictors than LSRF. FIML and 3SLS generally were ranked poorest.

When multicollinearity was used with program-parameter 10 set at 9, the results were quite similar to those obtained without multicollinearity. The structural equations were, if anything, estimated more poorly. The medians of the estimates of the reduced-form coefficients were quite far from the true values. Though usually significantly biased, they were often closer to the true values than in the experiment with the misspecification but without multicollinearity. DLS improved its standing relative to other methods to a small extent.

---

[183] That is, the experiment with program-parameter 10 set equal to 7, reported in the previous section.

The variances of the residuals of the misspecified equations, especially of the second equation, were smaller when multicollinearity was present. The medians of the variances of the residuals of the second equation were only two-thirds or less the size of the medians when there was no multicollinearity. With all variables being highly correlated with each other, one might guess that the variables would act as better substitutes for each other. Effects which were falsely specified not to exist would be attributed to other variables with the result that the false structure would give a better explanation of the endogenous variables. This, in part, seems to have happened.

Summary. The structural coefficients estimated in the misspecified experiments were not similar to the true coefficients. The false structures did not fit the data well, though with multicollinearity the fit was apt to improve considerably. The results of this section and the preceding one are to be contrasted with those of Section 6 where it was found that specifying that unimportant non-zero coefficients be zero did not lead to poor performances by the estimators.

9. Omission of an Equation

The last type of misspecification investigated was the effect of omitting one of the structural equations which generated the data from the structure estimated. When program-parameter 10 was set at 13, the last structural equation was not estimated. The variable "explained" by this equation was also deleted. A value of 14 for program-parameter 10 also led to the omission of the last equation. The last endogenous variable, however, was included in the structure to be estimated where it was treated as a predetermined variable.

Two special structures were used in these experiments, MOSPEC1 and MOSPEC2. These structures are shown in Table VIII-18. In MOSPEC1 the last endogenous variable played only a very small part in the first three structural equations. The other endogenous variables played large roles in the last structural equation. This structure was used in one experiment with program-parameter 10 set at 13. In MOSPEC2, the other endogenous variablesplayed a small role in the last equation, but the last endogenous variable was important in other equations. MOSPEC2 was used with program-parameter 10 set at 14 in one experiment. One structure which had been investigated earlier was used in experiments with program-parameter 10 set at 13 and 14.[184] Finally, MOSPEC1 and MOSPEC2 were used in experiments which had a variety of the special features of the generation of the data which were explored earlier, as well as misspecifications.[185]

The results of the experiments using MOSPEC1 and MOSPEC2 without special features of the generation of the data were very similar to the results of the basic experiment. However, LIML and FIML showed significantly

---

[184]The structure was MODEL46, shown in Table V-14. The experiments are Y-AB of the fourth set of experiments, Appendix C, Table 4.

[185]These are experiments AU and AV of Appendix C, Table 4.

TABLE VIII-18

Two Additional Structures

MOSPEC1

$$B = \begin{bmatrix} 1 & -.89 & -.16 & -.04 \\ -.54 & 1 & 0 & -.09 \\ 0 & -.29 & 1 & -.06 \\ -.42 & -.36 & .54 & . \end{bmatrix}$$

$$-\Gamma = \begin{bmatrix} 44 & .74 & 0 & 0 & .13 & 0 & 0 \\ 62 & 0 & .7 & 0 & .96 & 0 & .06 \\ 40 & 0 & .53 & .11 & 0 & .56 & 0 \\ 29 & 0 & 0 & 0 & 0 & 0 & .52 \end{bmatrix}$$

MOSPEC2

$$B = \begin{bmatrix} 1 & -.89 & -.16 & 0 \\ -.54 & 1 & 0 & -.4 \\ 0 & -.29 & 1 & -.6 \\ -.1 & 0 & 0 & 1 \end{bmatrix}$$

$$-\Gamma = \begin{bmatrix} 44 & .74 & 0 & 0 & .13 & 0 \\ 62 & 0 & .7 & 0 & .96 & 0 \\ 40 & 0 & 0 & .11 & 0 & .56 \\ 24 & .65 & 0 & .21 & 0 & .39 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 35.60 & 35.16 & 31.36 & 9.32 \\ & 37.32 & 30.16 & 11.32 \\ & & 40.80 & 8.80 \\ & & & 37.40 \end{bmatrix}$$

biased medians for the estimates of the structural coefficients slightly more frequently. In estimating the reduced-form coefficients, especially of MOSPEC1, the medians of all methods were significantly biased quite frequently. The rankings of the methods were but little affected. Overall they were the same as in the basic experiment except for reversals of order among methods which were found to be very similar to each other in the basic experiment.[186] The standard errors performed, by and large, as in the basic experiment.

When the structure which had been studied in earlier experiments was used with program-parameter 10 set at 13 or 14, the differences from the correctly-specified experiment using that structure again were slight. However, when program-parameter 10 was set at 13 there was little difference in the rankings of the full-model methods from those of the consistent k-class estimators. In both experiments, but especially when program-parameter 10 was set at 14, the usefulness of the standard errors was less than when there was no misspecification.

More pronounced differences from other experiments occurred when lagged-endogenous variables, time-trends, auto-correlated disturbances, multicollinearity and heteroskedastic disturbances were all used with MOSPEC1 and MOSPEC2 with program-parameter 10 set at 13 and 14 respectively.

---

[186] When program-parameter 10 was set at 14, the calculated reduced-form used as the true value of the reduced-form treated the last endogenous variable as an exogenous variable, as in the structural specification.

When the last endogenous variable was omitted from MOSPEC1, the main difference from earlier experiments was that FIML did poorly in estimating the structural coefficients and was ranked behind both 3SLS and 2SLS. The standard errors of all methods were much poorer. FIML and 3SLS had especially unreliable standard errors. There was very little difference between the methods in estimating the reduced form. DLS did about as well as or better than other k-class estimators.

When MOSPEC2 was used with all the other special features, DLS was ranked as the best estimator of the structural and reduced-form coefficients. It was followed fairly closely by FIML and 3SLS. The other k-class estimators did much more poorly than these methods and the differences between the two groups of methods was stronger than the differences observed in the basic experiment. No method had very good standard errors. 2SLS, UBK and LIML were particularly over-optimistic about the dispersions of the estimates. 3SLS erred on the other side. All its standard errors grossly over-stated the variability of the estimates of the structural coefficients. None of the 3SLS t-ratios was greater than two.

Summary. Omitting a structural equation had little effect on the performances of the estimators when other special features of the data were not present. When these other features were present, the misspecification affected the relative performances and the standard errors of the methods. The effects depended on whether the misspecification was the omission of both the equation and an endogenous variable or the treatment of an endogenous variable as exogenous. In the latter case, DLS performed remarkably well while other k-class estimators performed especially poorly. This was one of the few experiments where there were pronounced differences in the relative performances of the methods.

10. Conclusion

The results of the experiments seemed to depend in part on which type of misspecification was studied, but even more important was the seriousness of the misspecification, in the sense of how greatly the structure estimated differed from the true structure. If the failures of specification were slight, so were the changes in the performances of the methods. For example, the specification that coefficients whose values were almost zero be zero, the omission of an unimportant variable, the inclusion of an extraneous variable, or the estimation of a zero-valued coefficient which was not important in identifying the structure had very little effect. When more serious failures of specification occurred, the changes in the performances of all estimators were marked. When the structure estimated was of form quite different from that generating the data or where an equation was either very weakly identified or not identified at all, the estimates were far from the true values.

When a misspecification affected only one equation of the structure, the FIML and 3SLS estimates of the other equations were affected and usually became poorer. The changes in their performances were not great and did not usually lead to these methods being judged poorer estimators of these equations than the k-class estimators. The full-model methods were usually slightly more sensitive to misspecifications than other estimators when judged by their estimates of the reduced form, but they were still apt to be the best estimators of it.

All methods produced what might seem to be reasonable estimates of seriously misspecified equations most of the time. Even though the estimates must be judged as very poor estimates of the true values of structural coefficients, their orders of magnitude could hardly be judged ridiculous nor were the residuals of the misspecified equations very large. Without knowledge of how well the equations should explain the data, it seems unlikely that the equations would have been recognized as being inadequate. In these cases, the standard errors would not usually have indicate the weakness of the specifications. FIML, in cases of weak specification, did occasionally produce sufficiently wild estimates that they might be recognized as useless. Against this feature of the method must be weighed the very poor estimates it made in some of the correctly specified experiments reported in Chapters V and VI.

CHAPTER IX

Conclusion

The Monte Carlo studies conducted for this study led to many interesting findings. Earlier chapters have discussed the results of separate experiments and groups of experiments. These showed how the estimators of systems of simultaneous equations were sensitive to the specific types of data used. The estimates of the structures made in the course of the experiments also suggested several more general conclusions about the performances of the estimators.

The differences in the performances of the estimators were not great. The most striking and important feature of most of the experiments was the low values of Kendall's W for the totals of the ranks of the deviations from the true values of the individual estimates of the parameters. These low values of W suggest that the advantage to be gained from using any one of the estimators rather than another one is slight. The fact that in any one experiment the number of parameters for which the dispersion about the true values of the estimates of one method were significantly different from those of another was usually small indicates that the differences between the methods are not very important.

The frequency with which significant differences were found in the relative performances of the methods was another important feature of the experiments. The number of times one method came closer to the true values of the parameters than another was apt to vary significantly over the parameters of any one experiment. The frequency with which one method came closer to the true value of a particular parameter in an experiment with one set of data quite often differed significantly from the frequencies in experiments using other sets of data. Even when changes in the data were made which one would not expect to affect the performances of the methods, such as the use of different sets of exogenous data drawn from the same populations or different values for the structural coefficients, significant differences were found between the experiments. The finding of these significant differences indicates that it would be impossible to make a general assessment of the probability of doing better by using one method for estimating an econometric model than by using another one.

A weak ranking of the methods was suggested by many of the experiments. This ranking placed FIML and 3SLS in the first places. Next came 2SLS, UBK and LIML. In the last places were found DLS and (when used) LSRF. The closeness of 3SLS and FIML and of 2SLS, UBK and LIML to each other was one of the striking features of the experiments. But it must be emphasized that this grouping together of the methods did not occur in all the experiments and that the ranking of the methods was not always found. Indeed, in the experiments conducted, it would have been possible to find almost any ranking of the estimators, at least according to some of the criteria examined. Considering the closeness of the structures and the data used in most of the experiments to each other, it would be fairer to say that the experiments suggested no general ranking of the estimators rather than to say that they established a ranking of the methods into three groups.

Comparisons of the distributions of the estimates of each of the methods shed some light on the problem of choosing an estimator for econometric models. They indicated that the choice of an estimator is a difficult matter which could not be settled by the Monte Carlo experiments alone. Quite frequently in the experiments, the distributions of the estimates had the appearance of Figure IX-1. (For clarity of drawing, Figure IX-1 exaggerates considerably the differences between the methods.) The estimates of the full-model methods were usually more highly concentrated about the true values of the parameters than the consistent k-class estimates. At the same time, it was often found that the distributions for the full-model methods had thicker tails than those of the consistent k-class methods. That is, it was quite often the case that the probabilities of making estimates which were close to the true values and of making estimates which were far from the true values seemed to be higher for the full-model estimators than for the consistent k-class ones. DLS, unlike the other methods, was apt to show pronounced biases, but the dispersions of its estimates about their central tendencies were as good as those of other methods and were frequently better. While this sort of pattern was evident in many of the experiments, it must be emphasized that this was not always the case. There were experiments where the full-model methods did not have thicker tails than the k-class estimators.[187] There were other experiments where they were not more highly concentrated about the true values than the consistent single-equation methods. Even

[187] This was true for many of the distributions shown in Appendix E.

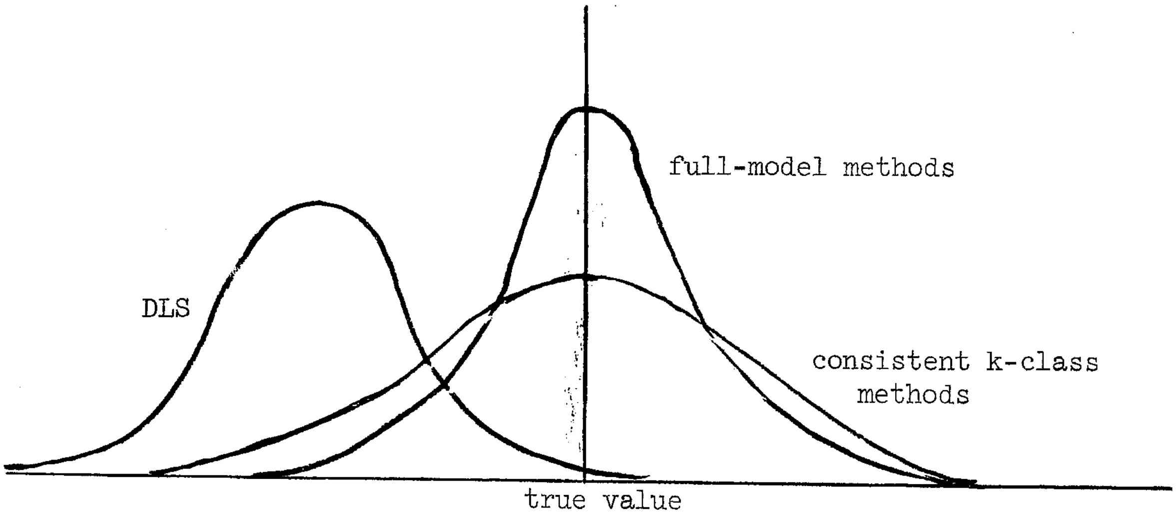


Figure IX-1

Typical Frequency Functions of the Estimates

if the patterns of Figure IX-1 were the ones always found, no one estimator is clearly better than the others. The choice of an estimator could be made only on the basis of the purpose for estimating a particular econometric model. It is worth noting that in the correctly specified experiments, LSRF often had distributions with smaller concentrations about the true values and thicker tails than any of the consistent structural estimators. It could be judged poorer than them for the estimates of the reduced-form coefficients and of the predictions on the basis of a comparison of the distributions.

The absolute performances of the estimators are hard to evaluate since judgment of the adequacy of the estimates must rest on the purposes for which an econometric model is built. However, it does seem true that, if the object of an econometric model is to get an idea of the orders of magnitude of the parameters of economic systems, the estimators may be useful. Usually the estimates were of the right size and were quite close to the true values. In addition, larger parameters were apt to have larger estimates than smaller ones. If great precision is needed, the estimators will not serve, at least not if small samples of data are used. It was rare to find inter-quartile ranges for the estimates of the structural or the reduced-form coefficients which were less than 20% of the true values being estimated. Usually they were a great deal larger. It is worth noting that the predictions were usually a good deal more accurate than the estimates of the coefficients.

The standard errors of the consistent methods were found to be useful for making inferences about the deviations of the estimates of the structural coefficients from the true values in most experiments. As a practical matter, it was found advantageous to adjust the standard errors for the number of coefficients to be estimated in each equation.[188] The hypothesis that the t-ratios thus adjusted follow the t-distribution with degrees of freedom equal to the number of observations minus the number of coefficients in each equation led to fairly reliable confidence intervals for the estimates in most of the experiments. However, there were several experiments where this hypothesis would not have led to reliable inferences about the dispersions of the estimates. Some of the most notable cases of the failure of the t-ratios occurred when the true value of a coefficient was zero. Probably the most common use of the

[188] Strictly speaking, the estimates of Σ should be adjusted for the "lost degrees of freedom" before being used to compute the standard errors. It was found that the adjustment would usually lead to more reliable estimates of Σ as well as to better standard errors.

standard errors is to test whether estimates differ significantly from zero. The failure of the standard errors when the true values were zero may well lead to one doubting the usefulness of the standard errors. When the standard errors were not reliable, it was frequently the case that the LIML standard errors were the least unreliable. The DLS standard errors did not lead to reliable inferences about the deviations of its estimates from the values being estimated.

One may wonder how firmly based are the conclusions of this study and whether further Monte Carlo experiments are needed to confirm them. Although many experiments were conducted, the number of independent sets of data used and the number of different structures examined were small. The size of the samples of estimates studied, fifty, is modest. If more definite conclusions had been reached about the performances of the estimators, more experiments might be called for. For example, if a definite ranking of the methods had been found, more experiments would be called for to see whether the ranking were peculiar to the data and structures used. However, more experiments would not seem likely to upset the conclusion that the ranking of the estimators does depend on the econometric model being estimated. If the estimators were found to be insensitive to the types of changes investigated, more experiments with much larger samples of data might be needed since the smallness of the sample might have prevented the finding of significant differences in the performances of the estimators. Again, this was not the conclusion. Finally, if the standard errors had always been found reliable, much larger samples would have been useful for checking the results and testing in more detail hypotheses about the distributions of the standard errors.

These arguments do not rule out the usefulness of other Monte Carlo experiments on the small-sample properties of the simultaneous-equation estimators.[189] Many questions were not examined at all. The experiments conducted only scratch the surface of the problems of mis-specification of the statistical model and the forms of the structural equations. Even though it seemed that identification problems may have led to some of the peculiar results, the usefulness of the tests for the identification of an equation was not examined. What is suggested is that it seems unlikely that a repetition of the types of experiment reported in this study would upset the findings of the experiments, though bad luck as well as poor design can have led to the wrong conclusions--as is true for any statistical findings.

The Monte Carlo experiments were designed to investigate two questions: are the performances of the estimators sensitive to the types of data used in different econometric models and would the small-sample properties of the estimators reveal that one method is clearly superior to others for use in estimating econometric models. Both questions have been answered by the study.

[189] And it may be remarked that if I were now undertaking the study *ab initio*, the design of the experiments might be substantially different.

APPENDIX A

The Computations

1. The Computer

The computations for the Monte Carlo experiments used a Control Data Corporation 1604 computer. This is a high-speed, stored-program, electronic computer. It uses binary arithmetic and has over 32,000 words of magnetic core storage. In single-precision, floating-point format, which was used for the computations, numbers have about ten decimal digit accuracy. The absolute size of the numbers can lie in the (approximate) range of 10^{-308} to 10^{308}. The principal input-output medium is magnetic tapes. These also served as auxiliary storage.

The size of core storage led to certain restrictions being placed on the experiments. No more than 28 structural coefficients could be estimated. Not more than 24 reduced-form coefficients could be analyzed. The programs could not handle structures with more than 7 equations or 9 exogenous variables. Predictions of each endogenous variable were limited to three. No more than 70 observations of the variables in each replication could be handled. The number of replications was fixed at 50. These restrictions were imposed to limit the use of magnetic tape as intermediate storage in the programs.

2. The Generation of the Data

Data for the experiments were generated using a random-number generator coded by Mr. G. J. Mitchell. The random-number generator was the particular solution to a difference equation of very large cycle. The pseudo-random numbers produced by it seemed, on investigation, to be random. Different sequences of numbers could be obtained by specifying different initial values to the generator. The generator produced random-bit configurations rather than random numbers. They were converted for use in the study in two ways. For generating the exogenous data they were transformed into rectangularly-distributed floating-point numbers, lying in the range $0 \leq x_i < 100$. A $T \times (K-1)$ matrix of such terms was generated to serve as the exogenous variables in an experiment. (T is the number of observations and K the number of exogenous variables. The other exogenous variable had the value of unity in all observations, being a dummy for the constant of the equations.) For the generation of the structural disturbances, the output of the pseudo-random number generator was converted to floating-point numbers of mean zero. Fifty such numbers were summed and multiplied by an appropriate factor to give approximately normally-distributed deviates of unit variance. A $T \times g$ matrix Q of such deviates was formed (where g is the number of equations). A $g \times g$ matrix, F, was specified by program-parameter 1, which also specified B and Γ. Let h be the value of the scale-factor for the structural disturbances (program-parameter 13). The structural disturbances, U, were then formed:[190]

$$U = hQF.$$

[190] $h^2 F'F = \Sigma$ since $1/T[E(U'U)] = I$ and $1/T[E(hF'Q'QFh)] = h^2 F'F$.

The endogenous variables, Y, were formed using the B- and Γ-matrices specified by program-parameter 1:

$$Y = -Z\Gamma'(B^{-1})' + U(B^{-1})' \quad .$$

The ways in which these basic data were altered to introduce the special features studied were described in the sections presenting the results of the experiments. When other stochastic elements were desired, such as errors in the exogenous variables, they were generated as were Q.

3. Estimation

Having obtained the data Y and Z, estimates were made from them. First, the LSRF estimates were calculated:[191]

$$\hat{\Pi}' = (Z'Z)^{-1}Z'Y \quad .$$

The LSRF residuals, $\hat{V}$, were also calculated:

$$\hat{V} = Y - Z(Z'Z)^{-1}Z'Y \quad .$$

[191] In this appendix we use the notation of Chapter II unless explicit exception is made. Inversion of matrices was performed with a program coded by R. E. Quandt. See R. E. Quandt, "Some Small Sample Properties of Certain Structural Equation Estimators," (Princeton: Econometric Research Program, Research Memorandum 48, 1962), p. 15. The inversion method was designed to work well with ill-conditioned matrices. In each inversion, a check was made to see if the matrix was singular (or rather if it appeared to be singular). If the program declared a matrix to be singular, the entire sample which gave rise to that matrix was abandoned, even though the estimates for some methods could be computed. Singular matrices were found most often in the course of estimating FIML. The results of the experiments do not reflect the fact that other methods could make estimates when FIML could not.

The k-class estimates, DLS, 2SLS, UBK and LIML were all calculated, after forming the matrices of the variables included in each equation, y_m, Y_m, and Z_m, by the formula:

$$\begin{bmatrix} \hat{\beta}_m' \\ \hat{\gamma}_m' \end{bmatrix}_k = \begin{bmatrix} Y_m'Y_m - k\hat{V}_m'\hat{V}_m & Y_m'Z_m \\ Z_m'Y_m & Z_m'Z_m \end{bmatrix}^{-1} \begin{bmatrix} Y_m - k\hat{V}_m & Z_m \end{bmatrix}' y_m \quad .$$

The residuals from each equation were then calculated:

$$(\hat{u}_m)_k = y_m - [Y_m \quad Z_m][\hat{\beta}_m \quad \hat{\gamma}_m]_k' \quad .$$

Standard errors were then calculated as the square roots of the elements on the main diagonal of

$$\frac{1}{T}(\hat{u}_m)_k'(u_m)_k \begin{bmatrix} Y_m'Y_m - kV_m'V_m & Y_m'Z_m \\ Z_m'Y_m & Z_m'Z_m \end{bmatrix}^{-1} \quad .$$

Following the computation of the estimates of the coefficients of each of the structural equations, the estimates of Σ were formed from

$$\hat{U}_k = \{\hat{u}_m\}_k \qquad (m = 1,\ldots,g) \quad ,$$

by

$$\hat{\Sigma}_k = \frac{1}{T}\hat{U}_k'\hat{U}_k \quad .$$

Note that neither the estimates of Σ nor the standard errors were adjusted for degrees of freedom. Next, the matrices $\hat{B}_k$ and $\hat{\Gamma}_k$ were formed by substituting the k-class estimates for the unknown elements of B and Γ. The reduced-form coefficients were then estimated by

$$\hat{\Pi}_k = -\hat{B}_k^{-1}\hat{\Gamma}_k \quad .$$

Three-stage least-squares estimates were calculated from the matrix

$$H = \begin{bmatrix} \sigma^{11}[Y_1 \quad Z_1]'Z(Z'Z)^{-1}Z'[Y_1 \quad Z_1] \ldots \sigma^{1g}[Y_1 \quad Z_1]'Z(Z'Z)^{-1}Z'[Y_g \quad Z_g] \\ \vdots \qquad\qquad\qquad\qquad \vdots \\ \sigma^{g1}[Y_g \quad Z_g]'Z(Z'Z)^{-1}Z'[Y_1 \quad Z_1] \ldots \sigma^{gg}[Y_g \quad Z_g]'Z(Z'Z)^{-1}Z'[Y_g \quad Z_g] \end{bmatrix}$$

and the vector

$$h = \begin{bmatrix} \sum_{i=1}^{g} \sigma^{1i}[Y_1 \quad Z_1]'Z(Z'Z)^{-1}Z'y_i \\ \vdots \\ \sum_{i=1}^{g} \sigma^{gi}[Y_g \quad Z_g]'Z(Z'Z)^{-1}Z'y_i \end{bmatrix} \quad .$$

The estimates were calculated:

$$\hat{\delta} = [\hat{\beta}_1, \hat{\gamma}_1, \ldots, \hat{\beta}_g, \hat{\gamma}_g]$$

$$= H^{-1}h \quad .$$

σ^{ij} is the i,j^{th} element of the inverse of the variance-covariance matrix of the 2SLS structural residuals. The standard errors were then calculated as the square roots of the elements on the main diagonal of H^{-1}.

Having calculated the structural coefficients, the estimates of B and Γ were formed and the estimates of Π were calculated as for the k-class estimates. The estimates of Σ were calculated:

$$\hat{\Sigma} = \frac{1}{T}(Y\hat{B}' + Z\hat{\Gamma}')'(Y\hat{B}' + Z\hat{\Gamma}') \quad .$$

FIML is the most complicated method to estimate. The computations were carried out after the procedures of Chernoff and Divinsky.[192]

The starting point of the computations, $\hat{\delta}_0$, were the 3SLS estimates. This choice was made after it appeared in trial runs that 3SLS tended to perform better than other methods and to be closer to the FIML estimates than those of other consistent methods.

In presenting the FIML estimates we use the following additional notation:

[192] Herman Chernoff and Nathan Divinsky, "The Computation of Maximum-Likelihood Estimates of Linear Structural Equations," in William C. Hood and Tjalling C. Koopmans, Studies in Econometric Method, (New York: Wiley, 1953), pp. 236-302.

$$X = [Y \quad Z]$$

$$c = \sum_{m=1}^{g} (g_m + k_m) \quad .$$

Computations for FIML start by forming the matrices $Q = X'X$, $F = X'Z(Z'Z)^{-1}Z'X$ and the basic matrix, Φ , of dimensions $c\times(K+g)$. Each row of Φ corresponds to a non-zero coefficient in the model and consists of zeroes except for a value of unity in the column corresponding to the variable to which the coefficient applies.[193] Φ can be partitioned into sub-matrices Φ_i corresponding to the separate equations. Sometimes we use only Φ_y , the columns of Φ corresponding to Y.

The matrices $V = \Phi Q\Phi'$ and $N = \Phi F\Phi'$ are formed. These can be partitioned into the $(g_i+h_i)\times(g_j+k_j)$ submatrices:

$$V_{ij} = \Phi_i Q\Phi_j'$$

and

$$N_{ij} = \Phi_i F\Phi_j \quad .$$

For each iteration three steps are involved. First we need to compute the vector of first derivatives of the likelihood function (2.15) at the values of the estimate in the previous iteration: $\hat{\Sigma}(j)$ and $\hat{A}(j) = [\hat{B}(j) \quad \hat{\Gamma}(j)]$. $(\hat{\Sigma}(j) = \frac{1}{T}[\hat{A}(j)X'X\hat{A}(j)'])$ This vector is formed

[193] For a discussion of the basic matrices, see T. C. Koopmans, H. Rubin and R. B. Liepnik, "Measuring the Equation Systems of Dynamic Economics," in T. C. Koopmans (ed.), Statistical Inference in Dynamic Economic Models, (New York: Wiley, 1950), pp. 160-166. In programming FIML, the basic matrices were used since this seemed to make the task easier and to save space in the computer. The simple form of the restrictions on the coefficients in the model, namely only that some are zero *a priori*, accounts for the simple form of Φ.

from two matrices. First the matrices $E = (\hat{B}^{-1})\ \Phi'_y$ and $H = \hat{\Sigma}(j)^{-1}P$ are formed where the $g \times c$ matrix P is formed:

$$P = \{P_{ik}\} = \hat{\delta}(j)_i V_{ik} \quad ,$$

$\hat{\delta}(j)_i$ being the estimates from the j^{th} iteration of the coefficients in the i^{th} equation. Next we obtain the $g \times c$ matrix:

$$M = E - H \ .$$

This can be partitioned into the g^2 vectors M_{ik}, with $(g_k + k_k)$ elements in the k^{th} vectors, according to the equation in which the coefficients corresponding to the columns of M are found. The vector n: $n = \{n_i\} = M_{ii}$ is the required vector of first derivatives of the likelihood function. The vector n is compressed to n_s by deleting elements corresponding to the coefficients on which the equations are normalized.

The next step is to choose a metric for the direction of steepest ascent and the calculation of the matrix T_s corresponding to it. Then the estimate at the end of the $(j+1)^{th}$ iteration is:

$$-\hat{\delta}\ (j+1) = -\hat{\delta}(j) + ed(j+1)$$

where $$d(j+1) = T_s^{-1} n_s$$

and e is the step size used.

Three metrics were used: 1) the "P" metric:

$$T_P\ \{T_{P(ij)}\} = \{(\hat{\Sigma}^{-1})_{ij} V_{ij}\} \ .$$

Υ_{Ps} is formed by deleting the rows and columns corresponding to coefficients on which the equations are normalized; 2) the "R" metric: $\Upsilon_R = \{\Upsilon_{R(ij)}\} = \{(\hat{\Sigma}^{-1})_{ij}N_{ij}$. Υ_{Rs} is formed as Υ_{Ps}; and 3) the "L" or Newtonian metric:

$$\Upsilon_L = -L_1 + L_2 + L_3 - L_4 \quad .$$

To form the four matrices needed, we first form the matrices

$$C = \{C_i\} = (B^{-1})'\Phi'_{iy}$$

$$H = \hat{\Sigma}^{-1}P$$

and

$$G = P'\hat{\Sigma}^{-1}P \quad .$$

C and H are partitioned as was M above. Then the matrices are formed:

$$L_{1(ij)} = \{C_{jk}\}'\{C_{ij}\}$$

$$L_{2(ij)} = \{H_{ji}\}'\{H_{ij}\}$$

$$L_3 = \{L_{3(ij)}\} = \{(\hat{\Sigma}^{-1})_{ij}G_{ij}\}$$

and

$$L_4 = \{L_{4(ij)}\} = \{(\hat{\Sigma}^{-1})_{ij}V_{ij}\} \quad .$$

The only remaining problems are the choices of the step-sizes and criteria for changing metrics and stopping the iterations. We started with the "P" metric. It was decided to take step sizes of unity and then to compute the ratio $q(j) = \frac{[\det B(j)]^2}{\det S(j)}$ which is $1/T^2$ times the square of the variable part of the concentrated likelihood function of the estimates on the j^{th} iteration. As criterion we then took $\ell = \frac{q(j) - q(j-1)}{q(j)}$. If $\ell < .009$, a switch was made to the "R" metric. If ℓ was greater than this threshold value, further unit steps with the same value of d were taken (up to the limit of six steps), until ℓ was found to decrease, whereupon a new iteration with the same metric was taken. The threshold for changing from the "R" to the "L" metric was .0004. Otherwise the procedure was identical with that for the "P" metric. For the "L" metric, step sizes of one only were used (except as noted below). When $\ell < .00003$, the estimate on that iteration was accepted as the FIML estimate.

In the event that the first step led to a negative value of ℓ, a revised procedure for that iteration was used. When this happened, the step size was shortened by 60% and the resulting estimate accepted for that iteration. If this happened repeatedly, all the step sizes were shortened by 40% every time it was found that five decreases in the likelihood function had occurred since the start of the iterations or since the last change in the step size. At the same time, the final threshold was lowered by 20%. It was still found, however, that it was possible with this schema to come very close to a maximum and start "bouncing" about in its neighborhood, without getting a positive increase small enough to satisfy the threshold. In these cases, it was decided to discontinue the iterations after 65 had been completed and accept the estimate on the last iteration as the FIML estimate. Since it was found in examining some of these cases that the differences among the estimates were slight and since one cause of the instability was rounding errors in inverting the large matrix T_L (as well as in other calculations needed), this procedure does not seem unreasonable.

When the iterations were completed, the standard errors were computed as the square roots of the elements on the main diagonal of $\frac{1}{T}(T_L^{-1})$.

4. Checking the Program

The program making the estimates was checked in several ways. First, portions of the program were checked by making hand calculations for several sets of data and comparing them with the calculations made by the program. Second, several small models were calculated using exactly the same procedures as did the program and the computations of the computer program were compared with the calculations at several points in the program. Finally, the Girshick and Haavelmo food model[194] was estimated by the program and the estimates compared with those which have been calculated by others.[195]

[194] M. A. Girshick and Trygve Haavelmo, "Statistical Analysis of the Demand for Food: Examples of Simultaneous Estimation of Structural Equations," in Hood and Koopmans (eds.), op. cit., pp. 92-111.

[195] Harry Eisenpress, "Note on the Computation of Full-Information Maximum Likelihood Estimates of Coefficients of a Simultaneous System," Econometrica, vol. 30, (1962), pp. 343-348.

APPENDIX B

Summary Tables Available

A number of tables was prepared to summarize the findings of the Monte Carlo experiments. Some of these tables are shown in Appendix D. (These tables are indicated by an asterisk in this appendix when they are described.) Others are available on request for the cost of Xerox reproduction.

The tables are arranged in four groups. The experiments contained in each group are shown in Appendix C.

The following types of tables are available for the estimates of the structural coefficients, the estimates of the reduced-form coefficients, the estimates of the elements of Σ,[196] and for the predictions.[197]

A. Tables of Typical Descriptive Statistics. These tables record for each experiment some of the typical descriptive statistics. Let s_{ij} be the descriptive statistic for the distribution of the estimates of the i^{th} parameter, θ_i, made by the j^{th} method. The typical statistics for each method S_j are defined as

$$S_j = \frac{1}{n} \sum_{i=1}^{n} \frac{s_{ij}}{|\theta_i|} .$$

[196]These are not presented for the C-set of tables.

[197]These are not presented for the D-set of tables.

The descriptive statistics for which these tables are available are:

1. Biases of the averages (differences of the averages from the true values)
2. Standard deviations
3. Root-mean-square errors
4. Biases of the medians*
5. Inter-quartile ranges*
6. Ranges
7. Median absolute deviations
8. Third-quartile absolute deviations

B. Statistics for the Ranking of the Estimates by their Absolute Deviations from the True Values. The rankings of the deviations of the estimates from the true values were summarized in two ways:

1. The totals of the ranks of the deviations of the estimates from the true values were summed over all replications and parameters and presented in tables.* These tables also show the values of Kendall's W for the totals and the value of the χ_r^2 statistic for Friedman's two-way analysis of variance. The significance of the χ_r^2 statistic is shown with the following coding:

0 - not significant at the .05 level

1 - significant at the .05 level but not at the .01 level

2 - significant beyond the .01 level.

The reasons for not placing emphasis on the findings of the χ_r^2 test for these totals of ranks were discussed in Chapter III, Section 11. The tables also record the number of separate coefficients for which the totals of the ranks of the estimates of the coefficients led to values of W which were significantly different from zero at the .05 level (referred to as "X95" in the tables) and at the .01 level ("X99").

2. The totals of the ranks of the estimates of each parameter were themselves ranked. These ranks were then summed over the parameters estimated. Summary tables containing these rank-totals are available.

C. Ranking of the Descriptive Statistics. The descriptive statistics of the estimates of the parameters were ranked. The totals of these ranks over the parameters estimated are available for

1. The standard deviations
2. The root-mean-square errors
3. The inter-quartile ranges
4. The ranges
5. The median absolute deviations
6. The third-quartile absolute deviations.

D. Counting Statistics. These tables show the number of times in the estimates of each experiment certain conditions arose. The tables are:

1. The total number of times each method had the estimate which was closest to the true values of the parameters being estimated,

2. The total number of times each method had the estimate which was farthest from the true values,

3. The number of times the estimates had the wrong signs,

4. The number of times the medians of the estimates of each method were significantly different from the true values being estimated,

5. The number of estimates falling within 20% of the true values.

E. Tables of Pairwise Comparisons of Methods. Some tables were prepared to show the performances of some of the methods relative to others. The methods compared were DLS with 2SLS, 2SLS with LIML, FIML with 3SLS and, where appropriate, DLS with LSRF. The tables show:

1. The number of times one method had estimates closer than the other for all the parameters estimated,

2. The number of parameters for which the sign test found that one method had significantly smaller dispersions than the other,

3. The number of parameters for which the Wilcoxon test found one method to have significantly smaller dispersions than the other,

4. The number of times the Wilcoxon test found that one method had significantly smaller central tendencies than the other,

5. The number of times the estimates of one method had a smaller descriptive statistic than those of the other. These are available for

a. The inter-quartile ranges,

b. The ranges,

c. The median absolute deviations,

d. The third-quartile absolute deviations.

Tables for the standard errors listed the proportion of the t-ratios for the estimates of all structural coefficients which fell outside the intervals studied. The intervals were:

1. The 95% confidence interval of the t-distribution with degrees of freedom equal to the number of observations minus the total number of structural coefficients to be estimated.* (This is referred to in the tables as "MODEL T". When this led to a negative number for the degrees of freedom, all the estimates were considered to lie outside the interval, since the hypothesis was meaningless.)

2. The 95% confidence interval for the t-distribution with degrees of freedom equal to the number of observations minus the number of coefficients to be estimated in the equation in which the coefficient for which the t-ratio was formed lay.

3. The interval -2 to +2.

APPENDIX C

The Experiments Conducted

This appendix lists the Monte Carlo experiments conducted. The experiments were divided into four parts. The tables of Appendix D follow the arrangement of experiments in the present appendix. To facilitate the comparison of the experiments, the first one in each set is the basic experiments (discussed in Chapter IV). The experiments were discussed in the study as follows:

Chapter V	-	Part 1, Experiments B-AE
Chapter VI	-	Part 1, Experiments AF-AK Part 2
Chapter VII	-	Part 3
Chapter VIII	-	Part 4.

The experiments are defined by the values of the program-parameters. The program-parameters and their effects were:

Program-Parameter	Description	Discussed in
1	Structure used	Chapter V, Section 4 Section 5 Section 6 Chapter VIII, Section 6 Section 9
2	Number of observations	Chapter VI, Section 2
3	Errors in the exogenous variables	Chapter VII, Section 2
4	Lagged-endogenous variables	Chapter VI, Section 4

Program-Parameter	Description	Discussed in
5	Time-structures in the exogenous variables	Chapter VI, Section 5
6	Multicollinearity	Chapter VI, Section 3
7	Proportionate errors	Chapter VII, Section 5
8	Stochastic coefficients	Chapter VII, Section 3
9	Auto-correlated disturbances	Chapter VII, Section 4
10	Misspecification	Chapter VIII
11	Not used	
12	Heteroskedastic disturbances	Chapter VII, Section 5
13	Disturbance size-factor	Chapter V, Section 7
14	Disturbance set	Chapter V, Section 2
15	Exogenous variables set	Chapter V, Section 3
16	Scale-factor for errors in the exogenous variables	Chapter VII, Section 2
17	Scale-factor for the stochastic additions to the structural coefficients.	Chapter VII, Section 3

In the tables of this appendix and the next one, experiments are referred to as "runs" to save space in the table headings.

Appendix C
Table 1

PARAMETERS OF THE RUNS

RUN	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
A	MOD37A2	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
B	MOD37A2	20	0	0	0	0	0	0	0	0	0	0	2	1	0	2	2
C	MOD37A2	20	0	0	0	0	0	0	0	0	0	0	2	2	0	2	2
D	MOD37A2	20	0	0	0	0	0	0	0	0	0	0	2	0	1	2	2
E	MOD37A2	20	0	0	0	0	0	0	0	0	0	0	2	0	2	2	2
F	MOD37A1	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
G	MOD37A3	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
H	MOD37A4	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
I	MOD37A5	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
J	MOD37AA	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
K	MOD37AB	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
L	MOD37AC	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
M	MOD37AD	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
N	MOD37AE	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
O	MODEL37B	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
P	MODEL37C	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
Q	MODEL37D	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
R	MODEL37E	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
S	MODEL5A	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
T	MODEL46	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
U	MODEL48	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
V	MODEL56	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
W	MODEL58	20	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
X	MOD37A2	20	0	0	0	0	0	0	0	0	0	0	1	0	0	2	2
Y	MOD37A2	20	0	0	0	0	0	0	0	0	0	0	4	0	0	2	2
Z	MOD37A2	20	0	0	0	0	0	0	0	0	0	0	6	0	0	2	2
AA	MOD37A2	20	0	0	0	0	0	0	0	0	0	0	8	0	0	2	2
AB	MOD37A2	20	0	0	0	0	0	0	0	0	0	0	10	0	0	2	2
AC	MOD37AB	20	0	0	0	0	0	0	0	0	0	0	8	0	0	2	2
AD	MODEL37D	20	0	0	0	0	0	0	0	0	0	0	8	0	0	2	2
AE	MODEL46	20	0	0	0	0	0	0	0	0	0	0	8	0	0	2	2
AF	MOD37A2	35	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
AG	MOD37A2	50	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
AH	MOD37A2	70	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
AI	MOD37AB	70	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
AJ	MODEL37E	70	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2
AK	MODEL48	70	0	0	0	0	0	0	0	0	0	0	2	0	0	2	2

Appendix C
Table 2

PARAMETERS OF THE RUNS

RUN	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
A	MOD37A2	20	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
B	MOD37A2	20	Ø	Ø	Ø	2	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
C	MOD37A2	20	Ø	Ø	Ø	3	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
D	MOD37A2	20	Ø	Ø	Ø	4	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
E	MOD37A2	20	Ø	Ø	Ø	5	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
F	MOD37A2	20	Ø	Ø	Ø	6	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
G	MOD37A2	20	Ø	Ø	Ø	7	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
H	MODEL46	20	Ø	Ø	Ø	3	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
I	MODEL37D	20	Ø	Ø	Ø	6	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
J	MOD37A5	20	Ø	Ø	Ø	7	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
K	MOD37AB	20	Ø	Ø	Ø	7	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
L	MOD37A2	20	Ø	Ø	Ø	5	Ø	Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	2	2
M	MOD37A2	70	Ø	Ø	Ø	5	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
N	MODEL58	35	Ø	Ø	Ø	5	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
O	MOD37A2	20	Ø	1	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
P	MOD37A2	20	Ø	2	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
Q	MOD37A4	20	Ø	2	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
R	MODEL37B	20	Ø	2	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
S	MOD37A2	20	Ø	2	Ø	5	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
T	MOD37A2	20	Ø	2	Ø	6	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
U	MOD37A4	20	Ø	3	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
V	MOD37A4	70	Ø	3	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
W	MOD37AE	20	Ø	3	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
X	MOD37A2	20	Ø	Ø	1	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
Y	MOD37A2	20	Ø	Ø	2	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
Z	MODEL56	35	Ø	Ø	2	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AA	MOD37A2	20	Ø	Ø	3	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AB	MODEL46	20	Ø	Ø	3	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AC	MOD37A2	20	Ø	2	3	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AD	MOD37A2	20	Ø	2	3	6	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AE	MOD37A4	20	Ø	3	3	6	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AF	MOD37AE	20	Ø	3	3	6	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AG	MOD37A2	20	Ø	Ø	5	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AH	MODEL37D	20	Ø	Ø	5	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AI	MOD37A2	20	Ø	2	5	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AJ	MOD37A2	20	Ø	Ø	6	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AK	MOD37A2	20	Ø	Ø	7	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AL	MODEL48	20	Ø	Ø	7	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AM	MOD37A2	20	Ø	2	7	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
AN	MOD37A2	20	Ø	2	7	6	Ø	Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	2	2
AO	MOD37A4	20	Ø	3	7	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2

Appendix C
Table 3

PARAMETERS OF THE RUNS

RUN	1	2	3	4	5	6	7	8	9	1Ø	11	12	13	14	15	16	17
A	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
B	MOD37A2	2Ø	1	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	1	2
C	MOD37A2	2Ø	1	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
D	MOD37A2	2Ø	1	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	4	2
E	MOD37A2	2Ø	1	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	8	2
F	MOD37A2	2Ø	1	Ø	Ø	5	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
G	MOD37A1	2Ø	1	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
H	MOD37AB	2Ø	1	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
I	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
J	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	Ø	Ø	2	Ø	Ø	2	4
K	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	Ø	Ø	2	Ø	Ø	2	8
L	MOD37AB	2Ø	Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
M	MODEL37E	2Ø	Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
N	MOD37A2	2Ø	1	Ø	Ø	Ø	Ø	1	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
O	MOD37A2	2Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
P	MOD37A2	2Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	2	2
Q	MODEL37B	2Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
R	MOD37A5	2Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
S	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	Ø	2	Ø	Ø	2	2
T	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	Ø	2	Ø	Ø	2	2
U	MOD37A3	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	Ø	2	Ø	Ø	2	2
V	MODEL37D	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	Ø	2	Ø	Ø	2	2
W	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	3	Ø	Ø	Ø	2	Ø	Ø	2	2
X	MODEL46	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	3	Ø	Ø	Ø	2	Ø	Ø	2	2
Y	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	1	2	Ø	Ø	2	2
Z	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	2	Ø	Ø	2	2
AA	MOD37A2	2Ø	Ø	Ø	Ø	5	Ø	Ø	Ø	Ø	Ø	2	2	Ø	Ø	2	2
AB	MODEL37C	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	2	Ø	Ø	2	2
AC	MOD37A2	2Ø	Ø	Ø	Ø	Ø	1	Ø	2	Ø	Ø	Ø	2	Ø	Ø	2	2
AD	MOD37A2	2Ø	Ø	Ø	2	Ø	Ø	Ø	2	Ø	Ø	Ø	2	Ø	Ø	2	2
AE	MOD37A2	2Ø	Ø	Ø	2	Ø	Ø	Ø	Ø	Ø	Ø	2	2	Ø	Ø	2	2
AF	MOD37A2	2Ø	Ø	Ø	3	Ø	Ø	Ø	2	Ø	Ø	Ø	2	Ø	Ø	2	2
AG	MOD37A2	2Ø	Ø	Ø	3	Ø	Ø	Ø	3	Ø	Ø	Ø	2	Ø	Ø	2	2
AH	MOD37A2	2Ø	Ø	Ø	7	Ø	Ø	Ø	3	Ø	Ø	Ø	2	Ø	Ø	2	2
AI	MODEL46	2Ø	Ø	Ø	7	Ø	Ø	Ø	3	Ø	Ø	Ø	2	Ø	Ø	2	2
AJ	MOD37A2	2Ø	Ø	Ø	7	Ø	Ø	Ø	3	Ø	Ø	2	2	Ø	Ø	2	2
AK	MOD37AB	2Ø	Ø	Ø	7	Ø	Ø	Ø	3	Ø	Ø	2	2	Ø	Ø	2	2
AL	MOD37A2	2Ø	Ø	2	Ø	Ø	Ø	Ø	2	Ø	Ø	Ø	2	Ø	Ø	2	2
AM	MOD37A2	2Ø	Ø	2	3	Ø	Ø	Ø	3	Ø	Ø	Ø	2	Ø	Ø	2	2
AN	MOD37A2	2Ø	Ø	2	5	Ø	Ø	Ø	3	Ø	Ø	Ø	2	Ø	Ø	2	2
AO	MOD37A2	2Ø	Ø	2	7	Ø	Ø	Ø	3	Ø	Ø	Ø	2	Ø	Ø	2	2
AP	MOD37A2	2Ø	Ø	2	7	Ø	Ø	Ø	3	Ø	Ø	2	2	Ø	Ø	2	2
AQ	MOD37A4	2Ø	Ø	3	7	Ø	Ø	Ø	3	Ø	Ø	2	2	Ø	Ø	2	2
AR	MOD37AE	2Ø	Ø	3	7	Ø	Ø	Ø	3	Ø	Ø	2	2	Ø	Ø	2	2
AS	MOD37AB	2Ø	Ø	2	7	Ø	Ø	Ø	3	Ø	Ø	2	2	Ø	Ø	2	2
AT	MOD37A2	2Ø	1	2	3	5	Ø	Ø	2	Ø	Ø	2	2	Ø	Ø	2	2
AU	MOD37A2	2Ø	Ø	2	3	5	Ø	Ø	2	Ø	Ø	Ø	2	Ø	Ø	2	2
AV	MODEL46	2Ø	1	2	3	5	Ø	Ø	2	Ø	Ø	Ø	2	Ø	Ø	2	2
AW	MOD37A2	2Ø	Ø	2	7	6	Ø	Ø	3	Ø	Ø	2	2	Ø	Ø	2	2

Appendix C
Table 4

PARAMETERS OF THE RUNS

RUN	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
A	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	2
B	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	2	Ø	Ø	2	2
C	MOD37A4	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	2	Ø	Ø	2	2
D	MODEL48	35	Ø	Ø	Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	2	Ø	Ø	2	2
E	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	Ø	Ø	2	2
F	MOD37A1	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	Ø	Ø	2	Ø	Ø	2	2
G	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	3	Ø	Ø	2	Ø	Ø	2	2
H	MOD37A1	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	3	Ø	Ø	2	Ø	Ø	2	2
I	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	4	Ø	Ø	2	Ø	Ø	2	2
J	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	4	Ø	Ø	2	Ø	1	2	2
K	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	4	Ø	Ø	1	Ø	Ø	2	2
L	MOD37A2	7Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	4	Ø	Ø	2	Ø	Ø	2	2
M	MOD37A1	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	4	Ø	Ø	2	Ø	Ø	2	2
N	MOD37A3	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	4	Ø	Ø	2	Ø	Ø	2	2
O	MOD37AB	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	4	Ø	Ø	2	Ø	Ø	2	2
P	MODEL37C	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	4	Ø	Ø	2	Ø	Ø	2	2
Q	MODEL37D	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	4	Ø	Ø	2	Ø	Ø	2	2
R	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	6	Ø	Ø	2	Ø	Ø	2	2
S	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	7	Ø	Ø	2	Ø	Ø	2	2
T	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	8	Ø	Ø	2	Ø	Ø	2	2
U	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	9	Ø	Ø	2	Ø	Ø	2	2
V	MODNOID	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	1Ø	Ø	Ø	2	Ø	Ø	2	2
W	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	11	Ø	Ø	2	Ø	Ø	2	2
X	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	12	Ø	Ø	2	Ø	Ø	2	2
Y	MOSPEC1	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	13	Ø	Ø	2	Ø	Ø	2	2
Z	MODEL46	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	13	Ø	Ø	2	Ø	Ø	2	2
AA	MOSPEC2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	14	Ø	Ø	2	Ø	Ø	2	2
AB	MODEL46	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	14	Ø	Ø	2	Ø	Ø	2	2
AC	MOD37A2	2Ø	Ø	2	Ø	Ø	Ø	Ø	Ø	1	Ø	Ø	2	Ø	Ø	2	2
AD	MOD37A2	2Ø	Ø	Ø	3	Ø	Ø	Ø	Ø	1	Ø	Ø	2	Ø	Ø	2	2
AE	MOD37A2	2Ø	Ø	Ø	5	Ø	Ø	Ø	Ø	1	Ø	Ø	2	Ø	Ø	2	2
AF	MOD37A2	2Ø	Ø	2	2	Ø	Ø	Ø	1	1	Ø	Ø	2	Ø	Ø	2	2
AG	MOD37A2	2Ø	Ø	Ø	1	Ø	Ø	Ø	Ø	1	Ø	2	2	Ø	Ø	2	2
AH	MOD37A2	2Ø	Ø	Ø	Ø	3	Ø	Ø	Ø	1	Ø	Ø	2	Ø	Ø	2	2
AI	MOD37A2	2Ø	Ø	Ø	Ø	5	Ø	Ø	Ø	1	Ø	Ø	2	Ø	Ø	2	2
AJ	MOD37A2	2Ø	Ø	Ø	Ø	7	Ø	Ø	Ø	1	Ø	Ø	2	Ø	Ø	2	2
AK	MOD37A2	2Ø	Ø	Ø	Ø	5	Ø	Ø	Ø	2	Ø	Ø	2	Ø	Ø	2	2
AL	MOD37A2	2Ø	Ø	Ø	Ø	5	Ø	Ø	Ø	6	Ø	Ø	2	Ø	Ø	2	2
AM	MOD37A2	7Ø	Ø	Ø	Ø	5	Ø	Ø	Ø	6	Ø	Ø	2	Ø	Ø	2	2
AN	MOD37A2	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	6	Ø	Ø	2	Ø	Ø	2	2
AO	MOD37A2	2Ø	Ø	Ø	3	5	Ø	Ø	Ø	6	Ø	Ø	2	Ø	Ø	2	2
AP	MOD37A2	2Ø	Ø	Ø	Ø	5	Ø	Ø	Ø	9	Ø	Ø	2	Ø	Ø	2	2
AQ	MODNOID	2Ø	Ø	Ø	Ø	5	Ø	Ø	Ø	1Ø	Ø	Ø	2	Ø	Ø	2	2
AR	MODNOID	2Ø	Ø	2	Ø	Ø	Ø	Ø	Ø	1Ø	Ø	Ø	2	Ø	Ø	2	2
AS	MODNOID	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	2	1Ø	Ø	Ø	2	Ø	Ø	2	2
AT	MODNOID	2Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	1Ø	Ø	2	2	Ø	Ø	2	2
AU	MOSPEC1	2Ø	Ø	2	3	5	Ø	Ø	2	13	Ø	2	2	Ø	Ø	2	2
AV	MOSPEC2	2Ø	Ø	2	3	5	Ø	Ø	2	14	Ø	2	2	Ø	Ø	2	2

APPENDIX D

Summary Statistics for the Experiments

This appendix contains tables to illustrate the orders of magnitude of the estimates and the rankings of the estimators found in the course of the Monte Carlo experiments. The statistics are described in Appendix B. The four sets of tables follow the division of the experiments into four parts in Appendix C. The letters identifying the statistics correspond to those given to the experiments in Appendix C.

APPENDIX D

TABLE A-1

ESTIMATED STRUCTURAL COEFFICIENTS

TYPICAL MEDIAN BIASES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	-.13	-.00	.01	.02	.00	.00
B	-.19	-.05	-.05	-.02	-.02	-.01
C	-.17	-.01	.00	.02	-.00	.01
D	-.37	-.07	-.05	-.03	-.04	.01
E	-.20	-.06	-.05	-.03	-.03	-.02
F	-.26	-.09	-.07	-.02	-.07	-.02
G	-.01	.02	.02	.02	-.00	-.00
H	-.10	-.01	.00	.02	-.02	.00
I	-.18	-.09	-.08	-.04	-.07	-.02
J	-.12	-.03	-.02	-.01	-.01	.01
K	-.11	-.02	-.01	-.01	-.02	-.01
L	-.13	-.05	-.05	-.02	-.03	-.01
M	-.07	-.00	.01	.02	-.01	.01
N	-.02	.01	.02	.02	.01	.01
O	-.25	-.06	-.04	-.02	-.06	-.00
P	-.07	-.02	-.01	.01	-.01	-.01
Q	-.10	-.03	-.02	-.01	-.03	-.01
R	-.01	-.01	-.01	-.01	-.00	-.01
S	.00	-.00	-.00	.01	-.01	-.02
T	-.09	-.01	-.01	.00	-.02	-.00
U	-.00	.01	.01	.01	.01	.02
V	-.15	-.02	-.02	-.01	-.01	.02
W	-.03	-.00	-.00	-.01	-.01	-.00
X	-.04	.00	.00	.01	.01	.00
Y	-.38	-.07	-.04	-.00	-.00	.02
Z	-.50	-.19	-.12	-.02	-.14	.03
AA	-.59	-.33	-.27	-.12	-.31	-.25
AB	-.61	-.39	-.34	-.13	-.35	-.19
AC	-.39	-.17	-.23	-.13	-.17	-.10
AD	-.39	-.24	-.21	-.08	-.24	-.07
AE	-.17	-.13	-.14	-.09	-.10	-.09
AF	-.17	-.03	-.02	-.01	-.02	-.01
AG	-.19	-.02	-.01	-.01	-.01	-.00
AH	-.20	-.03	-.02	-.02	-.01	-.00
AI	-.13	-.02	-.02	-.02	-.02	-.01
AJ	-.00	-.00	-.00	.00	.01	.01
AK	-.00	-.00	-.00	-.00	-.01	-.01

APPENDIX D

TABLE A-2
ESTIMATED STRUCTURAL COEFFICIENTS

TYPICAL INTERQUARTILE RANGE

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	.39	.44	.44	.45	.29	.29
B	.30	.39	.40	.40	.27	.27
C	.38	.44	.46	.46	.28	.26
D	.33	.43	.45	.43	.27	.27
E	.32	.37	.38	.39	.26	.25
F	.21	.31	.33	.35	.29	.28
G	.34	.47	.47	.45	.23	.24
H	.27	.34	.36	.40	.30	.35
I	.19	.31	.36	.41	.29	.43
J	.38	.43	.44	.45	.39	.41
K	.38	.44	.44	.46	.45	.46
L	.37	.41	.42	.43	.45	.46
M	.31	.40	.41	.44	.38	.44
N	.35	.39	.39	.41	.42	.43
O	.17	.29	.30	.32	.24	.28
P	.24	.40	.43	.47	.36	.40
Q	.16	.21	.22	.23	.15	.15
R	.23	.30	.31	.32	.26	.27
S	.27	.40	.40	.46	.32	.32
T	.27	.35	.35	.39	.30	.31
U	.29	.31	.31	.32	.29	.28
V	.35	.41	.42	.42	.37	.37
W	.32	.40	.41	.41	.33	.35
X	.22	.22	.22	.22	.14	.14
Y	.50	.82	.87	1.00	.56	.58
Z	.55	1.07	1.26	1.58	.74	1.01
AA	.58	1.34	1.57	2.09	1.12	1.46
AB	.61	1.32	1.56	2.40	1.14	1.88
AC	1.03	1.51	1.70	1.94	1.50	1.93
AD	.26	.43	.60	.90	.36	.58
AE	.53	.89	.95	1.33	.81	1.15
AF	.21	.28	.29	.30	.19	.19
AG	.17	.25	.26	.25	.15	.14
AH	.17	.19	.19	.19	.11	.11
AI	.17	.19	.19	.19	.20	.20
AJ	.12	.14	.14	.14	.11	.11
AK	.14	.16	.17	.17	.15	.15

APPENDIX D

TABLE A-3
ESTIMATED STRUCTURAL COEFFICIENTS

RANKINGS ON ALL ESTIMATES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	CHI SQ	W	SIG	X95	X99
A	3334.	2751.	2698.	2771.	2078.	2118.	419.6	.11	2	15	14
B	3514.	2708.	2649.	2652.	2063.	2164.	505.5	.13	2	14	14
C	3388.	2734.	2801.	2806.	2040.	1981.	539.0	.14	2	14	14
D	3662.	2728.	2684.	2670.	1995.	2011.	710.6	.19	2	12	11
E	3379.	2743.	2736.	2783.	2034.	2075.	484.4	.13	2	13	12
F	3538.	2731.	2519.	2601.	2258.	2103.	481.4	.13	2	14	13
G	3074.	2768.	2771.	2847.	2103.	2187.	288.4	.08	2	12	11
H	3111.	2701.	2759.	2793.	2223.	2163.	252.6	.07	2	14	13
I	3315.	2618.	2479.	2554.	2443.	2341.	234.8	.06	2	15	13
J	3108.	2656.	2600.	2645.	2356.	2385.	139.1	.04	2	9	9
K	2877.	2480.	2488.	2565.	2621.	2719.	44.1	.01	2	5	4
L	2929.	2494.	2484.	2505.	2656.	2682.	56.4	.02	2	7	6
M	2838.	2559.	2664.	2756.	2425.	2508.	46.5	.01	2	8	6
N	2614.	2455.	2507.	2649.	2722.	2803.	32.2	.01	2	7	5
O	3644.	2675.	2558.	2572.	2176.	2125.	571.3	.15	2	14	14
P	3518.	2499.	2498.	2557.	2328.	2350.	380.2	.10	2	14	14
Q	2733.	2295.	2205.	2195.	1657.	1515.	474.9	.16	2	11	10
R	2819.	2450.	2555.	2615.	2145.	2116.	154.7	.04	2	11	11
S	2942.	2506.	2554.	2712.	1979.	2007.	303.2	.09	2	12	11
T	3987.	2973.	2996.	2948.	2477.	2469.	514.1	.12	2	15	15
U	4454.	3922.	3911.	3913.	3444.	3456.	181.2	.03	2	13	9
V	3945.	3583.	3645.	3734.	2951.	3142.	202.9	.04	2	14	11
W	5001.	4534.	4748.	4896.	3919.	4202.	193.8	.03	2	13	9
X	3177.	2787.	2757.	2800.	2092.	2137.	343.3	.09	2	12	11
Y	3604.	2739.	2572.	2662.	2087.	2086.	592.6	.16	2	15	14
Z	3393.	2679.	2572.	2748.	2129.	2229.	386.1	.10	2	14	13
AA	3186.	2561.	2490.	2659.	2146.	2708.	218.9	.06	2	13	13
AB	3056.	2528.	2487.	2708.	2214.	2757.	155.2	.04	2	14	13
AC	2635.	2294.	2629.	2745.	2550.	2897.	77.6	.02	2	10	10
AD	2501.	2057.	2103.	2246.	2050.	1643.	188.2	.06	2	12	12
AE	3608.	2798.	2728.	2928.	2808.	2980.	175.8	.04	2	15	14
AF	3661.	2644.	2665.	2587.	2094.	2099.	623.0	.17	2	14	14
AG	3718.	2581.	2652.	2651.	2077.	2071.	687.7	.18	2	14	14
AH	3889.	2639.	2605.	2630.	2028.	1959.	913.6	.24	2	15	15
AI	3363.	2553.	2509.	2508.	2377.	2440.	256.3	.07	2	8	8
AJ	3387.	2464.	2510.	2417.	2009.	1913.	557.4	.16	2	12	12
AK	5074.	3752.	3823.	3771.	3404.	3276.	530.7	.10	2	18	18

APPENDIX D

TABLE A-4
ESTIMATED STANDARD ERRORS

PROPORTION OF BIAS/S.E. OUTSIDE MODEL T

RUN	OLS	2SLS	UBK	LIML	3SLS	FIML
A	.20	.07	.06	.06	.05	.04
B	.18	.03	.03	.03	.04	.04
C	.18	.05	.05	.05	.04	.02
D	.32	.07	.06	.05	.05	.04
E	.18	.04	.03	.03	.03	.03
F	.43	.11	.09	.06	.13	.07
G	.10	.05	.05	.05	.04	.05
H	.24	.07	.07	.06	.06	.04
I	.42	.21	.16	.08	.23	.10
J	.14	.04	.04	.04	.05	.05
K	.07	.04	.04	.04	.05	.05
L	.07	.05	.04	.05	.06	.05
M	.12	.07	.07	.07	.07	.07
N	.07	.06	.06	.06	.07	.07
O	.36	.08	.08	.06	.08	.06
P	.30	.06	.05	.03	.06	.03
Q	.29	.10	.10	.07	.08	.08
R	.14	.06	.06	.06	.06	.06
S	.24	.07	.07	.06	.06	.05
T	.16	.02	.02	.01	.02	.01
U	1.00	1.00	1.00	1.00	1.00	1.00
V	1.00	1.00	1.00	1.00	1.00	1.00
W	1.00	1.00	1.00	1.00	1.00	1.00
X	.08	.06	.05	.06	.04	.04
Y	.40	.08	.07	.05	.08	.04
Z	.52	.13	.10	.06	.12	.10
AA	.57	.16	.13	.08	.17	.16
AB	.57	.19	.16	.09	.21	.18
AC	.21	.05	.05	.04	.09	.08
AD	.52	.27	.22	.10	.31	.17
AE	.48	.10	.08	.04	.14	.09
AF	.37	.05	.05	.05	.07	.07
AG	.49	.06	.05	.05	.06	.06
AH	.56	.05	.05	.05	.05	.05
AI	.35	.07	.07	.07	.07	.07
AJ	.48	.04	.04	.04	.05	.04
AK	.29	.04	.04	.04	.05	.05

APPENDIX D

TABLE A-5
ESTIMATED RESIDUALS

TYPICAL MEDIAN BIASES

RUN	OLS	2SLS	UBK	LIML	3SLS	FIML
A	-.49	-.36	-.35	-.33	-.28	-.21
B	-.46	-.33	-.32	-.29	-.25	-.20
C	-.42	-.27	-.25	-.22	-.19	-.12
D	-.50	-.35	-.34	-.31	-.23	-.16
E	-.47	-.34	-.33	-.30	-.24	-.18
F	-.64	-.48	-.45	-.35	-.39	-.21
G	-.41	-.33	-.33	-.31	-.28	-.22
H	-.55	-.40	-.38	-.34	-.33	-.22
I	-.73	-.58	-.56	-.42	-.50	-.26
J	-.44	-.34	-.33	-.31	-.27	-.23
K	-.85	-.26	-.24	-.20	-.23	-.15
L	-.31	-.28	-.27	-.24	-.27	-.28
M	-.15	-.38	-.36	-.32	-.33	-.24
N	-.49	-.30	-.30	-.26	-.28	-.26
O	-.56	-.41	-.39	-.35	-.34	-.24
P	-.58	-.42	-.40	-.34	-.37	-.28
Q	-.46	-.35	-.33	-.29	-.26	-.12
R	-.43	-.34	-.32	-.30	-.28	-.21
S	-.48	-.29	-.28	-.25	-.24	-.15
T	-.59	-.40	-.39	-.36	-.37	-.18
U	.44	-.37	-.36	-.34	-.32	-.28
V	-.44	-.29	-.27	-.24	-.28	-.22
W	-.49	-.38	-.37	-.34	-.28	-.16
X	-.37	-.33	-.33	-.32	-.25	-.20
Y	-.69	-.46	-.42	-.35	-.39	-.23
Z	-.27	-.57	-.52	-.36	-.52	-.18
AA	-.86	-.69	-.63	-.42	-.64	-.25
AB	-.89	-.75	-.70	-.46	-.73	-.30
AC	-1.41	-.73	-.99	-.35	-.77	-.38
AD	-.87	-.77	-.71	-.36	-.76	-.22
AE	-.82	-.74	-.74	-.58	-.74	-.29
AF	-.37	-.22	-.21	-.20	-.17	-.14
AG	-.29	-.08	-.07	-.06	-.06	-.05
AH	-.31	-.10	-.09	-.09	-.06	-.04
AI	.09	.12	.11	.12	.05	.07
AJ	-.28	-.09	-.08	-.07	-.06	-.05
AK	-.29	-.12	-.11	-.10	-.10	-.88

APPENDIX D

TABLE A-6
ESTIMATED RESIDUALS

TYPICAL INTERQUARTILE RANGE

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	.27	.41	.43	.45	.40	.45
B	.27	.39	.41	.43	.39	.44
C	.27	.45	.47	.45	.45	.46
D	.23	.39	.41	.43	.32	.41
E	.24	.37	.39	.41	.40	.43
F	.17	.36	.43	.52	.39	.49
G	.36	.43	.44	.47	.44	.44
H	.25	.39	.42	.44	.39	.49
I	.15	.31	.40	.51	.35	.79
J	.38	.44	.45	.47	.52	.57
K	1.84	2.04	2.19	2.23	2.30	2.23
L	1.12	1.25	1.28	1.35	1.47	1.53
M	.55	.47	.48	.55	.53	.58
N	1.85	2.00	2.01	2.09	2.19	2.59
O	.25	.44	.47	.49	.41	.51
P	.19	.46	.51	.60	.46	.64
Q	.30	.43	.47	.49	.45	.49
R	.29	.41	.43	.44	.47	.52
S	.27	.41	.41	.46	.47	.58
T	.65	1.10	.90	1.06	1.00	1.12
U	.45	.54	.56	.57	.43	.60
V	.85	1.00	1.02	1.06	1.09	1.20
W	.78	.97	.99	1.05	1.29	1.26
X	.32	.35	.36	.37	.37	.36
Y	.16	.43	.49	.65	.49	.63
Z	.14	.33	.44	.69	.39	1.09
AA	.08	.28	.38	.71	.31	3.34
AB	.06	.24	.47	.82	.27	3.53
AC	1.27	3.51	4.69	6.26	3.79	6.26
AD	.08	.21	.39	1.37	.23	1.21
AE	.19	.72	.85	1.80	.73	5.08
AF	.21	.31	.31	.32	.30	.31
AG	.20	.30	.31	.32	.32	.32
AH	.19	.29	.32	.30	.30	.30
AI	1.34	1.71	1.72	1.76	1.88	1.89
AJ	.20	.32	.32	.32	.31	.34
AK	.30	.49	.49	.50	.41	.41

APPENDIX D

TABLE A-7
ESTIMATED RESIDUALS

RANKINGS ON ALL ESTIMATES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	CHI SQ	W	SIG	X95	X99
A	1631.	1269.	1028.	896.	819.	657.	588.1	.39	2	6	6
B	1622.	1261.	1005.	931.	794.	687.	557.3	.37	2	6	6
C	1590.	1199.	1039.	895.	818.	759.	453.8	.30	2	6	6
D	1673.	1287.	1056.	917.	748.	619.	703.8	.47	2	6	6
E	1630.	1279.	1050.	898.	769.	674.	602.2	.40	2	6	6
F	1616.	1273.	1021.	918.	801.	671.	565.7	.38	2	6	6
G	1574.	1275.	1059.	938.	788.	666.	527.5	.35	2	6	6
H	1645.	1258.	1031.	889.	822.	655.	601.5	.40	2	6	6
I	1595.	1272.	1027.	846.	801.	759.	509.7	.34	2	6	6
J	1500.	1216.	1018.	894.	898.	774.	337.8	.23	2	6	6
K	1320.	1064.	931.	877.	107[illegible].	1038.	112.1	.07	2	6	6
L	1313.	1108.	989.	907.	1039.	944.	102.9	.07	2	5	5
M	1555.	1217.	1018.	879.	899.	732.	416.3	.28	2	6	6
N	1201.	1054.	936.	937.	1112.	1063.	50.[illegible]	.03	2	6	6
O	1657.	1222.	1305.	841.	875.	702.	568.4	.38	2	6	6
P	1580.	1253.	1036.	845.	849.	737.	478.8	.32	2	6	6
Q	1562.	1245.	1039.	925.	802.	727.	458.8	.31	2	6	6
R	1559.	1269.	1043.	912.	811.	706.	477.7	.32	2	6	6
S	1582.	1161.	1049.	859.	850.	799.	414.1	.28	2	6	6
T	2534.	1945.	1757.	1531.	1412.	1321.	570.8	.23	2	10	1[illegible]
U	2520.	2094.	1774.	1567.	1343.	1202.	692.1	.28	2	7	7
V	3344.	2376.	2626.	2536.	2211.	2157.	372.7	.10	2	14	12
W	3239.	2821.	2653.	2611.	2176.	2250.	289.[illegible]	.18	2	12	12
X	1603.	1333.	1074.	935.	757.	598.	657.8	.44	2	6	6
Y	1649.	1215.	1002.	893.	833.	708.	549.6	.37	2	6	6
Z	1589.	1201.	957.	868.	885.	800.	423.6	.28	2	6	6
AA	1534.	1188.	885.	805.	870.	1017.	355.8	.24	2	6	6
AB	1526.	1164.	876.	801.	916.	1017.	334.2	.22	2	6	6
AC	1142.	951.	1001.	1075.	1022.	1109.	24.3	.02	2	6	5
AD	1551.	1202.	930.	889.	996.	732.	398.5	.27	2	6	6
AE	2310.	1818.	1595.	1519.	15[illegible].	1758.	261.8	.10	2	1[illegible]	7
AF	1615.	1223.	993.	911.	844.	714.	502.[illegible]	.33	2	6	6
AG	1517.	1078.	950.	938.	884.	883.	274.4	.18	2	5	5
AH	1527.	1129.	1006.	966.	860.	812.	319.5	.21	2	6	6
AI	1291.	1028.	968.	960.	1028.	1025.	71.[illegible]	.05	2	5	4
AJ	1532.	1383.	964.	944.	901.	876.	29[illegible].[illegible]	.19	2	6	6
AK	2441.	1868.	1711.	1621.	1461.	1398.	409.7	.16	2	9	7

APPENDIX D

TABLE A-8
ESTIMATED REDUCED FORM

TYPICAL MEDIAN BIASES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	.10	.05	.04	.03	.01	.01	.28
B	.01	-.07	-.07	-.07	-.01	-.01	-.00
C	.02	-.02	-.02	-.02	-.01	-.02	.15
D	-.38	-.11	-.10	-.07	-.04	-.01	-.26
E	-.03	-.04	-.05	-.04	-.02	-.03	-.29
F	.12	.08	.04	.01	.04	-.01	.14
G	.24	.09	.08	.08	.02	-.03	.27
H	.34	.08	.07	.04	.06	-.00	.13
I	.78	.26	.20	.09	.31	.04	.15
J	.06	.02	.02	.01	.03	.02	.19
K	-.04	-.00	-.00	.00	.01	.01	.17
L	-.07	-.01	-.00	.01	-.00	.00	-.35
M	.25	.06	.05	.03	.03	-.02	.39
N	.11	.04	.03	.02	.03	.01	.06
O	.20	.04	.03	.02	.04	.00	.01
P	.40	.03	.01	-.02	.05	-.03	-.05
Q	.19	.05	.03	.02	.04	.00	.01
R	.24	.04	.03	.01	.04	.00	-.01
S	.30	-.03	-.03	-.06	.01	-.05	-.03
T	.20	.03	.02	.01	.04	-.00	-.04
U	.03	.03	.03	.02	.03	.03	-.01
V	.07	.00	.00	-.00	.01	-.00	.02
W	.13	.01	.00	-.02	.03	-.02	.02
X	.03	.02	.02	.02	.01	.00	.16
Y	.13	.04	.03	.01	.00	.00	.39
Z	.42	.16	.11	.05	.00	-.08	.82
AA	.76	.14	.10	.18	.12	-.17	1.20
AB	1.38	.21	.02	.02	.05	-.22	1.25
AC	.20	.05	-.09	-.04	.09	-.09	.33
AD	2.51	.66	.38	.04	.73	.01	.04
AE	3.23	.42	.36	.02	.47	-.01	-.02
AF	-.00	-.02	-.02	-.02	-.02	-.02	.03
AG	-.02	-.03	-.03	-.03	-.02	-.02	-.08
AH	-.01	-.03	-.03	-.03	-.00	-.01	-.11
AI	-.07	-.03	-.03	-.02	-.02	-.02	-.03
AJ	.24	-.01	-.01	-.01	.02	.01	-.01
AK	.01	.01	.01	.01	.01	.01	.02

APPENDIX D

TABLE A -9
ESTIMATED REDUCED FORM

TYPICAL INTERQUARTILE RANGE

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	.60	.59	.60	.59	.31	.30	2.01
B	.47	.46	.46	.46	.27	.27	1.76
C	.67	.61	.61	.61	.32	.28	2.40
D	.52	.48	.49	.49	.27	.27	2.06
E	.53	.53	.53	.53	.26	.25	1.81
F	.47	.43	.42	.46	.40	.35	1.21
G	.78	.95	.94	.93	.37	.41	2.45
H	.45	.48	.48	.48	.36	.39	1.40
I	.76	.63	.77	.76	.59	.60	1.65
J	.55	.56	.56	.57	.44	.43	1.95
K	.47	.52	.53	.54	.50	.54	1.65
L	.46	.44	.45	.47	.54	.57	1.24
M	.45	.48	.48	.49	.47	.47	1.29
N	.42	.44	.44	.44	.46	.47	1.19
O	.27	.30	.30	.30	.25	.26	.50
P	.40	.45	.46	.46	.41	.42	.78
Q	.28	.29	.32	.29	.21	.20	.57
R	.35	.40	.40	.41	.34	.32	.80
S	.40	.45	.45	.47	.35	.31	.75
T	.37	.37	.37	.39	.30	.30	.73
U	.32	.29	.29	.29	.27	.26	.74
V	.21	.22	.22	.23	.19	.21	.32
W	.33	.35	.35	.35	.29	.28	.37
X	.28	.29	.29	.28	.14	.14	1.04
Y	1.31	1.18	1.19	1.20	.71	.60	4.15
Z	2.66	2.09	1.99	1.79	1.19	.91	6.24
AA	4.21	3.10	3.05	2.09	2.01	1.39	8.13
AB	5.71	3.53	3.42	3.49	2.55	1.75	9.33
AC	2.69	2.44	2.60	2.73	2.36	2.38	6.36
AD	2.70	1.37	1.49	1.27	1.31	.85	2.20
AE	6.94	1.83	1.77	1.70	1.66	1.20	3.25
AF	.34	.39	.39	.39	.19	.21	1.31
AG	.29	.29	.29	.29	.16	.16	.99
AH	.26	.27	.27	.27	.12	.11	.88
AI	.24	.22	.22	.22	.24	.24	.64
AJ	.18	.15	.15	.15	.12	.11	.35
AK	.15	.15	.15	.14	.14	.14	.25

APPENDIX D

TABLE A-10

ESTIMATED REDUCED FORM

RANKINGS ON ALL ESTIMATES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	CHI SQ	W	SIG	X95	X99
A	5020.	4251.	4135.	4137.	3127.	3211.	5519.	929.1	.15	2	16	15
B	5139.	4150.	4026.	4052.	3311.	3391.	5331.	747.0	.12	2	19	19
C	4995.	4249.	4213.	4181.	3158.	3152.	5452.	895.2	.14	2	17	16
D	5103.	4038.	4090.	4153.	3229.	3136.	5651.	1027.8	.16	2	15	15
E	4931.	4208.	4261.	4207.	3213.	3114.	5466.	876.4	.14	2	17	17
F	5230.	4193.	3921.	3841.	3683.	3338.	5194.	666.5	.11	2	18	15
G	4700.	4171.	4224.	4279.	3188.	3436.	5402.	675.6	.11	2	15	14
H	5514.	4203.	4007.	3932.	3273.	3273.	5198.	928.6	.15	2	19	19
I	5499.	3960.	3908.	3769.	3950.	3395.	4919.	661.9	.11	2	18	18
J	4811.	4118.	4009.	4079.	3556.	3507.	5320.	526.6	.08	2	15	15
K	4409.	3818.	3841.	3991.	4065.	4056.	5220.	294.2	.05	2	11	11
L	4426.	3941.	3939.	3959.	4031.	4202.	4902.	156.3	.02	2	11	9
M	5094.	4127.	3988.	3909.	3543.	3635.	5104.	510.7	.08	2	16	15
N	4395.	3875.	3837.	3945.	4121.	4152.	5095.	235.4	.04	2	11	7
O	5832.	4133.	3890.	3820.	3526.	3258.	4961.	991.1	.16	2	18	18
P	5933.	3910.	3756.	3586.	3871.	3460.	4334.	976.6	.16	2	21	21
Q	5784.	4330.	4028.	3833.	3285.	2789.	5351.	1396.6	.22	2	21	21
R	5514.	4149.	3918.	3795.	3466.	3260.	5293.	938.9	.15	2	21	20
S	4223.	3485.	3524.	3751.	2833.	2940.	4444.	515.8	.10	2	13	13
T	6386.	4719.	4521.	4490.	4036.	3860.	5538.	827.5	.11	2	22	22
U	5877.	4629.	4500.	4500.	4108.	4212.	5974.	687.2	.10	2	19	18
V	5266.	4606.	4722.	5045.	4155.	4333.	5473.	251.4	.03	2	14	8
W	6142.	4826.	4805.	4899.	3993.	4000.	4935.	557.3	.08	2	22	22
X	4650.	4302.	4236.	4191.	3147.	3288.	5586.	831.8	.13	2	16	15
Y	5698.	4064.	3956.	3936.	3249.	3168.	5329.	1150.2	.18	2	21	20
Z	5987.	4355.	3767.	3777.	3593.	3143.	5078.	1191.3	.19	2	21	21
AA	5939.	4141.	3895.	3683.	3727.	3279.	4736.	968.8	.15	2	19	19
AB	5992.	4250.	3896.	3698.	3803.	3190.	4571.	994.6	.16	2	21	21
AC	4958.	3764.	4037.	4097.	3902.	3904.	4736.	258.7	.04	2	16	13
AD	7031.	4569.	3711.	2840.	4940.	2361.	3948.	2904.6	.46	2	21	21
AE	7569.	5103.	4354.	3928.	4691.	3505.	4450.	1880.3	.26	2	24	24
AF	5149.	4101.	4104.	4052.	3227.	3311.	5476.	880.2	.14	2	17	17
AG	5085.	4126.	4207.	4132.	3143.	3294.	5613.	1247.0	.17	2	19	19
AH	5398.	4170.	4147.	4165.	3081.	3098.	5341.	1063.0	.17	2	19	19
AI	4959.	3870.	3782.	3769.	4029.	3973.	5016.	366.4	.06	2	14	13
AJ	6328.	3836.	3969.	3973.	3113.	3108.	5068.	1610.4	.26	2	21	21
AK	7039.	4478.	4407.	4257.	4021.	4043.	5355.	1259.6	.17	2	22	21

APPENDIX D

TABLE A-11
PREDICTIONS

TYPICAL INTERQUARTILE RANGE

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	.04	.03	.03	.03	.03	.03	.05
B	.03	.03	.03	.03	.03	.02	.05
C	.03	.03	.03	.03	.03	.02	.05
D	.04	.04	.04	.04	.04	.04	.07
E	.03	.03	.03	.03	.03	.02	.03
F	.05	.04	.04	.05	.04	.04	.06
G	.03	.03	.03	.03	.02	.02	.04
H	.03	.04	.04	.04	.03	.03	.05
I	.04	.04	.04	.04	.03	.04	.05
J	.03	.03	.03	.03	.03	.03	.04
K	.03	.03	.03	.03	.03	.03	.04
L	.03	.02	.02	.02	.02	.02	.05
M	.04	.04	.04	.04	.03	.03	.05
N	.03	.03	.03	.03	.03	.03	.04
O	.03	.03	.03	.03	.03	.03	.04
P	.03	.03	.03	.03	.03	.02	.04
Q	.02	.02	.02	.02	.02	.02	.04
R	.03	.03	.03	.03	.02	.02	.04
S	.03	.03	.03	.03	.02	.02	.03
T	.02	.02	.02	.02	.02	.02	.03
U	.02	.02	.02	.02	.02	.02	.02
V	.02	.02	.02	.02	.02	.02	.02
W	.07	.06	.07	.07	.05	.05	.07
X	.02	.02	.02	.02	.02	.02	.02
Y	.08	.07	.07	.07	.06	.06	.10
Z	.15	.11	.10	.11	.10	.10	.14
AA	.22	.16	.16	.15	.15	.15	.20
AB	.31	.23	.23	.20	.21	.19	.26
AC	.12	.12	.12	.13	.12	.12	.14
AD	.18	.12	.12	.10	.10	.07	.16
AE	.23	.09	.09	.09	.09	.09	.10
AF	.02	.02	.02	.02	.02	.02	.02
AG	.02	.02	.02	.02	.01	.01	.02
AH	.01	.01	.01	.01	.01	.01	.02
AI	.01	.01	.01	.01	.01	.01	.01
AJ	.01	.01	.01	.01	.01	.01	.02
AK	.01	.01	.01	.01	.01	.01	.01

APPENDIX D

TABLE A-12

PREDICTIONS

RANKINGS ON ALL ESTIMATES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	CHI SQ	W	SIG	X95	X99
A	1899.	1768.	1782.	1825.	1560.	1496.	2270.	182.2	.07	2	7	5
B	1870.	1738.	1798.	1746.	1605.	1582.	2261.	147.5	.05	2	7	7
C	1901.	1778.	1821.	1809.	1473.	1436.	2382.	280.6	.10	2	9	9
D	1716.	1716.	1840.	1782.	1660.	1583.	2303.	159.9	.06	2	9	7
E	2178.	1801.	1823.	1756.	1497.	1429.	2136.	232.0	.09	2	9	9
F	2090.	1754.	1703.	1706.	1610.	1532.	2205.	179.2	.07	2	9	9
G	1899.	1771.	1784.	1798.	1508.	1587.	2253.	165.1	.06	2	7	5
H	2129.	1783.	1734.	1754.	1440.	1459.	2301.	291.4	.11	2	9	9
I	1965.	1749.	1738.	1659.	1771.	1568.	2150.	139.9	.04	2	8	5
J	1813.	1733.	1777.	1826.	1539.	1633.	2229.	124.7	.05	2	6	4
K	1783.	1662.	1695.	1792.	1736.	1793.	2137.	71.2	.03	2	4	3
L	1940.	1734.	1744.	1724.	1740.	1738.	1980.	34.6	.01	2	2	1
M	2031.	1750.	1701.	1721.	1588.	1614.	2195.	146.4	.05	2	7	5
N	1866.	1653.	1625.	1666.	1789.	1830.	2191.	112.3	.04	2	5	4
O	2637.	1727.	1589.	1606.	1494.	1431.	2116.	532.2	.20	2	9	9
P	1979.	1739.	1793.	1771.	1673.	1464.	2251.	186.6	.07	2	9	9
Q	2362.	1819.	1676.	1614.	1373.	1293.	2461.	591.1	.22	2	9	9
R	2053.	1788.	1808.	1790.	1419.	1401.	2341.	314.9	.12	2	9	9
S	1810.	1838.	1830.	1897.	1572.	1527.	2198.	159.4	.06	2	6	4
T	2661.	2434.	2452.	2435.	2140.	2055.	2673.	118.4	.03	2	7	5
U	2940.	2295.	2178.	2210.	2217.	2155.	2855.	232.6	.06	2	8	7
V	3098.	2999.	3081.	3362.	2657.	2721.	3362.	99.3	.02	2	6	5
W	3639.	3007.	3052.	3036.	2576.	2643.	3229.	236.3	.05	2	11	9
X	1846.	1811.	1828.	1846.	1511.	1477.	2281.	272.1	.07	2	6	5
Y	2054.	1725.	1770.	1776.	1473.	1511.	2291.	239.6	.09	2	9	9
Z	2106.	1681.	1743.	1753.	1566.	1562.	2139.	179.2	.07	2	7	7
AA	2078.	1793.	1846.	1703.	1535.	1574.	2074.	135.9	.05	2	7	6
AB	2082.	1753.	1877.	1724.	1574.	1552.	2231.	123.5	.05	2	7	7
AC	1895.	1674.	1798.	1772.	1739.	1740.	1982.	31.5	.01	2	3	[illegible]
AD	2871.	1936.	1574.	1235.	1923.	1269.	1742.	976.9	.36	2	9	9
AE	3398.	2405.	2238.	2230.	2329.	1973.	2257.	453.4	.13	2	12	11
AF	2045.	1745.	1734.	1696.	1676.	1656.	2343.	83.7	.03	2	3	3
AG	1944.	1799.	1867.	1850.	1433.	1524.	2133.	152.1	.06	2	5	5
AH	2196.	1755.	1756.	1760.	1445.	1435.	2253.	[illegible]	.11	2	8	9
AI	2125.	1626.	1615.	1638.	1636.	1766.	2144.	156.6	.06	2	8	7
AJ	2052.	1696.	1731.	1941.	1461.	1473.	2299.	261.5	.10	2	8	7
AK	3349.	2173.	2223.	2242.	2338.	2367.	2738.	[illegible]	.13	2	11	[illegible]

APPENDIX D

TABLE 8-1
ESTIMATED STRUCTURAL COEFFICIENTS

TYPICAL MEDIAN BIASES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	-.13	-.00	.01	.02	.00	.00
B	-.57	-.26	-.18	.01	-.16	-.05
C	-.33	-.09	-.06	.00	-.04	-.01
D	-.34	-.06	-.02	.04	-.00	.03
E	-.34	-.05	-.03	.02	-.00	.03
F	-.57	-.11	-.05	.03	-.05	.00
G	-.85	-.53	-.43	-.22	-.48	-.16
H	-.11	-.01	-.01	.01	-.00	.00
I	-.21	-.06	-.04	-.01	-.04	-.00
J	-.51	-.37	-.36	-.27	-.36	-.24
K	-.55	-.33	-.29	-.22	-.49	-.35
L	-.10	-.01	-.00	.01	.01	.02
M	-.40	-.04	-.03	-.03	-.02	-.01
N	0	-.01	-.01	-.01	-.02	-.01
O	-.08	-.01	-.00	.01	-.01	-.00
P	-.12	-.02	-.01	-.00	-.01	.02
Q	-.15	-.04	-.04	-.01	-.04	-.01
R	-.18	-.04	-.03	-.01	-.03	.00
S	-.23	-.04	-.03	-.01	-.01	.00
T	-.42	-.09	-.07	-.01	-.02	.01
U	-.10	-.03	-.02	-.00	-.02	-.01
V	-.11	-.00	-.00	.00	.00	.01
W	-.04	-.02	-.01	-.01	-.02	-.01
X	-.26	-.09	-.06	-.06	-.02	.00
Y	-.36	-.13	-.11	-.07	-.07	-.23
Z	-.16	.00	.01	.02	.01	.02
AA	-.23	-.02	-.00	.01	-.02	-.00
AB	-.11	-.00	.00	.02	-.01	.01
AC	-.15	-.03	-.02	-.01	-.01	.01
AD	-.56	-.12	-.09	.01	-.05	-.01
AE	-.13	-.00	.01	.04	.02	.03
AF	-.05	-.00	-.00	.00	-.01	.02
AG	-.18	-.02	-.01	.01	-.01	-.00
AH	-.05	-.01	-.01	.00	-.00	-.00
AI	-.12	-.02	-.01	-.00	-.01	.00
AJ	-.36	-.06	-.03	.01	-.04	.00
AK	-.74	-.33	-.25	-.09	-.28	-.11
AL	.14	.03	-.03	-.07	-.15	-.16
AM	-.41	-.11	-.09	-.01	-.05	-.01
AN	-.98	-.61	-.55	-.09	-.49	-.15
AO	-.25	-.29	-.07	.02	-.03	-.01

APPENDIX D

TABLE B-2
ESTIMATED STRUCTURAL COEFFICIENTS

TYPICAL INTERQUARTILE RANGE

RUN	OLS	2SLS	UBK	LIML	3SLS	FIML
A	.39	.44	.44	.45	.29	.29
B	.27	.52	.63	.88	.44	.50
C	.33	.46	.49	.52	.37	.40
D	.60	.84	.87	1.03	.67	.72
E	.34	.49	.51	.55	.50	.53
F	.34	.62	.70	.80	.48	.52
G	.29	.66	.94	1.34	.85	1.51
H	.31	.40	.40	.44	.34	.38
I	.19	.27	.29	.30	.19	.18
J	.23	.43	.51	.96	.47	1.11
K	.70	1.07	1.34	1.00	1.03	1.41
L	.24	.25	.26	.27	.25	.25
M	.18	.23	.24	.23	.22	.23
N	.33	.44	.54	.62	.43	.48
O	.28	.33	.33	.35	.24	.26
P	.22	.29	.29	.30	.25	.28
Q	.22	.32	.33	.37	.26	.30
R	.16	.23	.24	.27	.17	.20
S	.25	.36	.38	.41	.32	.34
T	.24	.51	.53	.54	.40	.42
U	.19	.28	.29	.30	.21	.21
V	.12	.15	.15	.15	.11	.11
W	.24	.25	.25	.25	.27	.28
X	.61	.71	.72	.73	.36	.36
Y	.62	.80	.80	.81	.46	.44
Z	.39	.41	.41	.44	.37	.38
AA	.35	.43	.44	.47	.24	.26
AB	.24	.34	.34	.35	.27	.28
AC	.22	.28	.28	.28	.23	.24
AD	.21	.56	.56	.56	.48	.52
AE	.22	.35	.37	.42	.27	.28
AF	.31	.35	.35	.36	.35	.37
AG	.32	.36	.36	.36	.23	.23
AH	.12	.14	.14	.14	.08	.06
AI	.21	.25	.26	.27	.21	.21
AJ	.53	.92	.98	1.06	.60	.62
AK	.56	1.22	1.40	1.93	.72	1.13
AL	.69	1.17	1.38	1.39	1.43	1.69
AM	.31	.51	.53	.63	.43	.44
AN	.37	.78	.93	2.11	.93	1.33
AO	.31	.54	.64	.81	.51	.61

APPENDIX D

TABLE B-3

ESTIMATED STRUCTURAL COEFFICIENTS

RANKINGS ON ALL ESTIMATES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	CHI SQ	W	SIG	X95	X99
A	3334.	2751.	2698.	2771.	2078.	2118.	419.6	.11	2	15	14
B	3702.	2756.	2475.	2598.	2158.	2061.	661.5	.18	2	15	14
C	3502.	2649.	2508.	2518.	2380.	2193.	396.8	.11	2	15	14
D	3703.	2730.	2478.	2589.	2125.	2125.	646.1	.17	2	15	15
E	3437.	2555.	2508.	2559.	2302.	2339.	362.7	.10	2	14	14
F	3720.	2691.	2517.	2630.	2104.	2038.	676.1	.18	2	14	14
G	3483.	2620.	2371.	2443.	2365.	2468.	352.8	.09	2	13	13
H	3856.	2915.	2929.	2946.	2554.	2650.	358.2	.08	2	12	11
I	3117.	2324.	2152.	2062.	1554.	1391.	899.7	.30	2	12	12
J	3146.	2463.	2394.	2478.	2590.	2679.	143.5	.04	2	13	7
K	2542.	2271.	2622.	2750.	2638.	2927.	91.1	.02	2	9	5
L	3211.	2605.	2574.	2615.	2390.	2355.	180.3	.05	2	8	7
M	4069.	2552.	2469.	2381.	2078.	2231.	1212.6	.27	2	14	13
N	5156.	4374.	4730.	4806.	4367.	4165.	192.5	.03	2	18	16
O	3287.	2691.	2649.	2777.	2095.	2251.	337.9	.09	2	14	13
P	3438.	2663.	2544.	2590.	2204.	2311.	368.4	.10	2	14	12
Q	3385.	2788.	2657.	2608.	2224.	2228.	428.3	.11	2	14	14
R	3469.	2794.	2677.	2721.	2046.	2343.	543.5	.14	2	12	12
S	3685.	2709.	2562.	2539.	2083.	2122.	641.2	.17	2	15	14
T	3874.	2733.	2621.	2539.	2072.	1731.	938.1	.25	2	15	15
U	3409.	2813.	2786.	2742.	2047.	1953.	562.2	.15	2	14	13
V	3878.	2562.	2597.	2567.	2122.	2224.	835.2	.22	2	15	15
W	2795.	2547.	2499.	2536.	2628.	2745.	27.9	.01	2	5	4
X	3439.	2758.	2635.	2745.	2116.	2097.	466.5	.12	2	15	14
Y	3522.	2857.	2644.	2653.	2064.	2712.	591.4	.16	2	15	15
Z	4233.	3587.	3592.	3646.	2928.	3016.	324.5	.06	2	14	13
AA	3519.	2728.	2694.	2764.	1988.	2057.	595.2	.16	2	15	15
AB	3887.	2969.	2943.	2968.	2482.	2576.	409.3	.12	2	15	14
AC	3475.	2661.	2596.	2650.	2144.	2224.	425.7	.11	2	14	13
AD	3935.	2769.	2521.	2462.	1927.	2136.	952.6	.25	2	15	15
AE	3563.	2756.	2688.	2674.	1934.	2143.	614.2	.16	2	14	13
AF	3009.	2496.	2426.	2457.	2679.	2673.	89.1	.02	2	4	4
AG	3363.	2776.	2725.	2799.	2021.	2066.	489.5	.13	2	15	15
AH	2717.	2331.	2254.	2259.	1530.	1531.	551.3	.13	2	12	12
AI	3465.	2715.	2593.	2636.	2119.	2225.	438.9	.11	2	14	13
AJ	3514.	2539.	2321.	2706.	2146.	2104.	462.5	.12	2	15	15
AK	3522.	2618.	2543.	2706.	2084.	2277.	469.2	.13	2	15	14
AL	4070.	3447.	3360.	4347.	3575.	[illegible]	139.1	.03	2	17	16
AM	3801.	2751.	2638.	2536.	1963.	2391.	811.5	.22	2	15	15
AN	3772.	2699.	2275.	2526.	2191.	2267.	669.1	.18	2	15	14
AO	3522.	2591.	2495.	2620.	2187.	2233.	445.5	.12	2	13	13

APPENDIX D

TABLE B-4

ESTIMATED STANDARD ERRORS

PROPORTION OF BIAS/S.E. OUTSIDE MODEL I

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	.20	.07	.06	.06	.05	.04
B	.61	.19	.16	.08	.17	.09
C	.40	.11	.10	.05	.13	.07
D	.53	.11	.10	.06	.11	.08
E	.35	.07	.05	.05	.09	.05
F	.57	.10	.09	.06	.09	.05
G	.73	.36	.28	.12	.40	.24
H	.16	.01	.01	.01	.02	.01
I	.47	.12	.11	.07	.10	.07
J	.71	.43	.36	.18	.51	.31
K	.37	.11	.10	.09	.23	.18
L	.13	.05	.04	.05	.07	.05
M	.70	.07	.07	.05	.07	.06
N	.25	.04	.04	.04	.10	.08
O	.14	.06	.06	.06	.05	.05
P	.19	.06	.06	.05	.07	.06
Q	.25	.07	.07	.06	.06	.03
R	.28	.07	.07	.06	.06	.07
S	.34	.08	.07	.06	.08	.07
T	.57	.11	.09	.07	.11	.06
U	.20	.07	.07	.06	.05	.04
V	.60	.06	.05	.05	.05	.05
W	.04	.04	.04	.04	.05	.05
X	.19	.07	.06	.05	.05	.03
Y	.23	.07	.07	.06	.06	.04
Z	.24	.08	.08	.08	.09	.09
AA	.22	.06	.05	.05	.05	.04
AB	.14	.02	.02	.01	.02	.02
AC	.23	.07	.06	.05	.07	.06
AD	.70	.13	.12	.07	.12	.08
AE	.43	.08	.08	.05	.07	.04
AF	.07	.03	.03	.03	.05	.04
AG	.16	.05	.05	.05	.05	.04
AH	.13	.07	.07	.06	.05	.06
AI	.19	.06	.06	.05	.07	.06
AJ	.37	.07	.07	.05	.08	.05
AK	.55	.11	.10	.06	.13	.10
AL	1.00	1.00	1.00	1.00	1.00	1.00
AM	.54	.09	.09	.06	.08	.06
AN	.77	.33	.27	.10	.32	.16
AO	.54	.19	.16	.09	.19	.13

APPENDIX D

TABLE B-5
ESTIMATED RESIDUALS

TYPICAL MEDIAN BIASES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	-.49	-.36	-.35	-.33	-.28	-.21
B	-.81	-.61	-.57	-.35	-.51	-.27
C	-.67	-.48	-.44	-.34	-.41	-.26
D	-.69	-.46	-.42	-.30	-.36	-.17
E	-.58	-.36	-.34	-.31	-.30	-.27
F	-.73	-.45	-.40	-.30	-.36	-.24
G	-.89	-.76	-.68	-.37	-.71	-.30
H	-.54	-.41	-.41	-.38	-.36	-.23
I	-.59	-.38	-.35	-.29	-.25	-.17
J	-.90	-.83	-.82	-.62	-.83	-.58
K	-1.07	-.81	-.55	-1.00	-.98	-.97
L	-.41	-.32	-.31	-.30	-.28	-.25
M	-.50	-.11	-.09	-.08	-.07	-.07
N	-.77	-.33	-.28	-.29	-.21	-.11
O	-.44	-.34	-.33	-.31	-.26	-.21
P	-.46	-.34	-.33	-.32	-.26	-.21
Q	-.54	-.39	-.37	-.33	-.31	-.22
R	-.49	-.35	-.34	-.31	-.27	-.20
S	-.54	-.36	-.35	-.30	-.27	-.19
T	-.68	-.41	-.38	-.30	-.31	-.20
U	-.46	-.32	-.31	-.28	-.23	-.18
V	-.33	-.08	-.07	-.06	-.04	-.04
W	-.53	-.40	-.40	-.38	-.42	-.35
X	-.50	-.38	-.37	-.35	-.29	-.23
Y	-.50	-.36	-.35	-.34	-.29	-.23
Z	-.39	-.23	-.22	-.20	-.19	-.17
AA	-.48	-.34	-.33	-.31	-.27	-.21
AB	-.56	-.42	-.41	-.39	-.31	-.22
AC	-.46	-.33	-.31	-.29	-.26	-.20
AD	-.76	-.44	-.40	-.30	-.33	-.21
AE	-.62	-.37	-.35	-.27	-.26	-.19
AF	-.55	-.25	-.24	-.19	-.21	-.15
AG	-.43	-.33	-.32	-.31	-.26	-.22
AH	-.32	-.26	-.25	-.24	-.16	-.12
AI	-.44	-.35	-.34	-.32	-.27	-.23
AJ	-.67	-.42	-.38	-.32	-.35	-.19
AK	-.82	-.58	-.51	-.24	-.49	-.27
AL	-.37	-.25	-.22	-.07	-.23	-.07
AM	-.70	-.41	-.37	-.28	-.31	-.21
AN	-.91	-.76	-.73	-.13	-.67	-.21
AO	-.82	-.63	-.55	-.26	-.46	-.19

APPENDIX D

TABLE B-6
ESTIMATED RESIDUALS

TYPICAL INTERQUARTILE RANGE

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	.27	.41	.43	.45	.4Ø	.45
B	.12	.31	.44	.95	.39	.64
C	.16	.36	.42	.47	.38	.52
D	.17	.38	.41	.56	.42	.54
E	.2Ø	.4Ø	.42	.46	.43	.45
F	.13	.4Ø	.48	.58	.45	.49
G	.Ø7	.25	.52	1.Ø9	.44	2.62
H	.62	.97	.99	1.Ø4	1.Ø5	1.16
I	.22	.4Ø	.45	.43	.34	.45
J	.Ø7	.17	.24	.87	.19	4.45
K	1.57	2.49	3.24	4.66	2.43	4.57
L	.27	.34	.34	.34	.35	.35
M	.14	.31	.32	.33	.31	.32
N	.61	1.Ø5	1.33	1.55	1.13	1.28
O	.28	.37	.37	.4Ø	.42	.46
P	.28	.33	.39	.41	.41	.46
Q	.25	.39	.42	.53	.43	.55
R	.25	.35	.37	.4Ø	.37	.38
S	.25	.39	.41	.44	.41	.45
T	.19	.42	.46	.51	.47	.52
U	.28	.36	.4Ø	.42	.38	.38
V	.17	.29	.29	.29	.27	.29
W	1.87	1.95	1.96	1.97	2.26	2.46
X	.29	.44	.45	.45	.45	.43
Y	.23	.37	.38	.41	.41	.44
Z	.64	.8Ø	.82	.87	.83	.7Ø
AA	.24	.37	.39	.38	.36	.37
AB	.65	.92	.93	.98	.94	.97
AC	.27	.38	.4Ø	.41	.39	.39
AD	.14	.4Ø	.48	.47	.41	.57
AE	.19	.39	.42	.41	.41	.48
AF	1.76	2.17	2.21	2.35	2.46	2.75
AG	.27	.36	.37	.37	.37	.36
AH	.3Ø	.32	.33	.34	.36	.38
AI	.29	.37	.38	.38	.36	.39
AJ	.18	.47	.54	.62	.49	.69
AK	.Ø9	.38	.55	.96	.42	.85
AL	.2Ø	.71	1.Ø9	1.63	1.1Ø	1.36
AM	.16	.44	.49	.67	.53	.66
AN	.Ø9	.3Ø	.47	2.14	.5Ø	1.54
AO	.12	.34	.47	.97	.41	.81

APPENDIX D

TABLE B -7

ESTIMATED RESIDUALS

RANKINGS ON ALL ESTIMATES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	CHI SQ	W	SIG	X95	X99
A	1631.	1269.	1028.	896.	819.	657.	588.1	.39	2	6	6
B	1637.	1265.	992.	930.	810.	666.	584.4	.39	2	6	6
C	1647.	1238.	985.	842.	848.	740.	548.7	.37	2	6	6
D	1636.	1259.	983.	847.	818.	707.	633.7	.42	2	6	6
E	1667.	1233.	973.	835.	838.	754.	570.4	.38	2	6	6
F	1679.	1270.	989.	862.	802.	698.	636.7	.42	2	6	6
G	1499.	1166.	918.	861.	920.	936.	283.9	.19	2	6	6
H	2468.	1943.	1753.	1521.	1450.	1365.	482.0	.19	2	10	10
I	1685.	1277.	1011.	847.	782.	698.	660.2	.44	2	6	6
J	1386.	1090.	940.	895.	967.	1020.	150.6	.10	2	5	4
K	1176.	970.	945.	1050.	1037.	1122.	36.8	.02	2	6	6
L	1644.	1299.	1015.	868.	792.	682.	620.2	.41	2	6	6
M	1667.	1093.	958.	904.	842.	836.	477.5	.32	2	6	6
N	3051.	2579.	2688.	2752.	2273.	2437.	142.9	.04	2	14	10
O	1613.	1278.	1024.	920.	781.	634.	564.6	.38	2	6	6
P	1653.	1272.	1001.	882.	807.	687.	603.3	.40	2	6	6
Q	1666.	1283.	1020.	878.	778.	670.	652.4	.43	2	6	6
R	1671.	1299.	1021.	891.	785.	633.	683.7	.46	2	6	6
S	1630.	1275.	992.	875.	773.	704.	646.1	.43	2	6	6
T	1696.	1271.	1013.	868.	759.	693.	678.8	.45	2	6	6
U	1643.	1272.	1051.	886.	770.	678.	613.9	.41	2	6	6
V	1575.	1283.	960.	953.	870.	854.	348.0	.23	2	6	6
W	1195.	1254.	957.	932.	1177.	1155.	44.7	.03	2	6	6
X	1628.	1290.	1039.	914.	780.	649.	613.3	.41	2	6	6
Y	1658.	1285.	1029.	849.	801.	678.	634.4	.42	2	6	6
Z	3435.	2859.	2566.	2415.	2248.	2227.	403.4	.11	2	13	13
AA	1658.	1272.	997.	899.	866.	668.	619.1	.41	2	6	6
AB	2505.	1966.	1751.	1559.	1422.	1317.	546.4	.22	2	10	13
AC	1638.	1244.	981.	871.	832.	692.	616.2	.41	2	6	6
AD	1636.	1232.	972.	832.	793.	735.	645.0	.43	2	6	6
AE	1695.	1278.	1021.	863.	724.	719.	685.4	.46	2	6	6
AF	1240.	1249.	985.	922.	1079.	1025.	55.4	.04	2	6	6
AG	1650.	1235.	1008.	878.	816.	663.	620.1	.41	2	6	6
AH	1572.	1271.	1027.	885.	777.	768.	479.2	.32	2	6	6
AI	1695.	1298.	1013.	867.	775.	652.	710.9	.47	2	6	6
AJ	1576.	1187.	938.	845.	891.	763.	545.6	.36	2	6	6
AK	1659.	1130.	903.	834.	872.	847.	502.4	.33	2	6	6
AL	2232.	1781.	1719.	1806.	1409.	1553.	224.3	.09	2	8	8
AM	1659.	1250.	1001.	905.	755.	730.	594.0	.40	2	6	6
AN	1573.	1172.	934.	917.	862.	842.	379.2	.25	2	6	6
AO	1610.	1249.	985.	884.	785.	787.	499.4	.33	2	6	6

APPENDIX D

TABLE B-8
ESTIMATED REDUCED FORM

TYPICAL MEDIAN BIASES

RUN	DLS	2SLS	LBK	LIML	3SLS	FIML	LSRF
A	.10	.05	.04	.03	.01	.01	.28
B	.30	.02	.03	.00	.02	.02	.29
C	.03	.06	.03	.02	.08	.01	-.30
D	.03	.07	.05	.07	.07	.05	2.88
E	-.07	.02	.02	.02	.03	.04	-.27
F	-.15	-.02	-.03	-.02	-.01	.02	.23
G	.13	-.15	-.22	-.17	-.12	-.08	.30
H	.11	.01	-.00	-.01	.03	.01	-.18
I	.01	-.02	-.02	-.02	-.00	-.01	-.05
J	-2.87	-.05	-.09	.09	.26	-.21	.59
K	-.00	-.05	-.08	-.07	-.10	-.05	.10
L	-.00	.02	.02	.02	.03	.03	-.14
M	-.16	-.04	-.04	-.03	-.02	-.02	-.15
N	.00	.00	-.00	.01	.01	.03	-.00
O	.12	.02	.02	.01	.00	-.00	.06
P	.06	.01	.01	.00	.01	-.00	-.05
Q	.17	.03	.03	.00	.06	.00	-.06
R	.07	.02	.01	.01	.01	-.00	.01
S	-.01	-.02	-.02	-.01	-.00	-.01	.11
T	-.04	-.02	-.02	-.03	-.02	-.01	.07
U	.09	.01	.01	.00	.01	.00	-.26
V	.09	-.01	-.01	-.01	.01	.00	.30
W	.05	.01	.01	.01	.01	-.00	-.04
X	-.27	-.15	-.14	-.13	-.02	-.00	-.04
Y	-.54	-.18	-.16	-.13	-.08	-.05	-.49
Z	.06	-.00	-.01	-.01	.01	.00	-.01
AA	-.12	-.01	-.00	-.00	-.02	.00	-.10
AB	.10	-.01	-.01	-.01	.02	.00	-.00
AC	.01	-.01	-.02	-.02	-.01	-.01	-.05
AD	-.05	-.00	-.01	-.02	-.01	-.01	-.02
AE	.19	.06	.04	.02	.04	.00	.14
AF	.18	.03	.02	.01	.22	.02	-.16
AG	-.10	-.01	-.01	.00	-.00	.01	-.12
AH	-.01	-.01	-.01	-.01	-.01	-.01	-.05
AI	-.00	-.01	-.01	-.01	-.01	-.01	-.03
AJ	-.01	-.02	-.02	-.06	-.00	-.06	.10
AK	-1.32	-.43	-.35	-.21	-.42	-.25	.34
AL	-1.21	-.26	-.18	-.05	-.28	.05	-.38
AM	.11	-.04	-.06	-.06	-.01	-.02	.02
AN	-.73	-.22	-.19	-.16	-.22	-.28	-1.86
AO	.61	.23	.17	-.01	.25	.04	.00

APPENDIX D

TABLE B -9

ESTIMATED REDUCED FORM

TYPICAL INTERQUARTILE RANGE

RUN	OLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	.60	.59	.63	.59	.31	.33	2.21
B	.83	.68	.74	.83	.43	.48	2.25
C	.71	.67	.72	.72	.50	.49	2.67
D	1.02	1.27	1.28	1.34	.95	.95	12.91
E	.53	.60	.60	.65	.56	.59	11.62
F	.77	.75	.79	.79	.56	.52	2.87
G	1.08	1.24	1.36	1.28	1.17	1.14	5.15
H	.44	.46	.46	.49	.39	.40	2.11
I	.30	.36	.37	.36	.21	.19	.88
J	6.65	3.35	3.55	2.85	3.93	1.93	5.72
K	2.43	1.64	1.46	1.51	1.58	1.39	3.79
L	.32	.31	.31	.31	.27	.28	5.81
M	.30	.29	.29	.29	.27	.30	4.69
N	.33	.50	.74	.80	.53	.60	1.14
O	.37	.40	.40	.40	.24	.25	1.52
P	.27	.27	.27	.28	.22	.23	1.46
Q	.36	.38	.39	.40	.30	.31	1.26
R	.18	.21	.21	.22	.14	.14	.44
S	.28	.31	.31	.34	.24	.25	1.56
T	.35	.43	.45	.47	.32	.34	2.27
U	.30	.31	.35	.36	.24	.22	1.23
V	.19	.21	.20	.19	.13	.13	.51
W	.29	.31	.30	.30	.29	.30	.84
X	1.38	1.25	1.25	1.22	.51	.40	3.56
Y	1.36	1.23	1.23	1.17	.72	.59	4.12
Z	.22	.23	.23	.23	.21	.21	.36
AA	.53	.49	.49	.48	.26	.26	1.83
AB	.32	.36	.36	.36	.27	.25	.77
AC	.26	.26	.26	.27	.17	.26	1.41
AD	.41	[illegible]	.46	.50	.35	.38	2.25
AE	.32	[illegible]	.44	.48	.32	.29	1.91
AF	.36	[illegible]	.35	.37	.37	.39	1.51
AG	.43	[illegible]	.42	.42	.25	.23	1.55
AH	.18	[illegible]	.18	.18	.10	.09	.47
AI	.24	[illegible]	.25	.26	.16	.17	1.28
AJ	1.58	1.41	1.39	1.41	.82	.74	4.33
AK	4.17	[illegible]	3.13	2.93	1.68	1.33	9.33
AL	1.76	[illegible]	1.76	2.13	1.50	1.22	3.17
AM	.64	[illegible]	.55	.59	.45	.45	3.51
AN	1.98	[illegible]	1.29	1.77	1.13	1.08	12.17
AO	.76	[illegible]	.72	.95	.75	.74	2.91

APPENDIX D

TABLE B-10
ESTIMATED REDUCED FORM

RANKINGS ON ALL ESTIMATES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	CHI SQ	W	SIG	X95	X99
A	5020.	4251.	4135.	4137.	3127.	3211.	5519.	929.1	.15	2	16	15
B	5054.	3798.	4110.	4173.	3461.	3315.	5489.	794.0	.13	2	18	18
C	4301.	4087.	4103.	4048.	3942.	3313.	5606.	588.9	.09	2	19	15
D	4441.	4034.	4095.	4233.	3594.	3325.	5678.	697.0	.11	2	18	18
E	4350.	3782.	3877.	3869.	3581.	3620.	6321.	1148.8	.18	2	19	17
F	4800.	3921.	4121.	4148.	3377.	3327.	5706.	847.8	.13	2	20	17
G	4771.	4029.	4219.	3784.	3973.	3646.	4978.	304.6	.05	2	11	10
H	5951.	4560.	4400.	4494.	4124.	4025.	6046.	758.2	.11	2	23	23
I	5361.	4208.	4148.	4061.	2952.	2675.	5995.	1729.6	.27	2	21	21
J	6022.	4133.	3730.	3586.	3999.	3503.	4427.	918.3	.15	2	20	18
K	5020.	4179.	3967.	3947.	4290.	3643.	4354.	231.3	.04	2	13	11
L	4135.	3901.	3929.	3908.	3691.	3513.	6323.	1120.5	.18	2	19	17
M	5253.	3785.	3739.	3754.	3324.	3538.	6007.	1257.8	.20	2	19	19
N	4860.	4046.	5282.	5544.	4101.	4248.	5519.	476.5	.07	2	23	20
O	4919.	4181.	4084.	4210.	3218.	3359.	5429.	757.7	.12	2	17	16
P	4937.	4127.	4220.	4055.	3300.	3454.	5507.	750.3	.12	2	15	12
Q	5497.	4194.	3961.	3885.	3547.	3111.	5205.	910.4	.14	2	20	18
R	4862.	4093.	4118.	4054.	3105.	3261.	5907.	1116.8	.18	2	21	20
S	4646.	4115.	4134.	4271.	3201.	3518.	5515.	695.5	.11	2	16	16
T	4564.	4072.	4315.	4356.	3159.	3354.	5580.	793.9	.13	2	20	18
U	5220.	4138.	4113.	4001.	3179.	3104.	5645.	1106.8	.18	2	20	20
V	5893.	4005.	4013.	3961.	3171.	3146.	5211.	1262.7	.20	2	19	19
W	4441.	3993.	3911.	3929.	3903.	3965.	5258.	310.3	.05	2	13	8
X	5068.	4235.	4109.	4015.	3154.	3091.	5728.	1113.5	.18	2	16	16
Y	5042.	4320.	4163.	4111.	3252.	3136.	5376.	846.2	.13	2	14	13
Z	5688.	4598.	4717.	4761.	4008.	4220.	5608.	438.3	.06	2	22	19
AA	4812.	4218.	4248.	4326.	3111.	3133.	5552.	927.6	.15	2	17	16
AB	5417.	4644.	4762.	4943.	4029.	4029.	5776.	458.6	.06	2	19	16
AC	4808.	4269.	4236.	4215.	3219.	3246.	5407.	756.2	.12	2	15	14
AD	4880.	3838.	4101.	4198.	3303.	3714.	5366.	613.0	.10	2	15	15
AE	5170.	4015.	4068.	4107.	3005.	3230.	5805.	1213.5	.19	2	21	21
AF	4569.	3824.	3739.	3810.	3915.	4062.	5481.	486.4	.08	2	16	15
AG	4618.	4299.	4285.	4303.	3176.	3191.	5528.	823.0	.13	2	16	16
AH	4380.	4239.	4389.	4257.	3072.	2974.	6091.	1312.0	.21	2	19	18
AI	4797.	4247.	4153.	4165.	3160.	3234.	5644.	910.6	.14	2	15	15
AJ	5462.	4019.	3943.	4061.	3406.	3286.	5223.	861.9	.14	2	21	21
AK	6273.	3902.	3826.	3885.	3369.	3159.	4986.	1432.1	.23	2	21	21
AL	5524.	4295.	4797.	5443.	4279.	3796.	5460.	520.7	.07	2	24	23
AM	5791.	3922.	3938.	3957.	3071.	3179.	5542.	1398.0	.22	2	21	21
AN	5160.	3531.	3912.	4165.	3537.	3574.	5521.	822.4	.13	2	15	17
AO	5372.	3828.	3894.	3861.	3578.	3466.	5401.	834.4	.13	2	21	20

APPENDIX D

TABLE B-11
PREDICTIONS

TYPICAL INTERQUARTILE RANGE

RUN	OLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	.04	.03	.03	.03	.03	.03	.05
B	.04	.04	.04	.04	.04	.04	.05
C	.04	.04	.04	.04	.04	.04	.05
D	.04	.04	.04	.04	.04	.03	.05
E	.02	.02	.02	.02	.02	.02	.03
F	.02	.02	.02	.02	.02	.02	.03
G	.04	.03	.02	.02	.03	.02	.02
H	.02	.02	.02	.02	.02	.02	.03
I	.01	.02	.02	.02	.01	.01	.03
J	.05	.03	.02	.02	.02	.03	.03
K	.03	.02	.02	.02	.02	.02	.02
L	.01	.01	.01	.01	.01	.01	.01
M	.01	.01	.01	.01	.01	.01	.01
N	.01	.02	.02	.02	.01	.01	.02
O	.03	.03	.03	.03	.02	.02	.03
P	.02	.02	.02	.02	.02	.02	.02
Q	.03	.03	.03	.03	.03	.03	.04
R	.00	.00	.00	.00	.00	.00	.01
S	.01	.01	.01	.01	.01	.01	.02
T	.01	.01	.01	.01	.01	.01	.02
U	.01	.01	.01	.01	.01	.01	.02
V	.01	.01	.01	.01	[illegible]	.01	.01
W	.01	.01	.01	.01	.01	.01	.01
X	.04	.04	.04	.04	.03	.03	.06
Y	.05	.04	.04	.04	.04	.04	.05
Z	.03	.03	.03	.03	.02	.02	.03
AA	.04	.04	.04	.04	.03	.03	.05
AB	.03	.03	.03	.03	.03	.03	.04
AC	.02	.02	.02	.02	.02	.02	.05
AD	.01	.01	.01	.01	.01	.01	.02
AE	.01	.01	.01	.01	.01	.01	.01
AF	.01	.01	.01	.01	.01	.01	.01
AG	.03	[illegible]	.03	.03	.02	.02	.03
AH	.02	.02	.02	.02	.02	.02	.03
AI	.02	.02	.02	.02	.02	.02	.02
AJ	.04	.04	.05	.05	.04	.04	.05
AK	.05	.05	.05	.06	.04	.04	.05
AL	.03	.02	.03	.04	.02	.02	.03
AM	.01	.01	.01	.01	.01	.01	.01
AN	.01	.01	.01	.01	.01	.01	.02
AO	.01	.01	.01	.02	.01	.01	.02

APPENDIX D

TABLE B-12
PREDICTIONS

RANKINGS ON ALL ESTIMATES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	CHI SQ	W	SIG	X95	X99
A	1899.	1768.	1782.	1825.	1560.	1496.	2270.	182.2	.07	2	7	5
B	2000.	1703.	1718.	1692.	1689.	1685.	2113.	91.1	.03	2	5	4
C	1848.	1827.	1793.	1680.	1760.	1549.	2143.	95.1	.04	2	6	4
D	1896.	1737.	1781.	1790.	1598.	1589.	2209.	126.8	.05	2	4	3
E	1780.	1674.	1724.	1830.	1709.	1687.	2196.	95.6	.04	2	5	5
F	1980.	1709.	1786.	1824.	1563.	1552.	2186.	146.7	.05	2	6	6
G	2099.	1785.	1821.	1596.	1782.	1659.	1858.	73.9	.03	2	5	4
H	2569.	2385.	2402.	2477.	2217.	2124.	2626.	69.8	.02	2	8	8
I	2327.	1767.	1691.	1708.	1362.	1316.	2429.	533.8	.20	2	9	9
J	2187.	1752.	1614.	1711.	1748.	1743.	1848.	96.8	.04	2	6	4
K	2147.	1852.	1693.	1647.	1974.	1587.	1700.	116.2	.04	2	7	4
L	1712.	1696.	1744.	1804.	1792.	1659.	2193.	93.4	.03	2	5	5
M	2165.	1678.	1761.	1808.	1514.	1557.	2117.	186.2	.07	2	8	5
N	3204.	2856.	3157.	3323.	2717.	2583.	3160.	134.5	.03	2	10	6
O	1865.	1746.	1759.	1847.	1563.	1698.	2122.	86.5	.03	2	6	2
P	1762.	1723.	1759.	1780.	1693.	1314.	2064.	42.9	.02	2	4	2
Q	2132.	1773.	1734.	1764.	1592.	1526.	2077.	148.9	.06	2	7	5
R	2083.	1797.	1753.	1701.	1499.	1588.	2179.	176.8	.07	2	9	7
S	1794.	1719.	1821.	1813.	1622.	1704.	2147.	80.0	.03	2	6	2
T	1850.	1759.	1798.	1821.	1610.	1730.	2032.	47.4	.02	2	2	1
U	1928.	1749.	1763.	1745.	1596.	1654.	2165.	104.5	.04	2	7	3
V	2233.	1717.	1761.	1750.	1519.	1482.	2138.	234.6	.09	2	8	7
W	1882.	1762.	1728.	1700.	1751.	1817.	1760.	24.6	.01	2	2	2
X	1901.	1755.	1732.	1739.	1497.	1546.	2330.	214.2	.08	2	9	7
Y	1978.	1817.	1806.	1738.	1413.	1484.	2314.	260.8	.10	2	9	8
Z	3156.	3050.	3127.	3149.	2460.	2486.	3572.	270.9	.06	2	11	8
AA	1926.	1804.	1858.	1911.	1455.	1439.	2207.	212.7	.08	2	7	7
AB	2586.	2486.	2331.	2374.	2204.	2164.	2735.	88.9	.02	2	4	4
AC	1807.	1687.	1778.	1826.	1677.	1722.	2173.	93.5	.03	2	6	5
AD	1874.	1705.	1818.	1777.	1649.	1816.	1761.	39.6	.01	2	6	2
AE	1957.	1801.	1828.	1829.	1497.	1587.	2101.	121.0	.04	2	6	3
AF	1892.	1671.	1728.	1791.	1739.	1814.	1765.	29.3	.01	2	1	1
AG	1756.	1813.	1873.	1930.	1514.	1539.	2125.	140.7	.05	2	5	3
AH	1798.	1749.	1805.	1799.	1562.	1633.	2284.	158.3	.06	2	5	3
AI	1650.	1709.	1827.	1878.	1686.	1739.	2091.	67.9	.03	2	5	3
AJ	1926.	1781.	1838.	1888.	1539.	1581.	2047.	96.4	.04	2	6	6
AK	2789.	1680.	1737.	1844.	1375.	1378.	1827.	648.3	.24	2	9	9
AL	2308.	2356.	2567.	2858.	2070.	2031.	2698.	171.8	.05	2	6	7
AM	2313.	1818.	1752.	1703.	1465.	1439.	2363.	262.7	.10	2	9	9
AN	1848.	1689.	1821.	1767.	1734.	1621.	2120.	73.8	.03	2	7	2
AO	1731.	1692.	1791.	1830.	1716.	1617.	2223.	112.8	.04	2	6	4

APPENDIX D

TABLE C-1

ESTIMATED STRUCTURAL COEFFICIENTS

TYPICAL MEDIAN BIASES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	-.13	-.00	.01	.02	.00	.00
B	-.13	-.01	.00	.01	-.01	-.00
C	-.15	-.01	-.00	.01	-.02	.00
D	-.20	-.02	-.01	.00	-.04	.00
E	-.29	-.07	-.06	-.04	-.07	-.02
F	-.34	-.05	-.03	-.00	-.01	.04
G	-.27	-.09	-.06	-.03	-.07	-.02
H	-.13	-.03	-.03	-.03	-.03	-.03
I	-.41	-.61	-.54	-.53	-.73	-.31
J	-.94	-1.04	-1.13	-1.34	-.79	-.69
K	-1.61	-2.14	-2.30	-2.05	-1.85	-1.96
L	-.51	-.64	-.65	-.06	-.76	-.74
M	-.09	-.03	-.02	-.01	-.04	.03
N	-.41	-.51	-.50	-.44	-.60	-.54
O	-.15	-.01	.00	.01	.03	.01
P	-.12	-.03	-.03	-.02	-.04	-.02
Q	-.24	-.06	-.05	-.02	-.05	-.01
R	-.19	-.09	-.07	-.04	-.06	-.04
S	.05	.00	-.00	-.00	.02	.04
T	-.05	.01	.01	.01	.00	.00
U	-.01	.01	.01	.01	-.01	-.00
V	-.03	-.01	-.01	-.01	-.01	.00
W	-.02	.00	.00	.01	-.00	-.00
X	-.02	-.01	-.01	-.00	-.00	-.01
Y	-.13	.00	.01	.02	-.01	.00
Z	-.14	-.01	-.00	.01	-.00	.01
AA	-.34	-.03	-.01	.03	.02	.05
AB	-.07	-.01	-.01	.01	-.01	-.00
AC	-.07	.00	.01	.01	-.00	-.00
AD	-.12	-.03	-.02	-.01	-.01	-.01
AE	-.35	-.09	-.07	-.04	-.05	.01
AF	-.09	-.00	.00	.01	-.02	-.00
AG	-.03	-.00	-.00	.00	-.01	-.02
AH	-.11	-.04	-.04	-.01	-.02	.01
AI	-.04	-.01	-.01	-.03	-.02	.01
AJ	-.05	.03	.03	.05	-.00	.01
AK	-.11	-.07	-.07	-.05	-.06	-.05
AL	-.03	.00	.00	.00	-.00	-.00
AM	-.01	.00	.00	.00	.01	.01
AN	-.01	-.00	-.00	-.00	[illegible]	-.00
AO	-.13	-.03	-.02	.03	-.02	-.01
AP	-.15	-.04	-.03	.03	-.00	.01
AQ	-.04	-.01	-.01	-.00	-.03	-.03
AR	-.02	-.01	-.01	-.01	-.01	-.00
AS	-.07	-.02	-.02	.00	-.02	.01
AT	-.15	-.03	-.03	-.01	-.02	-.01
AU	-.13	-.02	-.01	-.01	-.01	-.00
AV	-.10	-.01	-.00	.01	-.01	.01
AW	-.28	-.07	-.06	.03	-.05	-.02

APPENDIX D

TABLE C-2

ESTIMATED STRUCTURAL COEFFICIENTS

TYPICAL INTERQUARTILE RANGE

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	.39	.44	.44	.45	.29	.29
B	.38	.46	.47	.48	.32	.32
C	.39	.46	.47	.47	.38	.38
D	.44	.52	.54	.54	.49	.50
E	.50	.68	.71	.72	.66	.67
F	.36	.51	.54	.57	.52	.58
G	.23	.33	.36	.39	.30	.32
H	.39	.46	.46	.47	.47	.48
I	2.70	2.96	3.29	4.40	2.24	3.96
J	5.64	6.61	8.43	10.09	4.65	9.30
K	9.80	11.51	15.14	17.98	8.67	16.56
L	2.63	2.86	3.30	4.61	2.40	5.16
M	.77	1.25	1.44	1.85	1.50	2.75
N	2.66	2.91	3.04	4.63	2.21	4.63
O	.36	.42	.43	.45	.32	.32
P	.29	.34	.34	.35	.22	.22
Q	.18	.29	.31	.32	.27	.31
R	.20	.31	.38	.43	.32	.38
S	.50	.58	.58	.59	.44	.44
T	.24	.26	.27	.27	.17	.16
U	.31	.34	.34	.34	.14	.13
V	.11	.15	.15	.16	.13	.09
W	.13	.14	.14	.13	.09	.09
X	.14	.13	.13	.13	.10	.10
Y	.33	.41	.42	.45	.31	.32
Z	.36	.41	.42	.45	.29	.29
AA	.33	.49	.52	.57	.48	.53
AB	.23	.39	.42	.47	.38	.47
AC	.22	.25	.25	.26	.17	.17
AD	.68	.74	.74	.75	.36	.27
AE	.56	.72	.73	.75	.44	.46
AF	.29	.32	.33	.34	.17	.16
AG	.22	.22	.23	.22	.13	.11
AH	.75	.81	.81	.86	.59	.54
AI	.61	.72	.72	.65	.62	.56
AJ	.63	.65	.65	.71	.53	.46
AK	.61	.65	.66	.66	.61	.64
AL	.17	.18	.18	.18	.15	.16
AM	.10	.10	.10	.10	.09	.09
AN	.09	.10	.10	.12	.08	.08
AO	.43	.49	.50	.55	.42	.38
AP	.40	.50	.51	.56	.40	.33
AQ	.23	.28	.26	.32	.22	.20
AR	.27	.28	.29	.30	.28	.29
AS	.52	.53	.53	.54	.51	.50
AT	.20	.20	.22	.20	.16	.19
AU	.16	.20	.20	.20	.17	.17
AV	.17	.88	.91	1.24	.24	.22
AW	.40	.46	.47	.44	.36	.39

APPENDIX D

TABLE C-3

ESTIMATED STRUCTURAL COEFFICIENTS

RANKINGS ON ALL ESTIMATES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	CHI SQ	W	SIG	X95	X99
A	3334.	2751.	2698.	2771.	2078.	2118.	419.6	.11	2	15	14
B	3345.	271Ø.	2626.	2706.	2192.	2171.	352.7	.Ø9	2	15	15
C	3282.	2692.	2582.	2648.	223Ø.	2316.	262.9	.Ø7	2	14	12
D	3257.	2589.	2619.	2575.	23Ø6.	24Ø4.	211.Ø	.Ø6	2	1Ø	8
E	3172.	247Ø.	2597.	2596.	2358.	2557.	152.7	.Ø4	2	11	9
F	3511.	2522.	2469.	2498.	23Ø9.	2441.	369.4	.1Ø	2	15	12
G	3532.	2682.	2531.	2581.	2219.	22Ø5.	448.7	.12	2	12	12
H	2928.	2496.	2473.	258Ø.	259Ø.	2683.	52.6	.Ø1	2	6	5
I	2174.	2271.	2763.	3111.	2248.	3183.	395.2	.11	2	14	13
J	218Ø.	2327.	2941.	3Ø49.	2161.	3Ø92.	38Ø.9	.1Ø	2	13	13
K	21Ø6.	2336.	2941.	3118.	2145.	31Ø4.	44Ø.2	.12	2	15	12
L	212Ø.	2255.	2722.	3124.	2293.	3236.	432.5	.12	2	14	14
M	2119.	2ØØ9.	2531.	2753.	244Ø.	2848.	228.9	.Ø7	2	11	1Ø
N	2119.	2262.	269Ø.	3114.	2358.	32Ø7.	396.6	.11	2	14	14
O	3375.	2718.	2641.	269Ø.	2154.	2172.	382.3	.1Ø	2	15	14
P	3262.	2721.	2755.	2759.	2149.	21Ø4.	361.1	.12	2	13	12
Q	3599.	2676.	2539.	2546.	22Ø5.	2135.	5Ø8.5	.14	2	14	13
R	3275.	2545.	2432.	2473.	2465.	256Ø.	197.7	.Ø5	2	14	12
S	3146.	2637.	2591.	2651.	2318.	24Ø7.	158.2	.Ø4	2	1Ø	1Ø
T	3143.	285Ø.	2815.	2808.	2Ø67.	2Ø67.	385.2	.1Ø	2	14	1Ø
U	2977.	2846.	2916.	2955.	22ØØ.	2856.	411.7	.11	2	14	13
V	2584.	2356.	2318.	2257.	1589.	1496.	475.2	.16	2	11	11
W	3Ø55.	2796.	2852.	28Ø5.	2299.	2142.	3Ø8.Ø	.Ø8	2	13	11
X	355Ø.	3Ø51.	3Ø58.	3Ø14.	2534.	2643.	218.3	.Ø5	2	11	9
Y	3328.	262Ø.	2637.	2668.	217Ø.	2357.	295.3	.Ø8	2	14	12
Z	337Ø.	2732.	2655.	2739.	2145.	21Ø9.	412.5	.11	2	15	15
AA	3498.	2519.	2473.	2546.	2356.	2358.	36Ø.5	.1Ø	2	14	12
AB	3529.	2469.	243Ø.	2535.	2346.	2441.	38Ø.7	.1Ø	2	14	14
AC	311Ø.	2788.	2797.	28Ø4.	2142.	2111.	313.5	.Ø8	2	12	1Ø
AD	3252.	2921.	2799.	2842.	2Ø54.	1864.	545.7	.15	2	15	14
AE	3524.	277Ø.	2613.	2624.	2291.	2128.	518.7	.14	2	14	12
AF	3163.	2898.	2865.	2917.	1943.	1964.	536.7	.14	2	13	13
AG	3Ø93.	2953.	2882.	2925.	1996.	1921.	534.3	.14	2	14	14
AH	2874.	2768.	2698.	28Ø7.	2264.	2399.	112.Ø	.Ø3	2	9	8
AI	34Ø6.	3Ø26.	2967.	3174.	2665.	26Ø9.	153.4	.Ø4	2	12	11
AJ	2918.	2717.	2714.	2815.	2291.	2295.	136.7	.Ø4	2	11	8
AK	2676.	2644.	2679.	2733.	2479.	2539.	17.6	.ØØ	2	2	Ø
AL	3Ø6Ø.	2723.	2695.	2737.	2236.	2299.	18Ø.5	.Ø5	2	9	8
AM	2832.	2692.	2743.	2815.	2255.	2413.	136.4	.Ø3	2	1Ø	9
AN	2861.	2718.	27Ø2.	2738.	2315.	2416.	84.9	.Ø2	2	1Ø	1Ø
AO	3153.	268Ø.	2593.	2717.	24Ø7.	223Ø.	197.7	.Ø5	2	9	9
AP	33Ø2.	2818.	2639.	2681.	2275.	1985.	394.3	.11	2	13	8
AQ	3225.	28Ø6.	2721.	278Ø.	2169.	2749.	367.9	.1Ø	2	15	12
AR	2763.	2659.	2646.	2722.	2437.	2523.	28.9	.Ø1	2	3	1
AS	2959.	2651.	2625.	2577.	2467.	2471.	62.2	.Ø2	2	5	5
AT	3578.	2631.	2483.	2424.	2366.	2268.	443.2	.12	2	12	1Ø
AU	35Ø4.	2638.	257Ø.	2549.	2181.	23Ø8.	411.1	.11	2	12	1Ø
AV	2792.	3155.	3596.	4Ø4Ø.	2166.	21Ø1.	1Ø29.8	.24	2	15	15
AW	3577.	28Ø2.	2583.	2394.	2148.	2246.	519.6	.14	2	14	11

APPENDIX D

TABLE C-4

ESTIMATED STANDARD ERRORS

PROPORTION OF BIAS/S.E. OUTSIDE MODEL T

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	.20	.07	.06	.06	.05	.04
B	.20	.06	.06	.05	.06	.05
C	.21	.06	.05	.05	.06	.05
D	.21	.07	.06	.05	.06	.06
E	.24	.08	.08	.07	.09	.08
F	.35	.07	.07	.05	.09	.06
G	.43	.12	.10	.06	.13	.08
H	.08	.04	.04	.04	.05	.05
I	.20	.07	.06	.04	.12	.15
J	.19	.09	.08	.07	.12	.15
K	.17	.07	.06	.05	.09	.13
L	.19	.07	.07	.04	.11	.12
M	.11	.02	.02	.02	.05	.05
N	.19	.07	.07	.05	.13	.14
O	.20	.05	.05	.05	.05	.04
P	.11	.05	.05	.04	.03	.03
Q	.35	.08	.07	.06	.08	.07
R	.43	.20	.15	.08	.22	.14
S	.09	.05	.04	.04	.04	.03
T	.12	.06	.06	.06	.06	.04
U	.08	.07	.07	.07	.04	.03
V	.15	.11	.11	.09	.09	.09
W	.07	.05	.05	.05	.04	.04
X	.02	.01	.01	.01	.02	.01
Y	.18	.05	.05	.03	.04	.05
Z	.20	.06	.06	.05	.05	.04
AA	.35	.06	.05	.05	.08	.05
AB	.30	.05	.05	.03	.07	.03
AC	.12	.06	.05	.05	.05	.04
AD	.18	.14	.14	.13	.10	.05
AE	.20	.06	.06	.06	.05	.04
AF	.17	.12	.12	.11	.09	.07
AG	.22	.20	.20	.20	.19	.17
AH	.38	.37	.37	.37	.41	.35
AI	.30	.22	.22	.19	.25	.14
AJ	.35	.33	.33	.34	.37	.32
AK	.31	.31	.30	.30	.33	.29
AL	.15	.11	.10	.10	.09	.07
AM	.15	.15	.15	.16	.17	.17
AN	.17	.17	.17	.17	.18	.17
AO	.33	.23	.23	.21	.29	.27
AP	.38	.23	.23	.18	.24	.16
AQ	.24	.17	.17	.16	.16	.13
AR	.19	.18	.17	.18	.19	.18
AS	.26	.22	.22	.21	.22	.18
AT	.22	.05	.05	.05	.06	.05
AU	.22	.05	.05	.04	.05	.06
AV	.11	.18	.19	.19	.01	.03
AW	.47	.17	.17	.13	.19	.17

APPENDIX D

TABLE C-5
ESTIMATED REDUCED FORM

TYPICAL MEDIAN BIASES

RUN	OLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	.10	.05	.04	.03	.01	.01	.28
B	.08	.03	.03	.02	-.01	-.02	.18
C	.05	.03	.02	.02	-.03	-.01	.12
D	.02	-.02	-.03	-.04	-.03	-.02	-.12
E	-.06	-.12	-.13	-.14	-.13	-.13	-.49
F	-.09	-.00	.01	.01	-.01	.06	.11
G	.12	.06	.03	.00	.04	-.01	.19
H	-.05	-.04	-.03	-.04	-.03	-.01	-.02
I	.37	1.11	-.37	-.47	-1.08	-.62	.55
J	7.88	-3.04	2.56	-.45	.97	-.40	-.86
K	-4.07	1.48	.17	-6.66	-.37	-.75	4.43
L	-.44	1.39	.14	-.00	-1.45	-.44	1.22
M	.81	.17	.07	-.03	.39	-.21	.05
N	.86	.48	-.34	-.15	-1.08	-.75	1.59
O	.06	.03	.02	.02	.02	.02	.16
P	-.02	-.04	-.04	-.04	-.02	-.03	-.13
Q	.20	.03	.02	.01	.04	.00	.0[illegible]
R	.77	.26	.2[illegible]	.07	.31	.05	.14
S	.24	.05	.04	.12	.04	.00	.2[illegible]
T	.05	.03	.03	.03	.01	.00	.00
U	.05	.03	.03	.02	-.01	-.01	.28
V	.06	.01	.01	.00	.01	-.02	-.21
W	.01	.00	.00	.00	.00	-.03	.06
X	.01	-.01	-.01	-.01	-.00	-.02	-.01
Y	.12	.05	.04	.03	.02	.02	.25
Z	.08	.03	.03	.03	.01	.0[illegible]	.22
AA	-.06	.01	.02	.03	.03	.05	-.67
AB	.40	.06	.04	.00	.06	-.01	-.04
AC	.02	.01	.01	.01	.00	-.03	.02
AD	-.11	-.02	-.02	-.00	.03	-.00	-.17
AE	-.47	-.14	-.13	-.09	-.07	-.01	-.47
AF	-.04	-.20	.00	.00	-.01	-.01	-.21
AG	-.01	-.00	-.00	-.00	-.00	-.0[illegible]	.05
AH	-.10	-.07	-.07	-.06	-.05	-.02	.39
AI	.25	.08	.08	-.02	.12	-.01	-.05
AJ	.04	.04	.04	.06	.04	.01	.45
AK	-.20	-.12	-.11	-.08	-.08	-.06	.12
AL	.02	.01	.00	.00	.01	.02	.30
AM	-.01	-.01	-.01	-.01	-.01	-.01	-.04
AN	-.00	-.00	-.00	-.00	.00	.00	-.04
AO	.04	.01	.01	.00	.02	-.01	.03
AP	.04	-.02	-.02	.01	.01	-.02	.07
AQ	.08	.00	.02	-.01	.02	.02	-.04
AR	.02	-.01	-.01	-.01	-.01	-.01	-.03
AS	.05	.03	.03	.02	.00	.02	.13
AT	.01	.01	.00	.01	-.01	-.01	-.13
AU	-.00	-.01	-.01	-.01	-.01	-.01	-.06
AV	.02	-.03	-.03	-.05	.01	-.01	.01
AW	.00	.04	.04	.05	-.01	-.03	.25

APPENDIX D

TABLE C -6

ESTIMATED REDUCED FORM

TYPICAL INTERQUARTILE RANGE

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	.60	.59	.60	.59	.31	.30	2.01
B	.59	.61	.61	.61	.37	.33	1.94
C	.61	.59	.59	.59	.45	.43	1.89
D	.67	.69	.69	.69	.60	.58	1.75
E	.91	.83	.83	.85	.73	.72	2.13
F	.56	.65	.65	.68	.64	.67	6.56
G	.48	.43	.45	.46	.39	.37	1.21
H	.51	.54	.55	.54	.54	.55	1.41
I	39.97	19.21	21.42	20.39	26.78	11.85	26.49
J	321.04	162.37	120.19	102.36	136.94	86.90	134.38
K	168.56	94.70	81.97	87.52	104.40	43.16	93.50
L	35.29	19.39	21.40	16.75	25.23	9.27	25.85
M	1.87	1.63	1.78	1.98	1.70	1.84	2.68
N	39.80	21.81	20.06	14.32	27.79	9.93	26.29
O	.56	.56	.56	.58	.35	.35	1.87
P	.37	.35	.36	.37	.21	.23	1.58
Q	.26	.31	.31	.31	.26	.27	.48
R	.78	.62	.74	.79	.62	.73	1.63
S	.78	.73	.73	.73	.48	.45	1.86
T	.29	.30	.30	.30	.16	.15	1.24
U	.69	.66	.66	.67	.23	.19	1.66
V	.15	.17	.18	.17	.13	.12	.36
W	.16	.16	.16	.16	.09	.08	.92
X	.14	.12	.12	.12	.09	.09	.22
Y	.53	.55	.55	.56	.38	.38	2.07
Z	.58	.57	.57	.59	.31	.30	1.98
AA	.56	.61	.62	.61	.55	.60	10.33
AB	.39	.46	.48	.48	.43	.43	.88
AC	.28	.30	.30	.30	.18	.17	1.16
AD	1.30	1.26	1.25	1.25	.56	.41	2.75
AE	1.24	1.11	1.12	1.09	.66	.60	3.81
AF	.33	.34	.34	.34	.20	.17	1.15
AG	.27	.28	.28	.28	.15	.12	1.13
AH	1.24	1.21	1.20	1.19	.77	.52	2.77
AI	.85	.89	.89	.92	.72	.53	1.35
AJ	1.07	.99	.98	1.01	.69	.50	2.68
AK	.75	.77	.77	.77	.77	.78	1.98
AL	.20	.20	.20	.20	.15	.14	.68
AM	.12	.12	.12	.12	.10	.09	.37
AN	.11	.11	.11	.11	.08	.08	.34
AO	.55	.57	.57	.59	.44	.36	2.66
AP	.54	.56	.56	.56	.38	.29	2.76
AQ	.28	.33	.33	.36	.25	.22	.85
AR	.24	.26	.26	.26	.25	.26	.65
AS	.53	.50	.50	.51	.48	.45	2.00
AT	.21	.20	.20	.21	.17	.17	.97
AU	.18	.19	.19	.19	.15	.16	.78
AV	.18	.72	.73	.78	.23	.22	.46
AW	.48	.48	.49	.49	.36	.38	3.57

APPENDIX D

TABLE C-7
ESTIMATED REDUCED FORM

RANKINGS ON ALL ESTIMATES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	CHI SQ	W	SIG	X95	X99
A	5020.	4251.	4135.	4137.	3127.	3211.	5519.	929.1	.15	2	16	15
B	4949.	4162.	4105.	4112.	3289.	3305.	5478.	784.4	.12	2	17	16
C	4833.	4106.	4067.	4007.	3455.	3532.	5400.	593.0	.09	2	17	16
D	4686.	3997.	4141.	4016.	3545.	3775.	5240.	409.4	.06	2	16	15
E	4447.	3773.	4124.	4178.	3605.	4138.	5135.	302.4	.05	2	15	15
F	4356.	3730.	3887.	3991.	3592.	3894.	5950.	798.5	.13	2	16	15
G	5237.	4078.	3988.	3994.	3621.	3424.	5058.	581.9	.09	2	15	15
H	4383.	3883.	3949.	4063.	4025.	4029.	5068.	210.[illegible]	.03	2	9	9
I	5618.	4522.	3907.	3833.	4275.	3334.	3911.	647.8	.10	2	20	18
J	5659.	4644.	4139.	3737.	4271.	3121.	3829.	785.9	.12	2	20	19
K	5549.	4432.	3847.	3851.	4259.	3264.	4198.	612.2	.10	2	19	16
L	5515.	4452.	4048.	3864.	4378.	3179.	3964.	624.2	.10	2	20	18
M	5093.	3737.	4085.	4280.	3950.	3938.	4317.	240.1	.04	2	17	14
N	5658.	4434.	3954.	3661.	4453.	3219.	4021.	732.6	.12	2	21	20
O	5045.	4165.	4065.	4353.	3325.	3314.	5423.	775.2	.12	2	17	17
P	4806.	4272.	4235.	4180.	3209.	3235.	5465.	793.5	.13	2	17	13
Q	5787.	4065.	3870.	3835.	3571.	3421.	4831.	847.9	.13	2	18	18
R	5511.	3925.	3793.	3776.	3927.	3688.	4788.	574.0	.09	2	19	18
S	5084.	4165.	4022.	3953.	3514.	3503.	5159.	561.5	.09	2	15	14
T	4569.	4266.	4233.	4155.	3058.	3103.	6036.	1229.4	.23	2	18	18
U	4422.	4261.	4330.	4282.	3062.	3138.	5985.	1183.4	.19	2	19	18
V	5009.	4222.	4136.	4062.	3204.	2959.	5838.	1203.7	.19	2	21	21
W	4408.	4141.	4209.	4110.	3106.	3162.	6266.	1347.5	.21	2	21	21
X	5336.	4738.	4747.	4752.	4234.	4311.	5482.	235.7	.03	2	15	17
Y	5005.	4141.	4314.	3996.	3269.	3511.	5464.	748.3	.12	2	17	16
Z	5097.	4194.	4036.	4094.	3257.	3243.	5529.	920.6	.15	2	17	17
AA	4271.	3784.	3915.	3897.	3656.	3575.	6280.	1089.4	.17	2	18	18
AB	5860.	3992.	3784.	3600.	3966.	3499.	4699.	642.3	.13	2	21	21
AC	4472.	4166.	4197.	4173.	3393.	3223.	6077.	1181.[illegible]	.19	2	19	13
AD	4741.	4434.	4348.	4233.	3199.	2872.	5523.	998.4	.16	2	16	13
AE	4983.	4174.	4123.	4117.	3313.	3384.	5329.	686.5	.11	2	14	12
AF	4380.	4283.	4363.	4449.	2980.	2933.	6312.	1327.5	.21	2	21	21
AG	4321.	4242.	4204.	4262.	3266.	2982.	6363.	1194.1	.19	2	19	18
AH	4412.	4276.	4210.	4140.	3637.	3545.	5252.	476.4	.06	2	15	14
AI	5592.	4770.	4750.	5075.	4361.	3956.	5096.	333.4	.04	2	21	21
AJ	4454.	4234.	4190.	4136.	3627.	3527.	5232.	391.1	.06	2	12	11
AK	4211.	4055.	4242.	4155.	4209.	4135.	4783.	87.3	.01	2	11	8
AL	4640.	4273.	4211.	4266.	3402.	3332.	5276.	561.5	.09	2	15	13
AM	4141.	4224.	4411.	4423.	3449.	3451.	5371.	497.0	.08	2	21	21
AN	4335.	4292.	4366.	4392.	3418.	3405.	5192.	473.2	.08	2	15	14
AO	4730.	4242.	4137.	4149.	3738.	3426.	4978.	348.4	.06	2	17	16
AP	5016.	4404.	4257.	3763.	3458.	3024.	5278.	788.5	.13	2	19	15
AQ	4391.	4190.	4273.	4339.	3429.	3199.	5529.	722.1	.11	2	19	17
AR	4048.	3992.	4045.	4339.	3881.	4037.	5258.	198.5	.03	2	14	12
AS	4897.	4191.	3979.	3928.	3358.	3821.	4746.	241.4	.04	2	15	12
AT	4931.	4091.	4278.	4043.	3575.	3571.	5111.	449.4	.07	2	15	11
AU	4735.	4373.	4022.	4132.	3474.	3593.	5459.	599.1	.10	2	15	14
AV	3642.	3061.	3945.	6419.	3737.	3602.	4594.	1539.7	.21	2	24	23
AW	4067.	4166.	4258.	4252.	3546.	3626.	5555.	534.6	.08	2	22	23

APPENDIX D

TABLE C-8

PREDICTIONS

TYPICAL INTERQUARTILE RANGE

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	.04	.03	.03	.03	.03	.03	.05
B	.04	.03	.03	.03	.03	.03	.05
C	.04	.04	.04	.04	.03	.03	.05
D	.04	.04	.04	.04	.03	.04	.05
E	.04	.04	.04	.04	.04	.04	.05
F	.02	.02	.02	.02	.02	.02	.03
G	.05	.05	.05	.04	.04	.04	.06
H	.03	.03	.03	.03	.03	.03	.04
I	1.24	.65	.61	.63	.79	.42	.59
J	7.81	3.98	3.77	2.92	3.56	2.82	2.98
K	5.42	2.34	2.24	2.13	2.71	1.78	2.04
L	1.27	.75	.70	.66	.82	.42	.60
M	.13	.12	.12	.14	.13	.13	.16
N	1.19	.67	.57	.52	.80	.44	.60
O	.03	.03	.03	.04	.03	.03	.05
P	.02	.02	.02	.02	.02	.02	.03
Q	.03	.03	.03	.03	.03	.03	.04
R	.04	.04	.04	.04	.04	.04	.04
S	.04	.04	.04	.04	.03	.03	.04
T	.02	.02	.02	.02	.02	.02	.04
U	.02	.02	.02	.02	.02	.02	.03
V	.02	.02	.02	.02	.01	.01	.04
W	.02	.02	.02	.02	.02	.02	.03
X	.02	.02	.02	.02	.02	.02	.02
Y	.03	.03	.03	.03	.03	.03	.05
Z	.04	.04	.04	.04	.03	.03	.05
AA	.02	.02	.02	.02	.02	.02	.03
AB	.03	.03	.03	.03	.03	.02	.05
AC	.02	.02	.02	.02	.02	.02	.04
AD	.03	.03	.03	.03	.02	.02	.04
AE	.05	.05	.05	.05	.04	.04	.06
AF	.03	.03	.03	.03	.02	.02	.04
AG	.02	.02	.02	.02	.02	.01	.03
AH	.02	.02	.02	.02	.02	.02	.02
AI	.02	.02	.02	.02	.02	.02	.02
AJ	.02	.02	.02	.02	.02	.02	.02
AK	.01	.01	.01	.01	.01	.02	.02
AL	.01	.01	.01	.01	.01	.01	.01
AM	.01	.01	.01	.01	.01	.01	.01
AN	.01	.01	.01	.01	.01	.01	.01
AO	.01	.01	.01	.01	.01	.01	.01
AP	.01	.01	.01	.01	.01	.01	.01
AQ	.01	.01	.01	.01	.01	.01	.01
AR	.01	.01	.01	.01	.01	.01	.01
AS	.01	.01	.01	.01	.01	.01	.01
AT	.01	.01	.01	.01	.01	.01	.01
AU	.01	.01	.01	.01	.01	.01	.01
AV	.00	.00	.00	.00	.00	.00	.00
AW	.01	.01	.01	.01	.01	.01	.01

APPENDIX D

TABLE C-9

PREDICTIONS

RANKINGS ON ALL ESTIMATES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	CHI SQ	W	SIG	X95	X99
A	1899.	1768.	1782.	1825.	1560.	1496.	2270.	182.2	.07	2	7	5
B	1868.	1744.	1783.	1820.	1523.	1556.	2306.	190.8	.07	2	8	5
C	1845.	1714.	1749.	1800.	1542.	1630.	2320.	179.7	.07	2	8	6
D	1874.	1689.	1736.	1800.	1610.	1702.	2189.	104.2	.04	2	4	4
E	1983.	1681.	1732.	1739.	1596.	1760.	2109.	92.7	.03	2	6	5
F	1831.	1672.	1766.	1847.	1634.	1733.	2117.	73.0	.03	2	5	5
G	2032.	1757.	1749.	1690.	1630.	1561.	2181.	143.6	.05	2	7	7
H	1772.	1630.	1681.	1801.	1768.	1788.	2160.	83.1	.03	2	4	3
I	2345.	1932.	1671.	1612.	1892.	1452.	1696.	241.3	.09	2	8	9
J	2360.	1911.	1790.	1612.	1826.	1457.	1644.	240.0	.09	2	9	7
K	2341.	1831.	1676.	1626.	1853.	1509.	1764.	203.8	.08	2	8	7
L	2321.	1948.	1737.	1628.	1847.	1458.	1661.	221.6	.08	2	8	8
M	2094.	1639.	1735.	1756.	1752.	1671.	1953.	76.6	.03	2	7	6
N	2307.	1876.	1652.	1492.	1991.	1562.	1720.	228.2	.08	2	9	8
O	1890.	1807.	1850.	1840.	1480.	1478.	2255.	202.6	.08	2	7	7
P	1824.	1771.	1832.	1793.	1514.	1523.	2343.	217.1	.08	2	9	7
Q	2602.	1734.	1603.	1609.	1527.	1469.	2056.	463.1	.17	2	9	9
R	1955.	1708.	1674.	1645.	1755.	1757.	2106.	82.9	.03	2	5	4
S	2103.	1732.	1701.	1718.	1538.	1669.	2139.	149.4	.06	2	5	5
T	1783.	1772.	1812.	1805.	1484.	1534.	2410.	259.0	.10	2	9	8
U	1814.	1771.	1807.	1816.	1484.	1534.	2374.	238.8	.09	2	9	8
V	2041.	1789.	1732.	1679.	1501.	1433.	2425.	329.6	.12	2	9	9
W	1714.	1713.	1701.	1755.	1757.	1732.	2228.	103.1	.04	2	8	7
X	2566.	2460.	2464.	2475.	2135.	2127.	2573.	77.0	.02	2	7	4
Y	1842.	1735.	1759.	1515.	1543.	1647.	2259.	146.7	.05	2	7	5
Z	1886.	1796.	1804.	1812.	1537.	1488.	2277.	191.2	.07	2	7	7
AA	1812.	1666.	1741.	1835.	1672.	1615.	2259.	135.3	.05	2	6	7
AB	1936.	1745.	1774.	1756.	1652.	1579.	2237.	155.7	.06	2	9	8
AC	1753.	1762.	1802.	1813.	1483.	1527.	2455.	289.5	.11	2	9	8
AD	1908.	1817.	1832.	1826.	1332.	1443.	2442.	367.6	.14	2	9	9
AE	1939.	1814.	1828.	1806.	1426.	1518.	2269.	218.9	.08	2	9	8
AF	1817.	1836.	1928.	1938.	1373.	1322.	2386.	376.8	.14	2	8	8
AG	1779.	1749.	1766.	1922.	1514.	1545.	2425.	258.2	.10	2	9	8
AH	1861.	1687.	1611.	1576.	1862.	2094.	1929.	97.4	.04	2	6	4
AI	2618.	2540.	2542.	2545.	2317.	2263.	1978.	111.7	.03	2	7	5
AJ	1790.	1671.	1678.	1722.	1798.	1990.	1959.	47.2	.02	2	2	2
AK	1915.	1820.	1782.	1813.	1658.	1785.	1827.	16.8	.01	1	0	0
AL	1774.	1744.	1766.	1764.	1683.	1886.	1983.	29.2	.01	2	4	1
AM	1755.	1752.	1795.	1788.	1612.	1645.	2253.	129.1	.05	2	7	7
AN	1764.	1791.	1844.	1861.	1572.	1571.	2197.	128.1	.05	2	8	4
AO	1897.	1770.	1734.	1768.	1754.	1802.	1883.	12.2	.00	2	[illegible]	2
AP	2006.	1861.	1812.	1797.	1563.	1516.	2045.	115.8	.04	2	5	3
AQ	1772.	1774.	1802.	1823.	1621.	1726.	2102.	63.8	.02	2	5	2
AR	1819.	1825.	1818.	1866.	1685.	1626.	1961.	35.8	.01	2	4	3
AS	1920.	1797.	1756.	1839.	1632.	1681.	1975.	43.3	.02	2	1	1
AT	1795.	1732.	1791.	1792.	1700.	1682.	2118.	59.7	.02	2	3	3
AU	1788.	1796.	1852.	1834.	1656.	1677.	1997.	37.5	.01	2	3	3
AV	2076.	2617.	2832.	2889.	2031.	2210.	2145.	291.1	.08	2	11	9
AW	1955.	1839.	1783.	1881.	1452.	1584.	2106.	139.9	.05	2	7	9

APPENDIX D

TABLE D-1
ESTIMATED STRUCTURAL COEFFICIENTS

TYPICAL MEDIAN BIASES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	-.13	-.00	.01	.02	.00	.00
B	-.14	.00	.01	.02	-.05	-.06
C	-.10	-.00	-.00	.02	-.01	.00
D	.00	.04	.04	.04	.03	.04
E	-.11	.01	.01	.03	.01	.03
F	-.26	-.10	-.07	-.01	-.08	-.03
G	-.24	-.04	-.03	-.01	.00	.02
H	-.27	-.06	-.06	-.01	-.05	-.02
I	-.25	-.12	-.11	-.08	-.06	-.02
J	-.37	-.14	-.13	-.08	-.09	-.06
K	-.16	-.07	-.07	-.03	-.04	-.03
L	-.27	-.08	-.08	-.04	-.02	-.03
M	-.28	-.14	-.14	-.09	-.12	-.05
N	-.10	-.08	-.08	-.07	-.11	-.14
O	-.25	-.14	-.14	-.11	-.19	-.06
P	-.09	-.02	-.01	.00	-.02	-.01
Q	-.13	-.04	-.03	-.01	-.03	.00
R	-.21	-.12	-.12	-.11	-.12	-.11
S	1.27	1.32	1.33	1.34	1.36	1.21
T	1.13	1.48	1.52	3.01	1.37	1.47
U	1.69	1.85	1,86	2.30	1.68	1.64
V	-.12	.01	.02	.04	-.01	-.02
W	-.13	-.00	.00	.02	.00	.01
X	-.49	-.37	-.36	-.29	-.33	-.25
Y	-.12	.04	.04	.06	.07	.08
Z	-.11	.03	.04	.06	.05	.06
AA	-.15	-.06	-.05	.08	-.04	.01
AB	-.14	-.02	-.00	.04	.00	.01
AC	-.13	-.05	-.05	-.02	-.04	-.02
AD	-.24	.00	.00	.02	-.09	-.09
AE	-.17	-.01	-.01	.01	-.09	-.09
AF	-.01	-.03	-.03	-.03	-.02	-.02
AG	-.15	-.01	-.01	-.00	.01	.01
AH	-.32	-.07	-.06	.01	-.06	-.10
AI	-.35	-.04	-.03	.02	-.07	-.08
AJ	-.85	-.44	-.43	-.17	-.38	-.24
AK	-.30	-.04	-.02	.02	-.02	.03
AL	-.30	-.13	-.12	-.11	-.15	-.11
AM	-.45	-.22	-.22	-.22	-.22	-.21
AN	-.15	-.12	-.11	-.11	-.12	-.13
AO	-.44	-.18	-.15	-.12	-.20	-.14
AP	1.71	1.91	1.92	2.13	1.80	1.78
AQ	-.30	-.02	-.01	.04	-.02	-.06
AR	-.09	-.00	.00	.02	.01	.01
AS	-.04	.02	.02	.03	-.02	-.05
AT	-.13	.01	.02	.02	-.00	-.01
AU	-.04	.04	.04	.09	.08	.12
AV	-.10	18.11	19.08	20.84	-.08	.00

APPENDIX D

TABLE D-2
ESTIMATED STRUCTURAL COEFFICIENTS

TYPICAL INTERQUARTILE RANGE

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	.39	.44	.44	.45	.29	.29
B	.34	.39	.39	.41	.26	.30
C	.25	.36	.36	.41	.34	.45
D	.21	.24	.24	.25	.23	.24
E	.39	.44	.45	.47	.31	.30
F	.22	.30	.34	.34	.28	.29
G	.39	.44	.44	.48	.4[illegible]	.43
H	.22	.34	.34	.37	.34	.34
I	.33	.41	.41	.51	.42	.61
J	.26	.44	.46	.64	.55	.49
K	.20	.28	.28	.36	.27	.56
L	.14	.19	.19	.24	.20	.32
M	.20	.31	.32	.39	.33	.55
N	.37	.48	.48	.43	.44	.52
O	.37	.52	.53	.66	.58	.73
P	.24	.42	.45	.49	.38	.39
Q	.16	.23	.25	.25	.16	.18
R	.33	.39	.40	.44	.38	.46
S	.43	.48	.49	.54	.45	1.49
T	.73	.91	.94	2.45	.79	.57
U	.51	.58	.59	.93	.58	.59
V	.35	.41	.42	.43	.29	.5[illegible]
W	.39	.44	.45	.46	.36	.2[illegible]
X	.26	.53	.53	.91	.61	1.16
Y	.32	.42	.43	.45	.27	.29
Z	.16	.21	.22	.25	.22	.25
AA	.22	.32	.35	.45	.27	.27
AB	.23	.30	.32	.37	.28	.29
AC	.22	.28	.28	.29	.24	.27
AD	.33	.38	.39	.41	.22	.24
AE	.28	.31	.31	.32	.21	.2[illegible]
AF	.31	.32	.32	.32	.30	.3[illegible]
AG	.33	.35	.35	.37	.27	.29
AH	.32	.43	.45	.44	.38	.33
AI	.29	.43	.43	.47	.33	.34
AJ	.29	.67	.69	1.17	.79	.96
AK	.35	.51	.53	.57	.5[illegible]	.57
AL	.35	.58	.59	.64	.50	.66
AM	.22	.34	.34	.39	.3[illegible]	.38
AN	.23	.26	.26	.28	.25	.33
AO	.34	.53	.53	.65	.49	.77
AP	.62	.73	.73	.81	.65	.58
AQ	.30	.43	.46	.49	.49	.47
AR	.21	.25	.25	.28	.24	.28
AS	.24	.27	.27	.27	.17	.22
AT	.35	.37	.36	.41	.27	.31
AU	.16	.22	.2[illegible]	.21	.16	.17
AV	.21	479.19	516.1[illegible]	711.50	.54	.72

APPENDIX D

TABLE D-3
ESTIMATED STRUCTURAL COEFFICIENTS

RANKINGS ON ALL ESTIMATES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	CHI SQ	W	SIG	X95	X99
A	3334.	2751.	2698.	2771.	2078.	2118.	419.6	.11	2	15	14
B	3195.	2471.	2415.	2444.	2006.	2169.	339.9	.10	2	12	12
C	2918.	2381.	2374.	2497.	2238.	2292.	123.1	.04	2	12	12
D	4005.	3547.	3786.	3887.	3342.	3483.	89.9	.02	2	15	13
E	3340.	2793.	2691.	2761.	2074.	2091.	438.5	.12	2	15	14
F	3641.	2778.	2461.	2496.	2247.	2127.	567.7	.15	2	15	14
G	3441.	2534.	2500.	2521.	2344.	2410.	314.6	.08	2	12	11
H	3538.	2573.	2623.	2543.	2208.	2265.	436.8	.12	2	14	13
I	3189.	2514.	2528.	2427.	2515.	2577.	149.9	.04	2	12	7
J	3227.	2489.	2529.	2455.	2454.	2596.	171.1	.05	2	12	10
K	3175.	2519.	2534.	2457.	2450.	2615.	145.1	.04	2	7	5
L	3786.	2438.	2477.	2256.	2345.	2448.	628.8	.17	2	15	15
M	3365.	2564.	2606.	2537.	2237.	2441.	283.4	.08	2	14	13
N	2744.	2586.	2660.	2713.	2410.	2637.	27.1	.01	2	5	2
O	2757.	2392.	2431.	2752.	2776.	2642.	56.6	.02	2	4	4
P	3598.	2516.	2506.	2626.	2253.	2251.	476.6	.13	2	14	14
Q	2808.	2230.	2173.	2141.	1668.	1580.	467.7	.16	2	10	9
R	2791.	2521.	2462.	2602.	2541.	2833.	44.1	.01	2	10	6
S	2801.	2317.	2208.	2256.	2237.	2881.	191.1	.05	2	14	14
T	2365.	2272.	2283.	2662.	1938.	2130.	128.6	.04	2	13	12
U	2114.	2070.	2204.	2374.	1760.	2078.	96.7	.03	2	12	12
V	3255.	2723.	2673.	2698.	2115.	2286.	300.6	.08	2	11	10
W	3321.	2748.	2737.	2790.	2064.	2090.	434.4	.12	2	15	14
X	3203.	2652.	2577.	2702.	2288.	2328.	207.6	.06	2	10	8
Y	3446.	2719.	2661.	2587.	2170.	2167.	420.0	.11	2	14	13
Z	2444.	1809.	1836.	1934.	1715.	1812.	180.6	.07	2	10	13
AA	3376.	2638.	2530.	2851.	2194.	2161.	390.6	.10	2	13	9
AB	3102.	2350.	2171.	2202.	1933.	1892.	426.1	.13	2	13	11
AC	3207.	2561.	2335.	1967.	2369.	2261.	356.8	.10	2	12	10
AD	3324.	2452.	2449.	2463.	1948.	2064.	475.5	.14	2	14	13
AE	3191.	2501.	2423.	2420.	2069.	2096.	336.2	.10	2	12	11
AF	3155.	2453.	2276.	2107.	2369.	2340.	270.9	.08	2	11	9
AG	3275.	2526.	2447.	2446.	1990.	2016.	443.4	.13	2	14	13
AH	3321.	2517.	2344.	2278.	2217.	2023.	424.7	.12	2	14	13
AI	3429.	2427.	2333.	2418.	2038.	2055.	530.4	.15	2	14	13
AJ	3408.	2411.	2201.	2192.	2220.	2268.	462.8	.13	2	12	12
AK	3462.	2636.	2541.	2585.	2262.	2264.	370.1	.10	2	14	13
AL	2933.	2473.	2500.	2648.	2527.	2669.	55.5	.01	2	10	7
AM	3399.	2487.	2453.	2565.	2383.	2463.	280.4	.07	2	10	10
AN	2778.	2675.	2623.	2709.	2473.	2492.	28.1	.01	2	5	5
AO	3004.	2520.	2438.	2653.	2481.	2654.	80.8	.02	2	13	11
AP	1743.	1974.	2228.	2568.	1905.	2182.	201.7	.07	2	11	11
AQ	3473.	2489.	2420.	2413.	2436.	2519.	332.0	.09	2	14	11
AR	3193.	2683.	2589.	2616.	2294.	2375.	190.3	.05	2	12	11
AS	3065.	2696.	2612.	2574.	2175.	2628.	153.9	.04	2	13	12
AT	3274.	2689.	2617.	2699.	2139.	2332.	286.8	.08	2	11	11
AU	2829.	2596.	2613.	2759.	2352.	2601.	51.7	.01	2	15	15
AV	1395.	3106.	3916.	4168.	1599.	1566.	3034.6	.81	2	15	15

APPENDIX D

TABLE D-4
ESTIMATED STANDARD ERRORS

PROPORTION OF BIAS/S.E. OUTSIDE MODEL T

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	.20	.07	.06	.06	.05	.04
B	.21	.06	.06	.05	.06	.08
C	.23	.07	.07	.06	.07	.06
D	.13	.05	.05	.04	.05	.04
E	.11	.05	.05	.04	.06	.06
F	.38	.09	.08	.04	.11	.07
G	.25	.12	.12	.11	.14	.11
H	.39	.06	.06	.05	.09	.36
I	.43	.25	.25	.14	.31	[illegible]
J	.46	.18	.17	.07	.23	.14
K	.41	.17	.17	.08	.22	.12
L	.63	.19	.19	.08	.25	.17
M	.45	.21	.21	.12	.23	.14
N	.41	.35	.35	.27	.36	.26
O	.31	.07	.07	.06	.13	.08
P	.31	.06	.05	.04	.07	.04
Q	.37	.12	.11	.09	[illegible]	[illegible]
R	.26	.06	.06	.04	.07	[illegible]
S	.40	.27	.26	.26	.27	.43
T	.40	.35	.35	.30	.38	.62
U	.45	.41	.41	.39	.43	.65
V	.22	.08	.08	.07	.08	.06
W	.16	.05	.05	.05	.14	.05
X	.50	.22	.22	.12	.27	.17
Y	.23	.06	.06	.06	.07	.07
Z	.43	[illegible]	.06	.07	.11	.07
AA	.35	[illegible]	.09	.04	[illegible]	.06
AB	.38	.12	[illegible]	.08	.11	[illegible]
AC	.34	.24	.23	.14	.26	.14
AD	.26	.07	.07	.07	.08	.08
AE	.20	.06	.06	.06	[illegible]	.11
AF	.21	.12	.12	.10	.14	.13
AG	.24	.07	.07	.06	.06	.04
AH	.44	.15	.14	.06	.13	.09
AI	.41	.07	.06	.05	[illegible]	.12
AJ	.76	.33	.26	.14	.36	.23
AK	.21	.04	.04	.13	.06	.06
AL	.21	.03	.03	.02	.04	.02
AM	.71	.35	.35	.29	.37	.31
AN	.26	.07	.07	.06	.12	.02
AO	.22	.03	.03	.02	.04	.02
AP	.54	.51	.51	.53	.55	.62
AQ	.34	.07	.06	.05	.10	.07
AR	.25	[illegible]	.09	.09	.15	.12
AS	.16	.09	.09	.07	.12	.13
AT	.22	.03	.07	.07	.03	.07
AU	.20	.17	.16	.18	.24	.26
AV	.24	.74	.73	.57	[illegible]	[illegible]

APPENDIX D

TABLE D-6
ESTIMATED RESIDUALS

TYPICAL INTERQUARTILE RANGE

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML
A	.27	.41	.43	.45	.40	.45
B	.28	.42	.43	.45	.42	.46
C	.26	.47	.48	.53	.48	.62
D	.42	.63	.64	.66	.63	.64
E	.29	.43	.45	.47	.44	.44
F	.18	.34	.44	.51	.34	.48
G	.23	.43	.44	.51	.44	.50
H	.19	.43	.45	.59	.45	.48
I	.18	.31	.32	.47	.39	1.14
J	.14	.36	.37	.71	.44	.31
K	.19	.35	.35	.86	.42	2.34
L	.37	.28	.28	.51	.30	.75
M	.16	.30	.30	.50	.33	1.17
N	.20	.25	.25	.30	.27	.45
O	1.43	2.09	2.12	3.19	2.39	3.43
P	.18	.50	.53	.64	.47	.65
Q	.26	.41	.45	.49	.43	.57
R	.18	.24	.25	.28	.27	.44
S	.82	.99	1.02	1.22	1.08	12.22
T	1.72	1.97	2.10	13.92	2.21	4.31
U	1.92	2.25	2.29	8.17	2.56	4.39
V	.27	.42	.41	.44	.44	.51
W	.28	.43	.43	.44	.40	.43
X	.17	.32	.33	.96	.38	1.[illegible]
Y	.29	.48	.50	.56	.50	.65
Z	.29	.50	.54	.62	.55	.59
AA	.24	.42	.49	1.58	.62	.79
AB	.24	.34	.37	.37	.44	.49
AC	.31	.37	.37	.38	.4[illegible]	.62
AD	.24	.38	.38	.39	.36	.58
AE	.26	.34	.34	.35	.36	.39
AF	.77	.85	.85	.87	.88	.95
AG	.31	.46	.47	.43	.49	.48
AH	.16	.33	.38	.47	.30	.50
AI	.20	.40	.41	.46	.39	.41
AJ	.08	.27	.32	.83	.42	5.01
AK	.19	.35	.39	.41	.43	.44
AL	.13	.23	.24	.28	.27	.36
AM	.08	.19	.19	.24	.22	.27
AN	.06	.07	.07	.08	.12	.12
AO	.11	.22	.23	.27	.26	.39
AP	1.88	2.12	2.16	3.07	2.51	3.23
AQ	.20	.39	.41	.44	.42	.44
AR	.28	.36	.37	.41	.44	.52
AS	.12	.13	.15	.14	.16	.22
AT	.28	.41	.43	.50	.45	.43
AU	.16	.24	.24	.28	.27	.3[illegible]
AV	.191914.108128.326755.34				5.60	4[illegible].[illegible]3

APPENDIX D

TABLE D-7

ESTIMATED RESIDUALS

RANKINGS ON ALL ESTIMATES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	CHI SQ	W	SIG	X95	X99
A	1631.	1269.	1028.	896.	819.	657.	588.1	.39	2	6	6
B	1604.	1189.	1102.	893.	817.	695.	508.5	.34	2	6	6
C	1567.	1118.	1058.	863.	820.	874.	372.2	.25	2	5	5
D	2062.	1792.	1776.	1792.	1482.	1596.	112.6	.05	2	8	7
E	1632.	1295.	1047.	901.	783.	642.	627.3	.42	2	6	6
F	1656.	1309.	1001.	865.	831.	638.	655.9	.44	2	6	6
G	1619.	1219.	1037.	860.	811.	754.	507.9	.34	2	6	6
H	1605.	1215.	1070.	909.	835.	666.	523.1	.35	2	6	6
I	1540.	1221.	1033.	869.	799.	838.	390.8	.26	2	6	6
J	1524.	1204.	1028.	876.	815.	853.	355.4	.24	2	6	6
K	1592.	1233.	1012.	882.	739.	844.	472.6	.32	2	6	6
L	1555.	1159.	1011.	848.	796.	931.	369.4	.25	2	6	6
M	1545.	1238.	1060.	886.	784.	787.	426.0	.28	2	6	6
N	1430.	1173.	1034.	884.	874.	905.	227.9	.15	2	4	4
O	1269.	1022.	929.	1009.	1057.	1014.	63.2	.04	2	6	6
P	1607.	1213.	1065.	862.	862.	691.	511.1	.34	2	6	6
Q	1595.	1254.	1034.	891.	794.	732.	505.6	.34	2	6	6
R	1403.	1167.	1024.	931.	883.	922.	195.7	.13	2	6	5
S	1059.	942.	880.	922.	916.	1581.	339.9	.23	2	6	6
T	793.	815.	894.	1279.	1096.	1423.	323.1	.22	2	6	6
U	804.	834.	914.	1214.	1097.	1437.	290.0	.19	2	6	6
V	1635.	1293.	1054.	866.	768.	684.	617.7	.41	2	6	6
W	1630.	1260.	1063.	898.	813.	636.	601.5	.40	2	6	6
X	1496.	1190.	1025.	949.	821.	819.	319.2	.21	2	4	4
Y	1457.	1212.	1074.	957.	825.	775.	311.8	.21	2	6	6
Z	1317.	1092.	1017.	952.	959.	965.	94.4	.06	2	5	5
AA	1445.	1161.	1013.	1050.	821.	810.	266.4	.18	2	6	6
AB	1452.	1184.	1010.	883.	932.	839.	254.8	.17	2	6	6
AC	1451.	1192.	1111.	917.	758.	871.	304.5	.20	2	6	6
AD	1628.	1176.	1075.	858.	838.	725.	512.4	.34	2	6	6
AE	1639.	1214.	1094.	837.	808.	703.	568.2	.38	2	6	6
AF	651.	836.	952.	1135.	1266.	1462.	417.7	.28	2	6	6
AG	1576.	1188.	1092.	911.	831.	704.	461.3	.31	2	6	6
AH	1654.	1198.	1062.	811.	870.	705.	567.1	.38	2	6	6
AI	1672.	1195.	1039.	845.	834.	715.	579.9	.39	2	6	6
AJ	1559.	1131.	957.	862.	894.	897.	340.4	.23	2	6	6
AK	1703.	1298.	987.	853.	799.	660.	710.3	.47	2	6	6
AL	1425.	1150.	1004.	903.	908.	910.	203.9	.14	2	6	5
AM	1386.	1101.	1026.	940.	952.	895.	154.1	.10	2	5	4
AN	1484.	1261.	1080.	919.	807.	749.	381.5	.25	2	4	4
AO	1393.	1136.	1010.	925.	916.	920.	168.7	.11	2	6	6
AP	744.	841.	966.	1233.	1058.	1458.	328.0	.22	2	6	6
AQ	1666.	1240.	955.	814.	837.	788.	566.0	.38	2	6	6
AR	1648.	1247.	1018.	844.	771.	772.	566.7	.38	2	6	6
AS	1776.	1464.	1146.	869.	575.	470.	1240.5	.83	2	6	6
AT	1617.	1271.	1067.	880.	761.	704.	574.1	.38	2	6	6
AU	1676.	1375.	1121.	870.	726.	532.	865.0	.58	2	6	5
AV	650.	1244.	1572.	1655.	572.	607.	1200.8	.80	2	6	6

APPENDIX D

TABLE D-8

ESTIMATED REDUCED FORM

TYPICAL MEDIAN BIASES

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	.10	.05	.04	.03	.01	.01	.28
B	.06	.03	.03	.02	-.06	-.08	.48
C	.34	.08	.07	.03	.07	.02	.77
D	.01	0	-.00	-.00	.01	.00	.39
E	.11	.06	.05	.05	.03	.04	.40
F	.13	.11	.08	.02	.08	-.00	.15
G	.33	.04	.04	-.00	.08	.05	.29
H	.22	.08	.09	.05	.04	-.00	.16
I	.11	.11	.11	.06	.18	.05	.35
J	-.14	.04	.05	.00	.07	.05	-.20
K	.13	.05	.05	-.01	.11	-.01	.17
L	.15	.04	.04	.00	.12	.01	-.11
M	.09	.10	.10	.08	.08	-.03	.15
N	.73	.55	.54	.31	.51	-.10	.40
O	.00	.10	.11	-.05	.03	-.04	.18
P	.44	.03	.01	-.03	.07	-.01	-.05
Q	.19	.04	.03	.01	.04	-.01	.01
R	-.16	-.03	-.02	.00	-.01	.06	.48
S	1.11	.79	.77	.63	.78	-.18	.27
T	.69	.23	.22	-.55	.74	.63	.27
U	1.05	.43	.40	-.47	1.15	.99	.48
V	.26	.27	.27	.24	.10	.02	.70
W	.31	.11	.11	.08	.03	.01	.28
X	.29	.07	.06	-.00	.05	-.03	.29
Y	.31	.25	.25	.23	.28	.26	.27
Z	.25	.13	.12	.10	.13	.11	.00
AA	.33	.08	.05	-.12	.02	-.01	-1.97
AB	.14	.06	.06	.03	.05	.04	.09
AC	.24	.18	.17	.14	.15	.07	.51
AD	-.17	-.01	-.02	-.00	-.10	-.12	.25
AE	-.08	.01	.01	.02	-.08	-.10	.20
AF	.25	.18	.17	.17	.17	.15	.55
AG	.06	.05	.04	.04	.04	.03	.16
AH	.08	.17	.15	.04	.12	-.09	-.18
AI	-.09	.03	.03	.03	-.06	-.08	.21
AJ	-.01	-.05	-.10	-.11	-.02	-.11	.65
AK	-.01	.03	.03	.03	.02	.04	-.27
AL	-.11	.03	.04	-.08	.08	.07	-.33
AM	-.39	-.03	-.02	.00	.00	-.03	-.23
AN	-.01	.04	.05	.05	.02	.02	.07
AO	-.53	-.03	-.01	.05	-.01	.02	-.28
AP	.59	.34	.32	-.00	.69	1.20	.21
AQ	.06	.26	.28	.30	.19	-.01	1.40
AR	.39	.31	.30	.29	.26	.19	-.10
AS	.23	.25	.25	.24	.08	-.06	.25
AT	.25	.28	.27	.25	.10	.02	.61
AU	.21	.21	.21	.21	.21	.22	.24
AV	.12	-4.15	-4.22	-4.86	-.06	-.05	-2.21

APPENDIX D

TABLE D-9
ESTIMATED REDUCED FORM

TYPICAL INTERQUARTILE RANGE

RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF
A	.60	.59	.60	.59	.31	.30	2.01
B	.53	.49	.49	.47	.27	.25	1.55
C	.39	.46	.46	.45	.40	.45	1.20
D	.23	.23	.23	.24	.21	.22	.42
E	.62	.63	.63	.62	.38	.34	2.25
F	.53	.44	.44	.46	.39	.36	1.25
G	.77	.82	.83	.84	.72	.85	1.96
H	.45	.51	.52	.51	.50	.53	1.15
I	.80	.68	.68	.79	.69	.80	1.94
J	.69	.53	.53	.66	.95	.59	2.12
K	.35	.39	.39	.43	.34	.44	1.04
L	.29	.26	.26	.34	.32	.34	.92
M	.53	.41	.40	.53	.43	.54	1.21
N	1.75	2.07	2.02	2.19	1.93	1.99	2.81
O	.79	.86	.87	.91	.86	.79	1.64
P	.42	.45	.47	.47	.43	.44	.78
Q	.28	.31	.32	.31	.24	.25	.57
R	.92	.82	.83	.83	.82	.79	2.29
S	.76	.75	.76	.74	.53	.50	2.08
T	1.17	1.04	1.04	1.35	1.04	.53	2.08
U	.73	.71	.71	.77	.81	.62	1.55
V	.76	.76	.77	.76	.41	.42	2.86
W	.93	.93	.93	.93	.43	.38	2.01
X	.53	.58	.58	.78	.58	.49	2.08
Y	.53	.58	.59	.60	.34	.34	2.44
Z	.14	.17	.17	.18	.15	.16	.00
AA	.45	.41	.41	.33	.31	.27	1.89
AB	.24	.26	.26	.26	.19	.19	.13
AC	.32	.32	.32	.30	.29	.26	1.30
AD	.42	.43	.43	.42	.24	.22	1.40
AE	.35	.35	.35	.35	.18	.16	1.19
AF	.32	.32	.32	.32	.29	.29	1.16
AG	.48	.45	.45	.45	.27	.26	1.30
AH	.70	.64	.67	.65	.51	.35	2.45
AI	.39	.47	.47	.45	.29	.27	1.72
AJ	1.57	1.14	1.10	1.15	1.04	.83	4.64
AK	.61	.64	.64	.67	.57	.61	11.85
AL	1.38	1.66	1.63	1.87	1.81	1.51	10.21
AM	.97	.74	.74	1.03	.70	.78	4.63
AN	.47	.49	.49	.48	.48	.46	1.34
AO	1.86	1.72	1.67	2.16	1.78	1.74	10.97
AP	.61	.59	.59	.62	.68	.78	1.72
AQ	.57	.72	.74	.74	.73	.77	12.58
AR	.39	.38	.38	.40	.35	.45	2.48
AS	.39	.44	.44	.45	.25	.31	2.04
AT	.72	.74	.74	.77	.42	.42	2.68
AU	.21	.23	.23	.24	.16	.18	1.24
AV	.34	5.79	5.79	4.70	.69	.61	1.35

APPENDIX D

TABLE D-10

ESTIMATED REDUCED FORM

RANKINGS ON ALL ESTIMATES

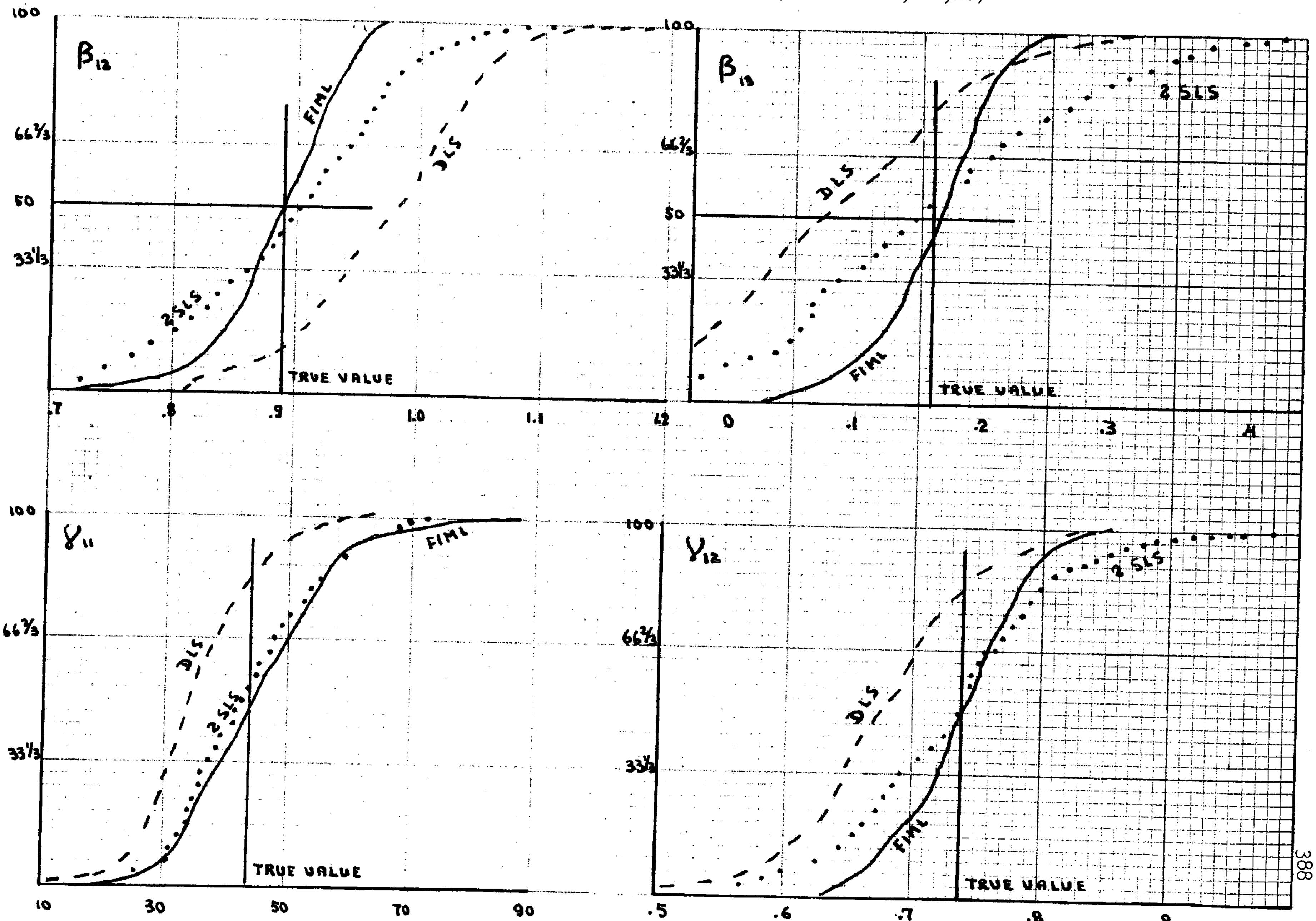
RUN	DLS	2SLS	UBK	LIML	3SLS	FIML	LSRF	CHI SQ	W	SIG	X95	X99
A	5020.	4251.	4135.	4137.	3127.	3211.	5519.	929.1	.15	2	16	15
B	4302.	3582.	3458.	3314.	2847.	3122.	4575.	557.4	.10	2	14	11
C	4737.	3454.	3291.	3157.	3027.	3183.	4351.	636.2	.12	2	14	13
D	5485.	3938.	4412.	4911.	3939.	4576.	6319.	797.4	.11	2	24	22
E	4895.	4250.	4137.	4069.	3245.	3276.	5528.	823.7	.13	2	15	15
F	5164.	4254.	3891.	3848.	3782.	3389.	5152.	617.7	.10	2	13	16
G	5449.	3916.	3879.	3804.	3513.	3616.	5223.	767.4	.12	2	17	16
H	5011.	4092.	4073.	3950.	3545.	3653.	5076.	457.7	.07	2	16	14
I	4730.	3897.	3989.	4024.	3860.	3631.	5269.	414.4	.07	2	15	13
J	4336.	3855.	4143.	4232.	3963.	3759.	5115.	251.2	.04	2	14	13
K	4927.	3945.	4080.	3913.	3631.	3678.	5228.	479.4	.08	2	15	13
L	5092.	3759.	3972.	3993.	3818.	3738.	5028.	434.7	.07	2	21	17
M	4791.	3783.	3871.	4140.	3751.	3878.	5181.	387.5	.06	2	18	14
N	4531.	4051.	4178.	4287.	3525.	3919.	4929.	251.7	.04	2	14	11
O	4354.	3950.	4136.	4312.	4109.	3864.	4675.	91.3	.01	2	4	4
P	6016.	3923.	3781.	3606.	3738.	3434.	4860.	1241.2	.17	2	21	21
Q	6043.	4239.	3927.	3713.	3294.	3025.	5162.	1395.9	.22	2	21	21
R	4586.	3952.	3884.	4172.	3585.	3991.	5238.	366.1	.06	2	12	11
S	5291.	4441.	4129.	3656.	4026.	3941.	3916.	352.5	.06	2	21	21
T	4913.	3840.	3567.	4303.	4618.	4664.	3474.	376.1	.06	2	17	17
U	3641.	3133.	3070.	3612.	4257.	4674.	2882.	584.6	.11	2	16	16
V	4491.	4022.	4091.	4128.	3502.	3822.	5366.	434.1	.07	2	16	15
W	5099.	4276.	4196.	4168.	3113.	3288.	5455.	977.3	.16	2	16	16
X	4820.	4042.	3980.	4361.	3564.	3433.	5232.	526.3	.08	2	16	14
Y	5258.	4312.	3946.	3611.	3745.	3452.	5476.	628.7	.10	2	21	21
Z	4168.	3414.	3379.	3273.	3215.	3225.	4596.	435.6	.09	2	13	18
AA	4803.	3665.	3636.	4265.	3233.	3139.	6659.	1895.3	.32	2	21	21
AB	4645.	3926.	3795.	3085.	3456.	3223.	6474.	1473.1	.23	2	21	21
AC	4470.	3664.	3283.	2900.	3418.	3551.	3914.	353.7	.07	2	16	16
AD	4165.	3547.	3536.	3418.	2876.	3072.	4616.	523.6	.10	2	15	12
AE	4000.	3615.	3555.	3352.	3012.	3117.	4599.	433.4	.08	2	14	12
AF	4526.	3832.	3369.	3164.	3242.	2924.	4143.	484.5	.09	2	12	11
AG	4454.	3604.	3424.	3373.	2878.	2924.	4543.	637.9	.12	2	16	13
AH	3775.	3812.	3519.	3177.	3437.	2855.	4623.	452.0	.08	2	15	13
AI	4136.	3553.	3406.	3392.	3061.	3105.	4547.	429.2	.08	2	12	10
AJ	4168.	3626.	3499.	3286.	3393.	3114.	4114.	232.2	.04	2	13	11
AK	4171.	3862.	3984.	4064.	3552.	3492.	6255.	1083.9	.17	2	17	17
AL	4465.	3754.	3638.	4265.	3723.	3627.	5358.	779.3	.12	2	16	15
AM	5950.	3792.	3555.	3096.	3263.	3367.	5479.	1407.3	.22	2	17	17
AN	4447.	4046.	3990.	4055.	3539.	3506.	5817.	751.7	.12	2	21	18
AO	5492.	3824.	3433.	3955.	3449.	3717.	5532.	1025.5	.16	2	20	17
AP	3295.	3172.	3202.	3605.	3854.	5146.	2925.	796.3	.15	2	18	18
AQ	3799.	3609.	3851.	3991.	3884.	4004.	6262.	1033.0	.16	2	19	17
AR	5369.	4172.	3736.	3630.	3934.	3923.	4631.	456.7	.07	2	16	16
AS	3977.	3866.	4027.	4074.	3749.	4229.	5487.	421.9	.07	2	17	17
AT	4573.	3985.	4033.	4024.	3544.	3266.	5275.	384.7	.06	2	15	17
AU	4051.	4083.	4153.	4284.	3942.	3941.	4948.	159.9	.02	2	21	21
AV	2198.	5372.	5599.	5803.	3211.	2423.	4724.	2938.2	.47	2	21	21

APPENDIX E

Cumulative Distributions of Some Estimates

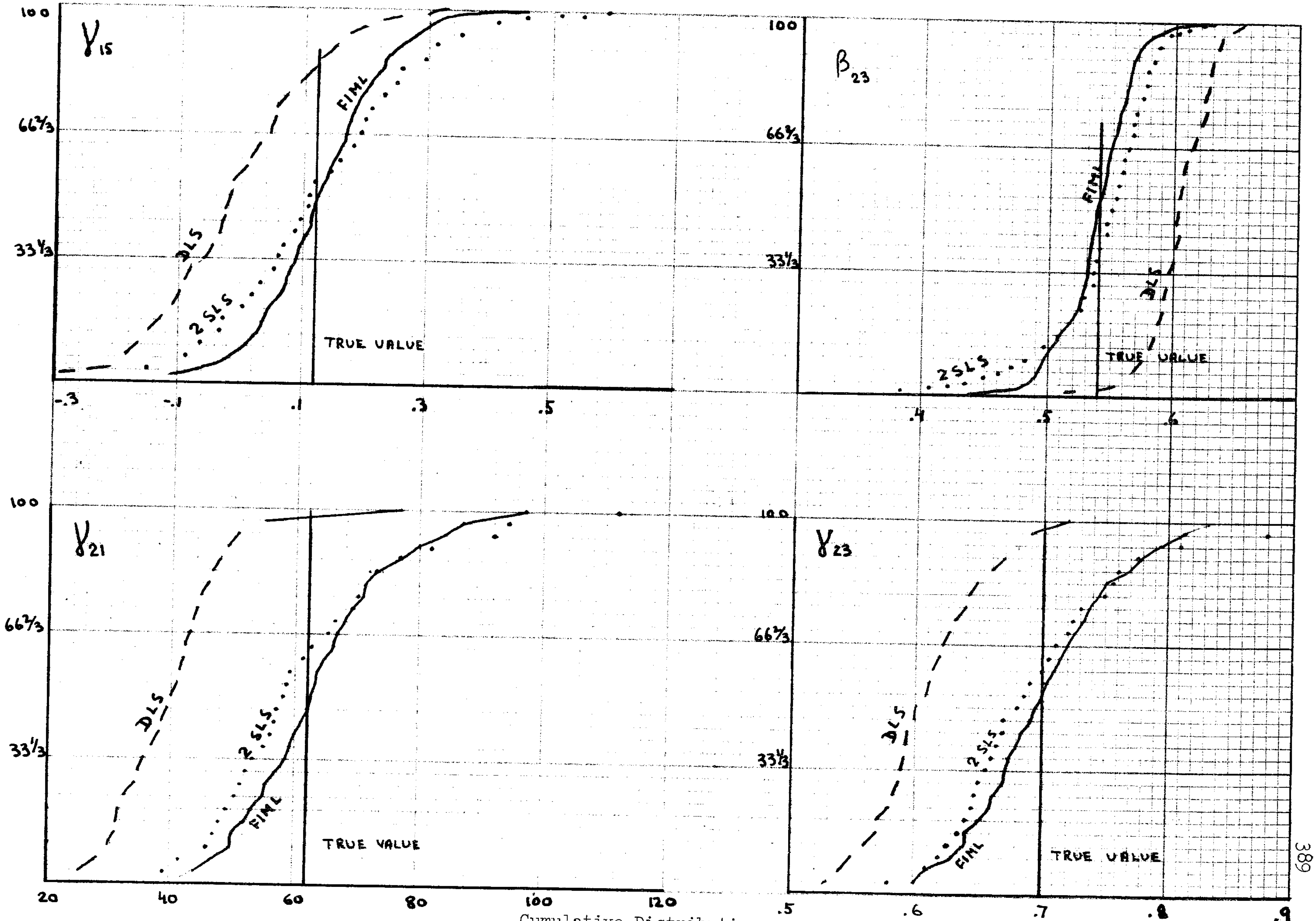
This appendix contains the cumulative distributions of the estimates made in the first three experiments. (These experiments used the basic structure and set of exogenous variables with different sets of disturbances giving 150 estimates of each parameter made by each method. The results of these experiments were discussed in Chapter IV and Chapter V, Section 2.) Usually only the cumulative distributions of DLS, 2SLS, FIML and, where appropriate, LSRF, were plotted. Other methods were omitted because it was found virtually impossible to distinguish the distributions of UBK and LIML from 2SLS in drawing the graphs and to separate 3SLS from FIML.

Appendix E, Figure 1 (where n = 1,...,16)



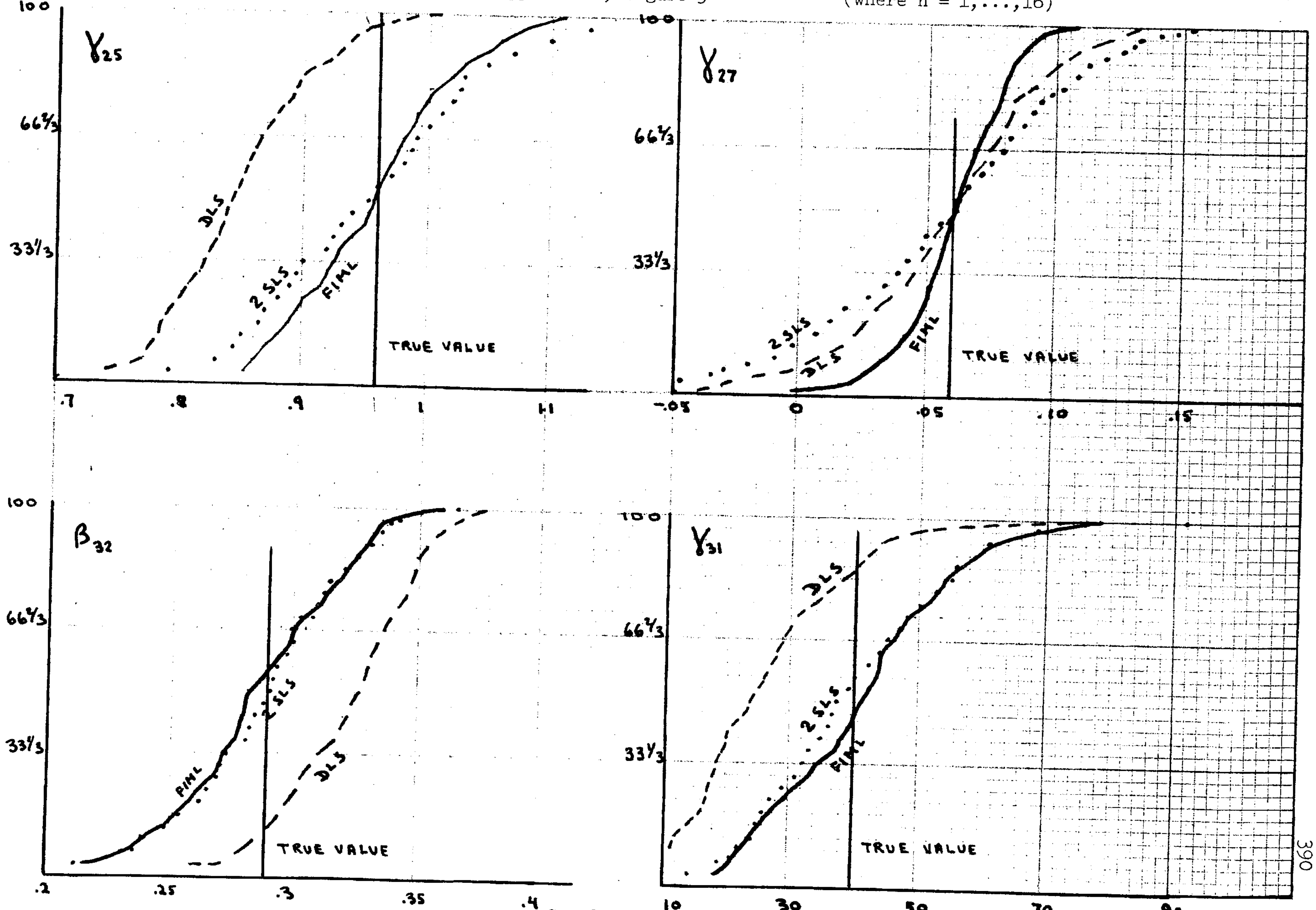
Cumulative Distributions

Appendix E, Figure 2 (where n = 1,...,16)



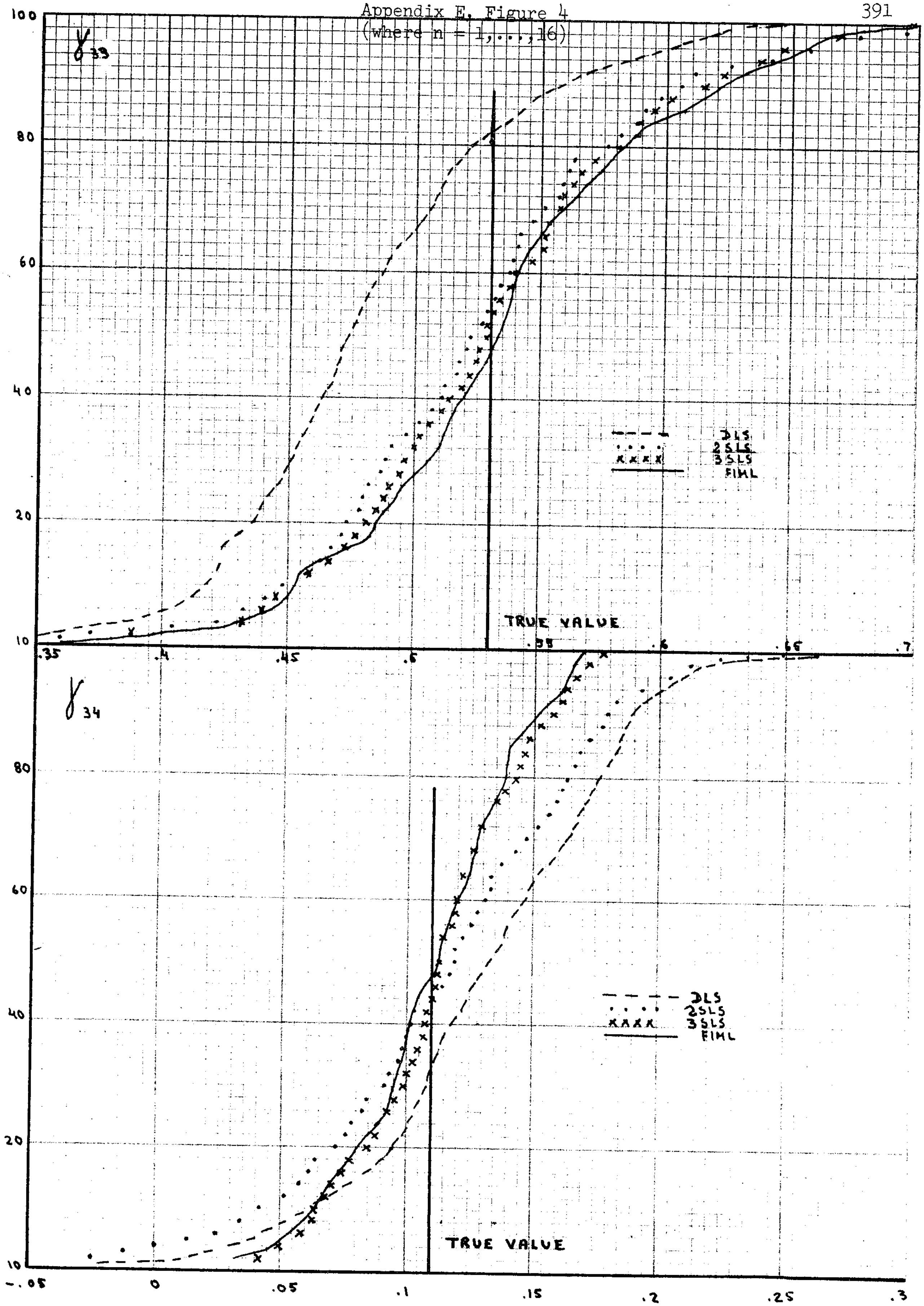
Cumulative Distributions

Appendix E, Figure 3 (where n = 1,...,16)



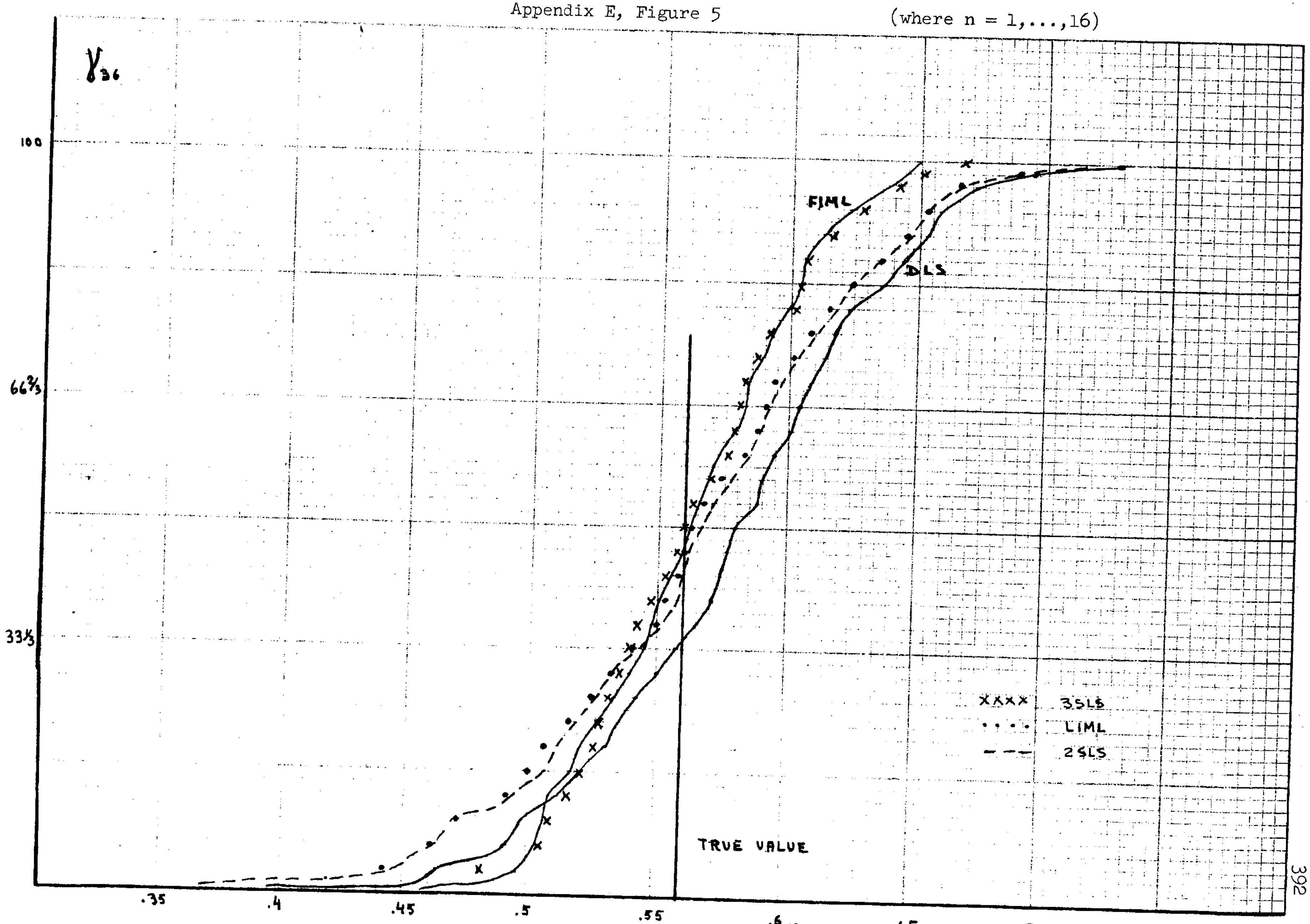
Cumulative Distributions

Appendix E, Figure 4
(where n = 1,...,16)



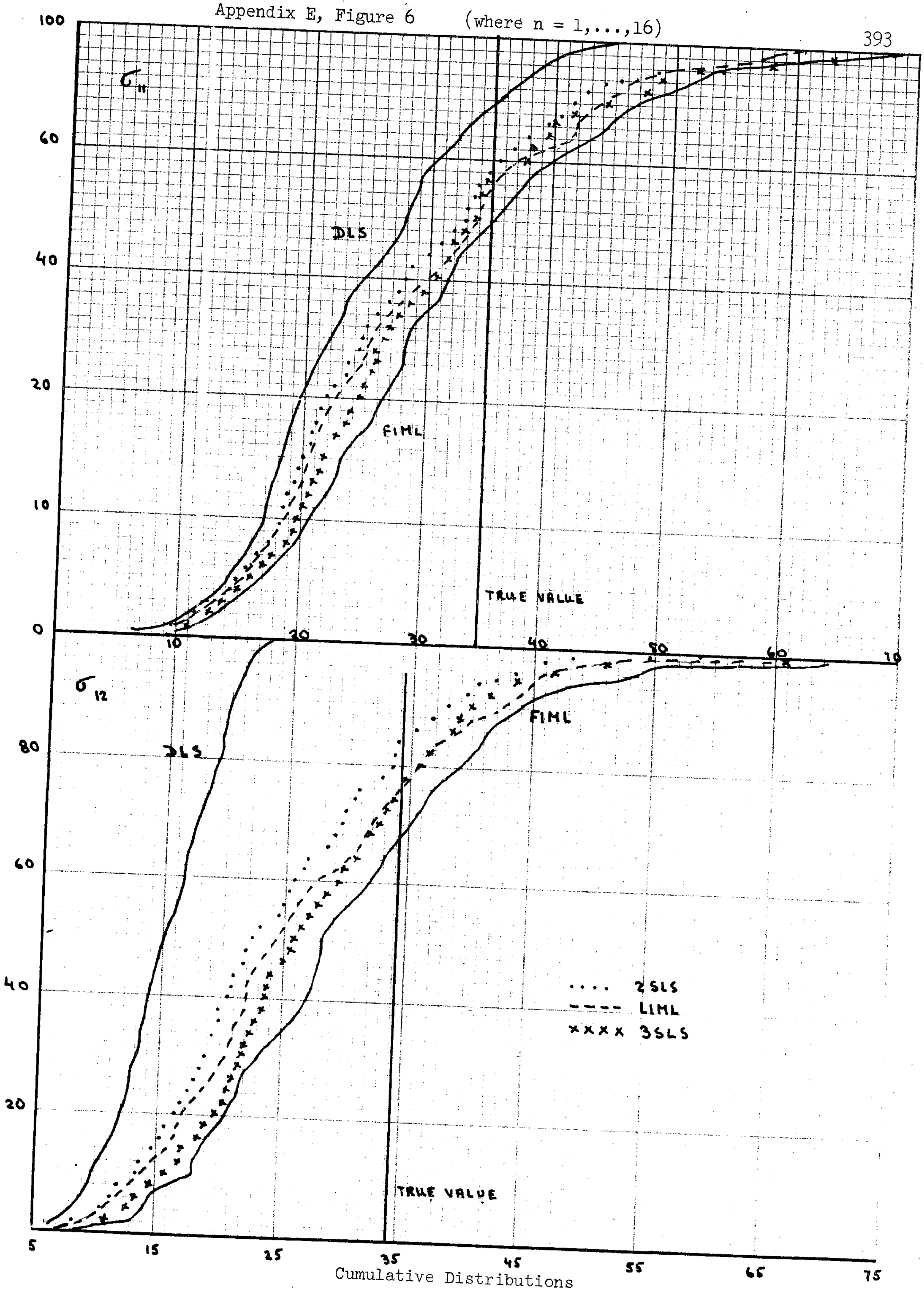
Cumulative Distributions

Appendix E, Figure 5 (where n = 1,...,16)



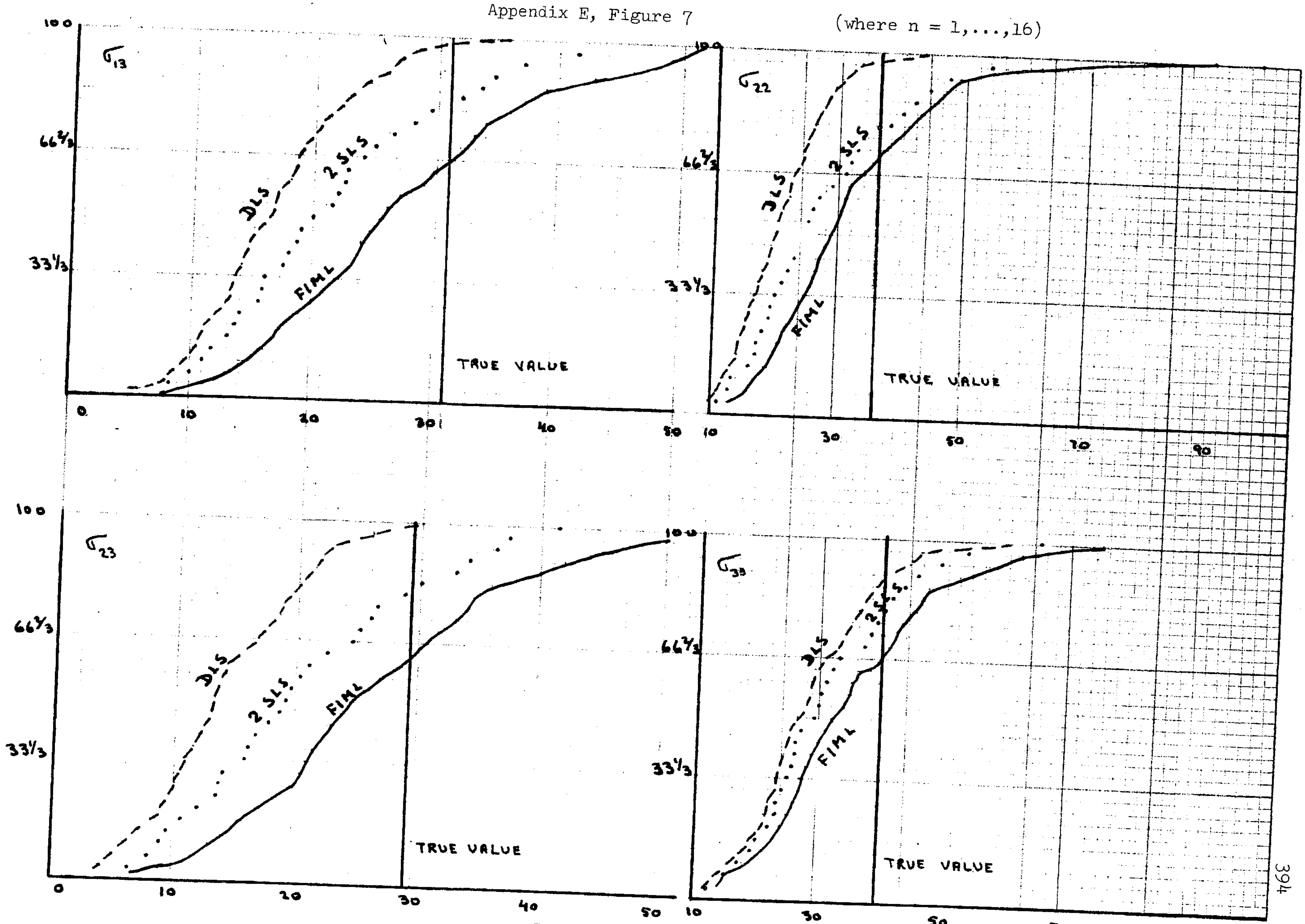
Cumulative Distributions

Appendix E, Figure 6 (where n = 1,...,16)



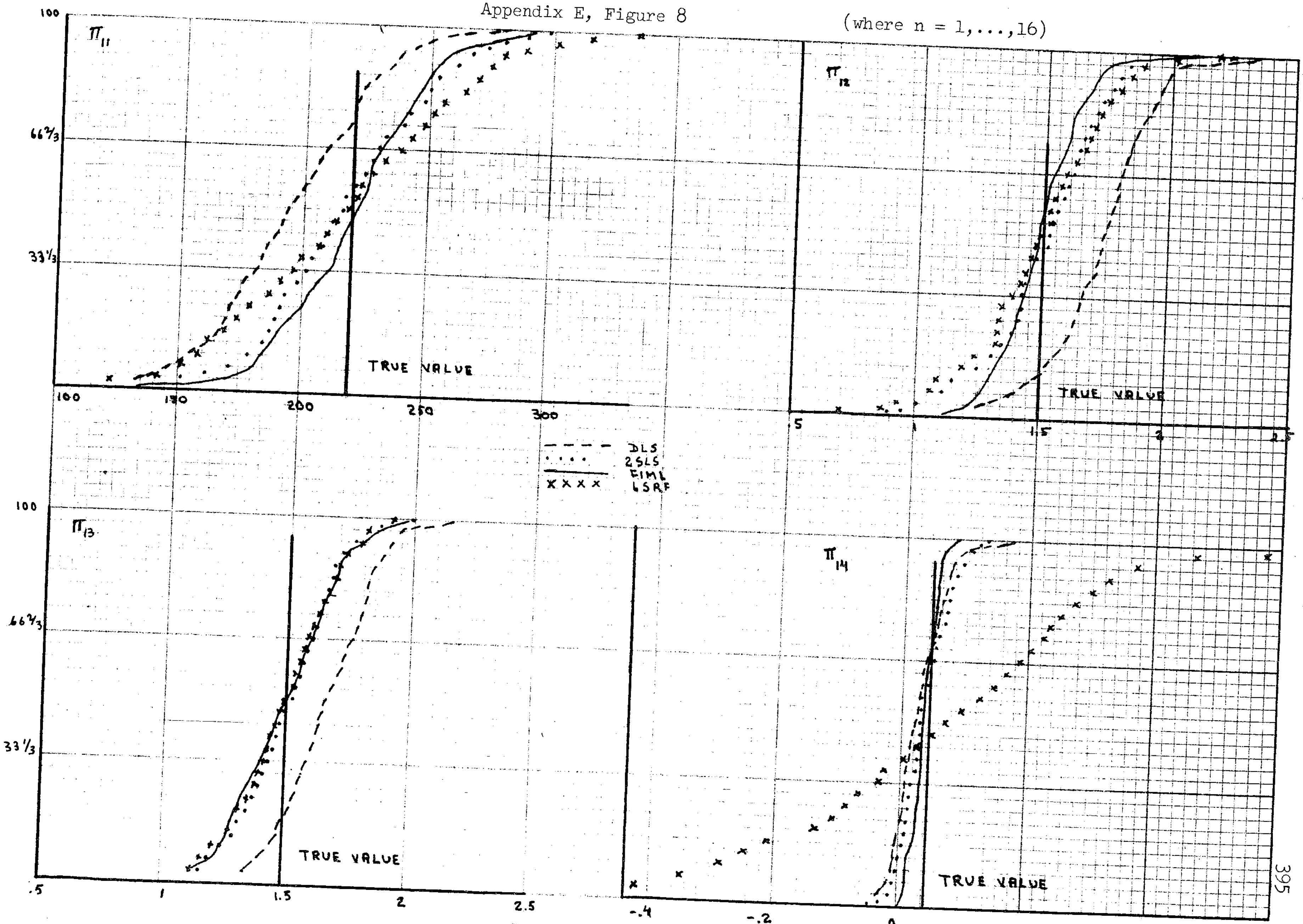
Cumulative Distributions

Appendix E, Figure 7 (where n = 1,...,16)



Cumulative Distributions

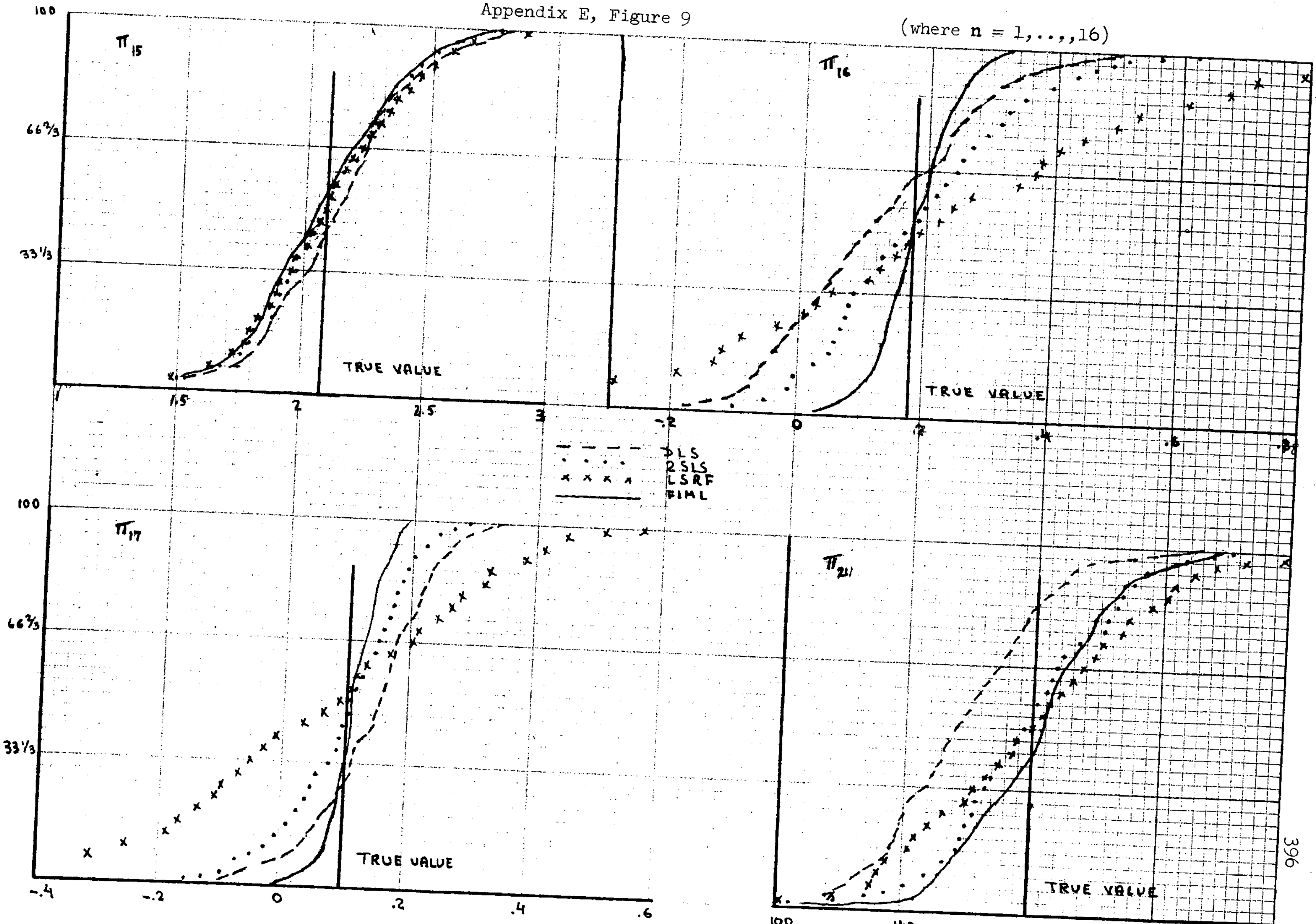
Appendix E, Figure 8 (where n = 1,...,16)



Cumulative Distributions

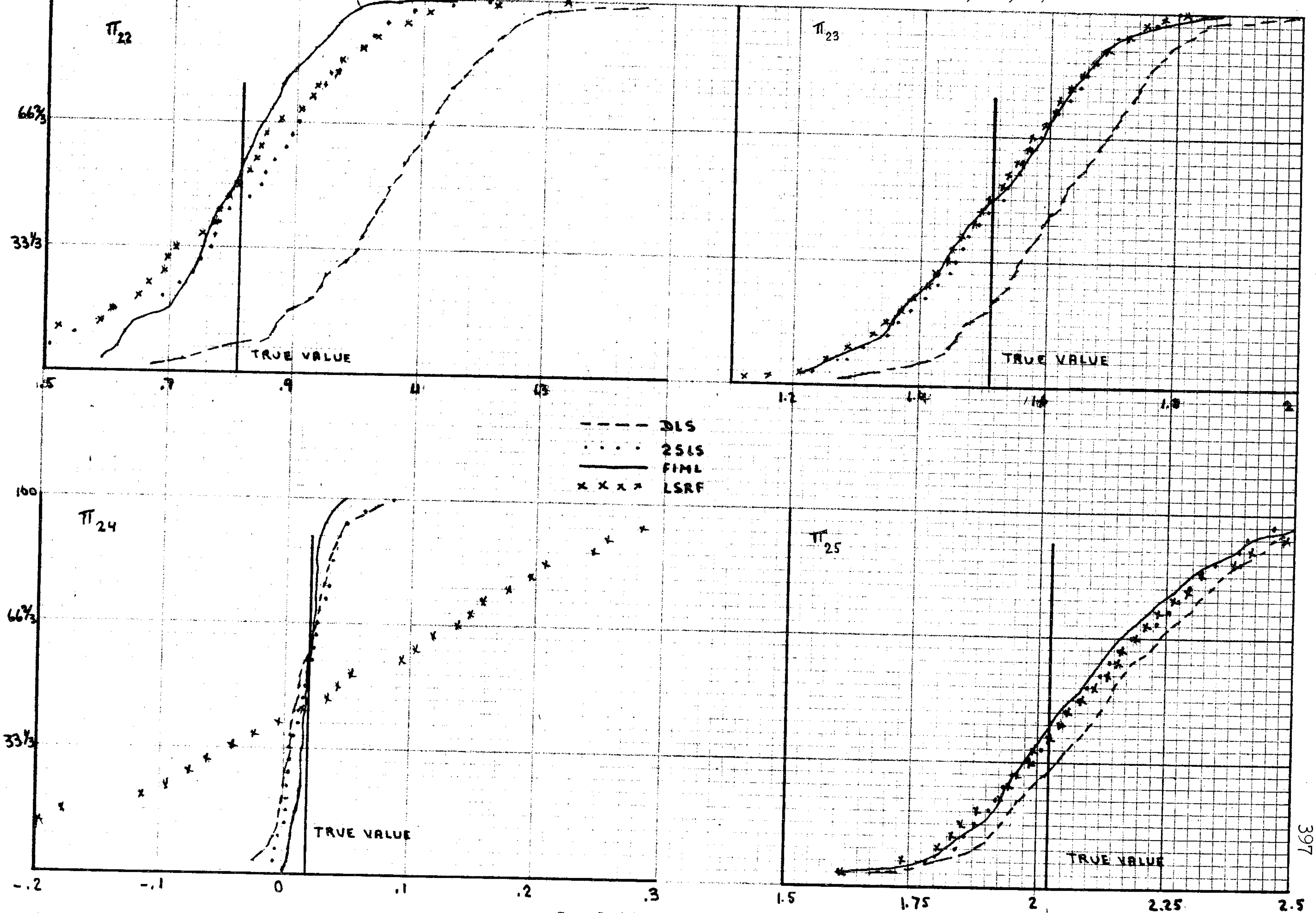
Appendix E, Figure 9

(where n = 1,...,16)



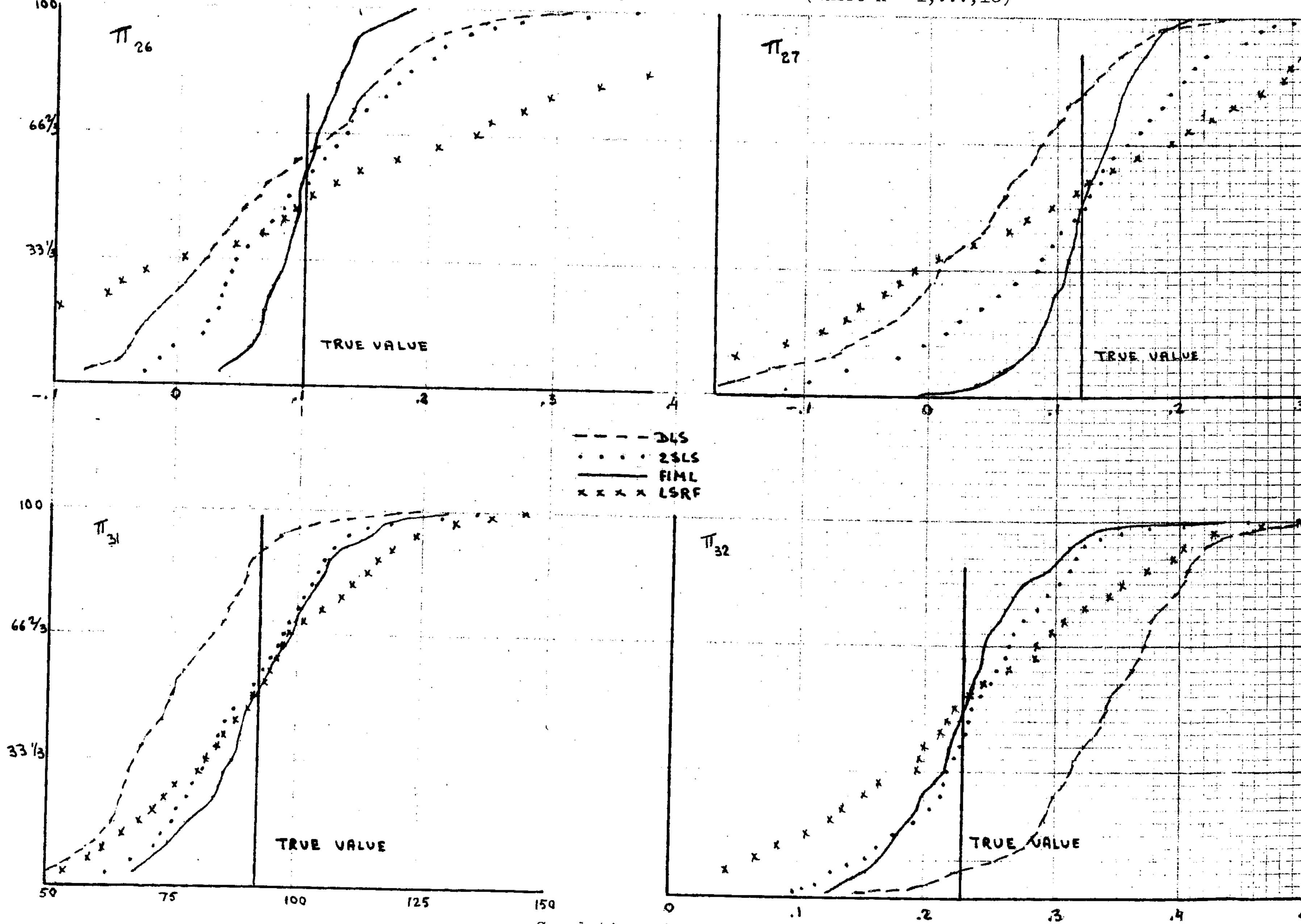
Cumulative Distributions

Appendix E, Figure 10 (where n = 1,...,16)



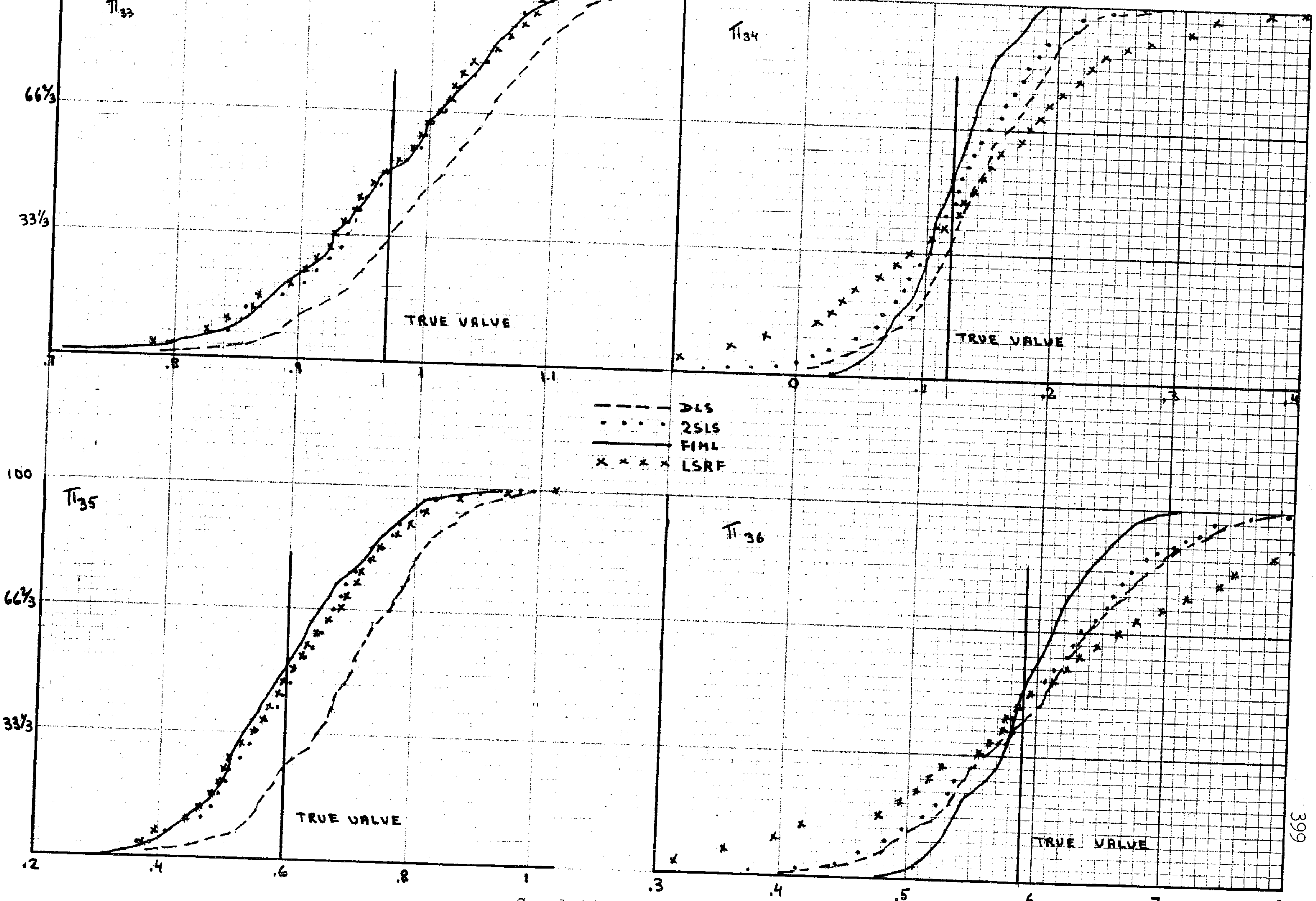
Cumulative Distributions

Appendix E, Figure 11 (where n = 1,...,16)



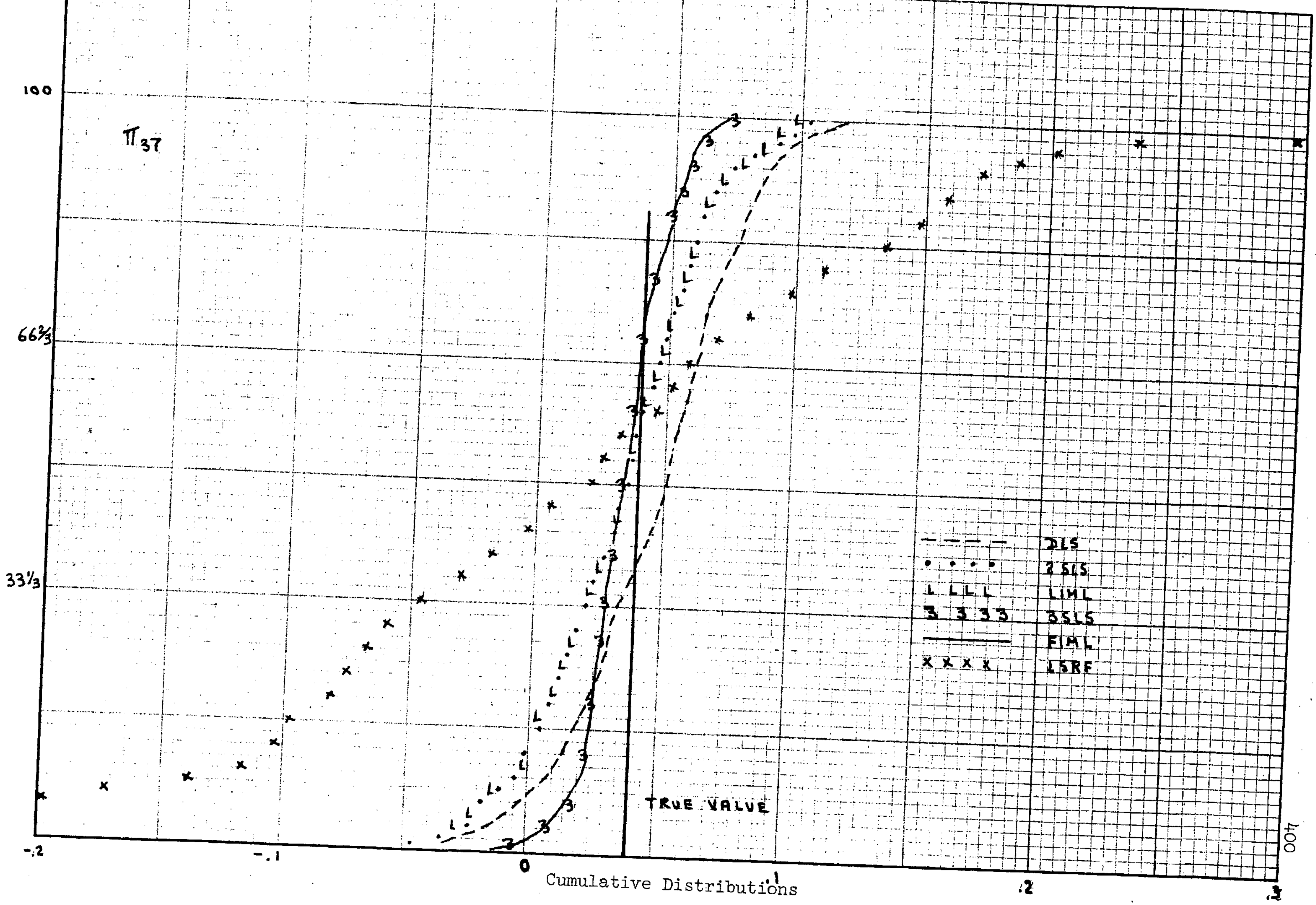
Cumulative Distributions

Appendix E, Figure 12 (where n = 1,...,16)

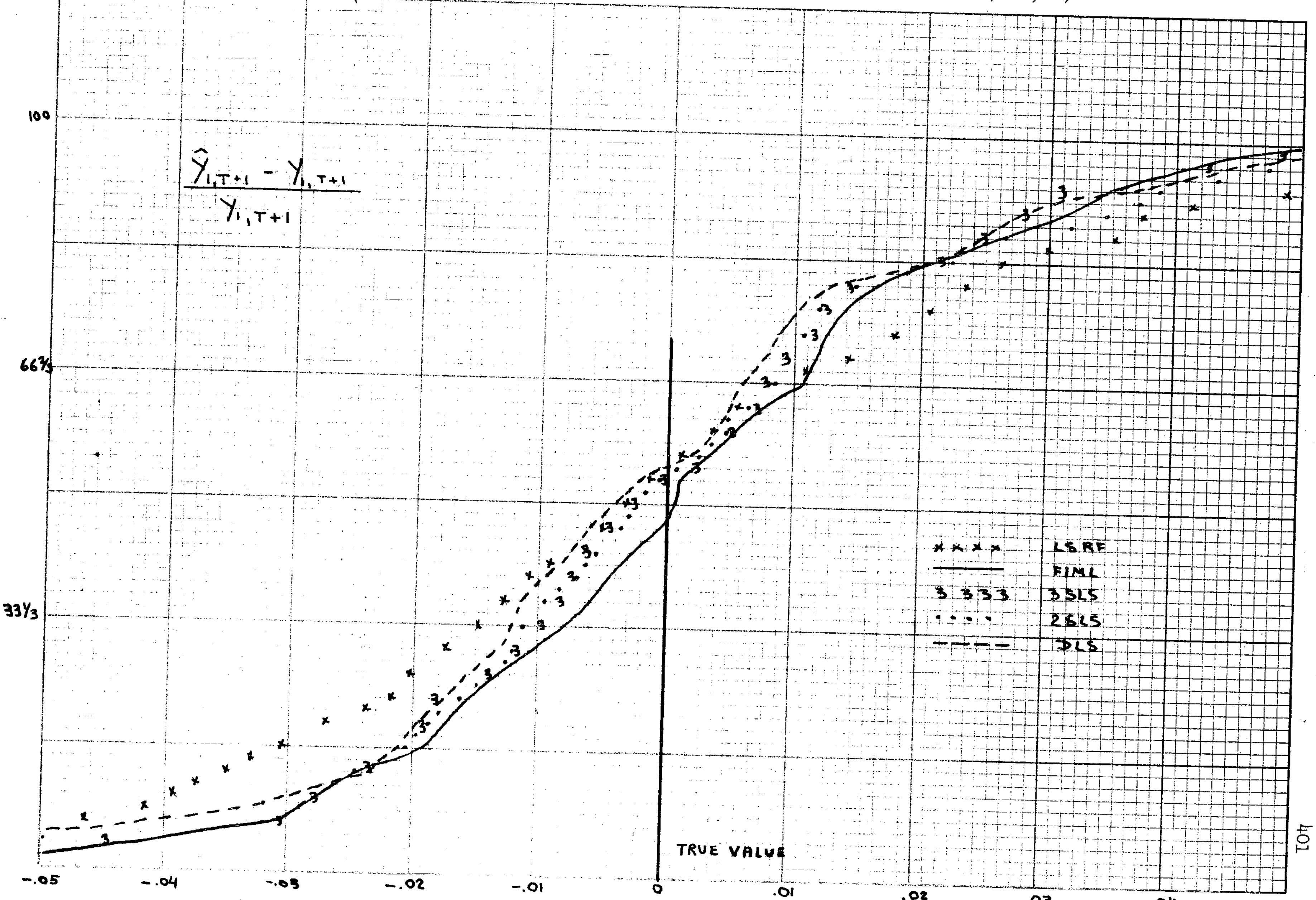


Cumulative Distributions

Appendix E, Figure 13 (where n = 1,...,16)

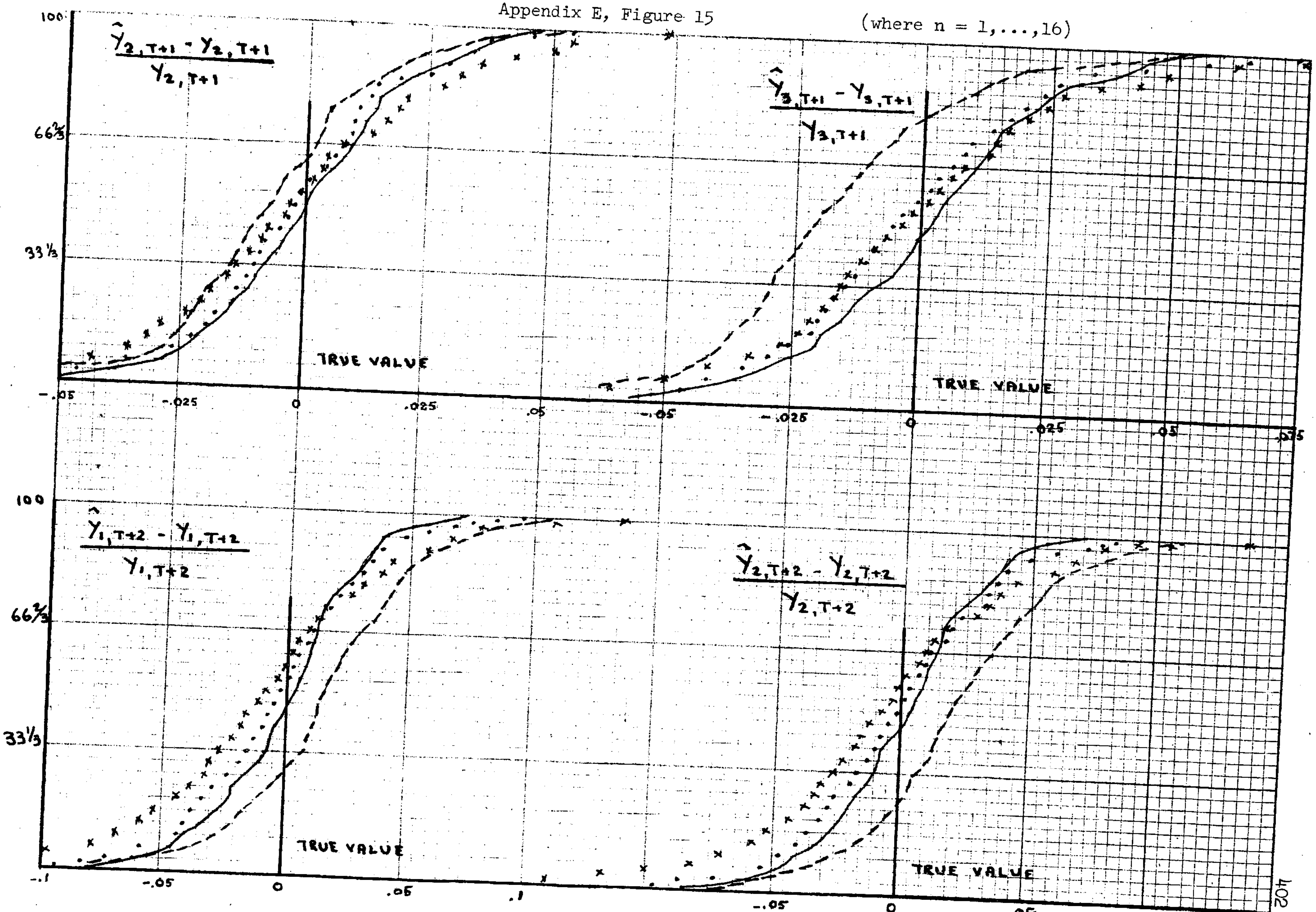


Appendix E, Figure 14 (where n = 1,...,16)



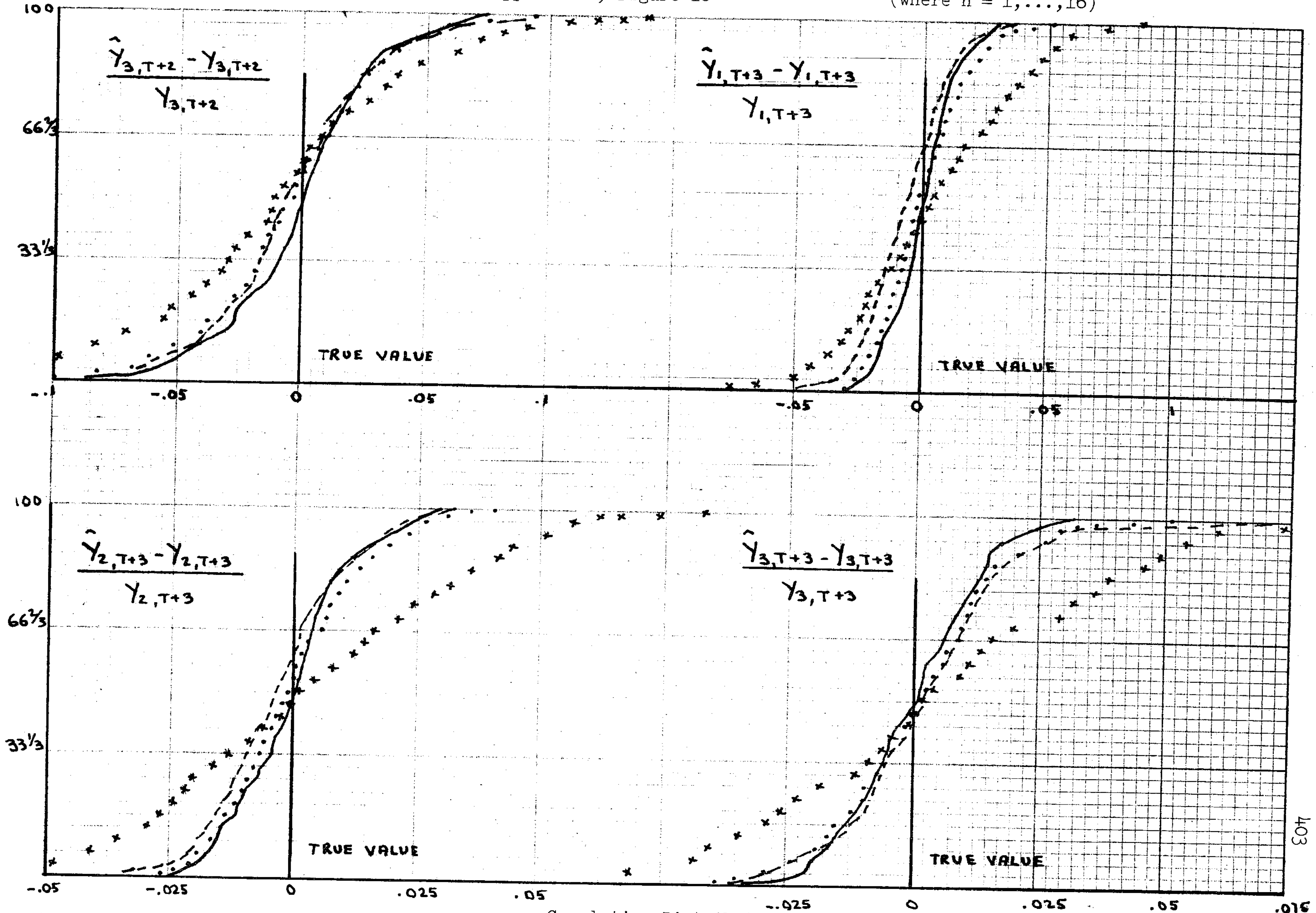
Cumulative Distributions

Appendix E, Figure 15 (where n = 1,...,16)



Cumulative Distributions

Appendix E, Figure 16 (where n = 1,...,16)



Cumulative Distributions

BIBLIOGRAPHY

Anderson, T. W., and Rubin, H., "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," Annals of Mathematical Statistics, vol. 20, 1949, pp. 46-63.

Anderson, T. W. and Rubin, H., "The Asymptotic Properties of Estimates of a Single Equation in a Complete System of Stochastic Equations," Annals of Mathematical Statistics, vol. 21, 1950, pp. 570-582.

Arrow, K. J., Social Choice and Individual Values, New York: Wiley, 1951.

Basmann, R. L., "A Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation," Econometrica, vol. 25, 1957, pp. 77-83.

Basmann, R. L., "An Expository Note on Estimation of Simultaneous Structural Equations," Biometrics, vol. 16, 1960, pp. 464-480.

Basmann, R. L., "A Note on the Exact Finite Sample Frequency of the Generalized Classical Linear Estimators in Two Leading Over-Identified Cases," Journal of the American Statistical Association, vol. 56, 1961, pp. 619-636.

Basmann, R. L., "On the Asymptotic Distribution of Generalized Linear Estimators," Econometrica, vol. 28, 1960, pp. 97-107.

Basmann, R. L., On the Exact Finite Sample Distribution of Generalized Classical Linear Structural Estimators, Santa Barbara: General Electric Company, Technical Military Planning Operation, Report SP-91, 1960.

Basmann, R. L., "The Computation of Generalized Classical Estimates of Coefficients in a Structural Equation," Econometrica, vol. 27, 1959, pp. 72-81.

Brown, T. M., "Simplified Maximum Likelihood and Comparative Structural Estimates," Econometrica, vol. 27, 1959, pp. 638-653.

Chow, Gregory C., A Comparison of Alternative Estimators for Simultaneous Equations, Yorktown Heights, N.Y.: International Business Machines Corporation, Thomas J. Watson Research Center, Report RC-781, 1962.

Christ, Carl F., "Simultaneous Equation Estimation: Any Verdict Yet?" Econometrica, vol. 28, 1960, pp. 835-845.

Durbin, J., "Incomplete Blocks in Ranking Experiments," The British Journal of Sociology (Statistical Section), 1951, pp. 85-90.

Eisenpress, Harry, "Note on the Computation of Full-Information Maximum Likelihood Estimates of Coefficients of a Simultaneous System," Econometrica, vol. 30, 1962, pp. 343-348.

Fisher, Franklin M., "On the Cost of Approximate Specification in Simultaneous Equation Estimation," Econometrica, vol. 29, 1961, pp. 139-170.

Friedman, M., "The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance," Journal of the American Statistical Association, vol. 37, 1937, pp. 675-701.

Goldberger, Arthur S., Econometric Theory, New York: Wiley, 1964.

Graybill, Franklin A., An Introduction to Linear Statistical Models, New York: McGraw-Hill, 1961.

Hildreth, Clifford, "Simultaneous Equation Estimation: Any Verdict Yet?" Econometrica, vol. 28, 1960, pp. 846-854.

Holtz, Fritz C., Economic Shock Models, Oslo: Norwegian Universities Press, 1962.

Hooper, John W., "Simultaneous Equations and Cannonical Correlation Theory," Econometrica, vol. 27, 1959, pp. 245-256.

Hood, William C., and Koopmans, Tjalling C. (eds.), Studies in Econometric Method, New York: Wiley, 1953.

Johnston, J., Econometric Methods, New York: McGraw-Hill, 1963.

Kendall, Maurice G., The Advanced Theory of Statistics, London: Griffin, 1950.

Kendall, Maurice G., Rank Correlation Methods, London: Griffin, 1948.

Klein, L. R., and Goldberger, A. S., An Econometric Model of the United States 1929-1952, Amsterdam: North-Holland, 1955.

Klein, L. R., A Textbook of Econometrics, Evanston, Illinois: Row, Peterson, 1953.

Klein, L. R., Essays in Economics and Econometrics, Chapel Hill: University of North Carolina Press, 1960.

Klein, L. R., "Single-Equation vs. Equation System Methods of Estimation in Econometrics," Econometrica, vol. 28, 1960, pp. 866-871.

Konijn, H. S., "Identification and Estimation in a Simultaneous Equations Model with Errors in the Variables," Econometrica, vol. 30, 1962, pp. 79-87.

Koopmans, T. C. (ed.), Statistical Inference in Dynamic Economic Models, New York: Wiley, 1950.

Kuenne, R. E., The Theory of General Economic Equilibrium, Princeton: Princeton University Press, 1963.

Kuh, Edwin, Capital Stock Growth: A Micro-Econometric Approach, Amsterdam: North-Holland, 1963.

Ladd, G. W., "Effects of Shocks and Errors in Estimation: An Empirical Comparison," Journal of Farm Economics, vol. 38, 1956, pp. 485-495.

Liu, Ta-Chung, "Underidentification, Structural Estimation, and Forecasting," Econometrica, vol. 28, 1960, pp. 855-865.

Madansky, Albert, "On the Efficiency of Three Stage Least Squares Estimates," Econometrica, vol. 32, 1964, pp. 51-57.

Meyer, H. A. (ed.), Symposium on Monte Carlo Methods, New York: Wiley, 1956.

Mood, A. M., Introduction to the Theory of Statistics, New York: McGraw-Hill, 1950.

Morgenstern, Oskar, On the Accuracy of Economic Observations, Princeton: Princeton University Press, 1963.

Nagar, A. L., "A Monte Carlo Study of Alternative Simultaneous Equation Estimators," Econometrica, vol. 28, 1960, pp. 573-590.

Nagar, A. L., Statistical Estimation of Simultaneous Economic Relationships, Rotterdam: privately printed, 1959.

Nagar, A. L., "The Bias and Moment Matrix of the General k-class Estimators of the Parameters in Simultaneous Equations," Econometrica, vol. 27, 1959, pp. 575-595.

Neiswanger, W. A., and Yancy, T. A., "Parameter Estimates and Autonomous Growth," Journal of the American Statistical Association, vol. 54, 1959, pp. 389-402.

Quandt, R. E., On Certain Small Sample Properties of k-class Estimators, mimeographed, 1963.

Quandt, R. E., Some Small Sample Properties of Certain Structural Equation Estimators, Princeton: Econometric Research Program, Research Memorandum 46, 1962.

Rothenberg, T. J., and Leenders, C. T., "Efficient Estimation of Simultaneous Equation Systems," Econometrica, vol. 32, 1964, pp. 57-76.

Rasmussen, P. Norregaard, "A Short Note on the Transmission of Shocks in Simultaneous Models," Econometrica, vol. 28, 1960, pp. 364-367.

Samuelson, P. A., Foundations of Economic Analysis, Cambridge, Mass.: Harvard, 1948.

Sargan, J. D., "The Estimation of Economic Relationships using Instrumental Variables," Econometrica, vol. 26, 1958, pp. 393-415.

Sargan, J. D., "Three-Stage Least-Squares and Full Maximum Likelihood Estimates," Econometrica, vol. 32, 1964, pp. 77-81.

Siegel, Sidney, Nonparametric Statistics for the Behavioral Sciences, New York: McGraw-Hill, 1956.

Summers, Robert, A Capital Intensive Approach to the Small Sample Properties of Various Simultaneous Equation Estimators, typescript.

Theil, H., Economic Forecasts and Policy, Amsterdam: North-Holland, 1961.

Valavanis, Stefan, Econometrics, New York: McGraw-Hill, 1959.

Wagner, Harvey M., "A Monte Carlo Study of Estimators of Simultaneous Equations," Econometrica, vol. 26, 1958, pp. 117-133.

Wilks, S. S., Mathematical Statistics, New York: Wiley, 1962.

Zellner, A., and Theil, H., "Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations," Econometrica, vol. 30, 1962, pp. 54-78.