International Finance Discussion Papers

Number 317

December 1987

MONTE CARLO METHODOLOGY AND THE FINITE SAMPLE PROPERTIES OF STATISTICS FOR TESTING NESTED AND NON-NESTED HYPOTHESES

by

Neil R. Ericsson

ABSTRACT

Using recently developed Monte Carlo methodology, this paper investigates the effect of dynamics and simultaneity on the finite sample properties of maximum likelihood and instrumental variables statistics for testing both nested and non-nested hypotheses. Numerical-analytical approximations (response surfaces) to the unknown finite sample size and power functions of those statistics are obtained for dynamic one- and two-equation models. The results illustrate the value of asymptotic theory in interpreting finite sample properties and certain limitations for doing so. Two practical finite sample results arise: the F form of the Wald statistic is strongly favored over its chi-squared form; and the effects of "large-σ" and a small effective sample size are particularly pronounced for Sargan's (1958) instrumental variables statistic and Ericsson's (1983) Cox-type instrumental variables statistic. Re-examining Pesaran and Deaton's (1978) empirical example illustrates the additional information gained from the instrumental variables statistics.

Keywords and phrases: asymptotic distributions. dynamics. econometrics. encompassing. evaluation criteria. finite sample properties. inference. Monte Carlo. non-nested hypotheses. power. response surfaces. simultaneity. simulation. test statistics.

Monte Carlo Methodology and the Finite Sample Properties of Statistics For Testing Nested and Non-nested Hypotheses

by

Neil R. Ericsson[1]

1. Introduction

Statistical inference has profoundly influenced econometric methodology and practice, both with regard to estimation and with regard to hypothesis testing. Mann and Wald (1943b), Haavelmo (1943, 1944), and Koopmans, Rubin, and Leipnik (1950) systematically exposit the framework for applying both aspects to the modeling of systems of economic relationships. Although the former (estimation) often has taken the more important role in econometrics, extensive testing of econometric models is becoming more common. Several reasons for that include a clearer understanding of the relationships between various econometric estimators, a marked reduction in the computing costs of estimating econometric models (those costs often having been a motivation for deriving different estimators), and a more widespread appreciation of the weaknesses of untested models. (E.g., see Hausman (1975) and Hendry (1976), Hendry and Srba (1980), and Hendry and Mizon (1978) and Sargan (1980a,b,d).) More extensive testing is of some comfort

[1]The author is a staff economist in the International Finance Division. This research was supported in part by grants from the Social Science Research Council to the Programme in Methodology, Inference and Modelling in Econometrics at the London School of Economics and to the project "The Modelling and Evaluation of Dynamic Econometric Systems" (HR8789) at Nuffield College, Oxford while I was at those institutions. The paper was prepared in part during a visit to the Center for Operations Research and Econometrics, Louvain-la-Neuve. I am grateful for the financial assistance from the S.S.R.C. although the views expressed in this paper are solely the responsibility of the author and should not be interpreted as reflecting those of the S.S.R.C., the Board of Governors of the Federal Reserve System, or other members of their staffs. Helpful suggestions and comments from Denis Sargan, David Hendry, Julia Campos, Raymond Chapman, Gary Chamberlain, Robert Summers, Ralph Tryon, and two anonymous referees are gratefully acknowledged. I am indebted to Heidi Lyss for preparing the graphs and to Frank Srba and David Hendry for the use of and their help in modifying their computer programs DAGER, GIVE, NAIVE, and PC-GIVE (Hendry and Srba (1977b, 1978, 1979, 1980) and Hendry (1987)). All numerical results were obtained using those programs.

to users of econometric models, particularly in light of the ease with which seemingly highly significant but nevertheless spurious regression results can be obtained with time series (nb. Yule (1926), Granger and Newbold (1974), Hendry (1980), and Phillips (1986b)). In recognition of the importance of hypothesis testing in econometric modeling, Pesaran (1982), Ericsson (1983), and Godfrey (1983) derive and analyze asymptotic properties of various statistics for testing nested and non-nested hypotheses in systems of economic relationships. However, both analytical and Monte Carlo studies indicate that the presence of dynamics and simultaneity may substantially influence the finite sample properties of statistics for testing nested hypotheses (cf. Phillips (1977, 1980), Sargan (1980c), Mizon and Hendry (1980), and Hendry (1984)); and even for simple static models, considerable discrepancies may exist between the finite sample and asymptotic properties of Cox's statistic for testing non-nested hypotheses (cf. Pesaran (1974, 1982)).

Using recently developed Monte Carlo methodology, this paper investigates the effects of dynamics and simultaneity on the finite sample properties of the statistics discussed in Pesaran (1982) and Ericsson (1983), including the Wald, Cox and F statistics. Section 2 describes the statistics and their asymptotic properties; Section 3, the class of econometric models to be investigated. Monte Carlo studies in econometrics often have been highly imprecise in estimating the underlying finite sample properties and specific to the particular parameter values and sample sizes chosen, so making any conclusions very tentative at best. Hendry (1984) presents a methodology reducing both imprecision and specificity and which aims to obtain "numerical-analytical formulae [response surfaces] which jointly summarize the experimental findings and known analytical results in order to help interpret empirical evidence and to compute outcomes at other points within the relevant parameter space" (p. 944). That methodology affects all aspects of Monte Carlo experimentation: design, simulation, and post-simulation analysis. Section 5 considers each of those aspects in turn for a Monte Carlo study of the properties of various instrumental variables test statistics in a dynamic simultaneous two-equation model, following a review in Section 4 of the role of response surfaces in analyzing the Monte Carlo simulations themselves. Section 6 provides a brief empirical example illustrating the potential practical value of these statistics.

2. The Test Statistics and Their Asymptotic Properties

This section summarizes existing analytical results for the statistics of interest.[2] Consider the two non-nested hypotheses

$$H_0: \quad y = X_0\beta_0 + u_0 \qquad u_0 \sim D(0, \sigma_0^2 \cdot I_T) \tag{1}$$

and

$$H_1: \quad y = X_1\beta_1 + u_1 \qquad u_1 \sim D(0, \sigma_1^2 \cdot I_T) \tag{2}$$

and the comprehensive hypothesis

$$H_2: \quad y = X_2\beta_2 + u_2 \qquad u_2 \sim D(0, \sigma_2^2 \cdot I_T) \tag{3}$$

where the dependent variable y is T×1, T being the econometric sample size; X_i is a $T \times k_i$ matrix of regressors and β_i the corresponding $k_i \times 1$ vector of coefficients (i=0,1,2); X_2 includes all the non-redundant variables in $(X_0 : X_1)$ with β_2 conformable; and u_i is a T×1 vector of disturbances distributed with mean zero and variance $\sigma_i^2 \cdot I_T$ (i=0,1,2). Two approaches have been suggested for testing H_0 against H_1: "direct" comparison of the non-nested hypotheses and comparison of each non-nested hypothesis with the comprehensive model. For the former, Pesaran (1974) proposes evaluating a modified likelihood ratio statistic for H_0 and H_1 when X_0 and X_1 are

---

[2] The statistics and most of their analytical properties appear in Pesaran (1974) and Ericsson (1983). Godfrey's (1983) statistic G_x is a linear function of t_6, so only one (t_6) is considered.

See MacKinnon, White, and Davidson (1983) for regularity conditions and MacKinnon (1983) for a review of the subject.

predetermined, following Cox (1961, 1962). Alternatively, because those hypotheses are nested in H_2, the restrictions implied by going from H_2 to H_0 (or H_1) can be tested using the F or Wald statistics.[3] Those Cox, F, and Wald statistics for testing H_0 are denoted D_0, f_2, and c_2. Under H_0 and under H_1 as a local alternative, they each are asymptotically distributed with the central and non-central distributions given in Table I.[4]

When simultaneity is present, Cox's test may be inconsistent unless the entire systems of equations from which H_0 and H_1 are drawn are specified and estimated (see Pesaran and Deaton (1978)); likewise, F and Wald tests using the least-squares estimator may be inconsistent. Instrumental variables (IV) statistics provide a convenient alternative. Sargan (1958, 1980c) proposes a χ^2-statistic (c_0) and the corresponding F-statistic (f_0) for testing the specification of an equation after estimation by instrumental variables: c_0 is the criterion function for the IV estimator, and it and f_0 are asymptotically distributed as $\chi^2(m-k_0,\cdot)$ and $F(m-k_0,T-m,\cdot)$, as given in Table I, with m being the number of instrumental variables Z^*.[5] IV generalizations of f_2 and c_2 may be used: those IV statistics also have the

---

[3]Cf. Cox (1961, pp. 105-106, 120-122), Dhrymes et al. (1972, pp. 316-317), and Cox and Hinkley (1974, pp. 327-328, 331-337) on the Cox statistic and Fisher (1922), Wald (1943, pp. 469, 479), Stroud (1971), Silvey (1975, pp. 115-116), and Phillips (1986a) on the F and Wald statistics.

[4]Ericsson (1983) presents explicit formulae for inter alia the asymptotic mean and variance of D_0 under H_1 (μ_0, ω_0) and the asymptotic non-centrality of f_2 and c_2 (λ_2). A second and minor approximation is made (i.e., in addition to the asymptotic one) to obtain those formulae. For brevity's sake, those approximate asymptotic distributions are referred to as "asymptotic" as well.

Pesaran's (1982) derivation using a different local alternative uses only the usual asymptotic approximation but requires at least as many regressors (total) under H_0 as regressors in H_1 but not in H_0. In many cases, that restriction is not satisfied and so Pesaran's formulae are not computable. For Pesaran's (1974) model, the two approximations are numerically similar; cf. Ericsson (1986) and Pesaran (1982, 1987).

[5]The statistics c_0 and f_0 and various of their properties are described in Sargan (1958, pp. 401-404; 1959, pp. 93-94, 99-100; 1964, pp. 28-29; 1976b, p. 19; 1980c, pp. 1124, 1136). See also Kiviet (1987, Chapter V).

Table I.

Asymptotic Properties of Statistics for Testing Nested and Non-nested Hypotheses

<table>
<tr><th colspan="3">Statistic</th><th colspan="2">Asymptotic Distribution</th><th colspan="2" rowspan="2">Conditions for Asymptotic Equivalences[a] under H_0 and H_1</th></tr>
<tr><th>Name</th><th>Type</th><th>Sources</th><th>H_0</th><th>H_1[b]</th></tr>
<tr><td>D_0</td><td>ML</td><td>Cox (1961), Pesaran (1974)</td><td>$N(0,1)$</td><td>$N(-\mu_0,\omega_0)$</td><td rowspan="2">$X_2 \subseteq Z^*$</td><td></td></tr>
<tr><td>t_6</td><td>IV</td><td>Ericsson (1983)</td><td>$N(0,1)$</td><td>$N(\mu_6,\omega_6)$</td><td rowspan="2">under H_0 only</td></tr>
<tr><td>t_4</td><td>IV</td><td>Ericsson (1983)</td><td>$N(0,1)$</td><td>$N(\mu_4,\omega_4)$</td><td></td></tr>
<tr><td>c_0</td><td>IV</td><td>Sargan (1958)</td><td>$\chi^2(m-k_0,0)$</td><td>$\chi^2(m-k_0,\lambda_2)$</td><td rowspan="2">always[c]</td><td rowspan="4">$m = k_2$</td></tr>
<tr><td>f_0</td><td>IV</td><td>Sargan (1980c)</td><td>$F(m-k_0,T-m,0)$</td><td>$F(m-k_0,T-m,\lambda_2)$</td></tr>
<tr><td>c_2</td><td>-[d]</td><td>Wald (1943)</td><td>$\chi^2(k_2-k_0,0)$</td><td>$\chi^2(k_2-k_0,\lambda_2)$</td><td rowspan="2">always[e]</td></tr>
<tr><td>f_2</td><td>-</td><td>Fisher (1922)</td><td>$F(k_2-k_0,T-k_2,0)$</td><td>$F(k_2-k_0,T-k_2,\lambda_2)$</td></tr>
</table>

Notes: a. Two statistics are said to be "asymptotically equivalent" under a given hypothesis if, when rescaled to be $O_p(1)$ (but not $o_p(1)$) and possibly after some nonlinear transformations, they differ only by a scale factor plus terms of $o_p(1)$ (nb. Mann and Wald's (1943a, p. 218 notation).

The asymptotic properties of the statistics under H_2 when H_2 is neither H_0 nor H_1 are discussed in Ericsson (1983).

b. The arguments μ_0, ω_0, μ_4, ω_4, μ_6, ω_6, and λ_2 are each a positive rational function of the parameters of H_1 and of the population second moments of the data.

c. The statistic f_0 is $[c_0/(m-k_0)]\cdot[(T-m)/(T-k_0)]/[1-c_0/(T-k_0)]$, which is $c_0/(m-k_0)$ with finite-sample adjustments arising from the finite-sample boundedness of c_0. f_0 is exactly distributed as an F-ratio when $Z^* = X_2$ and X_2 is fixed.

d. The Wald statistic and its F-transformation are applicable to testing hypotheses using a broad class of estimators, including ML and IV; cf. Stroud (1971).

e. The statistic f_2 is $c_2/(k_2-k_0)$, which is the classical F statistic for testing the exclusion from H_2 of those variables in X_2 but not in X_0. f_2 is exactly distributed as an F-ratio when $Z^* = X_2$ and X_2 is fixed.

asymptotic distributions given in Table I, but with λ_2 a more complicated function of the parameters and second moments. Using the IV criterion function c_0 in place of the likelihood function, Ericsson (1983) obtains IV statistics (denoted t_4 and t_6) resembling Cox's statistic D_0 and which are asymptotically equivalent to it under H_0 when the set of instrumental variables includes all the regressors and the regressors are predetermined. However, t_4 and t_6 are valid in the presence of simultaneity whereas D_0 may not be. Given a suitable set of instrumental variables, the statistics t_4 and t_6 are asymptotically distributed as normal variates: standardized under H_0, and with non-zero means and non-unit variances under H_1 as a local alternative. The final column of Table I gives conditions for asymptotic equivalences between the various test statistics. Whether they are equivalent or not, the asymptotic powers of t_4, t_6, c_0, c_2, f_0, f_2, and (if applicable) D_0 can be numerically calculated from the formulae in Table I, given a particular data generation process.[6] Similar statistics exist for testing the specification of (2) and are denoted by t_5, t_7, c_1, c_3, f_1, f_3, and D_1 (i.e., with incremented subscripts). Finally, the statistics for testing non-nested hypotheses may be interpreted as "variance-encompassing" test statistics or, equivalently, statistics for testing a certain scalar nonlinear restriction on the hypothesis H_2.[7]

3. The Data Generation Process

Using Monte Carlo techniques, Hendry and Harrison (1974) investigate the properties of single-equation estimators in the context of a dynamic

---

[6] In this paper, "the power of the statistic c_0" means "the power of an appropriate test based on c_0", and likewise for the other test statistics. This is done for brevity's sake, and no ambiguity should arise therefrom.

[7] For extensive discussion on the encompassing approach, see Davidson, Hendry, Srba and Yeo (1978), Davidson and Hendry (1981), Hendry and Richard (1982), Hendry (1983), Mizon (1984), and Mizon and Richard (1983, 1986).

simultaneous two-equation model.[8] Their model provides a convenient framework for analyzing the statistics discussed above: it is

$$y_t = bY_t + cZ_t + dy_{t-1} + u_t \qquad \begin{bmatrix} u_t \\ \nu_t \end{bmatrix} \sim \mathrm{NID}(0,\Sigma) \tag{4}$$

$$Y_t = ay_t + h'w_t \qquad + \nu_t \tag{5}$$

$$w_t = \qquad \Lambda w_{t-1} \qquad + v_t \qquad v_t \sim \mathrm{IID}(0,\Omega) \tag{6}$$

where $(y_t:Y_t)'$ and w_t are 2×1 and 4×1 vectors of endogenous and exogenous variables at time t (t=1,...,T); the $(i,j)^{th}$ element of Σ is σ_{ij}; that of Ω is ω_{ij}; $h' = (h_{21}:h_{22}:h_{23}:h_{24})$, $w_t' = (Z_t:w_{2t}:w_{3t}:w_{4t})$, and $h_{21} = 0$; Λ is a diagonal matrix $\mathrm{diag}(\rho_1:\rho_2:\rho_3:\rho_4)$; the latent root of (4)-(5) ρ_0 ($\equiv d/(1-ab)$) and all the latent roots of Λ lie within the unit circle; and $E(u_t v_{t*}') = E(\nu_t v_{t*}') = 0$ for all t and t*.[9]

The structure studied herein is the dynamic simultaneous two-equation model defined by (4)-(5) with non-zero a, b, d, and σ_{12}.[10] In order to study size as well as power, one of the non-nested hypotheses is assumed to be correctly specified and, without loss of generality, it is H_0. Thus,

$$H_0\colon \quad y_t = bY_t + cZ_t + dy_{t-1} + u_{0t} \quad , \tag{7}$$

---

[8] Cf. Hendry and Srba (1977a), Hendry (1979a), Maasoumi and Phillips (1982), Hendry (1982), and Kiviet (1985). Also, see Mizon and Hendry (1980) and Hendry (1984, pp. 971-972) on the influence of dynamics on the finite sample properties of the Wald statistic in a single-equation context, and Pesaran (1974, 1982), Godfrey and Pesaran (1983), and King and McAleer (1987) on the properties of the F and Cox statistics for a single static equation. See Cox (1962, pp. 414-415, 422-423), Jackson (1968), Atkinson (1970, p. 338), and Pereira (1977, 1978) for Monte Carlo analyses of the Cox statistic in the statistics literature.

[9] Equations (4)-(6) above correspond to equations (2)-(5) in Hendry and Harrison (1974), with some slight changes in notation.

Hendry and Harrison's model allows for autoregressive errors on the structural equations whereas (4)-(5) does not. However, noting that autoregressive errors imply a common factor restriction, enough equation dynamics are sufficient to account for such errors; cf. Appendix B.

Except for the inclusion of a constant term, the data generation process for Pesaran's (1974) Monte Carlo study is a particular case of the model in (4)-(6) with $b = d = \sigma_{12} = 0$ and $\rho_i = 0$ for i=1,...,4.

[10] A dynamic single-equation model served as a pilot study and is described in Appendix A.

in keeping with the notation of (1) and where (4) is the equation of interest. Although non-nested alternatives to (7) might involve mis-specification of dynamics or simultaneity, falsely included (or excluded) exogenous variables, or any combination thereof, attention is restricted to the (false) hypothesis that

$$H_1: \quad y_t = bY_t + h_{12}w_{2t} + u_{1t} \tag{8}$$

with $\gamma \equiv \text{corr}(Z_t, w_{2t}) > 0$. For γ close to unity, it may be difficult to detect which of the two exogenous variables, Z_t and w_{2t}, enters the correct specification. The comprehensive hypothesis (3) is

$$H_2: \quad y_t = bY_t + cZ_t + h_{12}w_{2t} + dy_{t-1} + u_{2t} \tag{9}$$

and will be used for constructing the Wald and F statistics.

The data generation process (or DGP) defined by (4)-(6) and the relationships of interest in (7)-(9) have certain implications for the properties of the statistics being examined. In (8), the exogenous variable w_{2t} is falsely included and Z_t is falsely excluded (as in the pilot study and in Pesaran (1974)), but also the lagged dependent variable y_{t-1} is falsely excluded (hence mis-specified dynamics). Pesaran's study and the one above differ also in the degrees of freedom for each statistic. In the former, the F statistic (asymptotically equivalent to the Wald statistic) has one degree of freedom in the numerator. In the latter, the Wald statistic has two degrees of freedom for (8) but only one for (7). The degrees of freedom for Sargan's statistic depend upon the number of instruments selected; but, for instance, with $(y_{t-1}:w_t')$ as instruments (i.e., two-stage least-squares), it is asymptotically distributed as a $\chi^2(2)$ for (7) and as a non-central $\chi^2(3)$ for (8). In addition to affecting the asymptotic powers of those statistics, the degrees of freedom may have a significant effect on their finite sample properties. Consistent estimation of the parameters in (7) requires some simultaneous equations estimation technique, as would be true for those in (8) if (8) were the DGP, so only the IV statistics are considered for that structure.

Various finite sample properties of the statistics might be analyzed (e.g., their means and variances; see Mizon and Hendry (1980, p. 40)), but their powers and sizes are viewed as being of primary importance, and as providing a simple way of summarizing their properties. Before turning to the experimental design, simulation, and results of this Monte Carlo study, I discuss the analysis of Monte Carlo data on powers and sizes.

4. Response Surface Methodology

Cox (1970, Chapters 3 and 6), in his discussion of the empirical logistic transform, implicitly provides the basis for developing response surfaces of estimated finite sample probabilities, including both estimated finite sample powers and estimated finite sample probabilities of type I error.[11] Consider a binary response variable for which the probability of "success" (or, later, acceptance or rejection by a particular test) is π $(0 < \pi < 1)$ and on which there are N observations $(N > 1)$, S being the number of "successes". Letting

$$A = [S(N-S)]/(N-1) , \tag{10}$$

$$L(\varsigma) = A^{1/2}\ln\left[\frac{\varsigma}{1-\varsigma}\right] \qquad 0 < \varsigma < 1, \tag{11}$$

and

$$L^*(\varsigma) = A^{1/2}\ln\left[\frac{\varsigma - (2N)^{-1}}{1 - \varsigma - (2N)^{-1}}\right] \qquad (2N)^{-1} < \varsigma < 1-(2N)^{-1} , \tag{12}$$

it can be shown that

$$\phi(s,\pi) \; = \; L^*(s) - L(\pi) \;\underset{A}{\sim}\; N(0,1) \tag{13}$$

[11]See Cochran and Cox (1957, pp. 335ff), Cox (1958, pp. 113-128), and the references in Cochran and Cox (1957, p. 369) on the use of response surfaces in statistical analyses. Their use in econometrics is relatively recent although Summers (1959) proposes using them; cf. Summers (1965) and Sowey (1973). Ericsson (1986) describes and uses response surfaces and other techniques for post-simulation analysis of Pesaran's (1974) Monte Carlo study of nested and non-nested hypothesis test statistics.

where $s \equiv S/N$ and $\underset{A}{\sim}$ denotes "converges in distribution to, as $N \to \infty$ ".[12] In the context of Monte Carlo studies of power, N is the number of replications in a particular experiment, S the number of replications for which the value of the test statistic lies in the critical region, and π the finite sample (i.e., finite econometric sample T) probability of the test statistic lying in the critical region. Below, π is treated as if it were "power" although all that is said applies equally for size.

Typically, π is some unknown function $g(\theta,T)$ (say) where θ is the vector of all parameters (except T) which define the model generating the binary random variable of interest, and the aim of finite sample research (whether using analytical or Monte Carlo techniques, or both) is to obtain a close approximation to it.[13] Even though it is unknown, $g(\theta,T)$ is implicitly defined by the computer program generating the Monte Carlo data. Further, approximations to $g(\theta,T)$ may be found and the accuracy of those approximations may be tested. As a first step to approximating $g(\theta,T)$, it is helpful to solve as much of the problem as possible analytically in order to minimize the imprecision and specificity arising from simulation. With that in mind, let

$$\left[\frac{\pi_T}{1 - \pi_T}\right] = \left[\frac{\pi_a}{1 - \pi_a}\right] \exp\{G^+(\theta,T)\} \tag{14}$$

without loss of generality, where π ($\equiv E(S/N)$) is subscripted by T so as to emphasize that it is a function of the econometric sample size; π_a is the (local) asymptotic (i.e., as $T \to \infty$) power of the test; and $G^+(\cdot,\cdot)$ is some appropriate function. By assumption, $\pi_T \to \pi_a$ as $T \to \infty$, so $G^+(\cdot,\cdot)$ is $o(1)$. Thus, (14) splits π_T into two components, an asymptotic term and a term

---

[12] That differs from the sense of "asymptotic" elsewhere in this paper, where it means "as $T \to \infty$ ". Unless otherwise noted, "asymptotic" and "finite sample" refer to T, not N. Ericsson (1986, Appendix) gives a proof of (13).

[13] If $g(\cdot,\cdot)$ were known, the exact finite sample probability (of "success", rejection) for any particular value of (θ,T) could be calculated directly, obviating any need for conducting Monte Carlo experiments to estimate π_T.

involving the deviation between the finite sample and asymptotic distributions. Because π_a can be calculated analytically for any (θ, T), the problem of directly simulating π_T (of $O(1)$) simplifies to one of simulating only $G^+(\cdot,\cdot)$ (of $o(1)$, and quite possibly $O(T^{-1/2})$).[14] In the analysis of an <u>estimator's</u> properties, an analogous partition is between its asymptotic value (its plim) and its finite sample bias (the deviation of the estimator from its plim).[15]

Using (14), (13) may be rewritten as

$$L^*(s) - L(\pi_a) = A^{1/2} \cdot G^+(\theta, T) + \epsilon \qquad \epsilon \underset{A}{\sim} N(0,1) , \tag{15}$$

providing a stochastic relationship between a feasible and unbiased estimator of π_T (i.e., s) and the known quantities π_a, θ, and T. However, the functional form of $G^+(\cdot,\cdot)$ remains unknown. From asymptotic theory, one expects that

$$G^+(\theta, T) \doteq T^{-1/2} G(\theta, T^{-1/2}) \tag{16}$$

where $G(\theta, T^{-1/2})$ is $O(T^0)$ (cf. Phillips (1977, p. 474; 1982), Sargan (1980c, p. 1120)). Thus, $G(\cdot,\cdot)$ might be expanded in powers of $T^{-1/2}$ (about $T = \infty$) and of the elements of θ. Truncating the series for $G(\theta, T^{-1/2})$, the coefficients of the powers and cross-products of θ and $T^{-1/2}$ may be estimated by least squares, correcting for heteroscedasticity using the weight $A^{1/2}$, i.e., from estimating

$$L^*(s) - L(\pi_a) = A^{1/2} T^{-1/2} H(\theta, T^{-1/2}) + e \tag{17}$$

where $H(\theta, T^{-1/2})$ is the weighted least squares approximation to $G(\theta, T^{-1/2})$ and the error term e is the combination of ϵ (the error from estimating π_T

---

[14] If an analytical approximation to π_T better than π_a is available (e.g., an Edgeworth expansion), it could appear in (14) in place of π_a, further reducing the order of the term being simulated; cf. Phillips (1982).

The functions $G^+(\cdot,\cdot)$, $G(\cdot,\cdot)$, and $H(\cdot,\cdot)$ in this section differ slightly from those identically labeled in Ericsson (1986). The change lends itself to a clearer exposition.

[15] See Campos (1986a) for a discussion on response surfaces for estimator biases and standard deviations and estimated asymptotic standard errors.

by s) and $A^{1/2}T^{-1/2}\{G(\cdot,\cdot)-H(\cdot,\cdot)\}$ (the error from approximating $G(\cdot,\cdot)$ by $H(\cdot,\cdot)$). The parameterization of θ is not unique and, before expanding $G(\cdot,\cdot)$, it may be worthwhile transforming "natural" parameters of the model into parameters which span the same range as $L^*(s)$ and which have econometrically interesting interpretations. For instance, it may be convenient to reparameterize θ to include a function of π_a such as $\ln\{\pi_a/(1-\pi_a)\}$. For the experimental design adopted in Section 5.1 below in which π_a is a design variable, that seems particularly appropriate.

A response surface like (17) summarizes a possibly vast array of Monte Carlo simulations in a relatively simple formula which may account for much of the variation in s across experiments and may be useful for predicting π_T at points within the parameter space of the experimental design (denoted $\Theta\times\Gamma$; see Section 5 below) but not included in the simulations. Further, the response surface may adequately approximate the underlying finite sample distribution function. One primary source of information exists for inferring how "good" a response surface like (17) is:

$$\epsilon \underset{A}{\sim} \text{NID}(0,1) \quad . \tag{18}$$

Using (18), many testable implications follow from the null hypothesis that $H(\cdot,\cdot)=G(\cdot,\cdot)$.

(A) $\underline{\sigma_e^2 = 1.}$ If $H(\cdot,\cdot)\neq G(\cdot,\cdot)$, then $\sigma_e^2 > 1$ because ϵ is uncorrelated with $A^{1/2}T^{-1/2}\{G(\cdot,\cdot)-H(\cdot,\cdot)\}$. The hypothesis $\sigma_e^2 = 1$ may be tested by noting that, under the null, the residual sum of squares from (17) is distributed as a χ^2 random variate with its degrees of freedom equal to the number of experiments less the number of regressors, provided N is large. Power under the alternative is directly related to the magnitude of $AT^{-1}\{G(\cdot,\cdot)-H(\cdot,\cdot)\}^2$ over the experiments.

(B) <u>The error e does not include any terms of $O(T^{-1/2})$ involving θ and $T^{-1/2}$.</u> By using OLS, e <u>can</u> <u>not</u> include any of the terms in $H(\cdot,\cdot)$. However, if $H(\cdot,\cdot)\neq G(\cdot,\cdot)$, e contains terms of a higher order than those included in $H(\cdot,\cdot)$ (cf. Maasoumi and Phillips (1982, p. 198) and Hendry (1982, p. 210)). By initially specifying a general formulation for $H(\cdot,\cdot)$ and simplifying, one can use an F statistic to test for the presence of such factors in the e's of the final specification.

(C) <u>The error e does not include any terms of $O(T^0)$ involving θ.</u> By construction from (14), $T^{-1/2}G(\cdot,\cdot)$ does not. However, if $H(\cdot,\cdot)\neq G(\cdot,\cdot)$, regressors of $O(T^0)$ in (17) such as a constant term or $\ln\{\pi_a/(1-\pi_a)\}$ may be "statistically significant", thereby revealing the mis-specification of $H(\cdot,\cdot)$. This hypothesis is particularly noteworthy, given the importance of the <u>in</u>significance of $\ln\{\pi_a/(1-\pi_a)\}$ in (17) vis-à-vis the analytical properties of the response surface.

(D) <u>The error e is normally distributed.</u>

(E) <u>The e's are serially independent for **any** ordering of experiments specified prior to simulation.</u> That follows from the independence of ϵ across experiments. If $H(\cdot,\cdot)\neq G(\cdot,\cdot)$ and experiments are ordered to be (e.g.) increasing in values of θ and T, terms in e involving θ and $T^{-1/2}$ may induce serial correlation and/or heteroscedasticity in the e's.

(F) <u>$H(\cdot,\cdot)$ is constant over regions of the parameter space which were not included in the estimation of (17).</u>

Table II lists most of the test statistics reported below; the convention used is that $\xi_i(q)$ and $\eta_i(q,p)$ denote statistics which have central $\chi^2(q)$ and $F(q,p)$ distributions respectively under a common null and against the i^{th} alternative. Thus, $\xi_9(q)$ and $\eta_9(q,K-m-q)$ both test for q^{th}-order residual autocorrelation. There are K experiments and n regressors in the response surface under the null hypothesis.

The extent to which (A)-(F) are not satisfied reflects the degree of approximation of the response surface to the underlying conditional probability formula (response function) although the <u>power</u> of tests of (A)-(F) depends crucially on the number of replications per experiment, on

Table II.
Criteria for Evaluating Response Surfaces

Null	Alternative	Statistic[a]	Sources
(A)	$\sigma_e^2 > 1$	$\xi_2(K-n)$	Theil (1971, pp. 137-138)
(B)	q invalid parameter restrictions	$\eta_3(q,K-n-q)$	Johnston (1963, p. 126)
(B)	q^{th}-order RESET	$\eta_4(q,K-n-q)$	Ramsey (1969)
(C)	$\psi \neq 0$[b]	$\eta_5(1,K-n-1)$	Hendry (1984, p. 962)
(D)	skewness (SK) and excess kurtosis (EK)	$\xi_6(2)$	Jarque and Bera (1980)
(D)	heteroscedasticity quadratic in regressors (q quadratic terms)	$\eta_7(q,K-n-q-1)$	White (1980a, p. 825), Nicholls and Pagan (1983)
(E)	q^{th}-order ARCH	$\xi_8(q)$, $\eta_8(q,K-n-q-1)$	Engle (1982)
(E)	first-order residual autocorrelation	dw	Durbin and Watson (1950, 1951), Farebrother (1980)
(E)	q^{th}-order residual autocorrelation	$\xi_9(q)$; $\eta_9(q,K-n-q)$	Box and Pierce (1970); Godfrey (1978), Harvey (1981, p. 173)
(F)	$H^+(\cdot,\cdot)$ not constant over j subsamples	$\eta_{10}((j-1)n,K-jn)$	Fisher (1922), Chow (1960, pp. 595ff)
(F)	predictive failure over a subset of q observations[c,d]	$\xi_1(q)$; $\eta_1(q,K-n-q)$	Hendry (1979b, p. 222); Chow (1960, pp. 594-595)

Notes: a. The value of q may differ across statistics, as may the number of regressors n and the number of experiments K across response surfaces and Monte Carlo studies.

b. ψ is the coefficient on $L^*(\pi_a)$ if the latter is included on the right-hand side of the response surface (17).

c. The Chow statistic is labeled $\eta_1(q,K-n-q)$. The covariance test statistic $\eta_{10}((j-1)n,K-jn)$ is often (and confusingly) referred to as the "Chow statistic" although Chow (1960, p. 592) was well aware of its presence in the literature.

d. Constancy may be tested using Chow's statistic, the covariance statistic, or the usual χ^2 statistic based upon the forecast errors. Often, an even more stringent test may be constructed by substituting unity for the estimated value of σ_e^2 in the relevant statistic, thereby testing the "absolute" accuracy of the response surface. Such statistics are designated as those above, but with a prime added, e.g., $\xi_1(q)$ becomes $\xi_1'(q)$.

the experimental design (i.e., the points in $\Theta\times\Gamma$ examined), and on the choice of DGP and $\Theta\times\Gamma$. Finally, even if any of (A)-(F) are rejected, the response surface still has certain desirable properties as an approximation to the unknown function $G(\cdot,\cdot)$ (White (1980b, pp. 155-157)), and it still may account for (and so summarize) much of the inter-experiment variation.

Instead of estimating response surfaces of the form (17), econometricians sometimes have estimated ones like:

$$s = h(\theta,T) + e \tag{19}$$

where $h(\theta,T)$ is the least squares approximation to $g(\theta,T)$ and e is the residual. Unlike (17), (19) does not account for the heteroscedasticity of s, conditional upon (θ,T), nor does it bound the range of $h(\theta,T)$, e.g., $h(\theta,T)$ could go outside the unit interval. Even so, White's standard errors are consistent; the response surface $h(\theta,T)$ <u>is</u> a least squares approximation to the underlying response function $g(\theta,T)$ and has the desirable properties that that entails; for this Monte Carlo study at least, some very simple response surfaces of the form (19) do very well at approximating π_T (as measured by the magnitude of the deviations $s-h(\theta,T)$); and, in so doing, those response surfaces succinctly summarize a large number of simulations. The more sophisticated response surfaces of the form (17) are more appealing theoretically and, with enough terms, can explain much of the remaining prediction error of the naive response surfaces, but the former lose in terms of summarizing the Monte Carlo results <u>because</u> of their complexity. For convenience, "naive" response surfaces (of the form (19)) are called type A; "sophisticated" ones (of the form (17)) are called type B.[16]

5. Simulation Evidence: A Two-equation Model

This section describes the experimental design, simulation, and post-simulation analysis of a Monte Carlo study of the nested and non-nested

---

[16] Type B response surfaces need not be complex, nor type A response surfaces simple. However, the latter's appeal lessens if they are complex.

hypothesis test statistics discussed in Section 2 for the dynamic simultaneous two-equation model in Section 3. At each stage, particular attention is given to techniques which will obtain as precise and general results as possible on the finite sample properties of those statistics.

5.1. Experimental Design

Following Hendry's (1984, p. 940) notation and terminology, the Monte Carlo design variables for the econometric model given in (4)-(6) are

$$\theta = (b, c, d, a, h', \sigma_{11}, \sigma_{12}, \sigma_{22}, (\mathrm{vec}\Lambda)', (\mathrm{vec}\Omega)')'$$
$$\in \Theta = \{\theta \mid |\rho_i|<1, i=0,\ldots,4;\ |\Sigma|>0;\ |\Omega|>0;\ \Sigma \text{ and } \Omega \text{ symmetric}\} \quad (20)$$

and

$$T \in \Gamma = [T_a, T_b] \quad (21)$$

where Γ is pre-assigned with T_a and T_b being the smallest and largest econometric sample sizes considered. Equations (4)-(6) are the data generation process (DGP); $\Theta\times\Gamma$ is the parameter space; equations (7)-(9) are the relationships of interest; and the objective of the Monte Carlo study is to determine the finite sample distributions of the statistics c_0, c_1, f_0, f_1, c_2, c_3, f_2, f_3, t_4, t_5, t_6, and t_7 as defined by the relationships of interest, within the specified parameter space of the DGP. More modestly, letting τ be any of those statistics and δ be the critical value associated with a test based on τ, the objective is to find the finite sample rejection frequency $\pi_T \equiv \mathrm{prob}(|\tau|\geq\delta)$. That probability depends upon θ and T and can be expressed as a conditional probability formula:

$$\pi_T \equiv \mathrm{prob}(|\tau|\geq\delta \mid \theta, T) = g(\theta, T) , \quad (22)$$

where π_T and $g(\theta,T)$ are precisely the probability and function discussed in Section 4.[17] Thus, we wish to know (or obtain a good approximation to) $g(\theta,T)$ over $\Theta\times\Gamma$, focusing on the effect on the statistics' finite sample power and type I error of dynamics, simultaneity, sample size, and (in the

---

[17] Implicitly, $g(\cdot,\cdot)$ is a function of δ as well. However, because δ is held constant for each of the statistics examined, its presence in $g(\cdot,\cdot)$ is ignored in the analysis below.

case of finite sample power) asymptotic power. Hence the key parameters in the experimental design are d, b, T, and π_a.

First, consider the <u>other</u> parameters. As in Hendry and Harrison (1974, pp. 164-166), the matrices Λ and Ω are held constant across experiments. In the present set of experiments, the diagonal elements of Λ and Ω are

$$(\rho_1 : \rho_2 : \rho_3 : \rho_4) = (.8 : .7 : .4 : .2) \tag{23}$$

$$(\omega_{11} : \omega_{22} : \omega_{33} : \omega_{44}) = (.5^2 : .5^2 : .7^2 : .7^2) \quad , \tag{24}$$

implying that $\mathrm{var}(w_{it})$ is virtually constant across i. Λ is a diagonal matrix (as in Hendry and Harrison (1974)); but Ω is not, with ω_{12} $(=\omega_{21})$ chosen such that $\gamma = .925$, and all other $\omega_{ij} = 0$ $(i\neq j)$. That implies that w_{1t} $(\simeq Z_t)$ and w_{2t} combined have a variance about twice that of w_{3t} and w_{4t} combined. All the exogenous variables are stochastic, i.e., varying across replications as well as across experiments. The parameters in the second equation are fixed across experiments, with $a = .3$, $h = (0 : 1 : 1 : 1)'$, and $\sigma_{22} = 1.0$. The error covariance σ_{12} is chosen such that $\mathrm{corr}(u_t, \nu_t) = .5$; and $c = 1.0$ (without loss of generality).

The values of the key parameters b, d, and T cover a range typical of econometric models estimated with actual data: $b = (-.5, .3)$, $d = (-.4, .2, .7)$, and $T = (20, 40, 80)$.[18] The number of replications N varies inversely with the econometric sample size ($N = 4000$ for $T = 20$, $N = 2000$ for $T = 40$, $N = 1000$ for $T = 80$), keeping computational costs virtually constant across sample sizes and giving more precise information on finite sample powers at smaller sample sizes (where the asymptotic approximations would be expected to provide less information about those finite sample powers). The error variance σ_{11} is the final parameter in the experimental design. Rather than assign it somewhat arbitrary values, possibly implying very high (or very low) finite sample powers, σ_{11} is set

---

[18] Cf. Hendry and Harrison (1974, p. 166) who chose a similar range for b, d, and T. See Klein (1969) and Hendry (1974) <u>inter alia</u> for estimated values of such parameters in empirical macro-economic models.

to obtain certain values of asymptotic power π_a, thus controlling to some extent the values of finite sample power π_T. So, σ_{11} is chosen such that

$$\pi_a = (.25, .5, .75, .90) \quad , \tag{25}$$

which seems a relevant range of powers, and one which ought to avoid having observed rejection frequencies too close to unity.[19] However, because several statistics are being considered, some having different asymptotic powers, it remains undecided as to which statistic the asymptotic power π_a in (25) corresponds. Because of the degrees of freedom involved and because H_0 is in fact the DGP, it is conjectured that (in general) t_7 would be most powerful, followed by c_3 and the asymptotically equivalent f_3, followed by c_1, with the placement of t_5 uncertain. So, to avoid any of the statistics having consistently high (or consistently low) power for all experiments, the asymptotic power of c_3 is the π_a in (25). Even so, the asymptotic power of t_7 implied by those values of b, d, T, and σ_{11} is always greater than one-half, and that for t_5 always less than .17, highlighting the difficulties of designing a Monte Carlo study for statistics with different asymptotic powers. (Note, however, that in the response surfaces below, π_a and ℓ ($\equiv \ln\{\pi_a/(1-\pi_a)\}$) are for whatever statistic is being examined, and not just for c_3.)

Given the choices of b, d, T, and π_a, a full factorial design is adopted, with 72 experiments in all. Three randomly selected experiments are retained from each econometric sample size for prediction. Estimation is by two-stage least squares.

5.2. Simulation and Computational Aspects

Noting the similarity between evaluating (for instance) D_0 under H_0 and H_1 and evaluating both D_0 and D_1 under H_0 only, and that the latter is computationally more efficient in these studies, only simulations under H_0

[19] See Appendix A, Mizon and Hendry (1980, p. 34), and Hendry (1984, p. 971) for counter-examples; cf. Poskitt and Tremayne (1981, pp. 266, 268).

were considered. Those Monte Carlo simulations were carried out with a modified version of Hendry and Srba's (1979, 1980) computer program NAIVE on the University of London's CDC 7600 computer. For a given set of parameters θ defining the DGP in (4)-(6), each (u_t, ν_t) was generated as a rescaled pair of normal pseudo-random numbers using Box and Muller's (1958) transformation on two uniform pseudo-random numbers.[20] Each v_{it} was generated as an appropriately rescaled sum of twelve pseudo-random uniform numbers from RNDM, which very closely approximates a pseudo-random normal number; see Hammersley and Handscomb (1964, pp. 39-40). The series for y_t, Y_t, and w_t were determined from those for u_t, ν_t, and v_t using (6) and the reduced form of (4)-(5). The relevant statistics were then calculated for each of N such replications.[21]

For a particular experiment, N replications were generated, of which S (dependent upon the statistic) were "successes" (e.g., the number of rejections; see Section 4 above). The fraction of successes s (=S/N) is an unbiased Monte Carlo estimator of the (unknown) finite sample rejection frequency π_T; and from those estimates, numerical-analytical approximations to π_T were obtained by estimating response surfaces as described in Section 4. To calculate the asymptotic powers of the statistics, the moments of the DGP for each experiment were obtained using Hendry and Srba's (1977b, 1980) program DAGER, from which the non-centrality of c_1, f_1, c_3, and f_3 and the asymptotic means and variances of D_1, t_5, and t_7 were

[20] Cf. Hendry and Harrison (1974, p. 153). The random number generators are Carrier, Atkins, and Taylor's (1969) mixed-congruential generator RNDM and NAg's (1977) multiplicative-congruential generator G05CAF. Different random number generators were used for each number in the pair of uniform pseudo-random numbers in order to avoid potential difficulties with Box and Muller's transformation: see Neave (1973). Nb. Hammersley and Handscomb (1964, Chapter 3), Kennedy and Gentle (1980, pp. 136ff). See Sowey (1972, 1973, 1978, 1986) and Sahai (1979) for bibliographies.

[21] The initial value for (y_t, Y_t, w_t) in each replication was its unconditional mean, so the first thirty observations generated for each replication were discarded to ensure stationarity of the series used for estimation and testing (cf. Hendry and Harrison (1974, p. 153)).

determined, using formulae in Ericsson (1983). The asymptotic powers of D_1, t_5, and t_7 were calculated assuming a symmetric two-sided test with critical values of ±1.96. The asymptotic powers of c_1 (f_1) and c_3 (f_3) were calculated assuming critical values corresponding to the 5% level, and approximating the non-central χ^2 (singly non-central F) by a <u>central</u> χ^2 (central F).[22]

5.3. <u>Post-simulation Analysis</u>

This subsection examines how well various analytical and numerical-analytical formulae approximate the underlying finite sample properties of the test statistics. To organize presentation, discussion centers around four formulae: the asymptotic (π_a), the F-adjusted asymptotic (π_x, explained below), and type A and B response surfaces ($\hat{\pi}_T$ and $\tilde{\pi}_T$). These are in order of increasing accuracy of approximation (and complexity) and are examined in that order. The value s is an unbiased estimate of π_T, so a natural measure of the degree of approximation of these formulae is their deviation from s. Graphs portray all this information concisely: results for (c_2, c_3), (f_2, f_3), (c_0, c_1), and (t_6, t_7) appear in Figures 1a-d, 2a-d, 3a-d, and 4a-d, respectively. For a given pair of statistics, Figures a and b graph the results for power and size, and Figures c and d plot the corresponding deviations of the formulae from s.[23] In the figures and elsewhere, the data are ordered by increasing values of T, then of π_a, of d, and of b. The remainder of this section describes the approximations obtained, interprets

---

[22] See Patnaik (1949), Johnson (1959), Kendall and Stuart (1973, pp. 237-240, 262-263), and Mizon and Hendry (1980, pp. 32-33) for further details.

[23] In all experiments, the observed rejection frequencies of t_5 are small (as are its asymptotic values), so no response surfaces are given for it. Although t_5 does not appear particularly useful in testing non-nested hypotheses when one of those hypotheses is the DGP, it shows promise for testing non-nested hypotheses when neither is the DGP (see Section 6 and Ericsson (1983, p. 294)).

The statistics f_0 and f_1 were not calculated in the Monte Carlo study, but there was little need to do so, given the results for π_x below.

them in light of existing finite sample theory, and proposes directions for further research.

The Asymptotic Approximation π_a. Figure 1a displays the asymptotic power (π_a) and the unbiased estimate of finite sample power (s) for c_3. As T increases, s converges to π_a, with the latter usually an upper bound of the former. The estimated finite sample size of c_2 generally is somewhat larger than its nominal value (.05) for small T, but tends to .05 as T increases (Figure 1b). The largest estimated finite sample size is 7.5%, occurring at the smallest sample size and with substantial dynamics and simultaneity present (T=20, b=-.5, d=.7). The properties of f_2 and f_3 resemble those of c_2 and c_3, except that discrepancies between s and π_a are typically smaller. That is particularly noticeable for the size, with s infrequently lying outside the interval [.035, .05], indicating how useful the F transformation is in small samples, even in the presence of dynamics and simultaneity. By contrast, the estimated finite sample sizes of c_0 and t_6 are almost always greater than 5%, and often exceed 15% (Figures 3b and 4b). The estimated finite sample power of c_1 generally exceeds its asymptotic power substantially: however, if its size were properly adjusted, its finite sample power would be considerably less, quite possibly bringing it more closely in line with its asymptotic values. The estimated finite sample power of t_7 deviates only slightly from its asymptotic power: adjusting its size would reduce its finite sample power considerably but generally increase deviations from its asymptotic values although the magnitude of those changes is difficult to predict.

The F-adjusted Asymptotic Approximation π_x. In certain circumstances, f_2 is exactly distributed as an F-ratio although c_2 remains only asymptotically χ^2 by failing to account for the variability in the estimated error variance used in its calculation. Because of the analytical relationship between c_2 and f_2, it is possible to calculate the size of the

Figure 1a. Asymptotic and Estimated Finite Sample Powers of c_3.

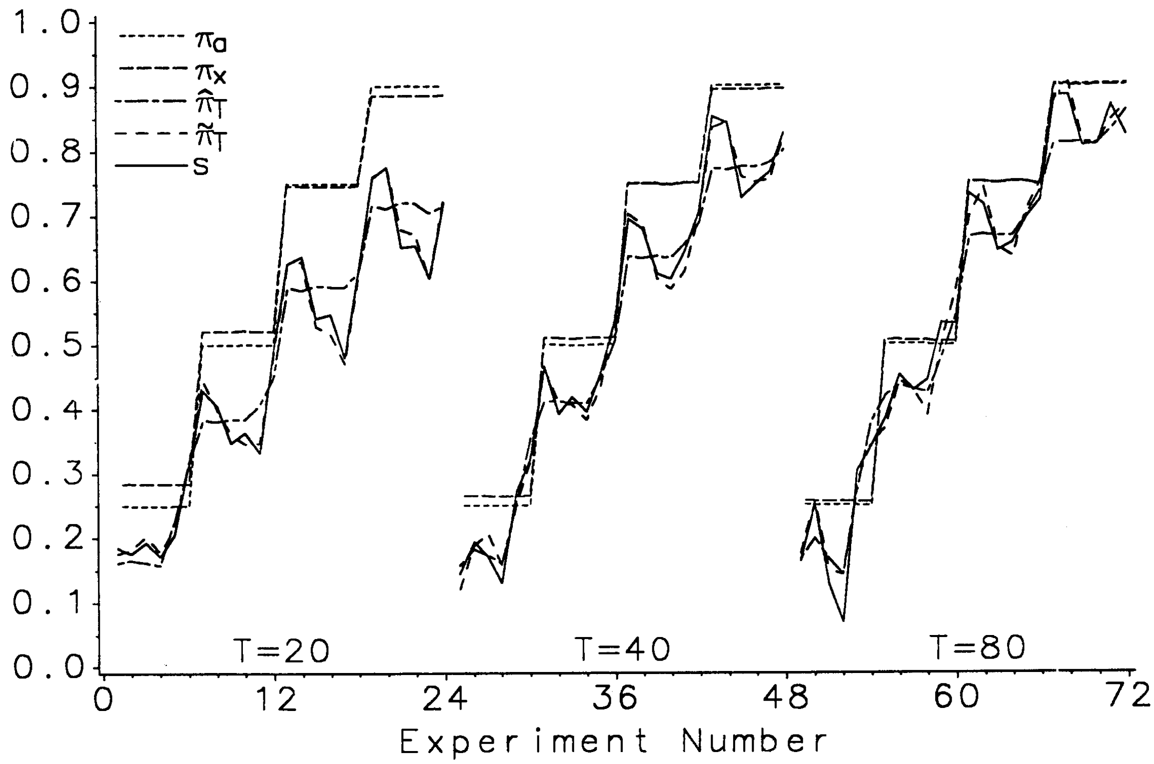


Figure 1b. Asymptotic and Estimated Finite Sample Sizes of c_2.

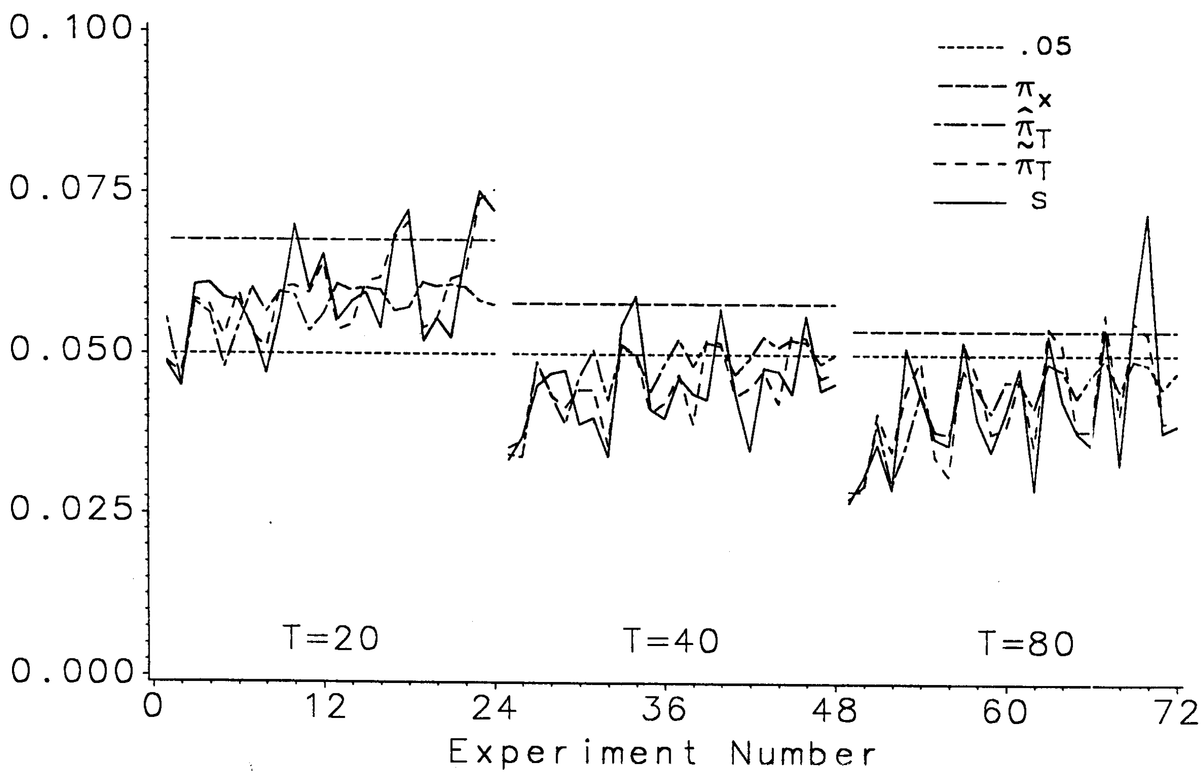


Figure 1c. Prediction Errors for Calculated Powers of c_3.

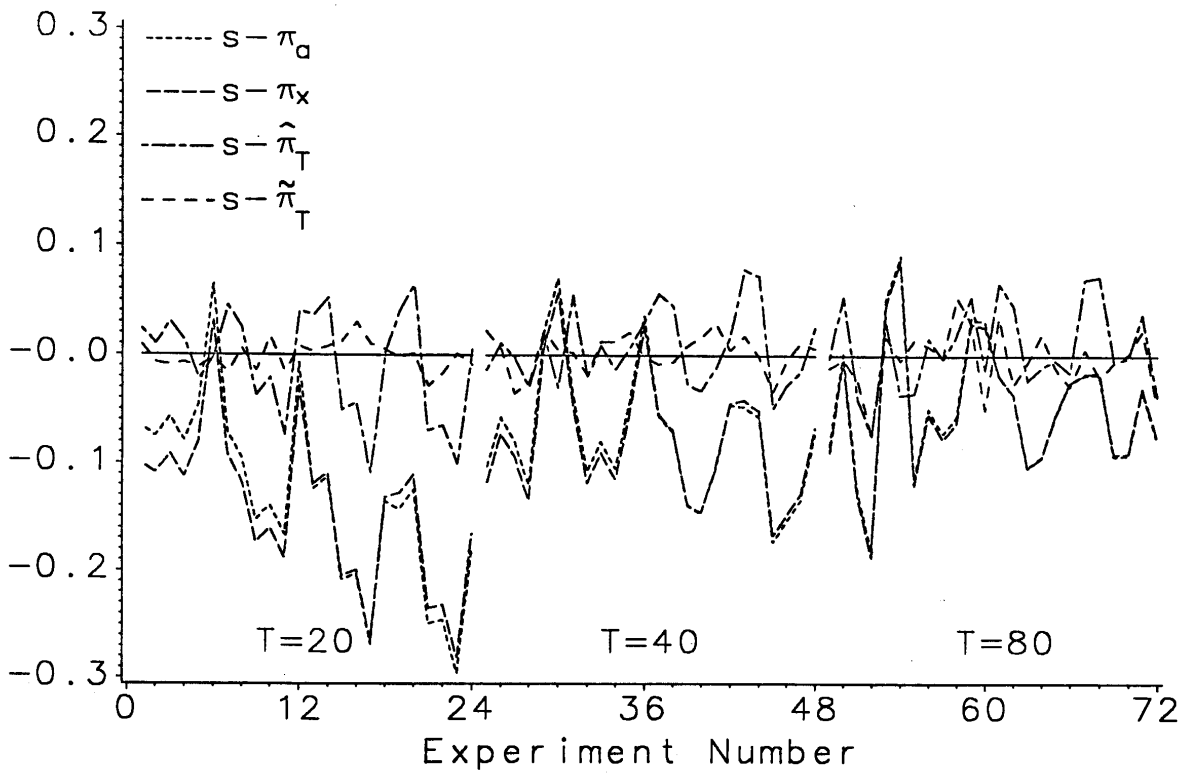


Figure 1d. Prediction Errors for Calculated Sizes of c_2.

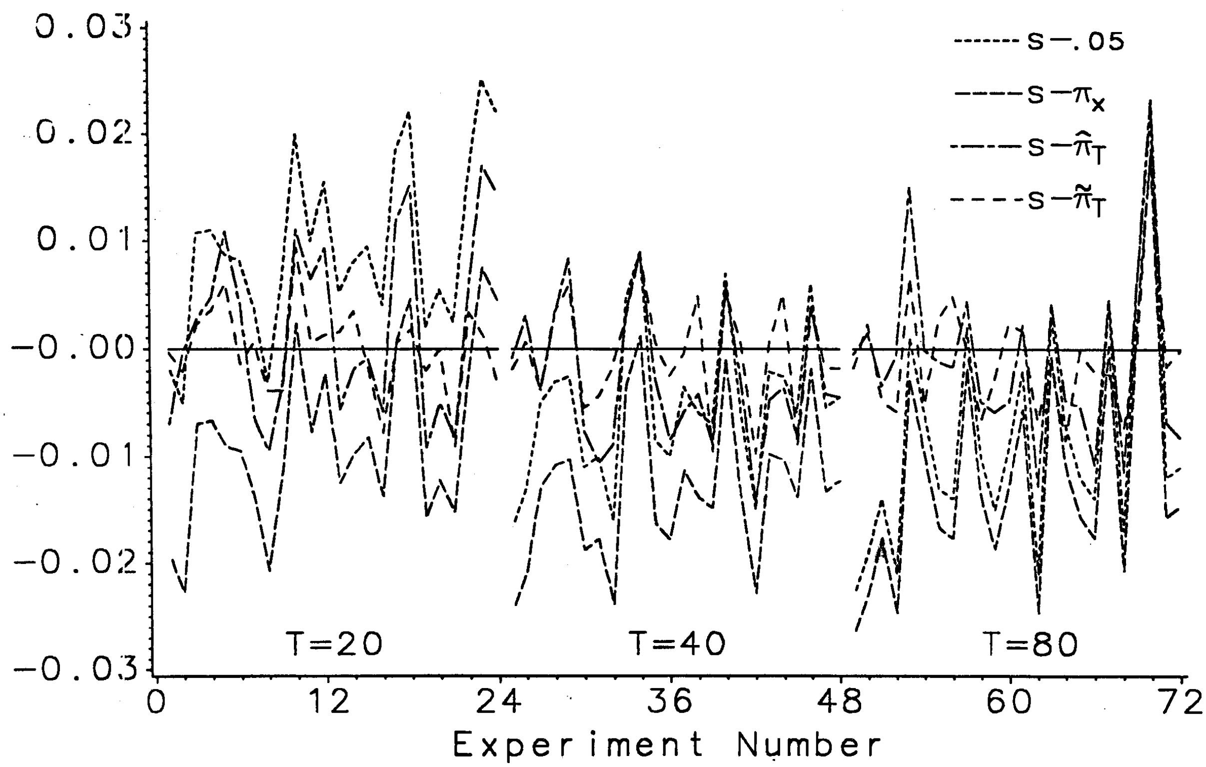


Figure 2a. Asymptotic and Estimated Finite Sample Powers of f_3.

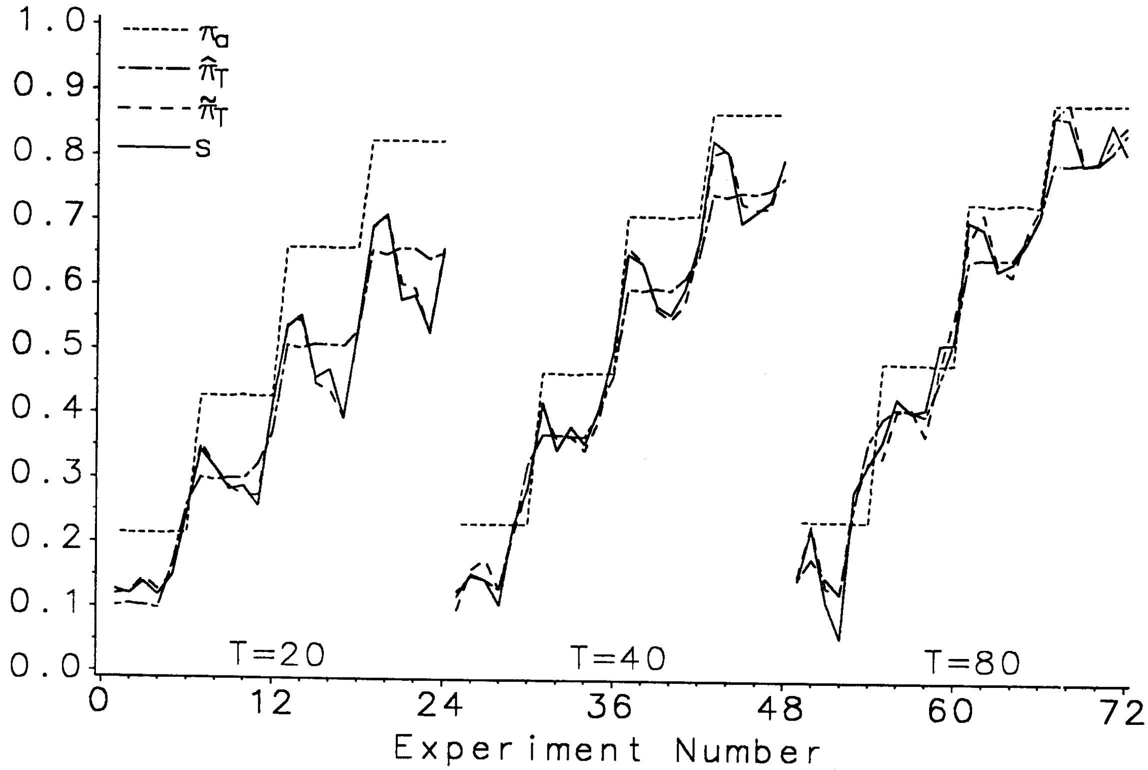


Figure 2b. Asymptotic and Estimated Finite Sample Sizes of f_2.

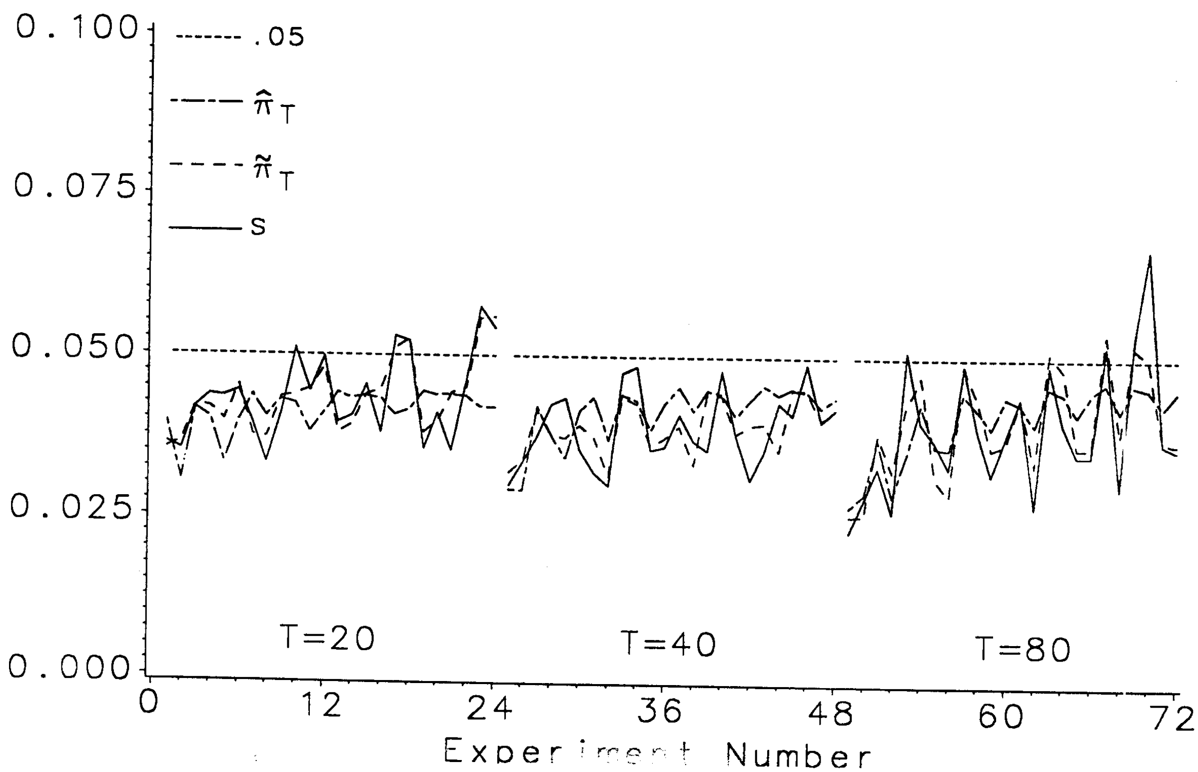


Figure 2c. Prediction Errors for Calculated Powers of f_3.

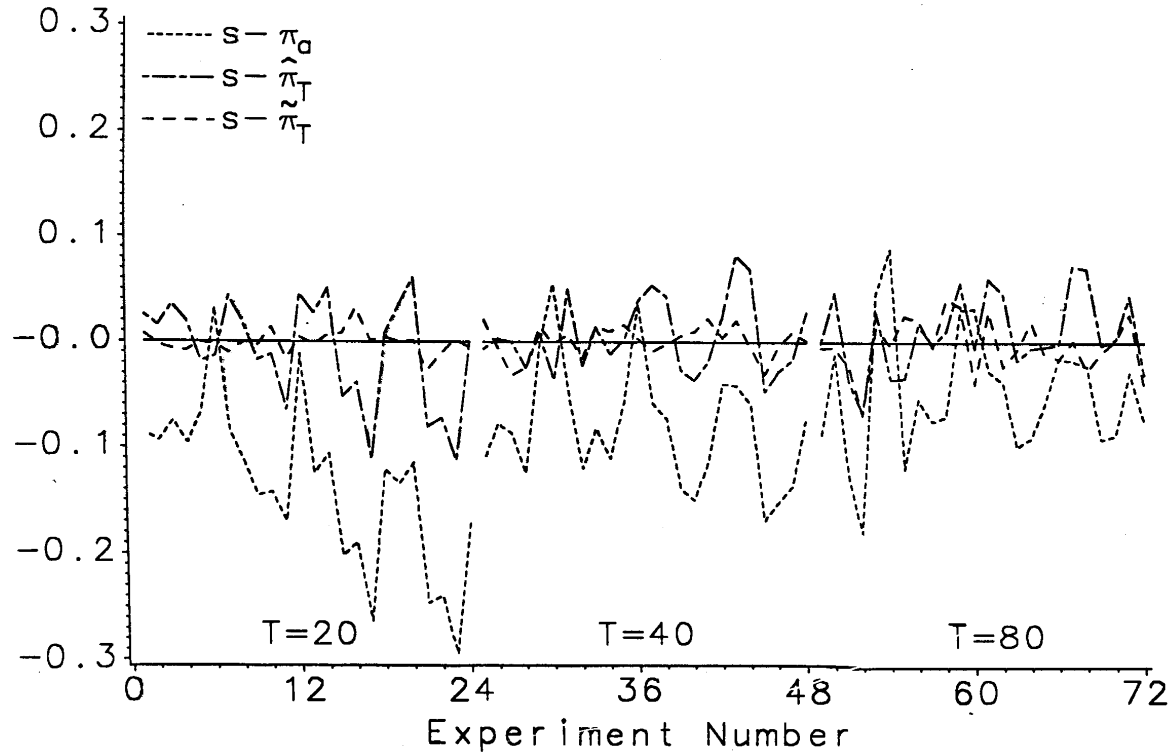


Figure 2d. Prediction Errors for Calculated Sizes of f_2.

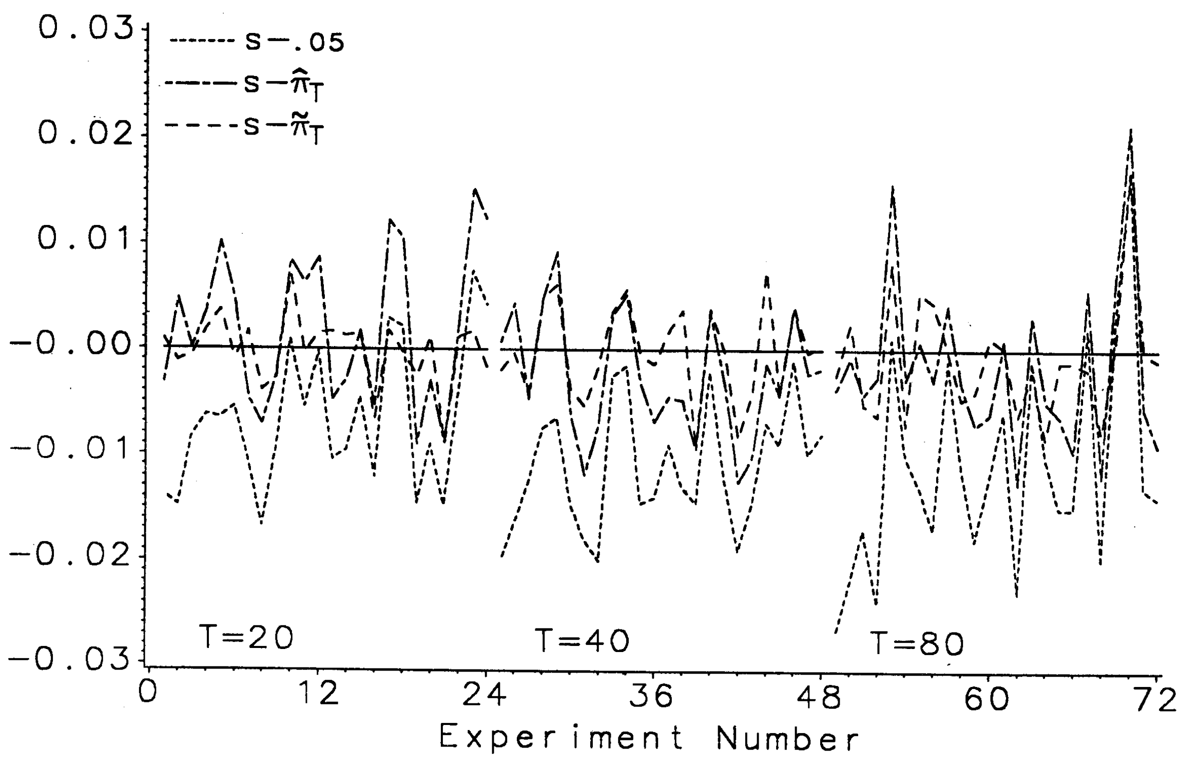


Figure 3a. Asymptotic and Estimated Finite Sample Powers of c_1.

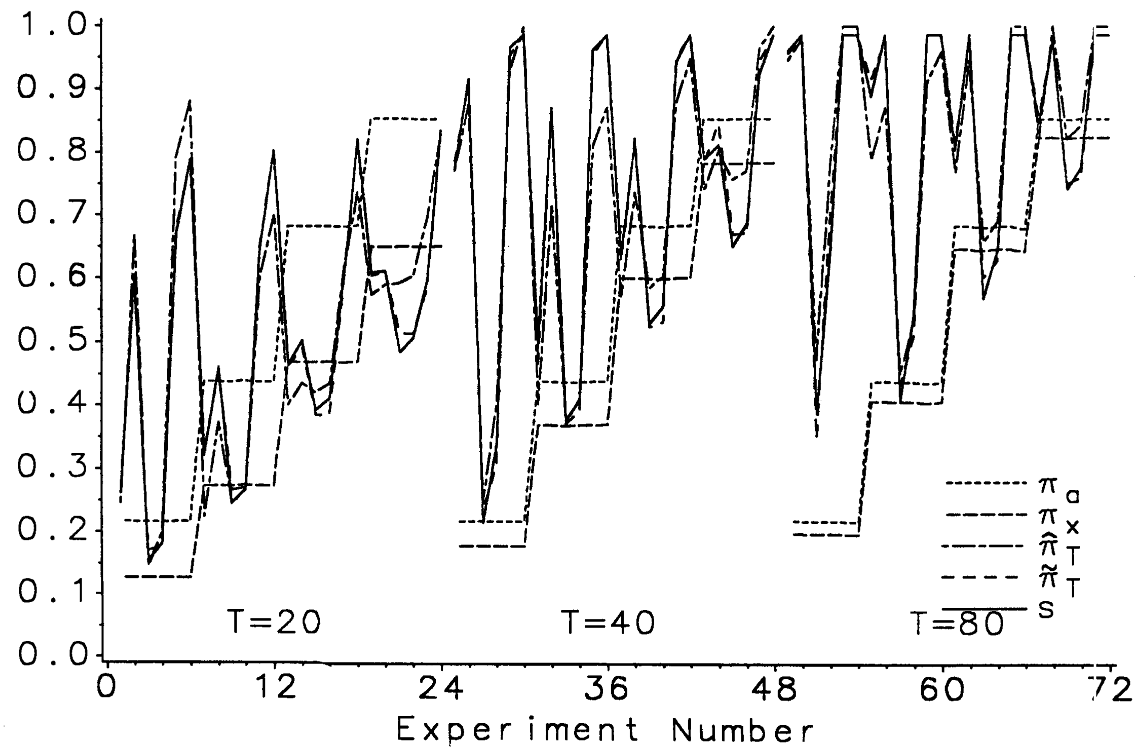


Figure 3b. Asymptotic and Estimated Finite Sample Sizes of c_0.

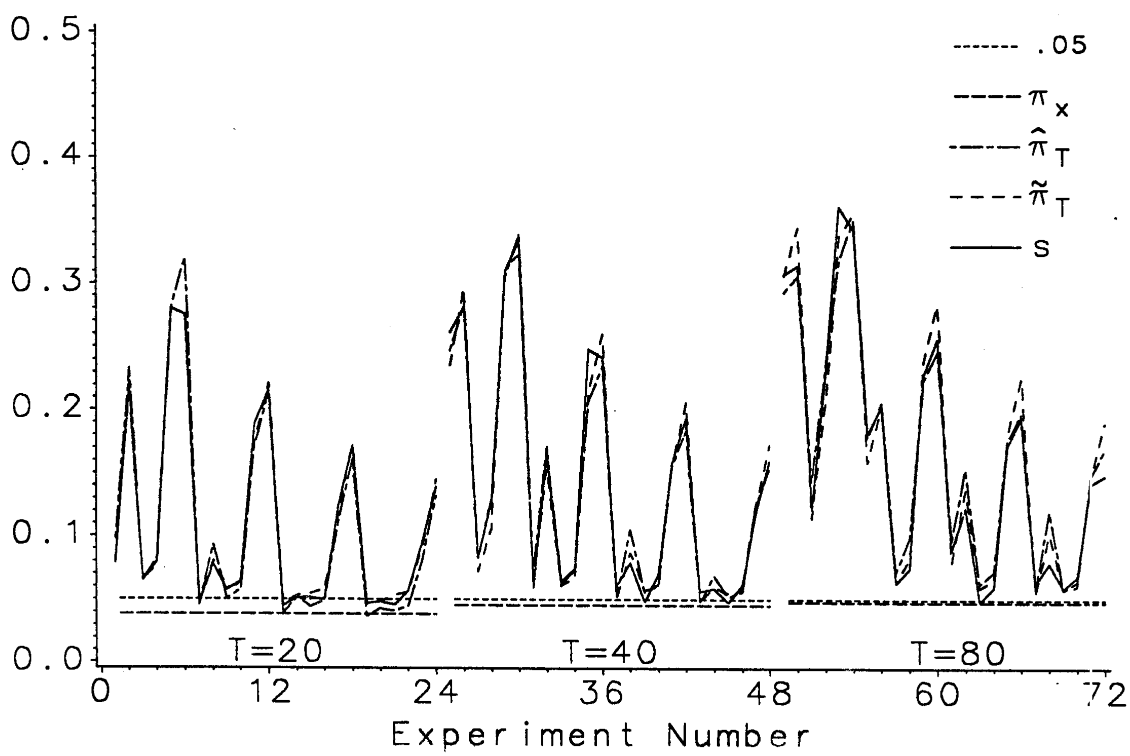


Figure 3c. Prediction Errors for Calculated Powers of c_1.

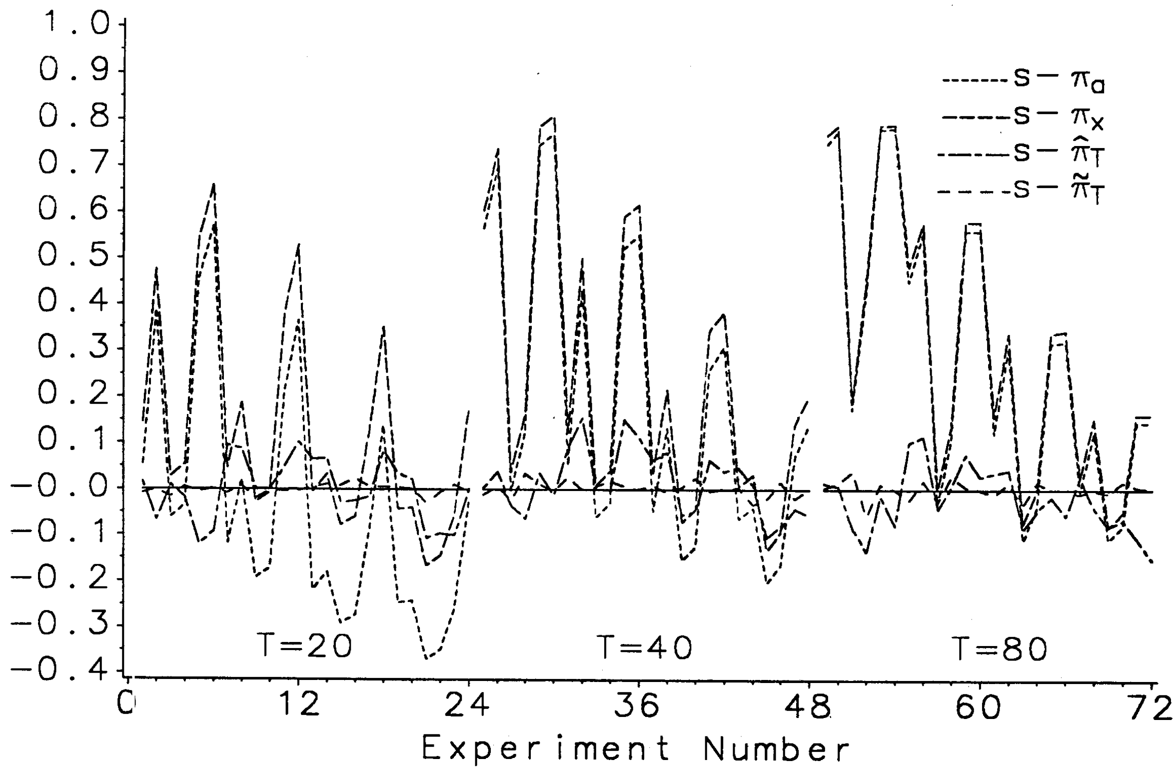


Figure 3d. Prediction Errors for Calculated Sizes of c_0.

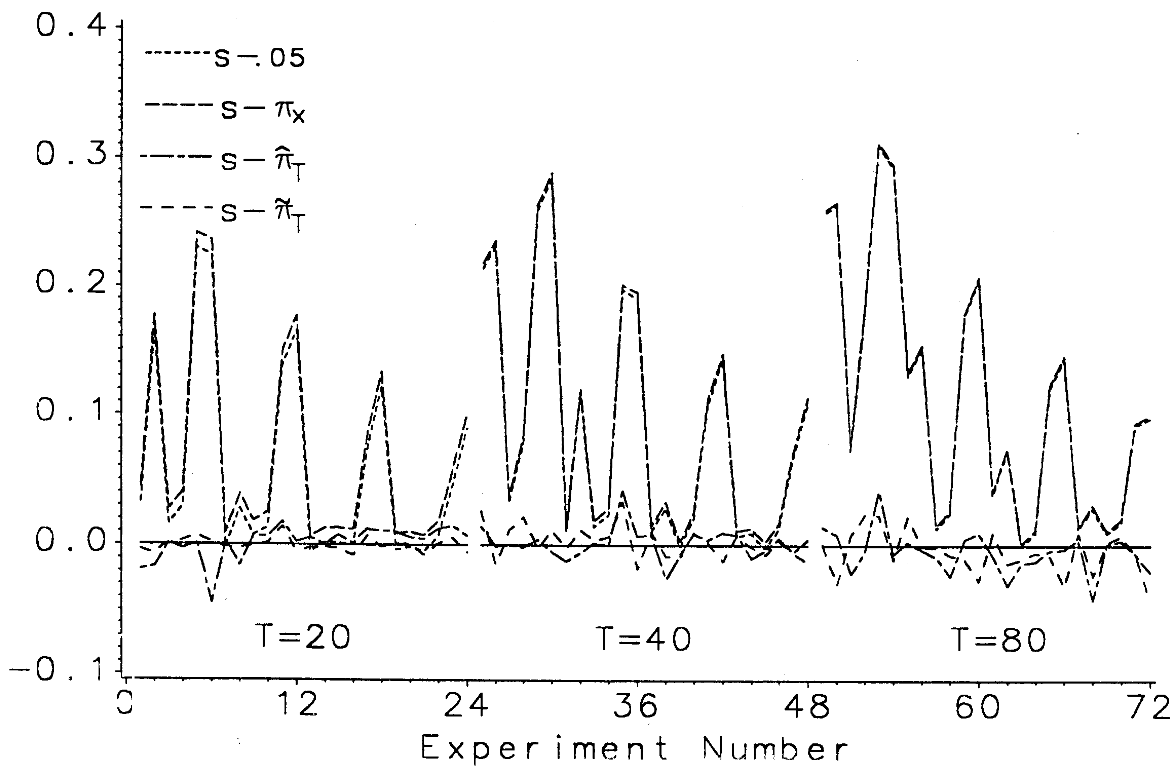


Figure 4a. Asymptotic and Estimated Finite Sample Powers of t_7.

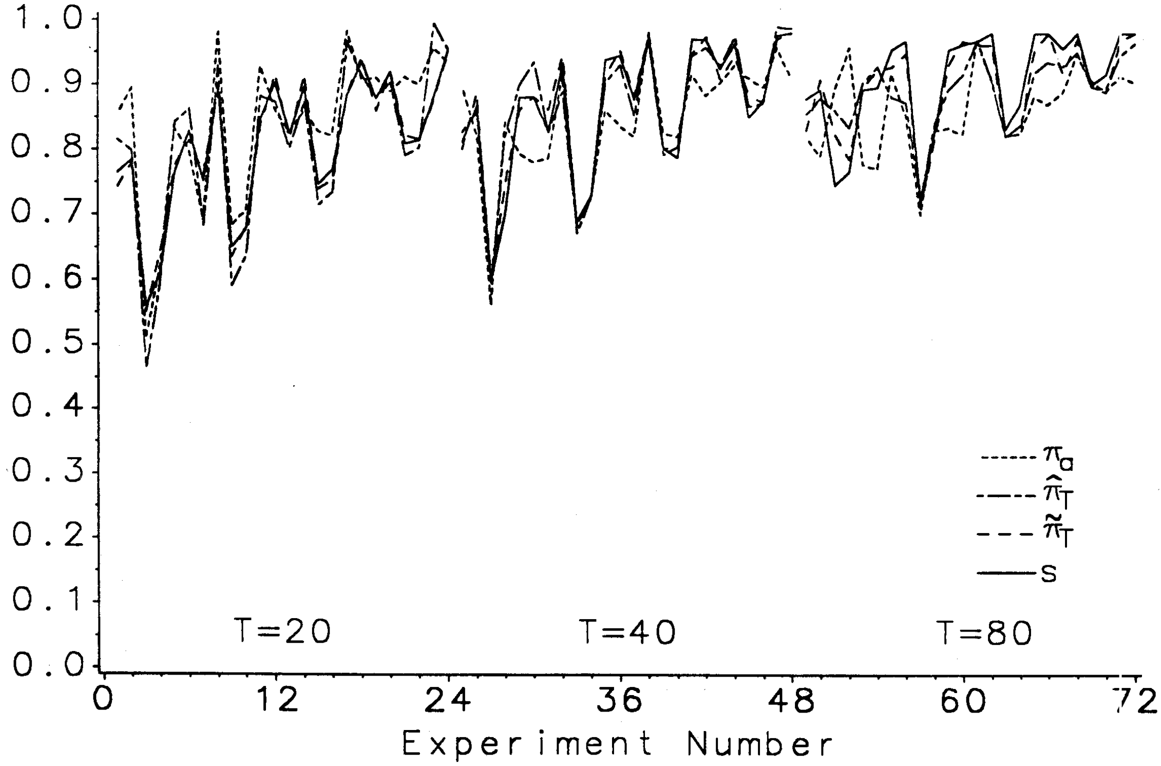


Figure 4b. Asymptotic and Estimated Finite Sample Sizes of t_6.

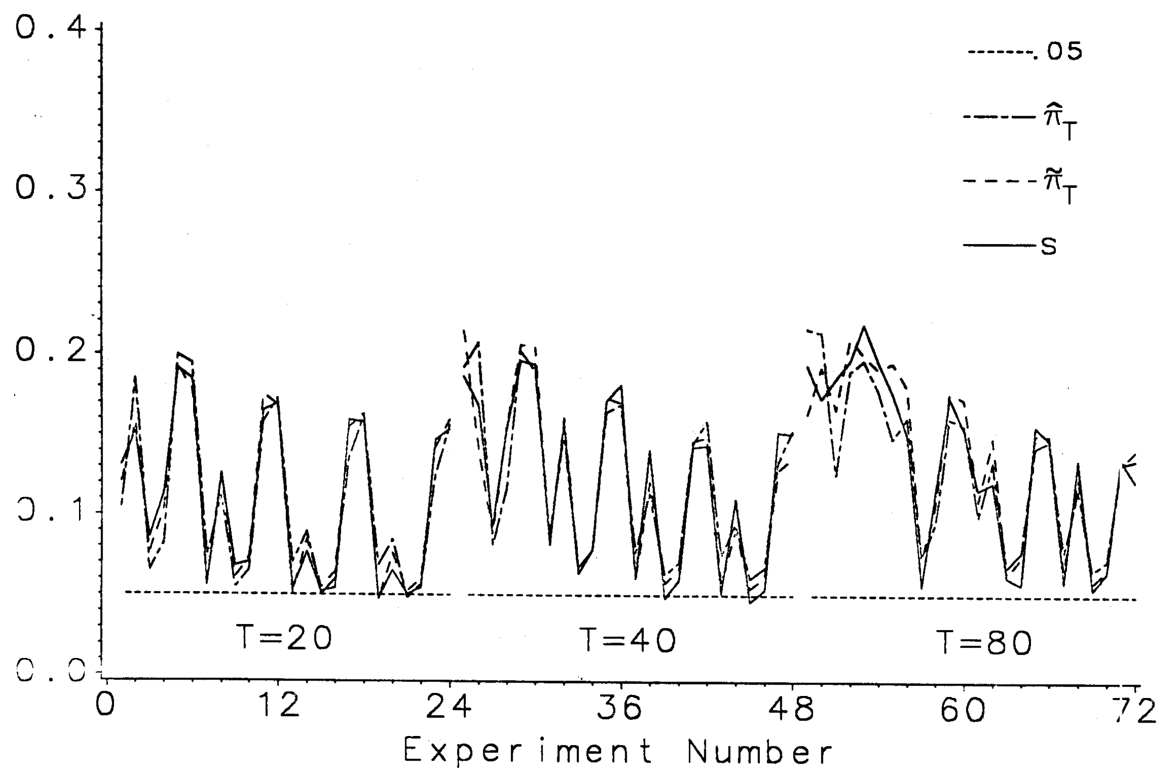


Figure 4c. Prediction Errors for Calculated Powers of t_7.

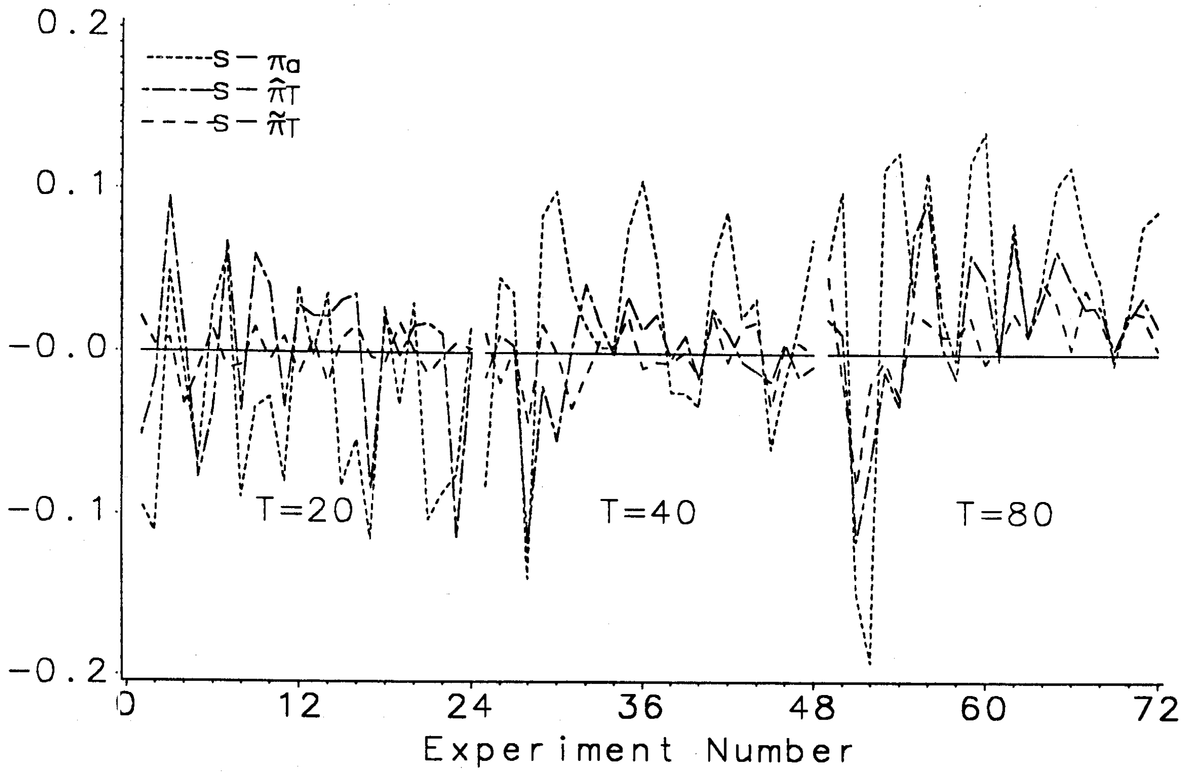
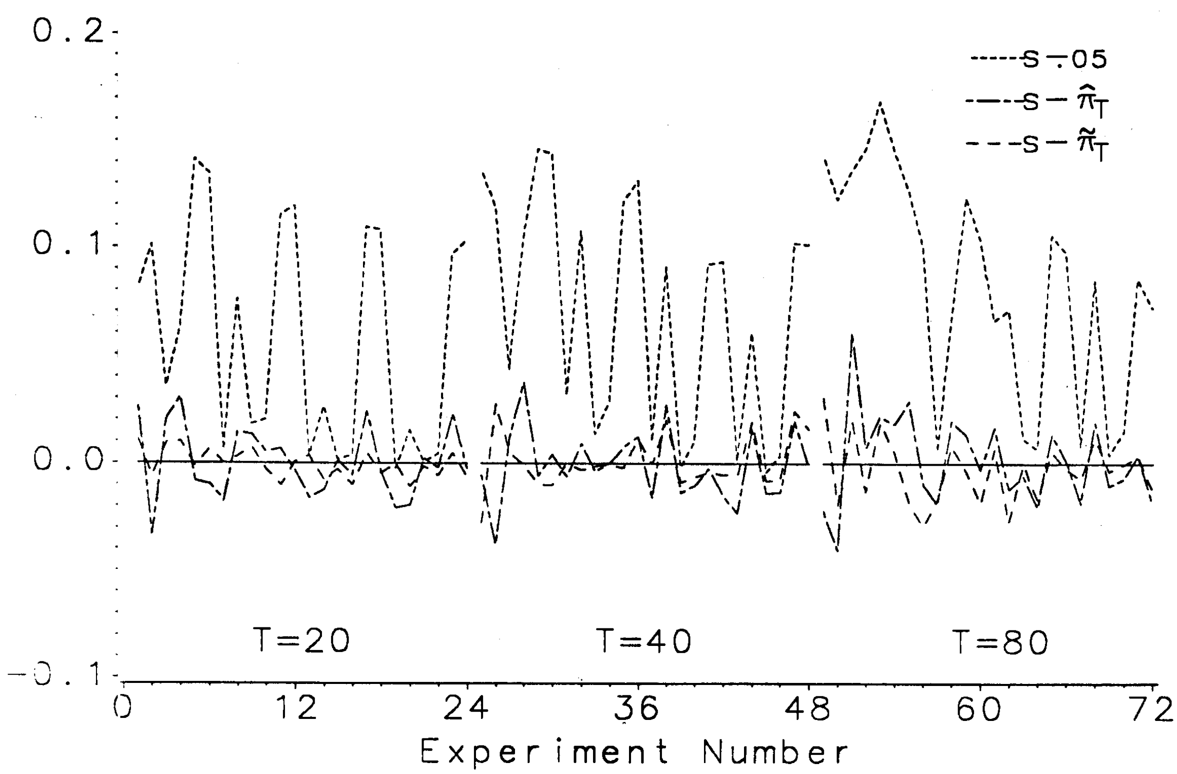


Figure 4d. Prediction Errors for Calculated Sizes of t_6.



test using c_2 (called the F-adjusted asymptotic size) that would result from taking that variability into account. A similar adjustment can be made for the asymptotic power: both are plotted in Figures 1a-b. Although typically $\pi_x - \pi_a$ for c_3 is small relative to the remaining discrepancy $s - \pi_x$, the F-correction captures a dominant finite sample term (of $O(T^{-1})$) in the distribution of c_2, as is apparent in Figure 1b. Because f_2 is not exactly an F-ratio in these experiments, the F-adjusted asymptotic approximation need not improve upon the regular asymptotic approximation although it is useful because it does.[24] Indeed, a similar correction for c_0 (and c_1) appears unimportant relative to remaining fluctuations in s.[25] Those fluctuations and similar ones for t_6 exhibit a pronounced pattern, inspiring the following digression.

The Effective Sample Size T^* and Large-σ Effects. The concepts of effective sample size and small-σ asymptotics are valuable for interpreting the finite sample fluctuations of all the statistics, and particularly those of c_0 and t_6.

Using a concept from Sims (1974), Hendry (1979a, 1984) develops the notion of an "effective sample size", that is, one which accounts for the lagged (and hence redundant) information accrued by each new observation of a dynamic process. For example, in an AR(1) process with autoregressive coefficient ρ, the effective sample size T^* is $T(1-\rho^2)$ where T is the number

[24] Kiviet (1986) studies a variety of mis-specification test statistics for a dynamic single equation and also finds the F form preferred to the χ^2 form.

[25] The F-adjustment to c_0 does aid analysis of further finite sample terms; cf. Table IV. Also, f_0 may be preferred over c_0 for other related reasons. c_0 is bounded from above by $(T-k_0)$, regardless of the DGP. Although that by itself may not induce large finite sample effects into the distribution of c_0 under H_0, it easily can under H_1 when λ_2 is sizable. Additionally, the numerator and denominator of c_0 in Ericsson (1983, eq. (10)) are positively correlated in finite samples, and that may lead to the finite sample distribution of c_0 deviating significantly from the χ^2 distribution. Nb. Sargan (1980c, pp. 1135-1137).

of observations. The greater the dynamics, the less new information on the process is gained with each additional observation, e.g., relative to a white-noise process: that is reflected in the measure T^*. In general, T^* involves all the latent roots of the dynamic system generating the data. However, for (4)-(6) and the experimental design in Section 5.1, ρ_0 is the only latent root that changes, so T^* is defined as $T(1-\rho_0^2)$. In static models, terms of $O(T^{-1/2})$ are often important, so in dynamic models the focus is on terms involving $(T^*)^{-1/2}$, denoted T for convenience.[26]

Most asymptotic results in econometrics are "large-sample", i.e., large T. Kadane (1970, 1971) proposes an alternative approach, small-σ asymptotics, in which T is held fixed and the equation error variance σ^2 is let to approach zero. Anderson (1977) discusses the relationship between large-T and small-σ asymptotics. Just as small-sample phenomena may appear when T is small enough, "large-σ" phenomena may exist for large enough σ. In (4), the error variance is σ_{11}, so $\sqrt{\sigma_{11}}$ is denoted σ.

The dominant finite sample term for c_0 appears to be $T\sigma$: that can be seen in several ways. In the cross-plot of s and $T\sigma$, their correlation is striking (Figure 5a). Even with no constant term and no correction for heteroscedasticity, the least-squares regression of (s-.05) on $T\sigma$ is:

$$(\bar{\pi}_T - .05) = \underset{\substack{(.0020)\\ [.0032]}}{.0890T\sigma} \qquad R^2 = .964 \qquad \bar{\sigma}_e = 2.388\% \tag{26}$$

where $(\cdot)$ and $[\cdot]$ denote conventionally calculated and White's (1980a, pp. 820-821; 1980b, p. 156) heteroscedasticity-consistent standard errors, R^2 is the unadjusted squared multiple correlation coefficient, σ_e is the square root of the residual variance, and $\bar{}$ denotes the least squares

[26]Also, Phillips (1977) shows that the first Edgeworth correction term to the normal distribution is O(T) for the t-ratio of the coefficient in an AR(1) process.

Figure 5a. Cross–plot of s (for c_0) and $T\sigma$.

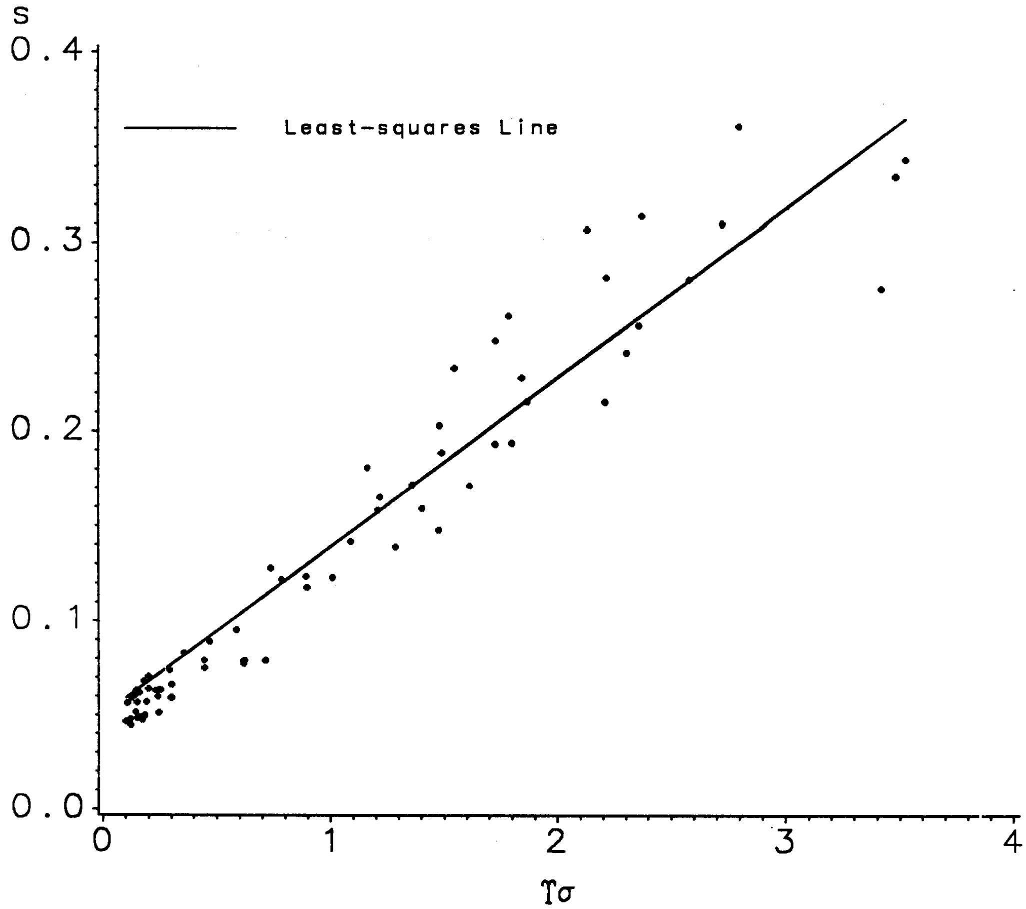
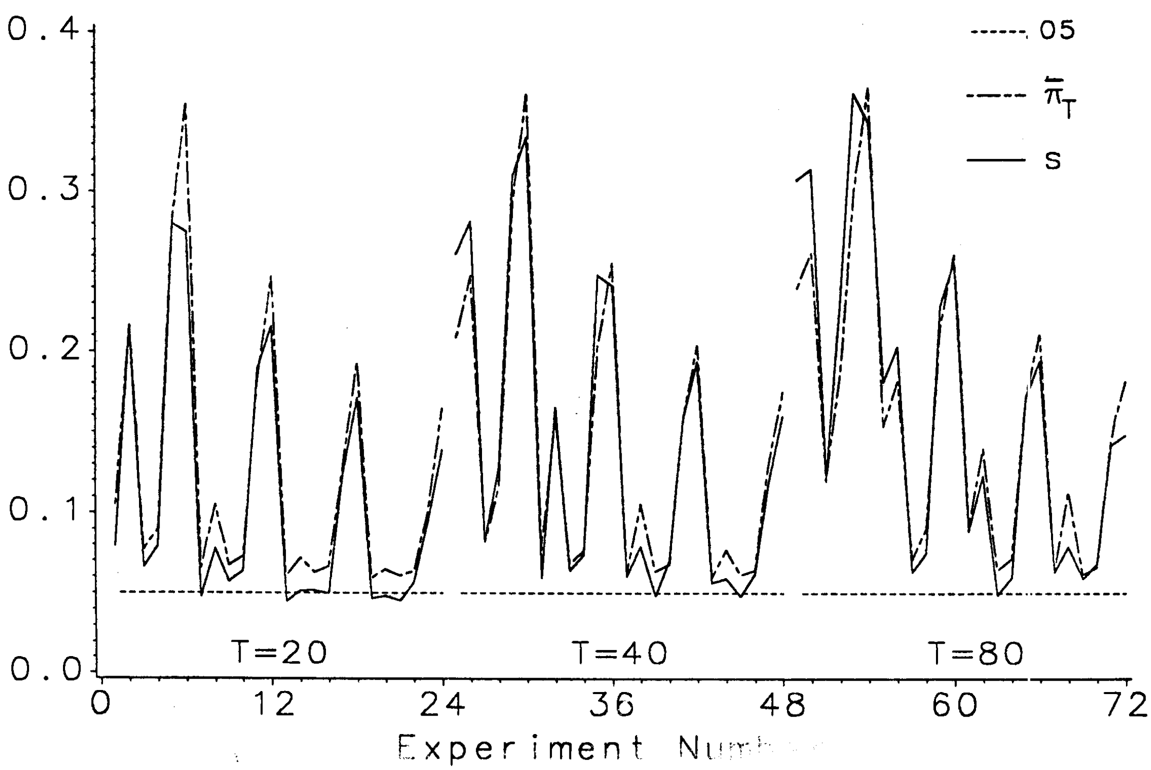


Figure 5b. The Nominal Size (.05), Unbiased Estimated Finite Sample Size (s), and a Simple Approximation to the Finite Sample Size ($\bar{\pi}_T$) of c_0.



estimate in this very simple type A response surface.[27] The estimated finite sample size s and the fitted values $\bar{\pi}_T$ from (26) appear in Figure 5b, along with the asymptotic value (.05). Although the standard error of the prediction errors is still 2.4%, it is small relative to the size of the fluctuations in π_T and is a remarkable reduction from 9.2%, the standard error of $s-\pi_a$. Individually, neither T nor σ approximate $\pi_T-.05$ nearly so well. Together they conveniently summarize effects of the sample size, dynamics, simultaneity, and goodness-of-fit, i.e., $T\sigma = \surd[\sigma_{11}/(T\{1-[d/(1-ab)]^2\})]$. Based upon this rather suggestive evidence and upon the analytics for effects from T and σ, θ is reparameterized to include T and σ in the response surfaces.

Before turning to the response surfaces, it is valuable to consider why T and (especially) σ so dominate this Monte Carlo study and not others. Two aspects of the experimental design are responsible. First, T^* and σ range widely, over [8, 78] and [.424, 20.2] respectively. For T^*, that arises because $|\rho_0|$ spans [.17, .77] and T, [20, 80]; σ is used to control π_a of c_2, given the values of (b,d,T) selected. Second, as <u>the</u> parameter for controlling π_a, σ effectively is one of the experimental design parameters, and one which varies over <u>all</u> experiments. Other investigators typically have normalized on σ, used it to control the population R^2, or included it explicitly as a design parameter but with a small number of values. Each approach has its merits, but none would be likely to elicit large-σ effects to the extent that the design in Section 5.1 does.

<u>Type A Response Surface Approximations ($\hat{\pi}_T$).</u> Tables III and IV give type A response surfaces for c_2, c_3, f_2, f_3, c_0, c_1, t_6, and t_7. They involve simple terms in T: T itself and/or T interacting multiplicatively with σ, κ ($\equiv |\rho_0|$), and (for powers) π_a or π_x. In simplifying from a

---

[27] Because there is no constant term in the regression, R^2 may lie outside the unit interval. However, σ_e, not R^2, is the appropriate measure of the goodness-of-fit for response surfaces.

Table III.
Estimates of Type A Response Surfaces for Finite Sample Sizes and Powers

Regressors and Diagnostic Statistics	Test Statistic τ and Dependent Variable[a]			
	Size		Power	
	f_2	c_2	f_3	c_3
	$s-.05$	$s-\pi_x$	$s-\pi_a$	$s-\pi_x$
T	-1.89	-2.29	-32.5	-36.4
	(.66)	(.70)	(9.5)	(10.9)
	[.67]	[.75]	[7.3]	[8.3]
$T\sigma$	-1.70	-1.87	-8.02	-7.82
	(.28)	(.30)	(1.74)	(1.80)
	[.21]	[.21]	[1.54]	[1.67]
$T\kappa\sigma$	2.08	2.23	17.8	17.7
	(.39)	(.42)	(2.3)	(2.4)
	[.29]	[.31]	[2.0]	[2.1]
$T\pi_a$ ($T\pi_x$ for c_3)			-.459	-.381
			(.125)	(.134)
			[.114]	[.117]
R^2	.661	.698	.868	.870
$\hat{\sigma}_e$	.743	.797	4.315	4.329
Means and Standard Deviations of Comparable Variables (as percentages)[b]				
s	4.04	4.86	47.36	51.51
	(.81)	(1.16)	(23.25)	(22.87)
$s-\pi_a$	-.96	-.14	-8.86	-8.48
	(.81)	(1.16)	(7.42)	(7.83)
$s-\pi_x$	-	-1.12	-	-9.03
		(.88)		(7.42)
$s-\hat{\pi}_T$	-.06	-.05	.20	.18
	(.73)	(.78)	(4.22)	(4.23)
$s-\tilde{\pi}_T$	-.03	-.04	-.08	-.08
	(.46)	(.49)	(1.88)	(2.07)
$s-\pi_T$	0	0	0	0
	(.47)	(.50)	(1.06)	(1.06)

Notes: a. The dependent variable (and π_a and π_x when they appear on the right-hand side) are rescaled by 100 to make them percentages and to achieve a reasonable scaling of coefficients. That implies that $\hat{\sigma}_e$ is a percentage; σ, however, is in its original units.

b. Noting that $s-\pi_T$ has zero mean and variance $\pi_T(1-\pi_T)/N$, for each experiment the latter is approximated as $s(1-\bar{s})/N$, with the values given above being based on averages across experiments.

Table IV.
Estimates of Type A Response Surfaces for Finite Sample Sizes and Powers

Regressors and Diagnostic Statistics	Test Statistic τ and Dependent Variable[a]			
	Size	Power	Size	Power
	c_0	c_1	t_6	t_7
	$s-\pi_x$	$s-\pi_x$	$s-.05$	$s-\pi_x$
T	-5.75 (1.29) [.98]	-48.3 (6.4) [5.5]		
$T\sigma$	14.24 (.54) [.55]	46.7 (2.7) [2.9]	11.35 (.66) [.89]	60.2 (11.0) [9.5]
$T\kappa$			64.9 (10.5) [8.4]	262.5 (40.1) [38.1]
$T\kappa\sigma$	-7.05 (.77) [.85]	-25.6 (3.8) [4.2]	-12.4 (1.2) [1.3]	-11.1 (4.4) [3.9]
T^2			-74.5 (17.4) [13.7]	-424.3 (57.7) [51.3]
$T\pi_a\sigma$				-.689 (.118) [.109]
R^2	.988	.963	.957	.636
$\hat{\sigma}_e$	1.457	7.221	1.876	4.645
Means and Standard Deviations of Comparable Variables (as percentages)[b]				
s	13.70 (9.15)	69.27 (25.18)	12.08 (5.25)	86.32 (10.48)
$s-\pi_a$	8.70 (9.15)	14.70 (32.08)	7.08 (5.25)	1.17 (7.39)
$s-\pi_x$	9.31 (9.06)	23.24 (28.60)	-	-
$s-\hat{\pi}$[illegible]	-.07 (1.43)	-.31 (7.11)	.03 (1.84)	.54 (4.36)
$s-\tilde{\pi}$[illegible]	-.09 (1.31)	.03 (1.63)	-.10 (1.21)	.26 (2.06)
$s-\pi_T$	0 (.83)	0 (.86)	0 (.79)	0 (.73)

Notes: See the notes for Table III.

regression in simple products of these parameters to a given response surface in Table III or IV, the primary criteria were parsimony and a small prediction error (not necessarily complementary criteria).

The size and power of the Wald statistic strongly resemble those of the F-statistic, once netted of its F-adjusted asymptotic approximation π_x rather than the simple asymptotic approximation π_a: the similarities are apparent both from the estimates for the respective response surfaces (Table III) and from the resulting predictions and actual values of s (Figures 1a-d and 2a-d). The terms in the response surfaces for the size of c_2 and f_2 are highly significant statistically, but the lower portion of Table III shows that they achieve only a moderate reduction in the residual standard deviation. That arises because the variances of $(s-\pi_x)$ and $(s-\pi_a)$ (for c_2, f_2) are close to their theoretical minima, i.e., the variances arising exclusively from the sampling of s, equivalent to the square of the residual standard deviation obtained if $g(\theta,T)$ were known (estimated by the last row in the table). That reflects the observation above for Figures 1b and 2b that estimated sizes for these statistics stay close to 5% over the entire range of sample sizes, dynamics, and simultaneity. The reduction is more substantial for powers, approximately threefold in the residual variance.

The type A response surfaces for c_0, c_1, t_6, and t_7 (Table IV) are similar in form to those for c_2, c_3, f_2, and f_3, but the magnitude and statistical significance of the estimated coefficients of the former are far greater than those of the latter and the signs of the estimated coefficients generally are reversed. For instance, the strong positive biases from $T\sigma$ for c_0, c_1, t_6, and t_7 are five to ten times the magnitude of comparable negative biases for c_2, c_3, f_2, and f_3; observed negative biases in the finite sample size of c_0 and t_6 are negligible, but positive ones are large and frequent. Because of the analytical relationships between c_2, c_3, f_2, and f_3, and, to a lesser extent, between c_0, c_1, t_6, and t_7, similarities in properties of statistics of either set are expected. However, neither the form, magnitude, nor sign of the finite sample effects was anticipated, nor was the considerable discrepancy between properties of statistics in different sets.

Type B Response Surface Approximations ($\tilde{\pi}_T$). Unrestricted type B response surfaces are estimated with regressors involving T, being multiplicative combinations of T, σ, ρ_0, and (for powers) ℓ, and combinations of powers thereof. Because of the experimental design, the factors T, σ, ρ_0, and ℓ, appear up to (and including) powers of three, two, one, and one, respectively. All combinations except for $T\sigma^2\rho_0\ell$, $T^2\sigma^2\rho_0\ell$, $T^3\sigma^2\rho_0\ell$ are included initially: that implies thirty-three regressors in the unrestricted response surfaces of estimated finite sample powers and eighteen in the unrestricted response surfaces of estimated finite sample type I errors.[28] Restricted versions of all response surfaces are presented in Tables I and II of Appendix C. No sets of restrictions are rejected at the 5% significance level. Unlike those for the type A response surfaces, the primary criteria here are those listed in Table I. Parsimony is not central in this framework, and so the response surfaces are more complex than those in Tables III and IV, while capturing more of the deviations between the finite sample and asymptotic properties of the test statistics.

In the response surfaces for c_2 and f_2, $\tilde{\sigma}_e$ is insignificantly different from unity, leaving little residual variation beyond that inherent from estimating π_T by s. With the additional complexity of these response surfaces, the standard deviations of the prediction errors for f_2 and c_2 fall from .73 and .78 (for s-$\hat{\pi}_T$) to .46 and .49 (for s-$\tilde{\pi}_T$) versus the

---

[28] There are fewer regressors in the latter because the unit vector is collinear with ℓ, which is in that case ln(.05/.95). Cf. Cochran and Cox (1957, pp. 148ff, 342ff) and Cox (1958, pp. 113-117) on factorial design and response surfaces.

$T\sigma^2\rho_0\ell$, $T^2\sigma^2\rho_0\ell$, $T^3\sigma^2\rho_0\ell$ are not included due to limitations in the number of regressors in PC-GIVE.

estimated theoretical lower bounds of .47 and .50. Similar moderate reductions in variation are achieved for c_0 and t_6 although substantial explainable variation remains unexplained for even type B response surfaces.

Type B response surfaces fare better for the powers of these tests, although still at the expense of greater complexity. Graphs c and d for Figures 1-4 illustrate the reductions in prediction error across the different predictors (π_a, π_x, $\hat{\pi}_T$, and $\tilde{\pi}_T$), with the most marked improvements generally being between the purely analytical formulae (π_a, π_x) and the response surfaces ($\hat{\pi}_T$, $\tilde{\pi}_T$). That does not discount the value of analytical formulae: to the contrary, the response surfaces serve to augment whatever analytical results are available. Further, the analytical formulae for powers frequently explain much of the <u>unconditional</u> inter-experiment variation in s, as is apparent both from Figures 1a-4a and from the first two rows of the lower half of Tables III and IV.[29] Also, although the finite sample terms explain over 95% of the variation in $L^*(s)-L(\pi_a)$ for c_1 and t_7, the value of $\tilde{\sigma}_e$ implies that under 25% of the remaining (residual) variation is due to sampling fluctuations (i.e., in estimating π_T by s), with over 75% being due to additional finite sample components. Better analytical approximations would be of considerable value here.

<u>Remarks.</u> In retrospect, several features of the experimental design are notable.

(a) Even though many of the response surfaces appear mis-specified, White's standard errors are consistent under the sorts of mis-specification present and the coefficient estimates are still useful for prediction within the population being investigated: see Hendry (1982, pp. 210-211) and White (1980b, pp. 155-157). In fact, both type A and type B response surfaces track the simulation estimates of the finite sample size and power

[29]Hendry (1973) convincingly argues the merits of analytical formulae in interpreting Monte Carlo studies.

remarkably well, in spite of mis-specification. That apparent contradiction has the following explanation. For a given set of experiments and their associated rejection frequencies, the estimates of the coefficients in a response surface are essentially invariant to the number of replications N, whereas ($\tilde{\sigma}_e^2$ - 1) is proportional to N times the square of any unexplained finite sample fluctuation.[30] The large values of N, both in Section 5 and Appendix A, magnify the effect on $\tilde{\sigma}_e$ of discrepancies between π_T and $\tilde{\pi}_T$ (or $\hat{\pi}_T$) although those discrepancies themselves are insensitive to N and appear quite small in general. Hence, a larger number of experiments with fewer replications per experiment would have been preferable.[31]

(b) For a third of the experiments, the largest latent root of (4)-(5) is .77, resulting in considerable dynamics affecting the statistics. Although that latent root is smaller for the other experiments, the largest latent root of the <u>entire</u> system (4)-(6) is always ρ_1 (=.8). Also, the actual sample size is only twenty for a third of the experiments, equaling the smallest sample size in Pesaran's (1974) experiments[32]; the <u>effective</u> sample size is sometimes as small as <u>eight</u>.[33] In light of that, the experimental design may be over-ambitious.

---

[30] Note that A = {S(N-S)}/(N-1) ≈ Ns(1-s) and that the rescaling factor $A^{1/2}$ is applied to <u>all</u> variables.

[31] Control variates for the estimated finite sample power might have been used to achieve more efficient Monte Carlo estimates of π_T: see Sargan (1976a, pp. 444-448) and Rothery (1982). However, their derivation appears practically intractable for most dynamic models: cf. Nankervis and Savin (1985).

[32] The Cox statistic departs significantly from its asymptotic properties in some of those experiments even though no dynamics or simultaneity is present and fewer instruments are used: see Ericsson (1986).

[33] Hendry and Neale's (1987) recently developed recursive Monte Carlo techniques permit rapid graphical analysis of estimator's properties for <u>every</u> feasible sample size up to the largest: parallel techniques for statistics would permit far more extensive analysis of their T- and T*-dependent properties than currently feasible.

(c) There is a clear need to limit the range of the (unknown) finite sample power of a statistic over experiments so as to avoid generating experiments with uninformative estimated powers (e.g., unity). Setting σ_{11} such that the asymptotic power π_a takes particular values goes some way to achieving that, although there is an inherent difficulty present when statistics with different asymptotic powers are being examined. Further control over the range of π_T may be possible, e.g., by using different critical values for statistics with different asymptotic powers.

(d) Small sample adjustments to statistics are a long-run objective of studying their finite sample properties and could take many forms in addition to the F-adjustment. Under the hypothesis (7), c_0 has more degrees of freedom than any of the other statistics (likewise, c_1 has more than any other, under the hypothesis (8)); and that may be partly responsible for the apparent relative poorness of the asymptotic approximation for c_0 (and for c_1).[34] Sargan's (1980c) transformation of the IV criterion may ameliorate that effect. Also, the properties of t_6 may improve from using that criterion rather than c_0 in constructing t_6. Godfrey and Pesaran (1983) propose bias adjustments to the numerator and denominator of the Cox statistic and thereby design a Cox-type statistic with a better finite sample size. Similar adjustments may be possible for the IV statistics, at least when $Z^* = X_2$. Finally, further analytical results on the statistics' finite sample properties, even for simple models, could be of value for deriving finite sample adjustments.

To summarize, the finite sample properties of the Wald and F statistics are quite closely in line with their asymptotic properties, with the F-adjustment capturing a dominant finite sample term of $O(T^{-1})$ in the distribution of c_2. The behavior of f_2 (versus that of c_2) favors use of

[34] See Sargan (1958, pp. 393, 400, 409, 414-415) and Sargan and Mikhail (1971, pp. 156-158) on similar considerations for the distributions of econometric estimators.

the F form of the Wald statistic rather than its χ^2 form, even for situations in which the Wald statistic has no known exact sampling distribution. A similar correction for c_0 appears unimportant relative to remaining fluctuations in π_T. The finite sample sizes of both c_0 and t_6 are typically strongly and positively biased: $T\sigma$ accounts for much of that bias. That contrasts with small and negative biases by $T\sigma$ for c_2 and f_2. Biases in the finite sample power of c_1 and t_7 are even larger, with T and σ being primary explanatory factors. Because of the analytical relationships between c_2, c_3, f_2, and f_3 and, to a lesser extent, between c_0, c_1, t_6, and t_7, properties of statistics within either set are generally similar: that helps to unify the results.

6. An Empirical Example

Pesaran and Deaton (1978) consider several (hypothesized) non-nested economic relationships between consumers' expenditure and income to demonstrate the application of Cox's maximum likelihood (ML) statistics in econometric modeling. This section re-examines their two linear models to illustrate the use of the IV statistics. Those models are:

$$H_0: \quad C_t = \beta_{00} + \beta_{01}Y_t + \beta_{02}W_t + u_{0t} \qquad u_{0t} \sim NID(0,\sigma_0^2) \qquad (27)$$

$$H_1: \quad C_t = \beta_{10} + \beta_{11}Y_t + \beta_{12}C_{t-1} + u_{1t} \qquad u_{1t} \sim NID(0,\sigma_1^2) \qquad (28)$$

where the data are quarterly, seasonally adjusted series in constant 1958 dollars for the United States (1954ii-1974iii) for consumers' personal expenditure (C), personal disposable income (Y), and personal wealth (W), with wealth measured at the beginning of each period. Throughout their analysis, Pesaran and Deaton assume that conditioning upon current income does not affect inference about the β_{ij}'s: IV estimation allows relaxation of that assumption. The IV statistics are calculated for several possible sets of instruments, namely:

$$\{1,\ W_t,\ ((C_{t-i},\ Y_{t-i}),\ i=1,\ldots,\iota)\} \qquad (29)$$

where ι varies from one to eight.[35]

Tables Va and Vb summarize the results for all values of ι. To focus discussion, consider the statistics and estimated coefficients for ι=5:

$$H_0: \quad \tilde{C}_t = \underset{(10.3)}{23.2} + \underset{(.038)}{.862}Y_t + \underset{(.00601)}{.00652}W_t \tag{30}$$

$$\tilde{\sigma}_0^2 = 18.078 \quad t_4 = -32.84 \quad t_6 = 86.41 \quad c_0 = 51.74 \quad c_2 = 50.18$$

$$H_1: \quad \tilde{C}_t = \underset{(2.13)}{2.65} + \underset{(.116)}{.067}Y_t + \underset{(.129)}{.930}C_{t-1} \tag{31}$$

$$\tilde{\sigma}_1^2 = 12.974 \quad t_4 = -7.59 \quad t_6 = -.19 \quad c_0 = 21.42 \quad c_2 = .04$$

where ~ denotes IV estimation and the values in parentheses are the IV estimated standard errors. The estimated coefficients in (30) are very similar to those obtained by ML, but those in (31) are not: that for Y_t is no longer significantly different from zero, consistent with simultaneity bias in the ML estimates, and that for C_{t-1} is now essentially unity.

The values of t_4, t_6, c_0, and c_2 in (30) all point to the mis-specification of H_0. c_2 indicates that C_{t-1} is significant if added to (30). c_0 shows that the instruments used are not independent of the residuals, most likely because inter alia C_{t-1} is an instrument and is not included in the specification of H_0. t_4 and t_6 are so significant because (31) markedly variance-dominates (30) (and that, because c_2 is significant).

With H_1 as the null hypothesis, only c_0 and t_4 appear significant: Sargan's statistic indicates that some of the instrumental variables are not valid (i.e., they are correlated with the residuals), suggesting that H_1 incorrectly omits certain lagged values of C and Y. The statistic t_4 also appears to detect that (although t_6 does not), possibly because H_0 excludes all lagged values of Y and C. The results for ι=5 are typical of all values of ι except ι=1, in which case none of the tests detect mis-specification.

[35] The behavior of the IV statistics is determined both by the number of instruments (here, $2\iota+2$) and by the lags at which variables appear in the instrument set: this analysis makes no attempt to separate those effects.

Table Va.

Values of the statistics for H_0: $C_t = \beta_{00} + \beta_{01}Y_t + \beta_{02}W_t + u_{0t}$

Estimation method[a]		t_4	t_6 $[D_0]$	c_0	c_2 $(=f_2)$
ML[b]		2.1	36.5 [-47.1]	29.6	46.8
IV[c]	$\iota=1$	3.3	157.7	46.5	45.7
	$\iota=2$	-13.8	107.1	46.9	50.1
	$\iota=3$	-23.4	102.6	49.3	49.9
	$\iota=4$	-24.1	101.7	51.6	51.1
	$\iota=5$	-32.8	86.4	51.7	50.2
	$\iota=6$	-35.3	82.7	51.9	49.7
	$\iota=7$	-40.8	81.9	53.6	49.3
	$\iota=8$	-56.4	80.5	58.1	47.9

Notes: See the notes for Table Vb.

Table Vb.

Values of the statistics for H_1: $C_t = \beta_{10} + \beta_{11}Y_t + \beta_{12}C_{t-1} + u_{1t}$

Estimation method[a]		t_4	t_6 $[D_0]$	c_0	c_2 $(=f_2)$
ML[b]		-.40	-.37 [.37]	.16	.16
IV[c]	$\iota=1$	-.03	-.03	.00	.00
	$\iota=2$	-3.22	-.13	8.08	.02
	$\iota=3$	-5.11	-.14	13.19	.02
	$\iota=4$	-5.31	-.14	14.07	.02
	$\iota=5$	-7.59	-.19	21.42	.04
	$\iota=6$	-8.26	-.21	23.66	.05
	$\iota=7$	-9.45	-.21	27.11	.05
	$\iota=8$	-12.77	-.22	36.38	.06

Notes: a. Under the null hypothesis (H_0 for Table Va, H_1 for Table Vb), the statistics t_4, t_6, and D_0 are asymptotically distributed as $N(0,1)$; c_0 is asymptotically distributed as $\chi^2(2\iota-1)$ for IV ($\chi^2(1)$ for ML); and c_2 is asymptotically distributed as $\chi^2(1)$.

b. The instruments for ML are $\{1, W_t, C_{t-1}, Y_t\}$.

c. The instruments for IV are $\{1, W_t, ((C_{t-i}, Y_{t-i}), i=1,\ldots,\iota)\}$.

These two models illustrate the potential value of both nested and non-nested hypothesis test statistics in practice. The IV statistics point to the possible importance of additional lags on income and consumers' expenditure in the equation for the determinants of consumers' expenditure, thus establishing a basis for the re-specification of that equation.[36] In general, IV statistics complement ML statistics by allowing for situations in which the specification of a complete set of simultaneous dynamic economic relationships is undesirable or impractical, as is typical in many existing econometric models.

7. Concluding Remarks

The finite sample properties of test statistics often deviate markedly from their asymptotic ones. Exact analytical results typically are not available for precisely those situations which are most interesting from a practical standpoint, e.g., dynamic, simultaneous, mis-specified models. This paper presents and implements an approach for obtaining numerical-analytical formulae (response surfaces) which integrate existing analytical knowledge with Monte Carlo (experimental) results. Response surfaces can help summarize and interpret Monte Carlo simulations, and may reasonably approximate the unknown finite sample conditional probability formulae of the statistics evaluating the relationships of interest, for the DGP considered. Cox (1970) provides the basis for assessing the closeness of that approximation and, more generally, for conducting inference about response surfaces. Applying this approach, this paper investigates the effect of dynamics and simultaneity on the finite sample properties of maximum likelihood and instrumental variables statistics for testing both nested and non-nested hypotheses for dynamic one- and two-equation models.

---

[36] However, the results themselves are more elucidative than substantive: both estimated equations and the comprehensive model exhibit considerable residual autocorrelation and parameter non-constancy, so further inferences are dubious.

The results demonstrate the value of asymptotic theory in interpreting finite sample properties and certain limitations for doing so. Response surfaces summarize the Monte Carlo results conveniently and provide simple formulae for obtaining reasonably accurate and computationally inexpensive predictions of finite sample rejection frequencies within a sizable parameter space. Two practical finite sample results arise. First, transforming the χ^2 Wald statistic to its F form eliminates a dominant term of $O(T^{-1})$. In fact, under the null hypothesis the resulting statistic is approximately an F-ratio and is virtually invariant to the degree of dynamics and simultaneity considered; under the alternative it is only moderately affected by those factors. Second, "large-σ" and a small effective sample size strongly affect the finite sample properties of Sargan's (1958) instrumental variables statistic and Ericsson's (1983) Cox-type instrumental variables statistic. Additional analytical results could help in specifying the functional dependence on σ and T and for deriving finite sample adjustments. Re-examination of Pesaran and Deaton's (1978) empirical example illustrates the additional information gained from the instrumental variables statistics. Although Monte Carlo experimentation can not replace analysis, the two can complement each other effectively to provide convenient formulae for interpreting empirical findings.

Appendix A. Simulation evidence: A single-equation model

This appendix describes the experimental design, simulation, and post-simulation analysis of a pilot Monte Carlo study of the Cox and Wald statistics for a dynamic single-equation model with autocorrelated regressors, used to assess the potential value of the asymptotic formulae in Section 2 and of the response surface methodology in Section 4. The model is defined by the restrictions $b = \sigma_{12} = 0$, so (4) and (5) are recursive. The correctly specified hypothesis H_0 is

$$H_0: \quad y_t = cZ_t + dy_{t-1} + u_{0t} \quad . \tag{A1}$$

The following (falsely specified) non-nested alternative is considered:

$$H_1: \quad y_t = h_{12}w_{2t} + dy_{t-1} + u_{1t} \tag{A2}$$

with $\gamma \equiv \mathrm{corr}(Z_t, w_{2t}) \neq 0$, so the comprehensive hypothesis H_2 is

$$H_2: \quad y_t = cZ_t + h_{12}w_{2t} + dy_{t-1} + u_{2t} \quad . \tag{A3}$$

The pair of non-nested hypotheses above is similar to that in Pesaran (1974) in that both have competing sets of exogenous variables. However, his DGP has no dynamics, whereas dynamics enter (A1) both directly ($d \neq 0$) and indirectly ($\rho_i \neq 0$ for $i = 1, 2$).

A.1. Experimental design

The Monte Carlo design variables of this study are

$$(\theta', T) \;=\; (c, d, \sigma_{11}, \rho_1, \rho_2, \omega_{11}, \omega_{22}, \gamma, T) \tag{A4}$$

where ω_{12} is chosen to give selected values of γ. Three parameters are normalized without loss of generality: $\sigma_{11} = 1$ and $\omega_{11} = \omega_{22} = 1/12$. Rather arbitrarily, $\rho_1 = \rho_2$ and $N = 1000$. The remaining parameters span ranges similar to those in Hendry and Harrison (1974, Section 6.1): $c = (1., 4.)$, $d = (.2, .7)$, $\rho_2 = (.3, .9)$, $\gamma = (.8, .9, .95)$, and $T = (20, 50, 80)$ with a full factorial design of seventy-two experiments.

A.2. Simulation

Given this study's exploratory nature and in order to minimize computational expenses, the exogenous variables Z_t and w_{2t} are non-stochastic (i.e., constant across replications, but <u>not</u> across experiments), and the v_{it} are uniform. Estimation is by OLS and only the Cox statistic and the χ^2 Wald statistic are evaluated. All other computational aspects are as in Section 5.

A.3. <u>Post-simulation analysis</u>

Cursory examination of the Monte Carlo results reveals that the Cox test generally rejects H_1 more frequently than the Wald test, parallelling Pesaran's (1974, 1982) Monte Carlo results for static models. It is unclear which test is more <u>powerful</u> because the estimated sizes of both tests are almost always larger than the asymptotic 5% level, with the Cox test the worse of the two. To evaluate powers as such, some control of the size would be necessary, e.g., by estimating the finite sample size by an order statistic based on the test statistic values from a Monte Carlo experiment when the assumed null hypotheses is true, or by the techniques discussed in Mehta (1979). Even without such adjustments, the simulation results are worthwhile evaluating because they use critical values which an applied econometrician typically would employ. In estimating response surfaces, all experiments for which the asymptotic power or the rejection frequency of H_1 is greater than .998 for either test, are excluded. Of the remaining fifty-two experiments, six randomly selected experiments are retained for the Chow test. In the spirit of Mizon and Hendry (1980), Type B response surfaces are estimated with ℓ, ℓ/T, λ_3/T, T^{-0}, $T^{-1/2}$, and T^{-1} as regressors for powers; ℓ, $T^{-1/2}$, T^{-1}, $dT^{-1/2}$, and dT^{-1} as regressors for sizes.[37] Parsimonious representations of those more general response surfaces appear in (A5)-(A8) with a selection of evaluation criteria.

[37] δ, the order of the second approximation for the asymptotic distribution of D_1, is also included in the unrestricted response surface for D_1, but proved insignificant. The data are ordered by sample size, increasing in λ_3 within each group. Detailed Monte Carlo results for Appendix A and Section 5 are available from the author upon request.

The restricted response surfaces for the type I error of the Wald (c_2) and Cox (D_0) statistics are

$$L^*(\tilde{s})-L(.05) \quad = \quad \underset{\substack{(.16)\\ [.20]}}{1.92}A^{1/2}T^{-1/2} \tag{A5}$$

$R^2 = .43 \quad \tilde{\sigma}_e = 1.659 \quad \eta_1(6,45) = .62 \quad \xi_2(51) = 140.4 \quad dw = 1.95$

$\eta_3(4,47) = 2.05 \quad \eta_7(1,49) = 10.23$

$$L^*(\tilde{s})-L(.05) \quad = \quad \underset{\substack{(.20)\\ [.24]}}{3.16}A^{1/2}T^{-1/2} \tag{A6}$$

$R^2 = .57 \quad \tilde{\sigma}_e = 2.240 \quad \eta_1(6,45) = .84 \quad \xi_2(51) = 255.9 \quad dw = 1.06$

$\eta_3(4,47) = .93 \quad \eta_7(1,49) = 11.58$

respectively. The response surfaces for the finite sample powers of the Wald (c_3) and Cox (D_1) statistics are

$$L^*(\tilde{s}) \quad = \quad \underset{\substack{(.11)\\ [.08]}}{.96}L(\pi_a) - \underset{\substack{(3.01)\\ [2.22]}}{6.65}L(\pi_a)/T - \underset{\substack{(.48)\\ [.51]}}{1.81}A^{1/2}(\lambda_3/T) \tag{A7}$$

$R^2 = .78 \quad \tilde{\sigma}_e = 7.490 \quad \eta_1(6,43) = 2.03 \quad \xi_2(49) = 2749.1 \quad dw = 2.11$

$\eta_3(3,46) = 2.75 \quad \eta_7(6,42) = 2.34$

$$L^*(\tilde{s}) \quad = \quad \underset{\substack{(.15)\\ [.11]}}{1.08}L(\pi_a) - \underset{\substack{(3.82)\\ [3.83]}}{11.68}L(\pi_a)/T \tag{A8}$$

$R^2 = .70 \quad \tilde{\sigma}_e = 8.069 \quad \eta_1(6,44) = 1.90 \quad \xi_2(50) = 3255.4 \quad dw = 2.17$

$\eta_3(5,45) = .41 \quad \eta_7(3,46) = 2.74$.

These response surfaces are in general agreement with the theory discussed earlier: the coefficients of L(.05) and $L(\pi_a)$ are statistically insignificantly different from unity, and $\tilde{\sigma}_e$ is close to unity for the response surfaces of type I errors. The coefficient on $T^{-1/2}$ in (A6) (versus that in (A5)) captures the larger positive finite sample bias in size for the Cox statistic. Also, the coefficients for ℓ/T and λ_3/T are negative in (A7), in line with Mizon and Hendry's (1980, pp. 35, 42) response surfaces for the Wald test of a common factor. However, although R^2 is large in (A7) and (A8), the corresponding values of $\tilde{\sigma}_e$ and η_7 suggest that much more variation present in $L^*(s)$ could be explained by additional terms in those response surfaces. Re-using random numbers across experiments could aid in estimating the response surface coefficients more precisely and in reducing the values of $\tilde{\sigma}_e$, but it also might create spurious correlation between experiments; cf. Mizon and Hendry (1980, pp. 34-37 and footnote 4) and Hendry (1984, p. 971). The lack of control over the range of π_T (or even π_a) and the resulting loss of one quarter of the experiments motivate the experimental design in Section 5.

The asymptotic formulae in Section 2 explain many features of the Monte Carlo simulations across experiments, with response surfaces providing a concise, useful method of analyzing the relationship between observed fluctuations, asymptotic approximations, and finite sample effects.

Appendix B. Hendry and Harrison's (1974) model

This appendix briefly describes Hendry and Harrison's (1974) model and its relationship to the model used in this paper. Their model can be expressed as

$$By_t + Cz_t + Dy_{t-1} = u_t \qquad u_t = Ru_{t-1} + \epsilon_t$$
$$z_t - \Lambda z_{t-1} = v_t \qquad (t=1,\ldots,T) \tag{B1}$$

where y_t and z_t are 2×1 and 4×1 vectors of endogenous and exogenous variables at time t; the normalizations are $b_{11} = b_{22} = -1.$; B, C, D, R, and Λ are matrices of dimension 2×2, 2×4, 2×2, 2×2, and 4×4, respectively, and all the latent roots of $B^{-1}D$, R, and Λ lie within the unit circle; $\epsilon_t \sim NID(0,\Sigma)$ and $v_t \sim IID(0,\Omega)$; and $E(\epsilon_t v_s') = 0$ for all t and s. Hendry and Harrison (1974, pp. 153-154) restrict C, D, and R such that the first line in (B1) may be written extensively as

$$y_{1t} = b_{12}y_{2t} + c_{11}z_{1t} + d_{11}y_{1,t-1} + u_{1t} \qquad u_{1t} = r_{11}u_{1,t-1} + \epsilon_{1t}$$
$$y_{2t} = b_{21}y_{1t} + \sum_{j=1}^{4} c_{2j}z_{jt} \qquad + \epsilon_{2t} \tag{B2}$$

with $c_{11}c_{21} = 0$ and with an implicit change in sign of the disturbances. Rewriting (B2) in a simpler notation (and with y_t and ϵ_t now denoting scalars),

$$y_t = bY_t + cZ_t + dy_{t-1} + u_t \qquad u_t = ru_{t-1} + \epsilon_t \tag{B3}$$
$$Y_t = ay_t + f'w_t \qquad + e_t \tag{B4}$$

where $f' = (c_{21}:c_{22}:c_{23}:c_{24})$ and w_t in (B4) equals z_t in (B1). Thus,

$$w_t = \Lambda w_{t-1} + v_t \quad . \tag{B5}$$

Equations (4)-(6) in Section 3 are equivalent to (B3)-(B5) with $r = 0$ and correspond to equations (2)-(5) in Hendry and Harrison (1974), but with some slight changes in notation: ρ_i (rather than λ_i) denotes the i^{th} diagonal element of Λ; h' and h_{ij} are their f' and c_{ij}; and the disturbances on the structural equations are u_t and ν_t rather than ϵ_t and e_t. Those changes are made to avoid confusion with other notation in this paper.

Hendry and Harrison's model allows for autoregressive errors on the structural equations whereas (4)-(5) do not. At first blush, including autoregressive disturbances ($r \neq 0$) might appear an interesting extension. Pesaran (1974, pp. 164-169) derives the Cox statistic for models with fixed regressors and first-order autoregressive errors; and Sargan's statistic c_0 generalizes to allow for models nonlinear in their parameters (and in particular for linear models with autoregressive errors, see Sargan (1959, pp. 101-105; 1964, pp. 25-29) and Campos (1986b)), from which an IV statistic for testing non-nested nonlinear hypotheses could be constructed. However, autoregressive errors can be regarded as arising from a common factor restriction on a more general dynamic model with serially uncorrelated errors, a restriction which is often invalid in empirical studies.[38] Hence, the restriction $r = 0$ is imposed in both Monte Carlo studies, but with $d \neq 0$. Note, though, that for that particular dynamic specification, (4) does not include a first-order autoregressive error as a special case except trivially so when $b = c = 0$.

---

[38] For both theoretical and empirical discussions of the common factor restriction, see Sargan (1959, pp. 91-92, 101-105), Durbin (1960a, pp. 150-153; 1960b, pp. 235-238), Sargan (1964, pp. 27, 39-41), Hendry (1974), Hendry and Mizon (1978), Mizon and Hendry (1980), Sargan (1980d), and Hendry, Pagan and Sargan (1984, pp. 1045-1047, 1078-1080).

Appendix C. Type B Response Surfaces

The dependent variable in each response surface is $L^{*}(s)-L(\pi_{*})$ where π_{*} is .05 for f_2 and t_6, π_a for f_3 and t_7, and π_x for c_2, c_3, c_0, and c_1. All right-hand side variables represent effects present only in finite samples.

Table C.I.
Estimates of Type B Response Surfaces for Finite Sample Sizes and Powers

Regressors and Diagnostic Statistics	Test Statistic τ			
	f_2	c_2	f_3	c_3
T	11.8 (3.2) [3.1]	9.7 (3.0) [2.9]	8.2 (2.8) [2.4]	6.1 (2.9) [2.6]
T^2	-130. (32.) [32.]	-106. (30.) [29.]	-77. (26.) [21.]	-54. (28.) [24.]
T^3	328. (81.) [79.]	268. (76.) [71.]	140. (63.) [50.]	89. (68.) [58.]
$T\sigma$	-2.10 (.48) [.38]	-1.76 (.46) [.35]	-2.44 (.35) [.42]	-2.17 (.36) [.38]
$T^2\sigma$	16.3 (5.0) [3.7]	12.5 (4.7) [3.4]	6.9 (1.6) [1.6]	4.8 (1.6) [1.6]
$T^3\sigma$	-36. (13.) [9.]	-27. (12.) [9.]		
$T\sigma^2$	.0197 (.0046) [.0044]	.0200 (.0044) [.0040]	.0400 (.0192) [.0166]	.0378 (.0204) [.0179]
$T^2\sigma^2$			.59 (.13) [.10]	.59 (.14) [.12]
$T\rho_0$	-17.2 (5.3) [3.5]	-13.7 (5.0) [3.2]	-26.5 (6.1) [4.7]	-24.0 (6.6) [5.5]
$T^2\rho_0$	185. (50.) [34.]	148. (46.) [30.]	184. (54.) [41.]	153. (58.) [48.]

Table C.I. (cont.)
Estimates of Type B Response Surfaces for Finite Sample Sizes and Powers

Regressors and Diagnostic Statistics	Test Statistic τ			
	f_2	c_2	f_3	c_3
$T^3\rho_0$	-443. (118.) [86.]	-352. (109.) [75.]	-366. (121.) [91.]	-283. (129.) [107.]
$T\rho_0\sigma$	2.15 (.66) [.45]	1.60 (.62) [.44]	5.32 (.81) [.68]	4.74 (.81) [.68]
$T^2\rho_0\sigma$	-21.3 (6.9) [4.5]	-15.6 (6.5) [4.5]	-29.8 (5.7) [4.6]	-24.1 (5.6) [4.6]
$T^3\rho_0\sigma$	48.4 (17.4) [11.4]	34.7 (16.3) [11.2]	44.3 (10.7) [8.3]	33.8 (10.7) [8.8]
$T\rho_0\sigma^2$			-.127 (.017) [.017]	-.122 (.017) [.016]
$T\ell$			-3.33 (.49) [.52]	-3.30 (.51) [.53]
$T^2\ell$			7.74 (1.97) [2.02]	6.61 (2.06) [2.10]
$T\ell\sigma$			.542 (.063) [.085]	.543 (.066) [.082]
$T^3\ell\sigma^2$			.530 (.092) [.067]	.559 (.108) [.090)
$T\rho_0\ell\sigma$			-.504 (.077) [.093]	-.485 (.082) [.092]

Table C.I. (concl.)
Estimates of Type B Response Surfaces for Finite Sample Sizes and Powers

Regressors and Diagnostic Statistics	Test Statistic τ			
	f_2	c_2	f_3	c_3
R^2	.850	.866	.985	.985
$\tilde{\sigma}_e$	1.080	1.092	2.069	2.256
n	13	13	19	19
Chow $\eta_1(9,63\text{-}n)$	.85	.93	.88	1.04
RSS $\xi_2(72\text{-}n)$	68.8	70.4	226.9	269.8
Parsimony $\eta_3(q,72\text{-}n\text{-}q)$	.50 {5}	1.23 {6}	.40 {14}	.88 {14}
Functional form $\eta_3(q,71\text{-}n\text{-}q)$	.46 {29}	.39 {29}	1.26 {28}	.75 {30}
RESET $\eta_4(4,68\text{-}n)$	2.92	3.32	1.01	.90
Unit coefficient $\hat{\psi}$	1.02 (.37)	1.36 (.34)	1.11 (.22)	1.13 (.23)
Unit coefficient $\eta_5(1,71\text{-}n)$	.00	1.10	.26	.35
Normality $\xi_6(2)$	.24	.60	7.23	3.82
Heteroscedasticity $\eta_7(q,71\text{-}n\text{-}q)$	.46 {26}	.45 {26}	.48 {38}	.49 {38}
ARCH $\eta_8(12,59\text{-}n)$	.59	.60	.15	.27
dw	2.31	2.25	1.86	1.86
AR residuals $\eta_9(12,60\text{-}n)$	.87	1.02	.81	.85

Note: The value of the degrees of freedom q appears in curly brackets {•}.

Table C.II.
Estimates of Type B Response Surfaces for Finite Sample Sizes and Powers

Regressors and Diagnostic Statistics	Test Statistic τ		Regressors and Diagnostic Statistics	Test Statistic τ	
	c_0	c_1		t_6	t_7
T	-.57 (.13) [.10]		$T^{-1/2}$	-38.5 (5.0) [5.4]	-109. (11.) [10.]
T^2		-17.90 (.84) [.92]	T	35.9 (4.7) [5.0]	113. (11.) [11.]
$T\sigma$	1.39 (.04) [.04]	-.72 (.33) [.35]	$TT^{-1/2}$		-28.9 (5.6) [5.5]
$T^2\sigma$	-.50 (.18) [.15]	15.9 (2.7) [2.8]	$T^{-1/2}\sigma$	17.2 (2.3) [2.7]	6.3 (1.4) [1.1]
$T^3\sigma$		-38.7 (6.1) [6.9]	$T\sigma$	-14.1 (2.1) [2.5]	-6.1 (1.2) [1.0]
$T\sigma^2$		.207 (.026) [.028]	$T^{-1/2}\sigma^2$	-.90 (.15) [.19]	
$T^2\sigma^2$	-.153 (.016) [.016]	-.080 (.126) [.114]	$T\sigma^2$	.82 (.14) [.17]	
$T\rho_0$	1.85 (.70) [.60]		$TT^{-1/2}\sigma^2$	-.520 (.055) [.072]	
$T^2\rho_0$	-3.01 (2.17) [1.92]	-45.0 (8.3) [7.5]	$T^{-1/2}\rho_0$	44.5 (8.5) [8.7]	146. (26.) [25.]
$T^3\rho_0$		128. (33.) [30.]	$T\rho_0$	-44.4 (8.2) [9.1]	-167. (25.) (25.)
$T\rho_0\sigma$	-.155 (.040) [.037]		$TT^{-1/2}\rho_0$	25.3 (12.5) [15.2]	137. (36.) [33.]
$T\rho_0\sigma^2$		-.077 (.027) [.030]	$T^{-1/2}\rho_0\sigma$	-17.5 (3.0) [3.5]	1.53 (.52) [.44]
$T^2\rho_0\sigma^2$		.63 (.15) [.15]	$T\rho_0\sigma$	17.0 (2.8) [3.3]	
$T\ell$		-2.00 (.13) [.15]	$TT^{-1/2}\rho_0\sigma$	-7.6 (2.5) [2.6]	

Table C.II. (cont.)
Estimates of Type B Response Surfaces for Finite Sample Sizes and Powers

Regressors and Diagnostic Statistics	Test Statistic τ: c_0	c_1	Regressors and Diagnostic Statistics	Test Statistic τ: t_6	t_7
$T\ell\sigma$		.34 (.10) [.12]	$T^{-1/2}\rho_0\sigma^2$	.96 (.19) [.23]	
$T\rho_0\ell$		-13.8 (2.6) [2.4]	$T\rho_0\sigma^2$	-.98 (.18) [.21]	
$T^2\rho_0\ell$		95. (18.) [16.]	$TT^{-1/2}\rho_0\sigma^2$	.87 (.17) [.18]	
$T^3\rho_0\ell$		-166. (32.) [26.]	$T^{-1/2}\ell$		30.3 (4.8) [3.8]
$T\rho_0\ell\sigma$		.46 (.11) [.11]	$T\ell$		-30.8 (4.6) [3.7]
$T\ell\sigma^2$		.081 (.014) [.015]	$T^{-1/2}\ell\sigma$		-7.38 (.93) [.74]
			$T\ell\sigma$		7.35 (.91) [.75]
			$TT^{-1/2}\ell\sigma$		-4.21 (.48) [.39]
			$T^{-1/2}\rho_0\ell$		-36.1 (11.3) [9.2]
			$T\rho_0\ell$		45.9 (11.0) [9.0]
			$TT^{-1/2}\rho_0\ell$		-53.1 (16.8) [14.6]
			$T\rho_0\ell\sigma$		-.88 (.26) [.22]

Table C.II. (concl.)

Estimates of Type B Response Surfaces for Finite Sample Sizes and Powers

Regressors and Diagnostic Statistics	Test Statistic τ		Regressors and Diagnostic Statistics	Test Statistic τ	
	c_0	c_1		t_6	t_7
R^2	.9958	.9947	R^2	.9923	.9777
$\tilde{\sigma}_e$	1.489	2.097	$\tilde{\sigma}_e$	1.703	3.056
n	7	17	n	16	18
Chow $\eta_1(9,63-n)$	.90	.81	Chow $\eta_1(9,63-n)$	.82	.66
RSS $\xi_2(72-n)$	144.0	241.9	RSS $\xi_2(72-n)$	162.4	504.3
Parsimony $\eta_3(q,72-n-q)$	.73 {11}	.94 {16}	Parsimony $\eta_3(q,72-n-q)$	.28 {2}	.60 {6}
Functional form $\eta_3(q,71-n-q)$	.83 {33}	.90 {31}	Functional form $\eta_3(q,71-n-q)$	1.27 {30}	.41 {33}
RESET $\eta_4(4,68-n)$	.34	3.91	RESET $\eta_4(4,68-n)$	7.74	4.84
Unit coefficient $\hat{\phi}$	.97 (.03)	.93 (.07)	Unit coefficient $\hat{\phi}$	1.02 (.04)	1.09 (.19)
Unit coefficient $\eta_5(1,71-n)$	1.26	1.11	Unit coefficient $\eta_5(1,71-n)$	.19	.22
Normality $\xi_6(2)$	1.00	1.34	Normality $\xi_6(2)$	1.97	1.42
Heteroscedasticity $\eta_7(q,71-n-q)$	1.08 {14}	.45 {34}	Heteroscedasticity $\eta_7(q,71-n-q)$	2.59 {31}	.54 {35}
ARCH $\eta_8(12,59-n)$	.37	.57	ARCH $\eta_8(12,59-n)$	.32	.39
dw	2.19	2.39	dw	2.13	1.84
AR residuals $\eta_9(12,60-n)$	2.24	1.23	AR residuals $\eta_9(12,60-n)$	.59	1.33

Note: The value of the degrees of freedom q appears in curly brackets {•}.

References

Anderson, T.W. (1977) "Asymptotic Expansions of the Distributions of Estimates in Simultaneous Equations for Alternative Parameter Sequences", Econometrica, 45, 2, 509-518.

Atkinson, A.C. (1970) "A Method for Discriminating Between Models", Journal of the Royal Statistical Society, Series B, 32, 3, 323-353 (with discussion).

Box, G.E.P. and M.E. Muller (1958) "A Note on the Generation of Random Normal Deviates", Annals of Mathematical Statistics, 29, 2, 610-611.

Box, G.E.P. and D.A. Pierce (1970) "Distribution of Residual Autocorrelations in Autoregressive-integrated Moving Average Time Series Models", Journal of the American Statistical Association, 65, 332, 1509-1526.

Campos, J. (1986a) "Finite-sample Properties of the Instrumental-variables Estimator for Dynamic Simultaneous-equation Subsystems with ARMA Disturbances", Journal of Econometrics, 32, 3, 333-366.

Campos, J. (1986b) "Instrumental Variables Estimation of Dynamic Simultaneous Systems with ARMA Errors", Review of Economic Studies, 53, 1, 125-138.

Carrier, N., E. Atkins, and C. Taylor (1969) "Report on the London Atlas Random Number Generator", mimeo, London, University of London Computing Centre.

Chow, G.C. (1960) "Tests of Equality Between Sets of Coefficients in Two Linear Regressions", Econometrica, 28, 3, 591-605.

Cochran, W.G. and G.M. Cox (1957) Experimental Designs, New York, John Wiley and Sons, Second Edition.

Cox, D.R. (1958) Planning of Experiments, New York, John Wiley and Sons.

Cox, D.R. (1961) "Tests of Separate Families of Hypotheses" in J. Neyman (ed.) Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, Volume I, 105-123.

Cox, D.R. (1962) "Further Results on Tests of Separate Families of Hypotheses", Journal of the Royal Statistical Society, Series B, 24, 2, 406-424.

Cox, D.R. (1970) The Analysis of Binary Data, London, Methuen.

Cox, D.R. and D.V. Hinkley (1974) Theoretical Statistics, London, Chapman and Hall.

Davidson, J.E.H. and D.F. Hendry (1981) "Interpreting Econometric Evidence: The Behaviour of Consumers' Expenditure in the UK", European Economic Review, 16, 1, 177-192.

Davidson, J.E.H., D.F. Hendry, F. Srba, and S. Yeo (1978) "Econometric Modelling of the Aggregate Time-series Relationship Between Consumers'

Expenditure and Income in the United Kingdom", Economic Journal, 88, 352, 661-692.

Dhrymes, P.J., E.P. Howrey, S.H. Hymans, J. Kmenta, E.E. Leamer, R.E. Quandt, J.B. Ramsey, H.T. Shapiro, and V. Zarnowitz (1972) "Criteria for Evaluation of Econometric Models", Annals of Economic and Social Measurement, 1, 3, 291-324.

Durbin, J. (1960a) "Estimation of Parameters in Time-series Regression Models", Journal of the Royal Statistical Society, Series B, 22, 1, 139-153.

Durbin, J. (1960b) "The Fitting of Time-series Models", Revue de l'Institut International de Statistique, 28, 3, 233-244.

Durbin, J. and G.S. Watson (1950) "Testing for Serial Correlation in Least Squares Regression. I", Biometrika, 37, 3 and 4, 409-428.

Durbin, J. and G.S. Watson (1951) "Testing for Serial Correlation in Least Squares Regression. II", Biometrika, 38, 1 and 2, 159-178.

Engle, R.F. (1982) "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", Econometrica, 50, 4, 987-1007.

Ericsson, N.R. (1983) "Asymptotic Properties of Instrumental Variables Statistics for Testing Non-nested Hypotheses", Review of Economic Studies, 50, 2, 287-304.

Ericsson, N.R. (1986) "Post-simulation Analysis of Monte Carlo Experiments: Interpreting Pesaran's (1974) Study of Non-nested Hypothesis Test Statistics", Review of Economic Studies, 53, 4, 691-707.

Farebrother, R.W. (1980) "The Durbin-Watson Test for Serial Correlation When There Is No Intercept in the Regression", Econometrica, 48, 6, 1553-1563; and (1981) "Erratum", Econometrica, 49, 1, 277.

Fisher, R.A. (1922) "The Goodness of Fit of Regression Formulae, and the Distribution of Regression Coefficients", Journal of the Royal Statistical Society, 85, 4, 597-612.

Godfrey, L.G. (1978) "Testing Against General Autoregressive and Moving Average Error Models When the Regressors Include Lagged Dependent Variables", Econometrica, 46, 6, 1293-1301.

Godfrey, L.G. (1983) "Testing Non-nested Models After Estimation by Instrumental Variables or Least Squares", Econometrica, 51, 2, 355-365.

Godfrey, L.G. and M.H. Pesaran (1983) "Tests of Non-nested Regression Models: Small Sample Adjustments and Monte Carlo Evidence", Journal of Econometrics, 21, 1, 133-154.

Granger, C.W.J. and P. Newbold (1974) "Spurious Regressions in Econometrics", Journal of Econometrics, 2, 2, 111-120.

Haavelmo, T. (1943) "The Statistical Implications of a System of Simultaneous Equations", Econometrica, 11, 1, 1-12.

Haavelmo, T. (1944) "The Probability Approach in Econometrics", Econometrica, 12, supplement, i-viii, 1-118.

Hammersley, J.M. and D.C. Handscomb (1964) Monte Carlo Methods, London, Chapman and Hall.

Harvey, A.C. (1981) The Econometric Analysis of Time Series, Oxford, Philip Allan.

Hausman, J.A. (1975) "An Instrumental Variable Approach to Full Information Estimators for Linear and Certain Nonlinear Econometric Models", Econometrica, 43, 4, 727-738.

Hendry, D.F. (1973) "On Asymptotic Theory and Finite Sample Experiments", Economica, 40, 158, 210-217.

Hendry, D.F. (1974) "Stochastic Specification in an Aggregate Demand Model of the United Kingdom", Econometrica, 42, 3, 559-578.

Hendry, D.F. (1976) "The Structure of Simultaneous Equations Estimators", Journal of Econometrics, 4, 1, 51-88.

Hendry, D.F. (1979a) "The Behaviour of Inconsistent Instrumental Variables Estimators in Dynamic Systems with Autocorrelated Errors", Journal of Econometrics, 9, 3, 295-314.

Hendry, D.F. (1979b) "Predictive Failure and Econometric Modelling in Macroeconomics: The Transactions Demand for Money" in P. Ormerod (ed.) Economic Modelling, London, Heinemann Education Books, 217-242.

Hendry, D.F. (1980) "Econometrics -- Alchemy or Science?", Economica, 47, 188, 387-406.

Hendry, D.F. (1982) "A Reply to Professors Maasoumi and Phillips", Journal of Econometrics, 19, 2/3, 203-213.

Hendry, D.F. (1983) "Comment", Econometric Reviews, 2, 1, 111-114.

Hendry, D.F. (1984) "Monte Carlo Experimentation in Econometrics", Chapter 16 in Z. Griliches and M.D. Intriligator (eds.) Handbook of Econometrics, Amsterdam, North-Holland, Volume II, 937-976.

Hendry, D.F. (1987) PC-GIVE: An Interactive Menu-driven Econometric Modelling Program for IBM-Compatible PC's, Version 5.0, Oxford, University of Oxford, Institute of Economics and Statistics and Nuffield College.

Hendry, D.F. and R.W. Harrison (1974) "Monte Carlo Methodology and the Small Sample Behaviour of Ordinary and Two-stage Least Squares", Journal of Econometrics, 2, 2, 151-174.

Hendry, D.F. and G.E. Mizon (1978) "Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England", Economic Journal, 88, 351, 549-563.

Hendry, D.F. and A.J. Neale (1987) "Monte Carlo Experimentation Using PC-NAIVE", Applied Economics Discussion Paper No. 20, Institute of Economics and Statistics, University of Oxford, Oxford, forthcoming in T. Fomby (ed.) Advances in Econometrics.

Hendry, D.F., A.R. Pagan, and J.D. Sargan (1984) "Dynamic Specification", Chapter 18 in Z. Griliches and M.D. Intriligator (eds.) Handbook of Econometrics, Amsterdam, North-Holland, Volume II, 1023-1100.

Hendry, D.F. and J.-F. Richard (1982) "On the Formulation of Empirical Models in Dynamic Econometrics", Journal of Econometrics, 20, 1, 3-33.

Hendry, D.F. and F. Srba (1977a) "The Properties of Autoregressive Instrumental Variables Estimators in Dynamic Systems", Econometrica, 45, 4, 969-990.

Hendry, D.F. and F. Srba (1977b) "Users' Manual for DAGER: Data Generation and Moment Calculation from Dynamic Econometric Systems", mimeo, London, London School of Economics, September.

Hendry, D.F. and F. Srba (1978) "Technical Manual for GIVE: General Instrumental Variables Estimation of Linear Equations with Lagged Dependent Variables and First Order Autoregressive Errors", mimeo, London, London School of Economics, October.

Hendry, D.F. and F. Srba (1979) "Technical Manual for NAIVE: Monte Carlo Investigation of GIVE Estimators (Numerical Analysis of Instrumental Variables Estimators)", mimeo, London, London School of Economics, January.

Hendry, D.F. and F. Srba (1980) "AUTOREG: A Computer Program Library for Dynamic Econometric Models with Autoregressive Errors", Journal of Econometrics, 12, 1, 85-102.

Jackson, O.A.Y. (1968) "Some Results on Tests of Separate Families of Hypotheses", Biometrika, 55, 2, 355-363.

Jarque, C.M. and A.K. Bera (1980) "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals", Economics Letters, 6, 3, 255-259.

Johnson, N.L. (1959) "On an Extension of the Connexion Between Poisson and χ^2 Distributions", Biometrika, 46, 3 and 4, 352-363.

Johnston, J. (1963) Econometric Methods, New York, McGraw-Hill.

Kadane, J.B. (1970) "Testing Overidentifying Restrictions When the Disturbances Are Small", Journal of the American Statistical Association, 65, 329, 182-185.

Kadane, J.B. (1971) "Comparison of k-Class Estimators When the Disturbances Are Small", Econometrica, 39, 5, 723-737.

Kendall, M.G. and A. Stuart (1973) The Advanced Theory of Statistics, Volume 2, Inference and Relationship, London, Griffin, Third Edition.

Kennedy, W.J., Jr., and J.E. Gentle (1980) Statistical Computing, New York, Marcel Dekker.

King, M.L. and M. McAleer (1987) "Further Results on Testing AR(1) Against MA(1) Disturbances in the Linear Regression Model", Review of Economic Studies, 54, 4, 649-663.

Kiviet, J.F. (1985) "Model Selection Test Procedures in a Single Linear Equation of a Dynamic Simultaneous System and Their Defects in Small Samples", Journal of Econometrics, 28, 3, 327-362.

Kiviet, J.F. (1986) "On the Rigour of Some Misspecification Tests for Modelling Dynamic Relationships", Review of Economic Studies, 53, 2, 241-261.

Kiviet, J.F. (1987) Testing Linear Econometric Models, Amsterdam, University of Amsterdam.

Klein, L.R. (1969) "Estimation of Interdependent Systems in Macroeconometrics", Econometrica, 37, 2, 171-192.

Koopmans, T.C., H. Rubin, and R.B. Leipnik (1950) "Measuring the Equation Systems of Dynamic Economics", Chapter 2 in T.C. Koopmans (ed.) Statistical Inference in Dynamic Economic Models, Cowles Commission Monograph No. 10, New York, John Wiley and Sons, 53-237.

Maasoumi, E. and P.C.B. Phillips (1982) "On the Behavior of Inconsistent Instrumental Variable Estimators", Journal of Econometrics, 19, 2/3, 183-201.

MacKinnon, J.G. (1983) "Model Specification Tests Against Non-nested Alternatives", Econometric Reviews, 2, 1, 85-110.

MacKinnon, J.G., H. White, and R. Davidson (1983) "Tests for Model Specification in the Presence of Alternative Hypotheses: Some Further Results", Journal of Econometrics, 21, 1, 53-70.

Mann, H.B. and A. Wald (1943a) "On Stochastic Limit and Order Relationships", Annals of Mathematical Statistics, 14, 3, 217-226.

Mann, H.B. and A. Wald (1943b) "On the Statistical Treatment of Linear Stochastic Difference Equations", Econometrica, 11, 3-4, 173-220.

Mehta, F. (1979) "A Monte Carlo Study of Tests of Dynamic Specification", mimeo, London, London School of Economics.

Mizon, G.E. (1984) "The Encompassing Approach in Econometrics", Chapter 6 in D.F. Hendry and K.F. Wallis (eds.) Econometrics and Quantitative Economics, Oxford, Basil Blackwell, 135-172.

Mizon, G.E. and D.F. Hendry (1980) "An Empirical Application and Monte Carlo Analysis of Tests of Dynamic Specification", Review of Economic Studies, 47, 1, 21-45.

Mizon, G.E. and J.-F. Richard (1983) "Comment", Econometric Reviews, 2, 1, 131-136.

Mizon, G.E. and J.-F. Richard (1986) "The Encompassing Principle and Its Application to Testing Non-nested Hypotheses", Econometrica, 54, 3, 657-678.

Nankervis, J.C. and N.E. Savin (1985) "The Student's t Approximation in a Stationary First Order Autoregressive Model", mimeo, Cambridge, Trinity College, December.

Neave, H.R. (1973) "On Using the Box-Muller Transformation with Multiplicative Congruential Pseudo-random Number Generators", Applied Statistics, 22, 1, 92-97.

Nicholls, D.F. and A.R. Pagan (1983) "Heteroscedasticity in Models with Lagged Dependent Variables", Econometrica, 51, 4, 1233-1242.

Numerical Algorithms Group (NAg) (1977) NAg FORTRAN Library Manual, Mark 6 (FLM6), Oxford, Numerical Algorithms Group Ltd.

Patnaik, P.B. (1949) "The Non-central χ^2- and F-distributions and Their Applications", Biometrika, 36, 1 and 2, 202-232.

Pereira, B. de B. (1977) "A Note on the Consistency and on the Finite Sample Comparisons of Some Tests of Separate Families of Hypotheses", Biometrika, 64, 1, 109-113.

Pereira, B. de B. (1978) "Empirical Comparisons of Some Tests of Separate Families of Hypotheses", Metrika, 25, 4, 219-234.

Pesaran, M.H. (1974) "On the General Problem of Model Selection", Review of Economic Studies, 41, 2, 153-171.

Pesaran, M.H. (1982) "Comparison of Local Power of Alternative Tests of Non-nested Regression Models", Econometrica, 50, 5, 1287-1305.

Pesaran, M.H. (1987) "Global and Partial Non-nested Hypotheses and Asymptotic Local Power", Econometric Theory, 3, 1, 69-97.

Pesaran, M.H. and A.S. Deaton (1978) "Testing Non-nested Nonlinear Regression Models", Econometrica, 46, 3, 677-694.

Phillips, P.C.B. (1977) "Approximations to Some Finite Sample Distributions Associated with a First-order Stochastic Difference Equation", Econometrica, 45, 2, 463-485; and (1982) "Erratum", Econometrica, 50, 1, 274.

Phillips, P.C.B. (1980) "Finite Sample Theory and the Distributions of Alternative Estimators of the Marginal Propensity to Consume", Review of Economic Studies, 47, 1, 183-224.

Phillips, P.C.B. (1982) "Best Uniform and Modified Padé Approximants to Probability Densities in Econometrics", Chapter 5 in W. Hildenbrand (ed.) Advances in Econometrics, Cambridge, Cambridge University Press, 123-167.

Phillips, P.C.B. (1986a) "The Exact Distribution of the Wald Statistic", Econometrica, 54, 4, 881-895.

Phillips, P.C.B. (1986b) "Understanding Spurious Regressions in Econometrics", Journal of Econometrics, 33, 3, 311-340.

Poskitt, D.S. and A.R. Tremayne (1981) "A Time Series Application of the Use of Monte Carlo Methods to Compare Statistical Tests", Journal of Time Series Analysis, 2, 4, 263-277.

Ramsey, J.B. (1969) "Tests for Specification Errors in Classical Linear Least-squares Regression Analysis", Journal of the Royal Statistical Society, Series B, 31, 2, 350-371.

Rothery, P. (1982) "The Use of Control Variates in Monte Carlo Estimation of Power", Journal of the Royal Statistical Society, Series C, 31, 2, 125-129.

Sahai, H. (1979) "A Supplement to Sowey's Bibliography on Random Number Generation and Related Topics", Journal of Statistical Computation and Simulation, 10, 1, 31-52.

Sargan, J.D. (1958) "The Estimation of Economic Relationships Using Instrumental Variables", Econometrica, 26, 3, 393-415.

Sargan, J.D. (1959) "The Estimation of Relationships with Autocorrelated Residuals by the Use of Instrumental Variables", Journal of the Royal Statistical Society, Series B, 21, 1, 91-105.

Sargan, J.D. (1964) "Wages and Prices in the United Kingdom: A Study in Econometric Methodology" in P.E. Hart, G. Mills, and J.K. Whitaker (eds.) Econometric Analysis for National Economic Planning, Colston Paper Volume 16, London, Butterworths, 25-54; reprinted as Chapter 10 in D.F. Hendry and K.F. Wallis (eds.) Econometrics and Quantitative Economics, Oxford, Basil Blackwell, 1984, 275-314.

Sargan, J.D. (1976a) "Econometric Estimators and the Edgeworth Approximation", Econometrica, 44, 3, 421-448; and (1977) "Erratum", Econometrica, 45, 1, 272.

Sargan, J.D. (1976b) "Testing for Misspecification after Estimating Using Instrumental Variables", mimeo, London, London School of Economics; forthcoming in E. Maasoumi (ed.) Contributions to Econometrics: The Seminal Works of J. Denis Sargan, Cambridge, Cambridge University Press.

Sargan, J.D. (1980a) "The Consumer Price Equation in the Post War British Economy: An Exercise in Equation Specification Testing", Review of Economic Studies, 47, 1, 113-135.

Sargan, J.D. (1980b) "A Model of Wage-price Inflation", Review of Economic Studies, 47, 1, 97-112.

Sargan, J.D. (1980c) "Some Approximations to the Distribution of Econometric Criteria Which Are Asymptotically Distributed as Chi-squared", Econometrica, 48, 5, 1107-1138.

Sargan, J.D. (1980d) "Some Tests of Dynamic Specification for a Single Equation", Econometrica, 48, 4, 879-897.

Sargan, J.D. and W.M. Mikhail (1971) "A General Approximation to the Distribution of Instrumental Variables Estimates", Econometrica, 39, 1, 131-169.

Silvey, S.D. (1975) Statistical Inference, London, Chapman and Hall.

Sims, C.A. (1974) "Distributed Lags", Chapter 5 in M.D. Intriligator and D.A. Kendrick (eds.) Frontiers of Quantitative Economics, Amsterdam, North-Holland, Volume II, 289-338 (with discussion).

Sowey, E.R. (1972) "A Chronological and Classified Bibliography on Random Number Generation and Testing", International Statistical Review, 40, 3, 355-371.

Sowey, E.R. (1973) "A Classified Bibliography of Monte Carlo Studies in Econometrics", Journal of Econometrics, 1, 4, 377-395.

Sowey, E.R. (1978) "A Second Classified Bibliography on Random Number Generation and Testing", International Statistical Review, 46, *, 89-102.

Sowey, E.R. (1986) "A Third Classified Bibliography on Random Number Generation and Testing", Journal of the Royal Statistical Society, Series A, 149, 1, 83-107.

Stroud, T.W.F. (1971) "On Obtaining Large-sample Tests from Asymptotically Normal Estimators", Annals of Mathematical Statistics, 42, 4, 1412-1424.

Summers, R. (1959) "A Strategy for Appraising Various Simultaneous Equation Estimating Methods by Means of Sampling Experiments", mimeo, New Haven, Cowles Foundation, January.

Summers, R. (1965) "A Capital Intensive Approach to the Small Sample Properties of Various Simultaneous Equation Estimators", Econometrica, 33, 1, 1-41.

Theil, H. (1971) Principles of Econometrics, New York, John Wiley and Sons.

Wald, A. (1943) "Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations Is Large", Transactions of the American Mathematical Society, 54, 3, 426-482; reprinted in T.W. Anderson et al. (eds.) Selected Papers in Statistics and Probability by Abraham Wald, New York, McGraw-Hill, 1955.

White, H. (1980a) "A Heteroskedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", Econometrica, 48, 4, 817-838.

White, H. (1980b) "Using Least Squares to Approximate Unknown Regression Functions", International Economic Review, 21, 1, 149-170.

Yule, G.U. (1926) "Why Do We Sometimes Get Nonsense-correlations Between Time-series? -- A Study in Sampling and the Nature of Time-series", Journal of the Royal Statistical Society, 89, 1, 1-69 (with discussion).

International Finance Discussion Papers

IFDP NUMBER	TITLES	AUTHOR(s)
	1987	
317	Monte Carlo Methodology and the Finite Sample Properties of Statistics for Testing Nested and Non-Nested Hypotheses	Neil R. Ericsson
316	The U.S. External Deficit: Its Causes and Persistence	Peter Hooper Catherine L. Mann
315	Debt Conversions: Economic Issues for Heavily Indebted Developing Countries	Lewis S. Alexander
314	Exchange Rate Regimes and Macroeconomic Stabilization in a Developing Country	David H. Howard
313	Monetary Policy in Taiwan, China	Robert F. Emery
312	The Pricing of Forward Exchange Rates	Ross Levine
311	Realignment of the Yen-Dollar Exchange Rate: Aspects of the Adjustment Process in Japan	Bonnie E. Loopesko Robert E. Johnson
310	The Effect of Multilateral Trade Clearinghouses on the Demand for International Reserves	Ellen E. Meade
309	Protection and Retaliation: Changing the Rules of the Game	Catherine L. Mann
308	International Duopoly with Tariffs	Eric O'N. Fisher Charles A. Wilson
307	A Simple Simulation Model of International Bank Lending	Henry S. Terrell Robert S. Dohner
306	A Reassessment of Measures of the Dollar's Effective Exchange Value	B. Dianne Pauls William L. Helkie
305	Macroeconomic Instability of the Less Developed Country Economy when Bank Credit is Rationed	David F. Spigelman
304	The U.S. External Deficit in the 1980s: An Empirical Analysis	William L. Helkie Peter Hooper

---

Please address requests for copies to International Finance Discussion Papers, Division of International Finance, Stop 24, Board of Governors of the Federal Reserve System, Washington, D.C. 20551.

International Finance Discussion Papers

IFDP NUMBER	TITLES	AUTHOR(s)
303	An Analogue Model of Phase-Averaging Procedures	Julia Campos Neil R. Ericsson David F. Hendry
302	A Model of Exchange Rate Pass-Through	Eric O'N. Fisher
301	The Out-of-Sample Forecasting Performance of Exchange Rate Models When Coefficients are Allowed to Change	Garry J. Schinasi P.A.V.B. Swamy
300	Financial Concentration and Development: An Empirical Analysis of the Venezuelan Case	Jaime Marquez Janice Shack-Marquez
299	Deposit Insurance Assessments on Deposits at Foreign Branches of U.S. Banks	Jeffrey C. Marquardt
	1986	
298	The International Debt Situation	Edwin M. Truman
297	The Cost Competitiveness of the Europaper Market	Rodney H. Mills
296	Germany and the European Disease	John Davis Patrick Minford
295	The United States International Asset and Liability Position: A Comparison of Flow of Funds and Commerce Department	Guido E. van der Ven John E. Wilson
294	An International Arbitrage Pricing Model with PPP Deviations	Ross Levine
293	The Structure and Properties of the FRB Multicountry Model	Hali J. Edison Jaime R. Marquez Ralph W. Tryon
292	Short-term and Long-term Expectations of the Yen/Dollar Exchange Rate: Evidence from Survey Data	Jeffrey A. Frankel Kenneth A. Froot
291	Anticipated Fiscal Contraction: The Economic Consequences of the Announcement of Gramm-Rudman-Hollings	Robert A. Johnson
290	Tests of the Foreign Exchange Risk Premium Using the Expected Second Moments Implied by Option Pricing	Richard K. Lyons