International Finance Discussion Papers

Number 276

Revised: March 1986

POST-SIMULATION ANALYSIS OF MONTE CARLO EXPERIMENTS: INTERPRETING PESARAN'S (1974) STUDY OF NON-NESTED HYPOTHESIS TEST STATISTICS

by

Neil R. Ericsson

ABSTRACT

"Monte Carlo experimentation in econometrics helps 'solve' deterministic problems by simulating stochastic analogues in which the analytical unknowns are reformulated as parameters to be estimated." (Hendry (1980)) With that in mind, Monte Carlo studies may be divided operationally into three phases: design, simulation, and post-simulation analysis. This paper provides a guide to the last of those three, post-simulation analysis, <u>given</u> the design and simulation of a Monte Carlo study, and uses Pesaran's (1974) study of statistics for testing non-nested hypotheses to illustrate the techniques described. A statistic is derived for testing for significant deviations between the asymptotic and (observed) finite sample properties. Further, that statistic provides the basis for analysing discrepancies between the finite sample and asymptotic properties using response surfaces. The results for Pesaran's study indicate the value of asymptotic theory in interpreting finite sample properties and certain limitations for doing so. Finally, a method is proposed for adjusting the finite sample sizes of different test statistics so that comparisons of their power may be made. Extensions to other finite sample properties are indicated.

Post-simulation Analysis of Monte Carlo Experiments: Interpreting Pesaran's (1974) Study of Non-nested Hypothesis Test Statistics

by

Neil R. Ericsson*

1. Introduction

Monte Carlo studies in econometrics often have been criticized for imprecision present in estimating the underlying finite sample properties investigated and for the specificity of the results from the particular parameter values and sample sizes chosen, so making any conclusions very tentative at best. Hendry (1984) presents a methodology reducing both and which aims to obtain "numerical-analytical formulae 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. Illustrating such post-simulation analysis, this paper re-examines Pesaran's (1974) Monte Carlo study of statistics for testing non-nested hypotheses.

---

*This research was supported in part by a grant from the Social Science Research Council to the Programme in Methodology, Inference and Modelling in Econometrics at the London School of Economics (1979-81) while I was at the LSE. 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. I have benefited from valuable discussions with and comments from Julia Campos, Raymond Chapman, David Hendry, Grayham Mizon, Hashem Pesaran, Denis Sargan, and Frank Srba. I am grateful to Hashem Pesaran for detailed information on his 1974 Monte Carlo experiments (including the results used for forecast tests herein), and for his permission to quote in equations (34) and (35) unpublished material from Pesaran (1981). All response surfaces were obtained with the computer program GIVE (Hendry and Srba (1978, 1980), Hendry, Morgan, and Srba (1982)). This is a revised and shortened version of Ericsson (1982b, Chapter 3) and is forthcoming in Ericsson (1986) (but without Appendices A, C, and D).

Pesaran (1974) compares the finite sample properties of the Cox (1961, 1962) and F statistics for testing the mis-specification of a static single equation model. As a basis for analysing his results, Section 2 describes the asymptotic properties of those statistics. Section 3 interprets Pesaran's Monte Carlo design in light of Hendry (1984) and notes further properties of the statistics when the data generation process is the one selected by Pesaran. Section 4 explains the valuable roles that a response surface can have as a convenient summary of Monte Carlo simulations and as an approximation to the underlying finite sample distribution function. Further, the section details how to assess the adequacy of a response surface. Section 5 compares Pesaran's experimental results directly with what asymptotic theory would predict, and, to the extent that significant discrepancies exist, attempts to explain them, employing the methodology from Section 4. Finally, Section 6 proposes adjustments such that the estimated finite sample powers of the Cox and F tests may be compared on a more equal footing in those instances where the estimated finite sample probability of type I error for the Cox statistic significantly exceeds its nominal value. Although the approach presented relates to finite sample rejection frequencies in particular, similar strategies exist for evaluating most properties of estimators and test statistics, including biases, simulation standard errors, and estimated asymptotic standard errors.

2. Statistics for testing non-nested hypotheses

This section summarizes existing analytical results for the two statistics examined by Pesaran (1974). Consider the two non-nested hypotheses

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

and

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

and the comprehensive hypothesis

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

where the dependent variable y is $T\times1$, T being the econometric sample size; X_i is a $T\times k_i$ matrix of regressors and β_i the corresponding $k_i\times1$ 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 $T\times1$ vector of disturbances 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 when X_0 and X_1 are predetermined: following Cox (1961, 1962), Pesaran (1974) proposes evaluating a modified likelihood ratio statistic for H_0 and H_1; 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 (cf. Silvey (1975, pp. 115-116)). Under H_0,

$$D_0 \;\tilde{a}\; N(0,1) \tag{4}$$

and

$$f_2 \;\tilde{a}\; F(k_2-k_0,\, T-k_2,\, 0) \tag{5}$$

where D_0 is the Cox statistic for which H_0 is assumed true, f_2 is the F statistic for testing H_0 against the general hypothesis H_2, and $\tilde{a}$ denotes "converges in distribution to, as $T \to \infty$". Pesaran (1982) and Ericsson (1983) derive the asymptotic properties of D_0 under H_0 and H_1; however, Pesaran (1974) estimates the finite sample rejection frequencies of D_0 and D_1 (the latter being the Cox statistic for which H_1 is assumed true) under H_0 <u>only</u>. Noting the similarity in those approaches, it follows from Ericsson (1983) that, under H_0 as a local alternative to H_1,

$$D_1 \;\overset{a}{\sim}\; N(\mu_1, \omega_1) \tag{6}$$

and

$$f_3 \;\overset{a}{\sim}\; F(k_2-k_1,\ T-k_2,\ \lambda_3) \tag{7}$$

where μ_1 and ω_1 are the asymptotic mean and variance of D_1, f_3 is the F statistic for testing H_1 against H_2, and λ_3 is the asymptotic non-centrality parameter of that F statistic.[1]

3. The data generation process and experimental design

Pesaran (1974, p. 160) generates his data in the following manner:

$$y_t = a_0 + b_0 x_t + u_{0t} \qquad u_{0t} \sim NID(0, \sigma_0^2) \tag{8}$$

$$\begin{bmatrix} x_t \\ z_t \end{bmatrix} \sim NID\left(\underset{\sim}{0}, \begin{bmatrix} 1 & r/(1-r^2)^{1/2} \\ r/(1-r^2)^{1/2} & (1-r^2)^{-1} \end{bmatrix}\right) \qquad t=1,\ldots,T \tag{9}$$

where r is the population correlation coefficient of x_t and z_t.[2] Given that process, he considers three specifications:

$$H_0:\ y_t = a_0 + b_0 x_t + u_{0t} \qquad u_{0t} \sim NID(0, \sigma_0^2) \tag{10}$$

$$H_1:\ y_t = a_1 + b_1 z_t + u_{1t} \qquad u_{1t} \sim NID(0, \sigma_1^2) \tag{11}$$

$$H_2:\ y_t = a_2 + b_2 x_t + c_2 z_t + u_{2t} \qquad u_{2t} \sim NID(0, \sigma_2^2)\ . \tag{12}$$

Letting R be the population multiple correlation coefficient for (8)

---

[1] Ericsson's (1983) derivation of an equation equivalent to (6) but for D_0 with H_1 as a local alternative involves an approximation additional to the standard asymptotic approximation. Pesaran (1982) only requires the usual asymptotic approximation but must have at least as many regressors (total) under H_0 as regressors in H_1 but not in H_0. (Note that the roles of H_0 and H_1 are reversed for D_1.) In practice, Ericsson's and Pesaran's approximations appear numerically similar; and, for the model in Pesaran (1974), Pesaran's (1982) asymptotic distribution for D_1^2 is just that of f_3. However, the formula in (6) is used throughout. See Appendix A for a detailed discussion of the statistical properties of D_0, D_1, f_2, and f_3.

For extensive discussions on the role and interpretation of non-nested hypothesis test statistics, see MacKinnon (1983) (including the comments thereon), the Journal of Econometrics issue on "non-nested models" (1983, Vol. 21, No. 1), Mizon (1984), and Mizon and Richard (1986). Godfrey and Pesaran (1983) suggest and evaluate analytical adjustments for improving the finite sample properties of the Cox statistic.

[2] Pesaran's (1974) notation of r_{xz} has been simplified to r, and n is now T.

($R^2=b_0^2/(b_0^2+\sigma_0^2)$), Pesaran (1974) calculates D_0, D_1, f_2, and f_3 for 500 replications of each of 400 points defined by all possible combinations of

$$T = (20, 40, 60, 80)$$
$$r^2 = (.90, .91, \ldots, .99)$$
$$R^2 = (.80, .81, \ldots, .89) \tag{13}$$

with $a_0=100$ and $b_0=2$. Following Hendry's (1984, p. 940) notation and terminology, the Monte Carlo design variables are

$$\theta = (a_0, b_0, r, R^2)' \, \varepsilon \, \Theta = \{\theta \mid r^2<1, R^2<1\} \tag{14}$$

and

$$T \, \varepsilon \, \mathrm{T} = [T^0, T^1] \tag{15}$$

where T is pre-assigned with $T^0=20$ and $T^1=80$.[3] Equations (8) and (9) are the data generation process (DGP); $\Theta \times \mathrm{T}$ is the parameter space; equations (10), (11), and (12) are the relationships of interest[4]; and the objective of the Monte Carlo study is to determine the finite sample distributions of the statistics D_0, D_1, f_2, and f_3 as defined by the relationships of interest, within the specified parameter space of the DGP. More modestly, letting τ be any of D_0, D_1, f_2, and f_3 and δ be the critical value associated with a test based on τ, the objective is to find the finite sample rejection frequency $\pi \equiv \text{prob}(|\tau|\geq\delta)$. That probability depends upon θ and T and can be expressed as a conditional probability formula:

$$\pi \equiv \text{prob}(|\tau|\geq\delta \mid \theta, T) = g(\theta, T) \, . \tag{16}$$

Thus, we wish to know (or obtain a good approximation to) $g(\theta,T)$ over $\Theta \times \mathrm{T}$.[5]

---

[3]In (14), θ could be defined as $\theta = (a_0, b_0, r, \sigma_0^2)'$, noting that (b_0, R^2) maps one-to-one onto (b_0, σ_0^2).

[4]Note that both H_0 and H_2 coincide with the DGP but H_1 does not: specifically, H_1 is a mis-specified model with x_t in H_0 replaced by z_t, a variable correlated with x_t.

[5]Implicitly, $g(\cdot,\cdot)$ is a function of δ as well. However, because δ is constant for each of the two types of statistics in Pesaran (1974), its presence in $g(\cdot,\cdot)$ is ignored in the analysis below.

The DGP defined by (8) and (9) and the relationships of interest in (10)-(12) have certain implications for the properties of the statistics being examined. First, the formulae for their asymptotic distributions have explicit representations: λ_3, μ_1, and ω_1 may be expressed as

$$\lambda_3 \quad = \quad Tb_0{}^2(1-r^2)/\sigma_0{}^2 \quad , \tag{17}$$

$$\mu_1 \quad = \quad -T^{1/2}b_0(1-r^2)^{1/2}(1+r^2)/(2\sigma_0 r^2) \quad , \tag{18}$$

and

$$\omega_1 \quad = \quad \{4r^4 + (1-r^2)(1+r^2)^2\}/(4r^6) \quad . \tag{19}$$

For r^2 close to unity (relevant for Pesaran's experimental design; see (13)), $\mu_1{}^2 \approx \lambda_3$ and $\omega_1 \approx 1$, so the distribution of $D_1{}^2$ is roughly a $\chi^2(1,\lambda_3)$. Because f_3 is asymptotically distributed as an $F(1, T-k_2, \lambda_3)$ (and hence asymptotically as a $\chi^2(1,\lambda_3)$), D_1 should have about the same power as f_3 for Pesaran's experiments. Second, for a given local alternative (i.e., for a constant non-zero $b_0\surd T$ in (17) and (18)), the asymptotic powers of the Cox and F tests tend to the nominal size as $r^2 \to 1$. Third, because X_2 is fixed, f_2 and f_3 are exactly distributed as $F(k_2-k_0, T-k_2, 0)$ and $F(k_2-k_1, T-k_2, \lambda_3^*)$ where the finite sample non-centrality parameter λ_3^* is

$$\lambda_3^* \quad = \quad b_0{}'X_0^{+}{}'(W_2-W_1)X_0^{+}b_0/\sigma_0{}^2 \tag{20}$$

with $W_i = X_i(X_i{}'X_i)^{-1}X_i{}'$ $(i=0,1,2)$ and $X_0^+ = (x_1, x_2, \ldots, x_T)'$. These analytical results are invaluable in interpreting the Monte Carlo simulation results, as will become apparent in the following sections.

4. Monte Carlo 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.[6] Consider a binary response variable for which the probability of "success" (or, frequently later, acceptance or rejection by a particular test) is π $(0 < \pi < 1)$ and on which there are N observations $(N > 1)$, S being "successes". Letting

$$A \quad = \quad [S(N-S)]/(N-1) \ , \tag{21}$$

$$L(\zeta) \quad = \quad A^{1/2}\ln\left[\frac{\zeta}{1-\zeta}\right] \qquad\qquad 0 < \zeta < 1, \tag{22}$$

and

$$L^*(\zeta) \; = \; A^{1/2}\ln\left[\frac{\zeta - (2N)^{-1}}{1 - \zeta - (2N)^{-1}}\right] \qquad (2N)^{-1} < \zeta < 1-(2N)^{-1} \tag{23}$$

in a notation similar to that found in Hendry (1984, pp. 957-961) and Mizon and Hendry (1980, p. 34), then, using results from Cox (1970, pp. 30-34, 41-42, 78-79), it can be shown that

$$\phi(s,\pi) \quad \equiv \quad L^*(s) - L(\pi) \quad \underset{A}{\sim} \quad N(0,1) \tag{24}$$

where $s \equiv S/N$ and $\underset{A}{\sim}$ denotes "converges in distribution to, as $N \to \infty$ ".[7] In the context of Pesaran's (1974) Monte Carlo study (and Monte Carlo studies of powers in general), 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.

Typically, π is unknown; and, as an initial step in analysing Monte Carlo results, it is of interest to test whether π equals the (local) asymptotic (i.e., as $T \to \infty$) power of the test (π_a, say). That is easily

---

[6]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; cf. Sowey (1973).

[7]A proof is given in Appendix B.

accomplished by replacing π by π_a in (24) and comparing the value of $\phi(s,\pi_a)$ with the appropriate critical values for the standardized normal distribution.[8]

Even when $\pi=\pi_a$ is not a good approximation, (24) still provides the basis for analysing the Monte Carlo results. Without loss of generality,

$$\left[\frac{\pi_T}{1-\pi_T}\right] = \left[\frac{\pi_a}{1-\pi_a}\right]^{\alpha} \exp\{G^+(\theta,T)\} \tag{25}$$

where θ is the vector of all the parameters (except T) which define the model generating the binary random variable of interest, $G^+(\cdot,\cdot)$ is some appropriate function, α is a parameter which itself may depend upon θ and T, and π is subscripted by T so as to emphasize that π ($\equiv E(S/N)$) is a function of the econometric sample size.[9] Using (25), (24) may be rewritten as

$$L^*(s) = \alpha\, L(\pi_a) + A^{1/2}\cdot G^+(\theta,T) + \varepsilon \qquad \varepsilon \underset{A}{\sim} N(0,1)\ . \tag{26}$$

---

[8]As indicated by Cox (1970, pp. 27-29), the logistic and normal distributions only differ slightly, primarily in the tail regions. So, provided that π is not too close to 0 or 1, the normal distribution should provide a good approximation to the distribution of (24). Furthermore, the use of the normal approximation is often justified because N is typically quite "large" in Monte Carlo studies (e.g., N = 500 in Pesaran (1974)).

It is assumed throughout that π and π_a each lie strictly within the unit interval.

[9]Clearly, (25) is equivalent to (16). However, the functional form of (25) ensures that predicted powers are within the unit interval. Further, (25) has the advantage of having split $g(\theta,T)$ in (16) into two components, an asymptotic term and a term involving the deviation between the finite sample and asymptotic distributions. The equivalent partition for (16) is $\pi_T = \pi_a + (\pi_T-\pi_a)$. By using asymptotic theory, we are able to simplify the problem of simulating π_T (of O(1)) directly to one of analytically calculating π_a and simulating only $(\pi_T-\pi_a)$ (of o(1), and quite possibly $O(T^{-1/2})$). That is in line with Hammersley and Handscomb's (1964, pp. 5, 59) precept that one should solve as much of the problem as possible analytically in order to minimize the imprecision and specificity arising from simulation. With that in mind, see Hendry (1973) and Nickell (1981) for two elegant examples in which analytical formulae greatly simplify the interpretation of previous Monte Carlo studies.

If an analytical approximation to π_T better than π_a is available (e.g., an Edgeworth expansion), it could appear in (25) in place of π_a, further reducing the order of the term being simulated.

In practice, α and the functional form of $G^{+}(\cdot,\cdot)$ are unknown although (e.g., in Pesaran (1974)) they are implicitly defined by the computer program generating the Monte Carlo data. If α and $G^{+}(\cdot,\cdot)$ were known, the exact finite sample probability (of "success", rejection) for any particular value of (θ,T) could be calculated directly from (25), obviating any need for conducting Monte Carlo experiments to estimate π_T. Even with α and $G^{+}(\cdot,\cdot)$ unknown, approximations to them may be found; and further, the accuracy of those approximations may be tested.[10]

From asymptotic theory, one expects that $\alpha \rightarrow 1$ as $T \rightarrow \infty$ and that

$$G^{+}(\theta,T) \;=\; T^{-1/2}G(\theta,T^{-1/2}) \tag{27}$$

where $G(\theta,T^{-1/2})$ is $O(T^{0})$ (cf. Phillips (1977, p. 474; 1982), Sargan (1980, p. 1120)). Thus, α might be expanded in powers of $T^{-1/2}$ about $T=\infty$:

$$\alpha \;=\; \alpha_0 \;+\; \alpha_1\, T^{-1/2} \;+\; \alpha_2\, T^{-1} \;+\; \ldots\;, \tag{28}$$

where α_0 is expected to be unity; and $G(\cdot,\cdot)$ might be expanded in powers of $T^{-1/2}$ and of the elements of θ.[11] Truncating both the series for α and the series for $G(\theta,T^{-1/2})$, the coefficients of the powers and cross-products of θ, $T^{-1/2}$, and $\ln\{\pi_a/(1-\pi_a)\}$ may be estimated by least squares, correcting for heteroscedasticity using the weight $A^{1/2}$, i.e., from estimating

---

[10]There are two distinct senses in which α and $G(\cdot,\cdot)$ can be known: for the particular experiments in which π_a is estimated, and for any values of (θ,T) in $\Theta\times T$. Clearly, the latter is far more useful. That distinction also emphasizes the value of choosing an experimental design which covers a wide range of $\Theta\times T$.

Aneuryn-Evans and Deaton (1980, pp. 284-285) suggest an alternative framework for analysing Monte Carlo results on statistics for testing non-nested hypotheses. However, their approach did not prove fruitful. See Ericsson (1982a) for details.

[11]The parameterisation of θ is not unique, and it may be worthwhile transforming "natural" parameters of the model into parameters which span the same range as $L^{*}(s)$ before expanding $G(\cdot,\cdot)$. For instance, a parameter bounded between -1 and 1 (ρ, say) might better appear in θ as $\rho/(1-\rho^2)$; a parameter bounded from below by zero (σ^2, say) might better appear in θ as $\ln(\sigma^2)$.

$$L^*(s) = \alpha_0 L(\pi_a) + \alpha_1 T^{-1/2} L(\pi_a) + \ldots + A^{1/2} T^{-1/2} H(\theta, T^{-1/2}) + e \qquad (29)$$

where $H(\theta, T^{-1/2})$ is the least squares approximation to $G(\theta, T^{-1/2})$ and the error term e is the combination of ε (the error arising from simulation) and $A^{1/2}T^{-1/2}\{G(\cdot,\cdot)-H(\cdot,\cdot)\}$ (the error from approximating $G(\cdot,\cdot)$ by $H(\cdot,\cdot)$).

A response surface like (29) 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 $\Theta \times T$ but not included in the simulations. Further, the response surface may adequately approximate the underlying finite sample distribution function. Two types of information are available for inferring how "good" a response surface like (29) is:

(A) asymptotic theory (i.e., $\alpha_0 = 1$), and

(B) $\varepsilon \underset{A}{\sim} \text{NID}(0,1)$

(cf. Hendry (1984, p. 962)). Although (B) is not directly testable, many testable implications follow from the null hypothesis that $H(\cdot,\cdot)=G(\cdot,\cdot)$.

(B_1) $\sigma_e^2 = 1$. 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 (29) 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_2) The error e does not include any terms involving θ, $T^{-1/2}$, and $\ln\{\pi_a/(1-\pi_a)\}$. By using OLS, e <u>can 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 rather general formulation for $H(\cdot,\cdot)$ and α and simplifying therefrom, the F statistic comparing the final specification against the general one helps test for the presence of such factors in the e's of the final specification.

(B_3) The error e is normally distributed.

(B_4) The e's are serially independent for <u>any</u> ordering of experiments specified prior to simulation. 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, $T^{-1/2}$, and $\ln\{\pi_a/(1-\pi_a)\}$ may induce serial correlation and/or heteroscedasticity in the e's.

(B_5) $H(\cdot,\cdot)$ is constant over regions of the parameter space which were not included in the estimation of (29).

Table I lists the bulk 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, $\xi_8(q)$ and $\eta_8(q,K-m-q)$ both test for q^{th}-order residual autocorrelation. There are K experiments and m regressors in the response surface under the null hypothesis.

The extent to which (A) and (B_1)-(B_5) 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) and (B_1)-(B_5) depends crucially on the number of replications per experiment, the experimental design (i.e., the points in $\Theta\times T$ examined), and the choice of DGP and $\Theta\times T$. Finally, even if any of (A) and (B_1)-(B_5) 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.

Table I. Criteria for evaluating response surfaces

Null	Alternative	Statistic[a]	Sources
(A)	$\alpha_0 \neq 1$	$\eta_2(1,K-m-1)$	Hendry (1984, p. 962)
(B_1)	$\sigma_e^2 > 1$	$\xi_3(K-m)$	Theil (1971, pp. 137-8)
(B_2)	q invalid parameter restrictions	$\eta_4(q,K-m-q)$	Johnston (1963, p. 126)
(B_3)	skewness (SK) and excess kurtosis (EK)	$\xi_5(2)$	Jarque and Bera (1980)
(B_3)	heteroscedasticity quadratic in regressors (q quadratic terms)	$\eta_6(q,K-m-q)$	White (1980a, p. 825), Nicholls and Pagan (1983)
(B_4)	first-order ARCH	$\xi_7(1)$	Engle (1982)
(B_4)	first-order residual autocorrelation	dw	Durbin and Watson (1950, 1951), Farebrother (1980)
(B_4)	q^{th}-order residual autocorrelation	$\xi_8(q)$; $\eta_8(q,K-m-q)$	Box and Pierce (1970); Godfrey (1978), Harvey (1981, p. 173)
(B_5)	$H^+(\cdot,\cdot)$ not constant over subsamples	$\eta_9(m,K-2m)$	Kendall (1946, pp. 242ff), Chow (1960, pp. 595ff)
(B_5)	predictive failure over a subset of q observations[b,c]	$\xi_1(q)$; $\eta_1(q,K-m-q)$	Hendry (1979, p. 222); Chow (1960, pp. 594-5)

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

b. We have labelled the Chow statistic $\eta_1(q,K-m-q)$ to highlight the pre-eminence of the issue of constancy. The covariance test statistic $\eta_9(m,K-2m)$ is often (and confusingly) referred to as the "Chow statistic" although Chow (1960, p. 592) was well aware of its presence in the literature.

c. 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 $\xi_1'(q)$.

5. Evaluation of Pesaran's estimated finite sample rejection frequencies

Equation (24) above is the foundation for evaluating Monte Carlo estimates of finite sample rejection frequencies. This section utilizes (24) both in assessing how close the simulation results are to what asymptotic theory would predict and in formulating response surfaces to explain significant deviations between Monte Carlo results and asymptotic theory.

The "closeness" of an estimated finite sample rejection frequency of a test statistic to its asymptotic value can be tested using $\phi(s,\pi_a)$ from (24) above where s is the fraction of replications for which the value of the statistic lies in the critical region and π_a is the probability of the statistic lying in the critical region asymptotically. The values of the estimated finite sample powers and estimated probability of type I error for the Cox and F statistics, their asymptotic values, and the values of $\phi(s,\pi_a)$ are given in Table II. Figure 1 strikingly displays how much of the inter-experimental variation in the finite sample rejection frequencies is explained by asymptotic theory. Now consider those results for each statistic in turn.

Pesaran (1974, p. 161) notes that the estimated probability of type I error for f_2 never differs significantly from .05, and the corresponding values of $\phi(s,.05)$ in Table II confirm that. Further, the sum of squares of $\phi(s,.05)$ over Pesaran's (1974) nine experiments is $\xi_3(9) = 6.45$, offering additional support to the hypothesis $\pi = .05$.[12]

The estimated probability of type I error for D_0 is significantly different from (and larger than) .05 in several experiments, although the magnitude of their difference decreases as T increases, in line with

[12]Note that $\xi_3(9)$ is distributed as a $\chi^2(9)$ for large N (provided that $\pi = .05$) and that $\text{prob}[\chi^2(9) \geq 16.92] = .05$.

Table II. The asymptotic powers (π_a) and estimated finite sample powers (s) of D_1 and f_3, the estimated type I errors (s) of D_0 and f_2, and, in each case, the value of a statistic $\phi(s,\cdot)$ for testing the closeness of the asymptotic and finite sample results.

		D_1			D_0		f_3			f_2	
Sample Size	r^2	s	π_a	$\phi(s,\pi_a)$	s^a	$\phi(s,.05)$	s	π_a	$\phi(s,\pi_a)$	s^a	$\phi(s,.05)$
T=20	.90	.818[b] (.017)[d]	.822	-.20[c]	.082 (.012)	3.18*	.656 (.021)	.768	-5.85*	.048 (.010)	-.30
	.95	.628 (.022)	.535	4.16*	.086 (.013)	3.58*	.436 (.022)	.458	-.99	.056 (.010)	.53
	.99	.196 (.018)	.149	2.91*	.072 (.012)	2.17*	.122 (.015)	.129	-.52	.052 (.010)	.11
T=40	.90	.968 (.008)	.979	-1.58	.076 (.012)	2.58*	.932 (.011)	.984	-8.38*	.048 (.010)	-.30
	.95	.830 (.017)	.815	.90	.080 (.012)	2.98*	.764 (.019)	.797	-1.81	.052 (.010)	.11
	.99	.306 (.021)	.248	2.98*	.056 (.010)	.53	.228 (.019)	.219	.46	.042 (.009)	-.92
T=80	.90	1.000 (.000)	1.000	-	.068 (.011)	1.76	1.000 (.000)	1.000	-	.046 (.009)	-.51
	.95	.980 (.006)	.979	.31	.050 (.010)	-.09	.968 (.008)	.987	-3.50*	.034 (.008)	-1.75
	.99	.476 (.022)	.436	1.80	.050 (.010)	-.09	.404 (.022)	.404	-.01	.038 (.009)	-1.34

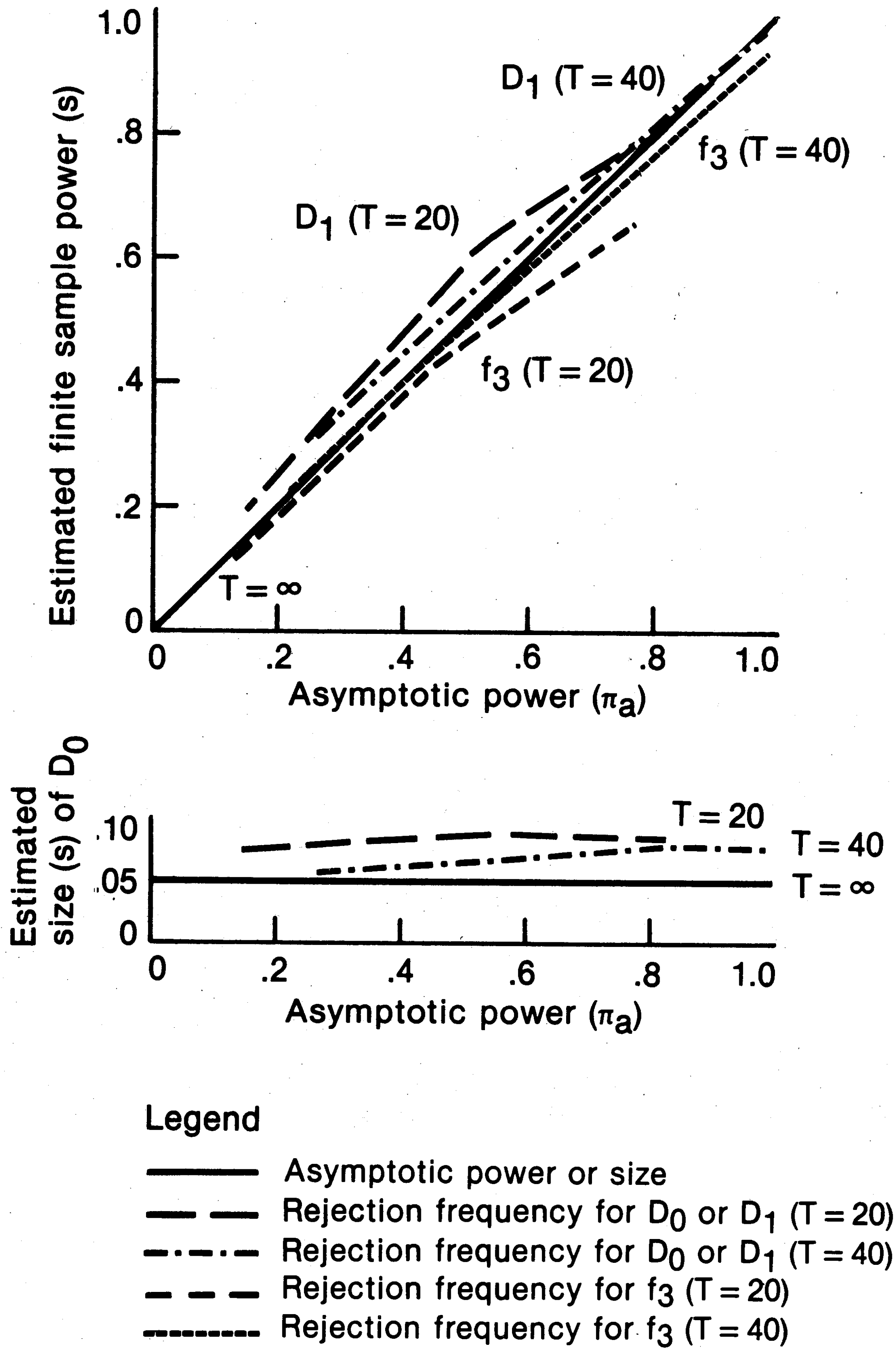
Notes: a. The nominal size of D_0 and f_2 is .05.

b. The total number of replications per experiment (N) is 500 for each experiment, with the multiple correlation coefficient of the DGP (R^2) being .80.

c. Under the null hypothesis that $\pi = \pi_a$ (or that $\pi = .05$ for D_0 and f_2), $\phi(s,\cdot) \underset{A}{\sim} N(0,1)$. Asterisks denote results for which the null hypothesis is rejected at the 5% level.

d. Simulation standard errors of estimated probabilities are in parentheses.

Figure 1
Asymptotic and estimated finite sample rejection frequencies for D_0, D_1, and f_3



asymptotic theory (cf. Figure 1). Further, for T=80, the three estimated probabilities are individually and jointly insignificantly different from .05.

The estimated finite sample power of D_1 differs significantly from its asymptotic power in three experiments although, as with the probability of type I error, the difference between estimated and asymptotic values decreases as T increases. Also, as with the probability of type I error, the estimated values are insignificantly different from their asymptotic values for T=80.[13]

That the estimated finite sample power for f_3 differs from its asymptotic power as markedly as it does in three experiments is a surprise at first blush, and requires some explanation. As noted in Sections 2 and 3, f_3 is exactly distributed as $F(1, T-k_2, \lambda_3^*)$ but only asymptotically as $F(1, T-k_2, \lambda_3)$, where λ_3^* and λ_3 are the finite sample and asymptotic non-centrality parameters in (20) and (17).[14] λ_3^* is <u>not</u> equal to λ_3

---

[13]No test can be made for $\{T=80, r^2=.90\}$ because S = N. See Cox (1970, pp. 33, 42, 78) on the treatment of experiments with S=0 or S=N.

[14]For Pesaran's model,

$$f_3 = \{y'(W_2-W_1)y/(k_2-k_1)\}/\{y'(I_T-W_2)y/(T-k_2)\} \quad .$$

Noting that $(I_T-W_2)X_2 = 0$ and $(W_2-W_1)X_2\beta_2 = (W_2-W_1)X_0^+\beta_0^+$ where β_0^+ is the sub-vector of parameters in β_0 corresponding to X_0^+, then, under H_0,

$$f_3 = \{u^{*\prime}(W_2-W_1)u^*/(k_2-k_1)\}/\{u^{*\prime}(I_T-W_2)u^*/(T-k_2)\}$$

where $u^* \equiv X_0^+\beta_0^+ + u_0$. If X_2 is fixed in repeated samples, or at least if X_2 and u_0 are independent in each sample (and so X_2 may be conditioned upon; see Schmidt (1976, pp. 93-94ff)), then, conditional upon X_2,

$$u^* \sim N(X_0^+\beta_0^+, \sigma_0{}^2\cdot I_T) \quad .$$

Because (W_2-W_1) and (I_T-W_2) are each idempotent and of ranks (k_2-k_1) and $(T-k_2)$, respectively,

$$u^{*\prime}(W_2-W_1)u^*/\sigma_0{}^2 \sim \chi^2(k_2-k_1, \lambda_3^*) \quad \text{and}$$
$$u^{*\prime}(I_T-W_2)u^*/\sigma_0{}^2 \sim \chi^2(T-k_2, 0) \quad \text{where}$$
$$\lambda_3^* = \beta_0^{+\prime}X_0^{+\prime}(W_2-W_1)X_0^+\beta_0^+/\sigma_0{}^2 \quad ,$$

noting that $(I_T-W_2)X_0^+ = 0$. Further, the two χ^2 variates above are independent because $(I_T-W_2)(W_2-W_1) = 0$ (see Rao (1973, p. 187)), so

$$f_3 \sim F(k_2-k_1, T-k_2, \lambda_3^*) \qquad \text{(exactly)} \; .$$

in general, so f_3 is not (generally) distributed as an $F(1, T-k_2, \lambda_3)$. Further, λ_3^* varies across replications because the x_t's and z_t's are generated for each replication. Because the distribution of f_3 is a non-linear function of λ_3^*, which is itself a non-linear function of X_0^+, the expected fraction of rejections by f_3 need not equal the probability of rejection associated with the given critical value for $F(1, T-k_2, \lambda_3)$.[15] However, (7) still holds, so, for large N, $\phi(s,\pi_a)$ should be approximately a standardized normal variate, provided that T is large enough. If X_0^+ were known for each replication (or if it had been held constant over replications within experiments and were known for each experiment), the exact probability of rejection for f_3 could be calculated for each replication (and so for each experiment). As it is, X_0^+ is not known, so only the asymptotic formula in (7) may be used. As noted for D_1, the difference between the estimated finite sample power of f_3 and its asymptotic power shrinks as T increases.[16]

Because Pesaran's (1974, p. 162) published results contain just nine experiments (effectively eight for D_1 and f_3), only highly restrictive response surfaces may be fitted to them. Expanding $G(\theta,T^{-1/2})$ and α in $T^{-1/2}$ only and truncating at the first derivative (all a priori and arbitrarily) leads to the rather simple response surface

---

[15]The discrepancies between the theoretical and estimated finite sample powers of f_3 might have arisen from the approximation used to calculate the probabilities associated with the non-central F distribution. However, Appendix C shows that the errors involved are not large enough to explain those discrepancies.

[16]A similar analysis applies to the distribution of f_2 under H_2, with $f_2 \sim F(k_2-k_0, T-k_2, \lambda_2^*)$ and $f_2 \underset{a}{\sim} F(k_2-k_0, T-k_2, \lambda_2)$ where λ_2^* and λ_2 are defined in an obvious manner. Under H_0 (i.e., with $\beta_1^+=0$), however, $\lambda_2^* = \lambda_2 = 0$; so the exact and asymptotic distributions of f_2 are identical, and not a function of X_2 at all. The results for f_2 in Table II are in line with that discussion.

$$L^*(s) = (\alpha_0 + \alpha_1 T^{-1/2})L(\pi_a) + A^{1/2}T^{-1/2}(\kappa_0 + \kappa_1 T^{-1/2}) + e$$
$$= A^{1/2}(\alpha_0 \ell + \alpha_1 \ell T^{-1/2} + \kappa_0 T^{-1/2} + \kappa_1 T^{-1}) + e \qquad (30)$$

where $\ell \equiv \ln\{\pi_a/(1-\pi_a)\}$. Equation (30) may be too general a model for the estimated probabilities of interest; so, in the usual way, coefficients of the regressors in (30) may be tested to be zero or to satisfy other constraints. The final response surfaces for the estimated finite sample powers of D_1 and f_3 are

$$L^*(\hat{s}) - L(\pi_a) = A^{1/2}T^{-1/2}(\underset{\substack{(.23)\\ [.24]}}{1.16} - \underset{\substack{(.17)\\ [.16]}}{.50}\ell) \qquad (31)$$

$K = 8$ $R^2 = .75$ $\hat{\sigma}_e = 1.02$ $\xi_1(8) = 10.2$ $\xi_1'(8) = 10.7$ $\eta_1(8,6) = 1.03$

$\eta_2(1,5) = .02$ $\xi_3(6) = 6.3$ $\eta_4(2,4) = .04$ $\xi_5(2) = .7$ $SK = -.5$ $EK = -.4$

$\eta_6(3,11) = .24$ $\xi_7(1) = 1.8$ $dw = 2.49$ $\eta_9(2,12) = .85$

and

$$L^*(\hat{s}) - L(\pi_a) = A^{1/2}T^{-1/2}(\underset{\substack{(.43)\\ [.32]}}{-1.07} - \underset{\substack{(.30)\\ [.34]}}{1.20}\ell) \qquad (32)$$

$K = 8$ $R^2 = .66$ $\hat{\sigma}_e = 1.96$ $\xi_1(8) = 6.8$ $\xi_1'(8) = 26.0$ $\eta_1(8,6) = .84$

$\eta_2(1,5) = 1.29$ $\xi_3(6) = 23.1$ $\eta_4(2,4) = .6$ $\xi_5(2) = .8$ $SK = -.0$ $EK = -1.2$

$\eta_6(3,11) = .70$ $\xi_7(1) = 1.8$ $dw = 2.23$ $\eta_9(2,12) = .01$

respectively. R^2 is the unadjusted squared multiple correlation coefficient[17], $\hat{\sigma}_e$ is the square root of the estimated residual variance, and $(\cdot)$ and $[\cdot]$ respectively denote conventionally calculated and

---

[17] R^2 may lie outside the unit interval because (30) has no constant term (see Schmidt (1976, pp. 3-5)). In particular, R^2 is less than zero in the general response surface for D_0 because $L^*(s)$ has a large negative mean. However, $\hat{\sigma}_e$, not R^2, is the appropriate measure of the goodness-of-fit for the response surfaces, so small or negative values of R^2 *per se* are not worrisome.

In this paper, "R^2" refers both to the unadjusted squared multiple correlation coefficient for response surfaces and to the squared population multiple correlation coefficient for the DGP (8). The uses are distinct and no confusion should arise from using a common notation.

White (1980a) coefficient standard errors. Tests of (B_5) use eight sets of results for which T and r^2 take the same values as the fitted results, but for which the R^2 of the DGP is .85.[18] Noting that the entire discussion of response surfaces above applies equally well for any estimated probability (including the estimated probability of type I error), response surfaces for the estimated probability of type I error for D_0 are estimated as well, the final specification of which is

$$L^*(\hat{s}) - L(.05) = \underset{\substack{(1.5)\\[1.1]}}{A^{1/2}T^{-1}10.5} \tag{33}$$

$K = 8$ $R^2 = .65$ $\hat{\sigma}_e = .89$ $\xi_1(8) = 6.5$ $\xi_1^{\prime}(8) = 5.2$ $\eta_1(8,7) = .80$

$\eta_2(1,5) = .00$ $\xi_3(7) = 5.5$ $\eta_4(2,5) = 1.70$ $\xi_5(2) = 1.2$ $SK = .4$ $EK = -1.1$

$\eta_6(1,14) = 2.07$ $\xi_7(1) = 1.0$ $dw = 1.80$ $\eta_9(1,14) = .12$

The estimated probability of type I error for f_2 need not be analysed further because it does not deviate significantly from its asymptotic value.

The summary statistics indicate that the response surfaces for D_1 and D_0 show no signs of mis-specification nor are the restrictions that $\alpha_0 = 1$ and $\kappa_1 = 0$ (for D_1) and $\alpha_0 = 1$ and $\kappa_0 = 0$ (for D_0) rejected. Given the data available, those response surfaces support the conjectures that $\alpha_0 = 1$, $\sigma_e^2 = 1$, and that very simple expansions of α and $G(\theta,T^{-1/2})$ are sufficient to explain the observed Monte Carlo experiments. Further, both restricted response surfaces predict the eight out-of-sample observations

---

[18]See Appendix D for all the results. The results for {T=80, r^2=.90} are not included in <u>any</u> of the response surfaces although they could have been in those for D_0. Experiments are ordered as in Table II.

In general, experiments for prediction could be chosen by random selection or stratified selection. Cf. Cox (1958), Wilks (1962).

The statistics $\xi_5(\cdot)$, SK, EK, and $\eta_6(\cdot,\cdot)$ use all sixteen experiments because of the degrees of freedom involved; the others (where appropriate) use only the eight published in Pesaran (1974) (i.e., for which $R^2 = .80$). However, their values alter only slightly if all sixteen are included.

with acceptable accuracy, no matter whether the relative or absolute measure of forecast accuracy is used.[19]

The response surface for f_3 presents more of a puzzle. The restrictions in (32) are not rejected, and there is no indication of serial correlation in either the unrestricted or restricted response surfaces (and hence no indication that higher order terms from the expansions of α and $G(\theta,T^{-1/2})$ ought to be present in those response surfaces). However, σ_e^2 is significantly greater than unity in both response surfaces (see $\xi_3(6)$), possibly indicative of the problems with f_3 noted above. For comparison, Pesaran (1981) obtains the following response surfaces for D_1 and f_3 when using a DGP and relationships of interest similar to (8)-(12) but with the number of non-overlapping regressors in (10) and (11) varying from one to four.[20]

$$L^*(\hat{s}) = \underset{(.019)}{.967}L(\pi_a) + A^{1/2}(\underset{(.025)}{.034} - \underset{(.80)}{7.61}\ell T^{-1} + \underset{(.18)}{1.12}\lambda_3 T^{-1}) \tag{34}$$

$K = 108 \quad \bar{R}^2 = .982 \quad \hat{\sigma}_e = 1.30$

$\eta_2(1,104) = 3.02 \quad \xi_3(104) = 176. \quad dw = 1.07$

$$L^*(\hat{s}) = \underset{(.023)}{.977}L(\pi_a) + A^{1/2}(\underset{(.026)}{-.031} + \underset{(.90)}{1.62}\ell T^{-1} - \underset{(.15)}{2.18}\lambda_3 T^{-1}) \tag{35}$$

$K = 108 \quad \bar{R}^2 = .978 \quad \hat{\sigma}_e = 1.52$

$\eta_2(1,104) = 1.00 \quad \xi_3(104) = 240. \quad dw = 1.44$

These response surfaces concisely summarize results for over one hundred experiments and satisfy many of the criteria (A) and (B_1)-(B_5). However, as with (31) and (32), the response surface for f_3 fits worse (in terms of

---

[19]In the response surfaces for D_1, it was not necessary to include any "terms of $O(\delta^{-1})$", i.e., terms resulting from the additional approximation in (6) (see Ericsson (1983)). That suggests that the normal distribution approximates the asymptotic distribution of D_1 quite well.

[20]Pesaran (1982), the published version of Pesaran's (1981) working paper, unfortunately does not include these response surfaces.

$\hat{\sigma}_e$) than the one for D_1, so that issue may be worthwhile investigating further.

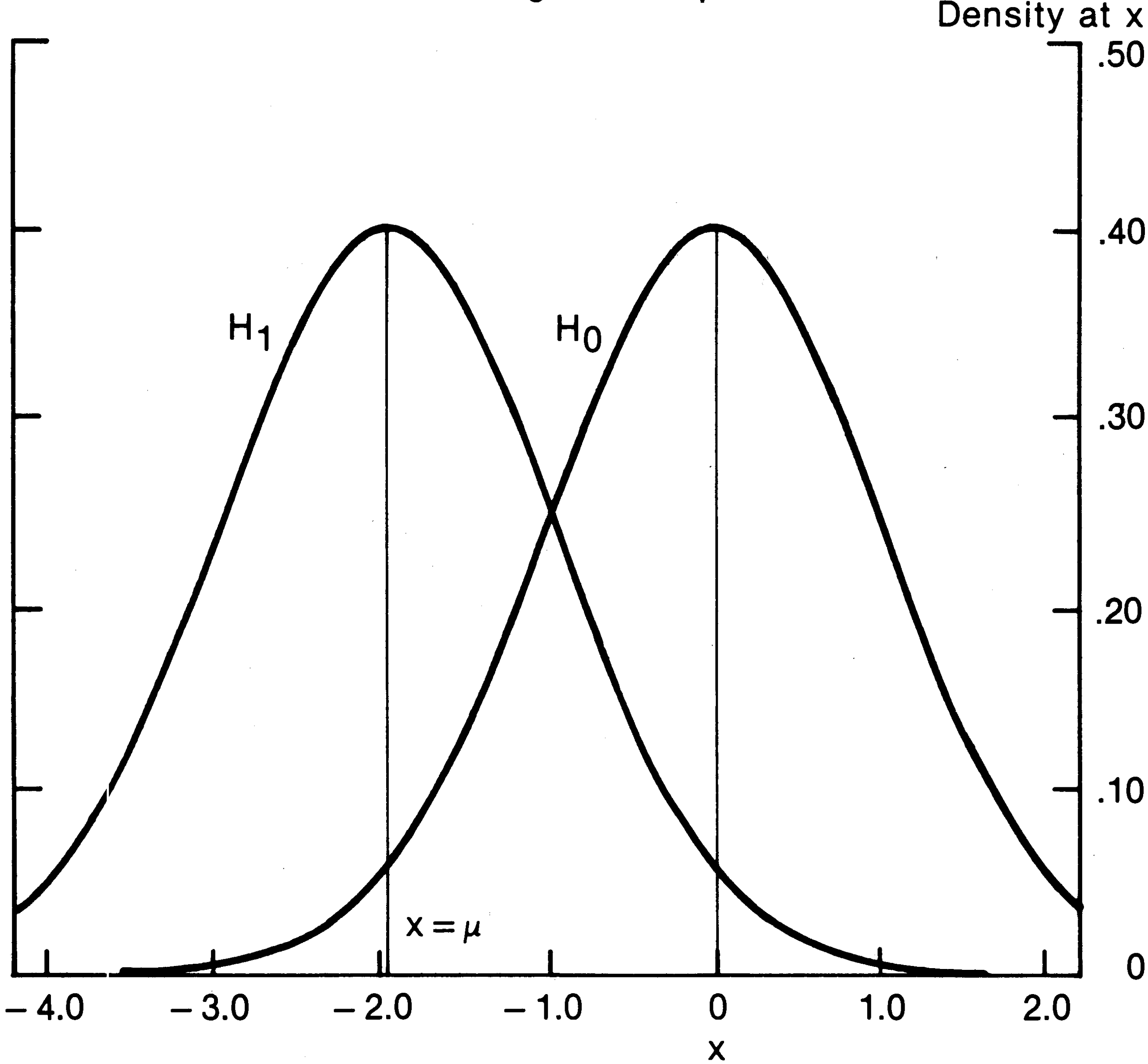
In general, response surfaces appear quite valuable in summarizing Pesaran's (1974) experimental results. Simple, well-determined response surfaces are obtained using only eight experiments for their estimation. Further, those response surfaces satisfactorily predict the outcomes of eight additional experiments for which one of the experimental design parameters (R^2) is considerably different from its value in all the experiments used for estimation, indicating the robust nature of those response surfaces. Even so, the response surfaces are limited by the relatively small number of experiments and the restrictiveness of the parameter space. Such specificity could be reduced markedly by performing more experiments over a more broadly defined parameter space (e.g., one including lagged dependent variables, several non-overlapping regressors, $k_0 \neq k_1$, and autocorrelated exogenous regressors), thereby allowing more sophisticated and (hopefully) reliable inferences to be drawn.

6. Adjustments for the estimated probability of type I error

In comparing the properties of the Cox and F statistics, it is only reasonable to consider both the probability of type I error and the probability of type II error. The estimated probability of type I error for D_0 significantly exceeds .05 in several instances (predominantly for smaller sample sizes and smaller values of r^2), whereas the estimated probability of type I error for f_2 never does. To account for both types of error, Pesaran (1974, p. 163) uses a simple linear loss function in which type I errors receive the same (or even twice as much) weight as type II errors, concluding that "the [Cox] test is preferable to the F-test when the sample size is small ($n \leqq 40$) and the correlation between the competing set [sic] of explanatory variables is large". Pesaran's conclusions may not depend so much upon the asymptotic nature of the Cox test (vs. the exact nature of the F test) as upon the comparison of estimated type I and type II errors for which the density of the statistic (for the type I error) at its critical values is a small fraction of the density of the statistic (for the type II error) at one of the critical values.

To see that, consider the extremely simplified example in which the test statistic τ is exactly distributed as $N(0,1)$ under H_0 and $N(\mu,1)$ under H_1 and the critical value (to be chosen) for a symmetric two-tailed test is δ_0 or δ_1 ($\delta_0 > \delta_1 > 0$) with δ_0 slightly greater than δ_1. The mean μ is non-zero and is assumed negative without loss of generality. Further, suppose that, for δ_0, power is in an intermediate range, so $\mu \approx -\delta_0$. Simplifying the example even more (but not appreciably weakening the argument), suppose that $\mu = -\delta_0 = -1.96$ (see Figure 2). At μ, the densities under H_0 and H_1 are approximately .058 and .399 respectively. Noting that the probability of τ being greater than +1.96 is negligible under H_1 and that the test is symmetric, it follows that a test using δ_1 would be preferred to one using δ_0, even if the probability of type I error were weighted three times as heavily as the probability of type II error. In the context of the Cox and F tests, because Pesaran was comparing statistics for which the probability of type I error was small ($\approx$ .05) and the power (and hence the probability of type II error) was usually between .10 and .90, his criterion might well have favoured the test statistic with the larger probability of type I error (i.e., the Cox statistic), even had the two test statistics had identical distributions but critical values corresponding to different probabilities of type I error.

Figure 2
The density of τ under H_0 and H_1



Ideally, one would choose critical values for D_0 and f_2 giving identical probabilities of type I error and use those critical values when estimating the finite sample powers of D_1 and f_3, but that is not feasible because the finite sample distribution of D_0 is not known. As an alternative, one could solve for the critical value of f_2 which would give it the same probability of type I error (α_f, say) as the estimated probability of type I error for D_0 (s_D, say).[21] (Table III lists those critical values for T = 20 and T = 40.) Then, using those critical values, the power of f_3 could be estimated for each Monte Carlo experiment. However, because that approach would depend upon knowing the values of the dependent variable and regressors for all replications of all experiments (information which is not currently available), a somewhat inferior method is adopted, namely, calculating the power of f_3 using its asymptotic non-centrality parameter (λ_3), evaluated at the adjusted critical value. Those (adjusted) powers for f_3 and the estimated powers of D_1 are presented in the last two columns of Table III and indicate much smaller differences between the powers of the tests based on D_1 and f_3 than are apparent in Pesaran's (1974, p. 162) Table 1. Even so, f_3 does appear slightly more powerful than D_1 at higher powers, and vice versa at lower powers.

7. Concluding remarks

This paper describes and implements an approach for obtaining numerical-analytical formulae (response surfaces) which integrate existing analytical knowledge with experimental results. Response surfaces can help summarize and interpret Monte Carlo simulations, and may reasonably approximate the unknown finite sample conditional probability formula for

[21]Cf. Sargan (1976, pp. 444ff) who suggests how to improve the efficiency of the estimated probability by using a control variate.

Table III. The powers of the F test (f_3) with non-centrality λ_3 when evaluated at critical values corresponding to $\alpha_f \approx .05$ and $\alpha_f \approx s_D$, α_f being the size of the F test and s_D the estimated size of the Cox test.

			$\alpha_f \approx .05$			$\alpha_f \approx s_D$				estimated finite sample power of D_1
Sample size	r^2	λ_3	critical value	α_f	asymptotic power of f_3	critical value	α_f	s_D	asymptotic power of f_3	
T=20	.90	8.0	4.45	.05003	.76755	3.4175	.08198	.082 (.012)[a]	.84731	.818 (.017)
	.95	4.0	4.45	.05003	.45798	3.3225	.08597	.086 (.013)	.57239	.628 (.022)
	.99	.8	4.45	.05003	.12934	3.6800	.07202	.072 (.012)	.16823	.196 (.018)
T=40	.90	16.0	4.10	.05014	.98435	3.3325	.07600	.076 (.012)	.99237	.968 (.008)
	.95	8.0	4.10	.05014	.79740	3.2400	.08002	.080 (.012)	.86375	.830 (.017)
	.99	1.6	4.10	.05014	.21897	3.8925	.05601	.056 (.010)	.23383	.306 (.021)

Note: a. Simulation standard errors of estimated probabilities are in parentheses.

the relationship(s) of interest and DGP considered. To evaluate the closeness of approximation, this paper presents a framework for inference about response surfaces. Illustrating this methodology, Pesaran's (1974) Monte Carlo simulations of the finite sample rejection frequencies of the Cox and F tests are re-examined. Further, adjustments are made in order that the estimated finite sample powers of various tests may be compared with each other on a more equal footing in those instances where the estimated finite sample probability of type I error for one of the tests differs significantly from the nominal value. Although this paper is concerned with tests in particular, the methodology is much more general: similar techniques have been successful in analysing a wide variety of estimators, both consistent and inconsistent, linear and non-linear (cf. Engle, Hendry, and Trumble (1985), Campos (1986), Hendry (1984), and references therein). Although Monte Carlo experimentation does not replace analysis, the two can complement each other effectively to provide convenient formulae for interpreting empirical findings.

Appendix A: The distributions of D_0, f_2, D_1, and f_3 under H_0: in general and for Pesaran's (1974) model in particular

This appendix derives the distributions of the test statistics of interest, using the formulae in Ericsson (1983). Further, for Pesaran's (1974) data generation process and relationships of interest, formulae explicitly in terms of his design parameters are found. The notation is the same as that in Ericsson (1983) (and that above) except that the econometric sample size is denoted by T (not n), the regression coefficients are β_i (not a_i), and N is the matrix $Z(Z'Z)^{-1}Z'$ (not the number of replications).

From Ericsson (1983),

$$D_0 \;\tilde{a}\; N(0,1) \tag{A1}$$

and

$$f_2 \;\tilde{a}\; F(k_2-k_0,\ T-k_2,\ 0) \tag{A2}$$

under H_0. Further,

$$D_0 \;\tilde{a}\; N(\mu_0,\omega_0) \tag{A3}$$

under H_1, where

$$\mu_0 = -(\zeta_0+\zeta_1)/(2\sigma_1\sqrt{\zeta_2}) \ ,$$

$$\omega_0 = \{(\zeta_0+\zeta_1)^2/(4\zeta_2)\}\cdot \mathrm{var}(q^*)/\sigma_1^2 \ , \quad \text{and}$$

$$\mathrm{var}(q^*)/\sigma_1^2 = 4(\zeta_0+\zeta_1-\zeta_2)/(\zeta_0+\zeta_1)^2 - 4(\zeta_2-\zeta_3)/\{(\zeta_0+\zeta_1)\zeta_2\} + (\zeta_3-\zeta_4)/\zeta_2^2 \ , \tag{A4}$$

with

$$\zeta_i = \phi_i - \phi_{i+1} \qquad (i \geqq 0) \ ,$$

$$\phi_0 = \mathrm{plim}_1\ \delta' X_1^{+\prime} N X_1^{+}\delta/T \ ,$$

$$\phi_i = \mathrm{plim}_1\ \delta' X_1^{+\prime}(Q_0Q_1)^{i-1}Q_0X_1^{+}\delta/T \qquad (i \geqq 1) \ ,$$

$$\delta = \beta_1^{+}\sqrt{T} \qquad (\delta \text{ a non-zero constant}) \ ,$$

$$Q_i = NP_i = NX_i(X_i'NX_i)^{-1}X_i'N \qquad (i = 0,\ 1,\ 2) \ , \quad \text{and}$$

$$N = Z(Z'Z)^{-1}Z' \ , \tag{A5}$$

Z being the T×m matrix of observations on m instruments, which is equal to the entire set of regressors X_2 in Pesaran (1974). Further, under H_1,

$$f_2 = \frac{y'(Q_2-Q_0)y/(k_2-k_0)}{y'(I_T-P_2)'(I_T-P_2)y/(T-k_2)}$$

$$= \frac{u'(Q_2-Q_0)u/(k_2-k_0)}{u'(I_T-P_2)'(I_T-P_2)u/(T-k_2)} \tag{A6}$$

where

$$u = X_1^+\beta_1^+ + u_1$$

and

$$u_1 \sim N(0, \sigma_1{}^2 \cdot I_T) \quad .$$

Hence

$$u'(Q_2-Q_0)u/\{\sigma_1{}^2(k_2-k_0)\} \mathrel{\underset{\sim}{a}} \chi^2(k_2-k_0,\lambda_2)/(k_2-k_0) \tag{A7}$$

and

$$u'(I_T-P_2)'(I_T-P_2)u/\{\sigma_1{}^2(T-k_2)\} \mathrel{\underset{\sim}{a}} \chi^2(T-k_2,0)/(T-k_2) \tag{A8}$$

where $\lambda_2 = \zeta_0/\sigma_1{}^2$. These random variates are independent asymptotically, so their ratio f_2 is asymptotically distributed as a singly non-central F,

$$f_2 \mathrel{\underset{\sim}{a}} F(k_2-k_0, T-k_2, \lambda_2) \quad . \tag{A9}$$

Assuming X_2 fixed in repeated samples, or at least assuming X_2 and u_1 independent and X_2 to be conditioned upon (see Schmidt (1976, pp. 93-94ff)), then

$$u \sim N(X_1^+\beta_1^+, \sigma_1{}^2 \cdot I_T) \quad ; \tag{A10}$$

and, with $Z = X_2$, the results in (A7), (A8), and (A9) are exact (i.e., with $\underset{\sim}{a}$ substituted by ~), with

$$\lambda_2^* = \beta_1^{+\prime}X_1^{+\prime}(Q_2-Q_0)X_1^+\beta_1^+/\sigma_1{}^2 \tag{A11}$$

replacing λ_2 in (A7) and (A9).

In Monte Carlo studies, it is often more convenient (and less expensive computationally) to evaluate (for instance) both D_0 and D_1 under H_0, than to evaluate D_0 under H_0 and H_1. Effectively, D_1 under H_0 behaves like D_0 under H_1 because each statistic is evaluated under a non-nested alternative. Using the formulae in (A3)-(A5), it readily follows that

$$D_1 \stackrel{a}{\sim} N(\mu_1, \omega_1) \tag{A12}$$

under H_0, where

$$\mu_1 = -(\zeta_{(0)}+\zeta_{(1)})/(2\sigma_0\sqrt{\zeta_{(2)}}) ,$$

$$\omega_1 = \{(\zeta_{(0)}+\zeta_{(1)})^2/(4\zeta_{(2)})\}\cdot \mathrm{var}(q^{(*)})/\sigma_0{}^2 , \tag{A13}$$

and

$$\begin{aligned}\mathrm{var}(q^{(*)})/\sigma_0{}^2 = & \; 4(\zeta_{(0)}+\zeta_{(1)}-\zeta_{(2)})/(\zeta_{(0)}+\zeta_{(1)})^2 \\ & - 4(\zeta_{(2)}-\zeta_{(3)})/\{(\zeta_{(0)}+\zeta_{(1)})\zeta_{(2)}\} \\ & + (\zeta_{(3)}-\zeta_{(4)})/\zeta_{(2)}{}^2 ,\end{aligned} \tag{A14}$$

with

$$\begin{aligned}\zeta(i) &= \phi(i) - \phi(i+1) \quad (i \geq 0) , \\ \phi(_0) &= \mathrm{plim}_0 \; \delta^{*\prime}X_0^{+\prime}NX_0^{+}\delta^*/T , \\ \phi(i) &= \mathrm{plim}_0 \; \delta^{*\prime}X_0^{+\prime}(Q_1Q_0)^{i-1}Q_1X_0^{+}\delta^*/T \quad (i \geq 1) , \text{ and} \\ \delta^* &= \beta_0^{+}\sqrt{T} \qquad (\delta^* \text{ a non-zero constant}) .\end{aligned} \tag{A15}$$

That is, the role of the parameters and variables in the model H_0 is exchanged with that of the parameters and variables in the model H_1. (Superfluous parentheses around subscripts and superscripts denote that those roles have been exchanged in the definition of the subscripted or superscripted variable.) Likewise, the distribution of f_3 under H_0 is

$$f_3 \stackrel{a}{\sim} F(k_2-k_1, T-k_2, \lambda_3) \tag{A16}$$

where $\lambda_3 = \zeta(_0)/\sigma_0{}^2$. If $Z = X_2$ and X_2 is fixed, f_3 is exactly distributed as an $F(k_2-k_1, T-k_2, \lambda_3^*)$ where

$$\begin{aligned}\lambda_3^* &= \beta_0^{+\prime}X_0^{+\prime}(Q_2-Q_1)X_0^{+}\beta_0^{+}/\sigma_0{}^2 \\ &= \delta^{*\prime}X_0^{+\prime}(Q_2-Q_1)X_0^{+}\delta^*/(\sigma_0{}^2T)\end{aligned} \tag{A17}$$

(and note that $\lambda_3 = \mathrm{plim}_0 \; \lambda_3^*$).

In Pesaran's (1974) model, the formulae above may be simplified because (i) $Z = X_2$ and (ii) $X_2 = (X_0^{+} : X : X_1^{+})$. Letting $W_i = X_i(X_i{}'X_i)^{-1}X_i{}'$ and $M_i = I_T-W_i$ (i=0,1,2), $Z = X_2$ implies that

$Q_i = P_i = W_i$ and $Q_2X_1^\dagger = NX_1^\dagger = W_2X_1^\dagger = X_1^\dagger$ (i=0,1). The expressions above for $\phi_{(i)}$ and λ_3^* simplify to

$$\phi_{(0)} = \text{plim}_0\ \delta^{*\prime}X_0^{+\prime}X_0^+\delta^*/T \quad ,$$

$$\phi_{(i)} = \text{plim}_0\ \delta^{*\prime}X_0^{+\prime}(W_1W_0)^{i-1}W_1X_0^+\delta^*/T \qquad (i \geq 1) \quad , \tag{A18}$$

and

$$\lambda_3^* = \delta^{*\prime}X_0^{+\prime}(I_T-W_1)X_0^+\delta^*/(\sigma_0{}^2T)$$

$$= \delta^{*\prime}X_0^{+\prime}M_1X_0^+\delta^*/(\sigma_0{}^2T) \quad . \tag{A19}$$

Adopting the following notation:

$$m_{++} = \text{plim}_0\ X_0^{+\prime}X_0^+/T \quad ,$$

$$m_{1+} = \text{plim}_0\ X_1{}'X_0^+/T \quad ,$$

$$m_{11} = \text{plim}_0\ X_1{}'X_1/T \quad ,$$

$$m_{10} = \text{plim}_0\ X_1{}'X_0/T \quad ,$$

$$m_{00} = \text{plim}_0\ X_0{}'X_0/T \quad , \tag{A20}$$

$m_{01} = m_{10}{}'$, and $m_{+1} = m_{1+}{}'$, then $\phi_{(i)}$ may be written as

$$\phi_{(0)} = \delta^{*\prime}m_{++}\delta^*$$

$$\phi_{(i)} = \delta^{*\prime}m_{+1}m_{11}{}^{-1}(m_{10}m_{00}{}^{-1}m_{01}m_{11}{}^{-1})^{i-1}m_{1+}\delta^* \qquad (i \geq 1) \quad . \tag{A21}$$

Given the population moments of the regressors, $\zeta_{(i)}$ (i=0,...,4) may be calculated; and hence so can be μ_1, ω_1, and λ_3, knowing $\sigma_0{}^2$. From those, the (approximate) asymptotic powers of the tests based on f_3 and D_1 are derived. (See Appendix C for details.)

The formulae for $\phi_{(i)}$ and λ_3 may be simplified further by noting two particular features of the structure of X_2: that X_2 is a T×3 matrix with X_0^+, X, and X_1^+ all being T×1 vectors; and that X_0^+, X, and X_1^+ satisfy the properties given in (8) and (9) where $X_0^+ = (x_1, x_2, \ldots, x_T)'$, $X = (1,1,\ldots,1)'$, and $X_1^+ = (z_1, z_2, \ldots, z_T)'$. The relevant moment matrix is

$$T^{-1}E(X_2'X_2) = T^{-1}E\begin{bmatrix} X_0^{+\prime}X_0^{+} & X_0^{+\prime}X & X_0^{+\prime}X_1^{+} \\ X'X_0^{+} & X'X & X'X_1^{+} \\ X_1^{+\prime}X_0^{+} & X_1^{+\prime}X & X_1^{+\prime}X_1^{+} \end{bmatrix}$$

$$= E\begin{bmatrix} x_t^2 & x_t & x_t z_t \\ x_t & 1 & z_t \\ x_t z_t & z_t & z_t^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \gamma \\ 0 & 1 & 0 \\ \gamma & 0 & \gamma^2+1 \end{bmatrix} . \tag{A22}$$

It follows that

$$m_{++} = [1]$$

$$m_{1+} = \begin{bmatrix} 0 \\ \gamma \end{bmatrix} \qquad m_{10} = \begin{bmatrix} 0 & 1 \\ \gamma & 0 \end{bmatrix}$$

$$m_{11} = \begin{bmatrix} 1 & 0 \\ 0 & \gamma^2+1 \end{bmatrix} \qquad m_{00} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\sigma_0 = \{(\beta_0^{+})^2(1-R^2)/R^2\}^{1/2}$$

$$\gamma = \{r^2/(1-r^2)\}^{1/2} \tag{A23}$$

for positive γ. The four-fold product of matrices for $\phi(i)$ (appearing in the parentheses of the formulae for $\phi(i)$ in (A21)) simplifies to

$$\begin{bmatrix} 1 & 0 \\ 0 & \gamma^2/(\gamma^2+1) \end{bmatrix} , \tag{A24}$$

so $\phi(i) = (\delta^*)^2(r^2)^i$, $i=0,\dots,5$. Substituting into (A13) and (A14),

$$\mu_1 = -\delta^*(1-r^2)^{1/2}(1+r^2)/(2\sigma_0 r^2) \quad , \tag{A25}$$

$$\omega_1 = \{4r^4 + (1-r^4)(1+r^2)\}/(4r^6) \quad , \text{ and} \tag{A26}$$

$$\lambda_3 = (\delta^*)^2(1-r^2)/\sigma_0^2 \quad . \tag{A27}$$

Those formulae for μ_1, ω_1, and λ_3 were used to calculate the asymptotic powers of f_3 and D_1 under H_0 from Pesaran's model.

It is interesting to note that

$$\mu_1{}^2 \;=\; (\delta^*)^2(1-r^2)(1+r^2)^2/(4\sigma_0{}^2r^4) \;; \tag{A28}$$

so that for r^2 close to unity,

$$\mu_1{}^2 \;\approx\; (\delta^*)^2(1-r^2)/\sigma_0{}^2 \;=\; \lambda_3 \tag{A29}$$

and

$$\omega_1 \;\approx\; 1 \;. \tag{A30}$$

Hence $D_1{}^2$ and f_3 have nearly identical asymptotic distributions; and, given a symmetric two-tailed test for D_1, D_1 and f_3 should have roughly the same power. Even when r^2 is not so close to unity, the approximations in (A29) and (A30) are still reasonable, as can be seen in Table A.I below for Pesaran's published results.

Table A.I. A comparison of the parameters in the approximate asymptotic distribution of D_1 (ω_1 and μ_1) with the one in the asymptotic distribution of f_3 (λ_3). R^2 for the DGP is .80.

		Parameter		
Sample size	r^2	ω_1	$\mu_1{}^2$	λ_3
n=20	.90	1.23	8.91	8.00
	.95	1.11	4.21	4.00
	.99	1.02	.81	.80
n=40	.90	1.23	17.83	16.00
	.95	1.11	8.43	8.00
	.99	1.02	1.62	1.60
n=80	.90	1.23	35.65	32.00
	.95	1.11	16.85	16.00
	.99	1.02	3.23	3.20

Appendix B: Proof of the asymptotic normality of the empirical logistic function.

It is to be shown that

$$\phi(s,\pi) \equiv A^{1/2}\{\ln\left[\frac{s-(2N)^{-1}}{1-s-(2N)^{-1}}\right] - \ln\left[\frac{\pi}{1-\pi}\right]\} \underset{A}{\sim} N(0,1). \quad (B1)$$

S is the number of "successes" out of a total of N observations ($N > 1$) on a binary response variable with probability of "success" equal to π ($0 < \pi < 1$), s ($\equiv S/N$) is the fraction of "successes", $A \equiv S(N-S)/(N-1)$, and $\underset{A}{\sim}$ denotes "converges in distribution to, as $N \to \infty$". Results in Cox (1970, pp. 30-34, 41-42, 78-79), Mann and Wald (1943), and Cramér (1946, pp. 254, 299-300) are used extensively.

Noting that $E(S) = N\pi$ and $\text{var}(S) = N\pi(1-\pi)$, define U such that

$$S = N\pi + U\sqrt{N} \quad . \quad (B2)$$

Hence $E(U) = 0$ and $\text{var}(U) = \pi(1-\pi)$. The term in braces in (B1) is

$$\{\cdot\} = \ln\left[\frac{\pi + (U/\sqrt{N}) - (2N)^{-1}}{\pi}\right] - \ln\left[\frac{1-\pi-(U/\sqrt{N})-(2N)^{-1}}{1-\pi}\right]$$

$$= U/\{\pi(1-\pi)\sqrt{N}\} + O_p(N^{-1}) \quad (B3)$$

(cf. Cox (1970, p. 33, (3.11)) with (his) a=-.5). The variable A in (B1) is

$$A = Ns(1-s) + O_p(N^0)$$

$$= N\pi(1-\pi) + O_p(N^{1/2}) \quad , \quad (B4)$$

so

$$A^{1/2} = N^{1/2}\{\pi(1-\pi) + O_p(N^{-1/2})\}^{1/2}$$

$$= \{N\pi(1-\pi)\}^{1/2} + O_p(N^0) \quad . \quad (B5)$$

Substituting (B3) and (B5) into (B1),

$$\phi(s,\pi) = U/\{\pi(1-\pi)\}^{1/2} + O_p(N^{-1/2}) \quad . \quad (B6)$$

Since S is the sum of N independent and identically distributed random variates, each with mean π and variance $\pi(1-\pi)$, and $U = (S - N\pi)/\sqrt{N}$, then $U/\{\pi(1-\pi)\}^{1/2}$ converges in distribution to $N(0,1)$ as $N \to \infty$ by the Lindeberg-Lévy variant of the central limit theorem (Cramér (1946, p. 215)). Hence (B6) converges in distribution to $N(0,1)$ as $N \to \infty$. ||

Appendix C: The calculation of probabilities associated with non-central χ^2 and F distributions, and with the normal distribution

The discrepancies between the theoretical and estimated finite sample powers of f_3 might have arisen from the approximation used to calculate the probabilities associated with the non-central F distribution. This appendix, which closely follows the substance and notation of Kendall and Stuart (1973, pp. 237-239, 241, 262-264), shows that the errors involved do not appear large enough to explain those discrepancies. For completeness and clarity in presentation, a related approximation (of the non-central χ^2 distribution by a central one) is given first. See Johnson (1959) on other approximations to the non-central χ^2 and F distributions.

The probability of a non-central χ^2 random variate exceeding $\chi^2_\alpha(\nu,0)$, the 100(1-α) per cent point of the central χ^2 distribution, is

$$P = \int_{\chi^2_\alpha(\nu,0)}^{\infty} d\chi^2(\nu,\lambda) \tag{C1}$$

where $\chi^2(\nu,\lambda)$ is the non-central χ^2 distribution with degrees of freedom ν and non-centrality parameter λ, as given in Kendall and Stuart (1973, p. 238, equation (24.18)). Equating the first two moments of the non-central χ^2 random variate to those of $\rho\chi^2(\cdot,0)$ (a central χ^2 random variate multiplied by a factor of proportionality ρ, to be determined), then

$$P \approx \int_{\chi^2_\alpha(\nu,0)/\rho}^{\infty} d\chi^2(\nu^*,0) \tag{C2}$$

where

$$\nu^* = \frac{(\nu + \lambda)^2}{\nu + 2\lambda} \tag{C3}$$

and

$$\rho = \frac{(\nu + 2\lambda)}{\nu + \lambda} \quad . \tag{C4}$$

Having solved for ν^* and the lower limit point of the integral in (C2), the integral itself may be calculated numerically, e.g., with the NAg (1977) routine G01BCF, linearly interpolating for P if ν^* is not integral.

Table C.I compares the approximation given in (C2) with exact values of the integral in (C1) for the values of ν and λ appearing in Patnaik (1949, p. 207, Table 1). Patnaik's own approximation to the exact probability involves not only linearly interpolating P for non-integral ν^*, but also interpolating P for the value of the lower limit of the integral in (C2). (C2) (using G01BCF) appears to be a better approximation to the exact probability than Patnaik's approximation, but only marginally so, with both approximating the exact probability quite well.

The probability of a singly non-central F random variate exceeding $F_\alpha(\nu_1,\nu_2,0)$, the 100(1-α) per cent point of the (central) F distribution, is

$$P = \int_{F_\alpha(\nu_1,\nu_2,0)}^{\infty} dF(\nu_1,\nu_2,\lambda) \tag{C5}$$

where $F(\nu_1,\nu_2,\lambda)$ is the singly non-central F distribution with the degrees of freedom in the numerator and denominator of the F-ratio being ν_1 and ν_2, respectively; and λ is the non-centrality parameter of the χ^2 random variate in the numerator, the non-centrality parameter of the χ^2 random variate in the denominator being zero (see Kendall and Stuart (1973, p. 262, equation (24.105))). Equating the first two moments of the non-central χ^2 random variate in the numerator to those of a central χ^2 random variate (as above), then

$$P \approx \int_{\{\nu_1/(\nu_1+\lambda)\}F_\alpha(\nu_1,\nu_2,0)}^{\infty} dF(\nu^*,\nu_2,0) \tag{C6}$$

where

$$\nu^* = \frac{(\nu_1 + \lambda)^2}{\nu_1 + 2\lambda} . \tag{C7}$$

From ν_1, ν_2, λ, and α, the approximation to P in (C6) may be calculated numerically, e.g., with the NAg (1977) routine G01BBF, linearly interpolating for P in (C6) if ν^* is not integral.

Table C.II compares the approximation given in (C6) with exact values of the integral in (C5) for values of ν_1, ν_2, and λ appearing in Patnaik (1949, p. 222, Table 7). Patnaik's own approximation to the exact probability involves both linearly interpolating P for non-integral ν^* and interpolating P for the value of the lower limit of the integral in (C6). The latter interpolation only slightly affects the values obtained, and both approximations perform well over a wide range of powers.

The probability of a normal variate with mean μ and variance σ^2 being less than a certain critical value z is

$$(2\pi)^{-1/2} \int_{-\infty}^{z^*} \exp(-u^2/2)\, du \qquad \text{(C8)}$$

where $z^* = (z-\mu)/\sigma$. This integral is calculated directly with the NAg (1977) routine S15ABF. The integral corresponding to the upper tail is calculated in a similar manner with NAg (1977) routine S15ACF.

Table C.I. A comparison of two approximations to the non-central $\chi^2(\nu,\lambda)$ distribution with points on the distribution.

ν	λ	$\chi^2_\alpha(\nu,0)$	Probability		
			Patnaik's approximation	Approximation from (C2)	Exact Probability
4	4	1.765	.0399	.0426	.0500
	4	10.000	.7191	.7180	.7118
	4	17.309	.9492	.9479	.9500
	4	24.000	.9913	.9909	.9925
	10	10.000	.3178	.3190	.3148
7	1	4.000	.1621	.1634	.1628
	1	16.004	.9499	.9495	.9500
	16	10.257	.0430	.0439	.0500
	16	24.000	.5947	.5945	.5898
	16	38.970	.9482	.9476	.9500
12	6	24.000	.8187	.8178	.8174
	18	24.000	.2936	.2926	.2901
16	8	30.000	.7895	.7895	.7880
	8	40.000	.9626	.9626	.9632
	32	30.000	.0590	.0594	.0609
	32	60.000	.8329	.8326	.8316
24	24	36.000	.1556	.1556	.1567
	24	48.000	.5333	.5333	.5296
	24	72.000	.9656	.9656	.9667

Table C.II. A comparison of two approximations to the singly non-central $F(\nu_1, \nu_2, \lambda)$ distribution with points on the distribution.

ν_1	ν_2	λ	$F_\alpha(\nu_1, \nu_2, 0)$	Probability: Patnaik's approximation	Probability: Approximation from (C6)	Probability: Exact probability
3	10	4	3.708	.752	.750	.745
		4	6.552	.919	.919	.918
		16	3.708	.203	.203	.206
		16	6.552	.520	.520	.517
3	20	4	3.098	.706	.707	.700
		4	4.938	.889	.888	.887
		16	3.098	.119	.119	.126
		16	4.938	.350	.350	.347
5	10	6	3.326	.731	.733	.731
		6	5.636	.913	.914	.914
		24	3.326	.157	.156	.158
		24	5.636	.463	.463	.461
5	20	6	2.711	.665	.668	.664
		6	4.103	.869	.871	.870
		24	2.711	.064	.064	.069
		24	4.103	.244	.244	.245
8	10	9	3.072	.715	.715	.714
		9	5.057	.909	.909	.908
		36	3.072	.117	.117	.119
		36	5.057	.409	.409	.408
8	30	9	2.266	.581	.581	.578
		9	3.173	.815	.815	.813
		36	2.266	.014	.014	.017
		36	3.173	.085	.085	.088

Appendix D: Pesaran's experimental results for the multiple correlation coefficient of the DGP (R^2) being .80 and .85.

Table D.I. Estimated probabilities for the Cox and F tests when R^2 = .80 .

		Probability of accepting H_0 and rejecting H_1		Probability of rejecting H_0 and accepting H_1		Probability of rejecting both H_0 and H_1		Probability of accepting both H_0 and H_1		Type I error: probability of rejecting H_0		Type II error: probability of accepting H_1	
T	r^2	D test	F test	D test	F test	D test	F test	D test	F test	D test	F test	D test	F test
20	.90	.786 (.018)[a]	.634 (.022)	.050 (.010)	.026 (.007)	.032 (.008)	.022 (.007)	.132 (.015)	.318 (.021)	.082 (.012)	.048 (.010)	.182 (.017)	.344 (.021)
	.95	.600 (.022)	.404 (.022)	.058 (.010)	.024 (.007)	.028 (.007)	.032 (.008)	.314 (.021)	.540 (.022)	.086 (.013)	.056 (.010)	.372 (.022)	.564 (.022)
	.99	.174 (.017)	.096 (.013)	.050 (.010)	.026 (.007)	.022 (.007)	.026 (.007)	.754 (.019)	.852 (.016)	.072 (.012)	.052 (.010)	.804 (.018)	.878 (.015)
40	.90	.916 (.012)	.902 (.013)	.024 (.007)	.018 (.006)	.052 (.010)	.030 (.008)	.008 (.004)	.050 (.010)	.076 (.012)	.048 (.010)	.032 (.008)	.068 (.011)
	.95	.802 (.018)	.738 (.020)	.052 (.010)	.026 (.007)	.028 (.007)	.026 (.007)	.118 (.014)	.210 (.018)	.080 (.012)	.052 (.010)	.170 (.017)	.236 (.019)
	.99	.280 (.020)	.202 (.018)	.030 (.008)	.016 (.006)	.026 (.007)	.026 (.007)	.664 (.019)	.756 (.019)	.056 (.010)	.042 (.009)	.694 (.021)	.772 (.019)
80	.90	.932 (.011)	.954 (.009)	.000 (.000)	.000 (.000)	.068 (.011)	.046 (.009)	.000 (.000)	.000 (.000)	.068 (.011)	.046 (.009)	.000 (.000)	.000 (.000)
	.95	.948 (.010)	.946 (.010)	.018 (.006)	.012 (.005)	.032 (.008)	.022 (.007)	.002 (.002)	.020 (.006)	.050 (.010)	.034 (.008)	.020 (.006)	.032 (.008)
	.99	.456 (.022)	.384 (.022)	.030 (.008)	.018 (.006)	.020 (.006)	.020 (.006)	.494 (.022)	.578 (.022)	.050 (.010)	.038 (.009)	.524 (.022)	.596 (.022)

Note: a. Simulation standard errors of estimated probabilities are in parentheses.

Table D.II. Estimated probabilities for the Cox and F tests when R^2 = .85 .

		Probability of accepting H_0 and rejecting H_1		Probability of rejecting H_0 and accepting H_1		Probability of rejecting both H_0 and H_1		Probability of accepting both H_0 and H_1		Type I error: probability of rejecting H_0		Type II error: probability of accepting H_1	
T	r^2	D test	F test	D test	F test	D test	F test	D test	F test	D test	F test	D test	F test
	.90	.854 (.016)[a]	.768 (.019)	.044 (.009)	.026 (.007)	.042 (.009)	.022 (.007)	.060 (.011)	.184 (.017)	.086 (.013)	.048 (.010)	.104 (.014)	.210 (.018)
20	.95	.688 (.021)	.540 (.022)	.054 (.010)	.024 (.007)	.034 (.008)	.032 (.008)	.224 (.019)	.404 (.022)	.088 (.013)	.056 (.010)	.278 (.020)	.428 (.022)
	.99	.254 (.019)	.116 (.014)	.050 (.010)	.026 (.007)	.022 (.007)	.026 (.007)	.674 (.021)	.832 (.017)	.072 (.012)	.052 (.010)	.724 (.020)	.858 (.016)
	.90	.922 (.012)	.940 (.011)	.006 (.003)	.004 (.003)	.072 (.012)	.048 (.010)	.000 (.000)	.008 (.004)	.078 (.012)	.052 (.010)	.006 (.003)	.012 (.005)
40	.95	.886 (.014)	.848 (.016)	.030 (.008)	.016 (.006)	.032 (.008)	.024 (.007)	.052 (.010)	.112 (.014)	.062 (.011)	.040 (.009)	.082 (.012)	.128 (.015)
	.99	.378 (.022)	.276 (.020)	.048 (.010)	.024 (.007)	.018 (.006)	.022 (.007)	.556 (.022)	.678 (.021)	.066 (.011)	.046 (.009)	.604 (.022)	.702 (.020)
	.90	.938 (.011)	.944 (.010)	.000 (.000)	.000 (.000)	.062 (.011)	.056 (.010)	.000 (.000)	.000 (.000)	.062 (.011)	.056 (.010)	.000 (.000)	.000 (.000)
80	.95	.926 (.012)	.948 (.010)	.004 (.003)	.008 (.004)	.070 (.011)	.042 (.009)	.000 (.000)	.002 (.002)	.074 (.012)	.050 (.010)	.004 (.003)	.010 (.004)
	.99	.528 (.022)	.476 (.022)	.034 (.008)	.030 (.008)	.022 (.007)	.024 (.007)	.416 (.022)	.470 (.022)	.056 (.010)	.054 (.010)	.450 (.022)	.500 (.022)

Note: a. Simulation standard errors of estimated probabilities are in parentheses.

References

Aneuryn-Evans, G. and A. Deaton (1980) "Testing linear versus logarithmic regression models", Review of Economic Studies, 47, 1, 275-291.

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. (1986) "Finite sample properties of the instrumental variables estimator for dynamic simultaneous equation sub-systems with ARMA disturbances", Journal of Econometrics, forthcoming.

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.

Cramér, H. (1946) Mathematical Methods of Statistics, Princeton, Princeton University Press.

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.

Engle, R.F., D.F. Hendry, and D. Trumble (1985) "Small-sample properties of ARCH estimators and tests", Canadian Journal of Economics, 18, 1, 66-93.

Ericsson, N.R. (1982a) "Testing linear versus logarithmic regression models: a comment", Review of Economic Studies, 49, 3, 477-481.

Ericsson, N.R. (1982b) Testing Non-nested Hypotheses in Systems of Linear Dynamic Economic Relationships, Ph.D. thesis, University of London.

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, forthcoming.

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.

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. 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.

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.

Hendry, D.F. (1973) "On asymptotic theory and finite sample experiments", Economica, 40, 158, 210-217.

Hendry, D.F. (1979) "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) "Monte Carlo experimentation in econometrics", Discussion Paper No. 8035, CORE, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, November.

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

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 Publishing Company, volume II, 937-976.

Hendry, D.F., M. Morgan, and F. Srba (1982) "GIVE guide", mimeo, London School of Economics.

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 School of Economics, October.

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.

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 Book Company.

Kendall, M.G. (1946) The Advanced Theory of Statistics, Volume II, London, Charles Griffin.

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

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.

Mann, H.B. and A. Wald (1943) "On stochastic limit and order relationships", Annals of Mathematical Statistics, 14, 3, 217-226.

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 (1986) "The encompassing principle and its application to testing non-nested hypotheses", Econometrica, forthcoming.

Nicholls, D.F. and A.R. Pagan (1983) "Heteroscedasticity in models with lagged dependent variables", Econometrica, 51, 4, 1233-1242.

Nickell, S. (1981) "Biases in dynamic models with fixed effects", Econometrica, 49, 6, 1417-1426.

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.

Pesaran, M.H. (1974) "On the general problem of model selection", Review of Economic Studies, 41, 2, 153-171.

Pesaran, M.H. (1981) "Comparison of local power of alternative tests of non-nested regression models", mimeo, Trinity College, Cambridge, June.

Pesaran, M.H. (1982) "Comparison of local power of alternative tests of non-nested regression models", Econometrica, 50, 5, 1287-1305.

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.

Rao, C.R. (1973) Linear Statistical Inference and Its Applications, New York, John Wiley and Sons, second edition.

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

Sargan, J.D. (1980) "Some approximations to the distribution of econometric criteria which are asymptotically distributed as chi-squared", Econometrica, 48, 5, 1107-1138.

Schmidt, P. (1976) Econometrics, New York, Marcel Dekker.

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

Sowey, E.R. (1973) "A classified bibliography of Monte Carlo studies in econometrics", Journal of Econometrics, 1, 4, 377-395.

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

Wilks, S.S. (1962) Mathematical Statistics, New York, John Wiley and Sons.

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.

International Finance Discussion Papers

IFDP NUMBER	TITLES	AUTHOR(s)
	1986	
276	Post-simulation Analysis of Monte Carlo Experiments: Interpreting Pesaran's (1974) Study of Non-nested Hypothesis Test Statistics	Neil R. Ericsson
275	A Method for Solving Systems of First Order Linear Homogeneous Differential Equations When the Elements of the Forcing Vector are Modelled as Step Functions	Robert A. Johnson
274	International Comparisons of Fiscal Policy: The OECD and the IMF Measures of Fiscal Impulse	Garry Schinasi
273	An Analysis of the Welfare Implications of Alternative Exchange Rate Regimes: An Intertemporal Model with an Application	Andrew Feltenstein David Lebow Anne Sibert
	1985	
272	Expected Fiscal Policy and the Recession of 1982	William H. Branson Arminio Fraga Robert A. Johnson
271	Elections and Macroeconomic Policy Cycles	Kenneth Rogoff Anne Sibert
270	Assertion Without Empirical Basis: An Econometric Appraisal of Monetary Trends in ... the United Kingdom by Milton Friedman and Anna J. Schwartz	David F. Hendry Neil R. Ericsson
269	Canadian Financial Markets: The Government's Proposal for Reform	Garry J. Schinasi
268	Was It Real? The Exchange Rate Interest Differential Relation, 1973-1984	Richard Meese Kenneth Rogoff
267	The U.K. Sector of the Federal Reserve's Multicountry Model: The Effects of Monetary and Fiscal Policies	Hali J. Edison
266	Optimal Currency Basket in a World of Generalized Floating: An Application to the Nordic Countries	Hali J. Edison Erling Vardal

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)
265	Money Demand in Open Economies: A Currency Substitution Model for Venezuela	Jaime Marquez
264	Comparing Costs of Note Issuance Facilities and Eurocredits	Rodney H. Mills
263	Some Implications of the President's Tax Proposals for U.S. Banks with Claims on Developing Countries	Allen B. Frankel
262	Monetary Stabilization Policy in an Open Economy	Marcus H. Miller
261	Anticipatory Capital Flows and the Behaviour of the Dollar	Arnold Kling
260	Simulating Exchange Rate Shocks in the MPS and MCM Models: An Evaluation	Arnold Kling
259	Trade Policy for the Multiple Product Declining Industry	Catherine Mann
258	Long Memory Models of Interest Rates, the Term Structure, and Variance Bounds Tests	Gary S. Shea
257	Currency Substitution and the New Divisia Monetary Aggregates: The U.S. Case	Jaime Marquez
256	The International Transmission of Oil Price Effects and OPEC's Pricing Policy	Jaime Marquez
255	U.S. Banks' Lending to Developing Countries: A Longer-Term View	Henry S. Terrell Rod Mills
254	Conditional Econometric Modelling: An Application to New House Prices in the United Kingdom	Neil R. Ericsson David F. Hendry
253	Low Pushing: Doctrine and Theory	William Darity, Jr.
	1984 (partial listing)	
252	Postwar Financial Policies in Taiwan, China	Robert F. Emery
251	Foreign Exchange Constraints and Growth Possibilities in LDCs	Jaime Marquez
250	The Determination of Front-end Fees on Syndicated Eurocurrency Credits	Rodney H. Mills Henry S. Terrell