COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 399

THE LIMITING DISTRIBUTION OF INCONSISTENT INSTRUMENTAL VARIABLES ESTIMATORS IN A CLASS OF STATIONARY STOCHASTIC SYSTEMS

David F. Hendry

August 19, 1975

The Limiting Distribution of Inconsistent Instrumental Variables Estimators in a Class of Stationary Stochastic Systems*

David F. Hendry
Cowles Foundation for Research in Economics,
Yale University and
London School of Economics
July 1975

I. The estimated specification of an econometric relationship can rarely be correct in practical applications and hence the resulting distributions of conventionally used estimators will not conform with those obtained under the assumption of a correct maintained hypothesis. The effects of such mis-specifications have not yet received the detailed consideration they deserve in view of their pervasiveness. Below we analyse the case of applying members of the class of Generalised Instrumental Variables Estimators (GIVE) which includes Ordinary and Two-Stage Least Squares (OLS and TSLS) to an equation where the instruments are not uncorrelated asymptotically with a non-spherical error, and also demonstrate the immense value of such asymptotic results in finite sample situations.

II. Consider a regression equation of the form

$$(1)\quad \underset{\sim}{W}\underset{\sim}{\alpha} = \underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta} = \underset{\sim}{y} - (\underset{\sim}{Y}b + \underset{\sim}{y}_1 c + \underset{\sim}{Z}d) = \underset{\sim}{u}$$

where $\underset{\sim}{W} = (\underset{\sim}{y}\ \underset{\sim}{X}) = (\underset{\sim}{y}\ \underset{\sim}{Y}\ \underset{\sim}{y}_1\ \underset{\sim}{Z})$ and $\underset{\sim}{\beta}' = (b\ c\ d)$ such that $\underset{\sim}{y}$ and $\underset{\sim}{Y}$ are endogenous, $\underset{\sim}{Z}$ is exogenous and $\underset{\sim}{y}_1$ denotes the lagged value of $\underset{\sim}{y}$. $\underset{\sim}{W}$ is $T\times(1 + m_1 + 1 + m_2)$ and there exists a set of m_3 excluded predetermined variables $\underset{\sim}{Z}^o$ such that all the parameters of (1) are overidentified ($m_3 \geq m_1 + 2$). Let $\underset{\sim}{Q} = (\underset{\sim}{y}_1\ \underset{\sim}{Z}\ \underset{\sim}{Z}^o)$ be the chosen set of Instruments and $\underset{\sim}{F} = (\underset{\sim}{W}\ \underset{\sim}{Z}^o)$ be the complete set of variables of relevance. Finally let

---

*Preliminary and Confidential; please do not quote without author's permission. I am indebted to Grayham E. Mizon and J. Denis Sargan for their advise and to Frank Srba for his invaluable help in programming the calculations. This work is part of the L.S.E. project on Econometric Methodology financed by a Grand from the Social Science Research Council; final preparation of the paper was done at the Cowles Foundation, Yale University with financial support from the National Science and Ford Foundations.

(2) $\underset{\sim}{u} = \rho \underset{\sim}{u}_{-1} + \underset{\sim}{\varepsilon}$ with $\underset{\sim}{\varepsilon} \sim NI(\underset{\sim}{0}, \sigma^2 \underset{\sim}{I})$.

If the investigator falsely assumes $\rho = 0$, the TSLS estimator of $\underset{\sim}{\beta}$ is

(3) $\tilde{\underset{\sim}{\beta}} = (\underset{\sim}{X}'\underset{\sim}{M}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{M}\underset{\sim}{y}$ where $\underset{\sim}{M} = \underset{\sim}{Q}(\underset{\sim}{Q}'\underset{\sim}{Q})^{-1}\underset{\sim}{Q}'$.

However, when $\rho \neq 0$, plim $T^{-1}\underset{\sim}{Q}'\underset{\sim}{u} \neq \underset{\sim}{0}$ so $\tilde{\underset{\sim}{\beta}}$ is inconsistent for $\underset{\sim}{\beta}$ and $\sqrt{T}(\tilde{\underset{\sim}{\beta}}-\underset{\sim}{\beta})$ is not distributed asymptotically as $N(\underset{\sim}{0},\sigma^2 \underset{\sim}{K}^{-1})$ where $\underset{\sim}{K} = \text{plim } T^{-1}\underset{\sim}{X}'\underset{\sim}{M}\underset{\sim}{X}$, thus invalidating inferences based on the estimates. In such a situation, we can derive the asymptotic distribution of $\sqrt{T}(\tilde{\underset{\sim}{\beta}}-\underset{\sim}{\beta}_{\ell})$ where plim $\tilde{\underset{\sim}{\beta}} = \underset{\sim}{\beta}_{\ell}$ as follows.

From (1) and (3),

(4) $(\tilde{\underset{\sim}{\beta}}-\underset{\sim}{\beta}) = \underset{\sim}{\Phi}\,\underset{\sim}{\phi}$ where $\underset{\sim}{\phi} = \frac{\underset{\sim}{Q}'\underset{\sim}{u}}{T}$, $\underset{\sim}{\Phi} = T(\underset{\sim}{X}'\underset{\sim}{M}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{Q}(\underset{\sim}{Q}'\underset{\sim}{Q})^{-1}$

and hence

(5) plim $(\tilde{\underset{\sim}{\beta}}-\underset{\sim}{\beta}) = (\underset{\sim}{\beta}_{\ell}-\underset{\sim}{\beta}) = \underset{\sim}{\Phi}_{\ell}\underset{\sim}{\phi}_{\ell} = \underset{\sim}{p}$ where $\underset{\sim}{\phi}_{\ell}$ = plim $T^{-1}\underset{\sim}{Q}'\underset{\sim}{u}$

and $\underset{\sim}{\Phi}_{\ell}$ = plim $\underset{\sim}{\Phi} = \underset{\sim}{K}^{-1}\underset{\sim}{L}$ for $\underset{\sim}{L}$ = plim $\underset{\sim}{X}'\underset{\sim}{Q}(\underset{\sim}{Q}'\underset{\sim}{Q})^{-1} = \underset{\sim}{R}\underset{\sim}{A}^{-1}$. Thus we have

(6) $\underset{\sim}{u} = \underset{\sim}{X}\underset{\sim}{p} + \underset{\sim}{e}$ where plim $T^{-1}\underset{\sim}{Q}'\underset{\sim}{e} = \mathcal{E}(\underset{\sim}{Q}'\underset{\sim}{e}) = \underset{\sim}{0}$

and the TSLS estimator of $\underset{\sim}{p}$ in (6) <u>is</u> consistent:

(7) $\tilde{\underset{\sim}{p}} = (\underset{\sim}{X}'\underset{\sim}{M}\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{M}\underset{\sim}{u}$ with plim $\tilde{\underset{\sim}{p}} = \underset{\sim}{p}$.

In Hendry and Harrison (1974) a Control Variable for TSLS is derived the distribution of which provides an <u>asymptotic approximation</u> to that of TSLS and we can use this result to facilitate obtaining the required limiting distribution of TSLS. From their expression (A22),

(8) $(\underset{\sim}{\beta}^{\dagger}- \underset{\sim}{\beta}) = \underset{\sim}{\Phi}_{\ell}\underset{\sim}{\phi} + \underset{\sim}{K}^{-1}(\Delta\underset{\sim}{L})\underset{\sim}{\phi}_{\ell} - \underset{\sim}{K}^{-1}(\Delta\underset{\sim}{K})\underset{\sim}{p}$

where the "Reduced Form" for $\underset{\sim}{X}$ can be written as

(9) $\underset{\sim}{X} = \underset{\sim}{Q}\underset{\sim}{L}' + \underset{\sim}{V}$ with $\mathcal{E}(\underset{\sim}{Q}'\underset{\sim}{V}) = \underset{\sim}{0}$

and

(10) $\Delta L = T^{-1}V'Q.A^{-1}$,

(11) $\Delta K = L(\frac{Q'X}{T}) + (\frac{X'Q}{T})L' - L(Q'Q)L' - K$

(ΔK should not include the $O(\frac{1}{T})$ term as in A(20) when it is used in (8))

But from (11)

(12) $K^{-1}(\Delta K)b = K^{-1}L(\frac{Q'X}{T})b + K^{-1}(\frac{X'Q}{T})L'b - K^{-1}L(\frac{Q'Q}{T})L'b - b$

$= \phi_\ell(\frac{Q'X}{T})b + K^{-1}(\frac{V'Q}{T})L'b - b$ using (9).

and hence

(13) $(\beta^{\dagger}-\beta) = b + \phi_\ell\phi + K^{-1}(\frac{V'Q}{T})\ (A^{-1}\phi_\ell - L'b) - \phi_\ell(\frac{Q'X}{T})b$

But

(14) $A^{-1}\phi_\ell = L'b$

and hence

(15) $(\beta^{\dagger} - \beta) = b + \phi_\ell\ (\phi - (\frac{Q'X}{T})b) = b + \phi_\ell(\frac{Q'e}{T})$

(greatly simplifying the interpretation and calculation of the Control Variable).

Finally we have that by construction of $(\beta^{\dagger} - \beta)$

(16) $(\tilde{\beta} - \beta) = (\beta^{\dagger} - \beta) + O(\frac{1}{T})$

so that, noting $\mathcal{E}\beta^{\dagger} = \beta_\ell$ and $\beta + b = \beta_\ell$

(17) $\sqrt{T}(\tilde{\beta} - \beta_\ell) = \phi_\ell\frac{Q'e}{\sqrt{T}} + O(\frac{1}{\sqrt{T}})$

and therefore the asymptotic distribution of $\sqrt{T}(\tilde{\beta} - \beta_\ell)$ is the same as that of $\phi_\ell\frac{Q'e}{\sqrt{T}}$ which has reduced the problem to a "canonical" form. Let $\mathcal{E}(e\ e') = \Omega = G\ G'$ where Ω is positive definite and let F be generated by a stationary stochastic process with zero mean.

Then $\mathcal{E}(G^{-1}\ e) = \mathcal{E}(e^{o}) = 0$ and $\mathcal{E}(G^{-1}\ e\ e'\ G'^{-1}) = I$. Finally, letting $Q^{o} = GQ$ from the results of Mann and Wald (1943)

(18) $$\frac{1}{\sqrt{T}} Q^{o'} e^{o} \underset{A}{\sim} N(0, \text{plim } T^{-1} Q^{o'} Q^{o}) = N(0, \text{plim } T^{-1} Q' \Omega Q)$$

and hence

(19) $$\sqrt{T}(\tilde{\beta} - \beta_{\ell}) \underset{A}{\sim} N(0, \Phi_{\ell} \Sigma \Phi'_{\ell}) \quad \text{when } \Sigma = \text{plim } T^{-1} Q' \Omega Q.$$

However, $e = W\alpha_{\ell}$ where $\alpha_{\ell}' = (1 - \beta_{\ell}')$

and we can rewrite $W\alpha_{\ell}$ as $(\alpha_{\ell}' \otimes I)(W')^{\nu}$ where $^{\nu}$ denotes the operation of vectoring (stacking columns) so that $(W')^{\nu'} = (W_1' \ldots W_T')$ and $\otimes$ denotes the Kronecker product: $A \otimes B = (b_{ij} A)$. Then

(20) $$\Omega = (\alpha_{\ell}' \otimes I)\, \mathcal{E}((W')^{\nu} (W')^{\nu'})\, (\alpha_{\ell} \otimes I).$$

More precise results can be obtained for the large class of systems in which the complete set of m variables F can be transformed such that they are generated by a first order vector autoregressive process with "white noise" errors:

(21) $$F' = D F_{-1}' + E' \qquad \mathcal{E}(F_{-1}' E) = 0$$

where all the latent roots of D have moduli less than unity, and are distinct.

Let $N = \mathcal{E}(T^{-1} F'F)$; then, since

(22) $$F' = D^{t} F_{-t}' + \Sigma_{j=0}^{t-1} D^{j} E_{-j}'$$

(23) $$\mathcal{E}(F\theta\theta'F') = \begin{pmatrix} \theta'N\theta & \theta'ND'\theta & \ldots & \theta'ND^{T-1'}\theta \\ \theta'DN\theta & \theta'N\theta & \ldots & \theta'ND^{T-2'}\theta \\ \vdots & & \ddots & \vdots \\ \theta'D^{T-1}N\theta & \ldots & \ldots & \theta'N\theta \end{pmatrix} = \Omega^{*}$$

for any $m \times 1$ vector θ. Consider plim $T^{-1} F'\Omega^{*}F$ and let

$$J = \begin{pmatrix} 0 & 0 \\ I_{T-1} & 0 \end{pmatrix} \quad \text{so that } J^{k} = \begin{pmatrix} 0 & 0 \\ I_{T-k} & 0 \end{pmatrix} \quad \text{and}$$

$(J^{k}F)' = F_{-k}'$ apart from asymptotically negligible "end corrections" or "initial values". Thus from (23)

(24) $$\Omega^{*} = \Sigma_{t=0}^{T-1} \theta'D^{t}N\theta\, (J^{t} + J^{t'}) - \theta'N\theta I$$

and

(25) $$F'\Omega^*F = \Sigma_{t=0}^{T-1} \ \theta' D^t N\theta (F'J^t F + F'J^{t'} F) - \theta' N\theta F'F$$

$$= \Sigma_{t=0}^{T-1} \ \theta' D^t N\theta \ (D^t F_t' F_t + F_t' F_t D^{t'}) - \theta' N\theta F'F$$

using (22), and dropping asymptotically negligible terms.

We can factorise D as $B\Delta B^{-1}$ where Δ is the diagonal matrix of eigenvalues $|\lambda_i|<1$ and hence

(26) $$\theta' D^t N\theta D^t F_t' F_t = B(\theta' B)\Delta^t (B^{-1} N\theta) \Delta^t B^{-1} F_t' F_t$$

$$= B(\delta' \Delta^t \zeta) \Delta^t \ B^{-1} F_t' F_t \quad = \quad B[\Sigma_j \ \delta_j \ \zeta_j (\lambda_j \lambda_\ell)^t] \ B^{-1} F_t' F_t$$

where [ ] denotes the ℓ^{th} element of a diagonal matrix ($\delta' = \theta' B$ etc.)

Thus as $\lim D^t = 0$

(27) $$\text{plim } T^{-1} F'\Omega^* F = B\left[\Sigma_j \frac{\delta_j \zeta_j}{1-\lambda_j \lambda_\ell}\right] B^{-1} N + NB^{-1'} \left[\frac{\Sigma_j \delta_j \zeta_j}{1-\lambda_j \lambda_\ell}\right] B' - \theta N\theta N.$$

For example, in the dynamic, simultaneous equations model with autocorrelated errors used by Hendry and Harrison (1974), $m_1 = m_2 = 1$, $m_3 = 3$ and

(28) $$D = \begin{pmatrix} D_1 & D_2 \\ 0 & D_3 \end{pmatrix}$$, $D_3 = dg(\mu_1 \ldots \mu_4)$ and after transforming (1) to eliminate the autocorrelation

(29) $$(D_1 : D_2) = \frac{1}{1-ab} \begin{pmatrix} (d+\rho) & -\rho b & -d\rho & c(\mu_1-\rho) & b\Psi_2\mu_2 & b\Psi_3\mu_3 & b\Psi_4\mu_4 \\ a(d+\rho) & -a\rho b & -ad\rho \,:\, & ac(\mu_1-\rho) & \Psi_2\mu_2 & \Psi_3\mu_3 & \Psi_4\mu_4 \\ 1-ab & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Thus $B = \begin{pmatrix} B_1 & B_2 \\ 0 & B_3 \end{pmatrix}$ where $(B_1 : B_2)$ is

(30) $$\begin{pmatrix} d/(1-ab) & 0 & \rho & \mu_1 & \mu_2 & \mu_3 & \mu_4 \\ ad/(1-ab) & -d/b & a\rho \,:\, & a\mu_1 & (\mu_2-d)/b & (\mu_3-d)/b & (\mu_4-d)/b \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

and $B_3 = dg(\ (\mu_1(1-ab)-d)/c \quad (\mu_2(1-ab)-d)/b\Psi_2 \ \ldots \ (\mu_4(1-ab)-d)/b\Psi_4)$.

$(b, c, \Psi_i \neq 0)$ where the first row of $\underset{\sim}{B}$ is the vector of eigenvalues of $\underset{\sim}{D}$. In (27), taking $\underset{\sim}{\theta}' = (\underset{\sim}{\alpha}_\ell' : \underset{\sim}{0}')$ and selecting the relevant elements of $\underset{\sim}{N}$ yields $\underset{\sim}{\Sigma}$ in (19). (Relative to the notation in Hendry and Harrison, P.154, we have $\rho = r$, $\Psi_i = f_i$, $\mu_i = \lambda_i$). A number of important special cases are now easily derived.

(i) <u>Correct instruments, no autocorrelated errors</u> $(\underset{\sim}{p} = 0, \rho = 0)$

Now $\underset{\sim}{\alpha}_\ell = \underset{\sim}{\alpha}$, $\underset{\sim}{\theta}'\underset{\sim}{N}\underset{\sim}{\theta} = \sigma^2$ and $\underset{\sim}{\theta}'\underset{\sim}{D} = \underset{\sim}{0}'$ so that

(31) $\sqrt{T}(\tilde{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) \underset{A}{\sim} N(\underset{\sim}{0}, \sigma^2 \underset{\sim}{K}^{-1})$, the standard result.

(ii) <u>Valid instruments, autocorrelation</u> $(\rho \neq 0, \underset{\sim}{Q} = (\underset{\sim}{Z}\ \underset{\sim}{Z}^o), \underset{\sim}{p} = \underset{\sim}{0})$

The main changes from (i) are that $\underset{\sim}{\theta}'\underset{\sim}{D} = \rho\underset{\sim}{\theta}'$ and $\underset{\sim}{\theta}'\underset{\sim}{N}\underset{\sim}{\theta} = \sigma^2/(1-\rho^2)$

hence

$$\underset{\sim}{\Omega}^* = \frac{\sigma^2}{1-\rho^2}\begin{pmatrix} 1 & \rho & \cdots & \rho^{T-1} \\ \rho & 1 & & \vdots \\ \vdots & & \ddots & \\ \rho^{T-1} & \cdots & \cdots & 1 \end{pmatrix} = \mathcal{E}(\underset{\sim}{u}\underset{\sim}{u}') \text{ with}$$

(32) $$\text{plim } T^{-1}\underset{\sim}{F}'\underset{\sim}{\Omega}^*\underset{\sim}{F} = \frac{\sigma^2}{1-\rho^2}\{(\underset{\sim}{I} - \rho\underset{\sim}{D})^{-1}\underset{\sim}{N} + \underset{\sim}{N}(\underset{\sim}{I} - \rho\underset{\sim}{D}')^{-1} - \underset{\sim}{N}\}$$

from which $\underset{\sim}{\Sigma}$ is easily selected for use in (19) with $\underset{\sim}{\beta}_\ell = \underset{\sim}{\beta}$. Otherwise one must calculate (27) to obtain $\underset{\sim}{\Sigma}$ for (19) with $\rho \neq 0$, $\underset{\sim}{\beta}_\ell \neq \underset{\sim}{\beta}$. To complete the analysis we must also calculate the plims of the estimated asymptotic variances based on the conventional formula $\tilde{\sigma}^2(\underset{\sim}{X}'\underset{\sim}{M}\underset{\sim}{X})^{-1}$ where $\tilde{\underset{\sim}{u}} = \underset{\sim}{y} - \underset{\sim}{X}\tilde{\underset{\sim}{\beta}}$. Thus plim $\tilde{\sigma}^2 = \sigma_\ell^2 = \mathcal{E}(\underset{\sim}{\alpha}_\ell'(\frac{W'W}{T})\underset{\sim}{\alpha}_\ell)$ and plim $\tilde{\sigma}^2 T(\underset{\sim}{X}'\underset{\sim}{M}\underset{\sim}{X})^{-1} = \sigma_\ell^2 \underset{\sim}{K}^{-1}$ which can be compared (numerically) with the variance of the limiting distribution.

The GIVE class includes the possible choice of $\underset{\sim}{X}$ as the set of instruments which yields OLS denoted ^, <u>equivalent</u> to $\underset{\sim}{M} = \underset{\sim}{I}$. Thus, matching (i) we have $\sqrt{T}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) \underset{A}{\sim} N(\underset{\sim}{0}, \sigma^2\underset{\sim}{H}^{-1})$ where $\underset{\sim}{H} = \text{plim } T^{-1}\underset{\sim}{X}'\underset{\sim}{X}$ (another sub-matrix of $\underset{\sim}{N}$) again reproducing the standard result. Matching (ii), $\sqrt{T}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) \underset{A}{\sim} N(\underset{\sim}{0}, \underset{\sim}{H}^{-1}\underset{\sim}{\Sigma}\underset{\sim}{H}^{-1})$, with $\underset{\sim}{\Sigma}$ given by (32) and hence if neither $\underset{\sim}{Y}$ nor y_1 is a regressor and $m_2 = 1$ with $\underset{\sim}{z} = \mu \underset{\sim}{z}_{-1} + \underset{\sim}{\omega}$ and

(33) $\underset{\sim}{\omega} \sim NI(\underset{\sim}{0}, \gamma^2 \underset{\sim}{I})$ we obtain $\sqrt{T}(\hat{c} - c) \underset{A}{\sim} N\left(0, \left(\frac{\sigma^2}{\gamma^2} \cdot \frac{1-\mu^2}{1-\rho^2} \cdot \frac{1+\mu\rho}{1-\mu\rho}\right)\right)$

as anticipated. The most general case ($\underset{\sim}{p} \neq \underset{\sim}{0}$, with or without $\rho \neq 0$) is (plim $\hat{\beta} = \beta_\ell$).

(34) $\sqrt{T}(\underset{\sim}{\hat{\beta}} - \underset{\sim}{\beta}_\ell) \underset{A}{\sim} N(\underset{\sim}{0}, \underset{\sim}{H}^{-1} \underset{\sim}{\Sigma} \underset{\sim}{H}^{-1})$ with $\underset{\sim}{\Sigma}$ selected from (27), corresponding to $\underset{\sim}{H}$ from $\underset{\sim}{N}$. For the model defined by (28) and (29) exact parametric expressions can be obtained for $\underset{\sim}{H}^{-1} \underset{\sim}{\Sigma} \underset{\sim}{H}^{-1}$ and $\underset{\sim}{\Phi}_\ell \underset{\sim}{\Sigma} \underset{\sim}{\Phi}_\ell'$; these are unenlightening in such a general case but alternative specialisations similar to (33) are reasonably easy to obtain and highlight the main determinants of the asymptotic variances of the estimators around their plims. Further, relative Mean Square Error comparisons are possible and so on.

An important question is the value of these results in samples of the size usually available in Econometrics and we can provide considerable evidence on this from the simulation results obtained by Hendry and Harrison (1974). Firstly, their regression results (as in their Table 4) strongly confirm the closeness of the inconsistency and the finite sample bias (see the Appendix below, however). A similar analysis for the plims of the estimated asymptotic variances and their small sample equivalents is also possible as all the required population moments are available. Thus we obtained the following results from 36 of their 40 experiments (set III) by regressing the mean direct simulation estimates of $\mathcal{E}(\tilde{\sigma}^2)$ and $\mathcal{E}(\hat{\sigma}^2)$ (based on 100 replications) on their <u>respective</u> plims in log linear equations:

(35) $$\ln \tilde{\sigma}^2_\ell = \underset{(.008)}{1.000} \ln \sigma^2_\ell - \underset{(.27)}{.41}/T$$

$$R^2 = .998 \qquad \chi^2_4 = 10.3$$

(36) $$\ln \hat{\sigma}^2 = \underset{(.008)}{1.006} \ln \sigma^2_\ell - \underset{(.23)}{.42}/T$$

$$R^2 = .998 \qquad \chi^2_4 = 0.6$$

The χ_4^2 statistic is an approximate test of the validity of each regression in predicting the outcomes in the remaining four experiments. These are shown in Table I.

Table I Estimates of σ^2 in four experiments

		TSLS		OLS	
T	σ^2	$\tilde{\sigma}^2$	σ_ℓ^2	$\hat{\sigma}^2$	σ_ℓ^2
15	4.	3.46 (.19)	4.33	2.53 (.11)	2.63
55	.25	0.32 (.01)	0.34	0.32 (.01)	0.33
55	4.	3.94 (.09)	4.00	3.51 (.07)	3.56
75	4.	10.23 (.41)	9.96	5.85 (.14)	5.89

(Simulation Standard Errors in parentheses)

It is clear that σ_ℓ^2 provides a very good explanation of estimates of σ^2 even in quite small samples. Similarly for the mean Estimated Standard Errors (ESE)(based on "conventional" formulae like $\tilde{\sigma}^2(\underset{\sim}{X}'\underset{\sim}{M}\underset{\sim}{X})^{-1}$) for the three parameters in (1) ($m_1 = m_2 = 1$, $m_3 = 3$) for the same 36 experiments. The results are shown in Table II, which records estimates of regressions of the form

(37) $$\ln ESE_{ij} = \gamma_{1i} \ln PE_{ij} + \gamma_{2i}/T_j \quad \begin{matrix}(j=1,\ldots 36)\\(i=1,\ldots 3)\end{matrix}$$ where PE_i is the plim of ESE_i.

Table II ESE Regressions

	TSLS				OLS			
	γ_1	γ_2	R^2	χ_4^2	γ_1	γ_2	R^2	χ_4^2
b	1.01 (.007)	2.40 (.40)	.994	5.3	1.01 (.003)	1.80 (.19)	.998	0.3
d	1.00 (.004)	2.27 (.27)	.996	0.7	1.00 (.003)	1.72 (.21)	.997	5.8
c	1.00 (.005)	3.58 (.20)	.998	1.7	1.00 (.004)	3.25 (.17)	.998	10.8

(Regression Coefficient Standard Errors in parentheses)

Table III records the outcomes for the same 4 experiments as in Table I.

Table III   Estimates of ESE and PE in four Experiments

	TSLS						OLS					
	b		d		c		b		d		c	
T	ESE	PE	ESE	PE	ESE	PE	ESE	PE	ESE	PE	ESE	PE
15	.40 (.21)	.41	.24 (.08)	.21	1.16 (.53)	.90	.24 (.08)	.22	.19 (.05)	.16	.96(.38)	.70
55	.07 (.01)	.07	.06 (.01)	.06	.15 (.03)	.14	.06 (.01)	.06	.06 (.01)	.05	.14(.03)	.14
55	.25 (.06)	.24	.14 (.01)	.14	.50(.13)	.46	.18 (.03)	.17	.13 (.01)	.13	.47(.11)	.43
75	.21 (.07)	.20	.06 (.02)	.06	.62 (.16)	.62	.10 (.01)	.10	.03 (.00)	.03	.47(.09)	.47

(Legend as Table I)

Again we convincingly confirm the value of asymptotic theory in explaining small sample outcomes. Finally we investigated the relationship between the actual sampling variances (e.g. $\mathcal{E}(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)'$) as estimated by the experiments (SV) and the variances of the asymptotic distributions based on (19) and (27) (denoted AV). Note that $\beta^{\dagger}$ has (19) as its finite sample distribution and therefore the rather complicated numerical calculation of AV can be checked by the simulation variance of the Control Variable. Further, the Monte Carlo efficiency gains from using $\beta^{\dagger}$ depend on AV being "close" to SV (i.e. on $\tilde{\beta}$ and $\beta^{\dagger}$ being highly correlated) and the results of Hendry and Harrison suggest that this will be the case (for their 20 experiments with T = 55 or 75, the average variance reduction achieved by the control variable was a factor of 10.6 for OLS and 7.6 for TSLS). Table IV records estimates of regressions of the form

$$(37) \quad \ln SV_{ij} = \gamma_{1i} \ln AV_{ij} + \gamma_{2i}/T_j \qquad \begin{matrix}(j=1,\ldots,36)\\(i=1,\ldots,3)\end{matrix}$$

and Table V records SV and AV in the remaining four experiments.

We conclude that for both the first and the second moments of $\tilde{\beta}$ and $\hat{\beta}$ the asymptotic results are an extremely good guide to the finite sample outcome confirming the value of the theory and casting considerable doubt on the net value of imprecise, specific and expensive Monte Carlo experiments. A similar conclusion holds for the estimated variances and estimated equation standard errors.

TABLE IV. SV Regressions

	TSLS				OLS			
	γ_1	γ_2	R^2	χ^2_4	γ_1	γ_2	R^2	χ^2_4
b	1.01 (.011)	1.72 (.66)	.982	1.7	1.01 (.008)	1.80 (.51)	.988	4.6
d	1.00 (.012)	1.93 (.71)	.973	2.5	1.01 (.013)	2.44 (.80)	.967	5.7
c	0.99 (.015)	3.21 (.55)	.983	8.9	1.00 (.015)	3.12 (.54)	.983	11.6

TABLE V. Estimates of SV and AV in 4 Experiments

	TSLS						OLS					
	b		d		c		b		d		c	
T	SV	AV	SV	AV	SV	AV	SV	AV	SV	AV	SV	AV
15	.393	.399	.299	.233	1.63	0.97	.250	.216	.250	.170	1.28	0.73
55	.079	.074	.078	.070	.199	.200	.057	.058	.077	.069	.196	.196
55	.263	.242	.122	.135	.526	.459	.185	.174	.112	.126	.492	.433
75	.187	.177	.055	.055	.346	.305	.146	.123	.044	.038	.401	.353

Appendix *

Maddala and Rao (1973) apparently disconfirm this result, but their finding is partly an artefact of their choice of parameter values and data generation process. They use

(A1) $Z_t = \mu Z_{t-1} + \omega_t$ and while $\mathcal{E}(Z_t^2) = \gamma^2/(1-\mu^2)$ this is <u>NOT</u>

equal to $\mathcal{E}(Z_t^2/Z_o)$ (the variance Z_t actually has in the Monte Carlo experiments). Maddala and Rao set $Z_o = 0$, generate 60 values and discard the first 20 and use $T = 40$ so that as

(A2) $$Z_t = \mu^t Z_o + \sum_{j=0}^{t-1} \mu^j \omega_{t-j}$$

(A3) $$\mathcal{E}\left(\frac{1}{40} \sum_{t=21}^{60} Z_t^2/0\right) = \frac{1}{40} \gamma^2 \sum_{t=21}^{60} \sum_{j=0}^{t-1} \mu^{2j} \approx \gamma^2(1 - \mu^{42}/40 - k(40,\mu))/(1-\mu^2)$$

and while the $\mu^{42}/40$ is negligible, if μ is close to unity, $k(T,\mu) = \mu^{T+4}/T(1-\mu^2)$ is not.

Indeed, $k(40, .98) \approx .25$ yielding a downward bias of about 25%, while $k(100, .98) \approx .04$ and $k(40, .8)$ is negligible. This exactly matches their findings. Thus for the sample sizes and parameters used, their actual signal-noise ratios did not have the values implied in the asymptotic calculations (whereas they would have obtained accurate results had they used (A3)). Whether this is regarded as a "failure" of asymptotic theory in near non-stationary situations, or an inappropriate experimental design is a moot point.

---

*I am grateful to James E. Davidson for bringing my attention to this point.

References

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

Maddala, G.S. and Rao, A. S. (1973) "Tests for Serial Correlation in Regression Models with Lagged Dependent Variables and Serially Correlated Errors," Econometrica, 41, 761-774.