COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 626

ON THE EXACT DISTRIBUTION OF LIML

Peter C. B. Phillips

March 26, 1982

ON THE EXACT DISTRIBUTION OF LIML

by

P. C. B. Phillips[*]

Cowles Foundation
Yale University

0. ABSTRACT

It is shown that the exact distribution of
the LIML estimator in a general and leading single
equation case is multivariate Cauchy. The corres-
ponding result for the IV estimator is a form of
multivariate t density where the degrees of
freedom depend on the number of instruments.

---

1. INTRODUCTION

Improvements in the algebraic machinery of multivariate analysis
have recently led to many advancements in our understanding of the finite
sample properties of statistical methods in econometrics, particularly
with regard to the simultaneous equations model. Modern multivariate
methods provide a convenient stepping stone to the solution of exact sampling
distribution problems through manageable algebraic representations of the
joint density functions of the matrices of sample moments upon which most
common econometric estimators depend. These matrix variates have, in
general, noncentral multivariate distributions whose algebraic forms and
properties have been intensively studied in mathematical statistics. Some
of the most important contributions in this area have been by Herz (1955),
Constantine (1963), James (1964) and Davis (1980a, 1980b), all of which
have substantially facilitated the development of econometric small sample
theory in recent years. A detailed account of the theoretical developments
that have taken place in econometrics, largely in conjunction with this
analytic progress in multivariate methods, may be found in Mariano (1982)
and Phillips (1980a, 1982a).

The purpose of the present paper is to focus on a simplified class
of problems within the simultaneous equations setting where standard methods
of multivariate analysis allow us to extract the exact distributions of
econometric estimators with relative ease.

2. LEADING CASES AND THE INSTRUMENTAL VARIABLE (IV) ESTIMATOR

We will work with the structural equation

$$(1) \qquad y_1 = Y_2\beta + Z_1\gamma + u$$

where $y_1 (T \times 1)$ and $Y_2 (T \times n)$ are an observation vector and observation matrix, respectively, of $n+1$ included endogenous variables, Z_1 is a $T \times K_1$ matrix of included exogenous variables, and u is a random disturbance vector. The reduced form of (1) is given by

$$(2) \qquad [y_1 \vdots Y_2] = [Z_1 \vdots Z_2]\begin{bmatrix} \pi_{11} & \Pi_{12} \\ \pi_{21} & \Pi_{22} \end{bmatrix} + [v_1 \vdots V_2] = Z\Pi + V ,$$

where Z_2 is a $T \times K_2$ matrix of exogenous variables excluded from (1). The rows of the reduced form disturbance matrix V are assumed to be independent, identically distributed, normal random vectors. We assume that the usual standardizing transformations (see Phillips (1982a)) have been carried out so that the covariance matrix of rows of V is the identity matrix and $T^{-1}Z'Z = I_K$ where $K = K_1 + K_2$. We also assume that $K_2 \geq n$ so that the necessary order conditions for (1) to be identified are satisfied.

There are two special categories of models such as (1) and (2) in which the exact density functions of the common single equation estimators of β in (1) can be extracted with relative ease. In the first category are the just identified structural models in which the usual consistent estimators all reduce to indirect least squares and take the form

$$(3) \qquad \beta_{ILS} = [Z_2'Y_2]^{-1}[Z_2'y_1]$$

of a matrix ratio of normal variates. In the two endogenous variable case

(where $n = 1$), this reduces to a simple ratio of normal variates whose

probability density function (p.d.f.) was first derived by Fieiller (1932)

and takes the following form here (see Mariano and McDonald (1979))

$$(4) \qquad pdf(r) = \frac{\exp\left\{-\frac{\mu^2}{2}(1+\beta^2)\right\}}{\pi(1+r^2)} {}_1F_1\left[1, \frac{1}{2}; \frac{\mu^2}{2} \frac{(1+\beta r)^2}{1+r^2}\right]$$

where $\mu^2 = T\Pi'_{22}\Pi_{22}$ is the scalar concentration parameter. In the general

case of $n+1$ included endogenous variables the density (4) is replaced by

a multivariate analogue in which the ${}_1F_1$ function has a matrix argument

(see Sargan (1976) and Phillips (1980b)).

The category of estimators that take the generic form of a matrix

ratio of normal variates, as in (3), also includes the general IV estimator

in the overidentified case provided the instruments are non-stochastic,

that is, if $\beta_{IV} = [W'Y_2]^{-1}[W'y_1]$ and the matrix W is non-stochastic,

as distinct from its usual stochastic form in the case of estimators like

2SLS in overidentified equations. This latter case has been discussed

by Mariano (1977). A further application of matrix ratios of normal var-

iates, related to (3), occurs in random coefficient models where the reduced

form errors are a matrix quotient of the form $A^{-1}a$ where both a and

the columns of A are normally distributed. Existing theoretical work

in this area has proceeded essentially under the hypothesis that det A

is non-random (see Kelejian (1974)) and can be generalized by extending

(4) to the multivariate case in much the same way as the exact distribution

theory for the IV estimator in the $n+1$ endogenous variable case.

The second category of special models that facilitate the develop-

ment of an exact distribution theory are often described as leading cases

of the fully parameterized simultaneous equations model. In these leading
cases, certain of the critical parameters are set equal to zero and the
distribution theory is developed under this null hypothesis. In the most
typical case, this hypothesis prescribes a specialized reduced form which
ensures that the sample moments of the data on which the estimator depends
have central rather than (as is typically the case) noncentral distribu-
tions. The adjective "leading" is used advisedly since the distributions
that arise from this analysis typically provide the leading term in the
multiple series representation of the true density that applies when the
null hypothesis itself no longer holds. As such the leading term provides
important information about the shape of the distribution by defining a
primitive member of the class to which the true density belongs in a more
general setting.

It is with such leading cases that the present paper is concerned.
We will consider, in particular, the leading subcase of (1) and (2) in which
$\Pi_{22} = 0$. Under this hypothesis the reduced form (2) becomes

(2') $[y_1 \vdots Y_2] = Z_1[\pi_{11} \vdots \Pi_{12}] + [v_1 \vdots V_2]$.

The statistical analysis of this leading case can now be simply illustrated
in terms of the following IV estimator of β :

(5) $\beta_{IV} = [Y_2'Z_3Z_3'Y_2]^{-1}[Y_2'Z_3Z_3'y_1]$

where $Z_3(T \times K_3)$ is a submatrix of Z_2 selected as instruments additional
to Z_1 and where it is assumed that $K_3 \geq n$. We note that the conditional
distribution of $(T^{-1}Y_2'Z_3Z_3'Y_2)^{-1/2}(T^{-1}Y_2'Z_3Z_3'y_1)$ given $Z_3'Y_2$ is $N(0, I_n)$.
This is independent of $Z_3'Y_2$ and is also, therefore, the unconditional

distribution. Further, $T^{-1}Y_2'Z_3Z_3'Y_2$ has a central Wishart distribution of order n with degrees of freedom K_3 and covariance matrix I_n. We may therefore write β_{IV} in the form

$$(6) \qquad \beta_{IV} = [(T^{-1}Y_2'Z_3Z_3'Y_2)^{-1/2}][(T^{-1}Y_2'Z_3Z_3'Y_2)^{-1/2}(T^{-1}Y_2'Z_3Z_3'y_1)]$$

$$\equiv [W_n(K_3, I_n)]^{-1/2}N(0, I_n)$$

so that β_{IV} is proportional to a multivariate t variate (see, for example, Dickey (1967)). The p.d.f. of β_{IV} is therefore given by

$$(7) \qquad pdf(r) = \frac{\Gamma\left(\frac{K_3+1}{2}\right)}{\pi^{n/2}\Gamma\left(\frac{K_3-n+1}{2}\right)(1+r'r)^{(K_3+1)/2}} = \frac{\Gamma\left(\frac{L+n+1}{2}\right)}{\pi^{n/2}\Gamma\left(\frac{L+1}{2}\right)(1+r'r)^{(L+n+1)/2}}$$

where $L = K_3 - n$ is the number of surplus instruments used in the estimation of β.

The density (7) specializes to the case of two stage least squares for $K_3 = K_2$ (where the result was given by Basmann (1974)) and to the case of ordinary least squares for $K_3 = T - K_1$ (where the result was given by Wegge (1971)). As shown in Phillips (1982b), (7) is in fact the leading term in the multiple series representation of the exact density of β_{IV} in the general single equation case where Π_{22} is not necessarily the zero matrix. Moreover, the leading marginal densities can be readily deduced from (7) (see Phillips (1982c)) and standard properties of the multivariate t confirm that integer moments exist up to the order L (i.e. the number of surplus instruments).

3. THE DISTRIBUTION OF LIML

The LIML estimator β_{LIML} of β minimizes the ratio $\beta_\Delta' W \beta_\Delta / \beta_\Delta' S \beta_\Delta$

where $\beta_\Delta' = (1, -\beta')$, $W = X'(P_Z - P_{Z_1})X$, $S = X'(I - P_Z)X$ and where

$X = [y_1 : Y_2]$ and $P_A = A(A'A)^{-1}A'$. Under the null hypothesis that

$\Pi_{22} = 0$ in (2) we deduce that W and S have central Wishart distribu-

tions $W_m(K_2, I)$ and $W_m(T-K, I)$ respectively, where $m = n+1$. Since

$(P_Z - P_{Z_1})(I - P_Z) = 0$, W and S are independent with joint p.d.f.

given by

$$(8) \quad pdf(W,S) = \frac{etr\left\{-\frac{1}{2}(W+S)\right\}(\det W)^{(K_2-m-1)/2}(\det S)^{(T-K-m-1)/2}}{2^{m(T-K+K_2)/2}\Gamma_m\left(\frac{K_2}{2}\right)\Gamma_m\left(\frac{T-K}{2}\right)}.$$

We now transform $W \to S^{-1/2}WS^{-1/2} = F$ giving the joint density

$$(9) \quad pdf(F,S) = \frac{etr\left\{-\frac{1}{2}S(I+F)\right\}(\det F)^{(K_2-m-1)/2}(\det S)^{(T-K+K_2-m-1)/2}}{2^{m(T-K_1)/2}\Gamma_m\left(\frac{K_2}{2}\right)\Gamma_m\left(\frac{T-K}{2}\right)}.$$

Using the following matrix variate gamma integral (Herz (1955))

$$(10) \quad \int_{S>0} etr(-SZ)(\det S)^{a-(m+1)/2}dS = \Gamma_m(a)(\det Z)^{-a}$$

for $Re(Z) > 0$ and $Re(a) > (m-1)/2$ we deduce from (9) the marginal matrix

density of F

$$(11) \quad pdf(F) = \frac{\Gamma_m\left[\frac{T-K_1}{2}\right] (\det F)^{(K_2-m-1)/2} [\det(I+F)]^{-(T-K_1)/2}}{\Gamma_m\left[\frac{K_2}{2}\right]\Gamma_m\left[\frac{T-K}{2}\right]}$$

$$= \left[B_m\left[\frac{K_2}{2}, \frac{T-K}{2}\right]\right]^{-1} (\det F)^{(K_2-m-1)/2} [\det(I+F)]^{-(T-K_1)/2}$$

which is a multivariate Beta density of the second kind (see, for example, Tan (1969)).

Expression (11) now leads simply to the p.d.f. of the LIML estimator β_{LIML} . We note that if λ_m is the smallest latent root of $F = S^{-1/2}WS^{-1/2}$ then β_{LIML} satisfies the system

$$(12) \quad (S^{-1/2}WS^{-1/2} - \lambda_m I)\beta_\Delta = (F - \lambda_m I)\beta_\Delta = 0$$

where $\beta_\Delta = (\beta_{\Delta 1}, \beta_{\Delta 2})'$ is the latent vector associated with the smallest latent root λ_m and $\beta_{LIML} = -\beta_{\Delta 2}/\beta_{\Delta 1}$.

In order to extract the exact distribution of β_{LIML} we introduce the orthogonal transformation H by which F is diagonalized so that $H'FH = \Lambda = diag(\lambda_1, \ldots, \lambda_n)$. This transformation is unique if we specify that $\lambda_1 > \lambda_2 > \ldots > \lambda_m$ and that the elements in the first row of H are positive. (The latter eliminates the possibility of multiplying columns of H by -1.) From (11) we can deduce the joint distribution of (Λ, H) . It is most convenient in the general multivariate case to work with the probability element $pdf(F)dF$. Under the transformation $F \rightarrow (\Lambda, H)$ we then have

$$pdf(\Lambda,H)d\Lambda(dH) = pdf(F)dF$$

$$(13) \qquad = pdf(F) \prod_{i<j} (\lambda_i - \lambda_j)(\prod_{i=1}^{m} d\lambda_i)(dH)$$

$$(14) \qquad = \left[B_n\left(\frac{K_2}{2}, \frac{T-K}{2}\right) \right]^{-1} (\prod_{i=1}^{m} \lambda_i)^{(K_2-m-1)/2}$$

$$\cdot [\prod_{i=1}^{m} (1+\lambda_i)]^{-(T-K_1)/2} \prod_{i<j} (\lambda_i - \lambda_j)(\prod_{i=1}^{m} d\lambda_i)(dH) .$$

Line (13) follows from Constantine (1963, equation (43), p. 1280) and line (14) follows directly from (11). (dH) is the invariant measure on the orthogonal group (the group of orthogonal matrices) that is normalized so that the measure over the whole group (restricted so that $h_{1j} \geq 0$) is unity. The distribution of H is called the conditional Haar invariant distribution (see Anderson (1958, p. 322)) and we see from (14) that H is distributed independently of the latent roots that form the diagonal elements of Λ . To find the distribution of β_{LIML} we now need to concentrate on the final column of H which we write as the (n+1)-vector h or in partitioned form as $h' = (h_1, h_2')$. The invariant distribution of H implies an invariant measure over the Stiefel manifold defined by $h'h = h_1^2 + h_2'h_2 = 1$. The latter is the unit sphere in (n+1)-dimensional Euclidean space and the invariant measure on this manifold is given by the exterior differential form

$$(15) \qquad (dh) = \bigwedge_{j=1}^{n} b_j'dh$$

where b_1, b_2, ..., b_n are orthonormal column vectors orthogonal to h

(see equation (5.1) of James (1954)). Using the parameterization of the manifold in which $h_1 = (1 - h_2'h_2)^{1/2}$ and restricting the region so that $h_1 > 0$, the invariant measure (15) can be written in the alternative form

$$(16) \qquad (dh) = \frac{k \, dh_2}{(1 - h_2'h_2)^{1/2}}$$

(see Farrell (1976), equations (7.7.3-4)) where the constant k is selected so that the measure over the restricted $(h_1 > 0)$ region of the unit sphere is unity. Since the measure over the entire unit sphere in R^{n+1} is $2\pi^{(n+1)/2}/\Gamma\left(\frac{n+1}{2}\right)$, that is the surface area of the sphere (see James (1954) equation (5.9)), the normalizing constant in the invariant measure (16) over the restricted region is (ignoring questions of sign in (16) since we are working with positive probability measures):

$$(17) \qquad k = \pi^{-(n+1)/2}\Gamma\left(\frac{n+1}{2}\right) .$$

We now renormalize the latent vector h to yield the LIML estimator. This involves the transformation $h_2 \to -r/(1+r'r)^{1/2}$ with $h_1 = (1 - h_2'h_2)^{1/2} = (1+r'r)^{-1/2}$. Taking differentials we deduce that

$$dh_2 = -(1+r'r)^{-1/2}[I+rr']^{-1}dr$$

and the modulus of the Jacobian of the transformation is $(1+r'r)^{-(n+2)/2}$. Thus, the invariant measure (16) defined over the appropriately restricted region (for which $h_1 > 0$) of the unit sphere in R^{n+1} transforms as follows:

$$(18) \qquad (dh) = \frac{\Gamma\left[\frac{n+1}{2}\right] dr}{\pi^{(n+1)/2}(1+r'r)^{(n+1)/2}} \ .$$

The p.d.f. of β_{LIML} then takes the form

$$(19) \qquad pdf(r) = \frac{\Gamma\left[\frac{n+1}{2}\right]}{\pi^{(n+1)/2}(1+r'r)^{(n+1)/2}} \ ,$$

that is, a multivariate Cauchy distribution.

In the two endogenous variable case (n = 1), (19) reduces to the univariate Cauchy, which provides the leading term in the multiple series representation of the exact density given by Mariano and McDonald (1979) in the general case. We deduce directly from (18) that β_{LIML} has no finite moments of integral order, as was originally shown by Mariano and Sawa (1972) and Sargan (1970).

4. CONCLUDING REMARKS

The exact distributions obtained above apply when $\Pi_{22} = 0$ and correspond, therefore, to a particular structure of the model in which the true coefficient vector β is not identifiable. Neither of the exact densities (7) or (19) actually involve β . Both are in fact centered around the origin. When β is itself zero, there is an absence of simultaneity in the model and in this case OLS (with $K_3 = T - K_1$ in (7)) is consistent. In both this case ($\beta = 0$) and for other values of β the exact densities of LIML and 2SLS (the latter with $K_3 = K_2$ in (7)) are invariant to changes in the sample size T . Thus, as T $\to \infty$ these distributions continue to demonstrate the uncertainty about β due to the lack of identification. It is of interest that in the more general case

where $\Pi_{22} \neq 0$ the exact distributions of 2SLS and LIML retain certain important properties (such as their tail area behavior) which apply in the primitive forms of these densities given in (7) and (19) above.

REFERENCES

Anderson, T. W. (1958). <u>An Introduction to Multivariate Statistical Analysis</u>. New York: J. Wiley & Sons.

Basmann, R. L. (1974), "Exact Finite Sample Distribution for Some Econometric Estimators and Test Statistics: A Survey and Appraisal," Chapter 4 in <u>Frontiers of Quantitative Economics</u>, Vol. 2 (M. D. Intriligator and D. A. Kendrick, eds.). North-Holland.

Constantine, A. G. (1963), "Some Noncentral Distribution Problems in Multivariate Analysis," <u>Annals of Mathematical Statistics</u>, 34, 1270-1285.

Davis, A. W. (1980a), "Invariant Polynomials with Two Matrix Arguments Extending the Zonal Polynomials," in P. R. Krishnaiah (ed.) <u>Multivariate Analysis</u>, Amsterdam: North-Holland.

Davis, A. W. (1980b), "Invariant Polynomials with Two Matrix Arguments Extending the Zonal Polynomials: Applications to Multivariate Distribution Theory," <u>Annals of Institute of Statistical Mathematics</u>, forthcoming.

Dickey, J. M. (1967), "Matricvariate Generalisations of the Multivariate t Distribution and the Inverted Multivariate t Distribution," <u>Annals of Mathematical Statistics</u>," 38, 511-518.

Farrell, R. H. (1976), <u>Technique of Multivariate Calculation</u>, Springer: New York.

Fieiller, E. C. (1932), "The Distribution of the Index in a Normal Bivariate Population," <u>Biometrika</u>, 24, 428-440.

Herz, C. W. (1955), "Bessel Functions of Matrix Argument," <u>Annals of Mathematics</u> 61, 474-523.

James, A. T. (1954), "Normal Multivariate Analysis and the Orthogonal Group," Annals of Mathematical Statistics, 25, 40-75.

James, A. T. (1964), "Distribution of Matrix Variates and Latent Roots Derived from Normal Samples," Annals of Mathematical Statistics, 35, 475-500.

Kelejian, H. H. (1974), "Random Parameters in a Simultaneous Equation Framework: Identification and Estimation," Econometrica, 42, 517-527.

Mariano, R. S. (1977), "Finite-Sample Properties of Instrumental Variable Estimators of Structural Coefficients," Econometrica, 45, 487-496.

Mariano, R. S. (1982), "Analytical Small-Sample Distribution Theory in Econometrics: The Simultaneous-Equations Case," International Economic Review (forthcoming).

Mariano, R. S. and J. McDonald (1979), "A Note on the Distribution Functions of LIML and 2SLS Coefficient Estimators in the Exactly Identified Case," Journal of the American Statistical Association, 74, 847-848.

Mariano, R. S. and T. Sawa (1972), "The Exact Finite-Sample Distribution of the Limited Information Maximum Likelihood Estimator in the Case of Two Included Endogenous Variables," Journal of American Statistical Association, 67, 159-163.

Phillips, P. C. B. (1980a), "Finite Sample Theory and the Distributions of Alternative Estimators of the Marginal Propensity to Consume," Review of Economic Studies, Vol. 47, No. 1, January 1980, pp. 183-224.

Phillips, P. C. B. (1980b), "The Exact Finite Sample Density of Instru-
mental Variable Estimators in an Equation with $n+1$ Endogenous
Variables," Econometrica, Vol. 48, No. 4, May 1980, pp. 861-868.

Phillips, P. C. B. (1982a), "Exact Small Sample Theory in the Simultaneous
Equations Model," in M. D. Intriligator and Z. Griliches (Eds.),
Handbook of Econometrics, North Holland: Amsterdam.

Phillips, P. C. B. (1982b), "Marginal Densities of Instrumental Variable
Estimators in the General Single Equation Case," Advances in
Econometrics (forthcoming).

Sargan, J. D. (1970), "The Finite Sample Distribution of FIML Estimators,"
Paper presented to The World Congress of the Econometric Society,
Cambridge.

Sargan, J. D. (1976), "Econometric Estimators and the Edgeworth Approxi-
mation," Econometrica 44, 421-428, and (1977) "Erratum,"
Econometrica, 15, 272.

Tan, W. Y. (1969), "Note on the Multivariate and the Generalizad Multi-
variate Beta Distributions," Journal of the American Statistical
Association, 64, 230-241.

Wegge, L. L. (1971), "The Finite Sampling Distribution of Lease Squares
Estimators with Stochastic Regressors," Econometrica, 38, 241-251.