Cowles Foundation Paper #638

INTERNATIONAL ECONOMIC REVIEW
Vol. 27, No. 1, February, 1986

THE DISTRIBUTION OF FIML IN THE LEADING CASE*

BY P. C. B. PHILLIPS**

1. INTRODUCTION

In a recent article [1984a] I showed that the distribution of the limited information maximum likelihood (LIML) estimator of the coefficients of the endogenous variables in a single structural equation is multivariate Cauchy in the leading (totally unidentified) case. The purpose of the present note is to show that the same result holds for the full information maximum likelihood (FIML) estimator. Our proof relies on the theory of invariant measures on a Stiefel manifold. This approach provides a major simplification of the derivation of the LIML result given in the earlier article and extends to the FIML case without difficulty. We start by illustrating its use for LIML.

2. THE LIML DISTRIBUTION

Our notation and assumptions are the same as those of the earlier article [1984a]. The structural equation has the form:

$$[y_1 \vdots Y_2]b = Z_1\gamma + u; \quad X = [y_1^1 \vdots Y_2^n] \tag{1}$$

where X is a $T\times m$ matrix of observations of $m=n+1$ 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 X is given by:

$$X = [Z_1 \vdots Z_2]\begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{bmatrix} + V = Z\Pi + V \tag{2}$$

where Z_2 is a $T\times K_2$ matrix of exogenous variables excluded from (1). This equation is assumed to be in standardized form [see Phillips (1983)], so that the rows of V are i.i. $N(0, I_m)$ and $Z'Z=I_K$ where $K=K_1+K_2$. It is assumed that $K_2\geq n$ and consequently (1) is apparently identified by order conditions to the investigator. However, under the leading case hypothesis $\Pi_{22}=0$, the structural equation (1) is totally unidentified. It is this hypothesis that enables us to achieve the simplicity of the central distribution theory that follows.

The LIML estimator of b minimizes the ratio

$$\frac{b'Wb}{b'Sb}; \quad W = X'(P_Z-P_{Z_1})X, \quad S = X'(I-P_Z)X \tag{3}$$

* Manuscript received January, 1983; revised June, 1985.

** My thanks go to Glena Ames for typing the manuscript of this paper and to the NSF and the John Simon Guggenheim Foundation for financial support.

under a normalization rule, which we take to be $b'b=1$. $P_a=a(a'a)^{-1}a'$ for any matrix a of full column rank.

W and S have independent Wishart distributions with the same covariance matrix I_m and with degrees of freedom K_2 and $T-K$ respectively. Both Wishart distributions are central in the leading case. Since the LIML criterion (3) depends only on (W, S), so too does the optimizing value of b. Denote this optimizing value of b by the random m-vector h, which of course lies on the unit sphere (or Stiefel manifold $V_{1,m}$). We now note that the distribution of (W, S), and hence the criterion (3), is invariant under the transformation $(W, S)\rightarrow(H'WH, H'SH)$ for any $H\in O(m)$, the orthogonal group of $m\times m$ matrices. We also note that the criterion (3) is invariant under the transformation $b\rightarrow bk$ where $k\in O(1)$ (that is, $k^2=1$). It follows that the distribution of the random vector h is invariant under the simultaneous transformations $h\rightarrow Hhk$ for any $H\in O(m)$ and $k\in O(1)$. Now the only distribution on the Stiefel manifold $V_{1,m}$ which is invariant under both left and right (groups of) orthogonal transformations is given by the normalized invariant measure on the manifold. This measure produces a volume element on the (hyper) surface of the manifold and is usually represented by an exterior differential form as in equation (12) of the earlier article [1984a]. The reader is referred to James [1954] for a full discussion. The distribution so generated is the uniform distribution on the sphere $h'h=1$.

The argument of the previous paragraph applies equally well to the LIMLK estimator, which is obtained by minimizing the quadratic form $b'Wb$ subject to the condition $b'b=1$. This verifies that the distribution of LIML and LIMLK are the same in this leading case.

When the structural equation (1) is rewritten as

$$y_1 = Y_2\beta + Z_1\gamma + u \tag{4}$$

the LIML estimator of β may be obtained from the partition of $h'=(\overset{1}{h_1'}, \overset{n}{h_2'})$ by taking the ratio $\beta_{\text{LIML}}=-h_2/h_1=r$, say. When h is uniform over the sphere, the distribution of r is multivariate Cauchy, as shown in the earlier article [1984a].

3. THE FIML DISTRIBUTION

We write the system of structural equations as:

$$YB + ZC = U, \quad \text{or equivalently} \quad XA = U \tag{5}$$

where $X=[Y : Z]$, $A'=[B' : C']$, $Y'=[y_1,\ldots, y_T]$ is a $g\times T$ matrix of observations of g endogenous variables and U is a matrix of structural disturbances. The reduced form of (5) is

$$Y = Z\Pi + V \tag{6}$$

corresponding to (2) above; but now we allow for the possibility of more endogenous variables in the system so that $g\geq m=n+1$. The first column of (5) yields the structural equation (1). Thus, upon the appropriate ordering of the variables we may partition the coefficient matrices as:

$$B = \overset{\;\;1\;\;\;\;\; g-1}{[b_1 \vdots B_2]}, \quad C = \overset{\;\;1\;\;\;\;\; g-1}{[c_1 \vdots C_2]} \tag{7}$$

where

$$b_1' = (b', 0), \quad c_1' = (\gamma', 0). \tag{8}$$

As before we may assume that the system is in standardized form with the rows of V now being i.i. $N(0, I_g)$ and $Z'Z = I_K$. We also assume that each equation of the structural system (5) is apparently identified by exclusion restrictions. The leading case hypothesis is:

$$H_0: \quad \Pi = 0 \tag{9}$$

and each equation of the structural system is, in effect, totally unidentified, as in the case of (1) above.

Structural FIML estimates are obtained by maximizing the concentrated log-likelihood

$$\text{const.} + ln|\det B| - (1/2) ln \det(A'X'XA), \tag{10}$$

under a normalization rule which we take to be $b_i'b_i = 1 (i = 1, \ldots, g)$. For the first structural equation this becomes $b'b = 1$ as before. We will denote the FIML estimator of b by the random m-vector h.

The distribution of the FIML criterion (10) depends on the distribution of Y. Now Y is matrix $N_{T,g}(0, I_{Tg})$ which is invariant under the simultaneous transformations $Y \to K_1 Y K_2$ for any $K_1 \in O(T)$ and $K_2 \in O(g)$. The distribution of the criterion (10) and hence that of the optimizing value of A is also invariant under the same transformations. We consider a particular group of transformations in which $b \to Hbk$ for $H \in O(m)$ and $k \in O(1)$; this is equivalent to $b \to Lb$ for $L = kH \in O(m)$. We define $K \in O(g)$ as

$$K = \begin{bmatrix} L & 0 \\ 0 & P \end{bmatrix}, \quad \text{for any} \quad P \in O(g-m)$$

and transform $Y \to YK'$. Correspondingly, in (10)

$$B \to KB = \left[\begin{array}{c|c} Lb & \\ \cdots & KB_2 \\ 0 & \end{array} \right]$$

and the FIML estimator of Lb is Lh. However, since the distribution of Y is invariant under the transformation $Y \to YK'$ it follows that the distribution of h is invariant under transformation $h \to Lh$. This argument holds for any choice of $L \in O(m)$. As in the case of LIML, the only distribution on the manifold $V_{1,m}$ which is invariant under this group of orthogonal transformations is given by the normalized invariant measure on the manifold. Thus, the FIML estimator

h has a uniform distribution on the sphere $h'h=1$.

Upon renormalization of the first equation as in (4) we deduce as before that the FIML estimator of the structural coefficient vector β has a multivariate Cauchy distribution.

4. SOME FURTHER RESULTS

The distribution of the criterion (10) and the optimizing value of A is also invariant under the transformation $Y'Z \to Y'ZJ$ for any $J \in O(K)$. The FIML estimator of C transforms as $C \to J'C$; and the distribution of the estimator is invariant under this transformation. Confining our attention to a column of C, and thereby to a particular structural equation of (5), we deduce that the FIML estimator of the coefficient vector of the exogenous variables in that equation has a distribution which is invariant under the group of orthogonal transformations.[1] This distribution is therefore spherically symmetric, by definition.

In the case of LIML we may take the argument further by using the explicit formulae for the estimator of the exogenous variable coefficient vector. Thus, for equation (1) the LIML estimator of γ (which we denote by s) is given by

$$s = Z_1'Xh,$$

where h is the LIML estimator of b. Since $Z_1'X$ and h are independent, the conditional distribution of s given h is $N(O, I_{K_1})$. Being independent of h, this is also the unconditional distribution. That is, s is $N(O, I_{K_1})$.

If we normalize the structural equation as in (4), the LIML estimator of γ is now

$$q = s(1+r'r)^{1/2}$$

where r is the LIML estimator of β in (4). The conditional distribution of q given r is $N(O, (1+r'r)I_{K_1})$. However, the marginal distribution of r is multivariate Cauchy. Thus, by an elementary integration we find that the density of the LIML estimator of γ under the leading hypothesis (9) is:

$$pdf(q) = \frac{B\left(\frac{n+1}{2}, \frac{K_1+1}{2}\right)\exp\left(-q'q/2\right)}{(2\pi)^{K_1/2}B\left(\frac{1}{2}, \frac{K_1+n+1}{2}\right)}\, {}_1F_1\left(\frac{n}{2}, \frac{n+K_1+1}{2}; \frac{1}{2}q'q\right).$$

The density is the same as that of the instrumental variable estimator of γ when there are no surplus instruments; it may be deduced from equation (15) of Phillips [1984b] by setting $L=0$ and $\gamma=0$.

[1] Note that to the extent that there are exclusion restrictions on a column of C the support of the distribution is confined to a linear subspace of $\boldsymbol{R}^K$. Transformations in $O(K)$ affect only the orientation of the support. The probability measure itself is invariant.

5. COMMENT

Extension of the noncentral results for LIML in Phillips [1984c] to FIML would appear to be very difficult. This is partly because the estimating equations for FIML involve more extensive nonlinearities associated with the full system of structural equations and partly because of the presence of the estimator error covariance matrix in these estimating equations.

Cowles Foundation, Yale University, U. S. A.

REFERENCES

JAMES, A. T., "Normal multivariate analysis and the orthogonal group," *Annals of Mathematical Statistics*, 25 [1954], 40–75.

PHILLIPS, P. C. B., "Exact small sample theory in the simultaneous equations model," in, M. D. Intriligator and Z. Griliches, eds., *Handbook of Econometrics*, North Holland: Amsterdam (1983) Ch. 8, 449–516.

———, "The exact distribution of LIML: I," *International Economic Review*, 25 [1984a], 249–261.

———, "The exact distribution of exogenous variable coefficient estimators," Journal of Econometrics, 26 [1984b], 387–398.

———, "The exact distribution of LIML: II," *International Economic Review* 26 (February, 1985), 21–36.