

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS  
AT YALE UNIVERSITY

Box 2125, Yale Station  
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 840

Note: Cowles Foundation Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. Requests for single copies of a Paper will be filled by the Cowles Foundation within the limits of the supply. References in publications to Discussion Papers (other than acknowledgment that a writer had access to such unpublished material) should be cleared with the author to protect the tentative character of these papers.

JOINT DISTRIBUTION THEORY FOR SOME STATISTICS BASED  
ON LIML AND TSLS

by

Grant Hillier

June 1987

## ABSTRACT

In the context of a single linear structural equation under classical assumptions, we derive the joint conditional density of the LIML endogenous coefficient estimator, and the usual characteristic root arising from the LIML procedure, given the OLS estimates of the reduced form coefficients for the excluded exogenous variables. This provides the joint distributions for various combinations of the statistics commonly used for inference in this model, and is hence an important stepping stone in the analysis of these procedures.

The main result also leads to a new derivation of the density of the LIML estimator itself, and provides a result which is directly comparable to earlier results for IV estimators, including OLS and TSLS. We also consider briefly the density of the LIML structural variance estimator, and the joint density of the LIML and TSLS estimators for the endogenous coefficients.

## 1. INTRODUCTION

The simple structural equation model--comprising a single linear structural equation, together with the reduced form equations for the endogenous variables involved--belongs to a class of models in which the dimension of the set of minimal sufficient statistics exceeds that of the parameter space. Classically, the setting is thus precisely that in which there is greatest doubt over how best to proceed, and this ambiguity can only be resolved by studying the sampling properties of the various suggested procedures.

Progress on the distribution theory for this model has been very slow, because the problems are so difficult, but has accelerated rapidly in the last few years (see [5], [6], [10], [11], [13], [14]). It is, nevertheless, fair to say that even the most basic inference problems for this simplest structural model remain unresolved. This is no doubt due in part to the complexity of the results that have emerged so far--they are certainly difficult to interpret, and are unsuitable, at least at present, for numerical work (see [11] for a discussion of this). But it also, I think, derives from the fact that attention so far has been focused on the marginal distributions of the various statistics, rather than on joint distributions. For hypothesis testing problems, in particular, the joint distributions are essential, and they can also be useful in estimator comparisons.

In this paper we obtain joint distributional results for a number of statistics that commonly form the basis for inference in this model. Specifically, we consider structural coefficient estimators and structural variance estimators based on limited information maximum likelihood and two-

stage least squares. These results should thus provide a basis for analyzing the relative merits of methods based on these two estimation procedures. The key result will be found in Section 3, and Section 4 discusses some more immediate aspects of this result, but the present paper does not attempt a complete study of the uses to which it may be put.

Instead, we use the results in Section 3 to obtain some distribution results for limited information maximum likelihood statistics. Phillips, in [12] and [14], has obtained expressions for the density of the vector of coefficients of the endogenous variables in an operator form, and Rhodes [15] has studied the density of the relevant characteristic root. In Section 5 we obtain considerably simpler forms of both of these results. In particular, our result for the coefficient estimator is directly comparable with the corresponding result for IV estimators (see [5] and [10]). A companion paper will discuss the interpretation of these results for the coefficient estimators and show that, properly interpreted, these highly complex formulae can in fact shed light on the comparative properties of the various estimators.

In Section 6 we outline some of the results that are accessible from the earlier results, and give some special cases. We begin by describing the model, the statistics, and the canonical forms in terms of which the main results will be stated.

## 2. MODEL AND CANONICAL FORMS OF SOME STATISTICS

We consider the structural equation

$$y = Y\beta^* + Z_1\gamma + u \quad (1)$$

and associated reduced form

$$(y, Y) = (Z_1, Z_2) \begin{pmatrix} \pi_1 & \Pi_1 \\ \pi_2 & \Pi_2 \end{pmatrix} + (v, V) \quad (2)$$

where  $y$  is  $T \times 1$ ,  $Y$  is  $T \times n$ ,  $Z_1$  is  $T \times K_1$ ,  $Z_2$  is  $T \times K_2$ , etc., and we assume that  $Z = (Z_1, Z_2)$  ( $T \times K$ ;  $K = K_1 + K_2$ ) is fixed and of full column rank  $K$ . We assume that the rows of  $(v, V)$  in (2) are independent normal vectors with zero means and common covariance matrix

$$\Omega = \begin{pmatrix} \omega_{11} & \omega'_{21} \\ \omega_{21} & \Omega_{22} \end{pmatrix} \begin{matrix} 1 \\ n \end{matrix}, \quad (3)$$

and that  $K_2 \geq n$ . As is well-known, the compatibility of (1) and (2) entails the relations

$$\pi_2 = \Pi_2\beta^*, \quad (4)$$

$\pi_1 = \Pi_1\beta^* + \gamma$ , and  $u = v - V\beta^*$ . From the last of these we find

$$\text{var}(u_t) = \sigma^2 = \omega_{11} - 2\omega'_{21}\beta^* + \beta^{*'}\Omega_{22}\beta^* = \omega^2(1 + \beta'\beta), \quad (5)$$

where  $\omega^2 = \omega_{11} - \omega'_{21}\Omega_{22}^{-1}\omega_{21}$  and

$$\beta = \Omega_{22}^{1/2} (\beta^* - \Omega_{22}^{-1} \omega_{21}) / \omega, \quad (6)$$

and  $u_t$  is independent of the  $t$ -th row of  $Y$  in (1) if and only if  $\beta$  in (6) is zero.

For any matrix  $A$  of full column rank, let  $\bar{P}_A = I - A(A'A)^{-1}A'$ , and define the  $K_2 \times (n+1)$  matrix

$$(x, X) = (Z_2' \bar{P}_{Z_1} Z_2)^{-1/2} Z_2' \bar{P}_{Z_1} [(y - Y \Omega_{22}^{-1} \omega_{21}) / \omega, Y \Omega_{22}^{-1/2}]'. \quad (7)$$

Note that  $(x, X)$  is a simple transformation of the least squares estimator for  $(\pi_2, \Pi_2)$  in (2),

$$(\hat{\pi}_2, \hat{\Pi}_2) = (Z_2' \bar{P}_{Z_1} Z_2)^{-1} Z_2' \bar{P}_{Z_1} (y, Y)'$$

The rows of  $(x, X)$  are independent normal vectors with covariance matrix  $I_{n+1}$  and  $E(x, X) = M(\beta, I_n)$ , where  $M = (Z_2' \bar{P}_{Z_1} Z_2)^{1/2} \Pi_2 \Omega_{22}^{-1/2}$ .

The Two-Stage Least Squares (TSLS) estimator for  $\beta^*$  in (1) is given by

$$b_2 = (Y' (\bar{P}_{Z_1} - \bar{P}_Z) Y)^{-1} Y' (\bar{P}_{Z_1} - \bar{P}_Z) y.$$

Hence, from (7), if we define

$$r_2 = (X'X)^{-1} X'x, \quad (8)$$

the canonical form for  $b_2$ , we have

$$r_2 = \Omega_{22}^{1/2} (b_2 - \Omega_{22}^{-1} \omega_{21}) / \omega . \quad (9)$$

Thus,  $r_2$  is a simple transformation of  $b_2$  and is related to  $b_2$  in exactly the manner in which  $\beta$  is related to  $\beta^*$  in (6). Likewise, one estimator for  $\sigma^2$  in (5) based on TSLS uses the quadratic form

$$q = (y - Yb_2)' (\bar{P}_{Z_1} - \bar{P}_Z) (y - Yb_2) = \omega^2 (x - Xr_2)' (x - Xr_2) = \omega^2 s^2 , \quad (10)$$

where  $s^2 = (x - Xr_2)' (x - Xr_2) = x' \bar{P}_X x$ . Note that  $q$  yields an inconsistent estimator (cf. Phillips [11], pp. 482-484).

Next define

$$S = ((y - Y\Omega_{22}^{-1}\omega_{21})/\omega, Y\Omega_{22}^{-1/2})' \bar{P}_Z ((y - Y\Omega_{22}^{-1}\omega_{21})/\omega, Y\Omega_{22}^{-1/2}) , \quad (11)$$

so that  $S$  is a simple transform of the matrix  $S^* = (y, Y)' \bar{P}_Z (y, Y)$  from which the usual estimator for  $\Omega$  in (2) is constructed.<sup>1</sup> The matrix  $S$  has the central Wishart distribution  $W_{n+1}(m, I_{n+1})$ , where  $m = T-K$ , and is independent of the matrix  $(x, X)$  in (7). The Limited Information Maximum Likelihood (LIML) estimator for  $\beta^*$  in (1) may be defined in terms of a characteristic vector  $\beta_\Delta^*$  that satisfies

$$[S^* - f_1(\hat{\pi}_2, \hat{\Pi}_2)' Z_2' \bar{P}_{Z_1} Z_2(\hat{\pi}_2, \hat{\Pi}_2)] \beta_\Delta^* = 0 ,$$

---

<sup>1</sup>The statistics  $S^*$ ,  $(\hat{\pi}_2, \hat{\Pi}_2)$ , and  $(\tilde{\pi}_1, \tilde{\Pi}_1) = (X_1' X_1)^{-1} X_1' (y, Y)$ , are (jointly) minimal sufficient for  $\Omega$ ,  $\Pi_1$ ,  $\Pi_2$ ,  $\gamma$  and  $\beta$ . As mentioned in the Introduction, the dimension of this set exceeds that of the parameter space by  $(K_2 - n)$ , the degree of overidentification of the equation.

where  $f_1$  is the largest root of the corresponding determinental equation.

Partitioning  $\beta_{\Delta}^* = (\beta_{\Delta 1}^*, \beta_{\Delta 2}^*)'$ , with  $\beta_{\Delta 2}^* n \times 1$ , we define

$b_1 = -\beta_{\Delta 2}^*/\beta_{\Delta 1}^*$ . Thus, if  $\beta_{\Delta}$  is a characteristic vector satisfying

$$[S - f_1(x, X)'(x, X)]\beta_{\Delta} = 0, \quad (12)$$

where  $f_1$  is the largest root of  $|S - f(x, X)'(x, X)| = 0$ , and we define

$r_1 = -\beta_{\Delta 2}/\beta_{\Delta 1}$  as the canonical form for the LIML estimator, where  $\beta_{\Delta}$  is partitioned as  $\beta_{\Delta} = (\beta_{\Delta 1}, \beta_{\Delta 2})'$ , with  $\beta_{\Delta 2} n \times 1$ , then

$$r_1 = \Omega_{22}^{1/2}(b_1 - \Omega_{22}^{-1}\omega_{21})/\omega, \quad (13)$$

exactly as before for the TSLS estimator.

It is straightforward to check that the LIML estimator for  $\sigma^2$  in (5) is a multiple of

$$q_1 = (y - Yb_1)' \bar{P}_{Z_1} (y - Yb_1) = \omega^2(1 + f_1)(x - Xr_1)'(x - Xr_1) = \omega^2 s_1^2,$$

say, where

$$s_1^2 = (1+f_1)(x - Xr_1)'(x - Xr_1) = (1+f_1)[s^2 + (r_1 - r_2)'X'X(r_1 - r_2)]. \quad (14)$$

Notice that, from (14),  $s_1^2 \geq s^2$ , where  $s^2$  is given in (10).

### 3. CONDITIONAL DISTRIBUTIONS

In this section we derive the conditional joint density of  $r_1$  and  $f_1$  given  $(x, X)$  or, what is the same thing, given  $W = (x, X)'(x, X)$ , and the corresponding marginal densities  $\text{pdf}(r_1|W)$  and  $\text{pdf}(f_1|W)$ . This result will provide a good deal of insight into the distribution theory for the

model (1)-(2), and is the key to a number of otherwise very difficult distribution problems involving inference based on Limited Information Maximum Likelihood.

Let

$$T = \begin{bmatrix} s & , & 0 \\ (X'X)^{1/2}r_2 & , & (X'X)^{1/2} \end{bmatrix} ,$$

so that  $T'T = W = (x, X)'(x, X)$  (see (8) and (10) above). From (12),  $T\beta_\Delta$  is a characteristic vector corresponding to the largest root,  $f_1$ , of the matrix  $R = T'^{-1}ST^{-1}$ . From Muirhead [9, Theorem 3.2.5], given  $W$ ,  $R \sim W_{n+1}(m, (TT')^{-1})$ , so that

$$\text{pdf}(R|W) = C_1 \text{etr}\left\{-\frac{1}{2}TT'R\right\} |R|^{(m-n-2)/2} |W|^{m/2} \quad (15)$$

where  $C_1 = [2^{m(n+1)/2} \Gamma_{n+1}(m/2)]^{-1}$ ,  $\text{etr}(A)$  denotes  $\exp\{\text{tr}(A)\}$ , and

$$\Gamma_r(t) = \pi^{r(r-1)/4} \prod_{i=1}^r \Gamma(t - \frac{1}{2}(i-1))$$

denotes the multivariate Gamma function.

We shall first obtain the joint conditional distribution of  $f_1$ , the largest characteristic root of  $R$ , and  $h$ , say, the corresponding characteristic vector, with  $h$  normalized so that  $h'h = 1$ . With  $h$  partitioned as  $h' = (h_1, h_2')$ ,  $h_2$   $n \times 1$ , we then define  $\tilde{r} = -h_1^{-1}h_2$  and transform  $h \rightarrow \tilde{r}$ . Since  $T\beta_\Delta = h$  and  $r_1 = -\beta_{\Delta 1}^{-1}\beta_{\Delta 2}$  we have  $\tilde{r} = s^{-1}(X'X)^{1/2}(r_1 - r_2)$ , or

$$r_1 = r_2 + s(X'X)^{-1/2} \bar{r} . \quad (16)$$

That is, for fixed  $W$  (i.e., fixed  $s$ ,  $r_2$ , and  $X'X$ )  $r_1$  is a simple transformation of  $\bar{r}$  and we may then transform  $\bar{r} \rightarrow r_1$ , giving the conditional joint density  $\text{pdf}(r_1, f_1|W)$ .

Thus, we first make a transformation from  $R$  to its characteristic roots and vectors:  $R \rightarrow (F, H)$ , where  $F = \text{diag}(f_1, f_2, \dots, f_{n+1})$ ,  $f_1 > f_2 > \dots > f_{n+1}$ , and  $H$  is an orthogonal matrix such that  $R = HFH'$ , so that the columns of  $H$  are orthonormal characteristic vectors of  $R$ . The transformation  $R \rightarrow (F, H)$  is made one-to-one by imposing the requirement that the elements in the first row, say, of  $H$  are positive, and the volume element transforms as

$$(dR) = \prod_{1 \leq j} (f_1 - f_j) \prod_{i=1}^{n+1} df_i (H'dH) ,$$

where we follow the notational conventions in Muirhead [9] (see, in particular, [9 pp. 103-105]). The expression  $(H'dH)$  here denote the unnormalized invariant measure on the orthogonal group  $O(n+1)$ .

Now partition  $H = (h, H_1)$ , where  $H_1$  is an element of the Stiefel manifold  $V_{n+1, n}$  orthogonal to  $h$ . Following Constantine and Muirhead [2, Lemma 1] (see also the argument in James [8, pp. 59-60]), we may generate the Stiefel manifold orthogonal to  $h$  by setting  $H_1 = GK$ , where  $G$  is a fixed matrix such that  $G'G = I_n$  and  $GG' = I - hh'$ , and  $K$  ranges over the orthogonal group  $O(n)$ . The invariant differential form  $(H'dH)$  decomposes as

$$(H'dH) = (h'dh)(K'dK) ,$$

where  $(h'dh)$  denotes the (unnormalized) invariant measure on the surface of the unit sphere  $S_{n+1}$ , and  $(K'dK)$  that on the orthogonal group  $O(n)$ .

Transforming  $R \rightarrow (F, h, H_1)$  in (15), setting  $H_1 = GK$ , and integrating over  $O(n)$  using James [8, Equation (23)] we obtain

$$\text{pdf}(F, h|W) = C'_1 \exp\left\{-\frac{1}{2}f_1 h' T T' h\right\} |F|^{(m-n-2)/2} |W|^{m/2} {}_0F_0^{(n)}\left[-\frac{1}{2}G' T T' G, F_2\right] \prod_{i < j} (f_i - f_j) \quad (17)$$

where we have partitioned  $F$  into

$$F = \begin{bmatrix} f_1 & 0 \\ 0 & f_2 \end{bmatrix},$$

and  ${}_0F_0^{(n)}(\cdot)$  denotes the two-matrix-argument generalized hypergeometric series (see Muirhead [9, p. 259]). The constant  $C'_1 = [\pi^{n^2/2} C_1 / \Gamma_n(n/2)]$  because, in integrating over  $O(n)$  the result has been divided by  $2^n$  to take account of the initial sign restriction on the elements of  $H$ . Although we have retained the abbreviation  $\text{pdf}(\cdot)$  on the left in (17) it should be noted that (17) denotes a probability measure on  $F \times S_{n+1}$ , where  $F$  denotes the set  $\{f_1, \dots, f_{n+1}; f_1 > f_2 > \dots > f_{n+1} > 0\}$  and  $S_{n+1}$  denotes the surface of the unit sphere in  $n+1$  dimensions, evaluated with respect to ordinary Lebesgue measure on  $F$  and the invariant measure on  $S_{n+1}$ .

Next we integrate out  $F_2$ . Define  $\bar{F}_2$  by  $F_2 = f_1 \bar{F}_2$  (Jacobian  $f_1^n$ ), so that  $1 > \bar{F}_2 > \dots > \bar{F}_{n+1} > 0$ , and note that

$$\prod_{i < j} (f_i - f_j) = f_1^{n(n+1)/2} \prod_{i=2}^{n+1} (1 - \bar{f}_i) \prod_{2 \leq i < j} (\bar{f}_i - \bar{f}_j) .$$

Using Phillips [14, Equation (27)] we obtain, on integrating over the set  $1 > \bar{f}_2 > \dots > \bar{f}_{n+1} > 0$ ,

$$\begin{aligned} \text{pdf}(f_1, h|W) &= C_2 \exp\left\{-\frac{1}{2}f_1 h' T T' h\right\} f_1^{m(n+1)/2-1} |W|^{m/2} \\ &\quad {}_1F_1((m-1)/2, (m+n+2)/2; -\frac{1}{2}f_1 G' T T' G) \\ &= C_2 \text{etr}\left\{-\frac{1}{2}f_1 W\right\} f_1^{m(n+1)/2-1} |W|^{m/2} \\ &\quad {}_1F_1((n+3)/2, (m+n+2)/2, \frac{1}{2}f_1 G' T T' G) \end{aligned} \quad (18)$$

where  $C_2 = [\Gamma_n((n+3)/2) \Gamma_n((m-1)/2) / 2^{m(n+1)/2} \Gamma_n((m+n+2)/2) \Gamma_{n+1}(m/2)]$ , and the second line of (18) follows from the first on using the Kummer relation for the confluent hypergeometric function (James [8, Equation (51)]).

The marginal conditional density of  $f_1$  given  $W$  is easily obtained from (18) by integrating out  $h$  (remembering that  $h_1 > 0$  by construction). This gives

$$\begin{aligned} \text{pdf}(f_1|W) &= C'_2 \text{etr}\left\{-\frac{1}{2}f_1 W\right\} f_1^{m(n+1)/2-1} |W|^{m/2} \\ &\quad \sum_{\ell=0}^{\infty} \sum_{\lambda}^* \frac{((n+3)/2)_{\lambda} (n/2)_{\lambda}}{\ell! ((n+1)/2)_{\lambda} ((m+n+2)/2)_{\lambda}} C_{\lambda} \left(\frac{1}{2}f_1 W\right) . \end{aligned} \quad (19)$$

Here  $\lambda = (\ell_1, \ell_2, \dots, \ell_{n+1})$ ,  $\ell_1 \geq \ell_2 \geq \dots \geq \ell_{n+1} \geq 0$ ,  $\sum_{i=1}^{n+1} \ell_i = \ell$ , denotes an ordered partition of  $\ell$  into  $n+1$  or less parts, but the notation  $\sum_{\lambda}^*$  indicates that the sum is to be taken over only those partitions  $\lambda$  with  $\ell_{n+1} = 0$ , i.e., of  $n$  or less parts. Also,

$$c_2' = [\pi^{(n+1)/2} c_2 / \Gamma((n+1)/2)] .$$

Now, partition  $h' = (h_1, h_2)$ , with  $h_2$   $n \times 1$ , and define  $\tilde{r} = -h_1^{-1} h_2$  ( $n \times 1$ ). The transformation of the invariant measure  $(h' dh)$  on  $S_{n+1}$  is given in the following:

Lemma: Let  $(h' dh)$  denote the invariant measure on the  $S_{n+1} : h'h = 1$ , and define  $\tilde{r} = -h_1^{-1} h_2$ , with  $h' = (h_1, h_2)$  and  $h_2$   $n \times 1$ . Then

$$(h' dh) = 2(1 + \tilde{r}' \tilde{r})^{-(n+1)/2} \bigwedge_{i=1}^n d\tilde{r}_i . \quad (20)$$

Proof: Let

$$\tilde{v} = \begin{pmatrix} \tilde{r}' \\ I_n \end{pmatrix} (I + \tilde{r} \tilde{r}')^{-1/2} , \quad (21)$$

and note that under the transformation  $\tilde{r} = -h_1^{-1} h_2$ ,

$$h = \varepsilon \begin{pmatrix} 1 \\ -\tilde{r} \end{pmatrix} (1 + \tilde{r}' \tilde{r})^{-1/2} ; \quad \varepsilon = \pm 1, \quad (22)$$

because  $h'h = 1$  implies  $h_1 = (1 + \tilde{r}' \tilde{r})^{-1/2}$  (up to sign). Note that  $\tilde{v}' h = 0$ . Differentiating (22),

$$dh = \varepsilon (1 + \tilde{r}' \tilde{r})^{-3/2} \begin{pmatrix} -\tilde{r}' \\ \tilde{r} \tilde{r}' - (1 + \tilde{r}' \tilde{r}) I_n \end{pmatrix} d\tilde{r} ,$$

so that

$$\bar{V}'dh = \epsilon(1 + \bar{r}'\bar{r})^{-1/2} [I + \bar{r}\bar{r}']^{-1/2} d\bar{r}.$$

By definition (cf. James [7]),  $(h'dh)$  is the exterior product of the  $n$  terms on the left, and this yields the right side of (20) on allowing for  $\epsilon = +1$  and  $\epsilon = -1$ .

Using the Lemma we may transform  $h \rightarrow \bar{r}$  in (18), but note that the expression on the right in (20) must in our case be halved because  $h_1 > 0$  by construction. Also, note that the argument matrix  $G'TT'G$  may be replaced by  $T'GG'T = T'(I - hh')T$ , because these matrices have the same characteristic roots, and that

$$I - hh' = (\bar{r}, I_n)'(I + \bar{r}\bar{r}')^{-1}(\bar{r}, I_n).$$

Hence we have

$$\begin{aligned} \text{pdf}(\bar{r}, f_1 | W) &= C_2 \text{etr} \left\{ -\frac{1}{2} f_1' W \right\} f_1^{m(n+1)/2-1} |W|^{m/2} (1 + \bar{r}'\bar{r})^{-(n+1)/2} \\ &\quad {}_1F_1((n+3)/2, (m+n+2)/2; \frac{1}{2} f_1' (\bar{r}, I) T T' (\bar{r}, I)' (I + \bar{r}\bar{r}')^{-1}) . \end{aligned} \quad (23)$$

Finally we transform from  $\bar{r}$  to  $r_1$  using (16), the Jacobian being  $(s^2)^{-n/2} |X'X|^{1/2}$ . Combining the Jacobian and the term  $(1 + \bar{r}'\bar{r})^{-(n+1)/2}$  in (23) we have

$$\begin{aligned} (s^2)^{-n/2} |X'X|^{1/2} (1 + \bar{r}'\bar{r})^{-(n+1)/2} &= |W|^{1/2} [s^2 + (r_1 - r_2)' X' X (r_1 - r_2)]^{-(n+1)/2} \\ &= |W|^{1/2} [(x - Xr_1)' (x - Xr_1)]^{-(n+1)/2} \\ &= |W|^{1/2} [(1 + r_1' r_1) v' W v]^{-(n+1)/2} \end{aligned} \quad (24)$$

where we have put

$$v = \begin{pmatrix} 1 \\ -r_1 \end{pmatrix} (1 + r_1' r_1)^{-1/2} \in S_{n+1}. \quad (25)$$

Also, note that, using (16),

$$(\tilde{r}, I_n)' = (\tilde{r}, I) \begin{bmatrix} s & 0 \\ (X'X)^{1/2} r_2 & (X'X)^{1/2} \end{bmatrix} = (X'X)^{1/2} (r_1, I_n)' ,$$

and

$$I + \tilde{r}\tilde{r}' = I + \frac{(X'X)^{1/2} (r_1 - r_2)(r_1 - r_2)' (X'X)^{1/2}}{s^2} .$$

Hence, the argument of the confluent hypergeometric function in (23) may be replaced by

$$\begin{aligned} & (I + r_1 r_1')^{1/2} \left[ (X'X)^{-1} + \frac{(r_1 - r_2)(r_1 - r_2)'}{s^2} \right]^{-1} (I + r_1 r_1')^{1/2} \\ &= (I + r_1 r_1')^{1/2} \left[ (r_1, I) W^{-1} (r_1, I)' \right]^{-1} (I + r_1 r_1')^{1/2} \\ &= \left[ v' W^{-1} v \right]^{-1} \end{aligned} \quad (26)$$

where

$$v = \begin{pmatrix} r_1' \\ I_n \end{pmatrix} (I + r_1 r_1')^{-1/2} , \quad (27)$$

so that  $(v, v) \in O(n+1)$  (see (25) above). Since

$$[(v, v)' W (v, v)]^{-1} = (v, v)' W^{-1} (v, v) ,$$

we may also write

$$\begin{aligned} (v' W^{-1} v)^{-1} &= v' W v - v' W v (v' W v)^{-1} v' W v \\ &= (I + r_1 r_1')^{1/2} X' \left[ I - \frac{(x - X r_1)(x - X r_1)'}{(x - X r_1)'(x - X r_1)} \right] X (I + r_1 r_1')^{1/2} . \end{aligned} \quad (28)$$

Using (24) and (26), (23) becomes, on transforming from  $\tilde{r} \rightarrow r_1$  ,

$$\begin{aligned} \text{pdf}(r_1, f_1 | W) &= C_2 \text{etr} \left\{ -\frac{1}{2} f_1' W \right\} f_1^{m(n+1)/2-1} |W|^{(m+1)/2} (1 + r_1' r_1)^{-(n+1)/2} \\ &\quad (v' W v)^{-(n+1)/2} {}_1F_1((n+3)/2, (m+n+2)/2; \frac{1}{2} f_1' (v' W^{-1} v)^{-1}) . \end{aligned} \quad (29)$$

The marginal conditional density of  $r_1$  itself may be obtained from (29) by a simple integration over  $f_1 > 0$  .

#### 4. DISCUSSION OF THE CONDITIONAL RESULT

It is instructive at this point to temporarily recast the distribution theory in terms of the "raw" LIML estimator  $b_1$  using equation (13). At the same time we shall replace the conditioning matrix  $W$  by the "raw" moment matrix

$$\bar{W} = (\hat{\pi}_2, \hat{\Pi}_2)' Z_2' \bar{P}_{Z_1} Z_2 (\hat{\pi}_2, \hat{\Pi}_2) , \quad (30)$$

where  $\bar{W}$  is related to  $W$  by

$$W = \begin{bmatrix} \omega^{-1} & , & 0 \\ -\omega^{-1}\alpha & , & \Omega_{22}^{-1/2} \end{bmatrix}' \bar{W} \begin{bmatrix} \omega^{-1} & , & 0 \\ -\omega^{-1}\alpha & , & \Omega_{22}^{-1/2} \end{bmatrix}$$

(see (7) above), with  $\alpha = \Omega_{22}^{-1}\omega_{21}$ .

Transforming  $r_1 \rightarrow b_1$  in (29), the Jacobian is  $\omega^{-n}|\Omega_{22}|^{1/2}$  and we find

$$\begin{aligned} \text{pdf}(b_1, f_1 | \bar{W}) &= C_2 \text{etr} \left\{ -\frac{1}{2} f_1 \Omega^{-1} \bar{W} \right\} |\Omega|^{-m/2} f_1^{m(n+1)/2-1} |\bar{W}|^{(m+1)/2} \\ &\quad [(1 + b_1' b_1) v_1' \bar{W} v_1]^{-(n+1)/2} \\ &\quad \times {}_1F_1 \left[ (n+3)/2, (m+n+2)/2; \frac{1}{2} f_1 v_1' \Omega^{-1} v_1 (v_1' \bar{W}^{-1} v_1)^{-1} \right] \end{aligned} \quad (31)$$

where  $v_1' = (1 + b_1' b_1)^{-1/2} (1, -b_1')$  and

$$v_1 = \begin{bmatrix} b_1' \\ I_n \end{bmatrix} (I + b_1 b_1')^{-1/2}.$$

The matrix  $\bar{W}$  here has the non-central Wishart distribution with  $K_2$  degrees of freedom, covariance matrix  $\Omega$ , and matrix of non-centrality parameters  $\Omega^{-1}(\beta^*, I)' \Pi_2' Z_2' \bar{P} Z_2 \Pi_2(\beta^*, I) \Omega^{-1}$ .

Several aspects of the result in equation (31) deserve comment. First, it is apparent from (31) that, given  $\bar{W}$ , the LIML quantities  $(b_1, f_1)$  contain no direct information about the structural coefficient vector  $\beta^*$ ; this information is transmitted entirely through  $\bar{W}$ . However, since information about  $\Omega$  is clearly pertinent to the estimation of  $\beta^*$ , and (31) does depend on  $\Omega$ , the LIML quantities  $(b_1, f_1)$  can be expected to

contribute relevant information even after  $\bar{W}$  is known.<sup>2</sup> Indeed, (31) implies that  $(b_1, f_1)$  contain information on  $\Omega$  that is independent of  $\beta^*$ , unlike  $\bar{W}$  itself (and hence  $q$  and  $b_2$  which are functions of  $\bar{W}$ ). So, at this heuristic level, (31) suggests that  $b_1$  (in particular) may well contain valuable information not available from knowledge of  $\bar{W}$  alone.

The fact that the conditional density (31) does not depend on  $\beta^*$  also says something about hypothesis tests concerning  $\beta^*$ . For, in the joint density

$$\text{pdf}(b_1, f_1, \bar{W}) = \text{pdf}(b_1, f_1 | \bar{W}) \text{pdf}(\bar{W}), \quad (32)$$

only the second factor depends on  $\beta^*$ . Therefore, for fixed values of the nuisance parameters  $\Omega$  and  $\Pi_2$ , the likelihood ratio for hypotheses concerning  $\beta^*$  will not involve the LIML quantities  $(b_1, f_1)$ . Hence, by the Neyman-Pearson Lemma, the best tests about  $\beta^*$  based on functions of  $b_1$ ,  $f_1$ , and  $\bar{W}$  never involve the LIML quantities  $(b_1, f_1)$ . Since the TSLS estimator  $b_2$ ,  $q$  in (10) (the quadratic form used to construct the inconsistent TSLS-based variance estimator), and  $\bar{W}_{22} = \hat{\Pi}_2' Z_2' \bar{P}_{Z_1} Z_2 \hat{\Pi}_2$  are one-to-one functions of  $\bar{W}$ , this says that, among tests based on  $(b_1, f_1, b_2, q, \bar{W}_{22})$ , the best tests depend only upon the TSLS quantities  $(b_2, q, \bar{W}_{22})$ .<sup>3</sup>

This, at first, seems a striking result: it seems to say that the LIML

---

<sup>2</sup>The situation parallels that in which a pair of statistics is jointly sufficient for a pair of parameters, but neither is individually sufficient.

<sup>3</sup>Note that  $\bar{W}_{22}$  is the matrix that is used to construct a consistent estimate of the asymptotic covariance matrix of both  $b_1$  and  $b_2$ .

quantities  $(b_1, f_1)$  can be ignored in testing  $\beta^*$ . However, the limitation on the class of tests considered is critical: all functions of the reduced form error moment matrix  $S^* = (y, Y)' \bar{P}_Z (y, Y)$  other than  $(b_1, f_1)$  are excluded. Thus, although the result does cover some t-type tests based on the asymptotic distribution of  $b_1$  and  $b_2$ , tests such as the (limited information) likelihood ratio tests, and those that use the consistent TSLS-based variance estimator (a multiple of  $(1, -b_2) S^* (1, -b_2)'$ ) are excluded.

Thus, while our result does have something to say about testing, the support it offers TSLS-based procedures can only be regarded as very tentative at best.

##### 5. UNCONDITIONAL DENSITIES

In this section we derive the unconditional densities  $\text{pdf}(f_1)$  and  $\text{pdf}(r_1)$  from the conditional results (19) and (29). A version of the former result is essentially given in Rhodes [15], but Rhodes's [15] result involves some unknown constants; the result given here does not.

The density of the matrix  $W$  is given by

$$\text{pdf}(W) = C_3 \text{etr} \left\{ -\frac{1}{2} W \right\} |W|^{(K_2 - n - 2)/2} {}_0F_1 \left[ K_2/2; \frac{1}{4} (\beta, I)' M' M (\beta, I) W \right] \quad (33)$$

$$\text{where } C_3 = \left[ \text{etr} \left\{ -\frac{1}{2} M' M (I + \beta \beta') \right\} / \Gamma_{n+1}(K_2/2) 2^{(n+1)K_2/2} \right].$$

Multiplying (19) by (33) and integrating over  $W > 0$  gives the unconditional result

$$\begin{aligned}
\text{pdf}(f_1) &= C_4 f_1^{m(n+1)/2-1} (1 + f_1)^{-(n+1)(m+K_2)/2} \\
&\sum_{j,l=0}^{\infty} \sum_{\alpha,\lambda} \frac{(n/2)_{\lambda} ((n+3)/2)_{\lambda} (n/2)_{\alpha}}{j! l! ((n+1)/2)_{\lambda} ((m+n+2)/2)_{\lambda} (K_2/2)_{\alpha} ((n+1)/2)_{\alpha}} f_1^l (1 + f_1)^{-(j+l)} \\
&c(\alpha, \lambda) C_{\alpha} \left[ \frac{1}{2} M' M (I + \beta \beta') \right] \quad (34)
\end{aligned}$$

where

$$c(\alpha, \lambda) = \sum_{\rho \in \alpha, \lambda} \frac{((m + K_2)/2)_{\rho} ((n+1)/2)_{\rho}}{(n/2)_{\rho}} e^{(\theta_{\rho}^{\alpha, \lambda})^2} C_{\rho}(I_n) / C_{\alpha}(I_n) , \quad (35)$$

$\alpha$  ,  $\lambda$  are now partitions with  $n$  or less parts, and

$$C_4 = [\Gamma_{n+1} ((m + K_2)/2)^2]^{(n+1)(m+K_2)/2} C_2' C_3 .$$

Note that, in the leading case characterized by  $M = 0$  (i.e.,  $\Pi_2 = 0$ ), (34) reduces to

$$\begin{aligned}
\text{pdf}(f_1) &= C_4 f_1^{m(n+1)/2-1} (1 + f_1)^{-(n+1)(m+K_2)/2} \\
&{}_2F_1 \left[ (n+3)/2, (m + K_2)/2, (m+n+2)/2, \frac{f_1}{1 + f_1} I_n \right] . \quad (36)
\end{aligned}$$

Turning now to the joint density, equation (29), we have,

$$\text{pdf}(r_1, f_1) = \int_{W>0} \text{pdf}(r_1, f_1 | W) \text{pdf}(W) (dW) ,$$

with  $\text{pdf}(W)$  given by (33) above. In this integral we first transform to  $\bar{W} = (v, V)' W(v, V)$  , with  $v$  and  $V$  given in equations (25) and (27) respectively. Partitioning  $\bar{W}$  as

$$\bar{W} = \begin{bmatrix} \bar{w}_{11} & \bar{w}'_{21} \\ \bar{w}_{21} & \bar{w}_{22} \end{bmatrix} \begin{matrix} 1 \\ n \end{matrix},$$

we then transform to  $B = \bar{w}_{22} - \bar{w}_{11}^{-1} \bar{w}_{21} \bar{w}'_{21}$ ,  $z = \bar{w}_{21}$ , and  $a = \bar{w}_{11}$  (the Jacobian is unity), giving

$$\begin{aligned} \text{pdf}(r_1, f_1, a, z, B) &= C_2 C_3 (1 + r_1' r_1)^{-(n+1)/2} f_1^{m(n+1)/2-1} \\ &\quad \text{etr} \left\{ -\frac{1}{2}(1+f_1)B \right\} \exp \left\{ -\frac{1}{2}(1+f_1)(a + a^{-1} z' z) \right\} a^{(m+K_2-2n-2)/2} \\ &\quad |B|^{(m+K_2-n-1)/2} {}_1F_1 \left[ (n+3)/2, (m+n+2)/2; \frac{1}{2} f_1 B \right] \\ &\quad {}_0F_1 \left[ K_2/2; \frac{1}{2} \bar{M}(\beta, I)(v, V) \begin{bmatrix} a & z' \\ z & B + a^{-1} z z' \end{bmatrix} (v, V)'(\beta, I)' \bar{M}' \right] \end{aligned} \quad (37)$$

where  $\bar{M}$  is any  $n \times n$  matrix such that  $\bar{M}' \bar{M} = \frac{1}{2} M' M$ . We need to integrate (37) over  $a > 0$ ,  $B > 0$ , and  $z \in R^n$ .

To facilitate the integration we write the last term in (37) as an inverse Laplace transform (cf. Herz [4], Constantine [1]):

$$\begin{aligned} &a_n \Gamma_n(K_2/2) \int_{\text{Re}(Z) > 0} \text{etr}(Z) |Z|^{-K_2/2} \text{etr} \left\{ \frac{1}{2} Q_{22} B \right\} \\ &\quad \exp \left\{ \frac{1}{2} (a q_{11} + 2 z' q_{21} + a^{-1} z' Q_{22} z) \right\} (dZ) \end{aligned} \quad (38)$$

where  $a_n = [2^{n(n-1)/2} / (2\pi i)^{n(n+1)/2}]$  and we have put

$$\begin{bmatrix} q_{11} & q'_{21} \\ q_{21} & Q_{22} \end{bmatrix} = (v, V)'(\beta, I)' \bar{M}' Z^{-1} \bar{M}(\beta, I)(v, V).$$

Using (38) in (37) it is straightforward to integrate over  $B > 0$ , which yields

$$2^{\frac{n(m+K_2)}{2}} \Gamma_n((m+K_2)/2) \left| (1+f_1)I - Q_{22} \right|^{-(m+K_2)/2} \\ \times {}_2F_1\left((m+K_2)/2, (n+3)/2, (m+n+2)/2; f_1[(1+f_1)I - Q_{22}]^{-1}\right), \quad (39)$$

and to integrate out  $z$  by completing the square, which yields

$$(2\pi)^{n/2} a^{n/2} \left| (1+f_1)I - Q_{22} \right|^{-1/2} \exp\left\{\frac{1}{2} a q'_{21} [(1+f_1)I - Q_{22}]^{-1} q_{21}\right\}. \quad (40)$$

Now, combining the exponential term in (40) with those involving  $a$  in (37) and (38) we have, in the exponent,

$$-\frac{1}{2} a \left\{ (1+f_1) - q_{11} - q'_{21} [(1+f_1)I - Q_{22}]^{-1} q_{21} \right\} \\ - \frac{1}{2} a \left\{ |(1+f_1)I_{n+1} - Q| / |(1+f_1)I_n - Q_{22}| \right\}.$$

Hence, integrating over  $a > 0$  yields the terms

$$2^{\frac{(m+K_2-n)}{2}} \Gamma((m+K_2-n)/2) \left| (1+f_1)I_{n+1} - Q \right|^{-\frac{(m+K_2-n)}{2}} \left| (1+f_1)I_n - Q_{22} \right|^{\frac{(m+K_2-n)}{2}}.$$

Collecting these results we have an expression for  $\text{pdf}(r_1, f_1)$  as an inverse Laplace transform:

$$\begin{aligned}
\text{pdf}(r_1, f_1) &= C_5 f_1^{m(n+1)/2-1} (1 + f_1)^{-(m+K_2)(n+1)/2} (1 + r_1' r_1)^{-(n+1)/2} \\
&\quad a_n \Gamma_n(K_2/2) \int_{\text{Re}(Z) > 0} \text{etr}(Z) |Z|^{-K_2/2} |I_n - \bar{M}(I + \beta\beta') \bar{M}' Z^{-1} / (1 + f_1)|^{-(m+K_2-n)/2} \\
&\quad |I_n - Q_{22}/(1+f_1)|^{-(n+1)/2} \\
&\quad {}_2F_1 \left[ (m+K_2)/2, (n+3)/2, (m+n+2)/2; f_1 [(1+f_1)(I - Q_{22}/(1+f_1))]^{-1} \right] (dZ) \quad (41)
\end{aligned}$$

where

$$C_5 = [2^{(n+1)(m+K_2)/2} \Gamma_{n+1}((m+K_2)/2) C_2 C_3] = \pi^{-(n+1)/2} \Gamma((n+1)/2) C_4 \quad (42)$$

(see equations (35) and (36) above).

In the leading case characterized by  $M = 0$  the joint density may be obtained at once from (41). For, in that case,  $Q_{22} = 0$  and  $\bar{M} = 0$ , so that the integral in (41) reduces to simply

$${}_2F_1 \left[ (m + K_2)/2, (n+3)/2, (m+n+2)/2; \frac{f_1}{1 + f_1} I_n \right]$$

and we have

$$\text{pdf}(r_1, f_1) = \text{pdf}(r_1) \text{pdf}(f_1), \quad (43)$$

with  $\text{pdf}(f_1)$  given by equation (36) above, and (from (42)),

$$\text{pdf}(r_1) = \Gamma((n+1)/2) [\pi(1 + r_1' r_1)]^{-(n+1)/2} \quad (44)$$

(cf. Phillips [12]). That is, in this leading case  $r_1$  and  $f_1$  are independent,  $f_1$  has the density (36), and  $r_1$  has the Cauchy distribution.

In the general case we first use Davis [3, Equations (6.19) and (2.2)] to reduce the last two terms in the integral in (41) to the form

$$\sum_{j, \ell=0}^{\infty} \sum_{\alpha, \lambda} \frac{((m + K_2)/2)_{\lambda} ((n+3)/2)_{\lambda}}{j! \ell! ((m+n+2)/2)_{\lambda} ((n+1)/2)_{\lambda}} f_1^{\ell} (1 + f_1)^{-(j+\ell)} a(\alpha, \lambda) C_{\alpha}(Q_{22}) \quad (45)$$

where

$$a(\alpha, \lambda) = \sum_{\rho \in \alpha \cdot \lambda} ((n+1)/2)_{\rho} (\theta_{\rho}^{\alpha, \lambda})^2 C_{\rho}(I_n) / C_{\alpha}(I_n) \quad (46)$$

Recall that, from the definition of  $Q_{22}$  below equation (38),  $C_{\alpha}(Q_{22})$  may be written as  $C_{\alpha}(\bar{M}(\beta, I) V V' (\beta, I)' \bar{M}' Z^{-1})$ , and so, on expanding the term

$\left| I_n - \bar{M}(I + \beta \beta') \bar{M}' Z^{-1} \right|^{-(m+K_2-n)/2}$  in (41) and completing the inverse Laplace transform, we have

$$\begin{aligned} \text{pdf}(r_1, f_1) &= C_5 f_1^{m(n+1)/2-1} (1+f_1)^{-(m+K_2)(n+1)/2} (1 + r_1' r_1)^{-(n+1)/2} \\ &\sum_{j, k, \ell=0}^{\infty} \sum_{\alpha, \kappa, \lambda} \sum_{\emptyset \in \alpha \cdot \kappa} \frac{((m + K_2)/2)_{\lambda} ((n+3)/2)_{\lambda} ((m + K_2 - n)/2)_{\kappa}}{j! k! \ell! ((m+n+2)/2)_{\lambda} ((n+1)/2)_{\lambda} (K_2/2)_{\emptyset}} a(\alpha, \lambda) \\ &f_1^{\ell} (1+f_1)^{-(j+k+\ell)} \theta_{\emptyset}^{\alpha, \kappa} C_{\emptyset}^{\alpha, \kappa} \left[ \bar{M}(\beta, I) V V' (\beta, I)' \bar{M}', \bar{M}(I + \beta \beta') \bar{M}' \right] \quad (47) \end{aligned}$$

Note that, in general,  $r_1$  and  $f_1$  are not independent. Equation (34), the marginal density of  $f_1$ , may be obtained from (47) by integrating out  $r_1$  but we omit the somewhat tedious details of this calculation.

Integrating over  $f_1 > 0$  in (47) we obtain the marginal density of  $r_1$  in the form

$$\text{pdf}(r_1) = C_6 (1 + r_1' r_1)^{-(n+1)/2} \sum_{\alpha, \kappa; \emptyset} \frac{((m + K_2 - n)/2)_{\kappa}}{j! k! (K_2/2)_{\emptyset}} g(\alpha, \kappa) \\ \theta_{\emptyset}^{\alpha, \kappa} C_{\emptyset}^{\alpha, \kappa} \left[ \bar{M} (I + \beta r_1') (I + r_1 r_1')^{-1} (I + r_1 \beta') \bar{M}', \bar{M} (I + \beta \beta') \bar{M}' \right] \quad (48)$$

where

$$g(\alpha, \kappa) = \frac{(K_2(n+1)/2)_f}{((m+K_2)(n+1)/2)_f} \sum_{\ell=0}^{\infty} \sum_{\lambda} \frac{((m+K_2)/2)_{\lambda} ((n+3)/2)_{\lambda} (m(n+1)/2)_{\ell} a(\alpha, \lambda)}{\ell! ((m+n+2)/2)_{\lambda} ((n+1)/2)_{\lambda} (f+(m+K_2)(n+1)/2)_{\ell}} \quad (49) \\ (f = j+k)$$

$$\text{and} \quad C_6 = [\Gamma(m(n+1)/2) \Gamma(K_2(n+1)/2) C_5 / \Gamma((m + K_2)(n+1)/2)] .$$

Equation (48) is an explicit version of the result given in Phillips [14]. It is directly comparable with the results for IV estimators given in [5] and [10] (in [10]  $\bar{M}$  is  $\bar{\Pi}_{22}$ , while in [5]  $\bar{M}$  is  $(TA)^{1/2}$ ; the other notation is the same).

When  $n = 1$  equations (47) and (48) simplify greatly, because in that case the invariant polynomials  $C_{\emptyset}^{\alpha, \kappa}(\cdot, \cdot)$  are simply powers of the two scalar arguments. Moreover, from (46) we have  $a(j, \ell) = (1)_{j+\ell}$ , and from (49) (when  $n = 1$ )

$$g(j, k) = \frac{(K_2)_f (1)_j}{(m + K_2)_f} {}_4F_3(2, m, j+1, (m+K_2)/2, 1, j+k+m+K_2, (m+3)/2; 1) . \quad (50)$$

Hence, from (47), when  $n = 1$

$$\begin{aligned}
\text{pdf}(r_1, f_1) &= C_5 f_1^{m-1} (1 + f_1)^{-(m+K_2)} (1 + r_1^2)^{-1} \\
&\sum_{j,k,\ell=0}^{\infty} \frac{((m+K_2)/2)_{\ell} (2)_{\ell} ((m+K_2-1)/2)_k (1)_{j+\ell}}{j! k! \ell! ((m+3)/2)_{\ell} (1)_{\ell} (K_2/2)_{j+k}} \\
&\left[ \mu^2 / (1+f_1) \right]^{j+k} \left[ f_1 / (1 + f_1) \right]^{\ell} \left[ (1 + r_1 \beta)^2 / (1 + r_1^2) \right]^j (1 + \beta^2)^k \quad (51)
\end{aligned}$$

where we have put  $\mu^2 = \bar{M}'\bar{M} = \frac{1}{2}M'M$  (a scalar), and, when  $n = 1$ ,

$$C_5 = [\Gamma((m+K_2)/2)\Gamma((m+K_2-1)/2)/\Gamma(1/2)\Gamma(K_2/2)\Gamma((K_2-1)/2)\Gamma(m/2)\Gamma((m+3)/2)].$$

Likewise, when  $n = 1$ , (48) reduces to

$$\begin{aligned}
\text{pdf}(r_1) &= C_6 (1 + r_1^2)^{-1} \sum_{j,k=0}^{\infty} \frac{((m+K_2-1)/2)_k (1)_j (K_2)_{j+k}}{j! k! (K_2/2)_{j+k} ((m+K_2)/2)_{j+k}} (\mu^2)^{j+k} (1 + \beta^2)^k \\
&\left[ (1+r_1\beta)^2 / (1+r_1^2) \right]^j {}_4F_3 \left[ 2, m, j+1, (m+K_2)/2, 1, j+k+m+K_2, (m+2)/2; 1 \right] \quad (52)
\end{aligned}$$

## 6. SOME FURTHER RESULTS

A number of statistics of interest are functions of  $(r_1, f_1, W)$ , and both joint and marginal distributions for these are readily obtainable from (29) and (33). In this section we give two special cases that will indicate the possibilities. The second of these--the joint density of  $r_1$  and  $r_2$ --will be used in the companion paper mentioned in the Introduction.

The LIML variance estimator in the leading case  $M = 0$

The conditional density (29) may be thought of as conditional upon  $(x, X)$ , rather than  $W$ , and written in the form (on using (24) and (28)):

$$\begin{aligned} \text{pdf}(r_1, f_1 | x, X) &= C_2 \text{etr} \left\{ -\frac{1}{2} f_1 (I + r_1 r_1') X' X \right\} f_1^{m(n+1)/2-1} \\ &\quad \exp \left\{ -\frac{1}{2} f_1 (z' z - 2z' X r_1) \right\} (z' z)^{(m-n)/2} |X' \bar{P}_z X|^{(m+1)/2} \\ &\quad {}_1F_1 \left[ (n+3)/2, (m+n+2)/2; \frac{1}{2} f_1 (I + r_1 r_1') X' \bar{P}_z X \right], \end{aligned} \quad (53)$$

where we have put  $z = x - X r_1$ . Also, when  $M = 0$ ,

$$\text{pdf}(x, X) = (2\pi)^{-(n+1)K_2/2} \text{etr} \left\{ -\frac{1}{2} (x, X)' (x, X) \right\}. \quad (54)$$

Hence, multiplying (53) and (54) together, and transforming  $(r_1, x, X) \rightarrow (r_1, z, X)$ , we have

$$\begin{aligned} \text{pdf}(r_1, f_1, z, X) &= C_7 \text{etr} \left\{ -\frac{1}{2} (1 + f_1) (I + r_1 r_1') X' X \right\} f_1^{m(n+1)/2-1} \\ &\quad \exp \left\{ -\frac{1}{2} (1 + f_1) (z' z - 2z' X r_1) \right\} (z' z)^{(m-n)/2} |X' \bar{P}_z X|^{(m+1)/2} \\ &\quad {}_1F_1 \left[ (n+3)/2, (m+n+2)/2; \frac{1}{2} f_1 (I + r_1 r_1') X' \bar{P}_z X \right] \end{aligned} \quad (55)$$

where  $C_7 = [(2\pi)^{-(n+1)K_2/2} C_2]$ . Note that, from (14), the LIML variance estimator is based on  $s_1^2 = (1 + f_1) z' z$ . Now put  $z = v c$ ,

$v = z / (z' z)^{1/2}$ ,  $c^2 = z' z$ ,  $\prod_{i=1}^{K_2} dz_i = \frac{1}{2} (c^2)^{K_2/2-1} (dc^2) (v' dv)$ , and let  $H_1$  be such that  $(v, H_1) \in O(K_2)$ . Then define  $x_1$  and  $X_1$  by

$$\begin{bmatrix} x'_1 \\ X_1 \end{bmatrix} = (v, H_1)'X .$$

Making these transformations in (55) we have

$$\begin{aligned} \text{pdf}(r_1, f_1, v, c^2, x_1, X_1) &= \frac{1}{2}C_7 \exp\left\{-\frac{1}{2}(1+f_1)c^2\right\} f_1^{m(n+1)/2-1} \\ &\quad (c^2)^{(m+K_2-n)/2-1} |X'_1 X_1|^{(m+1)/2} \text{etr}\left\{-\frac{1}{2}(1+f_1)(I+r_1 r'_1)X'_1 X_1\right\} \\ &\quad \exp\left\{-\frac{1}{2}(1+f_1)(x'_1(I+r_1 r'_1)x_1 + 2cx'_1 r_1)\right\} \\ &\quad {}_1F_1\left[(n+3)/2, (m+n+2)/2; \frac{1}{2}f_1(I+r_1 r'_1)X'_1 X_1\right] . \end{aligned} \quad (56)$$

Note that this does not depend on  $v$ . It is straightforward to integrate out  $v$ ,  $x_1$ , and  $X_1$  in (56), leaving

$$\begin{aligned} \text{pdf}(r_1, f_1, c^2) &= C_8 \exp\left\{-\frac{1}{2}(1+f_1)c^2/(1+r_1 r'_1)\right\} f_1^{m(n+1)/2-1} \\ &\quad (1+f_1)^{-n(m+K_2+1)/2} (c^2)^{(m+K_2-n)/2-1} (1+r_1 r'_1)^{-(m+K_2+1)/2} \\ &\quad {}_2F_1\left[(m+K_2)/2, (n+3)/2, (m+n+2)/2; f_1 I_n/(1+f_1)\right] \end{aligned} \quad (57)$$

$$\text{where } C_8 = \left[ 2^{n(m+K_2+1)/2} \Gamma_n((m+K_2)/2) C_2 / \Gamma(K_2/2) \Gamma_n((K_2-1)/2) 2^{(n+1)K_2/2} \right] .$$

Now put  $s_1^2 = (1+f_1)c^2$  (Jacobian  $(1+f_1)^{-1}$ ) and notice that the terms in  $f_1$  are exactly those in  $\text{pdf}(f_1)$  in (36) apart from the constant  $C_4$ . Hence, since (36) integrates to unity we have, upon integrating out  $f_1$ ,

$$\text{pdf}(r_1, s_1^2) = C_9 \exp\left\{-\frac{1}{2}s_1^2/(1+r_1'r_1)\right\} (s_1^2)^{(m+K_2-n)/2-1} (1+r_1'r_1)^{-(m+K_2+1)/2} \quad (58)$$

$$\text{with } C_9 = [\Gamma((n+1)/2)/\pi^{(n+1)/2} 2^{(m+K_2-n)/2} \Gamma((m+K_2-n)/2)] .$$

Note that (58) has the form

$$\text{pdf}(s_1^2, r_1) = \text{pdf}(s_1^2 | r_1) \text{pdf}(r_1) ,$$

with  $s_1^2/(1+r_1'r_1)$  conditionally  $\chi^2(m+K_2-n)$  and  $r_1$  Cauchy. Integrating out  $r_1$  in (58) leaves

$$\text{pdf}(s_1^2) = C'_9 \exp\left\{-\frac{1}{2}s_1^2\right\} (s_1^2)^{(m+K_2-n)/2-1} {}_1F_1\left[(n/2, (m+K_2+1)/2, \frac{1}{2}s_1^2\right] \quad (59)$$

with

$$C'_9 = [\Gamma((n+1)/2)\Gamma((m+K_2-n+1)/2)/\pi^{1/2}\Gamma((m+K_2+1)/2)\Gamma((m+K_2-n)/2) 2^{(m+K_2-n)/2}] .$$

It is straightforward to confirm that (59) integrates to unity, and that none of the moments of  $s_1^2$  exist. This last result carries over, of course, to the general case which can be dealt with in exactly the manner used above but is considerably more complicated.

#### Joint density of $r_1$ and $r_2$ : $n = 1$

Using the first line of (24), and the first line of (26), (29) may be expressed in terms of  $s^2 = x'\bar{P}_X x$ ,  $r_2 = (X'X)^{-1}X'x$ , and  $t^2 = X'X$ , all of which are scalars in the case  $n = 1$ , and we may think of the conditional density (29) as conditional upon  $s^2$ ,  $r_2$ , and  $t^2$ . Also, in the distribution of  $W$  (equation (33)) we may transform  $W \rightarrow (s^2, r_2, t^2)$  (the

Jacobian is  $t^2$  ) and hence, on multiplying the two together, obtain

$$\begin{aligned} \text{pdf}(r_1, r_2, s^2, t^2, f_1) = C_2 C_3 \exp\left\{-\frac{1}{2}(1+f_1)[s^2 + t^2(1+r_2^2)]\right\} \\ f_1^{m-1} (s^2)^{(m+K_2)/2-1} (t^2)^{(m+K_2)/2} [s^2 + t^2(r_1 - r_2)^2]^{-1} \\ {}_1F_1\left[2, (m+3)/2; \frac{1}{2}f_1 t^2 s^2 (1+r_1^2)/[s^2 + t^2(r_1 - r_2)^2]\right] \\ {}_0F_1\left[K_2/2; \frac{1}{2}\mu^2 [s^2 \beta^2 + t^2(1+r_2 \beta)^2]\right], \end{aligned} \quad (60)$$

where we have put  $\mu^2 = \frac{1}{2}M'M$ , also a scalar when  $n = 1$ . To obtain  $\text{pdf}(r_1, r_2)$  we merely have to integrate over  $s^2 > 0$ ,  $t^2 > 0$ ,  $f_1 > 0$ , and this is quite straightforward. Setting  $a^2 = s^2 + t^2(1+r_2^2)$  and  $b = t^2(1+r_2^2)/(s^2 + t^2(1+r_2^2))$  (Jacobian  $a^2/(1+r_2^2)$ ) and integrating over  $f_1 > 0$ ,  $a^2 > 0$  we obtain

$$\begin{aligned} \text{pdf}(r_1, r_2, b) = 2^{m+K_2} \Gamma(m) \Gamma(K_2) C_2 C_3 (1+r_2^2)^{-(m+K_2+2)/2} (1+r_1^2)^{-1} \\ b^{(m+K_2+2)/2-1} (1-b)^{(m+K_2)/2-1} (1-b\delta_1 - (1-b)\delta_2)^{-1} \\ {}_2F_1\left[m, 2, (m+3)/2; b(1-b)/(1+r_2^2)(1-b\delta_1 - (1-b)\delta_2)\right] \\ {}_1F_1\left[K_2, K_2/2; \mu^2(1+\beta^2)[b\Delta_1 + (1-b)\Delta_2]\right] \end{aligned} \quad (61)$$

where

$$\left. \begin{aligned} \delta_1 &= (1 + r_1 r_2)^2 / (1 + r_1^2)(1 + r_2^2) \\ \delta_2 &= r_1^2 / (1 + r_1^2) \\ \Delta_1 &= (1 + r_2 \beta)^2 / (1 + \beta^2)(1 + r_2^2) \\ \Delta_2 &= \beta^2 / (1 + \beta^2) \end{aligned} \right\} \quad (62)$$

Expanding the term  $[b\Delta_1 + (1-b)\Delta_2]$  in the confluent hypergeometric function binomially, and the term  $(1 - b\delta_1 - (1-b)\delta_2)^{-(\ell+1)}$  that occurs in the series expansion of the  ${}_2F_1$  function in (61) binomially (twice), we obtain

$$\begin{aligned} \text{pdf}(r_1, r_2) &= C_{10} (1 + r_1^2)^{-1} (1 + r_2^2)^{-(m+K_2+2)/2} \\ &\sum_{j,k,\ell,s=0}^{\infty} \frac{(K_2)_{j+k} (m)_{\ell} (2)_{\ell} (1)_{s+\ell} ((m+K_2+2)/2)_{j+\ell+s} ((m+K_2)/2)_{k+\ell}}{j!k!\ell!s!(K_2/2)_{j+k} ((m+3)/2)_{\ell} (1)_{\ell} (m+K_2+1)_{j+k+s+2\ell} (1 + r_2^2)^{\ell}} \\ &[\mu^2(1+\beta^2)]^{j+k} \Delta_1^j \Delta_2^k \delta_1^s {}_2F_1\left[s+\ell+1, k+\ell+(m+K_2)/2, j+k+s+2\ell+m+K_2+1; \delta_2\right] \quad (62) \end{aligned}$$

$$\text{with } C_{10} = \left[ 2^{m+K_2} {}_2F_1(m) \Gamma(K_2) \Gamma((m+K_2)/2) \Gamma((m+K_2+2)/2) C_2 C_3 / \Gamma(m+K_2+1) \right].$$

In this form the joint density does not seem specially helpful but, as we shall see in the companion paper mentioned earlier, the result (63) can be re-interpreted in a way that is quite illuminating. Of course, various other joint densities can be derived from expressions like (60), but these we defer to further work.

## REFERENCES

- [1] Constantine, A. G. (1963). "Some noncentral distribution problems in multivariate analysis," Annals of Mathematical Statistics, 34, 1270-1285.
- [2] Constantine, A. G., and Muirhead, R. J. (1976). "Asymptotic expansions for distributions of latent roots in multivariate analysis," Journal of Multivariate Analysis, 6, 369-391.
- [3] Davis, A. W. "Invariant polynomials with two matrix arguments extending the zonal polynomials: Applications to multivariate distribution theory," Annals of Institute of Statistical Mathematics, 31, Part A, 465-485.
- [4] Herz, C. S. (1955). "Bessel functions of matrix argument," Annals of Mathematics, 61, 474-523.
- [5] Hillier, G. H. (1985). "On the joint and marginal densities of instrumental variable estimators in a general structural equation," Econometric Theory, 1, 53-72.
- [6] Hillier, G. H., T. W. Kinal, and V. K. Srivastava (1984). "On the moments of ordinary least squares and instrumental variable estimators in a general structural equation," Econometrica, 52, 185-202.
- [7] James, A. T. (1954). "Normal multivariate analysis and the orthogonal group," Annals of Mathematical Statistics, 25, 40-75.
- [8] James, A. T. (1964). "Distribution of matrix variates and latent roots derived from normal samples," Annals of Mathematical Statistics, 35, 475-501.
- [9] Muirhead, R. J. (1982). Aspects of Multivariate Statistical Theory. New York: Wiley.
- [10] Phillips, P. C. B. (1980). "The exact finite sample density of instrumental variable estimators in an equation with  $n+1$  endogenous variables," Econometrica, 48, 861-878.
- [11] Phillips, P. C. B. (1983). "Exact small sample theory in the simultaneous equations model." In M. D. Intriligator and Z. Griliches, eds., Handbook of Econometrica. Amsterdam: North Holland.
- [12] Phillips, P. C. B. (1984). "The exact distribution of LIML: I," International Economic Review, 25, 249-261.
- [13] Phillips, P. C. B. (1984). "The exact distribution of exogenous variable coefficient estimators," Journal of Econometrics, 26, 387-398.

- [14] Phillips, P. C. B. (1985). "The exact distribution of LIML: II," International Economic Review, 26, 21-36.
- [15] Rhodes, G. F. (1981). "Exact density functions and approximate critical regions for likelihood ratio identifiability test statistics," Econometrica, 49, 1214-1226.