CHARACTERIZATION OF INVARIANT TESTS IN THE STRUCTURAL EQUATION MODEL

Randolph G. K. Tan

NANYANG BUSINESS SCHOOL

NANYANG TECHNOLOGICAL UNIVERSITY

SINGAPORE 639798

Abstract

The invariance principle is applied to the problem of testing the exclusion of endogenous variables from a general structural equation model. The main statistic was formulated by Hillier (1987). Here, we examine its invariant nature and derive its exact density function with the aid of the invariant polynomials of Chikuse and Davis (1986) in the case of an arbitrary number of restrictions. The leading case where one endogenous variable is tested corresponds to a traditional test of significance in classical regression theory and is discussed in detail. In a model with two endogenous variables, we show that probability values can be obtained without difficulty.

1. INTRODUCTION

Tremendous progress has been made in the statistical analysis of multivariate linear models (MLM) in recent years, evidenced by the cornucopia of literature on the subject. The more prominent texts dealing in this area include Anderson (1984), Rao (1973), Giri (1977), Muirhead (1982), and Eaton (1983). Some of these results can now be brought to bear on the study of structural equation models which play an important role in econometric modeling. While the difficulties in econometric analysis are peculiar to the field, we hope to point out in this paper an area of small-sample theory which has benefited from clearer insights into multivariate analysis.

Econometrics, as distinct from pure statistical analysis, engenders the added sophistication of structure. In his survey on simultaneous equation models, Hausman (1983) clearly regarded this to be a unique attribute which distinguished econometric analysis as a contributory specialization in its own right rather than a mere appendage to mainstream mathematical statistics. Structural forms, as opposed to reduced forms, imply a certain degree of infusion of understanding of the interactive nature of different economic agents in a system. In Lucas' (1976) terms, a structural description of an economic system refines the specification of the motion of that system by differentiating between the actions of private agents and that which he terms "nature". Accepting the inseparability of these two components of an economic system - while at the same time maintaining recognition of their distinctive influences - is a step towards more effective assessment of macroeconomic policy-making. This was Lucas' (1976) celebrated critique which challenged a future generation of researchers to master the complexities of comprehensive macro-modeling. It virtually ruled out the applicability of analysing the actions of different economic agents separately and then juxtaposing the separate piecewise analyses to offer an overall picture. In this sense, while the degree of interactiveness among economic agents offered by the MLM

$$Y = Z\Pi + V \tag{1}$$

may be adequate, the lack of simultaneity in actions would be deemed too simplistic. One possible solution is to assume that some of the regressors should also be simultaneously determined by the model even as it acts to determine the outcome of some dependent variable of interest.

The general structural equation model consists of the structural equation (written without an explicit normalization rule)

$$Y\beta = Z_1\gamma + u \tag{2}$$

where the underlying reduced form given for Y is (1). Note that $Z = (Z_1, Z_2)$, $\Pi = (\Pi_1', \Pi_2')'$. Further, Y is $T \times (n+1)$, Z_1 is $T \times K_1$, Z_2 is $T \times K_2$ and $K_1 + K_2 = K$. For the disturbance terms, u is $T \times 1$ and V is $T \times (n+1)$. Finally, β and γ are column vectors of length $(n+1)$ and K_1 respectively.

In aspiring towards a coherent analytical framework for such models, reliable tools for estimation and testing are required. This paper was motivated by the second of these. At the outset, it should be said that any additional insights offered by small-sample analysis are not always commensurate with the difficulties encountered. On the other hand, the seemingly comfortable tractability of asymptotic analysis may not live up to its promise either: in many commonly occurring situations, the asymptotic approximations do not ensure acceptable small-sample performance. One such instance occurs with tests of estimated coefficients in the general (un-normalized) structural equation model.

In addressing a special case of this testing problem, Anderson and Rubin (1949) is in fact credited with developing the small-sample approach to testing in the structural equation model. In that case, all the endogenous variables except one did not appear in the structural equation under the null. That paper provided an exact *F*-test (the Anderson-Rubin (AR) test; see also Phillips (1983)) for the estimated structural coefficients. Whereas Revanker and Mallela (1972) is intent on deriving the power function of the test, Hillier (1987) points out that the intrinsic formulation of the AR test constrains its ability to distinguish the alternative from the null.

Hillier (1987) also broaches the topic of testing linear restrictions on the coefficients of (2) based on an invariance principle. It suggests test statistics for three situations of interest, namely tests on a subset of γ, a subset of β and both subsets of β and γ respectively.

In this paper, we examine the second of these cases. After setting up the general testing problem in Section 3, we introduce in Section 4 the interpretation that a statistic proposed by Hillier (1987) for the second case is just the determinant of a maximal invariant for the problem at hand. This interpretation of maximal invariance comes from restricting the transformation group which leaves the problem invariant. With this result, we proceed in Section 5 to characterize a class of invariant tests by establishing the exact small-sample density function for this maximal invariant. In Sections 7 through 9, we consider the particular case where only one element of β is under test. There, Hillier's (1987) proposed statistic exactly equals the maximal invariant for the restricted group of transformations. We end by considering the simplest formulation for which the analysis is applicable. It should be mentioned that a maximal invariant in the conventional sense exists for the problem of testing linear restrictions on b, though details will have to be deferred to a separate paper.

Our approach here is to demonstrate the accessibility of the small-sample distributional results for the problem at hand.

Whilst small-sample analysis in econometrics is still very much a frontier science, remarkable inroads have been achieved. Hence, as far as estimation of single structural equation models are concerned, many of the classical problems have effectively been solved by the establishment of the functional form of the exact finite-sample densities of the OLS, 2SLS and instrumental variable estimators as well as those of their moments in Phillips (1980, 1983, 1984), Hillier, Kinal and Srivastava (1984), Hillier (1985) and Hillier and Skeels (1993). Hillier (1990) provided a unification of the properties of these estimators.

As far as standard tests of coefficient significance are concerned however, the issues are still relatively open to further examination. Basmann (1960, 1963a, b) considered the exact distributions of test statistics for identifying restrictions in a system of simultaneous equations estimated by 2SLS. McDonald (1972) derived the exact density for the likelihood ratio test statistic - based on limited information maximum likelihood - for identifiability for an equation containing exactly two endogenous variables while Rhodes (1981) generalised the results to the case of an arbitrary number of exogenous variables.

In deriving the small-sample properties of what is essentially a new test statistic, we exploit the methods developed in these studies. Some of the specific tools are introduced in the next section.

2. MATHEMATICAL PRELIMINARIES

In our derivations, we shall use several proven devices to facilitate the systematic handling of multivariate integrals. To obviate the tedium of excessive repetition later on, these are set out here in detail.

To begin, certain notational conventions are employed. We let $O(n)$ denote the group of $n \times n$ orthogonal matrices; $GL(n)$, the general linear group of degree n over the field of real numbers; $V_{m,n}$, the Stiefel manifold of $n \times m$ semi-orthogonal matrices when $m \leq n$; and $M_{n,m}$, the space of $n \times m$ real matrices. Further, S_n denotes the group of $n \times n$ positive-definite symmetric matrices. Invariant measures on probability spaces should be interpreted as exterior products of differentials. In particular, $(\mathrm{d}H)$ is the normalized invariant measure on $O(n)$. For any matrix F with full column rank, we write the projection matrix as $\overline{P}_F = I - F(F'F)^{-1}F'$.

Since elements of the Stiefel manifold $V_{m,n} = \left(V(n \times m){:}V'V = I_m\right)$ may be regarded as consisting of the first n columns of an orthogonal matrix $H \in O(\mathrm{n})$, integration over $V_{m,n}$ may be interpreted as integration over $O(\mathrm{n})$, so that the normalized volume measures effectively transform with unit Jacobian:

$$(\mathrm{d}H)=(\mathrm{d}V).$$

Whenever the normalized measure, denoted by $(V'\mathrm{d}V)$, is displayed, we can convert via

$$(V'\mathrm{d}V) = \frac{2^m \pi^{mn/2}}{\Gamma_m(n/2)}(\mathrm{d}V) \tag{3}$$

before applying the above device.

Often, an integral of the form

$$g(A,L)=\int_{V_{m,n}} etr(V'AV+V'L)(\mathrm{d}V) \tag{4}$$

will have to be evaluated. By noting that (dV) is the normalized Haar (invariant) measure, this integral can be seen to be invariant under the transformation

$$L \to LH\,, \qquad H\in \mathcal{O}(\mathrm{n}).$$

Then, as Hillier (1985) demonstrates, by averaging over the orthogonal group $\mathcal{O}$(n), we obtain:

$$\begin{aligned} g(A,L) &= \int_{\mathcal{O}(m)} g(A,L)(\mathrm{d}H) \\ &= \int_{\mathcal{O}(m)} g(A,LH)(\mathrm{d}H) \\ &= \int_{V_{m,n}} etr(V'AV)\,{}_0F_1\left(\frac{m}{2};\frac{1}{4}V'LL'V\right)(\mathrm{d}V) \end{aligned} \tag{5}$$

The complete resolution of our integrals almost always entails, as final step, integration over the Stiefel manifold of terms in VV' and $(I-VV')$. These are re-expressed using $V=HE$, $H\in\mathcal{O}$(n) where $E=[I_m \vdots \mathbf{0}_{m,n-m}]'$ is appropriately defined for selecting the first m columns of H to be in V. Thus

$$VV'=H\begin{bmatrix} I_m & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{n-m}\end{bmatrix}H'\,,\ I-VV'=H\begin{bmatrix} \mathbf{0}_m & \mathbf{0} \\ \mathbf{0} & I_{n-m}\end{bmatrix}H'\,.$$

Finally, we note that Chikuse and Davis (1986) have generalized the fundamental result for averaging a product of two zonal polynomials over the orthogonal group to the case of a product of r zonal polynomials, where r≥1. The result is a sum of products of pairs of zonal polynomials:

$$\begin{aligned} &\int_{\mathcal{O}(n)} C_{\alpha_1}(A_1H'B_1H)\cdot C_{\alpha_2}(A_2H'B_2H)\cdot\ldots\cdot C_{\alpha_r}(A_rH'B_rH)(\mathrm{d}H) \\ &= \sum_{\phi\in\alpha_1\cdot\ldots\cdot\alpha_r} C_{\phi}^{\alpha_1,\ldots,\alpha_r}(A_1,A_2,\ldots,A_r)\cdot C_{\phi}^{\alpha_1,\ldots,\alpha_r}(A_1,A_2,\ldots,A_r)\,/\,C_{\phi}(I) \end{aligned} \tag{6}$$

where ϕ is a partition of $f=j_1+j_2+\ldots+j_r$ into not more than n parts and "$\phi\in\alpha_1\cdot\alpha_2\cdot\ldots\cdot\alpha_r$" indicates that the summation on ϕ extends over all those partitions of ϕ for which the irreducible representation of the general linear group $GL(n)$ indexed by 2ϕ occurs in the decomposition of the Kronecker product $2\alpha_1 \otimes 2\alpha_2 \otimes \ldots \otimes 2\alpha_r$. In the specific case of an invariant polynomial with three arguments $C_{\phi}^{\alpha_1,\alpha_2,\alpha_3}(A_1, A_2, A_3)$, it is to be understood that α_1, α_2 and α_3 are lexicographically ordered partitions of j_1, j_2 and j_3 and $\phi \in \alpha_1 \cdot \alpha_2 \cdot \alpha_3$ (Chikuse and Davis(1986)).

3. SPECIFICATION OF THE TESTING PROBLEM

Let $\beta = (\beta_1', \beta_2')'$ be a partition into column vectors of $n_1 + 1$ and n_2 components. We are concerned with testing the hypothesis

$$H_0 : \beta_2 = 0 \tag{7a}$$

against the two-sided alternative

$$H_1 : \beta_2 \neq 0. \tag{7b}$$

The use of a MLM in defining the reduced form is also a deliberate attempt to focus attention on the fact that the single structural equation model is just a restricted form of the former. The restrictions, in this case, arise from the requirement that (1) and (2) should be compatible. The result is a set of rank restrictions on Π, by virtue of which we may interpret (1) and (2) together as a curved exponential model (see, for example Hosoya *et al* (1989) and van Garderen (1994)).

The basic distributional assumption for the reduced form error matrix is that

$$V \sim N(0, I_T \otimes \Omega). \tag{8}$$

Provided Z_1 and Z_2 have full column rank, (1) and (2) together imply

$$\Pi_1\beta + \gamma = 0, \tag{9}$$

and $\Pi_2\beta = 0$. (10)

Further, $u = V\beta$ so that $\mathrm{var}(u_t) = \beta'\Omega\beta$, t=1,2,...,T. From (10), we obtain a necessary and sufficient condition for the existence of a structural form unique up to normalization to be

$$rank(\Pi_2) = n. \tag{11}$$

Formally, the distribution of Y implied by (4) together with the restrictions on the parameter space implied by (5) and (6) provide the basis for inference for the parameters $(\Pi, \Omega, \beta, \gamma)$. In order to focus on the coefficients in (7a,b), we define the partitions of the reduced form parameters

$$\Pi = \begin{array}{cc} \begin{bmatrix} \Pi_{11} \\ \Pi_{12} \\ \Pi_2 \end{bmatrix} & \begin{array}{c} k_{11} \\ k_{12} \\ K_2 \end{array} \\ n+1 & \end{array} \qquad \Pi_{22} = \begin{array}{c} \begin{bmatrix} \Pi_{21} & \Pi_{22} \end{bmatrix} \\ \begin{array}{cc} n_1+1 & n_2 \end{array} \end{array} \qquad \Omega = \begin{array}{cc} \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} & \begin{array}{c} n_1+1 \\ n_2 \end{array} \\ \begin{array}{cc} n_1+1 & n_2 \end{array} & \end{array} \tag{12}$$

with $k_{11} + k_{12} = K_1$, $\Omega_{12} = \Omega'_{21}$ and $\Omega_{22\cdot1} = \Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12}$. Then, Hillier (1987) has shown that the problem in (7a,b) is equivalent to testing

$$H_0 : rank(\Pi_{21}) = n_1 \qquad (rank(\Pi_2) = n) \tag{13a}$$

against

$$H_1 : rank(\Pi_{21}) = n_1 + 1 \qquad (rank(\Pi_2) = n) \tag{13b}$$

The problem can be simplified without any loss of generality by transforming the statistics to canonical form. The application of these transformations would essentially be equivalent to using a canonical model as in Phillips (1983) and Hillier (1990). The maximum likelihood estimates (MLE) $\hat{\Pi} = (Z'Z)^{-1}Z'Y$ and $\hat{W} = Y'\overline{P}_Z Y$ are jointly sufficient for (Π, Ω), and are independently distributed as $N(\Pi, (Z'Z)^{-1} \otimes \Omega)$ and $W_{n+1}(\nu, \Omega)$ respectively (Muirhead (1982), Theorem 10.1.2) where $\nu = T - K$. Let $C_Z = (c_{ij})$, with $c_{11} = (Z_1'Z_1)^{-1}$, $c_{12} = c_{11}^{-1}Z_1'Z_2$, $c_{21} = 0$ and $c_{22} = (Z_2'\overline{P}_{Z_1}Z_2)^{1/2}$. Then, we define the transformed sufficient statistics $X = C_Z\hat{\Pi}\Omega^{-1/2}$

and $W = \Omega^{-1/2}\hat{W}\Omega^{-1/2}$, the transformations being equivalent to restricting both endogenous and exogenous variable matrices to have orthogonal columns. With the reduced form coefficient matrix transformed to $M = C_Z\Pi\Omega^{-1/2}$, hypotheses about Π such as H_0 and H_1 can restated in terms of M. Test statistics can be constructed from (X,W).

4. INVARIANCE PROPERTIES

Hillier (1987) shows that the testing problem in (13a,b) is invariant under the group of transformations

$$\begin{aligned} \mathsf{G} = \{(H, Q, M_{12}^{*}): \quad & H \in \mathsf{O}(K_2),\, M_{12}^{*} \in \mathsf{M}_{k_{12},n+1}, Q = \begin{pmatrix} Q_{11} & Q_{12} \\ 0 & Q_{22} \end{pmatrix}, \\ & Q_{11} \in \mathsf{GL}(n_1+1), Q_{12} \in \mathsf{M}_{n_1+1,n_2}, Q_{22} \in \mathsf{GL}(n_2)\} \end{aligned}$$

acting on the space of the statistic $(X_{12}, X_{21}, X_{22}, W)$ by

$$\mathsf{G}:(X_{12}, X_{21}, X_{22}, W) \rightarrow \left(X_{12}Q + M_{12}^{*}, HX_{21}Q_{11}, H[X_{21}Q_{12} + X_{22}Q_{22}], Q'WQ\right).$$

In light of this, the following theorem gives a characterization of the class of invariant tests in a weaker sense.

THEOREM 1. The statistic $\Phi = \Psi^{1/2}W_{22\cdot 1}^{-1/2}\Psi^{1/2}$, where $\Psi = X_{22}'\overline{P}_{X_{21}}X_{22}$, is a maximal invariant for the problem of testing H_0 against H_1 under a group of transformations G_1 where $\mathsf{G} = \mathsf{G}_1 \times \mathsf{G}_2$.

As a consequence, the following result obtains.

Corollary 1. The characteristic roots of Φ constitute a maximal invariant for the problem of testing H_0 against H_1 under the group G_1.

5. DENSITY OF THE MAXIMAL INVARIANT

It can be shown that the conditional distribution of Ψ given X_{21} is the non-central Wishart $W_{n_2}(K_2 - n_1 - 1, I_{n_2}, \Delta_{X_{21}})$ where $\Delta_{X_{21}} = M_{22}' \bar{P}_{X_{21}} M_{22}$. Further, $W_{22\cdot 1}$ has the (unconditional) Wishart distribution $W_{n_2}(\nu - n - 1, I_{n_2})$. The conditional distribution of Φ given X_{21} is a matricvariate non-central F. In the special case n_2=1, this will reduce to the univariate non-central F density, or equivalently, the distribution of a ratio of two independent χ^2 variates. Further exploration of this special case is carried out in Section 7.

In the general case, an expression for the conditional density of Φ is

$$\begin{aligned} pdf(\Phi|X_{21}) &= k_0 etr\left(-\frac{1}{2}\Delta_{X_{21}}\right)\det\Phi^{\nu_1-(n_2+1)/2} \\ &\quad \times \det(I+\Phi)^{-\nu_1-\nu_2}\,{}_1F_1\left(\nu_1+\nu_2;\nu_1;\frac{1}{2}\Delta_{X_{21}}\Phi(I+\Phi)^{-1}\right) \end{aligned} \tag{14}$$

where $k_0 = \dfrac{\Gamma_{n_2}(\nu_1+\nu_2)}{\Gamma_{n_2}(\nu_1)\Gamma_{n_2}(\nu_2)}$, and $\nu_2 = (\nu - n_1 - 1)/2$. The unconditional density is derived by averaging this with respect to the distribution of X_{21}. Since $X_{21} \sim N(M_{21}, I_{K_2} \otimes I_{n_2})$, we get

$$\begin{aligned} pdf(\Phi) &= \mu(\Phi) \int_{X_{21}\in M_{n_1+1,K_2}} {}_1F_1\left(\nu_1+\nu_2;\nu_1;\frac{1}{2}M_{22}'\bar{P}_{X_{21}}M_{22}\Phi(I+\Phi)^{-1}\right) \\ &\quad \times etr\left(-\frac{1}{2}M_{22}'\bar{P}_{X_{21}}M_{22}\right) etr\left(-\frac{1}{2}(X_{21}-M_{21})'(X_{21}-M_{21})\right)(\mathrm{d}X_{21}) \end{aligned} \tag{15}$$

where $\mu(\Phi) = \dfrac{\Gamma_{n_2}(\nu_1+\nu_2)\det\Phi^{\nu_1-(n_2+1)/2}\det(I+\Phi)^{-\nu_1-\nu_2}}{(2\pi)^{K_2(n_1+1)/2}\Gamma_{n_2}(\nu_1)\Gamma_{n_2}(\nu_2)}$.

As indicated, the integration is to be carried out over the range space of X_{21}. This range space is in fact the Lie group of $K_2 \times (n_1 + 1)$ matrices, and can be decomposed into factor spaces comprising the Stiefel manifold and symmetric space of positive matrices, in which the integral can be easily resolved. The factorization implies the transformation $X_{21} \rightarrow VR^{1/2}$ where the Jacobian is obtained from

$$(\mathrm{d}X_{21}) = \frac{\pi^{K_2(n_1+1)/2}}{\Gamma_{n_1+1}(K_2/2)} \det R^{(K_2-n_1-2)/2} (\mathrm{d}R)(\mathrm{d}V)\,. \tag{16}$$

The (unconditional) density function can therefore be explicitly written with the iterated integral

$$\begin{aligned} pdf(\Phi) = {} & \frac{\pi^{K_2(n_1+1)/2}}{\Gamma_{n_1+1}(K_2/2)} \mu(\Phi)\, etr\left(-\frac{1}{2} M'_{21} M_{21} \right) \\ & \times \int_{V_{n_1+1,K_2}} \int_{S_{n_1+1}} etr\left(-\frac{1}{2} M'_{22}(I - VV') M_{22} \right) etr\left(-\frac{1}{2} R + M'_{21} V R^{1/2} \right) \\ & \times \det R^{(K_2-n_1-2)/2} {}_1F_1\left(\nu_1 + \nu_2; \nu_1; \frac{1}{2} M'_{22}(I - VV') M_{22} \Phi (I+\Phi)^{-1} \right) (\mathrm{d}R)(\mathrm{d}V) \end{aligned} \tag{17}$$

To facilitate the evaluation of the integral over the Stiefel manifold, we use the fact that it is invariant under right orthogonal transformation of $M_{21}R^{1/2}$ with respect to the group $O(n_1+1)$. This implies that (17) is invariant under the transformation

$$etr\left(V' M_{21} R^{1/2} \right) \rightarrow {}_0F_1\left(\frac{n_1+1}{2}; \frac{1}{4} V' M_{21} R M'_{21} V \right).$$

The integral with respect to R becomes

$$\begin{aligned} & \int_{S_{n_1+1}} etr\left(-\frac{1}{2} R \right) \det R^{\nu_2 - 1/2} {}_0F_1\left(\frac{1}{2}(n_1+1); \frac{1}{4} V' M_{21} R M'_{21} V \right) (\mathrm{d}R) \\ & = 2^{K_2(n_1+1)/2} \Gamma_{n_1+1}(K_2/2) {}_1F_1\left(K_2/2; \frac{1}{2}(n_1+1); \frac{1}{2} M'_{21} V V' M_{21} \right) \end{aligned} \tag{18}$$

using Herz (1955). Inserting this in (17) produces the function

$$pdf\,(\Phi) = k_0 etr\left(-\tfrac{1}{2} M_2' M_2\right)\det\Phi^{\nu_1-(n_2+1)/2}\det(I+\Phi)^{-\nu_1-\nu_2}$$
$$\times\int_{V_{n_1+1,K_2}} etr\left(\tfrac{1}{2} M_{22}' VV' M_{22}\right){}_1F_1\left(\nu_1+\nu_2;\nu_1;\tfrac{1}{2} M_{22}'(I-VV')M_{22}\Phi(I+\Phi)^{-1}\right)$$
$$\times{}_1F_1\left(K_2/2;\frac{1}{2}(n_1+1);\frac{1}{2} M_{21}' VV' M_{21}\right)(\mathrm{d}V)$$

(19)

To simplify the evaluation, we have rewritten the invariant measure on the Stiefel manifold in normalized form (Muirhead (1982), pp.70-72). Since the columns of V can be regarded as the first (n_1+1) columns of a K_2-dimensional orthogonal matrix H, we may write $V = HE_1$ where $E_1 = (I_{n_1+1}\vdots\mathbf{0}_{n_1+1,K_2-n_1-1})'$ is the appropriate selection matrix. Equivalently, the integration over the Stiefel manifold in (19) can be performed over $O(K_2)$. The result is an infinite weighted sum of invariant polynomials whose arguments depend on the maximal invariant Φ and the test parameter matrix M:

$$\int_{O(K_2)} etr\left(\frac{1}{2} M_{22}' HE_1E_1'HM_{22}\right){}_1F_1\left(\nu_1+\nu_2;\nu_1;\frac{1}{2} M_{22}'H(I-E_1E_1')M_{22}\Phi(I+\Phi)^{-1}\right)$$
$$\times{}_1F_1\left(\frac{K_2}{2};\frac{n_1+1}{2};\frac{1}{2} M_{21}'HE_1E_1'H'M_{21}\right)(\mathrm{d}H)$$
$$=\sum_{\kappa_1,\kappa_2,\kappa_3;\phi}\frac{(K_2/2)_{\kappa_2}(\nu_1+\nu_2)_{\kappa_3}}{((n_1+1)/2)_{\kappa_2}(\nu_1)_{\kappa_3}\, j_1!\,j_2!\,j_3!}\int_{O(K_2)} C_{\kappa_1}\left(\frac{1}{2}E_1E_1'H'M_{22}M_{22}'H\right)$$
$$\times C_{\kappa_2}\left(\frac{1}{2}E_1E_1'H'M_{21}M_{21}'H\right)C_{\kappa_3}\left(\frac{1}{2}(I-E_1E_1')H'M_{22}\Phi(I+\Phi)^{-1}M_{22}'H\right)(\mathrm{d}H)$$
$$=\sum_{\kappa_1,\kappa_2,\kappa_3;\phi}\frac{(K_2/2)_{\kappa_2}(\nu_1+\nu_2)_{\kappa_3}(1/2)^f}{((n_1+1)/2)_{\kappa_2}(\nu_1)_{\kappa_3}\, j_1!\,j_2!\,j_3!\,C_\phi(I)}C_\phi^{\kappa_1,\kappa_2,\kappa_3}(E_1E_1',E_1E_1',I-E_1E_1')$$
$$\times C_\phi^{\kappa_1,\kappa_2,\kappa_3}(M_{22}M_{22}',M_{21}M_{21}',M_{22}\Phi(I+\Phi)^{-1}M_{22}')$$

(20)

The first equality gives the result after expanding the exponential and confluent hypergeometric functions in the expression in terms of zonal polynomials. We use $\sum_{\kappa_1,\kappa_2,\kappa_3;\phi}$ to denote the multiple summation $\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty}\sum_{j_3=0}^{\infty}\sum_{\kappa_1}\sum_{\kappa_2}\sum_{\kappa_3}\sum_{\phi\in\kappa_1\cdot\kappa_2\cdot\kappa_3}$.

Here, a_1, a_2, a_3 and f are ordered partitions of f_1, f_2, f_3 and $(f_1+f_2+f_3)$ into K_2 parts respectively. The summation on f extends over all those partitions of $(f_1+f_2+f_3)$ for which the irreducible representation of the general linear group $GL(K_2)$ indexed by $2\mathrm{f}$ occurs in the decomposition of $2\mathrm{a}_1 \otimes 2\mathrm{a}_2 \otimes 2\mathrm{a}_3$. The factor

$$\mathrm{k}_{\mathrm{f}}^{\mathrm{a}_1,\mathrm{a}_2,\mathrm{a}_3} = C_{\mathrm{f}}^{\mathrm{a}_1,\mathrm{a}_2,\mathrm{a}_3}(E_1E_1', E_1E_1', I - E_1E_1')$$

whose arguments comprise only linear structures may be evaluated in the manner of Hillier (1986, p.60). Using Chikuse and Davis (1986, (3.6)), we can show that it equals

$$\sum_{\mathrm{s}^* \in \mathrm{a}_1 \cdot \mathrm{a}_2} \sum_{\mathrm{f}' \equiv \mathrm{f}} \mathrm{b}_{\mathrm{s}^*;\mathrm{f}'}^{\mathrm{a}_1,\mathrm{a}_2,\mathrm{a}_3;\mathrm{f}} \left(\frac{n_1+1}{2}\right)_{\mathrm{s}^*} (\mathrm{n}_1)_{\mathrm{a}_3} \mathrm{q}_{\mathrm{f}'}^{\mathrm{s}^*,\mathrm{a}_3} C_{\mathrm{f}'}(I_{K_2}) \big/ (K_2/2)_{\mathrm{f}'}$$

Details are contained in Appendix B, together with an explanation of the coefficients. Gathering these results, the complete expression for the exact density of the maximal invariant Φ is

$$\begin{aligned} pdf(\Phi) &= k_0 etr\left(-\frac{1}{2} M_2' M_2\right) \det \Phi^{\mathrm{n}_1-(N_2+1)/2} \det(I+\Phi)^{-\mathrm{n}_1-\mathrm{n}_2} \\ &\times \sum_{\mathrm{a}_1,\mathrm{a}_2,\mathrm{a}_3;\mathrm{f}} \frac{(K_2/2)_{\mathrm{a}_2} (\mathrm{n}_1+\mathrm{n}_2)_{\mathrm{a}_3} (1/2)^f}{((n_1+1)/2)_{\mathrm{a}_2} (\mathrm{n}_1)_{\mathrm{a}_3} j_1!\, j_2!\, j_3!\, C_{\mathrm{f}}(I)} \mathrm{k}_{\mathrm{f}}^{\mathrm{a}_1,\mathrm{a}_2,\mathrm{a}_3} \\ &\times C_{\mathrm{f}}^{\mathrm{a}_1,\mathrm{a}_2,\mathrm{a}_3}(M_{22}M_{22}', M_{21}M_{21}', M_{22}\Phi(I+\Phi)^{-1}M_{22}') \end{aligned} \qquad (21)$$

Two points are worth noting about this result. In the context of the linear regression model, the invariance principle has usually succeeded in reducing the number of parameters entering into the problem. Beyond this, however, the notion of invariance is important in that it expresses the basic symmetry contained in the problem that the inference should inherit. As expected, in the case of rank testing, the nonlinear nature of the restriction means that the complexity of the problem depends not just on the number of restrictions but more crucially, on the number of parameters involved. Intuitively speaking, the relevance of the information contained in (21) is gauged with the invariance criterion: a maximal invariant preserves the integrity of "relevant" sample information while ignoring the "irrelevant" bits. Thus, we observe although all the original parameters enter into the density (21), they appear only as functions of $M'_{21}M_{21}$, $M'_{21}M_{22}$ and $M'_{22}M_{22}$. These, of course, are just the matrices which appear in the statement of the null hypothesis.

Secondly, it is clear that the alternative hypothesis is favoured by "large values" of Φ^{-1}. By this, we mean that the critical region consists of matrix variates J such that ($J-\Phi^{-1}$) is positive-definite. This will enable test procedures based on univariate test statistics to be prescribed.

6. DERIVATION OF A TEST STATISTIC

Though the invariant polynomials are rather difficult to compute in practice, they afford tremendous notational economy. Further, it should be obvious that they are essential for the resolution of the matricvariate integrals that accompany the case where an arbitrary number of restrictions under the null. Although the form of (21) makes it unsuitable for computational work at present, we may make certain observations. Since the distribution of Φ depends on *all of the nuisance parameters*, that is all the parameters of the reduced form, it implies that a uniformly most powerful invariant (UMPI) test cannot be obtained through the formal application of the Neyman-Pearson lemma. It would also suggest that test procedures developed from Φ will lack sensitivity. Nonetheless, the exact expression in (21) is useful for developing invariant tests in special cases. We could consider averaging power over a range of the values of the nuisance parameters in order to extract a test criterion. However, not all of these parameters are free to vary by virtue of the maintained hypothesis and the averaging has to be carried out with care so as to preserve the basis for distinguishing between the null and the alternative.

To demonstrate this, consider the power on some manifold in the space of nuisance parameters, so that the most powerful test on such a manifold should ostensibly represent an admissible test. For instance, matricvariate distributions such as the von Mises-Fisher distribution, the matrix Bingham distribution, or more generally the matrix angular central Gaussian distributions (Chikuse, 1990a, b) may be considered suitable candidates as priors for some submatrix of *M*. Under the alternative, this suggests that we may endow the manifold defined by $M_2^{*\prime} M_2^{*} = \mathrm{x}^2 I_n$ with suitable uniform measure, where M_2^{*} is a selection of *n* columns of M_2 and x is a scalar. This would be equivalent to treating M_2^{*} as an orientation for which the angles between the columns are all equal to zero. Clearly, M_2^{*} could be chosen using prior information about the normalization of the structural form. Under the null, M_{21} has less than full rank given the maintained hypothesis, implying then the manifolds

$$M_{21}^{*\prime} M_{21}^{*} = \mathrm{x}^2 I_{n_1}, \qquad M_{22}' M_{22} = \mathrm{x}^2 I_{n_2} \tag{22}$$

in parameter space on which power can be maximized.

As to the question of what function of Φ to use as the test statistic, one answer would be the matrix trace. The density of *T*=tr(Φ) can be obtained by straightforward integration. Constantine's (1966, Theorem 4) density for Hotelling's T_o^2 statistic applies to the trace of a statistic which has a matricvariate F distribution. This is the form of the conditional density of *T*, given X_{21}. The unconditional density of *T*, obtained by averaging with respect to the density of X_{21}, is given in the following theorem.

THEOREM 2. *The exact density of* T=tr(Φ), *where* $0 \le T < 1$, *is*

$$pdf\,(T) = k_0 etr\left(-\frac{1}{2}M_{21}'M_{21}\right)T^{n_2\nu_1}\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{j_1}\frac{(-T)^{j_1}}{(n_2\nu_1)_{j_1}\,j_1\,!}$$
$$\times\sum_{\alpha_1}\sum_{\alpha_2}\binom{\alpha_1}{\alpha_2}\frac{(\nu_1+\nu_2)_{\alpha_1}(\nu_1)_{\alpha_1}C_{\alpha_1}(I)}{(\nu_1)_{\alpha_2}C_{\alpha_2}(I)}$$
$$\times\sum_{j_3=0}^{\infty}\sum_{\alpha_3}\sum_{j_4=0}^{\infty}\sum_{\alpha_4}\sum_{\phi\in\alpha_3\cdot\alpha_2\cdot\alpha_4}\frac{(K_2/2)_{\alpha_4}(1/2)^{j_3+j_4}}{([n_1+1]/2)_{\alpha_4}\,j_3\,!\,j_4\,!C_\phi\,(I)}$$
$$\times C_\phi^{\alpha_3,\alpha_2,\alpha_4}(-M_{22}M_{22}',-M_{22}M_{22}',M_{21}M_{21}')\cdot C_\phi^{\alpha_3,\alpha_2,\alpha_4}(I-E_1E_1',I-E_1E_1',E_1E_1')$$

Note that the null is rejected for small values of T.

7. THE CASE OF n_2=1

The case where only one of the endogenous variables of the structural equation is being tested corresponds to n_1=1. This case holds particular interest because it's motivation is precisely the significance testing that underlies the usual *t* test of significance in the linear regression model. The maximal invariant reduces to a scalar quantity and its density function may be obtained from (21). Substituting n_2=1 and n_1+1=n into the formula for Φ and appending a multiplicative factor gives the statistic

$$\Phi_1 = \left(\frac{\nu-n}{K_2-n}\right)\frac{X_{22}'\overline{P}_{X_{21}}X_{22}}{W_{22\cdot 1}}\,. \tag{23}$$

The motivation for including the factor becomes clear from observing that the distribution of Φ_1 conditional on a given value of X_{21} is the noncentral univariate F with (K_2-n) and (ν-n) degrees of freedom, denoted $F'(K_2-n,\nu-n,\delta^2(X_{21}))$. The distinction between the null and the alternative is embodied in the noncentrality parameter

$$\delta^2(X_{21}) = d'\overline{P}_{X_{21}}d\,,$$

with the vector M_{22} denoted d now. Unlike Anderson and Rubin's (1949) F-test, information about whether the null is true is incorporated into the distribution of the test statistic through the noncentrality parameter.

8. UNCONDITIONAL DISTRIBUTION OF Φ_1

The noncentral F can be represented as a mixture of F distributions

$$\Pr(F'(\nu_1,\nu_2,\delta^2) \leq f_0) = \sum_{j=0}^{\infty} p_j \Pr\left(F(\nu_1 + 2j,\nu_2) \leq \frac{\nu_1 f_0}{\nu_1 + 2j} \right) \tag{24}$$

where the weights p_j are the Poisson probabilities

$$p_j = \frac{e^{-\delta^2/2}\delta^{2j}}{2^j j!}, \quad j = 0,1,2,\ldots \tag{25}$$

Applying this to the conditional result and taking expectations with respect to the density of X_{21} yields the unconditional distribution for Φ_1

$$\Pr(\Phi_1 \leq f_1) = \sum_{j=0}^{\infty} p_j^* \Pr\left(F(K_2 - n + 2j,\nu - n) \leq \frac{(K_2 - n) f_1}{K_2 - n + 2j} \right) \tag{26}$$

with $p_j^* = E\left(\frac{e^{-\delta^2(X_{21})/2}(\delta^2(X_{21}))^j}{2^j j!} \right), \quad j = 0,1,2,\ldots.$

In order to evaluate the weights of the mixture expression, we make use of the following result from Hillier (1987).

THEOREM 3. (Hillier (1987)) *The weight* p_j^* *is the coefficient of* t^j *in the expansion of* $E\left(\exp\left(-\frac{1}{2}(1-t)\delta^2(X_{21}) \right) \right)$ *and can be evaluated as*

$$\frac{1}{j!} \frac{\partial^j}{\partial t^j} \left[E\left(\exp\left(-\frac{1}{2}(1-t)\delta^2(X_{21}) \right) \right) \right]_{t=0} .$$

The proof follows from replacing the exponential function by its series expansion and evaluating the partial derivative at *t*=0. The required expectation is evaluated with respect to the distribution $X_{21} \sim N(M_{21}, I_{K_2} \otimes I_n)$.

Thus, if $B(t) = \exp\left(-\frac{1}{2}(1-t)\delta^2(X_{21})\right)$, we require the solution to

$$\begin{aligned} E(B(t)) &= \int_{\mathcal{M}_{K_2,n}} \exp\left(-\frac{1}{2}(1-t)\delta^2(X_{21})\right) \cdot pdf(X_{21})(dX_{21}) \\ &= c_0 \int\limits_{\mathcal{M}_{K_2,n}} \exp\left(-\frac{1}{2}(1-t)d'\overline{P}_{x_{21}}d\right) etr\left(-\frac{1}{2}[X_{21}-M_{21}]'[X_{21}-M_{21}]\right)(dX_{21}) \end{aligned} \tag{27}$$

where $c_0 = (2\pi)^{-K_2 n/2}$.

As we have done previously, the transformation of the integral over the group $\mathcal{M}_{K_2,n}$ to a double integral over the product space of the Stiefel manifold and the space of positive definite symmetric matrices is based on applying Fubini's theorem.

After rearranging terms, we obtain

$$\begin{aligned} E(B(t)) &= \frac{c_0}{\Gamma_n(K_2/2)} etr\left(-\frac{1}{2}M_{21}'M_{21}\right)\exp\left(-\frac{(1-t)d'd}{2}\right) \\ &\times \int_{\mathcal{S}_n} \det R^{(K_2-n-1)/2} etr\left(-\frac{1}{2}\Omega_{11}^{-1}R\right) \\ &\times \int_{\mathcal{V}_{n,K_2}} etr\left(\frac{1}{2}(1-t)V'dd'V\right) etr(V'M_{21}R^{1/2})(dV)(dR) \end{aligned} \tag{28}$$

where $V \in \mathcal{V}_{n,K_2}$ and $R \in \mathcal{S}_n$. The problem here is similar to that in (17). Hence, by putting $V = HE_2$, we find

$$\begin{aligned} E(B(t)) &= etr\left(-\frac{1}{2}M_{21}'M_{21}\right) etr\left(-\frac{[1-t]dd'}{2}\right) \\ &\times \int\limits_{\mathcal{O}(K_2)} etr\left(\frac{[1-t]}{2}E_2E_2'H'dd'H\right) {}_1F_1\left(\frac{K_2}{2};\frac{n}{2};\frac{1}{2}E_2E_2'H'M_{21}M_{21}'H\right)(\mathrm{d}H) \end{aligned} \tag{29}$$

where $H \in \mathcal{O}(K_2)$ and $E_2 = [I_n \vdots \mathbf{0}_{n,K_2-n}]'$ is a $K_2 \times n$ selection matrix. If α_1 and α_2

are lexicographically ordered partitions of j_1 and j_2 into K_2 parts respectively, then resolving the right-hand side of (29) in terms of invariant polynomials in two matrix arguments yields

$$etr\left(\frac{1}{2}M_{21}'M_{21}\right)\sum_{\alpha_1,[j_2];\phi}\frac{a_\phi^{\alpha_1,[j_2]}}{j_1!\,j_2!}C_\phi^{\alpha_1,[j_2]}(M_{21}M_{21}',dd')(t-1)^{j_2} \tag{30}$$

where $a_\phi^{\alpha_1,[j_2]}=\frac{(K_2/2)_{\alpha_1}([K_2-n]/2)_{j_2}}{(K_2/2)_\phi}\theta_\phi^{\alpha_1,[j_2]}\left(\frac{1}{2}\right)^{j_1+j_2}$. In the summation, the second partition $[j_2]$ denotes the special case of the top-order partition $[j_2, 0,..., 0]$; the reason being that the index j_2 in this summation corresponds to the argument dd' in the invariant polynomial which is being summed. Therefore, the fact that dd' is of unit rank means that only the top-order partition will yield non-vanishing terms.

The coefficient of $(t-1)^{j_2}$ which appears in (30) is

$$z_{j_2}=etr\left(-\frac{1}{2}M_{21}'M_{21}\right)\sum_{j_1=0}^{\infty}\sum_{\alpha_1}\sum_{\phi\in\alpha_1\cdot[j_2]}\frac{a_\phi^{\alpha_1,[j_2]}}{j_1!\,j_2}C_\phi^{\alpha_1,[j_2]}(M_{21}M_{21}',dd') \tag{31}$$

This depends on all of the reduced form parameters as well as the degrees of over-identification K_2-n. From Theorem 3 and (30), we have the unconditional weights

$$p_j^*=\frac{1}{j!}\sum_{k=0}^{\infty}(-k)_j(-1)^k z_k \tag{32}$$

for the infinite series expression of the distribution of the maximal invariant.

9. DISTINGUISHING BETWEEN THE NULL AND THE ALTERNATIVE FOR n_2=1

In the leading case n_2=1, the basis of distinguishability between the null and the alternative is contained in the parameters of the unconditional expression (28). While this is a weighted sum of (central) *F* distributions, the result is not a non-central *F*, in general, because the weights are no longer Poisson probabilities.

Now, when H_0 is true, $M_2\mathrm{b} = M_{21}\mathrm{b}_1 + M_{22}\mathrm{b}_2 = 0$ implies that $M_{21}\mathrm{b}_1 = 0$. In turn, this implies that $rank(M_{21}) = n_1 = n - 1$. Hence, $M_{21}'M_{21}$ - a square matrix of dimension *n* - must be singular. Further, $M_{21}M_{21}'$ (of dimension K_2) must also be singular. The presence of these singular matrices are manifested through the terms $etr\left(-\frac{1}{2}M_{21}'M_{21}\right)$ and $C_{\mathrm{f}}^{\mathrm{a}_1,[j_2]}(M_{21}M_{21}', dd')$. Specifically, ($K_2$-*n*+1) of the latent roots of $M_{21}M_{21}'$ will be zero so that the invariant polynomial will vanish for partitions of j_1 which possess more than (K_2-*n*+1) non-zero parts.

When H_0 is false, the maintained hypothesis requires that $rank(M_{21}) = n$. Consequently, $M_{21}'M_{21}$ must be nonsingular. But $M_{21}M_{21}'$ is still singular with (K_2-*n*) of its characteristic roots now equal to zero. Hence, for partitions of j_1 into more than (K_2-*n*) non-zero parts, the invariant polynomials will vanish.

10. CRITICAL REGIONS WHEN *n*=1

In the case *n*=1, the null model is a classical linear regression and the problem of deriving a critical region based on Φ_1 is considerably simplified, mainly because the invariant polynomials reduce to elementary functions with scalar arguments. Under the null, $M_{21}M_{21}'$ will have rank zero, implying that

$$C_f^{j_1,j_2}(M_{21}M_{21}', dd') = C_{j_2}(dd')$$

(note that *dd'* has rank one under both the null and the alternative). Then, (31) reduces to

$$z_{j_2} = \frac{((K_2-1)/2)_{j_2}}{j_2!(K_2)/2)_{j_2}}(d'd)^{j_2} \equiv \frac{((K_2-1)/2)_{j_2}}{j_2!(K_2)/2)_{j_2}}(M'_{22}M_{22})^{j_2}\,. \tag{33}$$

We have used the fact that $q^{0,j_2}_{j_2} = 1$ (Chikuse and Davis (1986)). The test can be implemented by using a consistent estimate of $\Omega_{22}^{-1}M'_{22}M_{22}$ based on sample data. The prob-value of an observed test statistic can then be calculate using (32) and compared against the desired level of significance. The computational demands of the exact test are less imposing in this case, following from the fact that the null restriction results in a linear regression with only exogenous regressors.

APPENDIX A

Proofs of the theorems and proposition follow.

<u>To prove Theorem 1 and its corollary</u>:

Under the action of the group G, the following mappings take place:

$$X_{21} \to HX_{21}Q_{11},\ X_{22} \to HX_{21}Q_{12} + HX_{22}Q_{22} \text{ and } W \to Q'WQ.$$

These lead to the transformations

$$X'_{22}\overline{P}_{X_{21}}X_{22} \to Q'_{22}X'_{22}\overline{P}_{X_{21}}X_{22}Q_{22} \text{ and } W_{22\cdot1} \to Q'_{22}W_{22\cdot1}Q_{22}.$$

Thus, we have established the invariance of Φ under the action of G. Suppose, now, that $\Phi\left(X_{21}, X_{22}, W\right) = \Phi\left(X^*_{21}, X^*_{22}, W^*\right)$ where $\Phi(\cdot)$ is to be regarded as a function of arguments appearing within the parentheses. Setting

$$Q_{22} = (X'_{22}\overline{P}_{X_{21}}X_{22})^{-1/2}\left(X^{*\prime}_{22}\overline{P}_{X^*_{21}}X^*_{22}\right)^{1/2},$$

it follows that

$$X^{*\prime}_{22}\overline{P}_{X^*_{21}}X^*_{22} = Q'_{22}(X'_{22}\overline{P}_{X_{21}}X_{22})Q_{22}.$$

Clearly, $Q_{22} \in GL(n_2)$. To completely identify a member of G which has mapped $\left(X_{21}, X_{22}, W\right)$ to $\left(X^*_{21}, X^*_{22}, W^*\right)$ we must pick elements $H \in O(K_2)$, $Q_{11} \in GL(n_1+1)$ and $Q_{12} \in M_{n_1+1,n_2}$ which satisfy the remaining requirements of the equivalence relation between the two points.

Instead of this, we note that elements of G which simultaneously satisfy the conditions

$$H = I_{K_2}, Q_{11} = I_{n_1+1}, Q_{12} = 0$$

constitute a subgroup of G, which we call G_1. Then the above arguments suffice to show that Φ is maximal invariant with respect to G_1. Also, we can write $G = G_1 \times G_2$ where the G_1 and G_2 are subgroups defined by

$$G_1 = \left\{(H,Q)\in G : H = I, Q_{11} = I, Q_{12} = 0\right\} \text{ and } G_2 = \left\{(H,Q)\in G : Q_{22} = I\right\}.$$

The corollary to Proposition 1 is proved by observing that the characteristic roots of Φ are invariant under the action of G_1.

<u>To prove Theorem 3</u>:

From Constantine (1966), we immediately have

$$pdf\,(T|X_{21}) = k_0 etr\left(-\frac{1}{2}\Delta_{X_{21}}\right)T^{n_2\mathrm{n}_1-1}\sum_{j_1=0}^{\infty}\sum_{\mathrm{a}_1}\frac{(-T)^{j_1}}{(n_2\mathrm{n}_1)_{j_1}\,j_1!}(\mathrm{n}_1+\mathrm{n}_2)_{\mathrm{a}_1}L^{\mathrm{h}}_{\mathrm{a}_1}\left(\frac{1}{2}\Delta_{X_{21}}\right)$$

for $0 \le T < 1$, where $\mathrm{h} = \mathrm{n}_1 - \frac{n_2+1}{2}$. In this, the generalized Laguerre polynomial is defined by (see Muirhead (1982))

$$L^{\mathrm{h}}_{\mathrm{a}_1}\left(\frac{1}{2}\Delta_{X_{21}}\right) = \left(\mathrm{h}+\frac{n_2+1}{2}\right)_{\mathrm{a}_1} C_{\mathrm{a}_1}(I_{n_2})\sum_{j_2=0}^{j_1}\sum_{\mathrm{a}_2}\binom{\mathrm{a}_1}{\mathrm{a}_2}\frac{C_{\mathrm{a}_2}(\frac{1}{2}\Delta_{X_{21}})}{(\mathrm{h}+(n_2+1)/2)_{\mathrm{a}_2}C_{\mathrm{a}_2}(I_{n_2})}$$

The conditional density is obtained by evaluating the expectation with respect to the density of X_{21}. Essentially, the steps involved in (9)-(12) are repeated with in place of pdf(Φ). We have

$$\begin{aligned}pdf\,(T) = {} & k_0(2\mathrm{p})^{-K_2(n_1+1)/2}\, etr\left(-\frac{1}{2}M_{21}'M_{21}\right)T^{n_2\mathrm{n}_1-1}\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{j_1}\frac{(-T)^{j_1}}{(n_2\mathrm{n}_1)_{j_1}\,j_1!}\\ & \times\sum_{\mathrm{a}_1}\sum_{\mathrm{a}_2}\binom{\mathrm{a}_1}{\mathrm{a}_2}\frac{(\mathrm{n}_1+\mathrm{n}_2)_{\mathrm{a}_1}(\mathrm{n}_1)_{\mathrm{a}_1}C_{\mathrm{a}_1}(I)}{(\mathrm{n}_1)_{\mathrm{a}_2}C_{\mathrm{a}_2}(I)}\int_{X_{21}\in M_{K_2,n_1+1}} etr\left(-\frac{1}{2}M_{22}'\overline{P}_{X_{21}}M_{22}\right)\\ & \times C_{\mathrm{a}_2}\left(-M_{22}'\overline{P}_{X_{21}}M_{22}\right)etr\left(-\frac{1}{2}X_{21}'X_{21}+M_{21}'X_{21}\right)(\mathrm{d}X_{21})\end{aligned}$$

This is followed by the familiar steps of factorization of measure and evaluation of iterated integrals which are defined respectively over the Stiefel manifold and the space of positive matrices. Note that α_1 and α_2 are partitions of j_1 and j_2 into n_2 parts, and α_3 and α_4 are partitions of j_3 and j_4 into K_2 parts, while ϕ is a partition of $(j_1+j_2+j_3)$ into K_2 parts, occurring as defined in Section 1. $\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}$ is the generalized binomial coefficient (Muirhead, 1982, Ch 7).

APPENDIX B

This appendix provides details of evaluations involving invariant polynomials. In the transition from (12) to (13), we employed Chikuse and Davis (1986, (3.6)) to evaluate the invariant polynomials in three matrix arguments which have special block-diagonal structure in which the diagonal matrices comprised only identity or null matrices. We obtained

$$C_{\phi}^{\alpha_1,\alpha_2,\alpha_3}(E_1E_1', E_1E_1', I - E_1E_1') = \sum_{\sigma^*\in\alpha_1\cdot\alpha_2}\sum_{\phi'\equiv\phi}\beta_{\sigma^*;\phi'}^{\alpha_1,\alpha_2,\alpha_3;\phi}\, C_{\phi'}^{\sigma^*,\alpha_3}(E_1E_1', I - E_1E_1')$$

where $\beta_{\sigma^*;\phi'}^{\alpha_1,\alpha_2,\alpha_3;\phi} = \sum_{\sigma\equiv\sigma^*}\theta_{\sigma}^{\alpha_1,\alpha_2}\gamma_{\sigma;\phi'}^{\alpha_1,\alpha_2,\alpha_3;\phi}$ with $\theta_{\sigma}^{\alpha_1,\alpha_2} = C_{\sigma}^{\alpha_1,\alpha_2}(I,I)\big/C_{\sigma}(I)$ which may be zero, and $\gamma_{\sigma;\phi'}^{\alpha_1,\alpha_2,\alpha_3;\phi} = \delta_{\sigma,\phi}\delta_{\phi,\phi'}$, δ being the Kronecker delta. Finally, we can write

$$C_{\phi'}^{\sigma^*,\alpha_3}(E_1E_1', I - E_1E_1') = \frac{((n_1+1)/2)_{\sigma^*}(\nu_2)_{\alpha_3}}{(K_2/2)_{\phi'}}\theta_{\phi'}^{\sigma^*,\alpha_3}C_{\phi'}(I_{K_2})$$

Inserting this into (12) gives the result in (13).

REFERENCES

Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Analysis*. Wiley.

Anderson, T.W., and H. Rubin (1949). “Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations,” *Annals of Mathematical Statistics*, 20, 46-63.

Anderson, T.W., and H. Rubin (1950). “The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations,” *Annals of Mathematical Statistics*, 21, 570-582.

Chikuse, Y. (1990a). “Distributions of Orientations on Stiefel Manifolds,” *Journal of Multivariate Analysis*, 33, 247-264.

Chikuse, Y. (1990b). “The Matrix Angular Central Gaussian Distribution,” *Journal of Multivariate Analysis*, 33, 265-274.

Chikuse, Y. and A.W. Davis (1986). “A Survey on the Invariant Polynomials with Matrix Arguments in Relation to Econometric Distribution Theory,” *Econometric Theory*, 2, 232-248.

Constantine, A.G. (1966). “The Distribution of Hotelling's Generalized T_o^2,” *Annals of Mathematical Statistics*, 37, 215-225.

Cragg, J.G., and S.G. Donald (1993). “Testing Identifiability and Specification in Instrumental Variable Models,” *Econometric Theory*, 9, 222-240.

Eaton, M. L. (1983). *Multivariate Statistics: A Vector Space Approach*. Wiley.

Giri, N.C. (1977). *Multivariate Statistical Inference*. Academic Press, New York.

Hausman, J.A. (1983). “Specification and Estimation of Simultaneous Equation Models,” in Z. Griliches and M.D. Intriligator (eds.), *Handbook of Econometrics*, Vol 1, Chapter 7, 391-443. North Holland, Amsterdam.

Herz, C.S. (1955). “Bessel Functions of Matrix Arguments,” *Annals of Mathematics,* 61, 474-523.

Hillier, G.H. (1985). “On the Joint and Marginal Densities of Instrumental Varaibles Estimators in a General Structural Equation,” *Econometric Theory*, 1, 53-72.

Hillier, G.H. (1987). “Hypothesis Testing In A Structural Equation: Part I. Reduced Form Equivalences and Invariant Test Procedures,” unpublished manuscript, Monash University.

Hillier, G.H. (1990). “On The Normalization of Structural Equations: Properties of Direction Estimators,” *Econometrica*, 58, 1181-1194.

Hillier, G.H., T. 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.

Hillier, G.H. and C.L. Skeels (1993), “Some further exact results for structural equation estimators” in P.C.B. Phillips (ed.) *Models, Methods and Applications of Econometrics*. Blackwell.

Hosoya, Y., Y. Tsukuda and N. Terui (1989), “Ancillarity and the limited information maximum-likelihood estimation of a structural equation in a simultaneous equation system,” *Econometric Theory*, 5, 385-404.

Lucas, R.E. (1976). “Econometric Policy Evaluation: A Critique,” in K. Brunner and A.H. Meltzer (eds.) *The Phillips Curve and Labor Economics*, *Supplement* to the *Journal of Monetary Economics*, 1, 19-46.

Morimune, K. (1988). “Higher-Order Approximations to the Null Distributions of Test Statistics for Nonlinear Restrictions on Regression Coefficients,” *Econometric Theory*, 4, 275-299.

Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York.

Phillips, P.C.B. (1983). “Exact Small Sample Theory in the Simultaneous Equations Model,” in Z. Griliches and M.D. Intriligator (eds.), *Handbook of Econometrics*, Vol 1, Chapter 8, 449-516. North Holland, Amsterdam.

Rao, C.R. (1973). *Linear Statistical Inference and Its Applications*, Second Edition. Wiley.

Revanker, N.S. and P. Mallela (1972). “The Power Function of an F-Test in the Context of a Structural Equation,” *Econometrica*, 40, 913-916.

van Garderen, K.J. (1994). “Curved Exponential Models in Econometrics,” mimeo, Department of Applied Economics, University of Cambridge.