TESTS OF OVERIDENTIFICATION AND EXOGENEITY IN SIMULTANEOUS EQUATION MODELS

T. W. ANDERSON and NAOTO KUNITOMO

TECHNICAL REPORT NO. 36
January 1990

PREPARED UNDER
NATIONAL SCIENCE FOUNDATION GRANT DMS 89-04851

Econometric Workshop
Fourth Floor, Encina Hall
Stanford University
Stanford, California
TESTS OF OVERIDENTIFICATION AND EXOGENEITY IN SIMULTANEOUS EQUATION MODELS

T. W. ANDERSON* and NAOTO KUNITOMO

TECHNICAL REPORT NO. 36
January 1990

PREPARED UNDER
NATIONAL SCIENCE FOUNDATION GRANT DMS 89-04851

Econometric Workshop
Fourth Floor, Encina Hall
Stanford University
Stanford, California

* Some of this research was conducted as a visitor at the University of Southern California and as a Visiting Distinguished Professor of the Norwegian Council for Scientific and Industrial Research at the University of Oslo.
Abstract

A set of simultaneous equations that is a part of a larger complete system is considered. One hypothesis concerning this set of equations is that it is overidentified as a block by a block of zero restrictions. Another hypothesis is that a subset of the current endogenous variables is uncorrelated with the disturbances in the set of simultaneous equations ("exogeneity"). Likelihood ratio, Lagrange multiplier, and Wald tests are derived for the overidentification hypothesis against unrestricted alternatives, for the exogeneity hypothesis against unrestricted alternatives, and for the exogeneity hypothesis against the alternative of overidentification. The tests are derived on the basis of normality but are asymptotically valid under very general conditions. This study generalizes and unifies many test procedures proposed previously. Some new tests are developed.

Key words and phrases: simultaneous equation models, overidentification, exogeneity, likelihood ratio tests, Lagrange multiplier tests, Wald tests
1. Introduction

Two important underlying assumptions of the traditional simultaneous equation approach in econometrics are the identifying restrictions and predeterminedness (or exogeneity in some sense) of several variables in the system of structural equations. Although these assumptions are often made on a priori ground, in practice it may be advisable to examine these two conditions from a statistical point of view. In this respect a number of statistical testing procedures for these restrictions have been proposed by econometricians. For instance, the test procedures of Anderson and Rubin (1949), Koopmans and Hood (1953), Basman (1960), Byron (1972), Wu (1973), Revankar and Hartley (1973), Revankar (1978), Hausman (1978), Kariya and Hodoshima (1980), Hwang (1980a), Hillier (1987), and Revankar and Yoshino (1989) among many others have drawn attention and have been applied in empirical works. However, since many testing procedures have been introduced based on intuitive reasoning, it may be difficult to understand the meaning of the statistics proposed.

The main purposes of this paper are to derive systematically several test procedures for each condition and to obtain the relationships among the different test statistics. For these intentions we consider a subsystem of structural equations and regard the single equation method as a special case of our formulation. Then we shall derive three types of test procedures, namely, the likelihood ratio (LR) test, Lagrange Multiplier (LM) test, and the Wald test for the block identifiability restrictions and the predeterminedness restrictions in the subsystem of structural equations. In this framework the test statistics we shall derive include most of the test statistics mentioned above as special cases and give new interpretations to some of them. These interpretations also apply to some test statistics commonly known in multivariate statistical analysis.

In a subsequent paper we shall derive the asymptotic distributions of these test criteria under very general conditions, based on a new central limit theorem using a Lindeberg-type condition for martingale differences [Anderson and Kunitomo (1989a,b)].

In Section 2 we formulate a subsystem of structural equations. In Section 3 we derive several statistics for testing identifying restrictions. We also relate these statistics to the statistics in multivariate statistical analysis yielding new interpretations of statistics.
commonly known among statisticians. Subsequently, in Section 4 we derive a number of
test procedures for testing econometric predetermination restrictions. Finally, in Section
5 some concluding remarks are given. Useful lemmas are given in the appendices.

2. Two Hypotheses in a Subsystem of Structural Equations

2.1. The model.

We consider a subsystem of $G_0$ structural equations

\begin{equation}
YB = Z_1 \Gamma + U,
\end{equation}

where $Y$ is a $T \times G$ matrix of observations on the endogenous variables appearing in the
first $G_0$ structural equations, $Z_1$ is a $T \times K_1$ matrix of observations on the $K_1$ exogenous
variables, $B$ and $\Gamma$ are $G \times G_0$ and $K_1 \times G_0$ matrices of (unknown) parameters, respectively,
and $U$ is a $T \times G_0$ matrix of unobservable disturbances. When $G_0 = 1$, (2.1) is the usual
single structural equation. We require the columns of $B$ to be linearly independent; that
is, the rank of $B$ is $G_0$.

The reduced form equation for the endogenous variables $Y$ appearing in the first $G_0$
structural equations (2.1) with $K (K = K_1 + K_2)$ predetermined variables is

\begin{equation}
Y = Z \Pi + V,
\end{equation}

where $Z = (Z_1, Z_2)$ is a $T \times K$ matrix of predetermined variables ($T > K$) of rank $K$,
and $Z_2$ is a $T \times K_2$ matrix of predetermined variables that are not included in (2.1). The
predetermined variables may include lagged endogenous variables. $V$ is a $T \times G$ matrix of
disturbances whose $t$–th row is denoted by $v_t'$. We assume that

\begin{equation}
E(v_t) = 0,
\end{equation}

\begin{equation}
E(v_tv_t') = \Omega,
\end{equation}

where $\Omega$ is a $G \times G$ positive definite matrix.
In this paper we shall consider two hypotheses. One is that the set of $G_0$ equations (2.1) is identified as a block. That is, any matrix $B$ such that $Z\Pi B = Z_1\Gamma$ for some $\Gamma$ is obtained from any other by multiplication on the right by a nonsingular $G_0 \times G_0$ matrix. The other hypothesis that we consider is that a subset of the endogenous variables is uncorrelated with the disturbances in the block of equations.

2.2. Block identification.

The relationship between the reduced form and the structural equations involves

\begin{align}
(2.5) \quad \Gamma &= \Pi_1.B, \\
(2.6) \quad U &= VB.
\end{align}

where $\Pi$ has been partitioned as

\begin{equation}
(2.7) \quad \Pi = \begin{pmatrix} \Pi_1. \\ \Pi_2. \end{pmatrix}.
\end{equation}

Let $u'_t$ be the $t$-th row of $U$. From (2.3), (2.4), and (2.6) we obtain

\begin{equation}
(2.8) \quad E(u_t) = 0,
\end{equation}

\begin{equation}
(2.9) \quad E(u_tu'_t) = B'\Omega B = \Sigma,
\end{equation}

where $\Sigma$ is a $G_0 \times G_0$ positive definite matrix. The block identifiability conditions are expressed as

\begin{equation}
(2.10) \quad H_\xi : \xi = 0,
\end{equation}

where

\begin{equation}
(2.11) \quad \xi = \Pi_2.B.
\end{equation}

From (2.11) we obtain the rank condition for the identifiability of (2.1),

\begin{equation}
(2.12) \quad \text{rank}\Pi_2. = G - G_0.
\end{equation}
The order condition is

\[(2.13)\quad L = K_2 - (G - G_0) \geq 0.\]

In the above notation, \(L\) is called the degree of overidentification.

Let \(\nu_G \geq \cdots \geq \nu_1 \geq 0\) be the roots of

\[(2.14)\quad \left| \frac{1}{T} \Theta_T - \nu \Omega \right| = 0,
\]

where

\[(2.15)\quad \Theta_T = \Pi'_2 \cdot A_{22,1} \Pi_2.\]

\[(2.16)\quad A_{22,1} = Z'_1 Z_2 - Z'_2 Z_1 (Z'_1 Z_1)^{-1} Z'_1 Z_2.\]

The block identifiability conditions are equivalent to the hypothesis \(H_\nu : \nu_1 = \cdots = \nu_{G_0} = 0\) and \(\nu_{G_0+1} > 0\). The existence of a matrix \(B\) such that \(\xi = 0\) is equivalent to (2.12), which, in turn, is equivalent to \(H_\nu\).

Note that the model (2.1) and the hypothesis \(H_\xi\) are invariant with respect to linear transformations on the right; that is, (2.1) and (2.10) can be multiplied on the right by an arbitrary nonsingular matrix \(A\) to yield another set of structural parameters

\[(2.16')\quad \tilde{B} = BA, \quad \tilde{\Gamma} = \Gamma A, \quad \tilde{\Sigma} = A' \Sigma A.\]

It may be convenient for some purpose to select a particular triple \((B, \Gamma, \Sigma)\) by a suitable normalization of \(B\) such as requiring a submatrix of \(B\) to be \(I_{G_0}\). The test procedures are invariant with respect to the group of transformations and hence do not depend on the normalization.

2.3. Exogeneity.

An essential difference between a system of structural equations and regression models in the multivariate analysis is that in the former correlation may exist between the endogenous variables \(y_t\), which is the \(t\)-th row of \(Y\), that is, \(v_t\) and the corresponding
disturbance term $u'_t$, but in the latter some components of $y'_t$ and $u'_t$ may be uncorrelated. In order to state this hypothesis we partition $Y = (Y_1, Y_2)$ into $G_1$ and $G_2$ columns ($G = G_1 + G_2$), $V = (V_1, V_2)$, and

\begin{equation}
\Omega = \begin{pmatrix}
\Omega_{11} & \Omega_{12} \\
\Omega_{21} & \Omega_{22}
\end{pmatrix}.
\end{equation}

From (2.9) the covariance matrix of $v'_{2t}$ and $u'_t$ is

\begin{equation}
\eta = \text{Cov}(v'_{2t}, u'_t) = (\Omega_{21}, \Omega_{22})B.
\end{equation}

We define the econometric predeterminedness restriction considered in this paper to be the hypothesis $H_\eta : \eta = 0$. The two hypotheses $H_\xi$ and $H_\eta$ imply the hypothesis $H_{\xi,\eta} : \xi = 0, \eta = 0$. When the disturbance terms follow the multivariate normal distribution, the uncorrelatedness implies an independence between a subset of regressors $Y_2$ and disturbance terms in (2.1). This testing problem has been sometimes called the test of independence. The hypothesis of predeterminedness in this paper has also been called weak exogeneity in econometrics. There are several different concepts of econometric exogeneity in simultaneous equation systems. Engle, Hendry, and Richard (1983) surveyed this issue in a systematic way. See Holly (1987), also.

The hypothesis of exogeneity is also invariant with respect to linear transformations on the right; that is, $\eta$ in (2.18) can be multiplied on the right by a nonsingular matrix $A$. The test criteria for exogeneity are also invariant with respect to such transformations and hence with respect to normalization of $B$.

In Section 3 we obtain the likelihood ratio, the Lagrange multiplier, and Wald-type tests of over-identification. In Section 4 we find test criteria for predeterminedness.

3. Tests of Block Identifiability

In order to derive test statistics we assume that the disturbance terms \{\$u_t\$\} are independently and normally distributed. The derivation here is considerably simpler than alternatives already known.
3.1. Likelihood ratio test.

Under the assumption of normal disturbances, the log likelihood function for \( Y' = (y_1, \ldots, y_T) \) is

\[
\log L_1 = c_1 - \frac{1}{2} T \log |\Omega| - \frac{1}{2} \text{tr}(Y - Z\Pi)'(Y - Z\Pi)\Omega^{-1},
\]

where \( c_1 = -\frac{1}{2} GT \log(2\pi) \). To maximize \( L_1 \) with respect to the covariance matrix \( \Omega \), we use Lemma A.1. The concentrated likelihood function is

\[
\log L_2 = c_2 - \frac{1}{2} T \log |S|,
\]

where \( c_2 = c_1 + \frac{1}{2} GT \log T - \frac{1}{2} GT \) and \( S = (Y - Z\Pi)'(Y - Z\Pi) \). The log likelihood maximized under the alternative hypothesis \( H_A : \xi \neq 0 \) (that is, no restriction) is

\[
\log L_3 = c_2 - \frac{1}{2} T \log |Y'\tilde{P}_F Y|,
\]

where \( P_F = F(F'F)^{-1}F' \) denotes the projection operator onto the space spanned by the columns of \( F \) and \( \tilde{P}_F = I - P_F \) for any matrix \( F \) of full column rank.

To maximize the likelihood under the null hypothesis \( H_0 : \xi = 0 \) we define the \( G \times G \) matrix

\[
H = \begin{pmatrix} B & 0 \\ I_{G*} & \end{pmatrix},
\]

where \( G* = G - G_0 \). Since \( B \) is of rank \( G_0 \), there is a \( G_0 \times G_0 \) submatrix that is nonsingular. When the numbering of components of \( y_t \) is such that this matrix consists of the first \( G_0 \) rows of \( B \), then \( H \) is nonsingular. The concentrated likelihood is rewritten as

\[
\log L_2 = c_2 + \frac{1}{2} T \log |H'H| - \frac{1}{2} T \log |S^*|,
\]

where \( S^* = H'SH = (W - Z\Pi^*)(W - Z\Pi^*), W = (W_0, W_*) = YH = (YB, Y_*), \) and

\[
\Pi^* = (\Pi^*_0, \Pi^*_*) = \Pi H = \begin{pmatrix} \Gamma \\ \xi & \Pi^*_* \end{pmatrix}.
\]

The unrestricted least squares estimate of \( \Pi^* \) is

\[
\hat{\Pi}^* = (\hat{\Pi}^*_0, \hat{\Pi}^*_*) = (Z'Z)^{-1}Z'W.
\]
Then

\[(3.8)\quad S^* = (W - Z\Pi^*)(W - Z\Pi^*)
\]
\[= W'PZW + (\hat{\Pi}^* - \Pi^*)Z'Z(\hat{\Pi}^* - \Pi^*)
\]
\[= W'PZW + [Z(\hat{\Pi}_{0}^* - \Pi_{0}^*), Z(\hat{\Pi}_{*}^* - \Pi_{*}^*)]'[Z(\hat{\Pi}_{0}^* - \Pi_{0}^*), Z(\hat{\Pi}_{*}^* - \Pi_{*}^*)].
\]

By Lemma A.2 the minimum of \(|S^*|\) with respect to \(\Pi_{*} = \Pi_{*}^*\) is

\[(3.9)\quad \frac{|W'PZW| \cdot |W_0'PZW_0 + (\hat{\Pi}_{0}^* - \Pi_{0}^*)'Z'Z(\hat{\Pi}_{0}^* - \Pi_{0}^*)|}{|W_0'PZW_0|}.
\]

The second determinant in the numerator of (3.9) is \(|(W_0 - Z\Pi_{0}^*)(W_0 - Z\Pi_{0}^*)|\), which is \(|(W_0 - Z_1\Gamma)(W_0 - Z_1\Gamma)|\) if \(\xi = 0\). That determinant is minimized with respect to \(\Gamma\) at \(\hat{\Gamma} = (Z_1Z_1)^{-1}Z_1^tYB\). The log likelihood ratio criterion is the maximum with respect to \(B\) of

\[(3.10)\quad \frac{1}{2}T \log \frac{|S| \cdot |H'H| \cdot |W_0'PZW_0|}{|W_0'PZW_0|^2} = \frac{1}{2}T \log \frac{|B'Y'PZYB|}{|B'Y'PZ_1YB|}.
\]

Lemma A.3 implies that the maximum of (3.10) is \(T/2\) times the sum of the \(G_0\) smallest characteristic roots of \(Y'PZ_1Y(Y'PZY)^{-1}\). The log likelihood ratio times \(-2\) is

\[(3.11)\quad LR_1 = T \sum_{i=1}^{G_0} \log(1 + \lambda_i),
\]

where \(\lambda_G \geq \cdots \geq \lambda_1 \geq 0\) are the roots of

\[(3.12)\quad |Y'(P_Z - P_{Z_1})Y - \lambda Y'PZY| = 0.
\]

The above equation is a sample analogue of (2.14).

The likelihood ratio statistic (3.11) for \(G_0 = 1\) was derived by Anderson and Rubin (1949); \(LR_1\) corresponds to the smallest root in the limited information maximum likelihood (LIML) estimation method. When \(G_0 = 2\), \(LR_1\) is identical to the statistic proposed by Koopmans and Hood (1953) as the non-identification test. Anderson (1951) was the first to obtain the likelihood ratio criterion (3.11), which he did by differentiation.
3.2. Lagrange multiplier or score test.

The Lagrange Multiplier statistic, which is identical to the Rao score statistic, has been developed as a test statistic to test a hypothesis $H$ about a vector parameter $\theta$ in a likelihood $L$. In these general terms the criterion is

$$(3.13) \quad LM = \left( \frac{\partial \log L}{\partial \theta} \right)^T_H \left( -\frac{\partial^2 \log L}{\partial \theta \partial \theta'} \right)_H^{-1} \left( \frac{\partial \log L}{\partial \theta} \right)_H$$

where $H$ denotes the null hypothesis and the value of the parameter in (3.13) maximizes the likelihood under the null hypothesis. If the null hypothesis is

$$(3.14) \quad H : h(\theta) = 0$$

and $\lambda$ is vector of Lagrange multipliers

$$(3.15) \quad \left. \frac{\partial \log L}{\partial \theta} \right|_H = -(\lambda|_H)' \left. \frac{\partial h}{\partial \theta'} \right|_H.$$

Then

$$(3.16) \quad LM = (\lambda|_H)'(\text{est. asymp cov. of } \lambda)^{-1}(\lambda|_H).$$

In our problem $h(\theta) = \text{vec } \Pi_2.B$, $\lambda = \text{vec } \Lambda$, where $\Lambda$ is $K_2 \times G_0$, and the Lagrange form is

$$(3.17) \quad \log L_4 = c_1 - \frac{T}{2} \log |\Omega| - \frac{1}{2} \text{tr } \Omega^{-1}(Y - Z\Pi)'(Y - Z\Pi) + \text{tr } \Lambda'\Pi_2.B.$$

Setting to 0 the derivative of $\log L_4$ with respect to each element of $\Pi$, we obtain

$$(3.18) \quad Z'(Y - Z\Pi)\Omega^{-1} + \begin{pmatrix} 0 \\ \lambda \end{pmatrix} B' = 0.$$ 

The upper half part of (3.18) gives $Z_1'(Y - Z\tilde{\Pi}) = 0$, and we have

$$(3.19) \quad \tilde{\Pi}_1 = (Z_1'Z_1)^{-1}Z_1'(Y - Z_2\tilde{\Pi}_2).$$

Then

$$(3.20) \quad Y - Z\tilde{\Pi} = Y - Z_1\tilde{\Pi}_1 - Z_2\tilde{\Pi}_2.$$

$$= \bar{P}_Z(Y - Z_2\tilde{\Pi}_2).$$
Multiplying (3.18) on the right by $\Omega B$, we obtain

\[(3.21) \quad \begin{pmatrix} 0 \\ \Lambda \end{pmatrix} = -Z'\bar{P}_Z Y B \Sigma^{-1}. \]

Using Lemma A.4, we find the first and second derivatives of the log-likelihood function as

\[(3.22) \quad \frac{\partial \log L_4}{\partial \text{vec } \Pi} = \frac{\partial \log L_1}{\partial \text{vec } \Pi} + \text{vec} \left[ \begin{pmatrix} 0 \\ \Lambda \end{pmatrix} B' \right], \quad \frac{\partial^2 \log L_1}{\partial \text{vec } \Pi \partial (\text{vec } \Pi)' } = -\Omega^{-1} \otimes Z'Z. \]

Then we define an LM statistic by

\[(3.23) \quad LM_1 = \left( \frac{\partial \log L_1}{\partial \text{vec } \Pi} \bigg|_{\Pi=\hat{\Pi}, \Omega=\hat{\Omega}} \right)' \left( -\frac{\partial^2 \log L_1}{\partial \text{vec } \Pi \partial (\text{vec } \Pi)' } \bigg|_{\Pi=\hat{\Pi}, \Omega=\hat{\Omega}} \right)^{-1} \left( \frac{\partial \log L_1}{\partial \text{vec } \Pi} \bigg|_{\Pi=\hat{\Pi}, \Omega=\hat{\Omega}} \right); \]

$\Omega$ and $\Pi$ are evaluated at their maximum likelihood estimators under the null hypothesis. Using Lemma A.5, $\Sigma = B'\Omega B$, and

\[(3.24) \quad \frac{\partial \log L_1}{\partial \text{vec } \Pi} \bigg|_{\Pi=\hat{\Pi}, \Omega=\hat{\Omega}} = 0, \]

we have

\[(3.25) \quad LM_1 = \left\{ \text{vec} \left[ \begin{pmatrix} 0 \\ \Lambda \end{pmatrix} B' \right] \right\}' \left[ \Omega \otimes (Z'Z)^{-1} \right] \text{vec} \left[ \begin{pmatrix} 0 \\ \Lambda \end{pmatrix} B' \right]
\]

\[= \text{tr} \Sigma (0_\Lambda')(Z'Z)^{-1} \begin{pmatrix} 0 \\ \Lambda \end{pmatrix}, \]

where the unknown parameters in (3.25) are evaluated at their maximum likelihood estimators under the null hypothesis. (See Engle (1984), for instance.) The LM statistic in the form (3.23) is known as Rao's Score test statistic among statisticians. From (3.18) and $\bar{P}_Z P_Z \bar{P}_Z = P_Z - P_{Z_1}$, we further simplify $LM_1$ as

\[(3.26) \quad LM_1 = \text{tr} \hat{B}_H' Y' (P_Z - P_{Z_1}) Y \hat{B}_H \hat{\Sigma}_H^{-1}; \]

where $\hat{B}_H$ and $\hat{\Sigma}_H = (1/T)\hat{B}_H' Y' P_Z Y \hat{B}_H$ are the maximum likelihood estimators of $B$ and $\Sigma$ under the null hypothesis. Let $c_i$ satisfy

\[(3.27) \quad Y'(P_Z - P_{Z_1}) Y c = \lambda_i Y' \bar{P}_Z Y c, \]

9
\[(3.27') \quad c'Y'\bar{P}_ZYc = T,\]

where \(\lambda_i\) satisfies (3.12), \(i = 1, \ldots, G_0\), and let

\[(3.27'') \quad C = (c_1, \ldots, c_{G_0}).\]

Then \(\hat{B}_H = CA\) for arbitrary nonsingular \(A\) and \(\hat{\Sigma}_H = (1/T)\hat{B}_HY'\bar{P}_ZY\hat{B}_H\). (See Appendix B). When we use the roots of (3.12), this statistic is expressed as

\[(3.28) \quad LM_1 = T \sum_{i=1}^{G_0} \frac{\lambda_i}{1 + \lambda_i}.\]

When \(G_0 = 1\), this statistic \(LM_1\) is the LM statistic proposed by Byron (1972). However, his derivation of the statistic is different from ours. It should be also noted that (3.28) is an analogue of the Bartlett-Nanda-Pillai trace criterion, which is well known in multivariate statistical analysis. (See Anderson (1984), Chapter 8.) Our derivation yields a new interpretation of the Bartlett-Nanda-Pillai test.

### 3.3. Wald test.

In general terms the Wald test is based on the statistic

\[(3.29) \quad h(\hat{\theta})'[\text{asymp. cov. of } h(\hat{\theta})]^{-1}h(\hat{\theta}),\]

where \(\hat{\theta}\) is the maximum likelihood estimator of the parameter vector \(\theta\) under the alternative hypothesis. In our problem the null hypothesis is that the rank of \(\Pi_2\) is \(G - G_0 = G_*\). To express this in the form of \(h(\theta) = 0\) we partition \(\Pi_2\) into \(L = K_2 - G_*\) and \(G_*\) rows and \(G_0\) and \(G_*\) columns:

\[(3.30) \quad \Pi_2. = \begin{pmatrix} \Pi_{\ell_0} & \Pi_{\ell_*} \\ \Pi_{m_0} & \Pi_{m_*} \end{pmatrix}.\]

Since \(\Pi_2.\) is of rank \(G_*\), there is at least one square matrix \(\Pi_{m_*}\) of order \(G_*\) that is nonsingular. There exists a \(G \times G_0\) matrix

\[(3.31) \quad B = \begin{pmatrix} B_0 \\ -B_* \end{pmatrix},\]
where \( B_0 \) is \( G_0 \times G_0 \), such that \( \xi = \Pi_2 B = 0 \).

This equation can be partitioned as

\[
(3.32) \quad \xi_\ell = \Pi_{\ell 0} B_0 - \Pi_{\ell *} B_* = 0,
\]

\[
(3.33) \quad \xi_m = \Pi_{m 0} B_0 - \Pi_{m *} B_* = 0.
\]

The second equation yields

\[
(3.34) \quad B_* = \Pi_{m *}^{-1} \Pi_{m 0} B_0.
\]

Substitution into (3.32) yields

\[
(3.35) \quad \xi_\ell = (\Pi_{\ell 0} - \Pi_{\ell *} \Pi_{m *}^{-1} \Pi_{m 0}) B_0 = 0.
\]

Since \( B \) is of rank \( G_0 \), \( B_0 \) is also of rank \( G_0 \). (If \( B_0 \) were of lower rank, there would exist a vector \( c \) such that \( B_0 c = 0 \); then by (3.34) \( B c = 0 \).) Hence, if \( \Pi_2 \) is of rank \( G_* \), \( \Pi_{\ell 0} - \Pi_{\ell *} \Pi_{m *}^{-1} \Pi_{m 0} = 0 \). We take

\[
(3.36) \quad h(\theta) = \text{vec} (\Pi_{\ell 0} - \Pi_{\ell *} \Pi_{m *}^{-1} \Pi_{m 0}).
\]

Let \( Y = (Y_0, Y_*) \), \( V = (V_0, V_*) \), \( \Pi_1 = (\Pi_{10}, \Pi_{1*}) \), and \( Z_2 = (Z_\ell, Z_m) \). The reduced form is

\[
(3.37) \quad Y_0 = Z_1 \Pi_{10} + Z_\ell \Pi_{\ell 0} + Z_m \Pi_{m 0} + V_0,
\]

\[
(3.38) \quad Y_* = Z_1 \Pi_{1*} + Z_\ell \Pi_{\ell *} + Z_m \Pi_{m *} + V_*.
\]

A just-identified set of structural equations is

\[
(3.39) \quad Y_0 = Y_* B_* + Z_1 \Gamma + Z_\ell \xi_\ell + U,
\]
where \( B_* \) is given by (3.34) and \( B_0 = I \). Then a set of parameters is \( B_*, \Gamma, \xi, \Pi_1, \Pi_\xi, \Pi_m, \) and \( \Omega \); the null hypothesis is defined by \( \xi = 0 \). The maximum likelihood estimator (under the alternative) of \( B_*, \Gamma \), and \( \xi \) is the indirect least squares estimator, which is the solution to

\[
\begin{pmatrix}
\hat{Y}_t' \\
Z_t'
\end{pmatrix}
\begin{pmatrix}
\hat{Y}_*, Z_1, Z_t \\
\hat{B}_t
\end{pmatrix}
\begin{pmatrix}
\hat{f} \\
\hat{\xi}
\end{pmatrix}
= \begin{pmatrix}
\hat{Y}_t' \\
Z_t'
\end{pmatrix} Y_0,
\]

where \( \hat{Y}_* = Z(Z'Z)^{-1}Z'Y_*. \) Then solving the above equation with respect to \( \hat{\xi} \), we have

\[
Z_t'\hat{P}_{Y_*, Z_1} Z_t \hat{\xi} = Z_t'\hat{P}_{Y_*, Z_1} Y_0.
\]

Applying Lemma A.6 for \( \hat{P}_{Y_*, Z_1} \) and noting that \( Z_t = P_Z Z_t \), we write the estimator of \( \xi \) as

\[
\hat{\xi} = (Z_t'N Z_t)^{-1} Z_t' N Y_0,
\]

where

\[
N = \hat{P}_{Z_t} - \hat{P}_Z - (\hat{P}_{Z_t} - \hat{P}_Z) Y_* \left[ Y'_t (\hat{P}_{Z_t} - \hat{P}_Z) Y_* \right]^{-1} Y'_t (\hat{P}_{Z_t} - \hat{P}_Z).
\]

In the above we have utilized the fact that \( Z_t'N Z_t \) is nonsingular because the matrix \((Y_*, Z_1, Z_t)\) is of full rank (a.s.) and rank \((N) = \text{rank}(Z_t) = L \) (a.s.). Then the covariance matrix of the limiting normal distribution of \( \sqrt{T} \text{vec}(\hat{\xi} - \xi) \) is the probability limit of

\[
T \left[ I_{G_0} \otimes (Z_t'N Z_t)^{-1} Z_t'N \right] \left( \Sigma \otimes I_T \right) \left[ I_{G_0} \otimes (Z_t'N Z_t)^{-1} Z_t'N \right]' = \Sigma \otimes \left( \frac{1}{T} Z_t'N Z_t \right)^{-1}.
\]

We now define a Wald-type statistic by

\[
W_1 = \text{vec}(\hat{\xi})' \left[ \hat{\Sigma}_I \otimes (Z_t'N Z_t)^{-1} \right]^{-1} \text{vec}(\hat{\xi}),
\]

where \( \hat{\Sigma}_I = \hat{B}_t' \hat{\Gamma} \hat{B}_t, \hat{B}_t = (I_{G_0}, -\hat{B}_*'), \) and \( \hat{\Gamma} = (1/T)Y'\hat{P}_Z Y \). Since we have normalized \( B_1 = I_{G_1} \), the two–stage least squares estimator of \( B \) under \( H_\xi: \xi = 0 \) is \( \hat{B}_{TS} = (I_{G_0}, -\hat{B}_*) \)
and $\hat{B}_*$ satisfies $[Y'_*(P_Z - P_{Z_1})Y_*] \hat{B}_* = Y'_*(P_Z - P_{Z_1})Y_0$. Thus using Lemma A.6 again, we have

\begin{equation}
(3.45) W_1 = \left\{ \text{vec } [Z_t'N Z_t]^{-1} Z_t'NY_0 \right\}' \left[ \hat{\Sigma} \otimes (Z_t'NZ_t)^{-1} \right]^{-1} \text{vec } [(Z_t'NZ_t)^{-1} Z_t'NY_0] \\
= [\text{vec}(Z_t'NY_0)]' [I \otimes (Z_t'NZ_t)^{-1}] [\hat{\Sigma}^{-1} \otimes (Z_t'NZ_t)^{-1}] \text{vec}(Z_t'NY_0) \\
= [\text{vec}(Z_t'NY_0)]' [\hat{\Sigma}^{-1} \otimes (Z_t'NZ_t)^{-1}] \text{vec}(Z_t'NY_0) \\
= \text{tr} \left\{ \hat{\Sigma}^{-1} Y_0'NZ_t(Z_t'NZ_t)^{-1} Z_t'NY_0 \right\}.
\end{equation}

Since $N^2 = N$, there exists a $T \times L$ matrix $X$ such that $N = X(X'X)^{-1}X'$ and

\begin{equation}
(3.46) \quad NZ_t(Z_t'NZ_t)^{-1}Z_tN \\
= X(X'X)^{-1}X'Z_t[Z_t'X(X'X)^{-1}X'Z_t]^{-1}Z_t'X(X'X)^{-1}X' \\
= X(X'X)^{-1}X' = N.
\end{equation}

Then we have

\begin{equation}
(3.47) \quad W_1 = \text{tr } \left\{ \hat{\Sigma}^{-1} Y_0'NY_0 \right\} \\
= \text{tr } \left\{ \hat{\Sigma}^{-1} (Y_0 - Y_0 \hat{B}_*)'(P_Z - P_{Z_1})(Y_0 - Y_0 \hat{B}_*) \right\} \\
= \text{tr } \left\{ \hat{\Sigma}^{-1} \hat{B}_{TS}'Y'(P_Z - P_{Z_1})Y \hat{B}_{TS} \right\}.
\end{equation}

The above derivation is an extension of Hwang (1980b). The last expression in (3.47) shows that except for $\hat{\Sigma}$ the criterion does not depend on the selection of variables to define $\xi_t$ and $\xi_m$, but it does depend on the selection of variables to define $Y_0$ and $Y_1$. The criterion (3.47) can be modified by defining $\hat{\Sigma}$ as $\hat{B}'\hat{\Omega}\hat{B}$ with $\hat{B}$ as some other estimate of $B$. The estimate of $B$ obtained when $\xi_m = 0$ (that is, the alternative hypothesis holds) is $(I, -\hat{B}_*)'$, where $\hat{B}_*$ is defined by (3.40). If $\hat{B} = \hat{B}_{TS}$, then (3.47) is completely independent of the selection of variables in $\xi_t$ and $\xi_m$. This is the statistic derived by Wegge (1978) for $G_0 = 1$. Hwang (1980b) has shown that it is identical to the Wald statistic proposed by Byron (1974). If we use the maximum likelihood estimator of $\Sigma$ under the null hypothesis,

\begin{equation}
(3.48) \quad \hat{\Sigma} = \frac{1}{T} \hat{B}'Y'Y \hat{B},
\end{equation}

13
the resulting statistic reduces to the statistic proposed by Basmann (1960) for the case of \( G_0 = 1 \). It can be interpreted as a Wald–type statistic in the present context.

If \( \hat{B} = \hat{B}_{LI} \), the limited information maximum likelihood estimator under \( H_\xi : \xi = 0 \), and \( \hat{B}_{TS} \) in (3.47) is replaced by \( \hat{B}_{LI} \), the statistic is

\[
W_1' = T \text{tr}(\hat{B}_{LI} Y' \overline{P}_Z Y \hat{B}_{LI})^{-1} \hat{B}_{LI} Y'(P_Z - P_{Z_1}) Y \hat{B}_{LI}
\]

\[
= T \sum_{i=1}^{G_0} \lambda_i,
\]

where \( \lambda_1, \ldots, \lambda_{G_0} \) are the \( G_0 \) smallest roots of (3.12).

It should be also noted that \( W_1' \) is an analogue (or generalization) of the Lawley–Hotelling Trace Criterion, which is well–known in multivariate statistical analysis. (See Anderson (1984), Chapter 8.) Thus our derivation also gives a new interpretation to the Lawley–Hotelling type statistic.

### 3.4 An Inequality Among Statistics.

We have derived three types of statistics for the block identifying restriction in a subsystem of structural equations. There is a simple inequality among the statistics we have derived. Using Lemma A.7, we have

\[
0 \leq LM_1 \leq LR_1 \leq W_1'.
\]

This type of inequality among three different types of statistics has been well–known for testing linear restrictions in the multivariate regression model (Anderson (1984), Chapter 8, for instance.) If we use the same significance point (based on the asymptotic \( \chi^2 \) distribution), the Wald–type statistic tends to reject the hypothesis more often than the other statistics while the likelihood ratio statistic tends to reject the hypothesis more frequently than the \( LM \) statistics.

### 4. Tests of Predeterminedness

In this section we shall derive several tests of the null hypothesis of econometric predeterminedness \( H_{\xi, \eta} : \xi = 0, \eta = 0 \). We suppose that \( G_0 \leq G_1 \).
4.1. The likelihood ratio test.

We first find the likelihood function maximized under $H_{\xi, \eta}$. The hypothesis $H_\eta$ is

\begin{equation}
0 = (\Omega_{21}, \Omega_{22}) \begin{pmatrix} B_1 \\ -B_2 \end{pmatrix} \\
= \Omega_{21} B_1 - \Omega_{22} B_2 \\
= \Omega_{22}(\rho B_1 - B_2),
\end{equation}

where $\rho = \Omega_{22}^{-1} \Omega_{21}$ and $B'$ has been partitioned as $B' = (B_1', -B_2')$ with $B_1$ having $G_1$ rows. This fact suggests a re-parametrization since $B_2 = \rho B_1$ under $H_\eta$ and then $H_\xi$ is

\begin{equation}
0 = \Pi_{21} B_1 - \Pi_{22} B_2 = (\Pi_{21} - \Pi_{22} \rho) B_1,
\end{equation}

where $\Pi_{ij}$ denotes the $i, j$-th submatrix of $\Pi$ partitioned into $K_1$ and $K_2$ rows and $G_1$ and $G_2$ columns. Let

\begin{equation}
Y^*_1 = Y_1 - Y_2 \rho, \quad V^*_1 = V_1 - V_2 \rho,
\end{equation}

\begin{equation}
\Pi^{**} = (\Pi_{11}^{**}, \Pi_{22}^{**}) = \begin{pmatrix} \Pi_{11}^{**} & \Pi_{12}^{**} \\ \Pi_{21}^{**} & \Pi_{22}^{**} \end{pmatrix} = \begin{pmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{pmatrix} \begin{pmatrix} I & 0 \\ -\rho & I \end{pmatrix} \\
= \begin{pmatrix} \Pi_{11} - \Pi_{12} \rho & \Pi_{12} \\ \Pi_{21} - \Pi_{22} \rho & \Pi_{22} \end{pmatrix}.
\end{equation}

Then $H_\xi$ is $\Pi_{21}^{**} B_1 = 0$. The reduced form for $(Y^*_1, Y_2)$ is

\begin{equation}
(Y^*_1, Y_2) = Z \Pi^{**} + (V^*_1, V_2).
\end{equation}

The covariance matrix of each row of $(V^*_1, V_2)$ is

\begin{equation}
\begin{pmatrix} I & -\rho' \\ 0 & I \end{pmatrix} \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \begin{pmatrix} I & 0 \\ -\rho & I \end{pmatrix} = \begin{pmatrix} \Omega_{11} - \rho \Omega_{12} & 0 \\ 0 & \Omega_{22} \end{pmatrix},
\end{equation}

where $\Omega_{11:2} = \Omega_{11} - \Omega_{12} \Omega_{22}^{-1} \Omega_{21}$. The likelihood function of the parameters $\Omega_{22}, \Omega_{11:2}, \Pi^{**}$, and $\rho$ (subject to (4.2) or alternatively $\Pi_{21}^{**} B_1 = 0$) is

\begin{equation}
L_5 = c_1 |\Omega_{11:2}|^{-T/2} \exp \left[ -\frac{1}{2} \text{tr} \Omega_{11:2}^{-1} (Y^*_1 - Z \Pi^*_1)^t (Y^*_1 - Z \Pi^*_1) \right] \\
\times |\Omega_{22}|^{-T/2} \exp \left[ -\frac{1}{2} \text{tr} \Omega_{22}^{-1} (Y_2 - Z \Pi_{2})^t (Y_2 - Z \Pi_{2}) \right].
\end{equation}
The function $L_5$ is maximized with respect to $\Pi_2, \Omega_{22}$, and $\Omega_{11.2}$ at

(4.8) \[ \hat{\Pi}_2 = (Z'Z)^{-1}Z'Y_2, \quad T\hat{\Omega}_{22} = Y_2'\overline{P}_Z Y_2, \]

(4.9) \[
T\hat{\Omega}_{11.2} = (Y_1^* - Z\Pi_{11}^{**})'(Y_1^* - Z\Pi_{11}^{**}) \\
= \left[ Y_1 - (Y_2, Z_1) \left( \begin{array}{c} \rho \\ \Pi_{11}^{**} \end{array} \right) - Z_2\Pi_{21}^{**} \right]' \left[ Y_1 - (Y_2, Z_1) \left( \begin{array}{c} \rho \\ \Pi_{11}^{**} \end{array} \right) - Z_2\Pi_{21}^{**} \right],
\]

and the concentrated likelihood is

(4.10) \[ L_6 = c_2 |T\hat{\Omega}_{22}|^{-T/2} |T\hat{\Omega}_{11.2}|^{-T/2}. \]

Since $T\hat{\Omega}_{22}$ is a sample quantity, we want to minimize $|T\hat{\Omega}_{11.2}|$ with respect to $\rho, \Pi_{11}^{**},$ and $\Pi_{21}^{**}$ subject to $\Pi_{21}^{**} B_1 = 0.$

The general alternative to $H_{\xi, \eta}$ is that $\Pi$ and $\Omega$ are unrestricted; we term this as $H_A$. The problem of testing $H_{\xi, \eta}$ vs $H_A$ has been reduced to testing $\Pi_{21}^{**} B_1 = 0$ vs $\Pi_{21}^{**} B_1 \neq 0.$ The likelihood ratio criterion by the algebra of Section 3.1 is

(4.11) \[ LR_2 = T \sum_{i=1}^{G_0} \log(1 + \lambda_i^*), \]

where $\lambda_{G_1}^* \geq \cdots \geq \lambda_1^* \geq 0$ are the roots of

(4.12) \[ |Y_1'(P_{Y_2, Z} - P_{Y_2, Z_1})Y_1 - \lambda^* Y_1'\overline{P}_{Y_2, Z} Y_1| = 0. \]

Another possible alternative to $H_{\xi, \eta}$ may be $H_\xi$, which defines the structural equations with the block identifiability restrictions. Because $H_{\xi, \eta}$ is nested within $H_\xi$, the log-likelihood ratio criterion for $H_{\xi, \eta}$ vs $H_\xi$ is the difference between the statistic for $H_{\xi, \eta}$ vs $H_A$ and the statistic for $H_\xi$ vs $H_A$, namely

(4.13) \[
LR_3 = T \sum_{i=1}^{G_0} \log(1 + \lambda_i^*) - T \sum_{i=1}^{G_0} \log(1 + \lambda_i) \\
= T \sum_{i=1}^{G_0} \log \frac{1 + \lambda_i^*}{1 + \lambda_i}.
\]

16
When \( G_0 = 1 \) and \( G_1 \geq 1, LR_3 \) reduces to the statistic obtained by Hwang (1980a). Furthermore, \( LR_3 \) reduces to the statistic obtained by Kariya and Hodoshima (1980) when \( G_0 = G_1 = 1 \).

4.2. Lagrange multiplier test.

The Lagrange multiplier statistic for testing \( H_{\xi, \eta} : \xi = 0, \eta = 0 \) vs. \( H_A \) is the Lagrange multiplier statistic for testing \( H_\zeta : \zeta = \Pi_{21}^* B_1 = 0 \) vs. \( H_A \). It is

\[
LM_2 = T \sum_{i=1}^{G_0} \frac{\lambda_i^*}{1 + \lambda_i^*}.
\]

where \( \lambda_1^*, \ldots, \lambda_{G_0}^* \) are the \( G_0 \) smallest roots of (4.12). This statistic \( LM_2 \) does not seem to have been derived previously.

Now consider testing \( H_{\xi, \eta} \) vs. \( H_\xi \). Let \( \Lambda \) and \( \Lambda_0 \) be \( K_2 \times G_0 \) and \( G_2 \times G_0 \) matrices of Lagrange multiplier parameters for \( H_{\xi, \eta} : \xi = 0, \eta = 0 \), respectively. The Lagrange form in this case is written as

\[
(4.15) \log L_7 = \log c_1 - \frac{T}{2} \log |\Omega| - \frac{1}{2} \text{tr} \Omega^{-1}(Y - Z\Pi)'(Y - Z\Pi) + \text{tr} \Lambda'(\Pi_{21}, \Pi_{22})B \\
+ \text{tr} \Lambda_0'(\rho, J_{G_2})B.
\]

Setting to 0 the derivative of \( \log L_7 \) with respect to the components of \( B_2 \), we have

\[
(4.16) \quad \Pi_{22}^* \Lambda + \Lambda_0 = 0.
\]

Substituting this relation into \( \log L_7 \) and ignoring a constant term we obtain

\[
(4.17) \quad \log L_8 = -\frac{T}{2} \log |\Omega_{11,22}| ||\Omega_{22}|| - \frac{1}{2} \text{tr} \Omega^{-1}(Y - Z\Pi)'(Y - Z\Pi) \\
+ \text{tr} \Lambda'(\Pi_{21} - \Pi_{22}\rho)B_1,
\]

and \( H_\xi \) is \( (\Pi_{21} - \Pi_{22}\rho)B_1 = 0 \) under \( H_\eta \).

The derivatives with respect to the elements of \( \Pi \) are

\[
(4.18) \quad Z'Y\Omega^{-1} - Z'Z\Pi\Omega^{-1} + \begin{pmatrix} 0 & 0 \\ \Lambda B_1' & -\Lambda B_1'\rho' \end{pmatrix}.
\]
Setting this matrix to 0 yields

\[(4.19) \quad Z'Z\Pi = Z'Y + \begin{pmatrix} 0 & 0 \\ \Lambda B'_1\Omega_{11.2} & 0 \end{pmatrix}.\]

We can write

\[(4.20) \quad \Omega^{-1} = \begin{pmatrix} \Omega_{11.2}^{-1} & -\Omega_{11.2}^{-1}\rho' \\ -\rho\Omega_{11.2}^{-1} & \rho\Omega_{11.2}^{-1}\rho' + \Omega_{22}^{-1} \end{pmatrix}.\]

The derivatives of log \( L_9 \) with respect to the elements of \( \rho \) are

\[(4.21) \quad (Y_2 - Z_1\Pi_{12} - Z_2\Pi_{22})'(Y - Z\Pi)\begin{pmatrix} I_{G_1} \\ -\rho \end{pmatrix}\Omega_{11.2}^{-1} - \Pi_{22}'\Lambda B'_1.\]

When (4.21) is set to 0 and multiplied by \( \Omega_{11.2} \) we obtain

\[(4.22) \quad Y'_2\overline{P}_ZY_1 - Y'_2\overline{P}_ZY_2\rho = \Pi_{22}'\Lambda B'_1\Omega_{11.2}\]

by use of (4.19). From the first set of columns of (4.19) we obtain

\[(4.23) \quad \Lambda B'_1\Omega_{11.2} = Z'_2\overline{P}_Z[Z_2\Pi_{21} - Y_1].\]

When we use this in (4.22), we obtain

\[(4.24) \quad Y'_2\overline{P}_ZY\begin{pmatrix} I_{G_1} \\ -\rho \end{pmatrix} = \Pi_{22}'Z'_2\overline{P}_Z[Z_2\Pi_{21} - Y_1].\]

Multiply (4.24) on the right by \( B_1 \) and replace \( \Pi_{21}B_1 \) by \( \Pi_{22}\rho B_1 \). Then solve (4.24) multiplied by \( B_1 \) with respect to \( \rho B_1 \) using (4.19) to obtain

\[(4.25) \quad \rho B_1 = (Y'_2\overline{P}_ZY_2)^{-1}Y'_2\overline{P}_ZY_1B_1.\]

Lemma A.6 has been used. Multiply (4.23) on the right by \( B_1 \) using (4.25) to obtain

\[(4.26) \quad \lambda\Sigma = Z'_2\overline{P}_Z(Z_2\Pi_{22}\rho - Y_1)B_1 = -Z'_2\overline{P}_{Y_2,Z_2}Y_1B_1,\]

where \( \Sigma = B'_1\Omega_{11.2}B_1 \). In this derivation we have used Lemma A.6 for \( (Y_2, Z_1) \).
From (4.26) we find that the Lagrange multiplier matrix in the sample for $H_\eta : \eta = 0$ is

\begin{equation}
(4.27) \quad \Lambda_0 = -\Pi_2' \Lambda \\
= -Y_2' \bar{P}_{Z, Z_2} (Z_2' \bar{P}_Z, Z_2)^{-1} Z_2' \bar{P}_{Y_2, Y_1} \hat{B}_1 \hat{\Sigma}^{-1} \\
= -Y_2' (\bar{P}_{Z, Z_1} - \bar{P}_Z) \bar{P}_{Y_2, Y_1} \hat{B}_1 \hat{\Sigma}^{-1} \\
= -Y_2' \bar{P}_Z \bar{P}_{Y_2, Z_1} Y_1 \hat{B}_1 \hat{\Sigma}^{-1},
\end{equation}

where $\hat{\Sigma}$ and $\hat{B}_1$ are the maximum likelihood estimators of $\Sigma$ and $B_1$. Since $\hat{\Sigma}$ is a consistent estimator of $\Sigma$ and $(1/T) \Lambda_0 \xrightarrow{p} 0$, we consider the quantity

\begin{equation}
(4.28) \quad \Lambda_0^* = Y_2' \bar{P}_Z \bar{P}_{Y_2, Z_1} Y_1 \hat{B}_1 \Sigma^{-1} \\
= Y_2' \bar{P}_Z \bar{P}_{Y_2, Z_1} Y_1 B_1 \Sigma^{-1} + Y_2' \bar{P}_Z \bar{P}_{Y_2, Z_1} Y_1 (\hat{B}_1 - B_1) \Sigma^{-1}
\end{equation}

which is asymptotically equivalent to $\Lambda_0$. We note that $\bar{P}_{Y_2, Z_1} Y_1 B_1 = \bar{P}_{Y_2, Z_1} U$ under $H_\eta : \eta = 0$. Now we apply the method to derive the asymptotic distribution in Lemma 3 in Anderson and Kunitomo (1989b) and substitute $(Y_2, Z)$ and $(Y_2, Z_1)$ for $Z$ and $Z_1$, respectively. Let

\begin{equation}
(4.29) \quad R^* = [(Z \Pi_1 + V_2 \rho) J_1, Y_2, Z_1] \\
= \left[ (Y_2, Z) \left( \begin{array}{c} \rho \\ \Pi_1 \end{array} \right) \right] J_1, Y_2, Z_1 \\
= RF,
\end{equation}

where $R = (Y_2, Z)$, $J_1' = (0, I_{G_1 - G_0})$ is a $(G_1 - G_0) \times G_1$ matrix and $F$ is a $(G_2 + K) \times (G_1 + K_1)$ matrix defined by

\begin{equation}
(4.30) \quad F = \left[ \left( \begin{array}{c} \rho \\ \Pi_1 \end{array} \right) \left( \begin{array}{c} 0 \\ I_{G_1 - G_0} \end{array} \right) \right].
\end{equation}

Let normalize $B_0 = I_{G_0}$ and partition $Y_1 = (Y_0, Y_{11})$ into $T \times (G_0 + (G_1 - G_0))$ submatrices. Then from (4.29) we write

\begin{equation}
(4.31) \quad (Y_{11}, Y_2, Z_1) = R^* + (V_1^* J_1, 0).
\end{equation}
Under the hypothesis $H_0 : \eta = 0$ we have

\[(4.32)\]

\[U = V_1 B_1 - V_2 B_2 = (V_1 - V_2 \rho)B_1 = V_1^* B_1.\]

Since each row of $R^*$ and $R$ in (4.29) is uncorrelated with each row of $V_1^* = V_1 - V_2 \rho$, $R^*$ is uncorrelated with $U$. Then by using the same argument as in the proof of Theorem 5 in Anderson and Kunitomo (1989b), the asymptotic distribution of $\text{vec}(\Lambda_0^*)$ is equivalent to the asymptotic distribution of

\[(4.33)\]

\[\text{vec}(Y_2' P_Z \overline{P}_{RF} U \Sigma^{-1}).\]

We write (4.33) as

\[(4.34)\]

\[\text{vec}(Y_2' P_Z \overline{P}_{RF} U \Sigma^{-1}) = (\Sigma^{-1} \otimes Y_2' P_Z \overline{P}_{RF}) \text{vec}(U).\]

Then conditional on $Y_2$ the covariance matrix of $\text{vec}(\Lambda_0^*)$ is

\[(4.35)\]

\[\Sigma^{-1} \otimes Y_2' P_Z \overline{P}_{RF}(\Sigma \otimes I_T)(\Sigma^{-1} \otimes \overline{P}_{RF} P_Z Y_2) = \Sigma^{-1} \otimes Y_2' P_Z \overline{P}_{RF} P_Z Y_2.\]

We now define an $LM$ statistic by

\[(4.36)\]

\[LM_3 = (\text{vec} \Lambda_0)' (\hat{\Sigma}^{-1} \otimes Y_2' P_Z \overline{P}_{RF} P_Z Y_2)^{-1} (\text{vec} \Lambda_0).\]

Then by the use of Lemma A.5, we rewrite (4.36) as,

\[(4.37)\]

\[LM_3 = \text{tr}\{ \hat{B}_1 Y_1' \overline{P}_{Y_2, Z_1} \overline{P}_Z Y_2 (Y_2' \overline{P}_Z \overline{P}_{RF} \overline{P}_Z Y_2)^{-1} Y_2' \overline{P}_Z \overline{P}_{Y_2, Z_1} Y_1 \hat{B}_1 \hat{\Sigma}^{-1} \},\]

where we have used the relation $\overline{P}_{Y_2, Z_1} \overline{P}_Z Y_2 = -\overline{P}_{Y_2, Z_1} P_Z Y_2$ and $F$ is evaluated at its maximum likelihood estimator.

When $G_1 = G_0$, we have $\overline{P}_{RF} = \overline{P}_{Y_2, Z_1}$. Using Lemma A.6, we obtain the expression

\[(4.38)\]

\[LM_3 = \text{tr}\{ Y_1' (\overline{P}_{Y_2, Z_1} - \overline{P}_X) Y_1 \hat{\Sigma}^{-1} \} = T \sum_{i=1}^{G_0} \frac{\lambda_i^{**}}{1 + \lambda_i^{**}}.\]
where \( \lambda_i^{**} \) are the characteristic roots of

\[
(4.39) \quad |Y_1'(P_X - P_{Y_2,Z_1})Y_1 - \lambda^{**}Y_1'\overline{P}_X Y_1| = 0,
\]

and \( \overline{X} = (Y_2, Z_1, \overline{P}_2 Y_2) \) is a \( T \times (G_2 + K_1 + G_2) \) matrix. The second line of (4.38) implies that \( \hat{\Sigma} = (1/T)\hat{B}_1 Y_1'\overline{P}_X Y_1 \hat{B}_1 \). In the present formulation of the LM test, \( \hat{\Sigma} \) should be based on the maximum likelihood estimator of \( \Sigma \) under the null hypothesis:

\[
(4.40) \quad \hat{\Sigma} = \hat{\Omega}_{11:2} = \frac{1}{T}Y_1'\overline{P}_{Y_2,Z_1} Y_1.
\]

However, in practice, several estimators of \( \Sigma \) could be used. For instance, instead of (4.40), we may use

\[
(4.41) \quad \hat{\Sigma} = \frac{1}{T - 2G_2 - K_1}Y_1'\overline{P}_X Y_1.
\]

In particular, \( \text{LM}_3 \) with (4.41) reduces to the statistic proposed by Wu (1973) and Wu (1974) when \( G_0 = G_1 = 1 \).

On the other hand, consider a testing problem for

\[
(4.42) \quad Y_1B_1 = Y_2B_2 + Z_1\Gamma + E_3B_3 + U,
\]

where \( B_3 \) is a \( G_2 \times G_0 \) vector of unknown parameters, and \( E_3 \) is the least squares residuals \( E_3 = Y_2 - \bar{Y}_2 = \overline{P}_2 Y_2 \). Hausman (1978) proposed the usual \( F \) test for \( H_0 : B_3 = 0 \) against \( H_1 : B_3 \neq 0 \) as a specification test when \( G_0 = G_1 = 1 \) and \( B_0 = 1 \). From (4.37) it is clear that \( \text{LM}_3 \) is proportional to Hausman's statistic in this case. In fact, Nakamura and Nakamura (1980) has shown this equivalence between Wu's test and Hausman's test for \( G_1 = 1 \). They also pointed out that a statistic proposed by Durbin (1954) is similar to them. Our derivation of statistics shows that these statistics can be interpreted as LM test procedures. Hwang (1985) also has shown the equivalence of Hausman's test and an LM test for \( G_0 = 1 \) and \( G_1 \geq 1 \) by a different method.

Another possibility of an estimator of \( \Sigma \) may be

\[
(4.43) \quad \hat{\Sigma} = \frac{1}{T - K - G_2}Y_1'\overline{P}_{Y_2,Z} Y_1.
\]
because it is an unrestricted sum of squares from the regression residuals. Then, the statistic $LM_3$ with (4.43) reduces to the one proposed by Revankar (1978) when $G_0 = G_1 = 1$. Therefore, we can also reinterpret Revankar’s test as an LM test procedure.

4.3. Wald test.

Now we consider Wald-type statistics for the present testing problems. For this purpose we first consider the null hypothesis $H_{\xi, \eta} : \xi = 0, \eta = 0$ vs the alternative hypothesis $H_A : \xi \neq 0$. In this case, our derivation of a Wald test is similar to Section 3.3. Thus a Wald-type statistic is

$$W_2 = T \sum_{i=1}^{G_0} \lambda_i^*,$$

(4.44)

where $\lambda_i^*$ are the characteristic roots of (4.44).

When $G_0 = G_1 = 1$, $W_2$ reduces to the statistic proposed by Revankar and Hartley (1973). Although their derivation was different from ours, we can interpret their statistic as a Wald test for $H_{\xi, \eta}$ against $H_A$. $W_2$ may be called the generalized Revankar-Hartley test.

We now derive a Wald-type statistic for the null hypothesis $H_{\xi, \eta} : \xi = 0, \eta = 0$ against the alternative $H_{\xi} : \xi = 0$. We note that from (3.18) under $H_{\xi}$

$$T \hat{\Omega} = (Y - Z\hat{\Pi})(Y - Z\hat{\Pi})$$

$$= Y'\hat{P}_Z Y + \hat{\Omega}\hat{B} \left( \begin{array}{c} 0 \\ \Lambda \end{array} \right)' (Z'Z)^{-1} \left( \begin{array}{c} 0 \\ \Lambda \end{array} \right) \hat{B}'\hat{\Omega}. \quad (4.45)$$

Using (3.21), we have

$$T \hat{\Omega}\hat{B} = Y'\hat{P}_Z Y \hat{B} + \hat{\Omega}\hat{B} \hat{C}_0^{-1} \hat{B}'(P_Z - P_{Z_1}) Y \hat{B}. \quad (4.46)$$

Because $T \hat{\Sigma}_H = \hat{B}_H Y'\hat{P}_{Z_1} Y \hat{B}_H$, where $\hat{B}_H = C \hat{C}_0^{-1}$, we obtain an unrestricted estimator of $\eta$ as

$$\hat{\eta} = J_2' \hat{\Omega}_H \hat{B}_H \quad (4.47)$$

$$= J_2' \frac{1}{T} Y'\hat{P}_Z Y \hat{B}_H C_0 (I + \hat{\Lambda}) \hat{C}_0^{-1}$$

$$= J_2' \frac{1}{T} Y'\hat{P}_Z Y \hat{B} + J_2' \frac{1}{T} Y'\hat{P}_Z Y \hat{C}_0 \hat{\Lambda} \hat{C}_0^{-1},$$
where $J'_2 = (0, I_{G_2})$ and $\hat{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_{G_0})$. Since $\lambda_i = o_p(1/T^{1-\varepsilon})$ for any $\varepsilon > 0$, $\sqrt{T}C_0\hat{\Lambda}C_0^{-1} \xrightarrow{P} 0$. (See Anderson and Kunitomo (1989b).) Then the limiting distribution of $\sqrt{T}(\hat{\eta} - \eta)$ is the same as the limiting distribution of

$$
\text{vec}(J_2'\Omega\sqrt{T}(\hat{B} - B)) = \text{vec}[(\Omega_{21}, \Omega_{22})J_*, 0]\sqrt{T}
\begin{bmatrix}
-(\hat{B}_* - B_*) \\
-(\hat{\Gamma} - \Gamma)
\end{bmatrix}
$$

is the limiting distribution of

$$
-I_{G_0} \otimes [(\Omega_{21}, \Omega_{22})J_*, 0](D'MD)^{-1}D'\frac{1}{\sqrt{T}}\text{vec}(Z'U),
$$

where $J'_* = (0, I_{G_*})$ and

$$
D = \begin{bmatrix}
\Pi_*, & (I_{K_1} \\
0
\end{bmatrix}
$$

is a $K \times (G_* + K_1)$ matrix.

The limiting distribution of $\sqrt{T}\text{vec}[(1/T)Y'\hat{P}_2Y - \Omega]$ is $N(0, \Omega \otimes \Omega)$. Hence the limiting distribution of

$$
\sqrt{T}\text{vec}
\begin{bmatrix}
J'_2 \left(\frac{1}{T}Y'\hat{P}_2Y - \Omega\right)B
\end{bmatrix}
$$

is $N(0, \Sigma \otimes \Omega_{22})$. From (4.50) and (4.52), the asymptotic covariance matrix of $\text{vec}\eta^*$ is given by

$$
\Sigma \otimes (\Omega_{22}(\rho J_1, I_{G_2}, 0)(D'MD)^{-1}(\rho J_1, I_{G_2}, 0)'\Omega_{22} + \Omega_{22}),
$$

where $J'_1 = (0, I_{G_1-G_0})$. Hence we define a Wald statistic by

$$
W_3 = T(\text{vec}\hat{\eta})'\{\hat{\Sigma} \otimes [T\hat{\Omega}_{22}(\hat{\rho} J_1, I_{G_2}, 0)(\hat{D}'Z'Z\hat{D})^{-1}(\hat{\rho} J_1, I_{G_2}, 0)'\hat{\Omega}_{22} + \hat{\Omega}_{22}]\}^{-1}(\text{vec}\hat{\eta})
$$

where $\hat{\Sigma}$, $\hat{\rho}$, and $\hat{\Omega}_{22}$ are the maximum likelihood estimators of $\Sigma$, $\rho$, and $\Omega_{22}$, and $\hat{D}$ is the maximum likelihood estimator of $D$, respectively. Again, using Lemma A.5, we have
\[(4.55) \quad W_3 = \text{tr}\{\hat{\mathcal{B}}'Y'\mathcal{P}_Z Y_2 [T\hat{\Omega}_{22}(\hat{\rho}J_1, I_{G_2}, 0)(\hat{D}'Z'Z\hat{D})^{-1}(\hat{\rho}J_1, I_{G_2}, 0)'T\hat{\Omega}_{22}
\quad + T\hat{\Omega}_{22}]^{-1}Y_2'\mathcal{P}_Z Y_{\hat{\mathcal{B}}}\Sigma^{-1},\]

where $\hat{\mathcal{B}}$ is the maximum likelihood estimator of $B$ under $H_\xi$.

This Wald-type statistic is similar to the Wald statistic proposed by Smith (1985) when $G_0 = 1$. Although it is complicated in general, it can be further simplified in the case when the subsystem of structural equations is just-identified as the alternative hypothesis. In this case, since $\lambda_i = 0$, $i = 1, \ldots, G_0$, in (3.13) we have $T\hat{\Omega}_{22} = Y_2'\mathcal{P}_Z Y_2$, $\hat{\mathcal{L}}_2 = (Z'Z)^{-1}Z'Y_2$, $T\hat{\Sigma} = \hat{\mathcal{B}}'Y'\mathcal{P}_Z Y_{\hat{\mathcal{B}}}$. Then, in particular, when $G_0 = G_1 = 1$, it can be shown that $W_3$ in (4.55) is equivalent to the statistic proposed by Wu (1973) and Wu (1974) except $\hat{\Sigma}$. This may give the Wu test procedure another new interpretation.

4.3. An inequality among statistics.

We have derived three types of statistics for the predeterminedness restriction in a subsystem of structural equations. There is a simple inequality among the statistics we have derived for $H_{\xi,\eta}: \xi = 0, \eta = 0$ vs $H_A : \xi \neq 0$. Using Lemma A.7, we have

\[(4.56) \quad 0 \leq LM_2 \leq LR_2 \leq W_2.\]

This inequality is an analogue to (3.50) for the problem of testing of the block identifying restriction in Section 3. However, a similar inequality can not be obtainable for the testing problem of $H_{\xi,\eta}: \xi = 0, \eta = 0$ vs $H_\xi : \xi = 0$.

5. Conclusion

In this paper we have derived systematically a number of procedures for testing the block identifiability condition and the predeterminedness condition in a subsystem of structural equations. We generalized the test statistics proposed previously and derived the LR test, LM test, and Wald test for these two problems. This formulation enables us to give new interpretations to a number of testing procedures. We explored the relationship
between test statistics in econometrics and those in multivariate statistical analysis and obtained some new interpretations for some test statistics commonly known in multivariate statistical analysis.

Among three types of test statistics discussed in this paper, the LR test procedures have often turned out to be considerably simpler than the other two procedures. It is especially evident for the econometric exogeneity hypothesis when $G_0 < G_1$. This finding may be important for practical implementation of the testing procedure in empirical studies.
Appendix A

In this appendix we present some useful lemmas. Most of these lemmas are known in multivariate statistical analysis and their proofs can be found in the works of Anderson (1984) or Rao (1973). We shall present only the proof of Lemma A.2, which may be new in econometrics.

Lemma A.1: Let $D$ and $G$ be $p \times p$ positive definite matrices. Then the function

$$f(G) = -N \log |G| - \text{tr}(G^{-1} D)$$

is maximized at $G = (1/N)D$.

Lemma A.2: Let a $p \times p$ positive definite matrix $A$ be decomposed into $(p_1 + p_2) \times (p_1 + p_2)$ submatrices $A = (A_{ij})$. For any $q \times p_1$ matrix $B$ and $q \times p_2$ matrix $C$,

$$\min_C \left| A + \begin{pmatrix} B' \\ C' \end{pmatrix} (B, C) \right| = \frac{|A|}{|A_{11}|} |A_{11} + B'B|$$

$$= |A_{11} + B'B| \left| A_{22} - A_{21} A_{11}^{-1} A_{12} \right|$$

and the minimum occurs at $C = BA_{11}^{-1} A_{12}$.

Proof: Let $D = (B, C)$. Then

$$|A + D'D| = \begin{vmatrix} A & -D' \\ D & I_q \end{vmatrix} = |A| |I_q + DA^{-1} D'|.$$

Let also the inverse matrix $A$ be decomposed into $(p_1 + p_2) \times (p_1 + p_2)$ submatrices $A^{-1} = (A_{ij})$. Then

$$DA^{-1} D' = (C + BA^{12}(A^{22})^{-1})A^{22}(C + BA^{12}(A^{22})^{-1})' + B(A^{11} - A^{12}(A^{22})^{-1} A^{21})B'$$

$$\geq B(A^{11} - A^{12}(A^{22})^{-1} A^{21})B' = BA_{11}^{-1} B'.$$

Hence,

$$|A + D'D| \geq |A| |I_q + BA_{11}^{-1} B'|.$$

Finally, we obtain (A.2) by using (A.3).
Lemma A.3: Let $A$ be a $p \times p$ positive semidefinite matrix and $0 \leq \lambda_1 \leq \cdots \leq \lambda_p$ be its characteristic roots. Let $B$ be a $p \times q$ ($p > q$) matrix. Then

\begin{equation}
\min_{B' B = I} |B' AB| = \prod_{i=1}^{q} \lambda_i.
\end{equation}

**Proof.** If $A$ is singular, the left-hand and right-hand sides of (A.5) are 0. For $A$ positive definite, we use

\begin{equation}
ch_i(UV) \geq ch_j(U)ch_k(V)
\end{equation}

for $j + k \leq i + 1$, $U$ positive definite, and $V$ positive semidefinite. Here $ch_i(W)$ is the $i$-th smallest characteristic root of $W$. [See, for example, Theorem 2.2 of Anderson and Das Gupta (1963).] For any $B$ such that $B' B = I_q$

\begin{equation}
ch_{p-q+i}(BB') = ch_i(B'B) = ch_i(I_q) = 1, \quad i = 1, \ldots, q.
\end{equation}

In (A.6) let $U = A$, $V = BB'$, $i = p - q + j$ and $k = p - q + 1$ to obtain

\begin{equation}
ch_j(B'AB) = ch_{p-q+j}(BB') \geq ch_j(A)ch_{p-q+1}(BB') = ch_j(A), \quad j = 1, \ldots, q.
\end{equation}

Then

\begin{equation}
|B' AB| = \prod_{j=1}^{q} ch_j(B'AB) \geq \prod_{j=1}^{q} ch_j(A) = \prod_{j=1}^{j} \lambda_j.
\end{equation}

Equality is obtained when the columns of $B$ are the characteristic vectors of $A$ corresponding to the roots $\lambda_1, \ldots, \lambda_q$. \hfill \Box

Lemma A.4:

\begin{equation}
\frac{\partial \text{tr}(AB)}{\partial B} = A', \quad \frac{\partial \text{tr}(B'ABC)}{\partial B} = ABC + A'BC'.
\end{equation}

Lemma A.5: For any $m \times n$ matrix $A = (a_1, \ldots, a_n)$, we define an $mn \times 1$ vector $\text{vec} \ A = (a_1', \ldots, a_n')'$. Then for any conformable matrices,

\begin{equation}
\text{vec} \ (BXC) = (C' \otimes B)(\text{vec} \ X),
\end{equation}

27
(A.12) \[ \text{tr} (BCD) = (\text{vec } B')' (I \otimes C)' (\text{vec } D), \]

(A.13) \[ \text{tr} (BX'CXD) = (\text{vec } X)' (DB \otimes C') (\text{vec } X), \]

where each \( K \) is a commutation matrix defined by \( \text{vec } (C) = K \text{vec } (C') \) for an arbitrary matrix \( C \) of suitable order.

**Lemma A.6:**

(A.14) \[ \overline{P}_{B,C} = \overline{P}_B - \overline{P}_B C (C' \overline{P}_B C)^{-1} C' \overline{P}_B, \]

where \( D^{-1} \) stands for the generalized inverse matrix of any matrix \( D \).

**Lemma A.7:** For non-negative \( \lambda_i, i = 1, \ldots, p, \)

(A.15) \[ \sum_{i=1}^{p} \frac{\lambda_i}{\lambda_i + \lambda_i} \leq \log \prod_{i=1}^{p} (1 + \lambda_i) \leq \sum_{i=1}^{p} \lambda_i. \]
Appendix B

Maximum Likelihood Estimators

Maximum likelihood estimators of $\Pi$, $B$, and $\Omega$ under $H_{\xi}$ as well as the likelihood ratio test of $H_{\xi}$ were developed by Anderson (1951). This appendix summarizes the results needed in the present paper. The exposition will refer to the table of correspondence between the notation of the present paper and that of Anderson (1951) at the end of this appendix. The reduced form model is specified in B.1 of the table and the basic statistics in B.2.

A matrix $B$ satisfying $\Pi_2 B = 0$ can be multiplied on the right by an arbitrary nonsingular matrix $F$ to obtain $BF$, which also satisfies the condition, $\Pi_2 (BF) = 0$. The maximum likelihood estimators of $B$ similarly can be transformed by multiplication on the right by an arbitrary nonsingular matrix. Any such estimator is composed of linear combinations of the $G_0$ characteristic vectors of

\[(Y'\hat{P}_Z Y)^{-1}Y'(P_Z - P_{Z_1})Y\]

corresponding to the $G_0$ smallest characteristic roots. With the normalization defined in B.3 of the table the matrix is $C$. When $B$ is normalized so

\[B = \begin{pmatrix} I \\ -B_* \end{pmatrix},\]

the maximum likelihood estimator is

\[\hat{B}_H = CC_0^{-1} = \begin{pmatrix} I \\ -C_* C_0^{-1} \end{pmatrix}.\]

In Anderson (1951) the likelihood was maximized under the condition that $\Gamma' \Sigma \Gamma = I$, but it was shown that the maximum of the likelihood function was independent of the normalization. The estimator of $\Gamma$ was normalized by

\[\frac{1}{N} \hat{\Gamma}' \hat{A} \hat{\Gamma} = \hat{\Gamma}' H \hat{\Gamma} = (I + \Phi^*)^{-1}.\]

Then

\[\hat{\Gamma} = C(I + \Phi^*)^{-\frac{1}{2}},\]

where $\Phi^*$ is the diagonal matrix composed of the smallest characteristic roots of (B.1). The maximum likelihood estimators of $B$, $\Omega$, $\Sigma$, and $\Pi_2$ are given in B.4.
Table. Correspondence of Notations

<table>
<thead>
<tr>
<th>This paper</th>
<th>Anderson (1951)</th>
</tr>
</thead>
</table>

**B.1. Model**

\[ Y = Z \Pi + V \]
\[ T \times G T \times K K \times G T \times G \]
\[ \varepsilon v_i v_i' = \Omega \]
\[ Z = (Z_1, Z_2), \Pi = \begin{pmatrix} \Pi_1 \\Pi_2 \end{pmatrix} \begin{pmatrix} K_1 \\ K_2 \end{pmatrix} \]
\[ \Pi_2. B = 0 \]
\[ K \times G G \times G_0 K \times G_0 \]
\[ K, G, G_0, T \]
\[ \varepsilon X = \bar{B} Z \]
\[ p \times N p \times q q \times N \]
\[ \Sigma \]
\[ \bar{B} = (\bar{B}_1, \bar{B}_2), Z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} q_1 q_2 \]
\[ \Gamma' \bar{B}_2 = 0 \]
\[ m \times p p \times q m \times q \]
\[ q, p, m, N \]

**B.2. Descriptive Statistics**

\[ (Z'Z)^{-1} Z'Y \]
\[ S = \bar{V}' \bar{V} \]
\[ = Y' \bar{P}_Z Y \]
\[ B = XZ'(ZZ')^{-1} \]
\[ A = NH = (N - q)S \]
\[ = (X - BZ)(X - BZ)' \]
\[ = X \bar{P}_Z X' \]
\[ A_{22.1} = Z_2'Z_2 - Z_1'Z_1(Z_1'Z_1)^{-1}Z_1'Z_2 \]
\[ = Z_2' \bar{P}_{Z_1} Z_2 \]
\[ Q = Z_2Z_2' - Z_2Z_1'(Z_1Z_1')^{-1}Z_1Z_2' \]
\[ = Z_2 \bar{P}_{Z_1} Z_2' \]
\[ Y'(P_Z - P_{Z_1})Y \]
\[ B_2'QB_2 \]

**B.3. Determinantal Roots and Associated Matrices**

\[ |Y'(P_Z - P_{Z_1})Y - \lambda Y' \bar{P}_Z Y| = 0 \]
\[ |B_2QB_2' - \phi A| = 0 \]
\[ \lambda_G \geq \lambda_{G-1} \geq \cdots \geq \lambda_1 > 0 \]
\[ \phi_1 \geq \phi_2 \geq \cdots \geq \phi_p > 0 \]
\[ [Y'(P_Z - P_{Z_1})Y - \lambda_i Y' \bar{P}_Z Y]c = 0 \]
\[ \frac{1}{c} c' Y' \bar{P}_Z Y c = 1 \]
\[ (B_2QB_2' - \phi_i A)c = 0 \]
\[ \frac{1}{c} c' Ac = c' H c = 1 \]
\[ c_i \]
Table. Correspondence of Notations—continued

<table>
<thead>
<tr>
<th>This paper</th>
<th>Anderson (1951)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C = (c_1, \ldots, c_{G_0}) = \begin{pmatrix} C_0 \ -C_+ \end{pmatrix} G_0$</td>
<td>$C = (c_{p-m+1}, \ldots, c_p)$</td>
</tr>
<tr>
<td>$\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_{G_0})$</td>
<td>$\Phi^* = \text{diag}(\phi_{p-m+1}, \ldots, \phi_p)$</td>
</tr>
<tr>
<td>$C^t \frac{1}{T} Y' P Z Y C = I$</td>
<td>$C^t \frac{1}{N} A C = C' H C = I$</td>
</tr>
<tr>
<td>$Y'(P Z - P Z_1) Y C = Y' \tilde{P} Z Y \Lambda$</td>
<td>$B_2 Q B'_2 C = A C \Phi^*$</td>
</tr>
<tr>
<td>$\frac{1}{T} C^t Y'(P Z - P Z_1) Y C = \Lambda$</td>
<td>$\frac{1}{N} C^t B_2 Q B'_2 C = \Phi^*$</td>
</tr>
</tbody>
</table>

**B.4. Estimators**

$$\hat{B}_H = C C_0^{-1} = \begin{pmatrix} I \\ -C_+ C_0^{-1} \end{pmatrix}$$

$$\hat{\Gamma} = C (I + \Phi^*)^{-\frac{1}{2}}$$

$$\tilde{\Omega} = \frac{1}{T} Y' \tilde{P} Z Y + \frac{1}{T} Y' \tilde{P} Z Y C \Lambda C^t \frac{1}{T} Y' \tilde{P} Z Y$$

$$= \frac{1}{T} Y' \tilde{P} Z Y + \frac{1}{T} Y'(P Z - P Z_1) Y C \Lambda C^t \frac{1}{T} Y' \tilde{P} Z Y$$

$$\hat{\Sigma}_H = \tilde{\Omega} \hat{B}_H \hat{B}_H$$

$$= (C_0^{-1})^{-1} (I + \Lambda) C_{C_0^{-1}}^{-1}$$

$$= \frac{1}{T} \tilde{B}_H Y' \tilde{P} Z_1 Y \tilde{B}_H$$

$$\hat{\Omega} C = \frac{1}{T} Y' \tilde{P} Z Y C + \frac{1}{T} Y'(P Z - P Z_1) Y C$$

$$= \frac{1}{T} Y' \tilde{P} Z Y C$$

$$= \frac{1}{T} Y' \tilde{P} Z Y C (I + \Lambda)$$

$$\hat{\Pi}_2 = (0, I_{K_2}) (Z' Z)^{-1} Z' Y (I - C C' \frac{1}{T} Y' \tilde{P} Z Y)$$

$$\hat{B}_2 = (I - \hat{\Sigma} \hat{\Gamma} \hat{\Gamma}') B_2$$
References


32


