88-F-2

Asymptotic Robustness of Tests in Simultaneous Equation Systems*

by

T.W. Anderson
Stanford University

and

Naoto Kunitomo
University of Tokyo

February 1988

*This paper was written while the second author was visiting the Department of Economics, Stanford University. The research was partly supported by Nomura Foundation Grant and Grant-in-Aid 62301071 of the Ministry of Education at the Faculty of Economics, University of Tokyo. We thank Yasunori Fujikoshi, Haruo Yanai, and Yoshihiko Tsukuda for their comments on the preliminary draft of this paper.

Abstract

We systematically derive a number of test procedures for testing the block identifiability condition and the predeterminedness condition in a subsystem of structural equations. We generalize the test statistics proposed in the past for these hypotheses and give them some new interpretation. We explore the relationship between test statistics in econometrics and those in multivariate statistical analysis. Our results give some new interpretation to a number of test statistics. We also obtain the asymptotic distributions of statistics under a set of fairly general conditions on the disturbance terms. For this purpose, we derive a new Martingale Central Limit Theorem, which is based on the Lindeberg condition for the Martingale difference sequences and hence useful for applications. Our results on the asymptotic distributions of test statistics give justification for using the test procedures commonly known even if some of the usual underlying assumptions for their derivations are not satisfied.

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 based 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 by Anderson and Rubin (1949), Hood and Koopmans (1953), Basmann (1960), Wu (1973), Byron (1972), Revanker and Hartley (1973), Revanker (1978), Hausman (1978), Kariya and Hodoshima (1980), and Hwang (1980a) among many others have drawn some 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 systematically derive several test procedures and to obtain the relationships among 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, that is, the likelihood ratio (RL) test, the 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 all test statistics mentioned above as special cases and give new interpretation to some of them. Incidentally, we shall also give some new interpretation to the test statistics commonly known in multivariate statistical analysis.

We shall also derive the asymptotic distributions of test statistics under a set of fairly general conditions on the disturbance terms. For this purpose, we shall obtain a new Martingale Central Limit Theorem based on the Lindeberg condition for the Martingale difference sequences and apply it to the present situation. We allow that there are a finite number of lagged endogenous variables and the disturbance terms are not necessarily independent. We shall show that the limiting distributions of test statistics considered in this paper are the non-central χ^2 distributions under local alternative hypotheses and are the central χ^2 distributions under the null-hypotheses when the disturbances are the Martingale difference sequences. Because test statistics have been often proposed under a set of relatively restrictive assumptions, it may be important to show that the assumptions usually made are not essential for the testing procedures in practice.

In Section 2, we formulate a subsystem of structural equations. In Section 3, we shall derive several statistics for testing identifying restrictions. We shall also relate those statistics to the statistics in multivariate statistical analysis yielding some new interpretation of statistics commonly known among statisticians. Subsequently, in Section 4, we shall derive a number of test procedures for testing econometric predeterminedness restrictions. In Section 5, we give a new Martingale Central Limit Theorem and present some general results on the asymptotic distributions of test statistics introduced in the previous two sections. Finally, in Section 6, some concluding remarks are given. Useful lemmas and some detailed proofs of theorems are given in the Appendices.

2. Two Hypotheses in a Subsystem of Structural Equations

We consider a subsystem of G_0 structural equations

(2.1) $\quad [Y_1, Y_2] B = Z_1 \Gamma + U,$

where $Y^* = (Y_1, Y_2, Y_3)$ consists of all endogenous variables in the complete system of $G_1 + G_2 + G_3$ equations, $Y = (Y_1, Y_2)$ is a $T \times G$ $(G = G_1 + G_2)$ matrix of observations on the endogenous variables appearing in the first $G_0 (G_0 \leqq G_1)$ structural equations, Z_1 is a $T \times K_1$ matrix of observations on the K_1 exogenous variables, $B' = (B_1', -B_2')$ and Γ are $G_0 \times (G_1 + G_2)$ and $K_1 \times G_0$ matrices with unknown parameters, and U is a $T \times G_0$ matrix of unobservable disturbances. When $G_0 = 1$, (2.1) is usually called a single structural equation.

We consider the reduced form equation for the endogenous variables Y appearing in the first G_0 structural equations (2.1) with K $(K = K_1 + K_2)$ exogenous variables defined by

(2.2) $\quad Y = Z\Pi + V,$

where $Z = (Z_1, Z_2) = (z_{ti})$ is a $T \times K$ matrix of exogenous variables $(T > K)$ with full rank, and Z_2 is a $T \times K_2$ matrix of excluded exogenous variables in (2.1). We note that the exogenous variables include lagged endogenous variables. The reduced form coefficient matrix $\Pi = [\ \Pi_1\ \Pi_2\ ]$ is partitioned into $(K_1 + K_2) \times (G_1 + G_2)$ submatrices

$$\Pi = \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{bmatrix}$$

and $V = (V_1, V_2)$ is a $T \times (G_1 + G_2)$ matrix of disturbances whose t-th row is denoted by v_t'. We assume that

(2.3) $\qquad E(v_t \mid F_{t-1}) = 0\ ,$

(2.4) $\qquad E(v_t\, v_t' \mid F_{t-1}) = \Omega\ ,$

where F_{t-1} is the σ-field generated by the information from exogenous as well as endogenous variables available at the period t-1 and Ω is a $G \times G$ positive definite matrix partitioned into $(G_1 + G_2)\times(G_1 + G_2)$ submatrices

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

In order to relate (2.1) and (2.2), (2.1) is postmultiplied to produce

(2.5) $\quad \Gamma = (\Pi_{11}, \Pi_{12}) B ,$

(2.6) $\quad U = VB .$

Let u_t' be the t-th row of U. From (2.4) and (2.6), we have

(2.7) $\quad E(u_t | F_{t-1}) = 0,$

(2.8) $\quad E(u_t u_t' | F_{t-1}) = \Sigma,$

where Σ is a $G_0 \times G_0$ positive definite matrix.

We also note that the block identifiability conditions are expressed as H_ξ: $\xi = 0$, where

(2.9) $\quad \xi = (\Pi_{21}, \Pi_{22}) B .$

From (2.9) we obtain the rank condition of the identifiability for (2.1),

(2.10) $\quad \mathrm{rank}(\Pi_{21}\ \Pi_{22}) = G - G_0 ,$

and columns of B are linearly independent. The order condition is given by

(2.11) $\quad L_1 = K_2 - (G - G_0) \geqq 0.$

In the above notation, L_1 is often called the degree of overidentification.

Let $\nu_G \geqq \cdots \geqq \nu_1 \geqq 0$ be the characteristic roots of

(2.12) $\quad \left| \frac{1}{T} \Theta_T - \nu \Omega \right| = 0,$

where

(2.13) $\quad \Theta_T = \begin{pmatrix} \Pi_{21}' \\ \Pi_{22}' \end{pmatrix} A_{22.1} (\Pi_{21}, \Pi_{22})$

(2.14) $\quad A_{22.1} = Z_2'Z_2 - Z_2Z_1(Z_1Z_1)^{-1}Z_1Z_2 .$

Then from (2.10), it is clear that the block identifiability conditions are equivalent to the hypothesis H_{ν}: $\nu_1 = \cdots = \nu_{G_0} = 0$, which is the well-known hypothesis for the rank test in multivariate analysis.

There is an essential difference between a system of structural equations and regression models in the multivariate analysis. In the former, we allow some correlation between the endogenous variables y_{2t}', which is the t-th row of Y_2, and the corresponding disturbance term u_t'. We consider the condition that y_{2t}' and u_t' are conditionally uncorrelated given the information set F_{t-1} available at t-1. From (2.8), the covariance matrix of y_{2t}' and u_t' is given by

$$
\begin{aligned}
(2.15) \qquad \eta &= \mathrm{Cov}\,(y_{2t}, u_t \mid F_{t-1}) \\
&= (\Omega_{21}\ \Omega_{22})\ B\ .
\end{aligned}
$$

Then we define the econometric predeterminedness restrictions considered in this paper to be the hypothesis H_{η}: $\eta = 0$. We note that two hypotheses H_{ξ} and H_{η} imply the hypothesis $H_{\xi,\eta}$: $\xi = 0$, $\eta = 0$. When we take the unconditional covariance in (2.15) and disturbance terms follow the multivariate normal distribution, the uncorrelatedness implies an independence between any subset of regressor and disturbance terms in (2.1). This may be the reason why this testing problem has been sometimes called the test of independence. The hypothesis of predeterminedness in this papar has been sometimes called weak exogeneity in econometrics. There are several different concepts of econometric exogeneity in simultaneous equation systems. Engle et. al. (1983) surveyed this issue in a systematic way.

3. Tests of Block Identifiability

In this section we assume $G_0 = G_1 \geq 1$ and

(3.1) $\det B_1 \neq 0$

without loss of generality. In order to derive test statistics, we also assume that the disturbance terms $\{ v_t \}$ are independently and normally distributed. Later in Section 5, however, we shall obtain the asymptotic distribution of statistics in a more general situation. It should be noted that our derivation is considerably simpler than the methods already known.

(3.a) Likelihood Ratio Test

Under the assumption of normal disturbances, the likelihood function for $\{y_1, \ldots, y_T\}$ is given by

(3.2) $$L_1 = c_1 |\Omega|^{-T/2} \exp\{-\frac{1}{2}\,\mathrm{tr}(Y - Z\Pi)'(Y - Z\Pi)\Omega^{-1}\} ,$$

where $c_1 = 1/(\sqrt{2\pi})^{GT}$. To maximize L_1 with respect to the covariance matrix Ω, we use Lemma A.1. The concentrated likelihood function becomes

(3.3) $$L_2 = c_2 |S|^{-T/2},$$

where $c_2 = c_1\, T^{GT/2} \exp\,(-GT/2)$ and $S = (Y - Z\Pi)'(Y - Z\Pi)$.

Let a $G \times G$ matrix

(3.4) $$\Lambda_1 = (B, J_2),$$

where $J_2' = (O, I_{G_2})$ is a $G_2 \times G$ choice matrix. Because of (3.1), Λ is nonsingular. Then the concentrated likelihood is rewritten as

(3.5) $$L_2 = c_2 |\Lambda_1|^T |S^*|^{-T/2},$$

where $S^* = (W - Z\Pi^*)'(W - Z\Pi^*)$, $W = Y\Lambda_1$, and

$$\Pi^* = \begin{pmatrix} \Gamma & \Pi_{12} \\ \xi & \Pi_{22} \end{pmatrix} .$$

Let

(3.6) $\hat{\Pi}^* = (\hat{\Pi}_1^*, \hat{\Pi}_2^*) = (Z'Z)^{-1}Z'W.$

Then

(3.7) $S^* = W'\bar{P}_Z W + [({\Gamma \atop \xi}) - \hat{\Pi}_1^*, \Pi_2^* - \hat{\Pi}_2^*]'Z'Z[({\Gamma \atop \xi}) - \hat{\Pi}_1^*, \Pi_2^* - \hat{\Pi}_2^*]$,

where $P_F = F(F'F)^{-1}F'$ denotes the projection operator onto the space spanned by the column vectors of F and $\bar{P}_F = I_T - P_F$ for any full column matrix F. We note that the maximized likelihood function under the alternative hypothesis H_A: $\xi \neq 0$ is given by

(3.8) $L_3 = c_2 |Y'\bar{P}_Z Y|^{-T/2}.$

Thus, using Lemma A.2, the likelihood ratio criterion for H_ξ vs H_A is rewritten as

$$(3.9) \qquad R_1 = \frac{\max_{H_0} L_2}{\max_{H_1} L_2}$$

$$= \left(\min_B \frac{|B'Y'\bar{P}_{Z_1} YB|}{|B'Y'\bar{P}_Z YB|} \right)^{-T/2} .$$

By the use of Lemma A.3, we define the likelihood ratio test by

$$(3.10) \qquad LR_1 = -2 \log R_1$$

$$= T \sum_{i=1}^{G_0} \log(1 + \lambda_i),$$

where $\lambda_G \geq \cdots \geq \lambda_1 \geq 0$ are the characteristic roots of

$$(3.11) \qquad | Y'(P_Z - P_{Z_1})Y - \lambda Y'\bar{P}_Z Y | = 0.$$

The above equation is a sample analogue of (2.12).

When G_0 (= G_1) = 1, the likelihood ratio statistic (3.10) has been derived by Anderson and Rubin (1949). In this case, LR_1 corresponds to the smallest characteristic root in the limited information maximum likelihood (LIML) estimation method. When G_0 $(= G_1) = 2$, LR_1 is identical to the statistic proposed by Koopmans and Hood (1953) as the non-identification test. It should be also noted that Anderson (1951b) has obtained a likelihood ratio criterion in multivariate statistical analysis by a different method, which corresponds to (3.10) in the general case.

(3.b) Lagrange Multiplier Test

We shall derive a Lagrange Multiplier (LM) test statistic for the hypothesis of (2.7). Let λ be a $K_2 \times G_1$ matrix of Lagrange multiplier parameters. The Lagrange form in this case is written as

$$(3.12)\quad \log L_4 = \log c_1 - \frac{T}{2}\log|\Omega| - \frac{1}{2}\mathrm{tr}\ \Omega^{-1}(Y - Z\Pi)'(Y - Z\Pi)$$

$$+ \ \mathrm{tr}\ \lambda'(\Pi_{21},\ \Pi_{22})\ B\ .$$

Differentiating $\log L_4$ with respect to each element of Π, we get

$$(3.13)\quad Z'(Y - Z\Pi)\Omega^{-1} + \begin{pmatrix} O \\ \lambda \end{pmatrix} B' = O.$$

The upper half part of (3.13) gives $Z_1'(Y - Z\hat{\Pi}) = O$, and we have

$$J_1'\hat{\Pi} = (Z_1'Z_1)^{-1}Z_1'(Y - Z_2J_2'\hat{\Pi}),$$

where $J_1' = (I_{K_1}, O)$ is a $K_1 \times K$ choice matrix. Then

$$(3.14)\quad Y - Z\hat{\Pi} = Y - Z_1J_1'\hat{\Pi} - Z_2J_2'\hat{\Pi}$$

$$= \bar{P}_{Z_1}(Y - Z_2J_2'\hat{\Pi}).$$

Multiplying ΩB from the right hand side of (3.13), we obtain

$$(3.15)\quad \begin{pmatrix} O \\ \lambda \end{pmatrix} = - Z'\bar{P}_{Z_1} Y\ B\ \Sigma^{-1}.$$

Using Lemma A.4, we have the first and second derivatives of the likelihood function as follows

$$(3.16) \qquad \frac{\partial \log L_4}{\partial \mathrm{vec}\Pi} = \frac{\partial \log L_1}{\partial \mathrm{vec}\Pi} + \mathrm{vec}\left(\begin{matrix} 0 \\ \lambda \end{matrix}\right)B' = 0, \quad \frac{\partial^2 \log L_1}{\partial \mathrm{vec}\Pi\, \partial(\mathrm{vec}\Pi)'} = -\Omega^{-1} \otimes Z'Z.$$

Then we define a LM statistic by

$$(3.17) \qquad LM_1 = \left(\frac{\partial \log L_1}{\partial \mathrm{vec}\Pi}\right)'\left(-\frac{\partial^2 \log L_1}{\partial \mathrm{vec}\Pi\, \partial(\mathrm{vec}\Pi)'}\right)^{-1}\left(\frac{\partial \log L_1}{\partial \mathrm{vec}\Pi}\right).$$

Using Lemma A.5 and $\Sigma = B'\Omega B$, we have

$$(3.18) \qquad LM_1 = \left[\mathrm{vec}\left(\begin{matrix} 0 \\ \lambda \end{matrix}\right)B'\right]'\left(\Omega \otimes (Z'Z)^{-1}\right)\left[\mathrm{vec}\left(\begin{matrix} 0 \\ \lambda \end{matrix}\right)B'\right]$$

$$= \mathrm{tr}\, \Sigma(0\ \lambda')(Z'Z)^{-1}\left(\begin{matrix} 0 \\ \lambda \end{matrix}\right),$$

where the unknown parameters in (3.18) are evaluated at their maximum likelihood estimators. (See Engle (1984), for instance.) We note that LM statistic in the form (3.17) has been known as the Rao's Scoring test among statisticians. From (3.15) and $\bar{P}_{Z_1} P_Z \bar{P}_{Z_1} = P_Z - P_{Z_1}$, we further simplify LM_1 as

$$(3.19) \qquad LM_1 = \mathrm{tr}\, \hat{B}'Y'(P_Z - P_{Z_1})Y\hat{B}\hat{\Sigma}^{-1},$$

where $\hat{B}$ and $\hat{\Sigma}$ are the maximum likelihood estimators of B and Σ under the null hypothesis. If we use the characteristic roots of (3.11), this statistic has an expression of

(3.20) $$LM_1 = T \sum_{i=1}^{G_0} \frac{\lambda_i}{1+\lambda_i} .$$

When G_0 (= G_1) = 1, this statistic LM_1 is identical to the LM statistic proposed by Byron (1972). However, it should be noted that the derivation of his statistic is different from ours. It should be also noted that (3.20) is an analogue of the Bartlett-Nanda-Pillai Trace Criterion, which is well known in multivariate statistical analysis. (See Anderson (1984).) Our derivation yields a new interpretation of the Bartlett-Nanda-Pillai test.

(3.c) <u>Wald Test</u>

Let us consider a Wald type statistic for the present testing problem. For this purpose, we first consider a subset of structural equations using all exogenous variables appearing in the reduced form and write

(3.21) $$Y_1B_1 = Y_2B_2 + Z_1\Gamma + Z_2\xi + U.$$

We decompose matrix Z_2 into $T\times(L_1+ G_2)$ submatrices $Z_2 = (Z_{21}, Z_{22})$. Similarly, two matrices ξ and (Π_{21}, Π_{22}) are decomposed into $(L_1+ G_2)\times G_1$ and $(L_1+ G_2)\times(G_1+ G_2)$ submatrices

(3.22) $$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}, \quad (\Pi_{21}, \Pi_{22}) = \begin{bmatrix} \Pi_{21.1} & \Pi_{22.1} \\ \Pi_{21.2} & \Pi_{22.2} \end{bmatrix}.$$

In this formulation, the testing problem of overidentifying restrictions can be interpreted as the one in which the null hypothesis is H_ξ: $\xi_1 = 0$, $\xi_2 = 0$, and the alternative hypothesis is H_{ξ_2} : $\xi_1 \neq 0$, $\xi_2 = 0$. If we further assume that a $G_2 \times G_2$ reduced form coefficient matrix $\Pi_{22.2}$ is of full rank, H_ξ is expressed as zero restrictions because

$$(3.23) \quad \xi_1 = (\Pi_{21.1} - \Pi_{22.1}\Pi_{22.2}^{-1}\Pi_{21.2})B_1 .$$

Hence in principle we can construct a Wald type statistic from the sample analogue of ξ_1 based on the unrestricted least squares estimator of Π.

Allow partition of a $K_2 \times K$ choice matrix J_3' into $(L_1 + G_2) \times K$ submatrices

$$(3.24) \quad J_3' = (O,\ I_{K_2}) = \begin{bmatrix} J_{31}' \\ J_{32}' \end{bmatrix}.$$

The alternative hypothesis H_{ξ_2} implies that the subsystem of structural equations is just-identified. If we take $B_1 = I_{G_1}$ for normalization, the estimator of (B_2, Γ, ξ_1) can be obtained by regressing $(\hat{Y}_2, Z_1, Z_{21})$ on Y_1, where $\hat{Y}_2 = (Z'Z)^{-1}Z'Y_2$. Then from (3.23) and Lemma A.6, an estimator of ξ_1 could be written as

$$(3.25) \quad \hat{\xi}_1 = J_{31}'(Z'Z)^{-1}Z'\{I_T - Y_2[J_{32}'(Z'Z)^{-1}Z'Y_2]^{-1}J_{32}'(Z'Z)^{-1}Z'\}Y_1B_1$$

$$= (Z_{21}'N_1Z_{21})^{-1}Z_{21}'N_1Y_1B_1,$$

where

$$N_1 = \bar{P}_{Z_1} - \bar{P}_Z - (\bar{P}_{Z_1} - \bar{P}_Z)Y_2[Y_2'(\bar{P}_{Z_1} - \bar{P}_Z)Y_2]^{-1}Y_2'(\bar{P}_{Z_1} - \bar{P}_Z).$$

In the above we have utilized that $Z_{21}'N_1Z_{21}$ is nonsingular because the matrix (Y_2, Z_1, Z_{21}) is of full rank and $\mathrm{rank}(N_1) = \mathrm{rank}(Z_{21}) = L_1$. Then the asymptotic covariance matrix of $\mathrm{vec}(\hat{\xi}_1)$ is calculated as

$$(3.26) \quad [I_{G_1} \otimes (Z_{21}'N_1Z_{21})^{-1}Z_{21}'N_1](\Sigma \otimes I_T)[I_{G_1} \otimes (Z_{21}'N_1Z_{21})^{-1}Z_{21}'N_1]$$

$$= \Sigma \otimes (Z_{21}'N_1Z_{21})^{-1} .$$

We now define a Wald type statistic by

$$(3.27) \quad W_1 = (\text{vec } \hat{\xi}_1)'\{\hat{\Sigma}\otimes(Z_{21}'N_1Z_{21})^{-1}\}^{-1}(\text{vec } \hat{\xi}_1) .$$

where $\hat{\xi}_1$ and $\hat{\Sigma}$ are constructed from the maximum likelihood estimators of parameters when the subsystem of structural equations is just-identified. In the situation where we normalized $B_1 = I_{G_1}$, the maximum likelihood estimator of B is the two-stage least squares estimator of B. Thus using Lemma A.6 again, we have

$$(3.28) \quad W_1 = \text{tr}\{ \hat{\Sigma}^{-1}\hat{B}_1'Y_1'N_1Z_{21}(Z_{21}'N_1Z_{21})^{-1}Z_{21}N_1Y_1\hat{B}_1\}$$

$$= \text{tr}\{ \hat{\Sigma}^{-1}\hat{B}'Y'(P_Z - P_{Z_1})Y\hat{B}\}$$

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

where an estimator of $\hat{\Sigma}$ is constructed from $\hat{B}$ and $\hat{\Omega}$, and λ_i ($i=1,\ldots,G_1$) are the characteristic roots of (3.11).

When we use the unrestricted estimator of Ω,

$$(3.29) \quad \hat{\Omega} = \frac{1}{T} Y'\bar{P}_Z Y,$$

the resulting Wald statistic W_1 is identical to the one derived by Wegge (1978) for the case of G_0 (= G_1) = 1. In this case, Hwang (1980a) has shown that Wegge's statistic is identical to the Wald statistic derived by Byron (1974). When we use the maximum likelihood estimator of Ω under the null hypothesis

$$(3.30) \quad \hat{\Omega} = \frac{1}{T} Y'\bar{P}_{Z_1} Y ,$$

the resulting W_1 reduces to the statistic proposed by Basman (1960) for the case of G_0 (= G_1) = 1. It can be interpreted as a modified LM statistic in the present context. It should be also noted that (3.29) is an analogue of the Lawley-Hotelling Trace Criterion, which is well-known in multivariate statistical analysis. (See Anderson (1984).) Thus our derivation also gives a new interpretation to the Lawley-Hotelling type statistic.

(3.d) 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

$$(3.31) \quad 0 \leqq LM_1 \leqq LR_1 \leqq W_1 .$$

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), for instance.) If we use the same asymptotic χ^2 distribution, the Wald type statistic tends to reject the hypothesis more than other type statistics while the likelihood ratio statistic tends to reject the hypothesis more than LM statistics.

4. Tests of Predeterminedness

In this section we shall derive several statistics for the hypothesis of econometric predeterminedness $H_{\xi,\eta}$ we defined in Section 2. We assume that $G_0 \leqq G_1$ and the multivariate normal distribution for disturbance terms.

(4.a) Likelihood Ratio Test

Let a $G \times G$ matrix

(4.1) $$\Lambda_2 = \begin{pmatrix} I & 0 \\ -\rho & I \end{pmatrix} ,$$

where $\rho = \Omega_{22}^{-1}\Omega_{21}$. Multiplying Λ_2 to (2.2) from the right, we obtain

(4.2) $$(Y_1^*, Y_2) = Z(\Pi_1^{**}, \Pi_2) + (V_1^*, V_2) ,$$

where

(4.3) $$Y_1^* = Y_1 - Y_2\rho , V_1^* = V_1 - V_2\rho ,$$

(4.4) $$\Pi_1^{**} = \begin{pmatrix} \Pi_{11}^{**} \\ \Pi_{21}^{**} \end{pmatrix} = \Pi\begin{pmatrix} I \\ -\rho \end{pmatrix} .$$

By this transformation of the system, the covariance matrix of (V_1^*, V_2) is given by

(4.4) $$\Lambda_2'\Omega\Lambda_2 = \begin{pmatrix} \Omega_{11.2} & 0 \\ 0 & \Omega_{22} \end{pmatrix} ,$$

where $\Omega_{11.2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$. From (3.2), the likelihood function is rewritten as

(4.5) $$L_5 = c_1|\Omega_{11.2}|^{-T/2}\exp\{-\tfrac{1}{2}\mathrm{tr}\Omega_{11.2}^{-1}(Y_1^*-Z\Pi_1^{**})'(Y_1^*- Z\Pi_1^{**})\}$$

$$\times |\Omega_{22}|^{-T/2}\exp\{-\tfrac{1}{2}\mathrm{tr}\Omega_{22}^{-1}(Y_2- Z\Pi_2)'(Y_2 - Z\Pi_2)\}.$$

To maximize L_5 with respect to $\Omega_{11.2}$, Ω_{22}, and Π_2, we obtain

(4.6) $$\hat{\Pi}_2 = (Z'Z)^{-1}Z'Y_2 \quad , T\hat{\Omega}_{22} = Y_2'\bar{P}_ZY_2 ,$$

$$T\hat{\Omega}_{11.2} = [I_{G_1}, -\rho'](Y - Z\hat{\Pi})'(Y - Z\hat{\Pi})\begin{bmatrix} I_{G_1} \\ -\rho \end{bmatrix}.$$

Then the concentrated likelihood function is given by

(4.7) $$L_6 = c_2 |(Y_1^* - Z\Pi_1^{**})'(Y_1^* - Z\Pi_1^{**})|^{-T/2} | T\hat{\Omega}_{22}|^{-T/2}$$

Because the second term in L_6 can be regarded as a constant, we shall maximize the first term with respect to Π_1^{**}. From (2.9) and (2.16), the hypothesis $H_{\xi,\eta}$ implies H_ζ: $\zeta = 0$, where

(4.8) $$\zeta = \Pi_{21}^{**} B_1 .$$

Let a $G_1 \times G_1$ matrix

(4.9) $$\Lambda_3 = (B_1, J_4) ,$$

where $J_4' = (0, I_{G_1 - G_0})$ is a $(G_1 - G_0) \times G_1$ choice matrix. Then the concentrated likelihood function is rewritten as

(4.10) $$L_7 = c_2 |S_{11}^*|^{-T/2} | T\hat{\Omega}_{22}|^{-T/2}$$

where

(4.11) $$S_{11}^* = (W^* - R\Pi_1^\dagger)'(W^* - R\Pi_1^\dagger) , \quad W^* = Y_1 \Lambda_3 , \quad \Pi_1^\dagger = \begin{pmatrix} \rho B_1 & \rho J_4 \\ \Gamma & \Pi_{11}^{**} J_4 \\ \zeta & \Pi_{21}^{**} J_4 \end{pmatrix} ,$$

and $\Gamma = \Pi_{11}^* B_1$ and $R = (Y_2, Z)$ is a $T \times (G_2 + K)$ matrix.

Let $\lambda_{G_1}^* \geq \cdots \geq \lambda_1^* \geq 0$ be the characteristic roots of

(4.12) $$|Y_1'(P_R - P_L)Y_1 - \lambda^* Y_1' \bar{P}_R Y_1| = 0 ,$$

where $L = (Y_2, Z_1)$ is a $T \times (G_2 + K_1)$ matrix.

First, the alternartive hypothesis against $H_{\xi,\eta}$ we consider is the reduced form (2.2) ignoring the block identifiability restrictions, i.e., H_A: $\xi \neq 0$, $\eta \neq 0$. In this case the maximized likelihood was given by L_3 in (3.8). Using Lemma A.2 as in Section 3.a, the likelihood ratio statistic for $H_{\xi,\eta}$ vs H_A is given by

(4.13) $$LR_2 = T \log \prod_{i=1}^{G_0} (1 + \lambda_i^*).$$

where λ_i^* is the characteristic roots of equations (4.12).

Another possible alternative hypothesis against $H_{\xi,\eta}$ may be the the structural equations with the block identifiability restrictions, i.e., H_ξ: $\xi = 0$. In this case the maximized likelihood can be obtained from (3.10). Thus the likelihood ratio statistic for $H_{\xi,\eta}$ vs H_ξ is given by

(4.14) $$LR_3 = T \sum_{i=1}^{G_0} \log\left(\frac{1 + \lambda_i^*}{1 + \lambda_i}\right)$$

where λ_i^* and λ_i are the characteristic roots of equations (4.12) and (3.11), respectively.

When $G_0 = 1$ and $G_1 \geqq 1$, 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.b) Lagrange Multiplier Test

Let λ and λ_0 be a $K_2 \times G_1$ and $G_2 \times G_0$ matrices of Lagrange Multiplier parameters for $\xi = 0$ and $\eta = 0$, respectively. The Lagrange form in this case is written as

(4.15) $$\log L_8 = \log c_1 - \frac{T}{2} \log |\Omega| - \frac{1}{2} \mathrm{tr}\Omega^{-1}(Y - Z\Pi)'(Y - Z\Pi) + \mathrm{tr}\lambda'(\Pi_{21}\ \Pi_{22})B$$

$$+ \mathrm{tr}\ \lambda_0'(\rho,\ I_{G_2})B .$$

Differentiating $\log L_8$ with respect to B_2, we have

(4.16) $$\Pi_{22}'\lambda + \lambda_0 = 0.$$

Substituting this relation into $\log L_8$ and ignoring a constant term,

$$(4.17) \quad \log L_9 = -\frac{T}{2}\log|\Omega_{11.2}||\Omega_{22}| - \frac{1}{2}\mathrm{tr}\Omega^{-1}(Y - Z\Pi)'(Y - Z\Pi)$$

$$+ \mathrm{tr}\ \lambda'(\Pi_{21}^{**}B_1)\ .$$

where $\Pi_{21}^{**} = \Pi_{21} - \Pi_{22}\rho$ and $\zeta = \Pi_{21}^{**}B_1$. (See (4.4) and (4.8)).

By differentiating $\log L_9$ with respect to Π and ρ, we also have

$$(4.18) \quad Z'Z\hat{\Pi} = Z'Y + \begin{pmatrix} 0 & 0 \\ \lambda B_1'\Omega_{11.2} & 0 \end{pmatrix},$$

$$(4.19) \quad Y_2'\bar{P}_Z Y \begin{pmatrix} I_{G_1} \\ -\hat{\rho} \end{pmatrix} = \hat{\Pi}_{22}'\lambda\hat{B}_1'\hat{\Omega}_{11.2}\ .$$

The last equality was obtained as a result of the relations

$$(4.20) \quad \Omega^{-1} = \begin{pmatrix} 0 \\ I_{G_2} \end{pmatrix}\Omega_{22}^{-1}(0,\ I_{G_2}) + \begin{pmatrix} I_{G_1} \\ \rho \end{pmatrix}\Omega_{11.2}^{-1}(I_{G_1},\ \rho')\ ,$$

$$(4.21) \quad (Y - Z\Pi)'(Y - Z\Pi) = Y'\bar{P}_Z Y + \begin{pmatrix} 0 & \Omega_{11.2}B_1\lambda' \\ 0 & 0 \end{pmatrix}(Z'Z)^{-1}\begin{pmatrix} 0 & 0 \\ \lambda B'\Omega_{11.2} & 0 \end{pmatrix}.$$

Substituting the upper parts of (4.18) and (4.19) into their lower parts, we have

$$(4.22) \quad \lambda B_1'\Omega_{11.2} = Z_2\ \bar{P}_{Z_1}(Z_2\hat{\Pi}_{21} - Y_1)\ ,$$

and

$$(4.23) \quad Y_2'\bar{P}_Z Y\begin{pmatrix} I_{G_1} \\ -\hat{\rho} \end{pmatrix} = \hat{\Pi}_{22}\ Z_2'\ \bar{P}_{Z_1}(Z_2\hat{\Pi}_{21} - Y_1)\ .$$

Since $\hat{\Pi}_{21} = \hat{\Pi}_{22}\rho$ and $\hat{\Pi}_{22} = J_3'\hat{\Pi}_2$ in (4.6), we solve (4.23) with respect to ρ and

(4.24) $\hat{\rho} = (Y_2'\bar{P}_{Z_1}Y_2)^{-1}Y_2'\bar{P}_{Z_1}Y_1$.

Multiplying B_1 to (4.22) from the right, we have

$$(4.25) \quad \lambda\Sigma = Z_2' \bar{P}_{Z_1} (Z_2\hat{\Pi}_{22}\hat{\rho} - Y_1)B_1$$

$$= - Z_2' \bar{P}_L Y_1 B_1,$$

where $\Sigma = B_1'\Omega_{11.2}B_1$. In the above derivation, we have used Lemma A.7 for $L = (Y_2, Z_1)$.

We first consider the testing the hypothesis $H_{\xi,\eta}$ vs H_A. A similar derivation as LM_1 in Section 3-(b) can be used to the present testing problem. We note that $H_{\xi,\eta}$ implies $H_\zeta : \zeta = 0$ instead of H_ξ in Section 3-(b). Thus a Lagrange Multiplier statistic in this case may be given by

$$(3.26) \quad LM_2 = T \sum_{i=1}^{G_0} \frac{\lambda_i^*}{1+\lambda_i^*} ,$$

where λ_i^* are the characteristic roots of (4.12).
Although the underlying idea in (4.26) is straightforward, this type of statistic LM_2 has not been previously derived.

We now consider the testing the hypothesis H_ζ vs H_ξ. In this case, we confine ourselves to the case of $G_0 = G_1$ and use the normalization $B_1 = I_{G_1}$ mainly for simplicity. Although the general case ($G_1 \geq G_0$) can be treated in the same way as we shall do, the resulting statistic become rather complicated. From (4.25), we observe that the Lagrange multiplier matrix in the sample for $H_\eta : \eta = 0$ is

$$(4.27) \quad \lambda_0 = Y_2' P_Z \bar{P}_L Y_1 \hat{\Sigma}^{-1},$$

where $\hat{\Sigma}$ is the maximum likelihood estimators of Σ. Since $\hat{\Sigma}$ is consistent estimator of Σ, we consider the quantity

$$(4.28) \qquad \lambda_0^* = Y_2' \, P_Z \bar{P}_L \, Y_1 \Sigma^{-1},$$

which is asymptotically equivalent to λ_0. Using Lemma A.5, we rewrite (4.28) as

$$(4.29) \qquad \text{vec } \lambda_0^* = (\Sigma^{-1} \otimes Y_2' P_Z \bar{P}_L) \text{ vec } U \ .$$

Then the covariance matrix of vec λ_0^* is given by

$$(4.30) \qquad (\Sigma^{-1} \otimes Y_2' P_Z \bar{P}_L)(\Sigma \otimes I_T)(\Sigma^{-1} \otimes Y_2' P_Z \bar{P}_L)' = \Sigma^{-1} \otimes Y_2' P_Z \bar{P}_L P_Z Y_2 \ .$$

We now define a LM statistic by

$$(4.31) \qquad LM_3 = (\text{vec } \lambda_0)'(\hat{\Sigma}^{-1} \otimes Y_2' P_Z \bar{P}_L P_Z Y_2)^{-1} (\text{vec } \lambda_0) \ .$$

Then by the use of Lemma A.5 and (4.27), we rewrite (4.31) as,

$$(4.32) \qquad LM_3 = \text{tr}\{Y_1' \bar{P}_L \bar{P}_Z Y_2 (Y_2' \bar{P}_Z \bar{P}_L \bar{P}_Z Y_2)^{-1} Y_2' \bar{P}_Z \bar{P}_L Y_1 \hat{\Sigma}^{-1}\}.$$

Further, using Lemma A.6, we obtain the expression

$$(4.33) \qquad LM_3 = \text{tr}\{Y_1' (\bar{P}_L - \bar{P}_X) Y_1 \hat{\Sigma}^{-1}\}$$

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

where λ_i^{**} are the characteristic roots of

(4.34) $$|Y_1'(P_X - P_L)Y_1 - \lambda^{**}Y_1'\bar{P}_X Y_1| = 0 ,$$

and $X = (L, \bar{P}_Z Y_2)$ is a $T\times(G_2 + K_1 + G_2)$ matrix.

In the present formulation of the LM test, $\hat{\Sigma}$ should be based on the maximum likelihood estimator of Σ under the null hypothesis:

(4.35) $$\hat{\Sigma} = \hat{\Omega}_{11.2} = \frac{1}{T} Y_1'\bar{P}_L Y_1 .$$

However, in practice, several estimators of Σ could be used. For instance, instead of (4.35), we may use

(4.36) $$\hat{\Sigma} = \frac{1}{T - 2G_2 - K_1} Y_1'\bar{P}_X Y_1 .$$

In particular, LM_3 with (4.36) reduces to the statistic proposed by Wu (1973) and Wu (1974) when $G_1 = 1$.

On the other hand, Hausman (1978) considered a testing problem for

(4.37) $$Y_1 = Y_2 B_2 + Z_1\Gamma + E_3 B_3 + u,$$

where $G_1 = 1$, B_3 is a $G_3\times 1$ vector of unknown parameters, and E_3 is the least squares residuals $E_2 = Y_2 - \hat{Y}_2 = \bar{P}_Z Y_2$. Hausman (1978) proposed the usual F test for H_0: $B_3 = 0$ against H_1: $B_3 \neq 0$ as a specification test.

From (4.34), it is clear that LM_3 is proportional to Hausman's statistic when $G_1 = 1$. 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.

Another possibility of an estimator of Σ may be

(4.38) $$\hat{\Sigma} = \frac{1}{T - K - G_2} Y_1'\bar{P}_R Y_1$$

because it is an unrestricted sum of squares from the regression residuals. Then, the statistic LM_3 with (4.36) reduces to the one proposed by Revankar

(1978) when $G_1 = 1$. Therefore, we can also reinterprete Revankar's test as a LM test procedure.

(4.c) **Wald Test**

Let us derive Wald type statistics for the present testing problem. For this purpose, we first consider the null hypothesis $H_{\xi,\eta}$ vs hypothesis H_A. In this case, our derivation of a Wald test is similar to Section 3-(c). Thus a Wald type statistic may be given by

$$(4.39) \qquad W_2 = T \sum_{i=1}^{G_0} \lambda_i^* ,$$

where λ_i^* are the characteristic roots of (4.12).

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 turn to derive a Wald type statistic for $H_{\xi,\eta}$ against H_ξ. In this testing problem, we confine ourselves to the case of $G_0 = G_1$ and use the normalization $B_0 = I_{G_1}$ mainly for simplicity. Although the general case $(G_1 \geq G_0)$ can be treated in the same way as we shall do, the resulting statistics become rather complicated. We note that from (3.13)

$$(4.40) \qquad T\hat{\Omega} = (Y - Z\hat{\Pi})'(Y - Z\hat{\Pi})$$

$$= Y'\bar{P}_Z Y + \hat{\Omega}\hat{B} \binom{O}{\lambda}' (Z'Z)^{-1} \binom{O}{\lambda} \hat{B}'\hat{\Omega} .$$

Using (3.15), we have

(4.41) $T\hat{\Omega}\hat{B} = Y'\bar{P}_Z Y\hat{B} + \hat{\Omega}\hat{B}\hat{\Sigma}^{-1}\hat{B}'Y'(P_Z - P_{Z_1})Y\hat{B}$.

Because $T\hat{\Sigma} = \hat{B}'Y'\bar{P}_{Z_1}Y\hat{B}$, we have an unrestricted estimator of η,

(4.42) $\hat{\eta} = J_2'\hat{\Omega}\hat{B} = J_2'Y'\bar{P}_Z Y\hat{B}(I + \Lambda)$,

where $\Lambda = \text{diag}(\lambda_i)$ is a $G_0 \times G_0$ diagonal matrix with G_0 characteristic roots of (3.11) and $0 \leq \lambda_1 \leq \cdots \leq \lambda_{G_0}$. Since $\lambda_i = O_p(1/T)$ as shown in (B.12) in Appendix B, the asymptotic distribution of $\hat{\eta}$ is the same as that of

(4.43) $\eta^* = J_2'\Omega\sqrt{T}(\hat{B} - B) + J_2'\sqrt{T}[\frac{1}{T}Y'\bar{P}_Z Y - \Omega]B$.

Using Lemma A.5, we rewrite

(4.44) $\text{vec}(\eta^*) = (I_{G_0} \otimes \Omega_{22})\text{vec}\sqrt{T}(\hat{B}_2 - B_2) + (B' \otimes J_2')\text{vec}\sqrt{T}[\frac{1}{T}Y'\bar{P}_Z Y - \Omega]$.

Under the assumption of normal disturbances, the asymptotic covariance of $\text{vec}\sqrt{T}[\frac{1}{T}Y'\bar{P}_Z Y - \Omega]$ is given by

(4.45) $(I_{G_0G_0} + K_{G_0G_0})\Omega\otimes\Omega$,

where K is the commutation matrix, that is, $\text{vec}(A) = K\,\text{vec}(A')$ for any arbitrary matrix A. (See Anderson (1987), for instance.) By the use of Lemma B.1, the asymptotic covariance of η^* is given by

(4.46) $\Sigma\otimes[\Omega_{22}(\Pi_{22}'M_{22.1}\Pi_{22})^{-1}\Omega_{22} + \Omega_{22}] + (B'\otimes J_2')K_{G_0G_0}(\Omega B \otimes \Omega J_2)$.

Using (A.10) in Lemma A.5, the last term of (4.43) becomes a zero-matrix under the hypothesis H_0: $\eta = 0$. Hence, we define a Wald statistic by

(4.47) $W_3 = (\text{vec}\,\hat{\eta})'\{\hat{\Sigma} \otimes [T\hat{\Omega}_{22}(\hat{\Pi}_{22}A_{22.1}\hat{\Pi}_{22})^{-1}\hat{\Omega}_{22} + \hat{\Omega}_{22}]\}^{-1}(\text{vec}\,\hat{\eta})$, where $\hat{\Sigma}$, $\hat{\Omega}_{22}$, and $\hat{\Pi}_{22}$ are the maximum likelihood estimators of Σ, Ω_{22}, and Π_{22}, respectively. Again, using Lemma A.5, we have

$$(4.48) \quad W_3 = tr(I + \Lambda)\hat{\Sigma}^{-1}(I + \Lambda)\hat{B}'Y'\bar{P}_Z YJ_2[T\hat{\Omega}_{22}(\hat{\Pi}_{22}'A_{22.1}\hat{\Pi}_{22})^{-1}\hat{\Omega}_{22} + \hat{\Omega}_{22}]^{-1}J_2'Y'\bar{P}_Z Y\hat{B},$$

where $\hat{B}$ is the maximum likelihood estimator of B, where its elements are the characterestic vectors of (3.11).

This Wald type statistic has not been obtained before. Although it is complicated in general, there is a situation where it can be further simplified. We consider the case when the subsystem of structural equations is just-identified as the alternative hypothesis. In this case, since λ_i = 0 $(i=1,\ldots,G_1)$ in (3.11) and $\Lambda = 0$ in (4.47), we have $T\hat{\Omega}_{22} = Y_2'\bar{P}_Z Y_2$, $\hat{\Pi}_2 = (Z'Z)^{-1}Z'Y_2$, $T\hat{\Sigma} = \hat{B}'Y'\bar{P}_{Z_1}Y\hat{B}$. Then, in particular, when $G_0 = G_1 = 1$, it can be shown that W_3 in (4.47) 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.d) 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 $H_{\xi,\eta}$ vs H_{ξ}. Using Lemma A.7, we have

$$(4.49) \quad 0 \leq LM_2 \leq LR_2 \leq W_2 .$$

This inequality is an analogue to (3.31) for the testing problem of the block identifying restirction in Section 3. However, a similar inequality can not be obtainable for the testing problem of $H_{\varepsilon,\eta}$ vs H_{ε}.

5. Asymptotic Distributions of Statistics

We now return to the assumptions (2.3) and (2.4) on disturbances we made in Section 2. We note that the assumptions (2.3) and (2.4) imply (2.5) and (2.6) because $u_t' = Bv_t$. In the conditional expectation operator in (2.3)-(2.6), F_{t-1} is the information set available at t-1. The exogenous variables z_t include a finite number of past endogenous variables y_{t-1}, y_{y-2}, ..., y_{t-p}. In order to investigate the asymptotic distribution of test statistics in this situation, we obtain a new Martingale Central Limit Theorem.

<u>Theorem 5.1</u>: Let $\{z_t, v_t\}$ be a sequence of pairs of random vectors, and let F_{t-1} be the σ-field generated by z_1, v_1, ..., z_{t-1}, v_{t-1}, z_t. Suppose (i)

$$M_T = \frac{1}{T}\sum_{t=1}^{T} z_t z_t' \xrightarrow{p} M , \tag{5.1}$$

(ii) $E(v_t | F_{t-1}) = 0$ a.s. , $E(v_t v_t' | F_{t-1}) = \Omega$ a.s. ,

and (iii)

$$\sup_t \int_{v'v>c} v'v \, dG_t \longrightarrow 0 \quad \text{a.s.} \tag{5.2}$$

as $c \rightarrow +\infty$, where G_t is the conditional distribution function of v_t given F_{t-1}. Then

$$\mathrm{vec}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_t v_t'\right) \xrightarrow{L} N\,(0,\ \Omega \otimes M). \tag{5.3}$$

This theorem generalizes Theorem 5(i) of Lai and Robbins (1981), where the scalar v_t are independently identically distributed. Because it is relatively easy to check the conditions (i)-(iii), it may be useful for many applications.

From Assumption (i) in Theorem 5.1, we note that a $(G_2+K)\times(G_2+K)$ matrix

$$(5.4)\qquad \frac{1}{T} R'R \xrightarrow{p} Q = \begin{pmatrix} \Pi_2' \\ I_K \end{pmatrix} M (\Pi_2, I_k) + \begin{pmatrix} \Omega_{22} & 0 \\ 0 & 0 \end{pmatrix}$$

as $T \to +\infty$.

Consider a local alternative hypothesis for the identifiability restrictions,

$$(5.5)\qquad \Pi B = \begin{pmatrix} \Gamma \\ 0 \end{pmatrix} + \frac{1}{\sqrt{T}} \xi_1 ,$$

where ξ_1 is a non-zero $K \times G_0$ matrix. When $\xi_1 = 0$, (5.6) reduces to (2.7) and (2.8). Kunitomo (1987) discussed the formulation of these local alternatives in some detail. Subsequently, we obtain the next result, for which the proof is given in Appendix B.

<u>Theorem 5.2</u> : Under the assumptions we made in Theorem 5.1, the three statistics LR_1, LM_1, and W_1 are asymptotically distributed as noncentral χ^2 with $G_0 \times L_1$ degrees of freedom and the noncentrality

$$(5.6)\qquad \delta_1^2 = tr(\theta_1 \Sigma^{-1}) ,$$

where

$$(5.7)\qquad \theta_1 = \xi_1' \left[M - MD(D'MD)^{-1}D'M\right] \xi_1 ,$$

where D is a $K\times(G_2+K_1)$ matrix

$$D = \begin{pmatrix} \Pi_2 & \begin{matrix} I_{K_1} \\ 0 \end{matrix} \end{pmatrix}.$$

When $\xi_1 = 0$, the three statistics are distributed as χ^2 with $G_0 \times L_1$ degrees of freedom. This result for the case of $G_0 = 1$ has been obtained under the assumptions that disturbances are independently, identically, and normally distributed and there are no lagged endogenous variables in the explanatory variables.

Next, we consider a local alternative hypothesis for the predeterminedness condition,

$$(5.8) \qquad (\Omega_{21} \;\; \Omega_{22})\, B = \frac{1}{\sqrt{T}}\, \eta_1$$

where η_1 is a non-zero $G_2 \times G_1$ matrix. In this formulation of alternatives, we obtain the next result, for which the proof is given in Appendix B.

Theorem 5.3 : (i) Under the assumptions we made in Theorem 5.1, three statistics LM_3, LR_3, and W_3 are asymptotically distributed as noncentral χ^2 with $G_0 \times G_2$ degrees of freedom and the noncentrality

$$(5.9) \qquad \delta_2^2 = \mathrm{tr}(\Theta_2 \Sigma^{-1})$$

where

$$(5.10) \qquad \Theta_2 = \eta_1'(\Omega_{22} + \Omega_{22}(\Pi_{22}' M_{22.1} \Pi_{22})^{-1} \Omega_{22})^{-1} \eta_1,$$

$$(5.11) \qquad M_{22.1} = M_{22} - M_{21} M_{11}^{-1} M_{12} ,$$

and we decompose M into $(K_1 + K_2) \times (K_1 + K_2)$ submatrices

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}.$$

(ii) Under the same assumptions we made in Theorem 5.1, the three statistics LR_2 , LM_2, and W_2 are asymptotically distributed as a noncentral χ^2 with $G_0 \times K_2$ degrees of freedom and the noncentrality δ_2^2.

When $n_1 = 0$, the two statistics LM_3 and LR_3 are asymptotically distributed as χ^2 with $G_0 \times G_2$ degrees of freedom and statistics LR_2 and W_2 are asymptotically distributed as χ^2 with $G_0 \times K_2$ degrees of freedom. These results for the case of $G_0 = 1$ have been obtained under the assumptions that the disturbances are independently and identically distributed and there are no lagged endogenous variables. Furthermore, in this case it is known that Wu's statistic, Revankar's statistic, and Revankar-Hartley statistic adjusted by their degrees of freedom are distrubuted as F when the disturbances are normally distributed.

6. Conclusion

In this paper, we have systematically derived a number of test procedures for testing the block identifiability condition and the predeterminedness condition in a subsystem of structural equations. We generalized the test statistics proposed in the past and derived the LR test, LM test, and Wald test for these two testing problems. This formaulation 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.

We have also derived the asymptotic distributions of test statistics under a set of fairly general conditions on the disturbance terms. For this purpose, we have derived a new Martingale Central Limit Theorem and applied it to the present situation. We allow that there are a finite number of lagged endogenous variables and the disturbance terms are the Martingale difference sequences, which are not necessarily independent. We have shown that the limiting distributions of test statistics considered in this paper are the non-central χ^2 distributions under local alternative hypotheses and are the central χ^2 distributions under the null-hypotheses. Because test statistics have often been proposed under a set of relatively restrictive assumptions, it may be important to show that the assumptions usually made are not essential for the testing procedures. Our results on the asymptotic robustness of tests give justification for using the test procedures commonly known in econometrics and multivariate analysis even if some of the usual underlying assumptions for their derivations are not satisfied.

Appendix A:

In Appendix A, we present some useful lemmas. Most lemmas here are known in the multivariate statistical analysis and their proofs could 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

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

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,

$$\text{(A.2)} \qquad \min_{C} \left| A + \begin{pmatrix} B' \\ C' \end{pmatrix} (B,\ C) \right| = \frac{|A|}{|A_{11}|} \left| A_{11} + B'B \right|.$$

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

$$\text{(A.3)} \qquad |A + D'D| = \begin{vmatrix} A & -D' \\ D & I_q \end{vmatrix} = |A|\,|I_q + DA^{-1}D'|.$$

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

$$DA^{-1}D' = [C + B\,A^{12}(A^{22})^{-1}]A^{22}[C + B\,A^{12}(A^{22})^{-1}]'$$

$$+ B[A^{11} - A^{12}(A^{22})^{-1}A^{21}]B' \geq B\,A_{11}^{-1}B'.$$

Hence,

$$\text{(A.4)} \qquad |A + D'D| \geq |A|\,|I_q + B\,A_{11}^{-1}B'|.$$

Finally, we obtain (A.2) by using (A.3). ■

Lemma A.3: Let A be a $p\times p$ non-negative definite matrix and $0 \leq \lambda_1 \leq \cdots \leq \lambda_p$ be its characteristic roots. Let also B be a $p\times q$ $(p > q)$ matrix. Then

$$\text{(A.5)} \qquad \min_{B'B=I} |B'AB| = \prod_{i=1}^{q} \lambda_i, \quad \max_{B'B=I} |B'A\,B| = \prod_{i=0}^{q-1} \lambda_{p-i}.$$

Lemma A.4:

(A.6) $$\frac{\partial tr(AB)}{\partial B} = A', \quad \frac{\partial tr(B'ABC)}{\partial B} = ABC + A'BC'.$$

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

(A.7) $$\text{vec}(BXC) = (C'\otimes B)(\text{vec } X),$$

(A.8) $$tr(BCD) = (\text{vec } B')'(I \otimes C)(\text{vec } D),$$

(A.9) $$tr(BX'CXD) = (\text{vec } X)'(DB \otimes C')(\text{vec } X),$$

(A.10) $$(A \otimes B)\, K = K(B \otimes A),$$

where K is the commutation matrix defined by vec(C) = K vec(C') for any matrix C.

Lemma A.6: Let A = (B,C) be a $p \times (q_1 + q_2)$ matrix. Then

(A.11) $$\bar{P}_A = \bar{P}_B - \bar{P}_B C\,(C'\bar{P}_B C)^{-1} C'\bar{P}_B ,$$

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

Lemma A.7: For non-negative λ_i $(i = 1,\cdots,p)$,

(A.12) $$\sum_{i=1}^{p} \frac{\lambda_i}{1+\lambda_i} \leqq \log \prod_{i=1}^{p} (1+\lambda_i) \leqq \sum_{i=1}^{p} \lambda_i .$$

Appendix B

In Appendix B, we give the proofs of Theorems 5.1, 5.2, and 5.3 in Section 5.

Proof of Theorem 5.1: The conclusion holds if

(B.1) $$(\text{vec } A)'\,\text{vec}\{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} z_t v_t' \} = \frac{1}{\sqrt{T}}\, tr \sum_{t=1}^{T} z_t v_t' A$$

$$= \frac{1}{\sqrt{T}} \operatorname{tr} \sum_{t=1}^{T} z_t A v_t' \xrightarrow{L} N(0, \Omega \otimes M)$$

for every A. Let $x_t = A' z_t$. Then

$$\text{(B.2)} \qquad \frac{1}{T} \sum_{t=1}^{T} x_t x_t' \xrightarrow{p} A' M A = D ,$$

say. We want to show that

$$\text{(B.3)} \qquad \frac{1}{\sqrt{T}} \sum_{t=1}^{T} x_t' v_t \xrightarrow{L} N(0, \operatorname{tr} \Omega D) .$$

Because of (5.1),

$$\text{(B.4)} \qquad \frac{1}{T+1} \sum_{t=1}^{T+1} x_t' x_t - \frac{1}{T} \sum_{t=1}^{T} x_t' x_t = \frac{1}{T+1} x_{T+1}' x_{T+1} - \frac{1}{T(T+1)} \sum_{t=1}^{T} x_t' x_t \xrightarrow{p} 0$$

and hence $x_{T+1}' x_{T+1} / (T+1) \xrightarrow{p} 0$, which implies

$$\text{(B.5)} \qquad \max_{t=1,\ldots,T} \frac{x_t' x_t}{T} \xrightarrow{p} 0 .$$

See Lemma 2.6. of Anderson (1971), for instance. Let

$$\text{(B.6)} \qquad w_{Tt} = \frac{1}{\sqrt{T}} x_t I(||x_t||^2 \leq T) , \quad t=1,\ldots,T, \quad T=1,2,\ldots,$$

where I(.) is the indicator function. Then

$$\text{(B.7)} \qquad \Pr\{ w_{Tt} = \frac{1}{\sqrt{T}} x_t, \ t=1,\ldots,T \} \longrightarrow 1$$

as $T \longrightarrow \infty$. Then

$$\text{(B.8)} \qquad E\{w_{Tt}' v_t | F_{t-1}\} = 0 ,$$

$$\text{(B.9)} \qquad \sum_{t=1}^{T} E\{(w_{Tt}' v_t)^2 | F_{t-1}\} = \sum_{t=1}^{T} w_{Tt}' \Omega w_{Tt} \xrightarrow{p} \operatorname{tr} \Omega D .$$

Now consider

$$\text{(B.10)} \qquad \sum_{t=1}^{T} E\{(w_{Tt}' v_t)^2 I[(w_{Tt}' v_t)^2 > \delta] | F_{t-1}\}$$

$$= \sum_{t=1}^{T} w_{Tt}'w_{Tt}E\{(\frac{w_{Tt}}{||w_{Tt}||}v_t)^2 I((\frac{w_{Tt}}{||w_{Tt}||}v_t)^2 > \frac{\delta}{||w_{Tt}||^2})|F_{t-1}\}$$

$$\leq \sum_{t=1}^{T} w_{Tt}'w_{Tt}E\{v_t'v_t I(v_t'v_t > \frac{\delta}{||w_{Tt}||^2})|F_{t-1}\} .$$

But (B.5), which is equivalent to max $_{t=1,\ldots,T}||w_{Tt}||^2 \longrightarrow 0$, implies that for any $\varepsilon > 0$ if T is sufficiently large the probability is arbitrarily close to 1 that $||w_{Tt}||^2 < \varepsilon$. Hence, for sufficiently large T and with arbitrarily high probability the right-hand side of (B.10) is less than or equal to

$$\sum_{t=1}^{T} w_{Tt}'w_{Tt}E\{v_t'v_t I(v_t'v_t > \frac{\delta}{\varepsilon})|F_{t-1}\} \leq \frac{1}{T}\sum_{t=1}^{T} z_t'z_t \sup_t E\{v_t'v_t I(v_t'v_t > \frac{\delta}{\varepsilon})|F_{t-1}\}$$

which is arbitrarily small by Assumption (iii). Hence

$$\text{(B.11)} \qquad \sum_{t=1}^{T} E\{(w_{Tt}'v_t)^2 I((w_{Tt}'v_t)^2 > \delta)|F_{t-1}\} \xrightarrow{p} 0 .$$

The theorem follows from Theorem 2.2 of Dvoretzky (1972) or Colloraly 3.1 of Hall and Heyde (1980). ■

<u>Lemma B.1</u>: Let $\hat{B}$ be a matrix consisting of the characteristic vectors of (3.11) and normalize B such that $B_1 = I_{G_0}$ in Section 3. Then under the assumptions we made in Theorem 5.1, $\hat{B} \longrightarrow B$ in probability as $T \longrightarrow \infty$ and

$$\text{vec } \sqrt{T}\begin{pmatrix} \hat{B}_2 - B_2 \\ \hat{\Gamma} - \Gamma \end{pmatrix}$$

is asymptotically distributed as $N(0, \Sigma\otimes(D'MD)^{-1})$.

<u>Sketch of Proof of Lemma B.1</u>: The proof of this lemma is quite similar to that of the consistency and the asymptotic normality of the limited information maximum likelihood estimator of B when $G_0 = G_1 = 1$. Thus we only sketch the proof. Because of (B.12), this problem reduces to the proof of the consistency and the asymptotic normality of the two-stage least squares estimator. Then we apply Theorem 5.1 for this case. ■

<u>Proof of Theorem 5.2:</u> The proof consists of showing the following three propositions (B.12), (B.13), and (B.14).

(B.12) $$T^{1-\varepsilon}\lambda_i \xrightarrow{p} 0 \quad (i = 1, 2, \cdots, G_0)$$

as $T \to \infty$. Using (B.12), we have

(B.13) $$\frac{1}{T}\hat{B}'Y\bar{P}_Z Y\hat{B} \xrightarrow{p} \Sigma$$

as $T \to \infty$, where $\hat{B}$ is the maximum likelihood estimator of B under the block identifiability restrictions. By the use of Theorem 5.1, we have

(B.14) $$\hat{B}'Y'(P_Z - P_{Z_1})Y\hat{B} \xrightarrow{L} W_{G_0}(\Theta_1, L_1, \Sigma)$$

as $T \to \infty$. In the above notation, $W_{G_0}(\Theta_1, L_1, \Sigma)$ is the G_0-dimensional noncentral Wishart distribution with L_1 degrees of freedom and the noncentrality matrix is given by (5.7).

<u>Derivation of (B.12)</u>: Let $\mu_i = 1 + \lambda_i$ $(i = 1, \cdots, G_1)$. From (3.11), we have

$$\text{(B.15)} \qquad \prod_{i=1}^{G_1} \mu_i = \min_B \frac{| B'Y'\bar{P}_{Z_1} Y B |}{| B'Y'\bar{P}_Z Y B |} \leq \frac{| B'Y'\bar{P}_{Z_1} Y B |}{| B'Y'\bar{P}_Z Y B |} \equiv \ell .$$

Under the local alternative (5.5), we write

$$\text{(B.16)} \qquad \ell = \frac{| U'\bar{P}_{Z_1} U^* |}{| U'\bar{P}_Z U |} = \frac{| W_1 |}{| W_2 |} ,$$

where $U^* = U + (1/\sqrt{T})\, Z'\xi_1$. Let ℓ_i $(i = 1, \cdots, G_1)$ be the characteristic roots of

$$\text{(B.17)} \qquad | \frac{1}{T} W_1 - \ell_i \frac{1}{T} W_2 | = 0.$$

Using Theorem 5.1 and Lemma B.1, as $T \to \infty$

$$\text{(B.18)} \qquad \frac{1}{T} W_2 \xrightarrow{p} \Sigma ,$$

$$\text{(B.19)} \qquad W_1 - W_2 \xrightarrow{L} W_{G_1}(\Theta_3, K_2, \Sigma) ,$$

where the noncentrality matrix

$$\text{(B.20)} \qquad \Theta_3 = \xi_1' \{ M - \binom{M_{11}}{M_{21}} M_{11}^{-1} (M_{11}, M_{12}) \} \xi_1 .$$

Since Anderson (1951a) and (1963) have shown that $T(\ell_i - 1)$ converge to the characteristic roots of the noncentral Wishart matrix $W_{G_1}(\Theta_3, K_2, \Sigma)$, $T^{1-\varepsilon}(\ell_i - 1)$ converges to zero in probability for any $0 < \varepsilon < 1$. Then

$$(B.21)\qquad T^{1-\varepsilon}(\,\Omega - 1\,) = T^{1-\varepsilon}(\,\prod_{i=1}^{G_1} \Omega_i - 1\,)$$

$$= T^{1-\varepsilon}\{\prod_{i=1}^{G_1} (\Omega_i - 1) + \sum_{j=1}^{G_1} \sum_{i \neq j} (\Omega_i - 1)(\Omega_j - 1) + \cdots + \sum_{i=1}^{G_1} (\Omega_i - 1)\}$$

converge to zeros in probability. Because $\lambda_i \geq 0$ ($i = 1, \cdots, G_1$) we obtain (B.12). ■

<u>Derivation of (B.13)</u> Using Lemma B.1, $\hat{B} \xrightarrow{p} B$ as $T \to +\infty$. Noticing that

$$(B.22)\qquad \frac{1}{T} V'\bar{P}_Z V \xrightarrow{p} \Omega ,$$

we obtain (B.8). ■

<u>Derivation of (B.14)</u>: Using the fact that $\hat{B} = B + (\hat{B} - B)$, the right hand side of (B.14) is decomposed as

$$(B.23)\qquad B'Y'(\,P_Z - P_{Z_1})YB + B'Y'(\,P_Z - P_{Z_1})Y(\,\hat{B} - B\,)$$

$$+ (\,\hat{B} - B\,)'Y'(\,P_Z - P_{Z_1}\,)\,YB + (\,\hat{B} - B\,)'Y'(\,P_Z - P_{Z_1})Y(\,\hat{B} - B\,).$$

From (5.5), the first term of (B.23) becomes

$$(B.24)\qquad U^{*\prime}(\,P_Z - P_{Z_1})U^{*}.$$

By the standardization of $\hat{B}_1 = B_1 = I_{K_1}$, we write

$$(B.25)\qquad Y(\,\hat{B} - B\,) = (Y_2,\, Z_1)\begin{bmatrix} -(\,\hat{B}_2 - B_2) \\ -(\,\hat{\Gamma} - \Gamma\,) \end{bmatrix}$$

$$= - \left[\frac{1}{\sqrt{T}}ZD + \frac{1}{\sqrt{T}}(V_2, O)\right]\left[\begin{matrix}\sqrt{T}(\hat{B}_2 - B_2) \\ \sqrt{T}(\hat{\Gamma} - \Gamma)\end{matrix}\right] .$$

Using Lemma B.1, the second term of (B.25) is asymptotically equivalent to

$$\text{(B.26)} \qquad (D'M\, D)^{-1}D'\frac{1}{\sqrt{T}}Z'U^{*}$$

and so $Y(\hat{B} - B)$ is asymptotically equivalent to

$$\text{(B.27)} \qquad -\left(\frac{1}{\sqrt{T}}Z'D\right)\left(D'\frac{1}{\sqrt{T}}Z'Z\,D\right)^{-1}\left(\frac{1}{\sqrt{T}}Z'D\right)'U^{*}.$$

Let denote $E = ZD$. Then $P_Z P_E = P_{Z_1}$ and $P_{Z_1}P_Z = P_E$. Thus, the second term of (B.23) is asymptotically equivalent to

$$\text{(B.28)} \qquad -U^{*\prime}(P_Z - P_{Z_1})P_E U^{*} = -U^{*\prime}(P_E - P_{Z_1})U^{*}.$$

By a similar consideration on the third and fourth terms of (B.23), we find that (B.23) is asymptotically equivalent to

$$\text{(B.29)} \qquad U^{*\prime}(P_Z - P_E)U^{*} = \left[\frac{1}{\sqrt{T}}U^{*\prime}Z(\tfrac{1}{T}Z'Z)^{-1/2}\right]\left[I_K - (\tfrac{1}{T}Z'Z)^{1/2}D(\tfrac{1}{T}D'Z'Z\,D)^{-1}\right.$$
$$\left.\times\, D'(\tfrac{1}{T}Z'Z)^{1/2}\right]\left[(\tfrac{1}{T}Z'Z)^{-1/2}\frac{1}{\sqrt{T}}Z'U^{*}\right] .$$

Since the rank of the middle parenthesis of (B.29) is $K - (G_2 + K_1)$, we obtain (B.23). ∎

Extending the method by Anderson (1951b), we obtain the asymptotic distribution of the characteristic roots of the noncentral Whishart random variables. This completes the proof of Theorem 5.2. ∎

<u>Proof of Theorem 5.3</u>: Here we give the proof for the case of $G_0 = G_1$. However, it can be extended to the general case by modifying the proof given below. Let

(B.30) $V_1^* = V_1 - V_2\Omega_{22}^{-1}\Omega_{21}$, $Y_1^* = Y_1 - V_1^*$.

Then each row of V_1^* is conditionally uncorrelated with each row of V_2. Under the local alternative (5.7),

$$Y_1^* = Z\,\Pi_1 + V_2\Omega_{22}^{-1}\Omega_{21} \tag{B.31}$$

$$= (Y_2,\ Z_1)\begin{pmatrix} B_2 \\ \Gamma \end{pmatrix} + \frac{1}{\sqrt{T}} V_2\Omega_{22}^{-1}\eta_1 .$$

Since the first term of Y_1^* is orthogonal to L and X, the noncentrality matrix is given by

$$\Theta_2 = \operatorname*{plim}_{T\to\infty} \frac{1}{T}\, \eta_1'\, \Omega_{22}^{-1}V_2'(P_X - P_L)\, V_2\Omega_{22}^{-1}\eta_1 . \tag{B.32}$$

We notice that

$$P = \operatorname*{plim}_{T\to\infty} \frac{1}{T} X'X \tag{B.33}$$

$$= \begin{pmatrix} D'M\,D & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \Omega_{22} & 0 & \Omega_{22} \\ 0 & 0 & 0 \\ \Omega_{22} & 0 & \Omega_{22} \end{pmatrix},$$

$$\operatorname*{plim}_{T\to\infty} \frac{1}{T} X'V_2 = J_5\Omega_{22} , \tag{B.34}$$

where $J_5' = (I_{G_2},\ 0,\ I_{G_2})$ is a $G_2 \times (G_2 + K_1 + G_2)$ choice matrix. We also note that $L = X\,J_6$, where $J_6' = (I_{G_2+K_1},\ 0)$ is a $(G_2 + K_1)\times(G_2 + K_1 + G_2)$ choice matrix. Then

(B.35) $\qquad \theta_2 = \eta_1' J_5' [\, P^{-1} - J_6 (J_6' P J_6)^{-1} J_6' \,] \, J_5 \eta_1 .$

Using the inversion formula of partitioned matrices, we have

(B.36) $\qquad \theta_2 = \eta_1' [\, \Omega_{22}^{-1} - (\, \Pi_{22}' M_{22.1} \Pi_{22} + \Omega_{22})^{-1}] \, \eta_1$

$$= \eta_1' [\, \Omega_{22} + \Omega_{22} (\, \Pi_{22}' M_{22.1} \Pi_{22})^{-1} \Omega_{22}]^{-1} \eta_1 \; .$$

Since the rank of $P_X - P_L$ is $2G_2 + K_1 - (\, G_2 + K_1) = G_2$ and V_1^* is conditionally uncorrelated with V_2,

(B.37) $\qquad F = (\, Y_1^* + V_1)' (\, P_X - P_L)(\, Y_1^* + V_1^*)$

is asymptotically distributed with $W_{G_2}(\theta_2,\; G_2,\; \Omega_{11.2})$. Under the local alternative (5.8), the estimator (4.36) is written as

(B.38) $\qquad \hat{\Sigma} = \frac{1}{T} \, (\, V_1^* + \frac{1}{\sqrt{T}} \, V_2 \Omega_{22}^{-1} \eta_1)' \bar{P}_X (\, V_1^* + \frac{1}{\sqrt{T}} \, V_2 \Omega_{22}^{-1} \eta_1).$

As $T \longrightarrow \infty$, $\hat{\Sigma} \longrightarrow \Omega_{11.2}$ in probability, which is the covariance matrix of each row of V_1^*. By a similar argument, (4.35) and (4.38) converge to $\Omega_{11.2}$ in probability as $T \longrightarrow \infty$. Thus we obtain the asymptotic distribution of LM_3 statistics.

We now turn to derive the asymptotic distribution of W_2 and LR_3. We note that under the local alternative (5.8),

(B.39) $\qquad U = V_1^* B_1 + \frac{1}{\sqrt{T}} \, V_2 \Omega_{22}^{-1} \eta_1 \; .$

We shall first obtain the asymptotic distribution of

$$\text{(B.40)} \qquad \mathrm{vec}\left(\frac{1}{\sqrt{T}} R'U \right) = \mathrm{vec}\left(\frac{1}{\sqrt{T}} R'V_1^* B_1 \right) + \mathrm{vec}\left(\frac{1}{T} R'V_2 \Omega_{22}^{-1} \eta_1 \right).$$

Noting that $R = Z[\Pi_2, I_K] + [V_2, 0]$ and (5.4), the first term of (B.40) is asymptotically distributed as $N[0, \Omega_{11.2} \otimes Q]$. The second term of (B.40) converges to $\mathrm{vec}(\eta_1)$ in probability as $T \longrightarrow \infty$. Hence the asymptotic distribution of (B.40) can be denoted by $\mathrm{vec}(C) \sim N[\mathrm{vec}(\eta_1), \Omega_{11.2} \otimes Q]$. Because of $L = R J_5 = X J_6$, we have

$$\text{(B.41)} \qquad \mathrm{vec}\left[\frac{1}{\sqrt{T}} X'V_1^*\right] = \mathrm{vec}\left\{\begin{bmatrix} D' \\ 0 \end{bmatrix} \frac{1}{\sqrt{T}} Z'V_1^*\right\} + \mathrm{vec}\left\{\frac{1}{\sqrt{T}}\begin{pmatrix} V_2' \\ 0 \\ V_2'P_Z \end{pmatrix} - V_1^*\right\},$$

which is asymptotically distributed as $N[0, \Omega_{11.2} \otimes P]$ and $\mathrm{vec}\left[\frac{1}{\sqrt{T}} L'V_1^*\right]$ is asymptotically distributed as $N[0, \Omega_{11.2} \otimes J_6'PJ_6]$. Therefore, the asymptotic distribution of $Y_1'(P_R - P_L)Y_1$ is equivalent to

$$\text{(B.42)} \qquad U'(P_R - P_L)U$$

$$= \left(\frac{1}{\sqrt{T}} U'R\right)\left(\frac{1}{T} R'R\right)^{-1/2}\left[I_{G_2+K} - \left(\frac{1}{T} R'R\right)^{1/2} J_6 \left(J_6' \frac{1}{T} R'RJ_6\right)^{-1} J_6'\right.$$

$$\left. \times \left(\frac{1}{T} R'R\right)^{1/2}\right]\left(\frac{1}{T} R'R\right)^{-1/2}\left(\frac{1}{\sqrt{T}} R'U\right),$$

which is the same as the distribution of

$$\text{(B.43)} \qquad A_1 = C'Q^{-\frac{1}{2}}\left[I_{G_2+K} - Q^{\frac{1}{2}} J_6 (J_6'Q J_6)^{-1} J'Q^{\frac{1}{2}}\right] Q^{-\frac{1}{2}} C .$$

Since the rank of A_1 is $G_2 + K - (G_2 + K_1) = K_2$ and (4.38) is a consistent estimator of $\Omega_{11.2}$, we obtain the asymptotic distribution of W_2 as stated in Theorem 5.3.

Let $J_7' = [0, I_K]$ be a $K\times(G_2 + K)$ choice matrix and $Z = R J_7$. Then from the derivation of (B.12) Appendix B, the asymptotic distribution of (B.13) is equvalent to that of $U'(P_Z - P_E) U$, which is written as

(B.44) $$\left(\frac{1}{\sqrt{T}}U'R\right)J_7\left\{\left(J_7'\left(\frac{1}{T}R'R\right)J_7\right)^{-1} - D\left(D'J_7'\left(\frac{1}{T}R'R\right)J_7D\right)^{-1}D'\right\}J_7'\left(\frac{1}{\sqrt{T}} R'U\right),$$

where D is given in (5.7) and $E = Z D$. As $T \longrightarrow \infty$, (B.44) converges to the distribution of A_2,

(B.45) $$A_2 = C'J_7M^{-\frac{1}{2}}\left(I_K - M^{\frac{1}{2}}D(D'M D)^{-1}D'M^{\frac{1}{2}}\right) M^{\frac{1}{2}}J_7'C$$

$$= C'Q^{\frac{1}{2}}\left\{Q^{\frac{1}{2}}\begin{pmatrix} 0 & 0 \\ 0 & M^{-1}\end{pmatrix} Q^{\frac{1}{2}} - Q^{\frac{1}{2}}\begin{pmatrix} 0 & 0 \\ 0 & D(D'MD)^{-1}D'\end{pmatrix}Q^{\frac{1}{2}}\right\}Q^{-\frac{1}{2}}C .$$

Define P_1 and P_2 such that

(B.46) $$A_i = C'Q^{-1/2} P_iQ^{1/2}C$$

for $i = 1$ and 2 in (B.43) and (B.45). Subsequently, by some calculation, we can show that $P_1^2 = P_1$, $P_2^2 = P_2$, $P_1P_2 = P_2P_1 = P_2$. Since $\mathrm{tr}P_1 = G_2 + K - (G_2 + K_1)$, $\mathrm{tr}P_2 = K - (G_2 + K_1)$, and $\mathrm{tr}P_1 - \mathrm{tr}P_2 = G_2$, we know that $A_1 - A_2 \sim W_{G_1}(\theta_2, G_2, \Omega_{11.2})$. Because of (5.4), the asymptotic distribution of LR_3 is equivalent to that of

(B.47) $$T \sum_{i=1}^{G_1} (\lambda_i^* - \lambda_i).$$

Hence we obtain the asymptotic distribution of LR_3 as stated in Theorem 5.3. The asymptotic distribution of W_3 can be obtained by a similar argument. ■

References

Anderson, T.W. [1951a], "The Asymptotic Distribution of Certain Characteristic Roots and Vectors," J.Neyman ed., Proceedings of Second Berkeley Symposium on Mathematical Statistics and Probability, 103-130, University of California Press.

Anderson, T.W. [1951b], "Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions," Annals of Mathematical Statistics, Vol.34, 327-351.

Anderson, T.W. [1963], "Asymptotic Theory for Principal Component Analysis," Annals of Mathematical Statistics, Vol.34, 122-148.

Anderson, T.W. [1971], The Statistical Analysis of Time Series, John Wiley & Sons.

Anderson, T.W. [1984], An Introduction to Multivariate Statistical Analysis, Second Edition, John Wiley & Sons.

Anderson, T.W. [1987], "Multivariate Linear Relations," Unpublished Manuscript.

Anderson, T.W. and H. Rubin [1949], "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," Annals of Statistical Mathematics, Vol.21, 570-582.

Basmann, R.L. [1960], "On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics," Journal of the American Statistical Association, Vol.55, 650-659.

Brown, B.M. [1971], "Martingale Central Limit Theorems", Annals of Statistics, Vol.42, 59-66.

Byron, R.P. [1972], "Testing for Misspecification in Econometric Systems Using Full Information," International Economic Review, Vol.13, No.3, 745-756.

Byron, R.P. [1974], "Testing Structural Specification Using the Unrestricted Reduced Form," Econometrica, Vol.42, No.5, 869-883.

Dvoretzky, A. (1972), "Asymptotic Normality of Sums of Dependent Random Variables, " Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, 2, University of California Press, 513-535.

Durbin, J. [1954], "Errors in Variables," Review of the International Statistical Institute, Vol.22, 23-32.
Vol.44, No.3, 475-477.

Engle, R. F. (1984), "Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics," In HandBook of Econometrics Vo.2, Z. Griliches and Intrigator eds., North-Holland.

Engle, R., D. Hendry, and H. Richard [1983], "Exogeneity," Econometrica, Vol.51, 277-304.

Hall, P., and Heyde, C.C. (1980), Martingale Limit Theory and Its Application, Academic Press, New York.

Hausman, J.A. [1978], "Specification Tests in Econometrics," Econometrica, Vol.46, No.6, 1251-1271.

Hwang, H. [1980a], "Test of Independence between a Subset of Stochastic Regressors and Disturbances," International Economic Review, Vol.21, No.3, 749-760.

Hwang, H. [1980b]. "A Comparison of Tests of Overidentifying Restrictions," Econometrica, Vol.48, No.7, 1821-1825.

Kadane, J.B. [1974], "Testing a Subset of the Overidentifying Restrictions," Econometrica, Vol.42, No.5, 853-867.

Kadane, J.B., and T.W. Anderson [1977], "A Comment on the Test of Overidentifying Restrictions," Econometrica, Vol.45, No.4, 1027-1031.

Kariya, T., and J. Hodoshima [1980], "Finite Sample Properties of the Tests for Independence in Structural Systems and the LRT," Economic Studies Quarterly, Vol.31, No.1, 45-56.

Koopmans, T.C. and Hood, W.C. [1953], "The Estimation of Simultaneous Linear Economic Relationships," Ch.6, Studies in Econometric Method ed. by Hood and Koopmans, Yale University Press.

Kunitomo, N. [1987], "Approximate Distributions and Power of the Test Statistics for Overidentifying Restrictions in s System of Simultaneous Equations," Econometric Theory, Forthcoming.

Kunitomo, N., K. Morimune, and Y. Tsukuda [1983], "Asymptotic Expansions of the Distributions of the Test Statistics for Overidentifying Restrictions in a System of Simultaneous Equation," International Economic Review, Vol.2, 199-215.

Lai, Tze Leung, and Robbins, Herbert (1981), "Consistency and Asymptotic Efficiency of Slope Estimates in Stochastic Approximation Schemes," Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete, 56, 329-360.

Nakamura, A., and M. Nakamura [1982], "On the Relationships among Several Specification Error Tests Presented by Durbin, Wu and Hausman," Econometrica, Vol.49, No.6, 1583-1588.

Rao, C.R. [1973], Linear Statistical Inference and its Applications, 2nd Edition, John Wiley & Sons.

Revankar, N. [1978], "Asymptotic Relative Efficiency Analysis of Certain Tests of Independence in Structural Systems," International Economic Review, Vol.1, No.1, 165-179.

Revankar, N.S., and M.J. Hartley [1973], "An Independence Test and Conditional Unbiased Predictions in the Context of Simultaneous Equation Systems," International Economic Review, Vol.14, No.3, 625-631.

Wegge, L.L. [1978], "Constrained Indirect Least Squares Estimators," Econometrica, Vol.46, No.2, 435-449.

Wu, D-M. [1973], "Alternative Tests of Independence between Stochastic Regressors and Disturbances," Econometrica, Vol.41, No.4, 733-750.

Wu, D-M. [1973], "Alternative Tests of Independence between Stochastic Regressors and Disturbances : Finite Sample Results," Econometrica, Vol.42, No.3, 529-546.