Jean-Marie Dufour [1]

LINEAR WALD METHODS FOR INFERENCE ON COVARIANCES AND WEAK EXOGENEITY TESTS IN STRUCTURAL EQUATIONS

ABSTRACT

Inference about the vector of covariances between the stochastic explanatory variables and the disturbance term of a structural equation is an important problem in econometrics. For example, one may wish to test the independence between stochastic explanatory variables and the disturbance term. Tests for the hypothesis of independence between the full vector of stochastic explanatory variables and the disturbance have been proposed by several authors. When more than one stochastic explanatory variable is involved, it can be of interest to determine whether all of them are independent of the disturbance and, if not, which ones are. We develop simple large-sample methods which allow us to construct confidence regions and test hypotheses concerning any vector of linear transformations of the covariances between the stochastic explanatory variables and the disturbance of a structural equation. The main method described is a generalized Wald procedure which simply requires two linear regressions. No nonlinear estimation is needed. Consistent tests for weak exogeneity hypotheses are derived as special cases.

1. INTRODUCTION

Inference about the vector of covariances between the stochastic explanatory variables and the disturbance term of a structural equation is an important problem in econometrics. For example, one may wish to test whether a set of stochastic explanatory variables are statistically independent of the

[1] Département de sciences économiques, Université de Montréal, C.P. 6128, succursale A, Montréal, Québec, H3C 3J7

I. B. MacNeill and G. J. Umphrey (eds.), Time Series and Econometric Modelling, 317–338.

disturbance of a structural equation, i.e., whether the stochastic explanatory variables considered can be treated as "exogenous" (or "predetermined"). In particular, it is well known that independence between explanatory variables and disturbances is usually needed to ensure that standard inference procedures, like ordinary least squares or F-tests, are appropriate in linear models. Furthermore, a number of economic hypotheses can be formulated in terms of the independence between stochastic explanatory variables and disturbances.[2]

Tests for the hypothesis of independence between a vector of stochastic explanatory variables and a disturbance term were proposed by several authors; see Durbin (1954), Wu (1973, 1974), Revankar and Hartley (1973), Farebrother (1976), Hausman (1978), Revankar (1978), Kariya and Hodoshima (1980), Richard (1980), and Holly and Sargan (1982).[3] These articles deal especially with the problem of testing whether the full vector of stochastic explanatory variables is independent of the disturbance. When more than one stochastic explanatory variable are involved, it is often necessary to determine whether all of them are independent of the disturbances and, if not, which ones are. This can be useful, for example, to check the specification of a simultaneous equation model (e.g., block recursiveness assumptions) and to get more efficient estimators for such models.

Tests for the hypothesis of independence between a subset of stochastic explanatory variables and the disturbance in a structural equation have been proposed by a number of authors: Hwang (1980) and Smith (1984) studied likelihood ratio (LR) tests, Hausman and Taylor (1981a), Spencer and Berk (1981) and Wu (1983b) proposed extensions of the tests previously studied by Wu (1973) and by Hausman (1978), while Engle (1982) derived Lagrange multiplier (LM) tests.

Each of these procedures has important drawbacks, either practical or theoretical. Some of them require nonlinear estimation, e.g., LR tests and certain forms of the LM tests. All of them require a separate estimation for each null hypothesis tested. It is difficult to construct confidence intervals for the covariances of interest because covariance estimates or their standard errors are not typically produced.

---

[2] See Wu (1973). For an example of a structural equation where the stochastic explanatory variables can be treated as "exogenous", see Zellner *et al.* (1966).

[3] Further useful discussions and extensions of these tests are provided by Bronsard and Salvas-Bronsard (1984), Engle (1982, 1984), Gouriéroux and Trognon (1984), Hausman and Taylor (1980, 1981a,b), Holly (1980, 1982a,b, 1983), Holly and Monfort (1982), Nakamura and Nakamura (1980, 1981), Plosser *et al.* (1982), Reynolds (1982), Riess (1983), Ruud (1984), Tsurumi and Shiba (1984), Turkington (1980), White (1982), and Wu (1983a).

Hausman-type tests are better viewed as consistency tests. By comparing an efficient estimator under the null hypothesis with a consistent estimator under the alternative hypothesis, one checks whether the constrained estimator is consistent (see Holly, 1982a,b; Hausman and Taylor, 1980, 1981b). When testing exogeneity, this is not equivalent to testing independence between possible endogenous variables and the disturbance term: the condition tested is weaker (unless special assumptions hold) and the test may not be consistent. (This is easy to see from Hausman and Taylor (1980, 1981b) and Wu (1983b).) Even though this condition may be sufficient to ensure the consistency or the efficiency of the constrained estimator, it is not generally sufficient to guarantee the validity of inferences obtained from the model by treating the regressors whose exogeneity is in doubt as being exogenous: tests and confidence intervals pertaining to the various coefficients of the model may not have the correct levels, even asymptotically.[4] In many if not most practical situations, the relevant hypothesis is whether one can treat some stochastic explanatory variables as being exogenous for all purposes of inference (i.e., the independence assumption).

In this paper, we consider a single linear structural equation and develop a class of linear Wald-type procedures which allow us to construct confidence regions as well as to test any set of linear restrictions on the vector of covariances between the stochastic explanatory variables and the disturbance term in the equation. Besides a set of instrumental regressions, all that is needed is a simple linear regression which yields consistent estimates of both the structural coefficients in the equation and the relevant vector of covariances. The asymptotic covariance matrix of the coefficients is then easily obtained. Using these results, one can test any set of linear restrictions on the covariances and construct confidence regions. Cross-restrictions between the structural coefficients and the covariances may also be tested. Special cases of this family of tests include tests of zero restrictions on the covariances, either for individual covariances or subvectors of covariances. In particular, one can compute in a routine way asymptotic "t-values" for each covariance, an especially convenient instrument to explore the recursiveness properties of a model. All the tests suggested are consistent.

Because they are based on consistent asymptotically normal estimators different from the maximum-likelihood estimators (Wald, 1943), the tests developed here should be viewed as generalized Wald tests rather than Wald tests in the usual sense (see Stroud, 1971; Szroeter, 1983). We will not need the information matrix associated with the maximum likelihood estimators. As we shall see below, the tests proposed can be obtained as a byproduct of

[4] See White (1982, p. 16), and Breusch and Mizon (comment to Ruud, 1984, p. 249).

the estimation of a structural equation by any instrumental-variable method (including two-stage least squares). They thus have a natural complementary with the latter estimation method.

In Section 2, we formulate the model considered and the assumptions used. In Section 3, we describe the procedures proposed and formulate the theorems underlying them. In Section 4, we examine three important special situations: the case where we want to test independence between the full vector of stochastic explanatory variables and the disturbance term, the one where a subset of stochastic explanatory variables is taken *a priori* as being exogenous and the case where the matrix of instruments includes all the fixed (or exogenous) regressors in the equation considered. In Section 5, we discuss econometric applications. Finally, in Section 6, we provide the proofs of the theorems.

2. FRAMEWORK

We consider the model described by the following assumptions.

ASSUMPTION 1:

$$y = Y\beta + Z_1\gamma + u, \tag{2.1}$$

where y is a $T \times 1$ random vector, u is a $T \times 1$ vector of disturbances, Y is a $T \times G$ matrix of stochastic explanatory variables, Z_1 is a $T \times K_1$ non-stochastic matrix of rank K_1, β and γ are $G \times 1$ and $K_1 \times 1$ vectors of coefficients.

ASSUMPTION 2:

$$Y = Z\Pi + V, \tag{2.2}$$

where Z is a $T \times K$ non-stochastic matrix of rank K, Π is a $K \times G$ matrix of coefficients and V is a $T \times G$ matrix of disturbances. Furthermore, we will denote by y_k, Π_k and w_k the kth columns of the matrices Y, Π and V respectively $(1 \le k \le G)$:

$$Y = [y_1, \ldots, y_G],\ \Pi = [\Pi_1, \ldots, \Pi_G],\ V = [w_1, \ldots, w_G]. \tag{2.3}$$

ASSUMPTION 3: The rows (u_t, v_t'), $t = 1, \ldots, T$, of the matrix $[u : V]$ are independent and normally distributed with mean zero and non-singular covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_{00} & \delta' \\ \delta & \Sigma_{22} \end{bmatrix}, \tag{2.4}$$

where

$$\delta = (\sigma_{01}, \sigma_{02}, \ldots, \sigma_{0G})',\ \Sigma_{22} = [\sigma_{jk}]_{\ j,\ k=1,\ldots,G}. \tag{2.5}$$

ASSUMPTION 4: Let $Z = [Z_{11} : Z_2]$ and $\Pi = [\Pi'_{11} : \Pi'_2]'$, where Z_2 is the $T \times K_2$ matrix of non-stochastic variables excluded from equation (2.1), Π_2 is the $G \times K_2$ corresponding matrix of coefficients, Z_{11} is a set of variables included in Z_1, so that $Y = Z_{11}\Pi_{11} + Z_2\Pi_2 + V$, $\text{rank}(\Pi_2) = G$ and $T > 2G + K_1$.
[This condition ensures identification of the coefficients of equation (2.1); see Fisher (1966, p. 53). Note also that Z_1 is not constrained to be a submatrix of Z.]

ASSUMPTION 5: The matrix $\frac{1}{T}Z'Z$ converges, as $T \to \infty$, to a positive definite matrix Q_z.

ASSUMPTION 6: The matrices $\frac{1}{T}Z_1'Z_1$ and $\frac{1}{T}Z'Z_1$ converge, as $T \to \infty$, to the matrices Q_{11} and Q_1 respectively, where Q_{11} is positive definite.

We want to test some set of linear restrictions on the parameter vector δ, i.e., a hypothesis of the type

$$H_0 : H\delta = d_0, \tag{2.6}$$

where H is an $r \times G$ matrix of rank $r \leq G$ and d_0 is a fixed $r \times 1$ vector. Since the vectors $(u_t, v_t')'$, $t = 1, \ldots, T$, are i.i.d. normal, we obtain by regressing u_t on v_t:

$$u = Va + e, \tag{2.7}$$

where $a = \Sigma_{22}^{-1}\delta$ and the vector e is $N[0, \sigma_e^2 I_T]$ independent of all the elements of V.[5] Then, substituting (2.7) into (2.1), we get

$$y = Y\beta + Z_1\gamma + Va + e, \tag{2.8}$$

where the disturbance vector e is independent of all the regressors. The latter formulation illustrates clearly that the existence of correlation between some of the regressors and the disturbance term in an econometric relationship, as generated, for example, by simultaneous equations, may be viewed as a problem of omitted variables. If the matrix V were observed, we would test any set of linear restrictions on the coefficients β, γ and a in equation (2.8) by standard F-tests, and these tests would be exact in small samples. In particular, linear hypotheses regarding the parameter vector a could be tested by using the least squares estimate $\hat{a}$ obtained from (2.8). Furthermore, if Σ_{22} were also known, the transformation $\delta = \Sigma_{22}a$ would allow one to test $H_0 : H\delta = d_0$ by a standard F-test. The difficulty, of course, is that neither V nor Σ_{22} are known. We also note that, although hypotheses regarding δ have relatively direct and intuitive interpretations (e.g., in

[5] This transformation is also used by Revankar (1973), Revankar and Hartley (1973) and Reynolds (1982).

terms of independence), the auxiliary parameter vector $a = \Sigma_{22}^{-1}\delta$ itself may be of interest. One may wish to test linear restrictions on a directly in the reparameterized model (2.8). In any event, we will deal with both problems.

We will first consider the problem of testing arbitrary linear restrictions on the parameter vector $\alpha = (\beta', \gamma', a')'$ and then restrictions on the covariance vector δ. In each case, we will first define a vector of linear consistent asymptotically normal estimators, derive the asymptotic covariance matrix and propose generalized Wald tests. In particular, we will derive the asymptotic distribution of the covariance estimator $\hat{\delta}$ under both the null and the alternative hypotheses. As a special case, it will be straightforward to test zero restrictions on δ, for example, $H_0 : \delta_1 = 0$ where $\delta = (\delta_1', \delta_2')'$. In the context of the model considered here, the hypothesis $\delta_1 = 0$ is equivalent to the independence between Y_1 and u, where $Y = [Y_1 : Y_2]$, or the weak exogeneity of Y_1 inside equation (2.1).[6] Further, from the same results, it is easy to construct a confidence region for any element or subvector of δ or α.

3. DESCRIPTION OF THE TESTS

In equation (2.8), replace the disturbance matrix V by the corresponding ordinary least squares (OLS) residuals

$$\hat{V} = Y - Z\hat{\Pi}, \tag{3.1}$$

where $\hat{\Pi} = (Z'Z)^{-1}Z'Y$. We obtain in this way the equivalent equation

$$y = Y\beta + Z_1\gamma + \hat{V}a + e^* = X\alpha + e^*, \tag{3.2}$$

where $X = [Y : Z_1 : \hat{V}]$, $\alpha = (\beta', \gamma', a')'$ and

$$e^* = Z(\hat{\Pi} - \Pi)a + e. \tag{3.3}$$

Also let

$$\hat{\Sigma}_{22} = \hat{V}'\hat{V}/T. \tag{3.4}$$

Under Assumptions 2 through 6, we have

$$\text{plim}\frac{Z'V}{T} = 0, \quad \text{plim } \hat{\Sigma}_{22} = \Sigma_{22} \tag{3.5}$$

[6] For a general discussion of exogeneity and related notions, see Engle *et al.* (1983).

(where plim refers to the probability limit as $T \to \infty$), hence

$$Q_x = \text{plim}\frac{X'X}{T} = \begin{bmatrix} \Pi' Q_z \Pi + \Sigma_{22} & \Pi' Q_1 & \Sigma_{22} \\ Q_1'\Pi & Q_{11} & 0' \\ \Sigma_{22} & 0 & \Sigma_{22} \end{bmatrix} \tag{3.6}$$

and

$$Q_{zx} = \text{plim}\frac{Z'X}{T} = [Q_z\Pi : Q_1 : 0], \tag{3.7}$$

where $\text{rank}(Q_x) \equiv L = 2G + K_1$. Consider the OLS estimate of α obtained from (3.2):

$$\hat{\alpha} = (X'X)^{-1}X'y. \tag{3.8}$$

Under the assumptions made, this estimate is unique with probability one. Further, the asymptotic distribution of $\hat{\alpha}$ is given by the following theorem. (The proofs of the theorems are given in Section 5.)

Theorem 1. Suppose that Assumptions 1 through 6 are satisfied, and let matrix Q_x defined in (3.6) be non-singular. Then the estimator $\hat{\alpha}$ given in (3.8) is consistent for α and $\sqrt{T}(\hat{\alpha} - \alpha)$ has a normal limiting distribution with mean zero and covariance matrix

$$\begin{aligned} \Sigma_\alpha &= Q_x^{-1}\left[\sigma_e^2 Q_x + \rho Q_{zx}' Q_z^{-1} Q_{zx}\right] Q_x^{-1} \\ &= \sigma_e^2 Q_x^{-1} + \rho Q_x^{-1} Q_{zx}' Q_z^{-1} Q_{zx} Q_x^{-1}, \end{aligned} \tag{3.9}$$

where Q_{zx} is given by (3.7) and

$$\rho = a'\Sigma_{22}a = \delta'\Sigma_{22}^{-1}\delta. \tag{3.10}$$

Further, the statistics

$$\hat{\sigma}_e^2 = (y - X\hat{\alpha})'(y - X\hat{\alpha})/T \tag{3.11}$$

and

$$\hat{\Sigma}_\alpha = \left(\frac{X'X}{T}\right)^{-1}\left[\hat{\sigma}_e^2\left(\frac{X'X}{T}\right) + \hat{\rho}\left(\frac{X'Z}{T}\right)\left(\frac{Z'Z}{T}\right)^{-1}\left(\frac{Z'X}{T}\right)\right]\left(\frac{X'X}{T}\right)^{-1} \tag{3.12}$$

are consistent estimators of σ_e^2 and Σ_α, where $\hat{\rho} = \hat{a}'\hat{\Sigma}_{22}\hat{a}$, $\hat{\Sigma}_{22} = \hat{V}'\hat{V}/T$ and $\hat{a}$ is the estimate of a from $\hat{\alpha}$.

We can test any set of linear restrictions on the vector α, such as $M\alpha = m_0$, where M is a $\nu \times L$ matrix of rank $\nu \leq L$ and m_0 is a $\nu \times 1$ fixed vector, by using a critical region of the form $\{S(M, m_0) \geq c\}$, where

$$S(M, m_0) = T(M\hat{\alpha} - m_0)'(M\hat{\Sigma}_\alpha M')^{-1}(M\hat{\alpha} - m_0) \tag{3.13}$$

and c is a constant which depends on the level of the test. The asymptotic distribution of the test statistic $S(M, m_0)$ is chi-square with ν degrees of freedom under the null hypothesis.

Since the coefficient a is of special interest here, it will be useful to summarize the asymptotic properties of $\hat{a}$ by the following corollary.

Corollary 1.1. Under the assumptions of Theorem 1, the subvector $\hat{a}$ of $\hat{\alpha} = (\hat{\beta}', \hat{\gamma}', \hat{a}')'$ is a consistent estimator of a and $\sqrt{T}(\hat{a} - a)$ has a normal limiting distribution with mean zero and covariance matrix:

$$\Sigma_a = A_2 \left[\sigma_e^2 Q_x + \rho Q_{zx}' Q_z^{-1} Q_{zx}\right] A_2', \tag{3.14}$$

where $A_2 = \text{plim}(C_2)$ and C_2 is the $G \times (2G + K_1)$ matrix such that

$$\left(\frac{X'X}{T}\right)^{-1} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}. \tag{3.15}$$

Further, the submatrix

$$\hat{\Sigma}_a = C_2 \left[\hat{\sigma}_e^2 \left(\frac{X'X}{T}\right) + \hat{\rho} \left(\frac{X'Z}{T}\right) \left(\frac{Z'Z}{T}\right)^{-1} \left(\frac{Z'X}{T}\right)\right] C_2' \tag{3.16}$$

in (3.12) is a consistent estimator of Σ_a.

Of course, tests of linear restrictions on a are special cases of the tests given by (3.13). However, if our interest lies in δ rather than $a = \Sigma_{22}^{-1}\delta$, the estimator directly relevant to us is not $\hat{a}$. We need an estimator of δ. Since $\hat{a}$ and $\hat{\Sigma}_{22}$ are consistent estimators of a and Σ_{22}, $\hat{\delta} = \hat{\Sigma}_{22}\hat{a}$ is a consistent estimator of δ. The asymptotic distribution of δ is given by the following theorem.

Theorem 2. Under the assumptions of Theorem 1, the estimator $\hat{\delta} = \hat{\Sigma}_{22}\hat{a}$ is consistent for δ and the vector $\sqrt{T}(\hat{\delta} - \delta)$ has a normal limiting distribution, as $T \to \infty$, with mean zero and covariance matrix

$$\Sigma_\delta = \Sigma_{22}\Sigma_a\Sigma_{22} + \rho\Sigma_{22} + \delta\delta', \tag{3.17}$$

where Σ_a is given by (3.16). Further, a consistent estimator of Σ_δ is provided by

$$\hat{\Sigma}_\delta = \hat{\Sigma}_{22}\hat{\Sigma}_a\hat{\Sigma}_{22} + \hat{\rho}\hat{\Sigma}_{22} + \hat{\delta}\hat{\delta}', \tag{3.18}$$

where $\hat{\rho}$ and $\hat{\Sigma}_{22}$ are defined in Theorem 1.

Consequently, we can test the hypothesis $H_0 : H\delta = d_0$, where H is a $r \times G$ matrix of rank $r \leq G$ and d_0 is a fixed $r \times 1$ vector, by using a critical region of the form $\{W(H, d_0) \geq c\}$, where

$$W(H, d_0) = T(H\hat{\delta} - d_0)'(H\hat{\Sigma}_\delta H')^{-1}(H\hat{\delta} - d_0) \tag{3.19}$$

and c depends on the level of the test. The asymptotic distribution of the statistic $W(H, d_0)$, under H_0, is chi-square with r degrees of freedom. Again, this test is valid for large samples.

Concerning the power of the above tests, we can make the important observation that they are consistent whenever $M\alpha \neq m_0$ or $H\delta \neq d_0$ (see Section 6.5).[7] Besides, by considering complements of the critical regions described above, we can obtain confidence regions for $M\alpha$ or $H\delta$, for example confidence intervals for the individual covariances in δ.

3. SPECIAL CASES

We will now examine three cases of special interest. First, consider the situation where the null hypothesis is $H_0 : \delta = 0$ or equivalently, $H_0' : a = 0$. Under H_0, we can rewrite equation (3.2) as

$$y = Y\beta + Z_1\gamma + \hat{V}a + e, \tag{4.1}$$

where e follows a $N[0, \sigma_e^2 I_T]$ distribution and is independent of both Y and $\hat{V}$. Then the standard F-statistic for testing $a = 0$ is

$$F = \frac{\hat{a}'(\hat{V}'M_1\hat{V})\hat{a}/G}{\hat{e}'\hat{e}/(T - K_1 - 2G)}, \tag{4.2}$$

where $M_1 = I_T - X_1(X_1'X_1)^{-1}X_1'$ and $X_1 = [Y : Z_1]$; under H_0, F follows a Fisher distribution with $(G, T - K_1 - 2G)$ degrees of freedom. The resulting test is exact rather than asymptotic.[8] It is not equivalent (even asymptotically) to the test of $a = 0$ based on the statistic $S_0 = T\hat{a}'\hat{\Sigma}_a^{-1}\hat{a}$, obtained

[7] This property is especially important in view of Holly's (1982a) discussion of Hausman-type tests.

[8] One can see easily that this test is equivalent to a test suggested by Wu (1973, T_2 statistic) and, in a different form, by Hausman (1978, eq. 2.23), except that Z_1 is not necessarily a submatrix of Z. On alternative forms of the Wu-Hausman test, see Nakamura and Nakamura (1981).

from (3.13). The main difference is that $\hat{\rho}$ is set to zero in the estimator of Σ_a in (3.16). This restriction is justified under H_0, for then $\rho = a'\Sigma_{22}a = 0$. If we write F in the form

$$F = \frac{1}{G}\left(T\hat{a}'\tilde{\Sigma}_a^{-1}\hat{a}\right), \tag{4.3}$$

where

$$\tilde{\Sigma}_a = \hat{\sigma}_0^2(\hat{V}'M_1\hat{V}/T)^{-1}, \quad \hat{\sigma}_0^2 = \hat{e}'\hat{e}/(T - K_1 - 2G), \tag{4.4}$$

we see easily that the statistics F and S_0/G are asymptotically identical under H_0 (since $\hat{\rho} \to 0$). Nevertheless, under the alternative, this equivalence does not hold because $\hat{\rho}$ does not, in general, converge to zero.

The second problem we wish to examine is to test whether a subset of the variables in Y are independent of u, conditional on the assumption that the others are independent of u. More precisely, given $Y = [Y_1 : Y_2]$, we want to test whether Y_1 and u are independent, knowing that Y_2 and u are independent. To do this, we can simply include Y_2 in Z_1 and reshape equation (2.1) accordingly:

$$y = Y_1\beta_1 + Z_3\gamma_3 + u, \tag{4.5}$$

where $Z_3 = [Y_2 : Z_1]$, $\gamma_3 = (\beta_2', \gamma')'$, and $\beta = (\beta_1', \beta_2')'$ is the partition of β corresponding to $[Y_1 : Y_2]$. We then proceed as previously on the transformed model.

Finally, consider the important case where the matrix Z_1 is a submatrix of Z, say $Z = [Z_1 : Z_2]$. This is probably the most frequent situation when (2.1) is viewed as a "structural equation" (presumably inside some system of equations) and (2.2) represents the "reduced-form equation" for the endogenous variables appearing on the right-hand side of (2.1). In this case, the estimates $\hat{\beta}$ and $\hat{\gamma}$, obtained from the regression given by (3.3), are the two-stage least squares (2SLS) estimates of β and γ. To see this, rewrite equation (3.2) as

$$y = \hat{Y}\beta + Z_1\gamma + \hat{V}a^* + e^*, \tag{4.6}$$

where $a^* = a + \beta$. By the orthogonality relations $\hat{V}'\hat{Y} = 0$ and $\hat{V}'Z_1 = 0$, the estimates of β and γ obtained by OLS from (4.6) are identical to those obtained from the regression

$$y = \hat{Y}\beta + Z_1\gamma + e^{**}. \tag{4.7}$$

They are thus identical to the 2SLS estimates of β and γ, showing clearly that the linear Wald tests described above have a natural complementary with the estimation of a structural equation by 2SLS.

In the same special case, the estimate $\hat{\delta}$ used in Theorem 2 may be derived in a second interesting manner. Using again the orthogonality relations, we see that

$$\hat{a} = \hat{\Sigma}_{22}^{-1}\left(\frac{1}{T}\hat{V}y\right),$$

hence

$$\hat{a} = \hat{a}^* - \hat{\beta}$$

and

$$\hat{\delta} = \frac{1}{T}\hat{V}'y - \hat{\Sigma}_{22}\hat{\beta}.$$

Further, substitute (2.2) into (2.1) to get the reduced-form equation for y:

$$y = Z\Pi\beta + Z_1\gamma + v_0,$$

where $v_0 = V\beta + u$. If we denote the tth element of v_0 by $v_{0t} = v_t'\beta + u_t$ and define $\omega_0 = E[v_t v_{0t}]$, we have

$$\delta = \omega_0 - \Sigma_{22}\beta.$$

Since ω_0 can be consistently estimated by $\hat{\omega} = \frac{1}{T}\hat{V}'\hat{v}_0$, where $\hat{v}_0$ is the vector of residuals from the regression of y on Z, this suggests the following estimate of δ:

$$\tilde{\delta} = \hat{\omega}_0 - \hat{\Sigma}_{22}\tilde{\beta},$$

where $\tilde{\beta}$ is a consistent estimate of β. Then, if we take $\tilde{\beta} = \hat{\beta}$, the 2SLS estimate of β, we see that

$$\tilde{\delta} = \frac{1}{T}\hat{V}'y - \hat{\Sigma}_{22}\hat{\beta} = \hat{\delta}, \tag{4.8}$$

which shows that the estimator $\hat{\delta}$ can be generated in a second natural manner.

5. ECONOMETRIC APPLICATIONS

As previously indicated, assumptions concerning the independence of various stochastic explanatory variables in a structural equation and the disturbance term can have important implications for the appropriate choice of method of inference. On the one hand, if all the stochastic explanatory variables are correlated with the disturbance term, OLS does not usually provide consistent estimates of the structural coefficients in the equation and, even more generally, standard inference techniques (like F-tests) are not valid; one should use a simultaneous equations technique (e.g., instrumental variables). On the other hand, if they are all independent of the disturbance term, standard linear regression techniques (OLS, F-tests) are appropriate. Furthermore, between these two extremes, several intermediate cases are possible. If some but not all stochastic explanatory variables are independent of the disturbance term, standard inference techniques are not generally valid. However, we can exploit this information to get a more efficient method. In particular, if we split the matrix of stochastic explanatory variables into two submatrices $Y = [Y_1 : Y_2]$, where Y_2 is independent of u, we can get more efficient consistent estimators and more powerful tests by treating Y_2 as exogenous: in particular, this can be done by using Y_2 as an additional set of instruments or, at least, by not replacing Y_2 by $\hat{Y}_2$ (see Maddala, 1977, pp. 477–478).

The procedures developed above allow one to test the exogeneity of each stochastic explanatory variable included in a given equation by looking at asymptotic t-values. It is easy to compute these in a routine way while estimating the equation by an instrumental-variable method. In this manner, one can get automatic indications on the simultaneity properties of a model and possible ways of improving estimation efficiency.

Finally, we may observe that a number of economic hypotheses can be formulated in terms of the independence between certain stochastic explanatory variables and the error term in an equation. Wu (1973) described a number of such cases, such as the permanent income hypothesis, the expected profit maximization hypothesis, and the recursiveness hypothesis in simultaneous equation models.

6. PROOFS

6.1 Proof of Theorem 1

First, from (3.2), (3.3) and (3.8), we have the identity:

$$\sqrt{T}(\hat{\alpha} - \alpha) = \left(\frac{X'X}{T}\right)^{-1} e_T, \tag{6.1}$$

where

$$e_T = \frac{1}{\sqrt{T}} X'e + \frac{X'Z}{T}\sqrt{T}(\hat{\Pi} - \Pi)a.$$

Moreover, we can see that

$$(\hat{\Pi} - \Pi)a = (Z'Z)^{-1}Z'Va;$$

hence $E[(\hat{\Pi} - \Pi)a] = 0$ and

$$E[(\hat{\Pi} - \Pi)aa'(\hat{\Pi} - \Pi)'] = \rho(Z'Z)^{-1},$$

where $\rho = a'\Sigma_{22}a$, for, from Assumption 3, it is easily verified that $E[Vaa'V'] = \rho I_T$. Consequently,

$$\sqrt{T}(\hat{\Pi} - \Pi)a \sim N\left[0, \rho\left(\frac{Z'Z}{T}\right)^{-1}\right]. \tag{6.2}$$

Also, since e is independent of Y, the distribution of $\frac{1}{\sqrt{T}}X'e$, conditional on Y, is $N[0, \sigma_e^2(\frac{X'X}{T})]$.

Consider the characteristic function of e_T,

$$\begin{aligned}\phi_T(\tau) &\equiv E\{\exp[i\tau' e_T]\}\\ &= E\left\{\exp\left[i\tau'\frac{1}{\sqrt{T}}X'e + i\tau'\left(\frac{X'Z}{T}\right)\sqrt{T}(\hat{\Pi} - \Pi)a\right]\right\},\end{aligned}$$

where $\tau \in R^{2G+K_1}$ and $i = \sqrt{-1}$. In order to get an explicit expression for $\phi_T(\tau)$, we first compute the expected value of $\exp\{i\tau' e_T\}$ conditional on Y. Since $\frac{1}{\sqrt{T}}X'e$ is normal for given Y, we have

$$R_T(\tau) \equiv E[\exp(i\tau' e_T) \mid Y] = R_T^{(1)}(\tau)R_T^{(2)}(\tau),$$

where

$$R_T^{(1)}(\tau) = \exp\left[-\frac{\sigma_e^2}{2}\tau'\left(\frac{X'X}{T}\right)\tau\right]$$

and

$$R_T^{(2)}(\tau) = \exp\left[i\tau'\left(\frac{X'Z}{T}\right)\sqrt{T}(\hat{\Pi} - \Pi)a\right].$$

Then, using (3.6) and (6.2), we see that

$$\text{plim } R_T^{(1)}(\tau) = \exp\left\{-\frac{1}{2}\sigma_e^2\tau' Q_x \tau\right\}.$$

Also, from (3.7), (6.2) and Assumption 5, we have

$$R_T^{(2)}(\tau) \xrightarrow{L} \exp\{i\tau' B\},$$

where $B \sim N[0, \rho Q'_{zx} Q_z^{-1} Q_{zx}]$ and $\xrightarrow{L}$ refers to convergence in distribution as $T \to \infty$. Consequently,

$$R_T(\tau) \xrightarrow{L} \exp\left(-\frac{1}{2}\sigma_e^2 \tau' Q_x \tau\right) \exp(i\tau' B);$$

hence, by the Helly-Bray Theorem,

$$\begin{aligned}\lim_{T\to\infty} E\{R_T(\tau)\} &= \exp\left(-\frac{1}{2}\sigma_e^2 \tau' Q_x \tau\right) E\{\exp(i\tau' B)\} \\ &= \exp\left\{-\frac{1}{2}\tau' \left(\sigma_e^2 Q_x + \rho Q'_{zx} Q_z^{-1} Q_{zx}\right) \tau\right\},\end{aligned}$$

for all τ. Since $\phi_T(\tau) = E\{R_T(\tau)\}$, it follows that $\phi_T(\tau)$ converges to the characteristic function of the $N[0, \sigma_e^2 Q_x + \rho Q'_{zx} Q^{-1} Q_{zx}]$ distribution. Therefore,

$$e_T \longrightarrow N[0, \sigma_e^2 Q_x + \rho Q'_{zx} Q_z^{-1} Q_{zx}] \tag{6.3}$$

and, using (6.1),

$$\sqrt{T}(\hat{\alpha} - \alpha) \longrightarrow N[0, \Sigma_\alpha], \tag{6.4}$$

where Σ_α is given by (3.9). The consistency of $\hat{\alpha}$ follows from (6.4). Concerning the estimator $\hat{\sigma}_e^2$, we can write

$$\hat{\sigma}_e^2 = \frac{e^{*\prime} e^*}{T} - \frac{1}{T}\left(\frac{X' e^*}{\sqrt{T}}\right)' \left(\frac{X'X}{T}\right)^{-1} \left(\frac{X' e^*}{\sqrt{T}}\right),$$

hence

$$\text{plim } \hat{\sigma}_e^2 = \text{ plim } \frac{e^{*\prime} e^*}{T}.$$

Moreover, by the definition of e^* in (3.3), we have

$$\frac{e^{*\prime} e^*}{T} = \frac{e'e}{T} + 2a'(\hat{\Pi} - \Pi)' \frac{Z'e}{T} + a'(\hat{\Pi} - \Pi)' \left(\frac{Z'Z}{T}\right) (\hat{\Pi} - \Pi) a;$$

hence, since plim $(Z'e/T) = 0$ and plim $(\hat{\Pi} - \Pi)a = 0$,

$$\text{plim } \hat{\sigma}_e^2 = \text{ plim } \frac{e'e}{T} = \sigma_e^2,$$

which shows that $\hat{\sigma}_e^2$ is a consistent estimator of σ_e^2. Finally, we can see that $\hat{\Sigma}_\alpha$ is a consistent estimator of Σ_α by considering the definitions of Q_x and Q_{zx}, and by noting that $\hat{\rho}$ and $\hat{\sigma}_e^2$ are consistent for ρ and σ_e^2. Q.E.D.

6.2 Proof of Corollary 1.1

The consistency of $\hat{a}$ follows from the consistency of $\hat{\alpha}$. The asymptotic distribution of $\sqrt{T}(\hat{a}-a)$ follows from the identity

$$\sqrt{T}(\hat{a}-a) = C_2\left(\frac{1}{\sqrt{T}}X'e^*\right) \tag{6.5}$$

and from (6.3). The consistency of $\hat{\Sigma}_a$ follows from the consistency of $\hat{\Sigma}_\alpha$ and the definition of A_2. Q.E.D.

6.3 Lemma

In order to obtain the asymptotic distribution of $\hat{\delta} = \hat{\Sigma}_{22}\hat{a}$, we will need the following lemma.

Lemma 1. Suppose that Assumptions 2, 3 and 5 are satisfied. Let σ_i' and $\hat{\sigma}_i'$ be the ith rows of Σ_{22} and $\hat{\Sigma}_{22}$, respectively ($i = 1,\ldots,G$), and

$$\sigma = (\sigma_1', \sigma_2', \ldots, \sigma_G')', \quad \hat{\sigma} = (\hat{\sigma}_1', \hat{\sigma}_2', \ldots, \hat{\sigma}_G')'.$$

Then, the vector $\sqrt{T}(\hat{\sigma}-\sigma)$ has a normal limiting distribution, as $T \to \infty$, with mean zero and covariance matrix

$$\Sigma_\sigma = \begin{bmatrix} V_{11} & \cdots & V_{1G} \\ \vdots & & \vdots \\ V_{G1} & \cdots & V_{GG} \end{bmatrix}, \tag{6.6}$$

where $V_{ij} = \sigma_{ij}\Sigma_{22} + \sigma_j\sigma_i'$. Furthermore, the vector $\sqrt{T}(\hat{\Sigma}_{22}-\Sigma_{22})c$, where c is any fixed $G \times 1$ vector, has a normal limiting distribution with mean zero and covariance matrix

$$\psi_c = (c'\Sigma_{22}c)\,\Sigma_{22} + (\Sigma_{22}c)\,(\Sigma_{22}c)'. \tag{6.7}$$

Proof. Let $\hat{w}_i = M_z w_i$, where w_i is the ith column of the matrix V and $M_z = I_T - Z(Z'Z)^{-1}Z'$. The (i,j)th element of $\hat{\Sigma}_{22}$ has the form $\hat{\sigma}_{ij} = \hat{w}_i'\hat{w}_j/T$; hence

$$\hat{\sigma}_{ij} = \frac{w_i'w_j}{T} - \frac{1}{T}\left(\frac{Z'w_i}{\sqrt{T}}\right)'\left(\frac{Z'Z}{T}\right)^{-1}\left(\frac{Z'w_j}{\sqrt{T}}\right).$$

Let $\tilde{\sigma}_{ij} = w_i'w_j/T, \tilde{\sigma}_i = (\tilde{\sigma}_{i1}, \tilde{\sigma}_{i2}, \ldots, \tilde{\sigma}_{iG})'$, $i,\ j = 1,\ldots,G$, and $\tilde{\sigma} = (\tilde{\sigma}_1', \tilde{\sigma}_2', \ldots, \tilde{\sigma}_G')'$. Then, using Assumptions 3 and 5, we get

$$\text{plim}\left[\sqrt{T}\,(\hat{\sigma}_{ij}-\sigma_{ij}) - \sqrt{T}\,(\tilde{\sigma}_{ij}-\sigma_{ij})\right] = 0, \quad i,\ j = 1,\ldots,G.$$

Thus, the vectors $\sqrt{T}(\hat{\sigma}-\sigma)$ and $\sqrt{T}(\tilde{\sigma}-\sigma)$ have the same limiting distribution.

Let w_{it} be the tth element of w_i and define $S_{it} = (w_{it}w_{1t}, w_{it}w_{2t}, \ldots, w_{it} w_{Gt})'$, $i = 1, \ldots, G$, $S_t = (S'_{1t}, S'_{2t}, \ldots, S'_{Gt})'$, $t = 1, \ldots, T$. It is clear that the vectors S_t, $t = 1, \ldots, T$, are independent and identically distributed with mean σ. Furthermore, since

$$\begin{aligned} E\left[(S_{it}-\sigma_i)(S_{jt}-\sigma_j)'\right] &= \left[\sigma_{ij}\sigma_{k\ell}+\sigma_{i\ell}\sigma_{jk}\right]_{k,\ \ell=1,\ldots,G} \\ &= \sigma_{ij}\Sigma_{22}+\sigma_j\sigma_i',\ i,j=1,\ldots,G, \end{aligned}$$

for all t (see Anderson, 1958, p. 39), the covariance matrix of S_t is Σ_σ, as given in (6.6). Thus, since

$$\sqrt{T}(\tilde{\sigma}-\sigma) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}(S_t-\sigma),$$

and using the Multivariate Central Limit Theorem (see Anderson, 1958, Theorem 4.2.3), we can conclude that the limiting distribution of $\sqrt{T}(\tilde{\sigma}-\sigma)$ is $N[0,\Sigma_\sigma]$. Furthermore, for any $G \times 1$ fixed vector c,

$$\sqrt{T}(\hat{\Sigma}_{22}-\Sigma_{22})c = (I_G \otimes c')\sqrt{T}\left(\hat{\sigma}-\sigma\right),$$

where I_G is the identity matrix of order G and $\otimes$ refers to the Kronecker product. Since $\sqrt{T}(\hat{\sigma}-\sigma)$ is asymptotically $N[0,\Sigma_\sigma]$, we can conclude that the vector $\sqrt{T}(\hat{\Sigma}_{22}-\Sigma_{22})c$ is asymptotically normal with mean zero and covariance matrix

$$\begin{aligned} \psi_c &\equiv (I_G\otimes c')\,\Sigma_\sigma\,(I_G\otimes c) \\ &= [c'V_{ij}c]_{i,\ j=1,\ldots,G}\,. \end{aligned}$$

We see easily that ψ_c reduces to the expression in (6.7). Q.E.D.

6.4 Proof of Theorem 2

First, note that the vector $\sqrt{T}(\hat{\delta}-\delta)$ can be decomposed as follows:

$$\begin{aligned} \sqrt{T}(\hat{\delta}-\delta) &= \sqrt{T}\hat{\Sigma}_{22}(\hat{a}-a)+\sqrt{T}(\hat{\Sigma}_{22}-\Sigma_{22})a \\ &= \hat{\Sigma}_{22}C_2\left[\frac{1}{\sqrt{T}}X'e+\left(\frac{X'Z}{T}\right)\sqrt{T}(\hat{\Pi}-\Pi)a\right] \\ &\quad +\sqrt{T}\left(\hat{\Sigma}_{22}-\Sigma_{22}\right)a, \end{aligned}$$

where we have used (6.5) and (3.3). Let

$$W_T \equiv \Sigma_{22}A_2\left[\frac{1}{\sqrt{T}}X'e+Q'_{zx}\sqrt{T}(\hat{\Pi}-\Pi)a\right]+\sqrt{T}\left(\hat{\Sigma}_{22}-\Sigma_{22}\right)a.$$

Since $\text{plim}(\hat{\Sigma}_{22}C_2) = \Sigma_{22}A_2$ and $\text{plim}(X'Z/T) = Q'_{zx}$, we have $\text{plim}[\sqrt{T}(\hat{\delta} - \delta) - W_T] = 0$; $\sqrt{T}(\hat{\delta} - \delta)$ and W_T must have the same asymptotic distribution (see Billingsley, 1968, p. 25, Theorem 4.1). Consider now the characteristic function of W_T,

$$\phi_T(\tau) = E\left\{\exp\left[i\tau' W_T\right]\right\},$$

where $\tau \in R^G$. Since e is independent of Y, $\frac{1}{\sqrt{T}}X'e \sim N[0, \sigma_e^2(\frac{X'X}{T})]$ for Y fixed and, by taking the expected value of $\exp(i\tau' W_T)$ conditional on Y, we get

$$S_T(\tau) \equiv E\left\{\exp\left(i\tau' W_T\right) \mid Y\right\} = S_T^{(1)}(\tau) S_T^{(2)}(\tau),$$

where

$$S_T^{(1)}(\tau) = \exp\left\{-\frac{1}{2}\sigma_e^2 \tau' \Sigma_{22} A_2 \left(\frac{X'X}{T}\right) A_2' \Sigma_{22} \tau\right\}$$

and

$$S_T^{(2)}(\tau) = \exp\left\{i\tau' \Sigma_{22} A_2 Q'_{zx} \sqrt{T}(\hat{\Pi} - \Pi)a + i\tau' \sqrt{T}\left(\hat{\Sigma}_{22} - \Sigma_{22}\right) a\right\}.$$

Furthermore, using (3.6),

$$\text{plim } S_T^{(1)}(\tau) = \exp\left\{-\frac{1}{2}\sigma_e^2 \tau' \Sigma_{22} A_2 Q_x A_2' \Sigma_{22} \tau\right\} \equiv S^{(1)}(\tau). \tag{6.8}$$

Consequently, by the Helly-Bray Theorem,

$$\lim_{T\to\infty} E[S_T(\tau)] = S^{(1)}(\tau) \lim_{T\to\infty} E\left[S_T^{(2)}(\tau)\right], \tag{6.9}$$

where the expectation E is taken over Y.

Each column $\hat{\Pi}_j$ of $\hat{\Pi}$ is independent of each column $\hat{w}_k$ of $\hat{V}$, since $E[(\hat{\Pi}_j - \Pi_j)\hat{w}_k'] = 0$, $j, k = 1, \ldots, G$. Therefore, $\hat{\Pi}$ and $\hat{\Sigma}_{22}$ are independent and

$$\begin{aligned} E\left\{S_T^{(2)}(\tau)\right\} &= E\left\{\exp\left[i\tau' \Sigma_{22} A_2 Q'_{zx} \sqrt{T}(\hat{\Pi} - \Pi)a\right]\right\} \\ &\quad \times E\left\{\exp\left[i\tau' \sqrt{T}\left(\hat{\Sigma}_{22} - \Sigma_{22}\right) a\right]\right\} \\ &= \exp\left[-\frac{1}{2}\rho\tau' \Sigma_{22} A_2 Q'_{zx} \left(\frac{Z'Z}{T}\right)^{-1} Q_{zx} A_2' \Sigma_{22} \tau\right] \\ &\quad \times E\left\{\exp\left[i\tau' \sqrt{T}\left(\hat{\Sigma}_{22} - \Sigma_{22}\right) a\right]\right\}, \end{aligned} \tag{6.10}$$

where the second identity comes from (6.2). By Lemma 1,

$$\sqrt{T}\left(\hat{\Sigma}_{22} - \Sigma_{22}\right) a \xrightarrow{L} N[0, \psi],$$

where $\psi = (a'\Sigma_{22}a)\Sigma_{22} + (\Sigma_{22}a)(\Sigma_{22}a)'$. Since $\rho = a'\Sigma_{22}a$ and $\delta = \Sigma_{22}a$, we see that $\psi = \rho\Sigma_{22} + \delta\delta'$. Thus, the limit characteristic function of $\sqrt{T}(\hat{\Sigma}_{22} - \Sigma_{22})a$ is

$$\lim_{T\to\infty} E\left\{\exp\left[i\tau'\sqrt{T}\left(\hat{\Sigma}_{22} - \Sigma_{22}\right)a\right]\right\} = \exp\left[-\frac{1}{2}\tau'\psi\tau\right]. \tag{6.11}$$

Using (6.8)–(6.11), we obtain

$$\begin{aligned}&\lim_{T\to\infty} E[S_T(\tau)]\\ &= \exp\left\{-\frac{1}{2}\tau'\left[\Sigma_{22}A_2\left(\sigma_e^2 Q_x + \rho Q_{zx}' Q_z^{-1} Q_{zx}\right)A_2'\Sigma_{22} + \psi\right]\tau\right\},\end{aligned}$$

which implies that the asymptotic distribution of $\sqrt{T}(\hat{\delta} - \delta)$ is normal with mean zero and covariance matrix $\Sigma_\delta = \Sigma_{22}\Sigma_a\Sigma_{22} + \psi$, where Σ_a is given by (3.14). Finally, the consistency of $\hat{\delta} = \hat{\Sigma}_{22}\hat{a}$ follows from that of $\hat{\Sigma}_{22}$ and $\hat{a}$ for Σ_{22} and a respectively, and the consistency of $\hat{\Sigma}_\delta$ follows from the consistency of $\hat{\Sigma}_{22}$, $\hat{\Sigma}_a$ and $\hat{\psi}$ for Σ_{22}, Σ_a and ψ respectively. Q.E.D.

6.5 Asymptotic Power

We will now show that the tests discussed above are consistent. The statistic $S(M, m_0)$ used to test $M\alpha = m_0$, where M is a $\nu \times L$ matrix of rank ν, can be decomposed in the following way:

$$S(M, m_0) = S_1(M, \alpha) + \sqrt{T}S_2(M, \alpha, m_0),$$

where

$$S_1(M, \alpha) = T(M\hat{\alpha} - M\alpha)'(M\hat{\Sigma}_\alpha M')^{-1}(M\hat{\alpha} - M\alpha)$$

converges to a chi-square distribution with ν degrees of freedom and

$$\begin{aligned}&S_2(M, \alpha, m_0)\\ &= \left[2\sqrt{T}M(\hat{\alpha} - \alpha) + \sqrt{T}(M\alpha - m_0)\right]'\left(M\hat{\Sigma}_\alpha M'\right)^{-1}(M\alpha - m_0).\end{aligned}$$

We will show that plim $S_2(M, \alpha, m_0) = +\infty$, whenever $M\alpha \neq m_0$.

Consider first the case where all the elements of the vector $M\alpha - m_0$ are different from zero. In the sum $[2\sqrt{T}M(\hat{\alpha} - \alpha) + \sqrt{T}(M\alpha - m_0)]$, the second term always dominates as $T \to \infty$ and $\sqrt{T}(\hat{\alpha} - \alpha)$ has a limiting distribution. Consequently

$$\text{plim}\; S_2(M, \alpha, m_0) = \text{plim}\; \sqrt{T}(M\alpha - m_0)'\left(M\hat{\Sigma}_\alpha M'\right)^{-1}(M\alpha - m_0) = +\infty,$$

where the fact that plim $(M\hat{\Sigma}_\alpha M')^{-1} = (M\Sigma_\alpha M')^{-1}$ is positive definite has been used. Second, for the case where $M\alpha \neq m_0$ but some elements of $(M\alpha - m_0)$ are zero, we can assume without loss of generality that these constitute the lower vector of $(M\alpha - m_0)$:

$$M\alpha - m_0 = (d_1', 0')',$$

where all the elements of the $\nu_1 \times 1$ vector d_1 are different from zero. Furthermore, let us partition $(\hat{\alpha} - \alpha)$ and $(M\hat{\Sigma}_\alpha M')^{-1}$ conformably with $(d_1', 0')'$:

$$\hat{\alpha} - \alpha = \begin{bmatrix} (\hat{\alpha} - \alpha)_1 \\ (\hat{\alpha} - \alpha)_2 \end{bmatrix}, \quad (M\hat{\Sigma}_\alpha M')^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

where $(\hat{\alpha} - \alpha)_1$ is a $\nu_1 \times 1$ vector and A_{11} is a $\nu_1 \times \nu_1$ positive definite matrix. Then

$$S_2(M, \alpha, m_0) = 2\sqrt{T}(\hat{\alpha}-\alpha)_1' M' A_{11} d_1 + 2\sqrt{T}(\hat{\alpha}-\alpha)_2' M' A_{21} d_1 + \sqrt{T} d_1' A_{11} d_1.$$

Since plim (A_{11}) is a positive definite matrix and $\sqrt{T}(\hat{\alpha} - \alpha)$ has a limiting distribution, we have

$$\text{plim } S_2(M, \alpha, m_0) = +\infty. \tag{6.12}$$

Thus, (6.12) holds whenever $M\alpha \neq m_0$,

$$\text{plim } S(M, m_0) = \text{ plim } \left[S_1(M, \alpha) + \sqrt{T} S_2(M, \alpha, m_0)\right] = +\infty,$$

whenever $M\alpha \neq m_0$, and

$$\lim_{T\to\infty} P\left[S(M, m_0) \geq c\right] = \begin{cases} \varepsilon, & \text{if } M\alpha = m_0 \\ 1, & \text{if } M\alpha \neq m_0, \end{cases} \tag{6.13}$$

where ε is the level of the test. This proves the consistency of the tests proposed for linear hypotheses regarding α. The consistency of tests regarding δ can be shown in a similar way.

ACKNOWLEDGMENTS

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, Foundation FCAR (Government of Quebec) and the Centre de recherche et développement en économique (Université de Montréal). The author thanks Arnold Zellner, Marcel Dagenais, Alberto Holly, Nicholas Kiefer, De-Min Wu and participants to the University of Chicago Econometrics and Statistics Colloquium for several very helpful comments. This paper is a revised version of Dufour (1979, Chapter 4; 1980).

REFERENCES

Anderson, T. W. (1958), *An Introduction to Multivariate Statistical Analysis*. New York: Wiley and Sons.

Billingsley, P. (1968), *Convergence of Probability Measures*. New York: Wiley and Sons.

Bronsard, C., and L. Salvas-Bronsard (1984), "On price exogeneity in complete demand systems." *Journal of Econometrics* **24**, 235–247.

Dufour, J.-M. (1979), "Methods for specification errors analysis with macroeconomic applications." Ph.D. Thesis, Department of Economics, University of Chicago.

Dufour, J.-M. (1980), "Tests of exogeneity." Cahier 8026, Département de sciences économiques, Université de Montréal.

Durbin, J. (1954), "Errors in variables." *Review of the International Statistical Institute* **22**, 23–32.

Engle, R. F. (1982), "A general approach to Lagrange multiplier diagnostics." *Journal of Econometrics* **20**, 83–104.

Engle, R. F. (1984), "Wald, likelihood ratio, and Lagrange multiplier tests in econometrics." In *Handbook of Econometrics*, Volume 2, ed. Z. Griliches and M. Intrilligator, pp. 775–826. Amsterdam: North-Holland.

Engle, R. F., D. F. Hendry, and J.-F. Richard (1983), "Exogeneity." *Econometrica* **51**, 277–304.

Farebrother, R. W. (1976), "A remark on the Wu test." *Econometrica* **44**, 475–477.

Fisher, F. M. (1966), *The Identification Problem in Econometrics*. New York: McGraw-Hill.

Gouriéroux, C., and A. Trognon (1984), "Specification pre-test estimation." *Journal of Econometrics* **25**, 15–28.

Hausman, J. (1978), "Specification tests in econometrics." *Econometrica* **46**, 1251–1272.

Hausman, J., and W. E. Taylor (1980), "Comparing specification tests and classical tests." Mimeo, Massachusetts Institute of Technology, Cambridge, MA.

Hausman, J., and W. E. Taylor (1981a), "Panel data and unobservable individual effects." *Econometrica* **49**, 1377–1398.

Hausman, J., and W. E. Taylor (1981b), "A generalized specification test." *Economics Letters* **8**, 239–245.

Holly, A. (1980), "Testing recursiveness in a triangular simultaneous equation model." Mimeo, Université de Lausanne.

Holly, A. (1982a), "A remark on Hausman's specification test." *Econometrica* **50**, 749–760.

Holly, A. (1982b), "Tests de spécification." *Cahiers du séminaire d'économétrie* **24**, 151–173.

Holly, A. (1983), "Une présentation unifiée des tests d'exogénéité dans les modèles à équations simultanées." *Annales de l'INSEE* **50**, April–June, 3–24.

Holly, A., and A. Monfort (1982), "Some useful equivalence properties of Hausman's test." Cahier 8201, Université de Lausanne.

Holly, A., and D. Sargan (1982), "Testing of exogeneity within a limited information framework." Cahier 8204, Université de Lausanne.

Hwang, H.-S. (1980), "Test of independence between a subset of stochastic regressors and disturbances." *International Economic Review* **21**, 749–760.

Kariya, T., and H. Hodoshima (1980), "Finite sample properties of the tests for independence in structural systems and the LRT." *The Economic Studies Quarterly* **31**, 45–56.

Maddala, G. S. (1977), *Econometrics*. New York: McGraw-Hill.

Nakamura, A., and M. Nakamura (1980), "On the usefulness of the Wu-Hausman test for detecting the least squares bias problem." Mimeo, University of Alberta.

Nakamura, A., and M. Nakamura (1981), "On the relationships among several specification error tests presented by Durbin, Wu and Hausman." *Econometrica* **49**, 1583–1588.

Plosser, C. J., G. W. Schwert, and H. White (1982), "Differencing as a test of specification." *International Economic Review* **23**, 535–552.

Rao, C. R. (1973), *Linear Statistical Inference and its Applications*, 2nd edition. New York: Wiley and Sons.

Revankar, N. S. (1978), "Asymptotic relative efficiency analysis of certain tests in structural systems." *International Economic Review* **19**, 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* **14**, 625–631.

Reynolds, R. A. (1982), "Posterior odds for the hypothesis of independence between stochastic regressors and disturbances." *International Economic Review* **23**, 479–490.

Richard, J.-F. (1980), "Models with several regimes and changes in exogeneity." *Review of Economic Studies* **17**, 1–20.

Riess, P. (1983), "Alternative interpretations of Hausman's test." Mimeo, Yale University.

Ruud, P. A. (1984), "Tests of specification in econometrics." *Econometric Reviews* **3**, 211–242. [with discussion by T. S. Breusch and G. E. Mizon, J. Hausman, L.-F. Lee and H. White, 243–276].

Smith, R. J. (1984), "A note on likelihood ratio tests for the independence between a subset of stochastic regressors and disturbances." *International Economic*

Review **25**, 263–269.

Spencer, D. E., and K. N. Berk (1981), "A limited-information specification test." *Econometrica* **49**, 1079–1085. [Erratum, **50**, 1087].

Stroud, T. W. F. (1971), "On obtaining large-sample tests for asymptotically normal estimators." *Annals of Mathematical Statistics* **42**, 1412–1424.

Szroeter, J. (1983), "Generalized Wald methods for testing nonlinear implicit and overidentifying restrictions." *Econometrica* **51**, 335–353.

Theil, H. (1971), *Principles of Econometrics*. New York: Wiley and Sons.

Thursby, J. G. (1982), "The common structure of regression-based specification error tests." Mimeo, Department of Economics, Ohio State University.

Tsurumi, H., and T. Shiba (1984), "Tests of exogeneity using restricted reduced form coefficients." Mimeo, Rutgers University.

Turkington, D. A. (1980), "A note on Hausman's test for the limited information model." Mimeo, Department of Economics, University of Western Australia (Nedlands, Australia).

Wald, A. (1943), "Tests of statistical hypotheses concerning several parameters when the number of observations is large". *Transactions of the American Mathematical Society* **54**, 426–482.

White, H. (1982), "Maximum likelihood estimation of misspecified models." *Econometrica* **50**, 1–26.

Wu, D.-M. (1973), "Alternative tests of independence between stochastic regressors and disturbances." *Econometrica* **41**, 733–750.

Wu, D.-M. (1974), "Alternative tests of independence between stochastic regressors and disturbances: finite sample results." *Econometrica* **42**, 529–546.

Wu, D.-M. (1983a), "Tests of causality, predeterminedness and exogeneity." *International Econometric Review* **24**, 547–558.

Wu, D.-M. (1983b), "A remark on a generalized specification test." *Economics Letters* **11**, 365–370.

Zellner, A., J. Kmenta, and J. Drèze (1966), "Specification and estimation of Cobb-Douglas production functions." *Econometrica* **34**, 784–795.