

A COMMENT ON THE TEST OF OVERIDENTIFYING RESTRICTIONS<sup>1</sup>

BY JOSEPH B. KADANE AND T. W. ANDERSON

THE TEST OF OVERIDENTIFYING restrictions of one equation in a simultaneous system proposed by Anderson and Rubin [1] and amplified by Koopmans and Hood [7] has been a source of some confusion in the literature. For instance, Liu and Breen [8] claimed that "It is . . . clear that the test does not really test the null hypothesis (of zero restrictions on the endogenous and exogenous variables)" because, they thought the restrictions on the endogenous variables were included in the computation of the likelihood under the alternative hypothesis. After Fisher and Kadane [3] gave a verbal argument showing that the test is consistent over a wide class of alternatives, Liu and Breen [9] withdrew their earlier view. Nonetheless there is a problem in that generally the null hypothesis is expressed in terms of the structural form, while generally consistency is a matter of the reduced form. Our purpose is to reexamine this problem, and prove two theorems showing the equivalence of various conditions in the literature. We suggest that the null hypothesis be extended.

We may write a single equation as

$$(1) \quad \beta_{\Delta} y'_{\Delta t} + \beta_{\Delta\Delta} y'_{\Delta\Delta t} + \gamma_* z'_{*t} + \gamma_{**} z'_{**t} = u_{1t},$$

where  $\beta_{\Delta}$ ,  $\beta_{\Delta\Delta}$ ,  $\gamma_*$ , and  $\gamma_{**}$  are (row) vectors with  $G_{\Delta}$ ,  $G_{\Delta\Delta}$ ,  $K_*$ , and  $K_{**}$  components, respectively. The reduced form may be written

$$(2) \quad \begin{aligned} y'_{\Delta t} &= \Pi_{\Delta*} z'_{*t} + \Pi_{\Delta**} z'_{**t} + v'_{\Delta t}, \\ y'_{\Delta\Delta t} &= \Pi_{\Delta\Delta*} z'_{*t} + \Pi_{\Delta\Delta**} z'_{**t} + v'_{\Delta\Delta t}. \end{aligned}$$

Anderson and Rubin ([1], p. 56) found that the likelihood ratio test of the null hypothesis that the rank of the  $G_{\Delta} \times K_{**}$  matrix  $\Pi_{\Delta**}$  is  $G_{\Delta} - 1$  against the alternative that the rank is  $G_{\Delta}$  consists of rejecting the null hypothesis when the root of a certain determinantal equation is greater than a suitable value. It was shown by Anderson and Rubin [2] on the basis of large-sample asymptotic theory that the value (appropriately normalized) can be obtained from the  $\chi^2$  distribution with  $K_{**} - G_{\Delta} + 1$  degrees of freedom. Later Kadane [4] showed on the basis of small-disturbance asymptotic theory that the value (appropriately normalized) can be obtained from an  $F$  distribution with  $K_{**} - G_{\Delta} + 1$  and  $T - K$  degrees of freedom.

Consider the identification of (1) by the zero restrictions

$$(3) \quad \beta_{\Delta\Delta} = \mathbf{0}, \quad \gamma_{**} = \mathbf{0}.$$

If  $K_{**} = G_{\Delta} - 1$ , the rank of  $\Pi_{\Delta**}$  is not greater than  $G_{\Delta} - 1$ , (1) is unidentified or just identified, and no test of zero restrictions is possible. Now suppose  $K_{**} \geq G_{\Delta}$ . The restrictions (3) imply the rank of  $\Pi_{\Delta**}$  is not greater than  $G_{\Delta} - 1$  because (3) implies

$$(4) \quad \beta_{\Delta} \Pi_{\Delta**} = \mathbf{0}$$

for some  $\beta_{\Delta} \neq \mathbf{0}$ . If the restrictions (3) are to effect identification, the rank of  $\Pi_{\Delta**}$  must be exactly  $G - 1$ . Koopmans and Hood [7] thus suggest using this same test statistic to test (3) against the alternative that (3) does not hold.

Thus the original Anderson and Rubin work had been done in terms of the reduced form, while Hood and Koopmans were thinking in terms of the structure. The relationship between the structural form and the reduced form is fundamental to the theory of linear

<sup>1</sup> This research was supported by National Science Foundation Grant SOC73-09243 at Carnegie-Mellon University and by National Science Foundation Grant SOC73-05453 at the Institute for Mathematical Studies in the Social Sciences, Stanford University.

systems of simultaneous equations. Most of economic intuition is expressed in terms of the structure, so the structure is often the object of interest for estimation and for testing. Yet the structure has the disadvantage that to a certain degree it is arbitrary, namely it can be multiplied by nonsingular linear transformations provided they do not disturb any special assumptions made about this structure. The reduced form does not share this disadvantage. Any nonsingular linear transformation of the structure leaves the reduced form invariant. For this reason the reduced form is convenient theoretically, but to be most useful, facts about it have to be translated back into structural statements.

The theorem given below accomplishes this task for the problem considered here.

**THEOREM 1:** *The following two conditions are equivalent: (i)  $\rho(\Pi_{\Delta^{**}}) \leq G_{\Delta} - 1$ ; (ii) there exists a nonsingular matrix  $F$  such that*

$$(5) \quad \bar{\beta}_{\Delta\Delta} = \mathbf{0}, \quad \bar{\gamma}_{**} = \mathbf{0},$$

where  $\bar{\beta}_{\Delta\Delta}$  consists of the last  $G_{\Delta\Delta}$  elements of the first row of  $\bar{B}$ , and  $\bar{\gamma}_{**}$  consists of the last  $K_{**}$  elements of the first row of  $\bar{\Gamma}$ , defined by

$$(6) \quad (\bar{B}, \bar{\Gamma}) = F(B, \Gamma).$$

The proof of Theorem 1 is given in the Appendix.

Condition (i) is stressed by Anderson and Rubin, while Condition (ii) gives a structural interpretation for (i). Together with the results of Anderson and Rubin [1, 2], Theorem 1 suggests that the null hypothesis for the test can be considered as (i) holding against the alternative that (i) does not hold or equivalently as (ii) holding against the alternative that (ii) does not hold. (Condition (ii) not holding can be interpreted as (3) not holding for any equation linearly derivable from (2).) This extends the null hypothesis of Hood and Koopmans to include all structures observationally equivalent to (3). (See Koopmans [6], p. 36.) Because they are observationally equivalent, this addition does not affect the significance level of the test.

These issues are illustrated by a special case. The simplest possible case that can be considered is the case of  $G_{\Delta} = G_{\Delta\Delta} = K_{**} = 1$  and  $K_{*} = 0$ . Then the structural equations are

$$(7) \quad \begin{aligned} \beta_{11}y_{1t} + \beta_{12}y_{2t} + \gamma_{11}z_{1t} &= u_{1t}, \\ \beta_{21}y_{1t} + \beta_{22}y_{2t} + \gamma_{21}z_{1t} &= u_{2t}. \end{aligned}$$

The matrix of coefficients of the jointly dependent variables is nonsingular; that is,  $\beta_{11}\beta_{22} - \beta_{12}\beta_{21} \neq 0$ . The reduced form is

$$(8) \quad \begin{aligned} y_{1t} &= \frac{\beta_{12}\gamma_{21} - \beta_{22}\gamma_{11}}{\beta_{11}\beta_{22} - \beta_{12}\beta_{21}} z_{1t} + \frac{\beta_{22}u_{1t} - \beta_{12}u_{2t}}{\beta_{11}\beta_{22} - \beta_{12}\beta_{21}}, \\ y_{2t} &= \frac{\beta_{21}\gamma_{11} - \beta_{11}\gamma_{21}}{\beta_{11}\beta_{22} - \beta_{12}\beta_{21}} z_{1t} + \frac{\beta_{11}u_{2t} - \beta_{21}u_{1t}}{\beta_{11}\beta_{22} - \beta_{12}\beta_{21}}. \end{aligned}$$

The restrictions (3) are

$$(9) \quad \beta_{12} = \gamma_{11} = 0;$$

hence, the first equation of (7) is overidentified when

$$(10) \quad \beta_{22} \neq 0, \quad \gamma_{21} \neq 0.$$

(In general, we define identification by zero restrictions as ‘‘overidentification’’ if there are at least two different ways of deleting a zero restriction so that the remaining zero restrictions effect identification.) The part of the matrix of coefficients in the reduced form

referring to the included jointly dependent variable and excluded predetermined variable is

$$(11) \quad \Pi_{\Delta^{**}} = \frac{\beta_{12}\gamma_{21} - \beta_{22}\gamma_{11}}{\beta_{11}\beta_{22} - \beta_{12}\beta_{21}},$$

which is 0 if the two zero restrictions hold.

When there are  $T$  observations  $(y_{1t}, y_{2t}, z_{1t})$ ,  $t = 1, \dots, T$ , the estimate of the reduced form coefficient of  $z_{1t}$ , in the equation for  $y_{1t}$ , is

$$(12) \quad p_{\Delta^{**}} = \frac{\sum_{t=1}^T y_{1t}z_{1t}}{\sum_{t=1}^T z_{1t}^2},$$

and the smallest (and only root) of the determinantal equation ((4.14) of Anderson and Rubin [1]) is

$$(13) \quad \frac{p_{\Delta^{**}}^2 \sum_{t=1}^T z_{1t}^2}{\sum_{t=1}^T y_{1t}^2 - p_{\Delta^{**}}^2 \sum_{t=1}^T z_{1t}^2}.$$

If  $y_{11}, y_{21}, \dots, y_{1T}, y_{2T}$  are normally distributed with means 0 and  $z_{11}, \dots, z_{1T}$  are exogenous,  $p_{\Delta^{**}}$  has a normal distribution with mean  $\Pi_{\Delta^{**}}$  and variance  $\omega_{11}/\sum_{t=1}^T z_{1t}^2$ , where  $\omega_{11}$  is the variance of  $y_{1t}$ ,  $t = 1, \dots, T$ . Moreover,  $(\sum_{t=1}^T y_{1t}^2 - p_{\Delta^{**}}^2 \sum_{t=1}^T z_{1t}^2)/\omega_{11}$  has a  $\chi^2$  distribution with  $T - 1$  degrees of freedom and statistically independent of  $p_{\Delta^{**}}$ . Then  $T - 1$  times (13) has a noncentral  $F$  distribution with 1 and  $T - 1$  degrees of freedom and noncentrality parameter

$$(14) \quad \frac{\Pi_{\Delta^{**}}^2 \sum_{t=1}^T z_{1t}^2}{\omega_{11}} = \frac{(\beta_{12}\gamma_{21} - \beta_{22}\gamma_{11})^2}{\beta_{22}^2\sigma_{11} - 2\beta_{22}\beta_{12}\sigma_{12} + \beta_{12}^2\sigma_{22}} \sum_{t=1}^T z_{1t}^2.$$

If

$$(15) \quad \beta_{12}\gamma_{21} = \beta_{22}\gamma_{11},$$

then  $\Pi_{\Delta^{**}} = 0$  and the distribution is the central  $F$  distribution. A test at significance level  $\alpha$  is a procedure to "reject" when  $(T - 1)$  times an observed value of (13) is greater than the  $\alpha$  significance point of the  $F$  distribution with 1 and  $T - 1$  degrees of freedom.

The properties of any test are summarized in its power function, which is the probability of "rejection" as a function of the parameters. In this case the power is a monotonically increasing function of the noncentrality parameter (14). In particular the power is  $\alpha$  (the significance level) for all values of the parameters such that  $\Pi_{\Delta^{**}} = 0$ , that is, for (15). The null hypothesis for which the test is appropriate is, therefore  $\Pi_{\Delta^{**}} = 0$ , that is, that the noncentrality parameter (14) is 0. If the noncentrality parameter is small, the probability of rejection is small; if the parameter is large, the probability of rejection is large.

When equation (15) holds,  $\Pi_{\Delta^{**}}$ , a  $1 \times 1$  matrix, is zero. Condition (i) of Theorem 1 obtains. Then Theorem 1 says that some  $F$  exists such that, in the equivalent system transformed by  $F$ , the null hypothesis

$$(16) \quad \tilde{\beta}_{12} = \tilde{\gamma}_{11} = 0$$

obtains. In fact we have already seen that transformation: it is the one that yields (8), the reduced form.

In examining large-sample properties such as consistency of a test, it is customary to make the following assumption:

ASSUMPTION 1:  $\text{plim}_{T \rightarrow \infty} T^{-1} \mathbf{Z}' \mathbf{Z} = \mathbf{M}$ , where  $\mathbf{M}$  is nonsingular and  $\mathbf{Z} = (z'_1, \dots, z'_T)$  is a  $K \times T$  matrix of observations on  $(z_{\Delta t}, z_{\Delta \Delta t})'$ .

With this assumption, we have the following theorem:

THEOREM 2: Under Assumption 1, Conditions (i) and (ii) are equivalent to the following condition:

(iii) there exist vectors  $\bar{\beta}_\Delta, \bar{\gamma}_*$ , not both zero vectors, such that

$$(17) \quad \text{plim}_{T \rightarrow \infty} T^{-1} [\bar{\beta}_\Delta \mathbf{Y}'_\Delta + \bar{\gamma}_* \mathbf{Z}'_*] \mathbf{Z} = \mathbf{0}.$$

The proof of Theorem 2 is in the Appendix.

The contrary of Condition (iii) is used by Fisher and Kadane [3] and Kadane [5] as a condition for consistency of the test. That work and Theorem 2 proves that when the null hypothesis is that (i) and (ii) hold, and the alternative is that they do not, under Assumption 1 this test is consistent.

*Carnegie Mellon University  
and  
Stanford University*

*Manuscript received September, 1975; revision received March, 1976.*

APPENDIX

PROOF OF THEOREM 1: Condition (ii) implies Condition (i). Suppose first that such an  $\mathbf{F}$  exists. The reduced form equations (2) are still valid. By assumption of (5) the equation

$$(18) \quad \bar{\beta}_\Delta y'_{\Delta t} + \bar{\gamma}_* z'_{*t} = \bar{u}_{1t}$$

holds. Multiplying the first equation of (2) by  $\bar{\beta}_\Delta \neq \mathbf{0}$  and applying a familiar argument yields

$$(19) \quad \bar{\beta}_\Delta \Pi_{\Delta**} = \mathbf{0}.$$

Hence  $\rho(\Pi_{\Delta**}) \leq G_\Delta - 1$ .

Condition (i) implies Condition (ii). Now suppose there is some vector  $\beta_\Delta$  such that (4) holds. Then take

$$(20) \quad \mathbf{F} = \begin{pmatrix} \beta_\Delta & \mathbf{0} \\ \mathbf{F}_3 & \mathbf{F}_4 \end{pmatrix} \mathbf{B}^{-1},$$

where  $\mathbf{F}_3$  and  $\mathbf{F}_4$  are arbitrary conformable matrices so that  $\mathbf{F}$  is nonsingular. Then

$$(21) \quad \begin{aligned} \mathbf{F}(\mathbf{B} \quad \Gamma) &= \begin{pmatrix} \beta_\Delta & \mathbf{0} \\ \mathbf{F}_3 & \mathbf{F}_4 \end{pmatrix} \mathbf{B}^{-1} (\mathbf{B} \quad \Gamma) \\ &= \begin{pmatrix} \beta_\Delta & \mathbf{0} \\ \mathbf{F}_3 & \mathbf{F}_4 \end{pmatrix} (\mathbf{I} \quad -\Pi) \\ &= \begin{pmatrix} \beta_\Delta & \mathbf{0} \\ \mathbf{F}_3 & \mathbf{F}_4 \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} & -\Pi_{\Delta*} & -\Pi_{\Delta**} \\ \mathbf{0} & \mathbf{I} & -\Pi_{\Delta\Delta*} & -\Pi_{\Delta\Delta**} \end{pmatrix} \\ &= \begin{bmatrix} \beta_\Delta & \mathbf{0} & -\beta_\Delta \Pi_{\Delta*} & -\beta_\Delta \Pi_{\Delta**} \\ \mathbf{F}_3 & \mathbf{F}_4 & -\mathbf{F}_3 \Pi_{\Delta*} - \mathbf{F}_4 \Pi_{\Delta\Delta*} & -\mathbf{F}_3 \Pi_{\Delta**} - \mathbf{F}_4 \Pi_{\Delta\Delta**} \end{bmatrix} \\ &= \begin{bmatrix} \beta_\Delta & \mathbf{0} & -\beta_\Delta \Pi_{\Delta*} & \mathbf{0} \\ \mathbf{F}_3 & \mathbf{F}_4 & -\mathbf{F}_3 \Pi_{\Delta*} - \mathbf{F}_4 \Pi_{\Delta\Delta*} & -\mathbf{F}_3 \Pi_{\Delta**} - \mathbf{F}_4 \Pi_{\Delta\Delta**} \end{bmatrix}; \end{aligned}$$

so an  $\mathbf{F}$  of the required form exists. This proves Theorem 1.

To prove Theorem 2, it is sufficient, in view of Theorem 1, to prove that (iii) and (i) are equivalent under Assumption 1.

Condition (iii) implies Condition (i). Suppose  $\bar{\beta}_\Delta, \bar{\gamma}_*$  are not both zero, and

$$\begin{aligned} (22) \quad \mathbf{0} &= \text{plim}_{T \rightarrow \infty} T^{-1} [\bar{\beta}_\Delta \mathbf{Y}'_\Delta + \bar{\gamma}'_* \mathbf{Z}'_*] \mathbf{Z} \\ &= \text{plim}_{T \rightarrow \infty} T^{-1} [\bar{\beta}_\Delta (\Pi_{\Delta*} \mathbf{Z}'_* + \Pi_{\Delta**} \mathbf{Z}'_{**} + \mathbf{V}'_\Delta) + \bar{\gamma}'_* \mathbf{Z}'_*] (\mathbf{Z}_*, \mathbf{Z}_{**}) \\ &= [(\bar{\gamma}'_* + \bar{\beta}_\Delta \Pi_{\Delta*}) \mathbf{M}_{**,*} + \bar{\beta}_\Delta \Pi_{\Delta**} \mathbf{M}_{***,*}, (\bar{\gamma}'_* + \bar{\beta}_\Delta \Pi_{\Delta*}) \mathbf{M}_{**,*} + \bar{\beta}_\Delta \Pi_{\Delta**} \mathbf{M}_{***,*}]. \end{aligned}$$

Now if  $\bar{\beta}_\Delta = \mathbf{0}$ , then  $\bar{\gamma}'_* \mathbf{M}_{**,*} = \mathbf{0}$ . Since  $\mathbf{M}_{**,*}$  is positive definite,  $\bar{\gamma}'_* = \mathbf{0}$ , which contradicts the hypothesis. Hence  $\bar{\beta}_\Delta \neq \mathbf{0}$ .

The equation (22) can be written

$$(23) \quad \mathbf{0} = (\bar{\gamma}'_* + \bar{\beta}_\Delta \Pi_{\Delta*}, \bar{\beta}_\Delta \Pi_{\Delta**}) \begin{pmatrix} \mathbf{M}_{**,*} & \mathbf{M}_{***,*} \\ \mathbf{M}_{***,*} & \mathbf{M}_{***,*} \end{pmatrix}.$$

The positive-definiteness of  $\mathbf{M}$  implies

$$(24) \quad \bar{\beta}_\Delta \Pi_{\Delta**} = \mathbf{0}.$$

Hence  $\rho(\Pi_{\Delta**}) \leq G_\Delta - 1$ .

Condition (i) implies Condition (iii). Suppose there is a vector  $\bar{\beta}_\Delta \neq \mathbf{0}$  such that (24) holds. Let  $\bar{\gamma}'_* = -\bar{\beta}_\Delta \Pi_{\Delta*}$ . Then by the computation above, (17) holds and  $(\bar{\beta}_\Delta, \bar{\gamma}'_*)$  are not both zero vectors. This proves Theorem 2.

REFERENCES

[1] ANDERSON, T. W., AND HERMAN RUBIN: "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, 20 (1949), 46-63.

[2] ———: "The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, 21 (1950), 570-582.

[3] FISHER, FRANKLIN M., AND JOSEPH B. KADANE: "The Covariance Matrix of the Limited Information Estimator and the Identification Test: Comment," *Econometrica*, 40 (1972), 901-903.

[4] KADANE, JOSEPH B.: "Testing Overidentifying Restrictions When the Disturbances are Small," *Journal of the American Statistical Association*, 65 (1970), 182-185.

[5] ———: "Testing a Subset of the Overidentifying Restrictions," *Econometrica*, 42 (1974), 853-867.

[6] KOOPMANS, TJALLING C.: "Identification Problems in Economic Model Construction," in *Studies in Econometric Method*, Cowles Commission Monograph No. 14, edited by William C. Hood and Tjalling C. Koopmans. New York: Wiley, 1953, 27-48.

[7] KOOPMANS, TJALLING C., AND WILLIAM C. HOOD: "The Estimation of Simultaneous Linear Economic Relationships," in *Studies in Econometric Method*, Cowles Commission Monograph No. 14, edited by William C. Hood and Tjalling C. Koopmans. New York: Wiley, 1953, 112-199.

[8] LIU, TA-CHUNG, AND WILLIAM J. BREEN: "The Covariance Matrix of the Limited Information Estimator and the Identification Test," *Econometrica*, 37 (1969), 222-227.

[9] ———: "The Covariance Matrix of the Limited Information Estimator and the Identification Test: A Reply," *Econometrica*, 40 (1972), 905-906.