Econometrica, Vol. 42, No. 5 (September, 1974)

TESTING A SUBSET OF THE OVERIDENTIFYING RESTRICTIONS

BY JOSEPH B. KADANE[1]

A family of tests of significance is developed for coefficients in a single equation of a simultaneous system. Different members of this family are distinguished by the k-class estimator on which they are based, and on the alternative hypothesis against which they test. The size of the test is found when the disturbances are small, and the test is shown to be consistent if plim $k = 1$ or $k = \lambda$, the limited information maximum likelihood value.

1. INTRODUCTION

SEVERAL KINDS of tests have been proposed for the estimated coefficients of an equation in a system of simultaneous linear equations.

Anderson and Rubin [**1**] proposed a test that the restrictions on the coefficients of the predetermined variables are correct against the alternative hypothesis that one or more of the restrictions is false. Hood and Koopmans [**9**] showed that the same test statistic can be used to test the null hypothesis that the restrictions on the coefficients of both predetermined and endogenous variables are correct against the alternative that one or more of the restrictions is false. This test statistic is closely associated with the limited information maximum likelihood estimator of the structural coefficients. More recently Basmann [**3**] proposed a test with the same null and alternative hypotheses as Koopmans and Hood, but with two-stage least squares estimators used in place of limited information maximum likelihood estimators. The consistency of these tests has been debated by Liu and Breen [**14** and **15**], and Fisher and Kadane [**6**].

A second kind of test has a much more restricted alternative, namely that a single specified restriction is false. Such tests are suggested in Dhrymes [**4**; **5**, pp. 272–277].

Dhrymes [**4**, pp. 222–223] also proposed a third kind of test which generalizes the two kinds above by allowing an arbitrary alternative hypothesis containing the null hypothesis. He also proposed a new test statistic based on two-stage least squares. This paper proposes a new family of statistics to test the coefficients of any k-class estimator, including limited information maximum likelihood against this more general alternative.

The tests here proposed have an especially close relationship to the first kind of test above. When the alternative is that one or more of the restrictions on the coefficients of either endogenous or exogenous variables is wrong, the limited information test statistic is that of Anderson and Rubin, and Hood and Koopmans, and the two-stage least squares test statistic is that of Basmann.

The distribution of these test statistics is also analyzed, using the small-σ method (Kadane [**10** and **11**]). The lowest order term, σ^0, does not depend on which k-class method was used to estimate, and yields a simple F distribution under the null hypothesis, generalizing the results of Kadane on the Anderson-

[1] Financial support from the Center for Naval Analyses and National Science Foundation Grant GS-38609 is gratefully acknowledged.

Rubin-Hood-Koopmans statistic and on the Basmann statistic. When the alternative allows only a single specified restriction, an intuitively appealing test would compare some function of the observations to a t distribution with degrees of freedom equal to the number of time periods (T) minus the number of parameters estimated. Dhrymes [**4**] was not able to justify such a test with his theory, but the theory here does give a test with the above degrees of freedom.

Finally, the consistency of these new tests is investigated. They are shown to be large-sample consistent, generalizing results of Fisher and Kadane [**6**] on the consistency of the Anderson-Rubin-Hood-Koopmans test.

An interesting research question would be to compare the two-stage least squares test proposed here to the Dhrymes test and the asymptotic test. Is the Dhrymes test consistent? Is one test universally more powerful than the others? Initial explorations in this direction have been conducted by Morgan and Vandaele [**17**], and by Maddala [**16**].

2. THE GENERAL DECISION PROBLEM

The development of a simultaneous equation econometric model is often shrouded in uncertainty. The specification of restrictions on various equations is never as easy as a text would lead one to believe. Hence a working econometrician may desire ways of testing models to see which specification best fits the data and whatever prior opinions he or she may have. There are several approaches to this problem currently advocated.

The most classical approach asks the user to specify a strict null hypothesis and an alternative hypothesis which is more general. Thus certain coefficients can be tested against zero. One contribution of this paper is to allow a more general assortment of alternative hypotheses than had been available. However, tests of significance suffer from certain criticisms, which apply here. First, the null hypothesis is not usually really believed. Why, then, should it be a subject of special study? Second, as the sample size grows, any consistent test of significance will reject the null hypothesis unless it is exactly, literally true. Thus even if the restrictions are nearly correct, sufficient data will permit rejection of the null hypothesis (see Kadane, Lewis, and Ramage [**12**] for a case where this consideration was important).

A second approach, the Bayesian method, asks for the econometrician's opinion in great detail. Since typical simultaneous equation models involve many parameters, eliciting such an opinion would be very difficult (Savage [**21**]). Phrased in Bayesian terms, the problem is to choose the right model. Research in this area has thus far been concentrated on single equation regression models. So far, reasonable situations have not been found in which the simpler model is chosen, since the more complicated model always has more information and thus fits better. Therefore to date most applications require the two models being compared to have the same number of parameters. (See [7 and **23**].) Further research is needed before adequate Bayesian procedures can be found, even with a solution to the elicitation problem.

A third approach is called "data analysis" [22]. It has much less formal structure, and consists mainly in trying various different models, estimators, and ways of looking at the data, and choosing a few to report. While some variant of this method, possibly tempered with some tests of significance, is probably the most widely used, it is hard to analyze and justify.

Thus although the results of this paper are in the mold of classical tests of significance, that should not be taken as evidence that I favor the first approach above. No doubt all three are legitimate areas for modern research in econometric methods.

3. DERIVATION OF THE LIKELIHOOD RATIO TEST AND ITS ANALOGS

Consider a complete system

$$YB + Z\Gamma + U = 0 \tag{1}$$

where Y is a $T \times G$ matrix of endogenous variables and Z is a $T \times K$ matrix of exogenous variables. B is a $G \times G$ non-singular matrix of parameters, Γ is a $K \times G$ matrix of parameters, and U is a $T \times G$ matrix of jointly normal residuals with zero means and covariances $Eu_{ti}u_{t'j} = \sigma_{ij}\delta_{tt'}$ where $\delta_{tt'}$ is 1 if $t = t'$ and zero otherwise. Note that no lagged endogenous variables are permitted in the system (1).

The null hypothesis specifies certain restrictions on the coefficients of the first equation, so that it may be written

$$y = Y_1\beta_1 + Z_1\gamma_1 + u \tag{2}$$

where Y is partitioned $Y = (y, Y_1, Y_2, Y_3) : T \times 1 + G_1 + G_2 + G_3$; Z is partitioned $Z = (Z_1, Z_2, Z_3) : T \times K_1 + K_2 + K_3$; and u is partitioned $u = (u, U^1) : T \times 1 + (G - 1)$. The vectors β_1 and γ_1 are conformable vectors of parameters.

The alternative hypothesis specifies a subset of the restrictions specified under the null hypothesis. The remaining restrictions—those specified by the null hypothesis but not by the alternative, are the hypotheses being tested.

With only the assumptions of the alternative hypothesis, the first equation may be written

$$y = Y_1\beta_1 + Y_2\beta_2 + Z_1\gamma_1 + Z_2\gamma_2 + u; \tag{3}$$

again the β's and γ's are conformable.

A test of the null hypothesis (2) against the alternative (3) may thus be thought of as a test of $\beta_2 = 0$ and $\gamma_2 = 0$. As examples, if a single β is being tested, then $K_2 = 0$ and $G_2 = 1$. Similarly if a single γ is being tested, $G_2 = 0$ and $K_2 = 1$. If all the β's and γ's are being tested, $K_3 = 0$, and $G_3 = 0$.

Under each hypothesis i, the concentrated log-likelihood function is well-known to be of the form

$$L^i = k' + \frac{T}{2}\log l_i \tag{4}$$

where k is a constant and l_1 and l_2 are given by

$$(5) \qquad l_1 = \min_{\beta_*^1} \frac{\beta_*^{1\prime} Y_*^{1\prime} \bar{P}_{Z_1} Y_*^1 \beta_*^1}{\beta_*^{1\prime} Y_*^{1\prime} \bar{P}_Z Y_*^1 \beta_*^1} = \frac{\hat{\beta}_*^{1\prime} Y_*^{1\prime} \bar{P}_{Z_1} Y_*^1 \hat{\beta}_*^1}{\hat{\beta}_*^{1\prime} Y_*^{1\prime} \bar{P}_Z Y_*^1 \hat{\beta}_*^1}$$

and

$$(6) \qquad l_2 = \min_{\beta_*^2} \frac{\hat{\beta}_*^{2\prime} Y_*^{2\prime} \bar{P}_{Z_1,Z_2} Y_*^2 \beta_*^2}{\beta_*^{2\prime} Y_*^{2\prime} \bar{P}_Z Y_*^2 \beta_*^2} = \frac{\hat{\beta}_*^{2\prime} Y_*^{2\prime} \bar{P}_{Z_1,Z_2} Y_*^2 \hat{\beta}_*^2}{\hat{\beta}_*^{2\prime} Y_*^{2\prime} \bar{P}_Z Y_*^2 \hat{\beta}_*^2}$$

where $Y_*^{1\prime} = (y', Y_1')$, $Y_*^{2\prime} = (y', Y_1', Y_2')$, and $\bar{P}_F = I - F(F'F)^{-1}F'$ for any matrix F.

The minimizing choice in (5) and (6), when normalized, can be written as $(-1, \hat{\beta}^{i\prime})$ $(i = 1, 2)$, where $\hat{\beta}^1$ is the limited information maximum likelihood estimator of β_1 in equation (2), and $\hat{\beta}^2$ is the limited information maximum likelihood estimator of (β_1, β_2) in equation (3). Note that when the alternative hypothesis allows (3) to be underidentified, $\hat{\beta}^2$ will not be unique. Nevertheless, any minimizing choice yields $l_2 = 1$. This causes no problems in the subsequent analysis.

The log-likelihood ratio is

$$(7) \qquad L^1 - L^2 = \frac{T}{2} \log (l_1/l_2),$$

which is equivalent to l_1/l_2. Hence the likelihood ratio statistic is equivalent to l_1/l_2.

THEOREM 1: *The likelihood ratio statistic for testing the null hypothesis* (2) *against the alternative hypothesis* (3) *is equivalent to* l_1/l_2, *where* l_1 *is given by* (5) *and* l_2 *is given by* (6).

In the special case where the alternative hypothesis specifies no restrictions on the equation, it is well-known that $l_2 = 1$. Hence the statistic in (7) simplifies to l_1, which is the Anderson-Rubin-Hood-Koopmans statistic.

Possibly some other estimator might be used instead of limited information maximum likelihood in (5) and (6). Let $l_1(k)$ be the expression in (5) with the k-class estimator with parameter k for equation (2) substituted in place of limited information maximum likelihood. Similarly let $l_2(k)$ be the expression in (6) with the k-class estimator with parameter k for equation (3) substituted in place of limited information maximum likelihood there.

The analogous statistic to (7) is then $l_1(k)/l_2(k)$. For any k-class estimator, when the model under the alternative hypothesis is just or underidentified, $l_2(k) \equiv 1$, and so $l_1(k)/l_2(k) = l_1(k)$. The Basmann test [3] is $l_1(1)$ with this alternative hypothesis.

4. DISTRIBUTION OF THE LIKELIHOOD RATIO STATISTIC AND ITS ANALOGS WHEN THE DISTURBANCES ARE SMALL

In [**10**] I introduced an approach to the distribution of econometric quantities based on asymptotic expansion in σ, the variance of the disturbances in the system. As $\sigma \to 0$, the system becomes less and less variable around the true regression. Applied to the distribution of the Anderson-Rubin-Hood-Koopmans statistic, the first order term, which does not depend on σ, gave a distribution which approaches, as $T \to \infty$, the large-sample results of Anderson and Rubin. Additionally the first order term of the distribution of Basmann's test statistic is the same F distribution found in a Monte Carlo study [**2**]. Applied to the statistic found in Section 3, we obtain:

THEOREM 2: *Asymptotically as $\sigma \to 0$, when the null hypothesis is true,*

$$\text{(8)} \qquad \frac{T - K + L_2}{L_1 - L_2}[(l_1(k)/l_2(k)) - 1]$$

has the Fischer variance ratio distribution $F_{L_1 - L_2, T - K + L_2}$ where $L_1 = K_2 + K_3 - G_1$ and $L_2 = [K_3 - G_1 - G_2]^+$ are the degrees of overidentification of the two models, and K is the number of exogenous variables in the system. This holds for every member of the k class with k fixed and for limited information maximum likelihood.

PROOF OF THEOREM 2: From Kadane [**10**], when the null hypothesis is true,

$$\text{(9)} \qquad l_1(k) = \frac{u'\bar{P}_{X^1}u}{u'\bar{P}_Z u} + 0_p(\sigma),$$

where $X^1 = [Z\Pi_1, Z_1]$.

Similarly $l_2(k)$ can be thought of as an analogous equation in a model with more included endogenous and exogenous variables (but some "true" coefficients are zero, even though they are estimated). Then, by exactly the same argument as before,

$$\text{(10)} \qquad l_2(k) = \frac{u'\bar{P}_{X^2}u}{u'\bar{P}_Z u} + 0_p(\sigma),$$

where $X^2 = [Z\Pi_1, Z\Pi_2, Z_1, Z_2]$.

Putting (9) and (10) together,

$$\frac{l_1(k)}{l_2(k)} = \frac{\dfrac{u'\bar{P}_{X^1}u}{u'\bar{P}_Z u} + 0_p(\sigma)}{\dfrac{u'\bar{P}_{X^2}u}{u'\bar{P}_Z u} + 0_p(\sigma)}$$

$$= \frac{u'\bar{P}_{X^1}u}{u'\bar{P}_{X^2}u} + 0_p(\sigma).$$

Since the columns of X^1 are included among those in X^2, $X^1\bar{P}_{X^2} = 0$. Hence $\bar{P}_{X^1}\bar{P}_{X^2} = \bar{P}_{X^2}$, so

$$\bar{P}_{X^2}[\bar{P}_{X^1} - \bar{P}_{X^2}] = 0. \tag{11}$$

Then

$$l_1(k)/l_2(k) = 1 + \frac{u'[\bar{P}_{X^1} - \bar{P}_{X^2}]u}{u'\bar{P}_{X^2}u} + 0_p(\sigma). \tag{12}$$

Here $u'[\bar{P}_{X^1} - \bar{P}_{X^2}]u$ and $u'\bar{P}_{X^2}u$ have independent χ^2 distributions, using (11). Their degrees of freedom are:

$$\operatorname{tr} \bar{P}_{X^2} = T - K + L_2 \qquad \text{and}$$
$$\operatorname{tr}[\bar{P}_{X^1} - \bar{P}_{X^2}] = T - K + L_1 - (T - K + L_2) = L_1 - L_2.$$

Therefore

$$[(l_1(k)/l_2(k)) - 1]\frac{T - K + L_2}{L_1 - L_2}$$

has an asymptotic (as $\sigma \to 0$) F distribution with degrees of freedom $L_1 - L_2$ and $T - K + L_2$. *Q.E.D.*

Less transformation is required in the following equivalent statement:

COROLLARY 1: *$l_2(k)/l_1(k)$ has an asymptotic (as $\sigma \to 0$) Beta distribution with parameters $(T - K + L_2)/2$ and $(L_1 - L_2)/2$. (See, for instance, Rao [**20**, p. 135]).*

In the special case in which the model is just or underidentified under the alternative hypothesis, $L_2 = 0$, $l_2 \equiv 1$, and Theorem 2 reduces to the Theorem of Kadane [**10**]. In the special case in which only a single restriction is being tested, $L_1 - L_2 = 1$, and Theorem 2 reduces to

$$(T - K + L_2)[(l_1/l_2) - 1] \sim F_{1,T-K+L_2}.$$

Notice that $T - K + L_2 = T - K_1 - K_2 - G_1 - G_2$ is a natural degrees of freedom parameter, namely the number of time periods minus the number of estimated parameters.

5. CONSISTENCY OF THE TESTS

One interesting property a test may have is to be consistent against certain alternative hypotheses, which means that under each of those alternative hypotheses, the power of the test approaches one as the sample size increases [**13**, p. 305]. This section reports on investigations into the consistency of the tests proposed in Sections 3 and 4.

Note that consistency as defined above is a *large-sample*, not a *small-σ* concept. In this section all probability limits are as $T \to \infty$. For a sophisticated treatment of the power of the Anderson-Rubin-Hood-Koopmans test in the small-σ context, see Ramage [**19**].

THEOREM 3: *Suppose*: (*i*) (*a*) *plim* $k = 1$ *under both the null and alternative hypotheses, or* (*b*) *that* $k = \lambda$, *the limited information maximum likelihood value*; (*ii*) *plim* $Z'Z/T = M$, *a positive definite matrix*; (*iii*) *plim* $Z'U/T = 0$; *and* (*iv*) *under the alternative hypothesis, plim* $T^{-1}Z'[y\tilde{\alpha} + Y_1\tilde{\beta} + Z_1\tilde{\gamma}] \neq 0$ *if* $\tilde{\alpha}$, $\tilde{\beta}$, *and* $\tilde{\gamma}$ *are not all zero.*

Then the test $l_1(k)/l_2(k)$ *is consistent.*

The assumptions of Theorem 3 deserve some comment. Assumption (i) limits the k-class estimators being considered, but includes two-stage least squares, $k = 1 + (L-1)T^{-1}$ discussed by Nagar [**18**], $k = 1 + (L-1)(T-K)^{-1}$ discussed by Kadane [**11**, p. 727], and limited information maximum likelihood. It excludes ordinary least squares. Assumptions (ii) and (iii) are standard for large-sample theory [**5**]. Assumption (iv) says, roughly, that under the alternative hypothesis there is no way "to manipulate the model to obtain a linear expression in Y_1^* and Z_1 which is equal to the disturbance term only" [**6**]. Thus the equation in question is overidentified under the alternative hypothesis.

PROOF OF THEOREM 3: First, some notation is needed. Let us partition $B^{-1} = (b, B_1, B_2, B_3)$ where B_i is $G \times G_i$ $(i = 1, 2, 3)$ so that

$$Y_i = -Z\Gamma B_i - UB_i = Z\Pi_i + V_i \qquad (i = 1, 2, 3).$$

Also partition M as follows:

$$M = \operatorname*{plim}_{T\to\infty} \frac{Z'Z}{T} = \operatorname*{plim}_{T\to\infty} \begin{bmatrix} \frac{Z_1'Z_1}{T} & \frac{Z_1'Z_2}{T} & \frac{Z_1'Z_3}{T} \\ \frac{Z_2'Z_1}{T} & \frac{Z_2'Z_2}{T} & \frac{Z_2'Z_3}{T} \\ \frac{Z_3'Z_1}{T} & \frac{Z_3'Z_2}{T} & \frac{Z_3'Z_3}{T} \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix}.$$

Finally let

$$\operatorname{plim} \frac{Z'Z_i}{T} = \begin{bmatrix} M_{1i} \\ M_{2i} \\ M_{3i} \end{bmatrix} = M_{.i} \qquad (i = 1, 2, 3)$$

and let $M_{i.} = M'_{.i}$. Next some consequences of assumption (iv) are recorded in Lemma 1.

LEMMA 1: *If assumption (iv) holds, then under the alternative hypothesis, (a) Π_1 is of full rank G_1; (b) $\{M\Pi_2\beta_2 + M_{.2}\gamma_2, M\Pi_1, M_{.1}\}$ are a set of $1 + G_1 + K_1$ linearly independent K-vectors.*

PROOF: Suppose $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\gamma}$ are not all zero. Then:

$$\begin{aligned}
0 &\neq \text{plim } T^{-1}Z'[(Y_1\beta_1 + Y_2\beta_2 + Z_1\gamma_1 + Z_2\gamma_2 + u)\tilde{\alpha} + Y_1\tilde{\beta}_1 + Z_1\tilde{\gamma}] \\
&= \text{plim } T^{-1}Z'[(Z\Pi_1\beta_1 + Z\Pi_2\beta_2 + Z_1\gamma_1 + Z_2\gamma_2)\tilde{\alpha} + Z_1\Pi_1\tilde{\beta} + Z_1\tilde{\gamma}] \\
&= (M\Pi_1\beta_1 + M\Pi_2\beta_2 + M_{.1}\gamma_1 + M_{.2}\gamma_2)\tilde{\alpha} + M\Pi_1\tilde{\beta} + M_{.1}\tilde{\gamma} \\
&= (M\Pi_2\beta_2 + M_{.2}\gamma_2)\tilde{\alpha} + M\Pi_1(\beta_1\tilde{\alpha} + \tilde{\beta}) + M_{.1}(\gamma_1\tilde{\alpha} + \tilde{\beta}).
\end{aligned}$$

Let $\tilde{\alpha}_1 = \tilde{\alpha}$, $\tilde{\beta}_1 = \beta_1\tilde{\alpha} + \tilde{\beta}$, and $\tilde{\gamma}_1 = \gamma_1\tilde{\alpha} + \tilde{\gamma}$. Then $(\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma})$ are not all zero if and only if $(\tilde{\alpha}_1, \tilde{\beta}_1, \tilde{\gamma}_1)$ are not all zero. Thus $\{M\Pi_2\beta_2 + M_{.2}\gamma_2, M\Pi_1, M_{.1}\}$ are a linearly independent set of vectors of length K. Finally, $\rho(M\Pi_1) = G_1$ by linear independence. Since M is non-singular, $\rho(\Pi_1) = G_1$. *Q.E.D.*

LEMMA 2: *Under the alternative hypothesis, and assuming (i) (a) of Theorem 3,*

$$plim\,(\beta_k - \beta) = (\Pi_1'C\Pi_1)^{-1}\Pi_1'C(\Pi_2\beta_2 + I_2\gamma_2)$$

where

$$C = M - M_{.1}M_{11}^{-1}M_{1.} \qquad \text{and} \qquad I_2 = \begin{array}{c} \begin{pmatrix} 0 \\ \cdots \\ I \\ \cdots \\ 0 \end{pmatrix} \end{array} \begin{array}{l} {}^{K_2}K_1 \\ \\ K_2 : K \times K_2. \\ \\ K_3 \end{array}$$

PROOF: The general k-class estimator is given by

$$\begin{aligned}
\begin{pmatrix} \beta_k \\ \gamma_k \end{pmatrix} &= \begin{pmatrix} Y_1'[I - k\bar{P}_Z]Y_1 & Y_1'Z_1 \\ Z_1'Y_1 & Z_1'Z_1 \end{pmatrix}^{-1} \begin{pmatrix} Y_1'[I - k\bar{P}] \\ Z_1' \end{pmatrix} y \\
&= \begin{pmatrix} \beta \\ \gamma \end{pmatrix} + \begin{pmatrix} Y_1'[I - k\bar{P}_Z]Y_1 & Y_1'Z_1 \\ Z_1'Y_1' & Z_1'Z_1 \end{pmatrix}^{-1} \begin{pmatrix} Y_1'[I - k\bar{P}_Z] \\ Z_1' \end{pmatrix} \\
&\quad \times (Y_2\beta_2 + Z_2\gamma_2 + u).
\end{aligned}$$

Therefore

$$\text{plim} \begin{pmatrix} \beta_k - \beta \\ \gamma_k - \gamma \end{pmatrix} = \text{plim} \begin{pmatrix} \dfrac{Y_1'[I - k\bar{P}_Z]Y_1}{T} & \dfrac{Y_1'Z_1}{T} \\ \dfrac{Z_1'Y_1}{T} & \dfrac{Z_1'Z_1}{T} \end{pmatrix}^{-1} \times \begin{pmatrix} \dfrac{Y_1'[I - k\bar{P}_Z](Y_2\beta_2 + Z_2\gamma_2 + u)}{T} \\ \dfrac{Z_1'(Y_2\beta_2 + Z_2\gamma_2 + u)}{T} \end{pmatrix}.$$

The probability limit of the inverse is already available in the literature (Dhrymes [**5**, p. 179]):

$$\operatorname{plim}\begin{pmatrix} \dfrac{Y_1'[I - k\bar{P}_Z]Y_1}{T} & \dfrac{Y_1'Z_1}{T} \\ \dfrac{Z_1'Y_1}{T} & \dfrac{Z_1'Z_1}{T} \end{pmatrix}^{-1} = \begin{bmatrix} \Pi_1'M\Pi_1 & \Pi_1'M_{.1} \\ M_{1.}\Pi_1 & M_{11} \end{bmatrix}^{-1}.$$

Now

$$\begin{aligned}
&\operatorname{plim}\frac{Y_1'[I - k\bar{P}_Z](Y_2\beta_2 + Z_2\gamma_2 + u)}{T} \\
&= \operatorname{plim}\frac{(\Pi_1'Z' + \sigma V_1')(I - k\bar{P}_Z)(Z\Pi_2\beta_2 + V_2\beta_2 + Z_2\gamma_2 + u)}{T} \\
&= \Pi_1'M\Pi_2\beta_2 + \Pi_1'M_{.2}\gamma_2 + \operatorname{plim}\frac{B_1'U'[I - k\bar{P}_Z]UB_2\beta_2}{T} \\
&\quad - \operatorname{plim}\frac{B_1U[I - k\bar{P}_Z]u}{T} \\
&= \Pi_1'M\Pi_2\beta_2 + \Pi_1'M_{.2}\gamma_2 + B_1'\Sigma B_2\beta_2 \operatorname{plim}\left(\frac{T - k(T - K)}{T}\right) \\
&\quad - B_1'\sigma_{.1} \operatorname{plim}\left(\frac{T - k(T - K)}{T}\right)
\end{aligned}$$

where $\sigma_{.1} = EU'u/T$, the first row of Σ. Now since

$$\operatorname{plim}\left(\frac{T - k(T - K)}{T}\right) = \operatorname{plim}\left(1 - k + k\frac{K}{T}\right) = 0,$$

$$\operatorname{plim}\frac{Y_1'[I - k\bar{P}_Z](Y_2\beta_2 + Z_2\gamma_2 + u)}{T} = \Pi_1'M\Pi_2\beta_2 + \Pi_1'M_{.2}\gamma_2.$$

Finally,

$$\begin{aligned}
&\operatorname{plim}\frac{Z_1'(Y_2\beta_2 + Z_2\gamma_2 + u)}{T} \\
&= \operatorname{plim}\frac{Z_1'[Z\Pi_2\beta_2 + V_2\beta_2 + Z_2\gamma_2 + \sigma u]}{T} \\
&= M_{1.}\Pi_2\beta_2 + M_{12}\gamma_2.
\end{aligned}$$

Then

$$\operatorname{plim}\begin{pmatrix}\beta_k-\beta\\ \gamma_k-\gamma\end{pmatrix}=\begin{bmatrix}\Pi'_1M\Pi_1 & \Pi'_1M_{.1}\\ M_{1.}\Pi_1 & M_{11}\end{bmatrix}^{-1}\begin{bmatrix}\Pi'_1M\Pi_2\beta_2+\Pi'_1M_{.2}\gamma_2\\ M_{1.}\Pi_2\beta_2+M_{12}\gamma_2\end{bmatrix}$$

$$=\begin{bmatrix}(\Pi'_1C\Pi_1)^{-1} & -(\Pi'_1C\Pi_1)^{-1}\Pi'_1M_{.1}M_{11}^{-1}\\ -M_{11}^{-1}M_{1.}\Pi_1(\Pi'_1C\Pi_1)^{-1} & M_{11}^{-1}+M_{11}^{-1}M_{1.}\Pi_1(\Pi'_1C\Pi_1)^{-1}\Pi'_1M_{.1}M_{11}^{-1}\end{bmatrix}$$

$$\begin{bmatrix}\Pi'_1M\Pi_2\beta_2+\Pi'_1M_{.2}\gamma_2\\ M_{1.}\Pi_2\beta_2+M_{12}\gamma_2\end{bmatrix}$$

$$=\begin{bmatrix}(\Pi'_1C\Pi_1)^{-1}\Pi'_1C(\Pi_2\beta_2+I_2\gamma_2)\\ M_{11}^{-1}M_{1.}(I-\Pi_1(\Pi'_1C\Pi_1)^{-1}\Pi'_1C)(\Pi_2\beta_2+I_2\gamma_2)\end{bmatrix}$$

Q.E.D. Lemma 2.

Then

$$\begin{aligned}\bar{P}_{Z_1}Y'_*\hat{\beta}_k &= \bar{P}_{Z_1}(y, Y_1)\begin{pmatrix}-1\\ \hat{\beta}_k\end{pmatrix}\\
&= \bar{P}_{Z_1}(-y+Y_1\hat{\beta}_k)\\
&= \bar{P}_{Z_1}(Y_1\hat{\beta}_k - Y_1\beta - Z_1\gamma_1 - Z_2\gamma_2 - Y_2\beta_2 - u)\\
&= \bar{P}_{Z_1}(Y_1(\hat{\beta}_k-\beta) - Z_2\gamma_2 - Y_2\beta_2 - u)\\
&= \bar{P}_{Z_1}((Z\Pi_1(\hat{\beta}_k-\beta) - Z_2\gamma_2 - Z\Pi_2\beta_2)\\
&\quad + (-UB_1(\hat{\beta}_k-\beta)+UB_2\beta_2-u))\\
&= \bar{P}_{Z_1}(N+S),\end{aligned}$$

where

$$N = Z\Pi_1(\hat{\beta}_k-\beta) - Z_2\gamma_2 - Z\Pi_2\beta_2$$

and

$$S = -UB_1(\hat{\beta}_k-\beta)+UB_2\beta_2-u.$$

Now

$$\begin{aligned}\operatorname{plim} l_1(k) &= \operatorname{plim}\frac{(N'+S')\bar{P}_{Z_1}(N+S)}{(N'+S')\bar{P}_Z(N+S)}\\
&= \frac{\operatorname{plim}\dfrac{(N'+S')\bar{P}_{Z_1}(N+S)}{T}}{\operatorname{plim}\dfrac{(N'+S')\bar{P}_Z(N+S)}{T}}\\
&= \frac{\operatorname{plim}\dfrac{N'\bar{P}_{Z_1}N}{T}+\operatorname{plim}\dfrac{S'\bar{P}_{Z_1}S}{T}}{\operatorname{plim}\dfrac{N'\bar{P}_ZN}{T}+\operatorname{plim}\dfrac{S'\bar{P}_ZS}{T}}.\end{aligned}$$

Now $N'\bar{P}_Z = 0$ by inspection, so that plim $N'\bar{P}_Z N/T = 0$. Let

$$e = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

a G-dimensional unit vector. Then

$$\begin{aligned} S &= -U[B_1(\hat{\beta}_k - \beta_1) + B_2\beta_2 - e] \\ &= -UD, \end{aligned}$$

where $D = B_1(\hat{\beta}_k - \beta_1) + B_2\beta_2 - e$, and let

$$\begin{aligned} \bar{D} &= \operatorname{plim} D \\ &= \operatorname{plim} B_1(\hat{\beta}_k - \beta_1) + B_2\beta_2 - e \\ &= B_1(\Pi_1' C \Pi_1)^{-1}\Pi_1' C(\Pi_2\beta_2 + I_2\gamma_2) + B_2\beta_2 - e, \end{aligned}$$

$$e = B^{-1}\begin{pmatrix} -1 \\ \beta_1 \\ \beta_2 \\ 0 \end{pmatrix} = (b, B_1, B_2, B_3)\begin{pmatrix} -1 \\ \beta_1 \\ \beta_2 \\ 0 \end{pmatrix} = -b + B_1\beta_1 + B_2\beta_2.$$

Hence $B_2\beta_2 - e = b - B_1\beta_1$. Then

$$\begin{aligned} \bar{D} &= b + B_1[(\Pi_1' C \Pi_1)^{-1}\Pi_1' C(\Pi_2\beta_2 + I_2\gamma_2) - \beta_1] \\ &= (b, B_1, B_2, B_3)\begin{pmatrix} 1 \\ (\Pi_1' C \Pi_1)^{-1}\Pi_1' C(\Pi_2\beta_2 + I_2\gamma_2) - \beta_1 \\ 0 \\ 0 \end{pmatrix}. \end{aligned}$$

Now if $\bar{D}$ were zero, then because of the invertibility of $B^{-1} = (b, B_1, B_2, B_3)$, as a consequence, it would be true that

$$\begin{pmatrix} 1 \\ (\Pi_1' C \Pi_1)^{-1}\Pi_1' C(\Pi_2\beta_2 + I_2\gamma_2) - \beta_1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

which is impossible. Hence $\bar{D} \neq 0$. Then

$$\operatorname{plim} \frac{S'\bar{P}_Z S}{T} = \bar{D}' \operatorname{plim}\left(\frac{U'\bar{P}_Z U}{T}\right)\bar{D}$$
$$= \bar{D}'\Sigma\bar{D} \operatorname{plim}\left(\frac{T-K}{T}\right)$$
$$= \bar{D}'\Sigma\bar{D} > 0.$$

Also, similarly,

$$\operatorname{plim} \frac{S'\bar{P}_{Z_1} S}{T} = \bar{D}'\Sigma\bar{D} \operatorname{plim}\left(\frac{T-K_1}{T}\right)$$
$$= \bar{D}'\Sigma\bar{D} > 0.$$

Finally it remains for us to analyze the term $\operatorname{plim} N'\bar{P}_{Z_1}N/T$.

$$N = Z\Pi_1(\Pi_1'C\Pi_1)^{-1}\Pi_1'C(\Pi_2\beta_2 + I_2\gamma_2) - Z_2\gamma_2 - Z\Pi_2\beta_2$$
$$= Z\{I - \Pi_1(\Pi_1'C\Pi_1) - 1\Pi_1'C\}(\Pi_2\beta_2 + I_2\gamma_2).$$

Then

$$\operatorname{plim} \frac{N'\bar{P}_{Z_1}N}{T} = (\gamma_2'I_2' + \beta_2'\Pi_2')[1 - C\Pi_1(\Pi_1'C\Pi_1)^{-1}\Pi_1']$$
$$\times C[I - \Pi_1(\Pi_1'C\Pi_1)^{-1}\Pi_1'C][\Pi_2\beta_2 + I_2\gamma_2]$$
$$= (\gamma_2'I_2' + \beta_2'\Pi_2')[C - C\Pi_1(\Pi_1'C\Pi_1)^{-1}\Pi_1'C](\Pi_2\beta_2 + I_2\gamma_2).$$

Now M is a square, K-dimensional positive definite matrix. Therefore it has a square-root matrix R such that $M = RR'$ and $M^{-1} = R'^{-1}R^{-1}$. Consider

$$C = M - M_{.1}M_{11}^{-1}M_{1.}$$
$$= R[I - R^{-1}M_{.1}M_{11}^{-1}M_{1.}R'^{-1}]R'.$$

Because

$$M_{1.}M^{-1}M_{.1} = M_{1.}M^{-1}M\begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix} = M_{1.}\begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix} = M_{11},$$

the quantity in brackets is idempotent. Since it is also symmetric, it is a projection. Also

$$\operatorname{tr}[I - R^{-1}M_{.1}M_{11}^{-1}M_{1.}R'^{-1}] = K - K_1$$

and

$$[I - R^{-1}M_{.1}M_{11}^{-1}M_{1.}R'^{-1}]R^{-1}M_{.1} = 0.$$

Finally $\rho(R^{-1}M_{.1}) = \rho(M_{.1}) = K_1$ (using Lemma 1), so

$$I - R^{-1}M_{.1}M_{11}^{-1}M_{1.}R'^{-1} = \bar{P}_{R^{-1}M_{.1}},$$

and

$$C = R\bar{P}_{R^{-1}M_{.1}}R'.$$

Consider now

$$C - C\Pi_1(\Pi_1'C\Pi_1)^{-1}\Pi_1'C = R\{\bar{P}_{R^{-1}M_{.1}} - \bar{P}_{R^{-1}M_{.1}}R'\Pi_1(\Pi_1'R\bar{P}_{R^{-1}M_{.1}}R'\Pi_1)^{-1}\Pi_1'R\bar{P}_{R^{-1}M_{.1}}\}R'.$$

The quantity in curly brackets is again easily seen to be symmetric and idempotent, and hence a projection. Its trace is $K - K_1 - G_1$, and it annihilates $(R'\Pi_1, R^{-1}M_{.1})$. Since $\rho(R'\Pi_1, R^{-1}M_{.1}) = \rho(M\Pi_1, M_{.1}) = G_1 + K_1$, again using Lemma 1, we have

$$C - C\Pi_1(\Pi_1'C\Pi_1)^{-1}\Pi_1'C = R\bar{P}_{R'\Pi_1, R^{-1}M_{.1}}R'.$$

Finally

$$\operatorname{plim}\frac{N'\bar{P}_{Z_1}N}{T} = (\gamma_2 I_2' + \beta_2'\Pi_2)R\bar{P}_{R'\Pi_1, R^{-1}M_{.1}}R'(\Pi_2\beta_2 + I_2\gamma_2).$$

Clearly this is a quadratic form which is non-negative, and which is positive if and only if $R'\Pi_1$, $R^{-1}M_{.1}$, and $R'(\Pi_2\beta_2 + I_2\gamma_2)$ are linearly independent vectors. But this is true if and only if M_{Π_1}, $M_{.1}$, and $M(\Pi_2\beta_2 + I_2\gamma_2)$ are linearly independent, which they are by Lemma 1. Hence

$$\operatorname{plim}\frac{N'\bar{P}_{Z_1}N}{T} = p > 0.$$

In conclusion, under the alternative hypothesis,

$$\operatorname{plim} l_1(k) = \frac{p + \bar{D}'\Sigma\bar{D}'}{\bar{D}'\Sigma\bar{D}'} > 1$$

for all k satisfying assumption (i)(a) of the theorem.

To show that $\operatorname{plim} l_1(\lambda) = \operatorname{plim}\lambda > 1$ under the alternative hypothesis, recall (see for example, [**8**, p. 343, equation 6.88] and Hood and Koopmans [**9**, p. 181])

$$|\Pi_{\Delta,**}^{*}{}^{*}M_{**,**}\Pi_{\Delta,**}' - \operatorname{plim}(\lambda - 1)\Omega_{\Delta\Delta}| = 0,$$

where ${}^{*}M_{**,**}$ and $\Omega_{\Delta\Delta}$ are certain positive definite matrices, and $\Pi_{\Delta,**}'$ is a submatrix of Π_1 of order $G_1 + 1$.

Clearly $\operatorname{plim}(\lambda - 1) = 0$ if $\rho(\Pi_{\Delta,**}) < G_1$ and $\operatorname{plim}(\lambda - 1) > 0$ if $\rho(\Pi_{\Delta,**}) = G_1$. Since Π_1 is of full rank, (Lemma 1 of Theorem 3), so are each of its submatrices. Hence $\rho(\Pi_{\Delta,**}) = G_1$, so $\operatorname{plim}(\lambda - 1) > 0$ under the alternative hypothesis.

Thus $\operatorname{plim} l_1(k) > 1$ for all the k-class estimators included under assumption (i).

The analysis of $l_2(k)$ is simplified by the observation that $\hat{\beta}_*^2$ is a consistent estimate of β_*^2 for equation (3) (see, for example, [**5**, p. 201]). Thus no equivalent to Lemma 2 is needed. Now

$$\begin{aligned}\bar{P}_{Z_1,Z_2}Y^2_*\hat{\beta}^2_{k*} &= \bar{P}_{Z_1,Z_2}(-y + Y_1\hat{\beta}_{1k} + Y_2\hat{\beta}_{2k}) \\ &= \bar{P}_{Z_1,Z_2}(Y_1(\hat{\beta}_{1k} - \beta_1) + Y_2(\hat{\beta}_{2k} - \beta_2) - Z_1\gamma_1 - Z_2\gamma_2 - u) \\ &= \bar{P}_{Z_1,Z_2}(Z\Pi_1(\hat{\beta}_{ik} - \beta_1) + Z\Pi_2(\hat{\beta}_{2k} - \beta_2) - u).\end{aligned}$$

Then

$$\begin{aligned}\operatorname{plim}\frac{\hat{\beta}^2_{k*'}Y^2_{*'}\bar{P}_{Z_1,Z_2}Y^2_*\hat{\beta}^2_{k*}}{T} &= \operatorname{plim}\frac{u'\bar{P}_{Z_1,Z_2}u}{T} \\ &= \sigma_{11}\operatorname{plim}\frac{T - K_1 - K_2}{T} = \sigma_{11}.\end{aligned}$$

Similarly

$$\operatorname{plim}\frac{\hat{\beta}^2_{k*'}Y^2_{*'}\bar{P}_ZY^2_*\hat{\beta}^2_{k*}}{T} = \sigma_{11};$$

thus $\operatorname{plim} l_2(k) = 1$, and hence

$$\operatorname{plim}\frac{l_1(k)}{l_2(k)} = \frac{\operatorname{plim} l_1(k)}{\operatorname{plim} l_2(k)} > 1.$$

As $T \to \infty$, when the null hypothesis is true,

$$\frac{T - K + L_2}{L_1 - L_2}([l_1(k)/l_2(k)] - 1)$$

approaches a χ^2 distribution with $L_1 - L_2$ degrees of freedom. Let χ_a be the ath percentile of a χ^2 distribution with $L_1 - L_2$ degree of freedom. Then the power of the test is the probability, under the alternative, that

$$\frac{T - K + L_2}{L_1 - L_2}[(l_1(k)/l_2(k) - 1] > \chi_a.$$

Under any alternative hypothesis of the type specified in the theorem,

$$\operatorname{plim}(l_1(k)/l_2(k) = \frac{\operatorname{plim} l_1}{\operatorname{plim} l_2} > 1.$$

Then

$$\operatorname{plim}\frac{T - K + L_2}{L_1 - L_2}[(l_1(k)/l_2(k) - 1] \to \infty,$$

so the power approaches 1. Therefore the test is consistent against the specified class of alternatives. *Q.E.D.*

Carnegie-Mellon University

Manuscript received July, 1970; revision received December, 1972.

REFERENCES

[1] ANDERSON, T. W., AND H. 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] BASMANN, R. L.: "An Experimental Investigation of Some Approximate Finite Sample Tests of Linear Restrictions on Matrices of Regression Coefficients," unpublished, 1959.
[3] ———: "On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics," *Journal of the American Statistical Association*, 55 (1960), 650–659.
[4] DHRYMES, P. J.: "Alternative Asymptotic Tests of Significance and Related Aspects of 2SLS and 3SLS Estimated Parameters," *Review of Economic Studies*, 36 (1969), 213–226.
[5] ———: *Econometrics: Statistical Foundations and Applications*. New York: Harper and Row, 1970.
[6] FISHER, F. M., AND J. B. KADANE: "The Covariance Matrix of the Limited Information Estimator and the Identification Test: Comment," *Econometrica*, 40 (1972), 901–903.
[7] GEISEL, M. S.: "Comparing and Choosing Among Parametric Statistical Models: A Bayesian Analysis with Macroeconomic Applications," unpublished Ph.D. dissertation, University of Chicago, 1970.
[8] GOLDBERGER, A. S.: *Econometric Theory*, New York: Wiley, 1964.
[9] HOOD, W. C., AND T. C. KOOPMANS: *Studies in Econometric Method*. Cowles Commission for Research in Economics, Monograph No. 14. New York: Wiley, 1953.
[10] KADANE, J. B.: "Testing Overidentifying Restrictions When the Disturbances are Small," *Journal of the American Statistical Association* 65 (1970), 182–185.
[11] ———: "Comparison of k-Class Estimators When the Disturbances are Small," *Econometrica*, 39 (1971), 723–737.
[12] KADANE, J. B., G. H. LEWIS, AND J. G. RAMAGE: "Horvath's Theory of Participation in Group Discussion," *Sociometry* 32 (1969), 348–361.
[13] LEHMANN, E. L.: *Testing Statistical Hypotheses*. New York: Wiley, 1959.
[14] LIU, T. C., AND W. J. BREEN: "The Covariance Matrix of the Limited Information Estimator and the Identification Test," *Econometrica*, 37 (1969), 222–227.
[15] ———: "The Covariance Matrix of the Limited Information Estimator and the Identification Test: A Reply," *Econometrica*, 40 (1972), 905–906.
[16] MADDALA, G. S.: "Some Small Sample Evidence on Tests of Significance in Simultaneous Equations Models," *Econometrica*, 42 (1974), 841–851.
[17] MORGAN, A., AND W. VANDAELE: "On Testing Hypotheses in Simultaneous Equation Models," mimeo, University of Chicago, 1972.
[18] NAGAR, A. L.: "The Bias and Moment Matrix of the General k-Class Estimators of the Parameters in Simultaneous Equations," *Econometrica*, 27 (1959), 575–595.
[19] RAMAGE, J. G.: "A Perturbation Study of the k-Class Estimators in the Presence of Specification Error," unpublished Ph.D. dissertation, Yale University, 1971.
[20] RAO, C. R.: *Linear Statistical Inference and Its Applications*. New York: Wiley, 1965.
[21] SAVAGE, L. J.: "Elicitation of Personal Probabilities and Expectations," *Journal of the American Statistical Association*, 66 (1971), 783–801.
[22] TUKEY, J.: "The Future of Data Analysis," *Annals of Mathematical Statistics*, 33 (1962), 1–67.
[23] ZELLNER, A.: *An Introduction to Bayesian Inference in Econometrics*. New York, Wiley, 1971.