The Annals of Statistics
1981, Vol. 9, No. 2, 381–390

TESTS FOR THE INDEPENDENCE BETWEEN TWO SEEMINGLY UNRELATED REGRESSION EQUATIONS.[1]

By Takeaki Kariya

Hitotsubashi University

When the error terms in two different regression equations are correlated, Zellner proposed an alternative estimator for the coefficients of each equation based on an estimated covariance matrix between the two error terms. However, since an estimated covariance matrix is used, the OLSE seems better than Zellner's estimator when the correlation of the two equations is close enough to zero. This paper considers the problem of testing the independence between two regression equations and derives a locally best invariant test for a one-sided alternative hypothesis and a locally best unbiased and invariant test for a two-sided alternative.

1. Introduction and summary. Let two different regression equations be

$$y_i = X_i\beta_i + u_i \qquad i = 1, 2, \tag{1.1}$$

where $X_i : n \times p_i$ is a fixed matrix with $\text{rank}(X_i) = p_i$ and $u_i = (u_{i1}, u_{i2}, \cdots, u_{in})'$ is an error term with mean $E(u_i) = 0$ and covariance matrix $E(u_i u_i') = \sigma_{ii} I_n$ and cross-covariance $E(u_1 u_2') = \sigma_{12} I_n (i = 1, 2)$. The relation (1.1) can be rewritten as a form of a multivariate regression model with prior information on the structure of the coefficient matrix:

$$Y = \tilde{X}B + U, \tag{1.2}$$

where $Y = [y_1, y_2]$, $\tilde{X} = [X_1, X_2]$, $U = [u_1, u_2]$ and

$$B = \begin{pmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix} : (p_1 + p_2) \times 2.$$

By letting $y = (y_1', y_2')'$, $\beta = (\beta_1', \beta_2')'$,

$$X = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} : 2n \times (p_1 + p_2)$$

and $u = (u_1', u_2')'$, it can be also rewritten as

$$y = X\beta + u. \tag{1.3}$$

We assume the joint normality for the error term u; $u \sim N(0, \Sigma \otimes I_n)$, where $\Sigma = [\sigma_{ij}] : 2 \times 2$ is a covariance matrix of u_{1t} and $u_{2t} (t = 1, \cdots, n)$ and $\Sigma \otimes I_n$ denotes the Kronecker product of Σ and I_n. From the normality of u, $\sigma_{12} = 0$ means the independence of the two equations.

Zellner (1962) has shown that, when $\sigma_{12} \neq 0$ and when Σ is known, the GLSE (generalized least squares estimator)

$$\hat{\beta}(\Sigma) = (X'[\Sigma \otimes I]^{-1}X)^{-1}X'[\Sigma \otimes I]^{-1}y \tag{1.4}$$

is more efficient than the OLSE (ordinary LSE) obtained from (1.1), provided $X_1 \neq X_2$.

Received January, 1978; revised December, 1978.

[1] This research was supported by the Ministry of Education in Japan "General D" (Kagaku Kenkyu-hi "Ippan D").

AMS 1970 *subject classifications.* Primary 62H15; secondary 62F05, 62J05.

Key ords and phrases. Seemingly unrelated regression, locally best test, invariance, correlation.

However, since Σ is in general unknown, using the OLS residuals in (1.1), Zellner replaced Σ in (1.4) by $\hat{\Sigma} = (\hat{\sigma}_{ij})$ with $\hat{\sigma}_{ij} = e_i'e_j/n (i, j = 1, 2)$, where

$$(1.5) \qquad e_i = N_i y_i \quad \text{and} \quad N_i = I_n - X_i(X_i'X_i)^{-1}X_i' \qquad i = 1, 2.$$

On the other hand, when the LS method is applied to (1.2) without using the prior information on the structure of the coefficient matrix in (1.2), another estimate for Σ is obtained from the multivariate LS residual matrix;

$$(1.6) \qquad S = Y'[I - \tilde{X}(\tilde{X}'\tilde{X})^{+}\tilde{X}']Y/n$$

where $(\tilde{X}'\tilde{X})^{+}$ is the Penrose inverse of $\tilde{X}'\tilde{X}$. The estimators $\hat{\beta}(\hat{\Sigma})$ and $\hat{\beta}(S)$ are known as the ZLSE (Zellner's LSE). When $X_1'X_2 = 0$, Zellner (1963) examined the gain in efficiency of the ZLSE over the OLSE and showed that the ZLSE is more efficient than the OLSE except when the correlation between the two equations

$$(1.7) \qquad \rho = \sigma_{12}/(\sigma_{11}\sigma_{22})^{-1/2}$$

is low and/or the sample size is small. For the case $X_1'X_2 \neq 0$, Kmenta and Gilbert (1968) obtained a similar result through a numerical method, and Revankar verified it analytically for a special case (see Section 3 below). Recently Mehta and Swamy (1976) verified it in the general case (see also Kunitomo (1977)). These results imply that when ρ is close enough to zero, the OLSE is more efficient than the ZLSE. Especially when it is zero, the OLSE will be preferred.

Motivated by these facts, we consider the problem of testing whether the correlation between the two equations is zero:

$$(1.8) \qquad H: \rho = 0 \quad \text{or equivalently} \quad \sigma_{12} = 0.$$

For testing H versus the one-sided alternative K_1: $\rho > 0$, an LBI (locally best invariant) test is derived, and for testing H versus the two-sided alternative K_2: $\rho \neq 0$, an LBIU (locally best invariant and unbiased) test is derived. Since there exists no uniformly most powerful invariant test (see Section 2), the local sensitivity of these tests near $\rho = 0$ is relevant to the problem of choice between the OLSE and the ZLSE. The null distribution of the LBI test statistic is also considered, but for that of the LBIU test statistic, a special case is only treated because of its difficulty.

Finally it is noted that a necessary and sufficient condition for which $\hat{\beta}(A) \equiv \hat{\beta}(I)$ for any given positive matrix A: 2×2 is $N_1 = N_2$, as Proposition 1 below shows. Hence throughout the paper, $N_1 \neq N_2$ is assumed since in case of $N_1 = N_2$ the above problems become uninteresting.

PROPOSITION 1. *Let* $b_i = (X_i'X_i)^{-1}X_i'y_i (i = 1, 2)$ *and* $\hat{\beta}(\Sigma)' = (\hat{\beta}_1(\Sigma)', \hat{\beta}_2(\Sigma)')$ *where* $\hat{\beta}_i(\Sigma)$: $p_i \times 1$ $(i = 1, 2)$. *Suppose* $\sigma_{12} \neq 0$. *Then a necessary and sufficient condition for* $\hat{\beta}_i(\Sigma) \equiv b_1$ *is* $X_2(X_2'X_2)^{-1}X_2'X_1 = X_1$, *and a necessary and sufficient condition for* $\hat{\beta}(\Sigma) \equiv \hat{\beta}(I)$ *is* $X_1(X_1'X_1)^{-1}X_1' = X_2(X_2'X_2)^{-1}X_2'$ *or* $N_1 = N_2$.

PROOF. Let σ^{ij} denote the (i, j) element of Σ^{-1} and let $\lambda = (\sigma^{12})^2/\sigma^{22}$ and $P_i = X_i(X_i'X_i)^{-1}X_i' (i = 1, 2)$. Then from (1.4)

$$(1.9) \qquad \hat{\beta}_1(\Sigma) = [\sigma^{11}X_1'X_1 - \lambda X_1'P_2X_1]^{-1}[\sigma^{11}X_1'y_1 - \lambda X_1'P_2y_1 + \sigma^{12}X_1'(I - P_2)y_2].$$

Since b_1 does not depend on y_2, it is easy to see that $\sigma^{12}X_1'(I - P_2)y_2 = 0$ is necessary for $\hat{\beta}_1(\Sigma) \equiv b_1$. Hence $P_2X_1 = X_1$ is obtained. Sufficiency is clear. For the second part, $\hat{\beta}_i(\Sigma) \equiv b_i (i = 1, 2)$ imply $P_2X_1 = X_1$ and $P_1X_2 = X_2$, which are equivalent to $P_1 = P_2$. This proves the proposition.

Revankar (1973) pointed out that when X_1 is a subset of X_2, i.e., $X_2 = [X_1, X_3]$, $\hat{\beta}_1(\Sigma) \equiv b_1$. This is a special case of Proposition 1.

2. LBI and LBIU tests. The problems stated in Section 1 are analyzed through invariance. Let $A_i = \{x \in R \mid x > 0\}$ and $G_i = R^{p_i} (i = 1, 2)$. We consider the group $G = A_1 \times A_2 \times G_1 \times G_2$, which leaves the problem invariant with the action;

$$y_i \to a_i y_i + X_i g_i, \tag{2.1}$$

$$\beta_i \to a_i \beta_i + g_i, \quad \text{and} \quad \sigma_{ij} \to a_i a_j \sigma_{ij} \qquad i, j = 1, 2, \tag{2.2}$$

where $a_i \in A_i$ and $g_i \in G_i$. Let $q_i = n - p_i (i = 1, 2)$. Here we state the main results but the proofs are deferred to the end of this paper.

THEOREM 1. *For testing H: $\rho = 0$ versus K_1: $\rho > 0$, an* LBI *test is given by the critical region*

$$W_1 \equiv [e_1' e_2 / (e_1' e_1 e_2' e_2)^{1/2}] > k_1, \tag{2.3}$$

where e_i's are given by (1.5). *For the alternative K_1': $\rho < 0$, the inequality is reversed.*

THEOREM 2. *For testing H: $\rho = 0$ versus K_2: $\rho \neq 0$, an* LBIU *test is given by the critical region*

$$W_2 \equiv q_1 q_2 W_1^2 - q_1[e_1' N_2 e_1 / e_1' e_1] - q_2[e_2' N_1 e_2 / e_2' e_2] > k_2, \tag{2.4}$$

where W_1 is defined by (2.3).

By an intuitive or constructive approach, the test statistic W_1 in Theorem 1 can be obtained since W_1 is nothing but the correlation between the two LS residuals $e_i (i = 1, 2)$. In this sense, Theorem 1 shows that the test based on this correlation is LBI. On the other hand, it seems difficult not only to obtain the test statistic W_2 in (2.4) through an intuitive approach but also to interpret it. One may propose for the alternative K_2 the critical region

$$W_1^2 > k_3. \tag{2.5}$$

Since the second term and the third term in (2.4) do not simultaneously become constant unless $N_1 = N_2$, which is excluded by assumption, the test defined by (2.5) is different from the test in (2.4). Further we can think of a third test based on S in (1.6);

$$W_3 \equiv [s_{12}^2 / s_{11} s_{22}] > k_4 \tag{2.6}$$

where s_{ij} is the (i, j) element of S. As is easily shown, this is equivalent to the LRT (likelihood ratio test) for testing H versus K_2 *in the case that* the prior information on the coefficient matrix B in (1.2) is *not* available. In this case the LRT is uniformly most powerful invariant (e.g., see Giri (1977) pages 194–195). However, in our problem where prior information is available, as asserted in Theorem 2, the test (2.4) locally dominates the LRT in power. On the other hand, the null distribution of $[(n - p_1 - p_2 - 1) W_3/(1 - W_3)]$ is $F(1, n - p_1 - p_2 - 1)$, F-distribution with degrees of freedom 1 and $n - p_1 - p_2 - 1$, while the distribution of W_2 is difficult to derive (see Section 3). Hence it will be interesting to compare these three tests by a numerical method and to see whether the test (2.4) is well approximated by the test (2.5) or the test (2.6). This work is left open in this paper.

Finally we remark that there exist no uniformly most powerful invariant tests in our problems. For a most powerful invariant test for testing $\rho = 0$ versus $\rho = \rho_0$ (fixed) is given by the critical region $f_T(t(z) \mid \rho_0)/f_T(t(z) \mid 0) > k$ where f_T is the density of a maximal invariant given in (4.14), but this most powerful test cannot be free from the fixed ρ_0.

3. The null distributions of the test statistics. Under H, the distribution of W_1 is first considered. Let L_i be an $n \times q_i$ matrix such that

$$L_i L_i' = I_n - X_i (X_i' X_i)^{-1} X_i' \quad \text{and} \quad L_i' L_i = I_{q_i} \qquad i = 1, 2, \tag{3.1}$$

and let $z_i = L_i' y_i (i = 1, 2)$. In terms of L_i and z_i, $W_1 = z_1' L_1' L_2 z_2 / (z_1' z_1 z_2' z_2)^{1/2}$. Let $\bar{z}_2 = L_1' L_2 z_2$ and let

$$V_1 = z_1' \bar{z}_2 / (z_1' z_1 \bar{z}_2' \bar{z}_2)^{1/2} \quad \text{and} \quad V_2 = [\bar{z}_2' \bar{z}_2 / z_2' z_2]^{1/2}. \tag{3.2}$$

Then $W_1 = V_1 V_2$. Since under H, z_i's are independently distributed as $N(0, \sigma_{ii} I_{qi}) (i = 1, 2)$, conditional on z_2, $z_1' \bar{z}_2 \sim N(0, \sigma_{11} \bar{z}_2' \bar{z}_2)$ and so $z_1' \bar{z}_2 (\bar{z}_2' \bar{z}_2)^{-1/2} \sim N(0, \sigma_{11})$ unconditionally. Hence V_1 is independent of V_2 and $(q_1 - 1)^{1/2} V_1 / (1 - V_1^2)^{1/2} \sim t(q_1 - 1)$ where $t(a)$ denotes the t-distribution with df a. This implies $V_1^2 \sim B(½, (q_1 - 1)/2)$, where $B(a, b)$ denotes the beta distribution with df a and b. Since the distribution of V_1 is symmetric about zero and since $V_2 > 0$, the distribution of W_1 is symmetric about zero under H. Hence $P(W_1 > k_1 | H) = ½ P(W_1^2 > k_1^2 | H)$ and so it suffices to consider the distribution of W_1^2 under H. While $V_1^2 \sim B(½, (q_1 - 1)/2)$ above, $V_2^2 = z_2' M_2 z_2 / z_2' z_2$ with $M_2 = L_2' L_1 L_1' L_2$ is of the same form as test statistics for serial correlation and its distribution is not easy to treat. Press (1969) reviews this problem. Here, as is often done (see, e.g., Press (1969) and Bloch and Watson (1967)), we approximate the distribution of V_2^2 by a beta distribution. From Press (1969) page 195,

$$V_2^2 / d_1 \sim B(a, b) \text{ approximately,} \tag{3.3}$$

where d_1 is the largest latent root of M_2 and

$$a = [E(V_2^2)/d_1][(E(V_2^2)/\text{Var}(V_2^2))(d_1 - E(V_2^2)) - 1], \tag{3.4}$$

and

$$b = [1 - E(V_2^2)/d_1][(E(V_2^2)/\text{Var}(V_2^2))(d_1 - E(V_2^2)) - 1]. \tag{3.5}$$

Since $z_2' z_2$ is a sufficient and complete statistic for σ_{22} and since V_2^2 is independent of σ_{22}, V_2^2 is independent of $z_2' z_2$ (Lehemann (1959) page 162). Hence

$$E(V_2^2) = E(z_2' M_2 z_2) / E(z_2' z_2) = \text{tr} M_2 / q_2 = \text{tr} N_1 N_2 / q_2, \tag{3.6}$$

and

$$\text{Var}(V_2^2) = [E(z_2' M_2 z_2)^2 / E(z_2' z_2)^2] - [E(V_2^2)]^2 = 2[q_2 \, \text{tr} M_2^2 - (\text{tr} M_2)^2] / q_2^2 (q_2 + 2). \tag{3.7}$$

Substituting these into (3.4) and (3.5), we obtain the approximation (3.3).
Hence for $0 \le x \le d_1$,

$$\begin{aligned} P(W_1^2 \le x | H) &= P(V_1^2 V_2^2 \le x, V_2^2 \le x | H) + P(V_1^2 V_2^2 \le x, V_2^2 > x | H) \\ &= P(V_2^2 / d_1 \le x/d_1 | H) + P(V_1^2 V_2^2 / d_1 \le x/d_1, V_2^2/d_1 > x/d_1 | H) \\ &= I(a, b : x/d_1) + \int_{x/d_1}^{1} I(½, (q_1 - 1)/2 : x/td_1) b(t : a, b) \, dt, \end{aligned} \tag{3.8}$$

where $I(\alpha, \beta : z) = \int_0^z b(t : \alpha, \beta) \, dt$ and $b(t : \alpha, \beta)$ denotes the density of beta distribution with df α and β. The second term in the last equation above is

$$\begin{aligned} &\int_{x/d_1}^{1} \left[\text{Be}\left(\frac{1}{2}, \frac{q_1 - 1}{2} \right) \right]^{-1} \Sigma_{j=0}^{\infty} \binom{(q_1 - 3)/2}{j} (-1)^j \\ &\left(j + \frac{1}{2} \right)^{-1} \left(\frac{x}{d_1} \right)^{j+1/2} \left(\frac{1}{t} \right)^{j+1/2} b(t : a, b) \, dt \\ &= \left[\text{Be}\left(\frac{1}{2}, \frac{q_1 - 1}{2} \right) \text{Be}(a, b) \right]^{-1} \Sigma_i \Sigma_j \binom{(q_1 - 3)/2}{j} \binom{a - j - 3/2}{i} \\ &(-1)^{i+j} \left[\left(j + \frac{1}{2} \right) (b + i) \right]^{-1} \left(\frac{x}{d_1} \right)^{j+1/2} \left(1 - \frac{x}{d_1} \right)^{i} \end{aligned} \tag{3.9}$$

where $\mathrm{Be}(a, b)$ denotes the beta function. In the above computation, we transformed t into $1 - t$ and expanded $(1 - t)^{a-j-3/2}$. Hence we obtain the approximation:

$$P(W_1^2 \le x \mid H) \fallingdotseq I(a, b{:}x/d_1) + (3.7). \tag{3.10}$$

For a given significance level α, the approximate significant point can be obtained from (3.10). We remark that various types of approximations to the distributions of such statistics as V_2^2 are also surveyed in Durbin and Watson (1971). Next we consider the two special cases; (1) $X_1'X_2 = 0$, and (2) $N_1N_2 = N_2$. For the case (1), it is easy to see that $V_2^2 \sim B((n - p_1 - p_2)/2, q_2/2)$, while the distribution of V_1^2 remains the same. Hence in the case of (1), (3.10) holds exactly with $a = (n - p_1 - p_2)/2$, $b = q_2/2$ and $d_1 = 1$. In the case (2), $V_2^2 = 1$ and hence $W_1^2 = V_1^2 \sim B(1/2, (q_1 - 1)/2)$ under H. In the case (1), Zellner (1963) considered the relative efficiency of his estimator over the OLSE. On the other hand, when X_1 is a subset of X_2, i.e., $X_2 = [X_1, X_3]$, in which the case (2) holds, Revankar (1974), 1976) analyzed the same problem. In Revankar (1974), an example in which $X_2 = [X_1, X_3]$ holds is indicated.

The distribution of the test statistic W_2 in (2.5) is more difficult to treat. Here we consider it only for the case $N_1N_2 = N_2$. As is shown above, in this case $V_2^2 = [e_2'N_1e_2/e_2'e_2] = 1$. Since $[e_1'N_2e_1/e_1'e_1] = [z_1'M_1z_1/z_1'z_1]$ where $M_1 = L_1'L_2L_2'L_1$, W_2 in (2.5) can be written as

$$W_2 = q_1\{z_1'[q_2M_0 - M_1]z_1/z_1'z_1\} - q_2 \tag{3.11}$$

where $M_0 = \bar{z}_2(\bar{z}_2'\bar{z}_2)^{-1}\bar{z}_2'$. Hence it suffices to consider the distribution of $W_2' \equiv z_1'[q_2M_0 - M_1]z_1/z_1'z_1$. Since M_i's are idempotent under $N_1N_2 = N_2 (i = 0, 1)$ and since $M_1M_0 = M_0$, there exists an orthogonal matrix $Q{:}q_1 \times q_1$ such that

$$Q[q_2M_0 - M_1]Q' = \mathrm{diag}\{(q_2 - 1), -1, \cdots, -1, 0, \underbrace{\ldots,}_{q1-q2} 0\},$$

where $\mathrm{diag}\{a_1, \cdot\cdot, a_n\}$ denotes the diagonal matrix with diagonal elements $a_1, \cdot\cdot, a_n$. Note that $N_1N_2 = N_2$ implies $q_1 \ge q_2$ and that $q_1 = q_2$ and $N_1N_2 = N_2$ imply $N_1 = N_2$. Hence our assumption $N_1 \ne N_2$ implies $q_1 > q_2$. Further we assume $q_2 > 1$. Now with $v = Qz_1$, $l = (q_2 - ½)$ and $c_2 = ½(q_2 - 1)$,

$$W_2' = (v_1^2/v'v) - [(v_2^2 + \cdots + v_{q2}^2)/v'v] \tag{3.12}$$

where $v = (v_1, \cdots, v_q)'$. Since $v \sim N(0, \sigma_{11}I)$, the distribution of $(T_1, T_2) \equiv (v_1^2/v'v, (v_2^2 + \cdots + v_{q2}^2)/v'v)$ is a Dirichlet distribution $D(½, c_2; c_3)$ with density $d(t_1, t_2) = Ct_1^{-1/2}t_2^{c_2-1}(1 - t_1 - t_2)^{c_3-1}$ where $c_3 = ½(q_1 - q_2)$ and $C = \Gamma(q_1/2)/[\Gamma(½)\Gamma(c_2)\Gamma(c_3)]$ (see Wilks (1962) pages 172–182). Further since given T_2, $T_1/(1 - T_2) \sim B(½, c_3)$ (Wilks (1962) page 180) and $T_2 \sim B(c_2, c_3 + ½)$,

$$P(W_2' \le x \mid H) = P\left(\frac{T_1}{1 - T_2} \le \frac{x + T_2}{l(1 - T_2)}\right) = E\left[I\left(\frac{1}{2}, c_3{:}\frac{x + T_2}{l(1 + T_2)}\right) \mid T_2\right] \equiv F(x). \tag{3.13}$$

We evaluate this. For $x \ge l$, $F(x) = 1$ since $(x + T_2)/l(1 - T_2) > 1$. For $l > x \ge 0$,

$$F(x) = \int_0^{(l-x)/(1+l)} I\left(\frac{1}{2}, c_3{:}\frac{(x + t_2)}{(1 - t_2)}\right) b\left(t_2; c_2, c_3 + \frac{1}{2}\right) dt_2.$$

For $0 > x \ge -1$,

$$F(x) = \int_{-x}^{(l-x)/(1+l)} I\left(\frac{1}{2}, c_3{:}\frac{(x + t_2)}{(1 - t_2)}\right) b\left(t_2; c_2, c_3 + \frac{1}{2}\right) dt_2.$$

And for $-1 > x$, $F(x) = 0$. From these we can tabulate $F(x)$ through a numerical method.

We remark on an approximation to the distribution of W_2 in the general case. Since W_2 is bounded, let the interval (x_1, x_2) be the support of the distribution of W_2. It is noted that the bounds x_1 and x_2 depend on the matrices L_1 and L_2 instead of the latent roots of N_1N_2. Let $\bar{W}_2 = (W_2 - x_1)/(x_2 - x_1)$ so that the support of $\bar{W}_2$ is (0, 1). The null distribution of $\bar{W}_2$ may be approximated by a Jaccobi polynomial. That is, the density of $\bar{W}_2$ may be expanded as the series $b(t{:}a, b)[1 + \Sigma_{j=1}\epsilon_j G_j(t)]$ where $G_j(t)$'s are Jaccobi polynomials (see, e.g., Durbin and Watson (1951) page 172 for the forms of G_j's and such an application). Here the $\epsilon_j = \int_0^1 G_j(t)f(t)\,dt/\int_0^1 G_{j(t)}^2 b(t{:}a, b)\,dt$, where $f(t)$ is the density of $\bar{W}_2$, and so ϵ_j's are computed from the moments of $\bar{W}_2$ or W_2. Setting $\epsilon_1 = \epsilon_2 = 0$ is equivalent to equating the first two moments with those of $B(a, b)$. A difficulty in this approximation is that it is not easy to find the bounds x_1 and x_2.

Let us take an example from Theil (1971) pages 295–302. There Zellner's method is applied to the estimation of investment functions of two corporations, General Electric and Westinghouse. From the nature of the problem, one-sided testing problem H: $\rho = 0$ versus K_1: $\rho > 0$ seems reasonable. From (1.10) in Theil (1971) page 295, $W_1^2 = 0.5314$. On the other hand, $d_1 = 1$, $\mathrm{tr}N_1N_2 = 16.8036$ and $\mathrm{tr}(N_1N_2)^2 = 16.6312$ are obtained from the data in Theil (1971) page 296. Hence from (3.8) and (3.9), $a = 82.567$ and $b = 0.965$ are obtained where $q_1 = q_2 = 17$. The following Table shows approximately that K_1 is significant at level 0.1%. Therefore the ZLSE will be more effective than the OLSE's obtained from each equation. Here it is well to note that $\mathrm{tr}N_1N_2 = 16.8036$ is close enough to $17 = q_1 = q_2$. Since $q_1 = q_2$ and $\mathrm{tr}N_1N_2 = q_1$ imply $N_1 = N_2$, Proposition 1 suggests that the ZLSE may not be so changed as we expect, compared to the OLSE's. This is not so in Theil's example (see (1.4) on page 295 and (1.13) on page 300).

4. PROOFS OF THEOREMS. The notation used in the preceding sections is effective in this section. For invariance, the readers are referred to Chapters 5 and 6 in Lehmann (1959).

LEMMA 1. *Under the group G defined in Section* 2, *a maximal invariant is* $(z_1/\|z_1\|, z_2/\|z_2\|)$ *and a maximal invariant parameter is* $\rho = \sigma_{12}/(\sigma_{11}\sigma_{22})^{1/2}$, *where* $\|z_i\| = (z_i'z_i)^{1/2}$.

PROOF. Invariance is clear. For maximality, suppose $L_i'y_i/\|L_i'y_i\| = L_i'y_i^*/\|L_i'y_i^*\|$. Then with $a_i = \|L_i'y_i\|/\|L_i'y_i^*\|$, $L_i'y_i = a_iL_i'y_i^*$, from which $y_i = a_iy_i^* + X_ig_i$ can be obtained for a certain g_i. Similarly for the maximal invariant parameter.

By Lemma 1, the distribution or density of a maximal invariant is free from the parameter β and it depends on $\Sigma = (\sigma_{ij})$ only through $\rho = \sigma_{12}/(\sigma_{11}\sigma_{22})^{1/2}$, hence under H: $\rho = 0$, it does not depend on any parameter. From this fact, we assume $\sigma_{11} = \sigma_{22} = 1$ without loss of generality. It is noted that the maximal invariant given in Lemma 1 is regarded as a maximal invariant under the group $A_1 \times A_2$ acting on the space of (z_1, z_2) by $z \to \bar{A}z$, where $z = (z_1', z_2')'$ and

$$\bar{A} = \begin{pmatrix} a_1 I_{q1} & 0 \\ 0 & a_2 I_{q2} \end{pmatrix} \quad \text{for} \quad (a_1, a_2) \in A_1 \times A_2. \tag{4.1}$$

TABLE 1

x	values of (3.8)	
0.1	0.801792	
0.2	0.938997	
0.3	0.982216	
0.4	0.995468	
0.5	0.999066	$q_1 = q_2 = 17$, $d_1 = 1$
0.6	0.999861	$a = 82$, $b = 1$
0.7	0.999988	
0.8	1.000000	
0.9	1.000000	

Conversely any maximal invariant under $A_1 \times A_2$ acting on (z_1, z_2) in this way is easily shown to be a maximal invariant under G with $z_i = L_i' y_i$. Hence considering the problems in terms of $y = (y_1', y_2')'$ acted upon by the group G is equivalent to considering them in terms of $z = (z_1', z_2')'$ acted upon by the group $A_1 \times A_2$. From the assumption of normality of y, the distribution of z is normal with mean 0 and covariance matrix

$$\Omega(\rho) = \begin{pmatrix} I_{q1} & \rho L_1' L_2 \\ \rho L_2' L_1 & I_{q2} \end{pmatrix}. \tag{4.2}$$

Now using Theorem 4 in Wijsman (1967), we derive the density of a maximal invariant under the transformation: $z \to \bar{A}z$. Wijsman's theorem is stated in terms of our problems as follows.

LEMMA 2. *Let $T = t(z)$ be a maximal invariant under the group $A_1 \times A_2$ acting on z by $z \to \bar{A}z$. Then the density of T with respect to the probability measure P_0^T induced by T under H: $\rho = 0$ is given by*

$$\frac{dP_\rho^T}{dP_0^T} = f_T(t(z)\,|\,\rho) = \frac{\int_{A_1 \times A_2} f(\bar{A}z\,|\,\Omega(\rho))\,|\bar{A}|\,d\nu(a_1, a_2)}{\int_{A_1 \times A_2} f(\bar{A}z\,|\,\Omega(0))\,|\bar{A}|\,d\nu(a_1, a_2)}, \tag{4.3}$$

where f is the normal density of z, $\bar{A}$ and $\Omega(\rho)$ are the matrices given in (4.1) *and* (4.2) *respectively, and ν is an invariant measure on $A_1 \times A_2$. Here P_ρ^T is the probability measure induced by T under ρ.*

Theorem 4 in Wijsman (1967) states the conditions for which (4.3) holds. Because checking the conditions is included in the proof of Theorem 2 in Wijsman (1967), it is omitted here. A direct proof can also be done by arguing as in Hajek and Sidāk (1967) pages 45–49 or Lehmann (1959) pages 248–249. An invariant measure, which is unique up to constant multiplication, is taken as $d\nu(a_1, a_2) = (a_1 a_2)^{-1}\, da_1\, da_2$, where da_i is the Lebesgue measure on A_i. After cancellation of constants, we let $K(\rho)$ be the numerator of the right-hand side of (4.3). Then from the assumption of normality,

$$K(\rho) = |\Omega(\rho)|^{-1/2} \int_0^\infty \int_0^\infty \exp\left[-\frac{1}{2} z'\bar{A}'\Omega(\rho)^{-1}\bar{A}z\right] |\bar{A}|(a_1 a_2)^{-1}\, da_1\, da_2. \tag{4.4}$$

Evaluating $\bar{A}'\Omega(\rho)^{-1}\bar{A}$ in (4.4) yields

$$K(\rho) = |\Omega(\rho)|^{-1/2} \int_0^\infty \int_0^\infty a_1^{q_1-1} a_2^{q_2-1} \exp\left(-\frac{1}{2} a' H(\rho) a\right) da_1\, da_2 \tag{4.5}$$

$(a = (a_1, a_2)'$,

where $a = (a_1, a_2)'$ and $H(\rho) = (h_{ij})$ is a 2×2 matrix with elements

$$h_{ii} \equiv h_{ii}(\rho) = z_i'(I_{q_i} - \rho^2 M_i)^{-1} z_i \qquad (i = 1, 2), \quad \text{and} \tag{4.6}$$

$$h_{21} = h_{12} \equiv h_{12}(\rho) = -\rho z_1'(I - \rho^2 M_1)^{-1} L_1' L_2 z_2.$$

Here $M_1 = L_1' L_2 L_2' L_1$ and $M_2 = L_2' L_1 L_1' L_2$ as is defined in Section 3. Let

$$r \equiv r(\rho) = -h_{12}/(h_{11}h_{22})^{1/2} \quad \text{and} \quad R = \begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}. \tag{4.7}$$

By changing a_i's into $x_1 = [h_{22}/h_{11}h_{22} - h_{22}^2)]^{-1/2} a_1$ and $x_2 = [h_{11}/(h_{11}h_{22} - h_{12}^2)]^{-1/2} a_2$ and after a little algebra, we obtain $K(\rho) = \kappa_1(\rho)\kappa_2(\rho)$, where

$$\kappa_1(\rho) = 2\pi |\Omega(\rho)|^{-1/2} (1 - r^2)^{-1/2(q_1+q_2-1)} h_{11}^{-q_1/2} h_{22}^{-q_2/2}, \tag{4.8}$$

and

$$\kappa_2(\rho) = \int_0^\infty \int_0^\infty [2\pi(1-r^2)^{1/2}]^{-1} x_1^{q_1-1} x_2^{q_2-1} \exp\left(-\frac{1}{2}x'R^{-1}x\right) dx_1\, dx_2 \tag{4.9}$$

$$x = (x_1, x_2)'.$$

By Kamat (1958) page 318, (4.9) becomes with $c_0 = 2^{-(q_1+q_2-6)}\pi^{-1}$

$$\begin{aligned}\kappa_2(\rho) = c_0\Big[\Gamma\left(\frac{q_1}{2}\right)\Gamma\left(\frac{q_2}{2}\right)F\left(-\frac{1}{2}(q_1-1), -\frac{1}{2}(q_2-1); \frac{1}{2}; r^2\right) \\ + 2r\Gamma\left(\frac{q_1+1}{2}\right)\Gamma\left(\frac{q_2+1}{2}\right)F\left(-\frac{1}{2}(q_1-2), -\frac{1}{2}(q_2-2); \frac{3}{2}; r^2\right)\Big],\end{aligned} \tag{4.10}$$

where

$$F(a, b; c; x) = \textstyle\sum_{j=0}^{\infty} \tau_j [x^j/j!] \quad \text{and} \quad \tau_j = \Gamma(a+j)\Gamma(b+j)\Gamma(c)/\Gamma(c+j)\Gamma(a)\Gamma(b). \tag{4.11}$$

Hence the denominator of the right-hand side of (4.4) is

$$K(0) = 2\pi d_{11}^{-q_1/2} d_{22}^{-q_2/2} c_0 \Gamma\left(\frac{q_1}{2}\right)\Gamma\left(\frac{q_2}{2}\right), \tag{4.12}$$

where

$$d_{ii} = h_{ii}(0) = z_i'z_i = y_i'N_iy_i = e_i'e_i \qquad i = 1, 2. \tag{4.13}$$

Since $f_T(t(z)\,|\,\rho) = K(\rho)/K(0)$ from (4.3), we obtain

THEOREM 3. *The density of a maximal invariant $T = t(z)$ with respect to the probability measure P_0^T is given by*

$$f_T(t(z)\,|\,\rho) = |\Omega(\rho)|^{-1/2}(1-r^2)^{-1/2(q_1+q_2-1)}[h_{11}/d_{11}]^{-1/2q_1}[h_{22}/d_{22}]^{-1/2q_2}\gamma(\rho), \tag{4.14}$$

where $\gamma(\rho) = \kappa_2(\rho)/c_0\Gamma(q_1/2)\Gamma(q_2/2)$. Naturally $f_T(t(z)\,|\,0) = 1$.

Now we prove Theorems 1 and 2. First we note that the power function of an invariant test ϕ is expressed as $\pi_\phi(\rho) = \int_0^\infty \phi(t(z)) f_T(t(z)\,|\,\rho)\, dP_0^T$. Since we show at the end of this section that the derivatives of the first and the second orders $\pi'(\rho)$ and $\pi''(\rho)$ in the neighborhood of $\rho = 0$ can be computed beneath the integral sign, as in Ferguson (1967) pages 235–238, for testing H: $\rho = 0$ versus K_1: $\rho > 0$, an LBI test is given by the critical region

$$\frac{\partial}{\partial\rho} f_T(t(z)\,|\,\rho)\,|_{\rho=0} > k f_T(t(z)\,|\,0), \tag{4.15}$$

and for testing H:$\rho = 0$ versus K_2:$\rho \neq 0$, an LBIU test is given by the critical region

$$\frac{\partial^2}{\partial\rho^2} f_T(t(z)\,|\,\rho)\,|_{\rho=0} > k^0 f_T(t(z)\,|\,0) + k^1 \frac{\partial}{\partial\rho} f_T(t(z)\,|\,\rho)\,|_{\rho=0}. \tag{4.16}$$

From Theorem 3, we write f_T as $f_T(t(z)\,|\,\rho) = \prod_{i=1}^5 Q_i(\rho)$ where

$$\begin{aligned} &Q_1(\rho) = |\Omega(\rho)|^{-1/2}, \quad Q_2(\rho) = (1-r^2)^{-1/2(q_1+q_2-1)}, \quad Q_3(\rho) = [h_{11}/d_{11}]^{-1/2q_1}, \\ &Q_4(\rho) = [h_{22}/d_{22}]^{-1/2q_2} \quad \text{and} \quad Q_5(\rho) = \gamma(\rho).\end{aligned} \tag{4.17}$$

By using the matrix formula, $(I-A)^{-1} = I + A(I-A)^{-1}$, it can be easily shown that $h_{ii}(0) = d_{ii}$, $h_{ii}'(0) = 0$, and $h_{ii}''(0) = d_{i+2,i+2}$ ($i = 1, 2$) where

(4.18) $$d_{33} = z'_1L'_1L_2L'_2L_1z_1 = e'_1N_2e_1 \quad \text{and} \quad d_{44} = z'_2L'_2L_1L'_1L_2z_2 = e'_2N_1e_2$$

and that $h_{12}(0) = 0$, $h'_{12}(0) = d_{12}$ and $h''_{12}(0) = 0$ where $d_{12} = z'_1L'_1L_2z_2 = e'_1e_2$. Hence $r(0) = 0$, $r'(0) = d_{12}/(d_{11}d_{22})^{1/2} = W_1$, $r''(0) = 0$, $Q'_i(0) = 0$ $(i = 1, \cdots, 4)$ and $Q'_5(0) = 2c_1W_1$ where $c_1 = \prod_{i=1}^2 [\Gamma((q_i + 1)/2)/\Gamma(q_i/2)]$. Thus $f'_T(t(z)\,|\,0) = [\prod_{i=1}^4 Q_i(0)]Q'_5(0) = 2c_1W_1$. Substituting this and $f_T(t(z)\,|\,0) = 1$ into (4.15), we obtain Theorem 1. To prove Theorem 2, we first note that $f''_T(t(z)\,|\,0) = \sum_{i=1}^5 [\prod_{j\neq 1} Q_j(0)]Q''_i(0)$ since $Q'_i(0) = 0$ $(i = 1, \cdots, 4)$. From the results above, it can be easily verified that $Q''_1(0)[\prod_{j=2}^5 Q_j(0)] = \text{const.}$, $Q''_2(0) = (q_1 + q_2 - 1)W_1^2$, $Q''_3(0) = q_1[d_{33}/d_{11}]$, $Q''_4(0) = q_2[d_{44}/d_{22}]$, $Q''_5(0) = (q_1 - 1)(q_2 - 1)W_1^2$. Hence

(4.19) $$f'_T(t(z)\,|\,0) = \text{const.} + q_1q_2W_1^2 - [q_1(d_{33}/d_{11}) + q_2(d_{44}/d_{22})],$$

and from (4.16) and LBIU test ϕ_2 is given by the critical region

(4.20) $$q_1q_2W_1^2 - [q_1(d_{33}/d_{11}) + q_2(d_{44}/d_{22})] > k^3 + k^4W_1.$$

Here k^3 and k^4 are so chosen that the size of this test is α, i.e., $E_0\phi_2 = \alpha$, and the first derivative of the power function at zero is zero or equivalently $E_0\phi_2W_1 = 0$ (see Ferguson (1967) page 238). But exactly following the argument in Ferguson (1967) pages 239–240, we can prove that $E_0\phi_2W_1 = 0$ if and essentially only if $k^4 = 0$ in (4.20). For (W_1, W_2) in our problem corresponds to (U, V) in Ferguson page 239, and the distribution of (W_1, W_2) is symmetric about $W_1 = 0$. Hence Theorem 2 is obtained.

To complete the proofs, we shall show that the derivatives can be computed beneath the integral sign. Since the latent roots of the matrix M_i lie between 0 and 1, it is easy to show that $1 \leq [h_{ii}/d_{ii}] \leq 1/(1 - \rho^2)$ $(i = 1, 2)$. Further since $h_{12}^2 \leq \rho^2z'_1(I - \rho^2M_1)^{-1}z'_1z'_2L'_2L_1(I - \rho^2M_1)^{-1}L'_1L_2z_2$, $[h_{12}^2/d_{11}d_{22}] \leq \rho^2/(1 - \rho^2)^2$. Hence $|r| \leq |\rho|/(1 - \rho^2)$ follows, and for $|\rho| < 1/3$, $|r| < 3/8$. This implies that $Q_j(\rho)$ is bounded for $|\rho| < 1/3$ $(j = 1, \cdots, 5)$. Next, since $|h'_{ii}/d_{11}d_{22}| \leq 2|\rho|/(1 - \rho^2)$ $(i = 1, 2)$, $|Q'_j(\rho)| \leq q_1|\rho|/(1 - \rho^2)$ $(j = 3, 4)$. And since $|h'_{12}/d_{11}d_{22}| \leq (1 - \rho^2)^{-1} + 2|\rho|/(1 - \rho^2)^2 \equiv \eta(\rho)$, $|r'| \leq \eta(\rho) + 2\rho^2/(1 - \rho^2)^3$ and so $Q'_j(\rho)$ is bounded for $|\rho| < 1/3$ $(j = 1, \cdots, 5)$. Therefore, since $f'_T(t(z)\,|\,\rho) = \sum_{i=1}^5 [\prod_{j\neq i} Q_j(\rho)]Q'_i(\rho)$ and since the hypergeometric functions in (4.10) are increasing in r^2 and converge absolutely for $r^2 < 1$, the boundedness of f'_T follows and $\pi'_\phi(\rho) = \int \phi f'_T\, dP_0^T$ is obtained at least for $|\rho| < 1/3$. Similarly since $Q''_j(\rho)$ is shown to be bounded for $|\rho| < 1/3$ and since $f''_T(t(z)\,|\,\rho) = \sum_{i=1}^5 [\prod_{j\neq i} Q_j(\rho)]Q''_i(\rho) + \sum_{i=1}^5 \{\sum_{j\neq i} [\prod_{k\neq j} Q_k(\rho)]Q'_j(\rho)\}Q'_i(\rho)$, $\pi''_\phi(\rho) = \int \phi f''_T\, dP_0^T$ is obtained for $|\rho| < 1/3$. This completes the proof of Theorems 1 and 2.

We remark that the essential uniqueness of the LBI test and the LBIU test can be proved by applying the necessary part of the generalized Neyman-Pearson lemma (Lehmann (1959) page 83).

Acknowledgment. The author is very grateful to the referees, the associate editors and Professor S. Nabeya of Hitotsubashi University for valuable comments and suggestions, and Kudan Computer Service for computing the table used in the numerical example.

REFERENCES

Bloch, D. A. and Watson, G. S. (1967). A Bayesian study of the multinomial distribution. *Ann. Math. Statist.* **38** 1423–1435.

Durbin, J. and Watson, G. S. (1950). Testing for serial correlation in least squares regression. I. *Biometrika* **37** 409–428.

Durbin, J. and Watson, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika* **38** 159–178.

Durbin, J. and Watson, G. S. (1971). Testing for serial correlation in least squares regression. III. *Biometrika* **58** 1–19.

Ferguson, T. S. (1967). *Mathematical Statistics: Decision Theoretic Approach.* Academic, New York.

GIRI, C. N. (1977). *Multivariate Statistical Inference.* Academic, New York.

HÁJEK, J. and SIDAK, Z. (1967). *Theory of Rank Tests.* Academic, New York.

KAKWANI, N. C. (1967). The unbiasedness of Zellner's seemingly unrelated regression equations estimators. *J. Amer. Statist. Assoc.* **62** 141–142.

KAMAT, A. R. (1958). Hypergeometric expansions for incomplete moments of the bivariate normal distribution. *Sankhyā* **20** 317–320.

KMENTA, J. and GILBERT, R. F. (1968). Small sample properties of alternative estimators of seemingly unrelated regression. *J. Amer. Statist. Assoc.* **63** 1180–1200.

KUNITOMO, N. (1977). A note on the efficiency of Zellner's estimator for the case of two seemingly unrelated regression equations. *Econ. Studies Quarterly* **28** 73–77.

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

MEHTA, J. S. and SWAMY, P. A. V. B. (1976). Further evidence on the relative efficiencies of Zellner's seemingly unrelated regressions estimator. *J. Amer. Statist. Assoc.* **71** 634–639.

PRESS, S. J. (1969). On serial correlation. *Ann. Math. Statist.* **40** 186–196.

REVANKAR, N. S. (1974). Some finite sample results in the context of two seemingly unrelated regression equations. *J. Amer. Statist. Assoc.* **69** 187–190.

REVANKAR, N. S. (1976). Use of restricted residuals in SUR systems: some finite results. *J. Amer. Statist. Assoc.* **71** 183–188.

THEIL, H. (1971). *Principles of Econometrics.* Wiley, New York.

WIJSMAN, R. A. (1967). Cross-sections of orbits and their application to densities of maximal invariants. *Fifth Berkeley Symp. Math. Statist. Prob.* **1** 389–400.

ZELLNER, A. (1962). An efficient method of estimating seemingly unrelated regression and tests for aggregation bias. *J. Amer. Statist. Assoc.* **57** 348–369.

ZELLNER, A. (1963). Estimators for seemingly unrelated regression equations: some exact sample results. *J. Amer. Statist. Assoc.* **58** 977–992.

WILKS, S. S. (1962). *Mathematical Statistics.* Wiley, New York.

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF PITTSBURGH
PITTSBURGH, PENNSYLVANIA 15260