Econometric Theory, **15**, 1999, 151–160. Printed in the United States of America.

PROBLEMS AND SOLUTIONS

PROBLEMS

99.1.1. *Minimum Distance Estimator with L1 Norm*, proposed by Oliver Linton. The method of minimum distance or minimum chi-squared is a well-established approach to imposing restrictions in estimation. Suppose that we have some consistent estimates $\hat{\theta}$ of a $p \times 1$ parameter vector θ and that we want to impose the restriction that $\theta = g(\delta)$, where δ is a $q \times 1$ vector with $q < p$. The idea is to let $\hat{\delta}$ be the minimizer of $\|\hat{\theta} - g(\delta)\|_2$ with respect to δ, where usually $\|x\|_2^2 = x'Vx$ for some matrix V, i.e., a Euclidean metric. The Cowles monograph by Rothenberg (1973) contains an excellent discussion of this method. Suppose instead that we use some other vector norm $\| \ \|$ to impose restrictions; what are the asymptotic properties of the resulting estimates of δ? We are accustomed to seeing the choice of norm as having little qualitative effect on the asymptotic properties of estimators, as in the comparison between median and mean, so that we might expect that the choice of norm would only affect the magnitude of the asymptotic variance and not the distributional form itself. Unfortunately, this is not true. Consider the special case that $\theta_{j0} = \delta_0$, $j = 1, \ldots, p$, and let $\hat{\delta}$ minimize

$$Q_n(\delta) = \|\hat{\theta} - \delta \cdot 1\|_1 = \sum_{j=1}^{p} |\hat{\theta}_j - \delta|,$$

where for simplicity suppose that $\sqrt{n}(\hat{\theta} - \theta_0) \Rightarrow N(0, I_p)$ and $p = 3$. Prove that

$$\Pr[\sqrt{n}(\hat{\delta} - \delta_0) \leq x] \to 6 \int_{-\infty}^{x} \Phi(t)\{1 - \Phi(t)\}\phi(t)\, dt,$$

where Φ and ϕ are, respectively, the c.d.f. and density functions of the standard normal.

REFERENCE

Rothenberg (1973) *Efficient Estimation with A Priori Information*. Cowles Foundation Monograph 23. New Haven: Yale University Press.

99.1.2. *An Interesting Property of Mahalanobis Distance*, proposed by Ali S. Hadi. Let $\mathbf{X}$ be a $(p+1) \times p$ matrix representing $p+1$ observations on p variates. Let $\bar{\mathbf{x}}$ and $\mathbf{S}$ be the mean vector and covariance matrix of the p variates in $\mathbf{X}$, respectively. Assume that $\mathbf{S}$ is invertible. Let

$$d_i^2 = (\mathbf{x}_i - \bar{\mathbf{x}})^T \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}), \qquad i = 1, \ldots, n$$

© 1999 Cambridge University Press 0266-4666/99 $9.50

https://doi.org/10.1017/S0266466699151089 Published online by Cambridge University Press

be the squared Mahalanobis distances for the $p+1$ observations in $\mathbf{X}$, where $\mathbf{x}_i$ is the transpose of the ith row of $\mathbf{X}$. Show that $d_i^2 = p^2/(p+1)$ for all i. Thus, all $p+1$ points (observations) must lie on the surface of the p-dimensional ellipsoid centered at $\bar{\mathbf{x}}$ and whose orientation (directions of axes) and shape are determined by the matrix $\mathbf{S}$. This property was observed in a method for the detection of outliers in multivariate data (Hadi,1994).

REFERENCE

Hadi, A.S. (1994) A modification of a method for the detection of outliers in multivariate samples. *Journal of the Royal Statistical Society B* 56, 393–396.

99.1.3. *An Application of Beinstein's Inequality*, proposed by Hua Liang. Using the following result:

Beinstein's Inequality. Let $V_1,\ldots,V_n$ be independent random variables with zero means and bounded ranges: $|V_i| \le M$. Then for each $\eta > 0$,

$$P\left(\left|\sum_{i=1}^{n} V_i\right| > \eta\right) \le 2\exp\left[-\eta^2 \Big/ \left\{2\left(\sum_{i=1}^{n} \operatorname{var} V_i + M\eta\right)\right\}\right].$$

Prove the following theorem.

THEOREM. *Let $V_1,\ldots,V_n$ be independent random variables with means zero and finite rth moment $(r \ge 2)$, i.e., $\sup_{1\le j\le n} E|V_j|^r \le C < \infty$. Assume $(a_{ki}, k, i = 1\ldots,n)$ is a sequence of positive numbers such that $\sup_{1\le i,k\le n}|a_{ki}| \le n^{-p_1}$ for some $0 < p_1 \le 1$ and $\sum_{j=1}^{n} a_{ji} = O(n^{p_2})$ for $p_2 \ge \max(0, 2/r - p_1)$. Then*

$$\max_{1\le i\le n}\left|\sum_{k=1}^{n} a_{ki}V_k\right| = O(n^{-s}\log n) \quad \text{for } s = (p_1 - p_2)/2. \quad a.s.$$

99.1.4. *Equality of Two IV Estimators*, proposed by Hailong Qian. Consider the following linear regression model:

$$y_i = x_i\beta + \varepsilon \equiv x_{1i}\beta_1 + x_{2i}\beta_2 + \varepsilon_i, \qquad i = 1,2,\ldots,N, \tag{1}$$

where the dimensions of x_i, x_{1i}, and x_{2i} are $1\times K$, $1\times K_1$, and $1\times K_2$, respectively, with $K = K_1 + K_2$. We allow the possibility that x_i may be correlated with ε_i, but we assume that we have instruments z_i such that $E(\varepsilon_i|z_i) = 0$ and $E(\varepsilon_i^2|z_i) = \sigma^2$. We partition the instruments into two subsets: $z_i = (z_{1i}, z_{2i})$, where the dimensions of z_i, z_{1i}, and z_{2i} are L, L_1, and L_2, with $L = L_1 + L_2$. We assume that $E(z_{1i}'x_i)$ has full column rank (so $L_1 \ge K$); this ensures that β can be estimated consistently using the subset of instruments z_{1i} only or using the entire set $z_i = (z_{1i}, z_{2i})$. We also assume that (y_i, x_i, z_i) is covariance stationary.

https://doi.org/10.1017/S0266466699151089 Published online by Cambridge University Press

Define $P_A = A(A'A)^{-1}A'$ and $M_A = I - P_A$ for any matrix A with full column rank. Let $X = (x_1', \ldots, x_N')'$, and similarly for X_1, X_2, y, Z_1, Z_2, and Z. Define $\hat{X} = P_{[Z_1]}X$ and $\hat{\beta} = (\hat{X}'\hat{X})^{-1}\hat{X}'y$, so that $\hat{\beta}$ is the instrumental variables (IV) estimator of (1) using Z_1 as instruments. Similarly, define $\tilde{X} = P_Z X$ and $\tilde{\beta} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'y$, so that $\tilde{\beta}$ is the IV estimator of (1) using Z as instruments.

Show that $\hat{\beta}_1 = \tilde{\beta}_1$ if $Z_2' M_{[Z_1]}[X_1 - X_2(X_2'P_1X_2)^{-1}X_2'P_1X_1] = 0$.

99.1.5. *The Justification of Multidimensional Scaling under Euclidean Conditions*, proposed by Michel van de Velden, Albert Satorra, and Heinz Neudecker. In multidimensional scaling (MDS) under Euclidean conditions one aims to find a configuration of (unknown) points $\mathbf{x}_i$ $(i = 1 \ldots n)$ using a known $n \times n$ (Euclidean) distance matrix $\mathbf{D} = \{d_{ij}\}$. For this purpose one defines

$$\mathbf{B} \equiv \mathbf{MAM},$$

where

$$\mathbf{A} = \{a_{ij}\} \equiv -\tfrac{1}{2}\{d_{ij}^2\}$$

and $\mathbf{M}$ is the usual centering matrix $\mathbf{I} - (1/n)\mathbf{11}'$. (The distances d_{ij} satisfy

$$d_{ij} = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)'(\mathbf{x}_i - \mathbf{x}_j)},$$

if and only if $\mathbf{B} \geq 0$.)

An $n \times k$ configuration $\mathbf{Z}$ of full column rank can then be obtained from the matrix $\mathbf{B}$ by means of a Schur decomposition, i.e.,

$$\mathbf{B} = \mathbf{T\Lambda T}' = \mathbf{T\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}\mathbf{T}' = \mathbf{ZZ}', \tag{1}$$

where

$$\mathbf{Z} \equiv \mathbf{T\Lambda}^{1/2}, \quad \text{and} \quad \mathbf{\Lambda} = \mathbf{\Lambda}_d > 0.$$

The n rows of $\mathbf{Z}$ give the coordinates of the n points (e.g., Mardia et al., 1994). Moreover, as a result of the semiorthogonality of $\mathbf{T}$,

$$\mathbf{Z}'\mathbf{Z} = \mathbf{\Lambda}. \tag{2}$$

Problem. Show that, given as $n \times p$ data matrix $\mathbf{X}$ such that $\mathbf{MX}$ is of rank k, the solution $\mathbf{Z}$ of MDS applied to the Euclidean distance matrix $\mathbf{D}$, i.e., the matrix $\mathbf{Z}$ obtained from $\mathbf{B}$, satisfies the equation

$$\mathbf{Z} = \mathbf{MXL},$$

with $\mathbf{L}$ a $p \times k$ semiorthogonal matrix.

REFERENCE

Mardia, K.V., J.T. Kent, & J.M. Bibby (1994) *Multivariate Analysis*. London: Academic Press.

https://doi.org/10.1017/S0266466699151089 Published online by Cambridge University Press

SOLUTIONS

98.1.1 *Similarity and Distance Matrices*—Solution, proposed by Heinz Neudecker and Michel van de Velden. We have to prove that

$$(c_{ii} + c_{jj} - 2c_{ij})^{1/2} \leq (c_{ii} + c_{kk} - 2c_{ik})^{1/2} + (c_{jj} + c_{kk} - 2c_{jk})^{1/2}$$

or equivalently

$$(\ell' J_{ij} C J_{ij}' \ell)^{1/2} \leq (\ell' J_{ik} C J_{ik}' \ell)^{1/2} + (\ell' J_{jk} C J_{jk}' \ell)^{1/2}, \tag{1}$$

where

$$C := \begin{pmatrix} c_{ii} & c_{ij} & c_{ik} \\ c_{ij} & c_{jj} & c_{jk} \\ c_{ik} & c_{jk} & c_{kk} \end{pmatrix},$$

$$J_{ij}' := (e_1, e_2), \qquad J_{ik}' := (e_1, e_3), \qquad J_{jk}' := (e_2, e_3),$$

e_1, e_2, and e_3 being three-dimensional unit vectors, and finally $\ell' := (1 \quad -1)$.

By using the vector norm $\|x\| := (x'x)^{1/2}$ we can rewrite (1) as

$$\|C^{1/2} J_{ij}' \ell\| \leq \|C^{1/2} J_{ik}' \ell\| + \|C^{1/2} J_{jk}' \ell\|.$$

Obviously $(J_{ik}' - J_{jk}')\ell = e_1 - e_2 = J_{ij}'\ell$. Hence by the vector triangle inequality $\|x + y\| \leq \|x\| + \|y\|$, the result follows. (The positive semidefiniteness of C is seen to be of strategic importance.)

98.1.2. *Lower Eigenbound for AR(1) Disturbance Covariance Matrix*—Solutions.[1] The following solutions have been proposed independently by Anton Granik and B.M. Pötscher. Each contains an interesting derivation.

1. Solution, proposed by Anton Granik. Given the form of the disturbance covariance matrix $\Omega(T \times T)$ in (1), it can be easily verified that

$$\Omega^{-1} = \begin{bmatrix} 1 & -\rho & 0 & & \dots & & & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & & & & . \\ 0 & -\rho & 1+\rho^2 & -\rho & 0 & & & . \\ . & 0 & -\rho & . & . & . & . & . \\ & & & . & . & & . & \\ . & & . & . & . & . & -\rho & 0 \\ . & & & . & & -\rho & 1+\rho^2 & -\rho \\ 0 & & \dots & & & 0 & -\rho & 1 \end{bmatrix}, \tag{3}$$

where $|\rho| < 1$ and $2 \leq T < \infty$.

https://doi.org/10.1017/S0266466699151089 Published online by Cambridge University Press

Let $\mu_i(\Omega^{-1})$ denote the eigenvalues of Ω^{-1} in nondecreasing order: $\mu_1(\cdot) \leq \mu_2(\cdot) \leq \cdots \leq \mu_T(\cdot)$ ($\mu_i > 0$ because Ω^{-1} is positive definite). Note that $\lambda_i(\Omega) = 1/\mu_i(\cdot)$, where $\lambda_i(\Omega)$ denote the eigenvalues of Ω in nonincreasing order. Let us put $\Omega^{-1} = (1+\rho^2)I_T - 2\rho C$, where

$$C = \frac{1}{2}\begin{bmatrix} \rho & 1 & 0 & . & . & 0 \\ 1 & 0 & 1 & & & . \\ 0 & 1 & . & . & & . \\ . & & . & . & . & 0 \\ . & & . & . & 0 & 1 \\ 0 & . & . & 0 & 1 & \rho \end{bmatrix}. \tag{4}$$

It is clear that the eigenvalues of Ω^{-1} are given by

$$\mu_i = 1 + \rho^2 - 2\rho\psi_i \qquad (i = 1,2,\ldots,T), \tag{5}$$

where ψ_i are the corresponding eigenvalues of C. Moreover, the largest eigenvalue of Ω^{-1} will correspond to the smallest eigenvalue of C when $\rho \geq 0$ and to the largest eigenvalue of C when $\rho < 0$. We note, further, that C can be written as

$$C = \frac{1}{2}\begin{bmatrix} 0 & 1 & 0 & . & . & 0 \\ 1 & 0 & 1 & & & . \\ 0 & 1 & . & . & & . \\ . & & . & . & . & 0 \\ . & & & . & & 1 \\ 0 & . & . & 0 & 1 & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \rho & 0 & . & . & . & 0 \\ 0 & 0 & . & & & . \\ . & . & . & . & & . \\ . & & . & . & . & . \\ . & & & . & 0 & 0 \\ 0 & . & . & . & 0 & \rho \end{bmatrix} = A + B.$$

Now, denote the eigenvalues of A, B, and C by α_j, β_j, and ψ_j, respectively, where all three sets are arranged in nonincreasing order. Because A and B are both *symmetric*, the following will be true (see Wilkinson, 1965):

$$\alpha_j + \beta_T \leq \psi_j \leq \alpha_j + \beta_1 \qquad (j = 1,2,\ldots,T).$$

The eigenvalues of A are given by $\alpha_j = \cos(\pi j/T+1)$ ($j = 1,2,\ldots T$) (see, e.g., Tanaka, 1996, Ch. 8, solution to ex. 1.3). The distinct eigenvalues of B are 0 and $\rho/2$. Therefore, when $\rho \geq 0$ ($\beta_T = 0$), $\psi_j \geq \cos(\pi T/T+1)$, $j = 1,2,\ldots,T$, and when $\rho < 0$ ($\beta_1 = 0$), $\psi_j \leq \cos(\pi/T+1)$, $j = 1,2,\ldots,T$. Also, because $\cos(\pi/T+1) = -\cos(\pi T/T+1)$, from (5) it follows that

$$\mu_i \leq 1 + |\rho|^2 + 2|\rho|\cos\left(\frac{\pi}{T+1}\right) < (1+|\rho|)^2 \qquad (i = 1,2,\ldots,T), \tag{6}$$

https://doi.org/10.1017/S0266466699151089 Published online by Cambridge University Press

which in turn implies that

(a)

$$\lambda_i(\Omega) \geq \frac{1}{1 + |\rho|^2 + 2|\rho|\cos\left(\dfrac{\pi}{T+1}\right)} > \frac{1}{(1+|\rho|)^2} \qquad (i = 1,2,\ldots,T)$$

(b) Because $\lim_{|\rho|\to 1} \lim_{T\to\infty}[1 + |\rho|^2 + 2|\rho|\cos(\pi/T + 1)] = 4$,

$$\lim_{|\rho|\to 1} \lim_{T\to\infty} \lambda_i(\Omega) \geq \tfrac{1}{4} \qquad (i = 1,2,\ldots,T)$$

2. Solution, proposed by B.M. Pötscher. Let $\alpha = (\alpha_1,\ldots,\alpha_T)'$, with $\alpha'\alpha = 1$, be an eigenvector of the covariance matrix Ω corresponding to the smallest eigenvalue λ_T. Then $\lambda_T = \alpha'\Omega\alpha = \int_{-\pi}^{\pi} |\alpha(e^{ix})|^2 f(x)\,dx$, where $f(x) = (2\pi)^{-1}|1 - \rho e^{ix}|^{-2} = (2\pi)^{-1}(1 + \rho^2 - 2\rho\cos x)^{-1}$ is the spectral density of the AR(1) process and $\alpha(e^{ix}) = \sum_{j=1}^{T} \alpha_j e^{ijx}$. It is obvious that $f(x) \geq (2\pi)^{-1}(1 + |\rho|)^{-2}$ for all $x \in [-\pi,\pi]$ with equality holding at most in two points. Consequently, $|\alpha(e^{ix})|^2 f(x) \geq |\alpha(e^{ix})|^2 (2\pi)^{-1}(1 + |\rho|)^{-2}$ for all $x \in [-\pi,\pi]$, with equality holding only in a finite number of points, because $\alpha(e^{ix})$, as a nontrivial polynomial in e^{ix}, can vanish only in finitely many points. This implies $\int_{-\pi}^{\pi} |\alpha(e^{ix})|^2 f(x)\,dx > (2\pi)^{-1}(1 + |\rho|)^{-2} \int_{-\pi}^{\pi} |\alpha(e^{ix})|^2\,dx = (1 + |\rho|)^{-2}$, because $\int_{-\pi}^{\pi} |\alpha(e^{ix})|^2\,dx = \alpha'\alpha = 1$. This completes the proof of part (a). Because the lower bound $(1 + |\rho|)^{-2}$ for the smallest eigenvalue λ_T is clearly larger than $\frac{1}{4}$ for $|\rho| < 1$, we immediately obtain $\inf_{|\rho|<1} \inf_{T\geq 2} \inf_{1\leq i\leq T} \lambda_i \geq \frac{1}{4}$, which takes care of part (b) (whether it is taken literally, in which case the limit over T can be dropped, or whether the subscript i is considered to be a typo and is replaced by T).

The preceding argument in fact establishes the well-known result that the smallest eigenvalue of the covariance matrix of T consecutive values of a weakly stationary process possessing a spectral density is not less than the infimum of the spectral density over $[-\pi,\pi]$, with the inequality being strict if the spectral density does not coincide with its infimum almost everywhere.

NOTE

1. An excellent solution has been proposed independently by S.K. Sapra, by Seuck Hean Song, by M. van de Velden and Jan Tuinstra, and by Kyung-Taik Han and Eric I. Im, the posers of the problem.

REFERENCES

Tanaka, K. (1996) *Time Series Analysis: Nonstationary and Noninvertible Distribution Theory.* New York: Wiley.

Wilkinson, J.M. (1965) *The Algebraic Eigenvalue Problem.* Oxford: Oxford University Press.

https://doi.org/10.1017/S0266466699151089 Published online by Cambridge University Press

98.1.3. *Equivalence of LR Test and Hausman Test*—Solution, proposed by Hailong Qian. Let $y_t = (y_{1t}, y_{2t})'$, $\theta = (\alpha', \sigma')'$ with $\alpha = (\alpha_1, \alpha_2)'$ and $\sigma = (\sigma_1^2, \sigma_2^2, \sigma_{12})'$. The log-likelihood function is

$$l(\theta) = -T\ln(2\pi) - \frac{T}{2}\ln|\Omega| - \frac{1}{2}\sum_{t=1}^{T}(y_t - \alpha)'\Omega^{-1}(y_t - \alpha). \tag{1}$$

Then, it is easy to show that the MLE of α, σ_1^2, σ_2^2, and σ_{12} are, respectively,

$$\hat{\alpha} = \bar{y}, \tag{2}$$

$$\hat{\sigma}_1^2 = S_1^2/T, \qquad \hat{\sigma}_2^2 = S_2^2/T, \tag{3}$$

$$\hat{\sigma}_{12} = S_{12}/T, \tag{4}$$

where $\bar{y} = \sum_{t=1}^{T} y_t/T$, $S_1^2 = \sum_{t=1}^{T}(y_{1t} - \hat{\alpha}_1)^2$, $S_2^2 = \sum_{t=1}^{T}(y_{2t} - \hat{\alpha}_2)^2$, and $S_{12} = \sum_{t=1}^{T}(y_{1t} - \hat{\alpha}_1)(y_{2t} - \hat{\alpha}_2)$. Thus, substituting (2)–(4) into (1), we have

$$l(\hat{\theta}) = -T[\ln(2\pi) + 1] - \frac{T}{2}\ln|\hat{\Omega}| = -T[\ln(2\pi) + 1] - \frac{T}{2}\ln(\hat{\sigma}_1^2\hat{\sigma}_2^2 - \hat{\sigma}_{12}^2). \tag{5}$$

Through straightforward but tedious calculation, we obtain the inverse of the information matrix:

$$I^{-1} = \left\{E\left[-\frac{\partial^2 l(\theta)}{\partial\theta\partial\theta'}\right]\right\}^{-1} = \begin{bmatrix} I_{\alpha\alpha}^{-1} & 0_{2\times 3} \\ 0_{3\times 2} & I_{\sigma\sigma}^{-1} \end{bmatrix}, \tag{6}$$

where

$$I_{\alpha\alpha}^{-1} = T^{-1}\Omega, \tag{7}$$

$$I_{\sigma\sigma}^{-1} = \left(\frac{2}{T-2}\right)\begin{bmatrix} \sigma_1^4 & \sigma_{12}^2 & \sigma_{12}\sigma_1^2 \\ \sigma_{12}^2 & \sigma_2^4 & \sigma_{12}\sigma_2^2 \\ \sigma_{12}\sigma_1^2 & \sigma_{12}\sigma_2^2 & \frac{1}{2}(\sigma_1^2\sigma_2^2 + \sigma_{12}^2) \end{bmatrix} \equiv \left(\frac{1}{T-2}\right)\Sigma. \tag{8}$$

Thus,

$$\sqrt{T}(\hat{\theta} - \theta) \to N(0, A) \tag{9}$$

with

$$A = p\lim\left(\frac{1}{T}I\right)^{-1} = \begin{bmatrix} \Omega & 0_{2\times 3} \\ 0_{3\times 2} & \Sigma \end{bmatrix}. \tag{10}$$

Under the assumption of $H_0: \sigma_1^2 = \sigma_2^2$, the log-likelihood function (1) becomes

$$l(\theta_r) = -T\ln(2\pi) - \frac{T}{2}\ln|\Omega_r| - \frac{1}{2}\sum_{t=1}^{T}(y_t - \alpha)'\Omega_r^{-1}(y_t - \alpha), \tag{11}$$

https://doi.org/10.1017/S0266466699151089 Published online by Cambridge University Press

where $\theta_r = (\alpha', \sigma_1^2, \sigma_1^2, \sigma_{12})'$ and Ω_r is the Ω under H_0. Then it is easy to show that the constrained MLE of α, σ_1^2, σ_2^2, and σ_{12} are, respectively,

$$\tilde{\alpha} = \bar{y}, \tag{12}$$

$$\tilde{\sigma}_1^2 = \tilde{\sigma}_2^2 = (S_1^2/T + S_2^2/T)/2, \tag{13}$$

$$\tilde{\sigma}_{12} = S_{12}/T. \tag{14}$$

Comparing (12)–(14) with (2)–(4), we see that

$$\tilde{\alpha} = \hat{\alpha}, \tag{15}$$

$$\tilde{\sigma}_1^2 = \tilde{\sigma}_2^2 = (\hat{\sigma}_1^2 + \hat{\sigma}_2^2)/2, \tag{16}$$

$$\tilde{\sigma}_{12} = \hat{\sigma}_{12}. \tag{17}$$

Substituting (12)–(14) into (11), we obtain

$$l(\tilde{\theta}_r) = -T[\ln(2\pi) + 1] - \frac{T}{2}\ln|\tilde{\Omega}| = -T[\ln(2\pi) + 1] - \frac{T}{2}\ln(\tilde{\sigma}_1^2\tilde{\sigma}_2^2 - \tilde{\sigma}_{12}^2). \tag{18}$$

We now wish to show the equivalence of the Hausman test and the likelihood ratio test. We start by deriving the Hausman test. We first notice that, using (15)–(17),

$$\hat{\theta} - \tilde{\theta}_r = \begin{bmatrix} \hat{\alpha} \\ \hat{\sigma}_1^2 \\ \hat{\sigma}_2^2 \\ \hat{\sigma}_{12} \end{bmatrix} - \begin{bmatrix} \tilde{\alpha} \\ \tilde{\sigma}_1^2 \\ \tilde{\sigma}_2^2 \\ \tilde{\sigma}_{12} \end{bmatrix} = \begin{bmatrix} 0_{2\times 1} \\ (\hat{\sigma}_1^2 - \hat{\sigma}_2^2)/2 \\ -(\hat{\sigma}_1^2 - \hat{\sigma}_2^2)/2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0_{2\times 1} \\ 1 \\ -1 \\ 0 \end{bmatrix} [(\hat{\sigma}_1^2 - \hat{\sigma}_2^2)/2] = R\hat{d}, \tag{19}$$

where

$$R \equiv (0, \quad 0, \quad 1, \quad -1, \quad 0)', \tag{20}$$

$$\hat{d} \equiv (\hat{\sigma}_1^2 - \hat{\sigma}_2^2)/2. \tag{21}$$

Thus, we have

$$V \equiv AV(\hat{\theta} - \tilde{\theta}_r) = R \cdot AV(\hat{d}) \cdot R' = \sigma_d^2 RR', \tag{22}$$

where

$$\sigma_d^2 \equiv AV(\hat{d}) = AV[(\hat{\sigma}_1^2 - \hat{\sigma}_2^2)/2] = (\sigma_1^4 - \sigma_{12}^2)/T, \tag{23}$$

using (9) and (10) and the null hypothesis that $\sigma_1^2 = \sigma_2^2$. Under H_0, σ_d^2 can be consistently estimated by

$$\hat{\sigma}_d^2 = (\hat{\sigma}_1^2\hat{\sigma}_2^2 - \hat{\sigma}_{12}^2)/T. \tag{24}$$

https://doi.org/10.1017/S0266466699151089 Published online by Cambridge University Press

Given R in (20) and V in (22), it is easy to verify that the generalized inverse of V is $V^+ = \sigma_d^{-2} RR'/4$, which can be consistently estimated by

$$\hat{V}^+ = \hat{\sigma}_d^{-2} RR'/4 \tag{25}$$

with $\hat{\sigma}_d^2$ as defined in (24). Therefore the Hausman test statistic is

$$\begin{aligned} H &= (\hat{\theta} - \tilde{\theta}_r)' \hat{V}^+ (\hat{\theta} - \tilde{\theta}_r) \\ &= (R\hat{d})' \cdot \hat{\sigma}_d^{-2} RR'/4 \cdot R\hat{d} \qquad \text{(using (19) and (25))} \\ &= \hat{d}^2/\hat{\sigma}_d^2 \end{aligned} \tag{26}$$

using $R'R = 2$, which is asymptotically distributed as $\chi^2(1)$.

We now turn to deriving the likelihood ratio test for $H_0: \sigma_1^2 = \sigma_2^2$. Using expressions of (5) and (18), we have

$$LR = 2[l(\hat{\theta}) - l(\tilde{\theta}_r)] = T \ln(|\tilde{\Omega}_r|/|\hat{\Omega}|) = T \ln[(\tilde{\sigma}_1^2 \tilde{\sigma}_2^2 - \tilde{\sigma}_{12}^2)/(\hat{\sigma}_1^2 \hat{\sigma}_2^2 - \hat{\sigma}_{12}^2)]. \tag{27}$$

Substituting (16) and (17) into the preceding expression and doing some simple algebra, we have

$$LR = T \ln[1 + (\hat{\sigma}_1^2 - \hat{\sigma}_2^2)^2/4(\hat{\sigma}_1^2 \hat{\sigma}_2^2 - \hat{\sigma}_{12}^2)]. \tag{28}$$

Under $H_0: \sigma_1^2 = \sigma_2^2$, using the consistency of MLE, we have

$$(\hat{\sigma}_1^2 - \hat{\sigma}_2^2)^2/4(\hat{\sigma}_1^2 \hat{\sigma}_2^2 - \hat{\sigma}_{12}^2) = o_p(1), \tag{29}$$

when the sample size is big enough. Thus, using (28) and (29) and the second-order Taylor expansion, we have

$$\begin{aligned} LR &= T\{[(\hat{\sigma}_1^2 - \hat{\sigma}_2^2)^2/4(\hat{\sigma}_1^2 \hat{\sigma}_2^2 - \hat{\sigma}_{12}^2)] + [(\hat{\sigma}_1^2 - \hat{\sigma}_2^2)^2/4(\hat{\sigma}_1^2 \hat{\sigma}_2^2 - \hat{\sigma}_{12}^2)]^2 \\ &\qquad + o_p([(\hat{\sigma}_1^2 - \hat{\sigma}_2^2)^2/4(\hat{\sigma}_1^2 \hat{\sigma}_2^2 - \hat{\sigma}_{12}^2)]^2)\} \\ &= T[(\hat{\sigma}_1^2 - \hat{\sigma}_2^2)^2/4(\hat{\sigma}_1^2 \hat{\sigma}_2^2 - \hat{\sigma}_{12}^2)] + o_p(1), \end{aligned} \tag{30}$$

where we used the following fact:

$$\begin{aligned} & T \cdot [(\hat{\sigma}_1^2 - \hat{\sigma}_2^2)^2/4(\hat{\sigma}_1^2 \hat{\sigma}_2^2 - \hat{\sigma}_{12}^2)]^2 \\ &= [T \cdot (\hat{\sigma}_1^2 - \hat{\sigma}_2^2)^2 /4(\hat{\sigma}_1^2 \hat{\sigma}_2^2 - \hat{\sigma}_{12}^2)][(\hat{\sigma}_1^2 - \hat{\sigma}_2^2)^2/4(\hat{\sigma}_1^2 \hat{\sigma}_2^2 - \sigma_{12}^2)] \\ &= \left\{ \left[\frac{1}{2} (\hat{\sigma}_1^2 - \hat{\sigma}_2^2) \right]^2 \Big/ \left[\frac{1}{T} (\hat{\sigma}_1^2 \hat{\sigma}_2^2 - \hat{\sigma}_{12}^2) \right] \right\} [(\hat{\sigma}_1^2 - \hat{\sigma}_2^2)^2/4(\hat{\sigma}_1^2 \hat{\sigma}_2^2 - \hat{\sigma}_{12}^2)] \\ &= \{\hat{d}^2/\hat{\sigma}_d^2\}[(\hat{\sigma}_1^2 - \hat{\sigma}_2^2)^2/4(\hat{\sigma}_1^2 \hat{\sigma}_2^2 - \hat{\sigma}_{12}^2)] \qquad \text{(using (21) and (24))} \\ &= O_p(1) \cdot o_p(1) \qquad \text{(using (26) and (29))} \\ &= o_p(1). \end{aligned} \tag{31}$$

https://doi.org/10.1017/S0266466699151089 Published online by Cambridge University Press

Thus, rewriting (30), we obtain

$$LR = \left[\frac{1}{2}(\hat{\sigma}_1^2 - \hat{\sigma}_2^2)\right]^2 \Big/ \left[\frac{1}{T}(\hat{\sigma}_1^2\hat{\sigma}_2^2 - \hat{\sigma}_{12}^2)\right] + o_p(1)$$
$$= \hat{d}^2/\hat{\sigma}_d^2 + o_p(1), \qquad \textbf{(32)}$$

using (21) and (24). Therefore, comparing (32) with (26), we conclude that the Hausman test is asymptotically equivalent to the likelihood ratio test.

https://doi.org/10.1017/S0266466699151089 Published online by Cambridge University Press