Econometric Theory, **16**, 2000, 451–458. Printed in the United States of America.

PROBLEMS AND SOLUTIONS

PROBLEMS

00.3.1. *Sampling Distribution of OLS under a Logit Model*, proposed by Badi H. Baltagi. In the spirit of Phillips and Wickens (1978) and Oksanen (1991), this problem considers a simple logit regression model

$$y_t = \Lambda(\beta x_t) + u_t \tag{1}$$

for $t = 1,2$, where $\Lambda(z) = e^z/(1 + e^z)$ for $-\infty < z < \infty$. Let $\beta = 1$, $x_1 = 1$, $x_2 = 2$ and assume that the u_t's are independent.

(a) Derive the sampling distribution of the least squares estimator of β, i.e., assuming a linear probit model when the true model is a logit model.

(b) Derive the sampling distribution of the least squares residuals and verify that the estimated variance of $\hat{\beta}_{ols}$ is biased.

REFERENCES

Oksanen, E.H. (1991) A simple approach to teaching generalized least squares theory. *American Statistician* 45, 229–233.

Phillips, P.C.B. & M.R. Wickens (1978) *Exercises in Econometrics*, vol. 1. Oxford: Phillip Allan/Ballinger.

00.3.2. *A Second-Degree Matrix Equation*, proposed by Heinz Neudecker. Let $XAX' = A$ for all matrices A, the unknown matrix X being nonsingular. Show then that $X = \pm I$.

SOLUTIONS

99.4.1. *The KPSS Test for Cointegration in Case of Bivariate Regressions with Linear Trends*—Solution, proposed by Uwe Hassler. In addition to the cointegration model

$$y_t = a + bx_t + e_t, \qquad b \neq 0, e_t \sim I(0), \qquad t = 1,\dots,T, \tag{1}$$

and the scalar $I(1)$ regressor with drift

$$x_t - x_{t-1} = \mu + u_t, \qquad \mu \in \Re\backslash\{0\}, \qquad u_t \sim I(0), \tag{2}$$

I assume the FCLT

$$T^{-0.5} \sum_{j=1}^{[rT]} e_j \Rightarrow B(r) = \omega W(r), \tag{3}$$

 0266-4666/00 $9.50

https://doi.org/10.1017/S0266466600163078 Published online by Cambridge University Press

where $W(r)$ is a standard Brownian motion and ω^2 is proportional to the spectral density of e_t at frequency zero. The KPSS statistic is defined as

$$\text{KPSS} = (\hat{\omega}T)^{-2} \sum_{t=1}^{T} S_t^2, \qquad S_t = \sum_{j=1}^{t} \epsilon_j,$$

where $\hat{\omega}$ is a consistent estimator computed from the OLS residuals ϵ_t. I shall show that

$$(\hat{\omega}^2 T)^{-0.5} S_{[rT]} \Rightarrow V_2(r), \tag{4}$$

where

$$V_2(r) = (2 - 3r)rW(1) + W(r) + 6r(r-1)\int_0^1 W(r)\, dr$$

is the second-level Brownian bridge from Kwiatkowski et al. (1992, eq. 16). By doing so, statements (i) and (ii) are proven at the same time because from (4) it follows that

$$\text{KPSS} \Rightarrow \int_0^1 V_2^2(r)\, dr.$$

To establish (4), I make use of the result by West (1988):

$$T^{1.5}(\hat{b} - b) \Rightarrow N\left(0, \frac{12\omega^2}{\mu^2}\right) \equiv \frac{12\omega}{\mu}\left(\int_0^1 r\, dW(r) - \frac{1}{2}W(1)\right). \tag{5}$$

Next, it is important to note that the linear trend dominates x_t from (2),

$$\begin{aligned} x_t &= x_0 + \mu t + \sum_{i=1}^{t} u_i \\ &= O(1) + O(T) + O_p(T^{0.5}), \end{aligned}$$

so that

$$T^{-2} \sum_{j=1}^{[rT]} x_j \Rightarrow \mu \int_0^r s\, ds = \frac{\mu r^2}{2}. \tag{6}$$

It is straightforward that (3), (5), and (6) sum up to

$$\begin{aligned} T^{-0.5} \sum_{j=1}^{[rT]} \epsilon_j &= T^{-0.5}(b - \hat{b}) \sum_{j=1}^{[rT]} (x_j - \bar{x}) + T^{-0.5} \sum_{j=1}^{[rT]} (e_j - \bar{e}) \\ &\Rightarrow -\frac{12\omega}{\mu}\left(\int_0^1 r\, dW(r) - \frac{1}{2}W(1)\right)\frac{\mu r}{2}(r-1) + \omega(W(r) - rW(1)) \end{aligned}$$

https://doi.org/10.1017/S0266466600163078 Published online by Cambridge University Press

or

$$(\hat{\omega}^2 T)^{-0.5} S_{[rT]} \Rightarrow -6r(r-1)\left(\int_0^1 r\,dW(r) - \frac{1}{2}W(1)\right) + W(r) - rW(1),$$

where this last expression equals $V_2(r)$ from (4) because

$$\int_0^1 r\,dW(r) \equiv W(1) - \int_0^1 W(r)\,dr.$$

This completes the proof.

REFERENCE

West, K.D. (1988) Asymptotic normality, when regressors have a unit root. *Econometrica* 56, 1397–1418.

99.4.2. *A Modified Estimator of the Error Variance in Linear Errors-in-Variables Models*—Solution, proposed by Hua Liang. A simple calculation deduces that $\beta_n - \beta = (\mathbf{W}^T\mathbf{W} - n\Sigma_{uu})^{-1}\{\mathbf{W}^T(\varepsilon - \mathbf{U}\beta) + \Sigma_{uu}\beta\}$ and $\hat{\varepsilon}_i = W_i^T(\beta - \beta_n) + (\varepsilon_i - U_i^T\beta)$. Note that, under the condition of the theorem, $\sqrt{n}(\beta_n - \beta)$ converges to a normal distribution with mean zero and covariance matrix $\Sigma^{-1}C\Sigma^{-1}$, where $\Sigma = E(XX^T)$ and C is the covariance matrix of $W(\varepsilon - U^T\beta)$. This fact and the law of large numbers imply that

$$\frac{1}{n}\sum_{i=1}^n \hat{\varepsilon}_i^k = \frac{1}{n}\sum_{i=1}^n \varepsilon_i^k + o_P(n^{-1/2}) \quad \text{for } k = 2,3,$$

so that ν_n and ξ_n^2 converge in probability to ν^2 and a^2, respectively.

In addition,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n \hat{\varepsilon}_i = \frac{1}{\sqrt{n}}\sum_{i=1}^n (\varepsilon_i - U_i^T\beta) + \frac{1}{n}\sum_{i=1}^n W_i^T\{\sqrt{n}(\beta - \beta_n)\},$$

where $\Sigma_n = 1/n(\mathbf{W}^T\mathbf{W} - n\Sigma_{uu})$. The second term is asymptotically negligible because $\sqrt{n}(\beta - \beta_n) = O_p(1)$ and $1/n\sum_{i=1}^n W_i^T \to EW_1^T = 0$.

It follows from the above arguments that

$$\sqrt{n}\hat{\sigma}_n^2 = \sqrt{n}\sigma_n^2 - \frac{\nu}{\sqrt{n}a^2}\sum_{i=1}^n (\varepsilon_i - U_i^T\beta) + o_P(1). \qquad \textbf{(2)}$$

Substituting σ_n^2 given by (1) in (2), we get the expression

$$\sqrt{n}(\hat{\sigma}_n^2 - \sigma^2) = \frac{1}{\sqrt{n}}\sum_{i=1}^n\left\{(\varepsilon_i - U_i^T\beta)^2 - (\sigma^2 + \beta^T\Sigma_{uu}\beta) - \frac{\nu}{a^2}(\varepsilon_i - U_i^T\beta)\right\} + o_P(1).$$

A central limit theorem and a direct simplification complete the proof of the theorem.

Remark. In the above arguments, we assume that Σ_{uu} is known. Sometimes it is unknown and must be estimated. The usual method of doing so is by partial replication, so that we observe $W_{ij} = X_i + U_{ij}$, $j = 1, \ldots, m_i$.

We consider here only the usual case that $m_i \leq 2$ and assume that a fraction δ of the data has such replicates. Let $\overline{W}_i$ be the sample mean of the replicates. Then a consistent, unbiased method of moments estimate for Σ_{uu} is

$$\hat{\Sigma}_{uu} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m_i} (W_{ij} - \overline{W}_i)(W_{ij} - \overline{W}_i)^T}{\sum_{i=1}^{n} (m_i - 1)}.$$

One estimator of β, say β_n^*, changes only slightly to accommodate the replicates, becoming

$$\beta_n^* = \left\{ \sum_{i=1}^{n} \overline{W}_i \overline{W}_i^T - n(1 - \delta/2)\hat{\Sigma}_{uu} \right\}^{-1} \times \sum_{i=1}^{n} \overline{W}_i Y_i.$$

By replacing β_n and Σ_{uu} by β_n^* and $\hat{\Sigma}_{uu}$, we similarly define a modified estimator $\hat{\sigma}_n^{*2}$, which possess the same advantage as $\hat{\sigma}_n^2$. The details are omitted.

99.4.3. *The Relative Efficiency of the Between-Estimator with Respect to the Within-Estimator*—Solution,[1] proposed by Shiferaw Gurmu. I submit a solution that is generally applicable when x_{it} is a vector of regressors. Let $D = I_{NT} \otimes e_T$ denote the $(NT \times N)$ matrix of individual dummy variables, where e_R is a matrix of ones of dimension R. Define $M_D = I_{NT} - P_D$ as the matrix that delivers deviations from the individual means. In general, $P_z = z(z'z)^{-1}z'$ shall denote the projection matrix on z. Rewrite (4) of the problem in matrix notation as

$$M_e y = M_D x\beta_w + Qx\beta_b + M_e u, \tag{1}$$

where $e = e_{NT}$, $M_e = I_{NT} - P_e$ obtains deviations of observations from the overall means, and $Q = (P_D - P_e) = (M_e - M_D)$.

(a) Observe that M_D and M_e are symmetric idempotent matrices and that $P_D P_e = P_e$. The key in establishing the result is to note that the vector of regressors $M_D x$ and Qx are orthogonal, i.e., $x'M_D Qx = 0$. This follows from $M_D Q = 0$. Hence, the least squares estimator of β_w from (4) in the problem is the within-estimator:

$$\hat{\beta}_w = (x'M_D x)^{-1} x'M_D y = W_{xy}/W_{xx}, \tag{2}$$

https://doi.org/10.1017/S0266466600163078 Published online by Cambridge University Press

where $W_{xy} = \sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it} - \bar{x}_{i.})(y_{it} - \bar{y}_{i.})$, and that of β_b is the between-estimator:

$$\hat{\beta}_b = [x'Qx]^{-1}x'Qy = B_{xy}^{*}/B_{xx}^{*}, \tag{3}$$

where $B_{xy}^{*} = \sum_{i=1}^{N}(\bar{x}_{i.} - \bar{x}_{..})(\bar{y}_{i.} - \bar{y}_{..})$ and $B_{xx}^{*} = \sum_{i=1}^{N}(\bar{x}_{i.} - \bar{x}_{..})^2$.

(b) The error term in (1) can be decomposed as $u = D\mu + \nu$ with $M_e D = 0$ and $M_e u = M_e \nu$. Hence, the covariance matrix of the disturbance term in (1) is $E(M_e \nu\nu' M_e) = \sigma_\nu^2 M_e$ with generalized inverse $\sigma_\nu^{-2}M_e$. Because $M_D M_e = M_D$ and $QM_e = Q$, the GLS on (1) gives

$$\begin{aligned}\hat{\beta}_{g1} &= (x'M_D\sigma_\nu^{-2}M_e M_D x)^{-1}x'M_D\sigma_\nu^{-2}M_e M_D y\\ &= (x'M_D x)^{-1}x'M_D y = \hat{\beta}_w\end{aligned}$$

and

$$\begin{aligned}\hat{\beta}_{g2} &= [x'Q\sigma_\nu^{-2}M_e Qx]^{-1}[x'Q\sigma_\nu^{-2}M_e Qy]\\ &= [x'Qx]^{-1}x'Qy = \hat{\beta}_b.\end{aligned}$$

Accordingly, GLS on (1) is equivalent to OLS on (1).

(c) Rewrite (1) of the problem in matrix form as $y = e\alpha + x\beta + D\mu + \nu$. Premultiplying both sides by $M_D M_e = M_D$, the within-regression equation is

$$M_D y = M_D x\beta + M_D \nu. \tag{4}$$

Least squares on this equation yields the within-estimator with

$$\operatorname{var}(\hat{\beta}_w) = \sigma_v^2(x'M_D x)^{-1} = \sigma_v^2/W_{xx}. \tag{5}$$

Next, consider the between-regression model $\bar{y}_{i.} = \alpha + \bar{x}_{i.}'\beta + \mu_i + \bar{\nu}_{i.}$ with $\operatorname{var}(\mu_i + \bar{\nu}_{i.}) = T^{-1}(T\sigma_\mu^2 + \sigma_\nu^2)$. Stacking over all N observations, this can be written as $\bar{y}_{.} = e_N\alpha + \bar{x}_{.}\beta + \mu + \bar{\nu}_{.}$, where $\bar{y}_{.} = (\bar{y}_{1.}, \bar{y}_{2.}, \ldots, \bar{y}_{N.})'$ and so on. Premultiplying both sides of the between-regression equation by $M_{e_N} = I_N - P_{e_N}$ yields

$$M_{e_N}\bar{y}_{.} = M_{e_N}\bar{x}_{.}\beta + M_{e_N}\mu + M_{e_N}\bar{\nu}_{.}, \tag{6}$$

where M_{e_N} obtains deviations of across time means from the overall means. Least squares on (6) yields the between-estimator with

$$\operatorname{var}(\hat{\beta}_b) = T^{-1}(T\sigma_\mu^2 + \sigma_\nu^2)(\bar{x}_{.}'M_{e_N}\bar{x}_{.})^{-1} = (T\sigma_\mu^2 + \sigma_\nu^2)/B_{xx}. \tag{7}$$

Accordingly, using (5) and (7) for the scalar case, the relative efficiency of the between-estimator with respect to the within-estimator is

$$\begin{aligned}\operatorname{var}(\hat{\beta}_w)/\operatorname{var}(\hat{\beta}_b) &= [B_{xx}/W_{xx}][\sigma_v^2/(T\sigma_\mu^2 + \sigma_\nu^2)]\\ &= [B_{xx}/W_{xx}][(1-\rho)/(T\rho + (1-\rho))].\end{aligned}$$

(d) Consider the test of linear restriction $\gamma_1 = R(\beta_w, \beta_b)' = 0$, where $R = (1 - 1)$. Because the covariance between the two estimators is zero, the square of the t statistic is

$$q_1 = (\hat{\beta}_w - \hat{\beta}_b)^2[\text{var}(\hat{\beta}_w) + \text{var}(\hat{\beta}_b)]^{-1}, \tag{8}$$

where the variance terms are given in (5) and (7). (Observe that least squares on (4) in the problem or (1) yields incorrect variance of the between-estimator!) The statistic (8) has an approximate $\chi^2(1)$ distribution. The usual Hausman test for panel data is based on the comparison of the GLS estimator, say $\hat{\beta}_g$, and the within-estimator, i.e. $H_0\colon \gamma_2 = \beta_g - \beta_w = 0$. It is well known that the GLS estimator can be expressed as the weighted average of the between- and within-estimators (see Hsiao, 1986, p. 36). By specializing to a single independent variable case, we have

$$\hat{\beta}_g = \Delta\hat{\beta}_b + (1 - \Delta)\hat{\beta}_w \quad \text{or} \quad \hat{\beta}_g - \hat{\beta}_w = \Delta(\hat{\beta}_b - \hat{\beta}_w),$$

where $\Delta = [(1 - \rho)B_{xx}/(W_{xx} + (1 - \rho)B_{xx})]$ is the weight. It follows that Hausman's test is equivalent to testing $\gamma_2 = \Delta(\beta_b - \beta_w) = 0$. The corresponding test statistic is

$$q_2 = [\Delta(\hat{\beta}_w - \hat{\beta}_b)]^2[\Delta^2(\text{var}(\hat{\beta}_w) + \text{var}(\hat{\beta}_b))]^{-1} = q_1. \tag{9}$$

This shows that q_1 and q_2, the Hausman's test for $\gamma_2 = 0$, are numerically identical.

Remarks. These results are applicable to models with k regressors, as appropriate. The solutions in parts (a) and (b) hold by replacing x by an $(NT \times k)$ matrix X and by suitably modifying W's and B's. For part (c), the variance–covariance matrices of the estimators can be compared, i.e.,

$$\text{var}(\hat{\beta}_w) - \text{var}(\hat{\beta}_b) = \sigma_v^2(X'M_D X)^{-1} - (T\sigma_\mu^2 + \sigma_\nu^2)(T\bar{X}_.'M_{e_N}\bar{X}_.)^{-1}.$$

Finally, by analogous minor modification of the proof in part (d), Hausman's test statistic can be shown to be numerically equivalent to the test statistics of $H_0\colon \Delta(\beta_b - \beta_w) = 0$ and $H_0\colon (\beta_b - \beta_w) = 0$, where Δ now is a nonsingular matrix.

NOTE

1. Excellent solutions have been proposed independently by H. Erlat and by B. Baltagi, the poser of the problem.

REFERENCE

Hsiao, C. (1986) *Analysis of Panel Data.* New York: Cambridge University Press.

99.4.4. *An Upper Bound for the Eigenvalues of the Product of a Symmetric Idempotent and a Non-Negative Definite Matrix*—Solution,[1] proposed by Heinz Neudecker. As $(n \times n)$ B is symmetric idempotent, it is expressible as $B = TT'$, with $T'T = I_k$, k being the rank of B. It is well known that ATT' and $T'AT$ have k eigenvalues in common. Apart from these, ATT' has $n - k$ additional zero eigenvalues. Let us consider the k common eigenvalues $\lambda_i(AB)$.

According to the Poincare separation theory,

$$\lambda_i(AB) = \lambda_i(T'AT) \leq \mu_i(A) \qquad (i = 1, \dots, k).$$

The other $n - k$ (zero) eigenvalues satisfy

$$0 = \lambda_i(AB) \leq \mu_i(A) \qquad (i = k+1, \dots, n)$$

as $A \geq 0$. This establishes the solution

$$\lambda_i(AB) \leq \mu_i(A) \qquad (i = 1, \dots, n).$$

NOTE

1. Excellent solutions have been proposed independently by S. Puntanen, G.P.H. Styan and H.J. Werner and by J. Graffelman and M. van de Velden, the posers of the problem.

99.4.5. *Asymptotic Bias of the OLS Estimator for a Censored Pareto Regression Model*—Solution, proposed by S.K. Sapra.

(a) Because OLS applied to observed data $(y_i', x_i)', i = 1,2,\dots,n$ yields the estimate

$$\hat{\beta} = (SS_{xx})^{-1} SS_{xy}, \tag{1}$$

we have

$$\text{plim}\, \hat{\beta} = [\text{var}(x)]^{-1} [\text{cov}(x, y)]. \tag{2}$$

Given that y^* and x have a joint $(m + 1)$ Pareto density function stated in the problem, we have from a generalization of eq. (20.1) in Johnson and Kotz (1972, p. 285) that

$$E(x|y^*) = (1 + a^{-1} y^*)e, \qquad e = (1,1,\dots 1)', \tag{3}$$

which is linear in y^* with slope a^{-1}. Therefore, it follows from eq. (2) and a proposition in Chung and Goldberger (1984, Sect. 3, p. 532) that

$$\text{plim}\, \hat{\beta} = \lambda\beta, \tag{4}$$

https://doi.org/10.1017/S0266466600163078 Published online by Cambridge University Press

where

$$\lambda = \operatorname{cov}(y, y^*)/\operatorname{var}(y^*)$$
$$= a^{-1}[k^{2-a}(a-1)^2(a-2) + (ak^{1-a} - a^2)(a-2)$$
$$+ a(1-k^{2-a})(a-1)^2], \qquad a > 2, \tag{5}$$

where the expression on the right-hand side of (5) is obtained by noting that the marginal density of y^* is the univariate standard Pareto density

$$f(y^*) = ay^{*-(a+1)}, \qquad y^* > 1, \tag{6}$$

and $y = \min(k, y^*)$. Hence, it follows from (4) that each element of $\hat{\beta}$ has the same proportional asymptotic bias with common proportionality constant λ given in (5).

(b) As $k \to \infty$, λ approaches 1 and therefore $\operatorname{plim} \hat{\beta} = \beta$. Furthermore,

$$\operatorname{plim} \hat{\alpha} = E(y) - (\operatorname{plim} \hat{\beta})' E(x). \tag{7}$$

Substituting for expectations on the right-hand side of (7) yields

$$\operatorname{plim} \hat{\alpha} = \lambda\alpha + \{(a-1)^{-1}[(1-\lambda)a - k^{1-a}]\}, \qquad a > 2, \tag{8}$$

which is equal to α as $k \to \infty$. Hence, $\hat{\beta}$ and $\hat{\alpha}$ are asymptotically unbiased as $k \to \infty$ and $n \to \infty$.

REFERENCES

Chung, C. & A.S. Goldberger (1984) Proportional projections in limited dependent variable models. *Econometrica* 52, 531–534.

Johnson, N.L. & S. Kotz (1972) *Distributions in Statistics: Continuous Multivariate Distributions.* New York: Wiley.

https://doi.org/10.1017/S0266466600163078 Published online by Cambridge University Press