Econometric Theory, **14**, 1998, 285–292. Printed in the United States of America.

PROBLEMS AND SOLUTIONS

PROBLEMS

98.2.1. *Consistency of OLS in the Presence of Lagged Dependent Variable and Serially Correlated Errors*, proposed by Jeffrey M. Wooldridge. Let $\{y_t : t = \ldots,-1,0,1,2,\ldots\}$ be a strictly stationary and ergodic stochastic process. Assume that $E(y_t^2) < \infty$ and $E(y_t) = 0$.

(i) Define $\beta \equiv \text{Cov}(y_{t-1}, y_t)/\text{Var}(y_t)$ and $u_t \equiv y_t - \beta y_{t-1}$, $t = 1,2,\ldots$, so that we can write

$$y_t = \beta y_{t-1} + u_t, \qquad t = 1,2,\ldots.$$

Show that the error u_t satisfies

$$E(u_t) = 0, \qquad \text{Cov}(y_{t-1}, u_t) = 0.$$

(ii) Show that the OLS estimator $\hat{\beta}$ from the regression

$$y_t \quad \text{on } y_{t-1}, \qquad t = 1,\ldots,n,$$

is consistent for β.

(iii) Show that, without further assumptions, $\{u_t\}$ is serially correlated.

(iv) A 1994 paper in the *Journal of Econometrics* leads with the statement "(i)t is well known that in linear regression models with lagged dependent variables, ordinary least squares (OLS) estimators are inconsistent if the errors are autocorrelated." (This statement, or a slight variation on it, appears in virtually all econometrics textbooks.) Reconcile this statement with your findings from parts (ii) and (iii), using the assumption that $\{y_t\}$ is asymptotically uncorrelated: $E(y_t y_{t-j}) \to 0$ as $j \to \infty$. (Hint: Show that $\{u_t\}$ cannot follow an AR(1) process.)

98.2.2. *Upper Bounds for Eigenvalues of Nonnegative Definite Matrices*, proposed by Eric Iksoon Im. Let $A(T \times T)$ be any n.n.d. matrix. Show that

$$\lambda_i(A) \leq T\alpha_i(A) \qquad (i = 1,2,\ldots,T) \qquad \textbf{(1)}$$

where $\lambda_i(\cdot)$ and $\alpha_i(\cdot)$ denote, respectively, the eigenvalues and diagonal elements in descending orders.

98.2.3. *Partial Observability in a Bivariate Logit Model*, proposed by S.K. Sapra. Poirier (1980) discussed a bivariate probit model with partial observability. Consider the following bivariate logit model with partial observability (model P, henceforth):

$$y_{1i}^* = \beta_1' x_i + u_{1i}, \qquad \textbf{(1)}$$

$$y_{2i}^* = \beta_2' x_i + u_{2i}, \qquad \textbf{(2)}$$

© 1998 Cambridge University Press 0266-4666/98 $9.50

https://doi.org/10.1017/S0266466698142068 Published online by Cambridge University Press

where the y_{ji}^* are unobservable, x_i is a $(k \times 1)$ vector of known constants, and β_1 and β_2 are $(k \times 1)$ vectors of unknown parameters. Define binary dependent variables y_{ji} as

$$y_{ji} = 1 \quad \text{if } y_{ji}^* > 0.$$
$$= 0 \quad \text{if } y_{ji}^* \leq 0, \qquad j = 1,2, \qquad i = 1,2,..,n. \tag{3}$$

The errors $(u_{1i}, u_{2i})'$ are i.i.d. with cumulative distribution function $F(z_1, z_2) = P(u_{1i} \leq z_1, u_{2i} \leq z_2) = 1/(1 + e^{-z_1} + e^{-z_2})$.

Suppose that instead of observing x_i, y_{1i}, and y_{2i}, $i = 1,2,\ldots,n$, we observe only x_i and w_i, where $w_i = y_{1i}y_{2i}$, $i = 1,2,\ldots,n$. That is,

$$w_i = 1 \quad \text{if } y_{1i} = 1 \quad \text{and} \quad y_{2i} = 1,$$
$$= 0 \text{ otherwise.} \tag{4}$$

Show that model P (equations (1)–(4)) is not locally identified under each of the following configurations: (i) $\beta_1 = \beta_2$, (ii) all components of β_1 and β_2 except the intercepts are zero, and (iii) for $k = 2$, $x_i = (x_{1i}, x_{2i})'$, $x_{1i} = 1$, and x_{2i} is a dummy variable that takes values 1 or 0, $i = 1,2,\ldots,n$.

REFERENCE

Poirier, D. (1980) Partial observability in bivariate probit model. *Journal of Econometrics* 12, 209–217.

SOLUTIONS

97.2.1. *A Simple Linear Trend Model with Error Components*—Solution,[1] proposed by Seuck Heun Song and Byoung Cheol Jung. In vector form, equation (1) can be written as

$$y = Z\delta + u, \tag{3}$$

where y is $NT \times 1$, $Z = [\iota_{NT}, \iota_N \otimes l_T]$, $\delta' = (\alpha, \beta)$ and ι_{NT} is a vector of ones of dimension NT and $l_T' = (1,2,\ldots,T)$. Also, equation (2) can be written as

$$u = (I_N \otimes \iota_T)\mu + \nu, \tag{4}$$

where $\mu' = (\mu_1,\ldots,\mu_N)$ and $\nu' = (\nu_{11},\ldots,\nu_{1T},\ldots,\nu_{N1},\ldots,\nu_{NT})$. The variance-covariance matrix of u is

$$\Omega = \sigma_\mu^2(I_N \otimes J_T) + \sigma_\nu^2(I_N \otimes I_T), \tag{5}$$

where J_T is a matrix of ones of dimension T.

Solution a. In the context of the simple model (3), we will show that the OLS, GLS, and FE estimators of β are identical.

(1) It is well known (see, e.g., Milliken and Albohali, 1984) that $\hat{\delta}_{OLS}$ is BLUE if and only if

$$Z'\Omega^{-1}\bar{P}_Z = 0, \tag{6}$$

https://doi.org/10.1017/S0266466698142068 Published online by Cambridge University Press

where $\bar{P}_Z = (I_{NT} - Z(Z'Z)^{-1}Z')$. The variance-covariance matrix given in (5) can be rearranged as

$$\Omega = \sigma_1^2(I_N \otimes \bar{J}_T) + \sigma_\nu^2(I_N \otimes E_T), \tag{7}$$

where $\sigma_1^2 = T\sigma_\mu^2 + \sigma_\nu^2$, $\bar{J}_T = J_T/T$, and $E_T = I_T - \bar{J}_T$. This is also the spectral decomposition of Ω (see Wansbeek and Kapteyn, 1982, 1983). This means that

$$\Omega^{-1} = \frac{1}{\sigma_\nu^2}(I_N \otimes I_T - (1-\theta)(I_N \otimes \bar{J}_T)) \quad \text{with } \theta = \sigma_\nu^2/\sigma_1^2. \tag{8}$$

Hence,

$$\begin{aligned} Z'\Omega^{-1} &= \frac{1}{\sigma_\nu^2}\begin{pmatrix} \iota'_{NT} \\ \iota'_N \otimes l'_T \end{pmatrix}(I_N \otimes I_T - (1-\theta)(I_N \otimes \bar{J}_T)) \\ &= \frac{1}{\sigma_\nu^2}\begin{pmatrix} \iota'_{NT} - (1-\theta)\iota'_N \otimes \iota'_T \\ \iota'_N \otimes l'_T - (1-\theta)(\iota'_N \otimes l'_T\bar{J}_T) \end{pmatrix} \\ &= \frac{1}{\sigma_\nu^2}\begin{pmatrix} \theta\iota'_{NT} \\ \iota'_N \otimes \left(l'_T - \dfrac{(1-\theta)(T+1)}{2}\iota'_T\right) \end{pmatrix}. \end{aligned}$$

Similarly,

$$Z'\Omega^{-1}Z = \frac{1}{\sigma_\nu^2}\begin{pmatrix} NT\theta & \dfrac{NT(T+1)}{2}\theta \\ \dfrac{NT(T+1)}{2}\theta & \dfrac{NT(T+1)(T-1)}{12} + \dfrac{NT(T+1)^2}{4}\theta \end{pmatrix}, \tag{9}$$

$$(Z'Z)^{-1} = \frac{1}{NT(T-1)(T+1)}\begin{pmatrix} 2(T+1)(2T+1) & -6(T+1) \\ -6(T+1) & 12 \end{pmatrix},$$

and so

$$\begin{aligned} Z'\Omega^{-1}Z(Z'Z)^{-1}Z' &= \frac{1}{\sigma_\nu^2}\begin{pmatrix} \theta & 0 \\ -\dfrac{(T+1)(1-\theta)}{2} & 1 \end{pmatrix}\begin{pmatrix} \iota'_{NT} \\ \iota'_N \otimes l'_T \end{pmatrix} \\ &= \frac{1}{\sigma_\nu^2}\begin{pmatrix} \theta\iota'_{NT} \\ \iota'_N \otimes \left(l'_T - \dfrac{(1-\theta)(T+1)}{2}\iota'_T\right) \end{pmatrix}. \end{aligned}$$

Therefore, $Z'\Omega^{-1} - Z'\Omega^{-1}Z(Z'Z)^{-1}Z' = 0$.

(2) If the μ_i's are fixed effects in model (1), we can obtain the FE estimator by premultiplying the model (3) by $Q = I_N \otimes E_T$ and performing OLS on the resulting transformed model:

$$y^* = X^*\beta + \nu^*, \tag{10}$$

where $y^* = Qy$, $X^* = Q(\iota_N \otimes l_T)$, and $\nu^* = Q\nu$. This uses the fact that $Q\iota_{NT} = Q(I_N \otimes \iota_T) = 0$. Note that this Q is symmetric, idempotent, and of rank $N(T-1)$.

https://doi.org/10.1017/S0266466698142068 Published online by Cambridge University Press

Hence $\mathrm{Cov}(\nu^*) = \sigma_\nu^2 Q$ is positive semidefinite and not positive definite as the NSC of Milliken and Albohali requires. However, a generalized inverse of Q is Q, and the NSC is applied with Q/σ_ν^2. In fact, condition (6) applied to (10) reduces to

$$X^{*\prime}(Q/\sigma_\nu^2)\bar{P}_{X^*} = ((\iota_N \otimes l_T)Q\bar{P}_{X^*})/\sigma_\nu^2 = (X^{*\prime}\bar{P}_{X^*})/\sigma_\nu^2 = 0.$$

Therefore, the FE estimator of β is also BLUE in this model.

Solution b.

(1) The variance of the OLS, GLS, and FE estimator of β is presented in the southeast corner of the inverse matrix of (9).

$$\mathrm{Var}(\tilde{\beta}_{GLS}) = \mathrm{Var}(\tilde{\beta}_{FE}) = \mathrm{Var}(\hat{\beta}_{OLS}) = \frac{12\sigma_\nu^2}{NT(T+1)(T-1)}. \quad \textbf{(11)}$$

(2) Let

$$D = \begin{bmatrix} -1 & 1 & 0 & . & \dots & . & 0 \\ 0 & -1 & 1 & . & \dots & . & 0 \\ . & . & . & . & \dots & . & . \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & . & \dots & -1 & 1 \end{bmatrix}$$

be the $(T-1) \times T$ difference matrix.

We can obtain the FD estimator from (3) by premultiplying the model by $(I_N \otimes D)$ and performing OLS on the resulting transformed model:

$$y^+ = X^+\beta + u^+, \quad \textbf{(12)}$$

where $y^+ = (I_N \otimes D)y$, $X^+ = (I_N \otimes D)(\iota_N \otimes l_T)$, and $u^+ = (I_N \otimes D)u$. The FD estimator is given by

$$\begin{aligned}\tilde{\beta}_{FD} &= (X^{+\prime}X^+)^{-1}X^{+\prime}y^+ \\ &= ((\iota_N' \otimes l_T')(I_N \otimes D')(I_N \otimes D)(\iota_N \otimes l_T))^{-1}(\iota_N' \otimes l_T')(I_N \otimes D')(I_N \otimes D)y \\ &= \frac{1}{N(T-1)}(\iota_N' \otimes \iota_{T-1}')(I_N \otimes D)y, \quad \text{because } Dl_T = \iota_{T-1}, \\ &= \frac{1}{N(T-1)}(\iota_N' \otimes \iota_{T-1}'D)y,\end{aligned}$$

where $\iota_{T-1}'D = (-1,0,\dots,0,1)$. Therefore,

$$\tilde{\beta}_{FD} = \frac{1}{N(T-1)}\sum_{i=1}^{N}(y_{iT} - y_{i1}). \quad \textbf{(13)}$$

https://doi.org/10.1017/S0266466698142068 Published online by Cambridge University Press

The variance of the FD estimator of β is

$$\begin{aligned}\mathrm{Var}(\tilde{\beta}_{FD}) &= (X^{*\prime}X^{*})^{-1}X^{*\prime}(\sigma_\nu^2 I_N \otimes DD')X^{*}(X^{*\prime}X^{*})^{-1} \\ &= \frac{\sigma_\nu^2}{N^2(T-1)^2}\,(\iota_N' \otimes \iota_{T-1}')(I_N \otimes DD')(I_N \otimes \iota_T) \\ &= \frac{\sigma_\nu^2}{N^2(T-1)^2}\,(\iota_N'\iota_N \otimes \iota_{T-1}'DD'\iota_{T-1}) \\ &= \frac{2\sigma_\nu^2}{N(T-1)^2}. \qquad \textbf{(14)}\end{aligned}$$

It is seen from the preceding variance formulas that $\mathrm{Var}(\tilde{\beta}_{FD}) = O(N^{-1}T^{-2})$, whereas the variance of the remaining estimators is $O(N^{-1}T^{-3})$. Thus it is verified that asymptotic relative efficiency of FD with respect to the other estimator of β, $\mathrm{Var}(\tilde{\beta}_{FD})/\mathrm{Var}(\tilde{\beta}_{*})$, where * denotes OLS, FE, or GLS, respectively, tends to ∞ as $T \to \infty$.

NOTE

1. Excellent solutions have been proposed independently by C. Kao and J. Emerson, and by B.H. Baltagi and W. Krämer, the posers of the problem.

REFERENCES

Milliken, G.A. & M. Albohali (1984) On the necessary and sufficient conditions for ordinary least-squares estimators to be best linear unbiased estimators. *American Statistician* 38, 298–299.

Wansbeek, T. & A. Kapteyn (1982) A simple way to obtain the spectral decomposition of variance components models for balanced data. *Communications in Statistics—Theory and Method* 11, 2105–2111.

Wansbeek, T. & A. Kapteyn (1983) A note on spectral decomposition and maximum likelihood estimation in ANOVA models with balanced data. *Statistics and Probability Letters* 1, 213–215.

97.2.2. *Estimation of a Triangular, Seemingly Unrelated, Regression System by OLS*—Solution,[1] proposed by Enrique Sentana. Let y_{1t}, y_{2t}, y_{3t} denote the three left-hand-side variables and x_{1t}, x_{2t}, x_{3t} the three sets of regressors. For simplicity of notation, let us partition $x_{2t}' = (x_{1t}', x_{2t}^{*\prime})$ and $x_{3t}' = (x_{2t}', x_{3t}^{*\prime}) = (x_{1t}', x_{2t}^{*\prime}, x_{3t}^{*\prime})$. The system can then be written as

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix} = \begin{pmatrix} \pi_{11} & 0 & 0 \\ \pi_{21} & \pi_{22} & 0 \\ \pi_{31} & \pi_{32} & \pi_{33} \end{pmatrix} \begin{pmatrix} x_{1t} \\ x_{2t}^{*} \\ x_{3t}^{*} \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{pmatrix}.$$

The proof is obtained by combining two existing results. First, Schmidt (1978) showed that in this framework, SUR applied to the first two equations is the same (for those equations) as SUR applied to all three equations. Furthermore, SUR applied to the first equation only is the same (for the first equation) as SUR

https://doi.org/10.1017/S0266466698142068 Published online by Cambridge University Press

applied to the first two equations. But SUR applied to the first equation only is just OLS applied to that equation, which proves (a).

Let us now consider the first two equations. Telser (1964) showed that iterative SUR applied to any two seemingly unrelated equations is equivalent to the following iterative procedure. Regress y_{1t} on x_{1t} and save the residuals, $\hat{u}_{1t}^{(1} = y_{1t} - \hat{\pi}_{11}^{(1} x_{1t}$. Add these residuals to the regression of y_{2t} on x_{2t}. The resulting estimated disturbances $\hat{u}_{2t}^{(1} = y_{2t} - \hat{\pi}_{2}^{(1} x_{2t}$ (where $\pi_2 = (\pi_{21}, \pi_{22})$) are then added to the regression of y_{1t} on x_{1t} to obtain new residuals $\hat{u}_{1t}^{(2} = y_{1t} - \hat{\pi}_{11}^{(2} x_{1t}$, which are again added to the regression of y_{2t} on x_{2t}, and so on. But because in our case OLS applied to the first equation is already SUR, it cannot be improved upon, and Telser procedure converges after one iteration, which proves (b).

Finally, Telser procedure applied to three equations involves adding the estimated disturbances from the other equations as regressors in an iterative manner. But because in our case, SUR applied to the first two equations is the same as SUR applied to all three equations, a similar argument shows that Telser procedure converges again in one iteration. Furthermore, by standard OLS properties, we can use in (c) either the estimated disturbances from the first two equations or the OLS residuals from (a) and (b).

NOTE

1. The Editor is grateful to W.M. Mikhail for pointing out that an alternative solution to this problem can be found in Srivastava and Giles (1987, pp. 25–26).

REFERENCES

Schmidt, P. (1978) A note on the estimation of seemingly unrelated regression systems. *Journal of Econometrics* 7, 259–261.

Srivastava, S. & D. Giles (1987) *Seemingly Unrelated Regression Equations Models.* New York: Marcel Dekker.

Telser, L.G. (1964) Iterative estimation of a set of linear regression equations. *Journal of the American Statistical Association* 59, 845–862.

97.2.3. *Two Matrix Inequalities Involving the Moore–Penrose Inverse*—Solution,[1] proposed by Heinz Neudecker and Michel van de Velden. We use the Schur decomposition $\mathbf{C} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}'$ with $\mathbf{T}'\mathbf{T} = \mathbf{I}$ and nonsingular diagonal matrix $\mathbf{\Lambda}$ containing the nontrivial eigenvalues of $\mathbf{C}$. Hence $\mathbf{C}^+ = \mathbf{T}\mathbf{\Lambda}^{-1}\mathbf{T}'$.

The two subproblems can be rewritten as

$$a(\mathbf{TT}' - \mathbf{T\Lambda T}')^2 + b(\mathbf{TT}' - \mathbf{T\Lambda}^{-1}\mathbf{T}')^2 \geq (\mathbf{TT}' - a\mathbf{T\Lambda T}' - b\mathbf{T\Lambda}^{-1}\mathbf{T}')^2$$

or, equivalently,

$$a(\mathbf{I} - \mathbf{\Lambda})^2 + b(\mathbf{I} - \mathbf{\Lambda}^{-1})^2 \geq (\mathbf{I} - a\mathbf{\Lambda} - b\mathbf{\Lambda}^{-1})^2 \qquad \textbf{(1)}$$

and

$$a(\mathbf{TT}' + \mathbf{T\Lambda T}')^+ + b(\mathbf{TT}' + \mathbf{T\Lambda}^{-1}\mathbf{T}')^+ \geq (\mathbf{TT}' + a\mathbf{T\Lambda T}' + b\mathbf{T\Lambda}^{-1}\mathbf{T}')^+$$

https://doi.org/10.1017/S0266466698142068 Published online by Cambridge University Press

or, equivalently,

$$a(\mathbf{I} + \mathbf{\Lambda})^{-1} + b(\mathbf{I} + \mathbf{\Lambda}^{-1})^{-1} \geq (\mathbf{I} + a\mathbf{\Lambda} + b\mathbf{\Lambda}^{-1})^{-1}. \quad \textbf{(2)}$$

Consider the corresponding scalar inequalities

$$a(1 - \lambda)^2 + b(1 - \lambda^{-1})^2 \geq (1 - a\lambda - b\lambda^{-1})^2 \quad \textbf{(1′)}$$

and

$$a(1 + \lambda)^{-1} + b(1 + \lambda^{-1})^{-1} \geq (1 + a\lambda + b\lambda^{-1})^{-1}, \quad \textbf{(2′)}$$

where λ is an arbitrary element on the diagonal of $\mathbf{\Lambda}$, hence an arbitrary nontrivial eigenvalue of $\mathbf{C}$.

In (1′) λ can take any value, and in (2′) λ can take any positive value. Using the relation $a + b = 1$, where $0 < a, b < 1$, and rewriting the two equations (1′) and (2′) as

$$a(1 - \lambda)^2 + b(1 - \lambda^{-1})^2 \geq (a(1 - \lambda) + b(1 - \lambda^{-1}))^2$$

and

$$a(1 + \lambda)^{-1} + b(1 + \lambda^{-1})^{-1} \geq (a(1 + \lambda) + b(1 + \lambda^{-1}))^{-1}$$

leads to the observation that proving the inequalities is equivalent to realizing that the functions $f(x) = x^2$ and $f(y) = y^{-1}$, for $y > 0$, are convex.

Both scalar inequalities become equalities when $\lambda = \lambda^{-1}$; hence $\lambda = \pm 1$ in (1′) and $\lambda = 1$ in (2′). The two matrix inequalities become equalities when the nontrivial eigenvalues of $\mathbf{C}$ are 1 or -1 in the first case and 1 in the second case. Thus, in the second case $\mathbf{C}$ is (symmetric) idempotent.

NOTE

1. An excellent solution has been proposed independently by S. Liu and W. Polasek, the posers of the problem.

97.2.4. *An Inequality Concerning the Hadamard Matrix Product*—Solution, proposed by Heinz Neudecker. Write $A \circ B = J'(A \otimes B)J$, where J is an appropriate semiorthogonal selection matrix and $\otimes$ denotes the Kronecker matrix product. (See Kollo and Neudecker, 1993.) Then,

$$(A \circ B)^2 = J'(A \otimes B)JJ'(A \otimes B)J,$$
$$A^2 \circ B^2 = J'(A^2 \otimes B^2)J = J'(A \otimes B)(A \otimes B)J.$$

Further,

$$A^2 \circ B^2 - (A \circ B)^2 = J'(A \otimes B)(I - JJ')(A \otimes B)J \geq 0.$$

The reason is that $I - JJ' = (I - JJ')^2 \geq 0$. Note that also $(A \circ B)^2$ and $A^2 \circ B^2$ are positive semidefinite, although A and B are not. For a proof of (1) with positive semidefinite A and B see Horn (1989).

https://doi.org/10.1017/S0266466698142068 Published online by Cambridge University Press

REFERENCES

Horn, R.A. (1989) The Hadamard product. In C.R. Johnson (ed.), *Matrix Theory and Applications*, *Proceedings of Symposium in Applied Mathematics*, vol. 40, pp. 87–169.

Kollo, T. & H. Neudecker (1993) Asymptotics of eigenvalues and unit-length eigenvectors of sample variance and correlation matrices. *Journal of Multivariate Analysis* 47, 283–300.

97.2.5. *Equivariance of an Instrumental Variable (IV) Estimator in the Linear Regression Model*—Solution,[1] proposed by Benoit Perron.

(a) The IV estimator is

$$\hat{\beta}(y) = (X'P_zX)^{-1}X'P_zy.$$

Because $y^* = dy + Xc$, the IV estimator for the transformed data is

$$\begin{aligned}\hat{\beta}(y^*) &= (X'P_zX)^{-1}X'P_zy^* \\ &= (X'P_zX)^{-1}X'P_z(dy + Xc) \\ &= d\hat{\beta}(y) + c\end{aligned}$$

after cancellation. Therefore $\hat{\beta}(y)$ is equivariant.

(b) The estimator of the variance is

$$\hat{\sigma}^2(y) = \frac{1}{n}[y - X\hat{\beta}(y)]'[y - X\hat{\beta}(y)],$$

so that

$$\hat{\sigma}^2(y^*) = \frac{1}{n}[y^* - X\hat{\beta}(y^*)]'[y^* - X\hat{\beta}(y^*)].$$

Using the preceding result and the definition of y^*, we obtain that

$$\begin{aligned}\hat{\sigma}^2(y^*) &= \frac{1}{n}[dy + Xc - X(d\hat{\beta}(y) + c)]'[dy + Xc - X(d\hat{\beta}(y) + c)] \\ &= \frac{1}{n}[dy - dX\hat{\beta}(y)]'[dy - dX\hat{\beta}(y)] \\ &= \frac{d^2}{n}[y - X\hat{\beta}(y)]'[y - X\hat{\beta}(y)] \\ &= d^2\hat{\sigma}^2(y)\end{aligned}$$

as required.

NOTE

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

https://doi.org/10.1017/S0266466698142068 Published online by Cambridge University Press