Econometric Theory, **14**, 1998, 381–386. Printed in the United States of America.

PROBLEMS AND SOLUTIONS

PROBLEMS

98.3.1. *Model Identification Based on Conditional Moments*, proposed by Xing Ming. Assume $\boldsymbol{\pi}(\mathbf{z})$ is an $s \times 1$ dimension functional vector of $t \times 1$ dimensional random variables $\mathbf{z}$ composed of the dependent variable y and the $k \times 1$ dimensional independent variables $\mathbf{x}$. The conditional moment model is specified as follows.

$$\mathsf{E}[\boldsymbol{\pi}(\mathbf{z})|\mathbf{x}] = \mathbf{g}(\mathbf{z}, \boldsymbol{\theta}), \tag{1}$$

where $\boldsymbol{\theta}$ is a $q \times 1$ column vector of parameters to be estimated.

DEFINITION. *Assume the true value of the parameters are* $\boldsymbol{\theta}^0$. *Then model* (1) *is identified if and only if* $\mathbf{g}(\mathbf{z}, \boldsymbol{\theta}^0) = \mathbf{g}(\mathbf{z}, \tilde{\boldsymbol{\theta}}) \Rightarrow \tilde{\boldsymbol{\theta}} = \boldsymbol{\theta}^0$. *Model* (1) *is not identified if and only if there exists* $\tilde{\boldsymbol{\theta}} \neq \boldsymbol{\theta}^0$ *such that* $\mathbf{g}(\mathbf{z}, \boldsymbol{\theta}^0) = \mathbf{g}(\mathbf{z}, \tilde{\boldsymbol{\theta}})$ *holds for all the admissible values of* $\mathbf{z}$.

The definition states that $\boldsymbol{\theta}$ is identified if and only if the moment function (1) can be uniquely defined by a single value of $\boldsymbol{\theta}^0$. Unfortunately, as also noted by Gallant (1977), such conditions are difficult to verify directly. But in special cases we can establish identification by inspection. Deb, Ming, and Trivedi (1997) used this definition to establish the identifiability of the moment-based finite mixture models.

Gallant (1977, p. 79) considered the identifiability of the following model.

$$a_0 + \ln y_1 + a_3 x = e_1, \tag{2}$$

$$b_0 + y_2 + b_1 y_1 + b_3 x = e_2, \tag{3}$$

where y_1 and y_2 are the endogenous variables, x is the exogenous variable, and e_1 and e_2 are the error terms. The parameters to be estimated are a_0, a_3, b_0, b_1, and b_3. Show that by using this definition of identification, we can obtain the same identification results as those obtained by Gallant (1977), i.e.,

1. equation (2) is identified;
2. when $a_3 \neq 0$, equation (3) is identified;
3. when $a_3 = 0$, equation (3) is *not* identified; specifically, b_0, b_1 cannot be identified and b_3 can be identified.

REFERENCES

Deb, P., X. Ming, & P.K. Trivedi (1997) Moment-Based Estimation of Finite Mixture Models of Event Counts. Working paper, Department of Economics, Indiana University, Bloomington.

Gallant, A.R. (1977) Three-stage least-squares estimation for a system of simultaneous, nonlinear, implicit equations. *Journal of Econometrics* 5, 71–88.

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

https://doi.org/10.1017/S0266466698143098 Published online by Cambridge University Press

98.3.2. *One–One Transformation of Multidimensional General Distribution to Multidimensional Uniform Distribution*, proposed by Xing Ming. It is well known that any continuous-type random variable X can be transformed into a uniformly distributed random variable on [0,1] by its own distribution function and that the transformation is $1-1$. The conclusion can be extended to a multidimensional random vector.

Assume $(X_1, X_2, \ldots, X_k)$ is a random vector valued in $\mathbf{R}^k$ and its distribution function $F(x_1, x_2, \ldots, x_k)$ is differentiable and strictly increasing in each argument. Prove that there exists a $1-1$ transformation,

$$Y_1 = f_1(X_1, X_2, \ldots, X_k)$$
$$Y_2 = f_2(X_1, X_2, \ldots, X_k)$$
$$\ldots$$
$$Y_k = f_k(X_1, X_2, \ldots, X_k),$$

such that $(Y_1, Y_2, \ldots, Y_k)$ is uniformly distributed on $[0,1]^k$.

98.3.3. *Prediction in the Equicorrelated Regression Model*, proposed by Badi H. Baltagi. Consider the following regression model,

$$y_t = x_t'\beta + u_t \qquad t = 1,2,\ldots,T.$$

where β is $k \times 1$ and the disturbances are equicorrelated with mean zero, $\text{var}(u_t) = \sigma^2$ and $\text{cov}(u_t, u_s) = \rho\sigma^2$ for $t \neq s$ with $-1/(T-1) \leq \rho \leq 1$. Assume that the regression has a constant.

a. Show that OLS is equivalent to GLS (see Kruskal, 1968).
b. Show that $E(s^2) = \sigma^2(1-\rho)$ and $0 \leq E(s^2) \leq [T/(T-1)]\sigma^2$ (see Dufour, 1986).
c. Applying Goldberger's (1962) results, show that the best linear unbiased predictor for y_{T+s} is $\hat{y}_{T+s} = x_{T+s}'\hat{\beta}_{\text{OLS}}$.

REFERENCES

Dufour, J.M. (1986) Bias of s^2 in linear regressions with dependent errors. *American Statistician* 40, 284–285.

Goldberger, A.S. (1962) Best linear unbiased prediction in the generalized linear regression model. *Journal of the American Statistical Association* 57, 369–375.

Kruskal, W. (1968) When are Gauss–Markov and least squares estimators identical? A coordinate-free approach. *Annals of Mathematical Statistics* 39, 70–75.

SOLUTIONS

97.3.1. *A Consistent Estimator for Truncated Poisson Models with Specification Error*—Solution, proposed by Joao Santos Silva. Whatever the value of $0 \leq$

https://doi.org/10.1017/S0266466698143098 Published online by Cambridge University Press

$\delta \leq 1$, a consistent estimator for β can be based on moment conditions of the form

$$E\{X_i' Y_i[Y_i - 1 - \exp(\alpha^* + X_i\beta)]|X, \delta\} = 0. \quad \textbf{(1)}$$

The consistency of a GMM estimator depends solely on the validity of the moment conditions on which it is based (see Hansen, 1982). Therefore, to prove the consistency of a GMM estimator based on (1), it is enough to show that this moment condition holds under the assumed sampling scheme.

Defining $\tilde{\lambda}_i = \exp(\alpha + X_i\beta)$ and noting that $E\{Y_i|X, \delta\} = \tilde{\lambda}_i/[1 - \delta P(Y_i = 0|X)]$ and $E\{Y_i^2|X, \delta\} = (\tilde{\lambda}_i + \eta^2\tilde{\lambda}_i^2)/[1 - \delta P(Y_i = 0|X)]$, it is very easy to show that (1) holds for $\alpha^* = \alpha + \log(1 + \eta^2)$.

REFERENCE

Hansen, L.P. (1982) Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029–1054.

97.3.2. *Coherency Conditions in a Simultaneous Equations Model with an Interval Censored Endogenous Variable*—Solution, proposed by S.K. Sapra. There are three reduced forms for model (1)–(2) in the problem as follows.

1. $y_{2t}^* \leq L_1$:

$$y_{1t} = \beta_1' X_{1t} + \gamma_1 L_1 + u_{1t}, \quad \textbf{(1)}$$

$$y_{2t}^* = (\gamma_2\beta_1')X_{1t} + (\gamma_1\gamma_2)L_1 + \beta_2' X_{2t} + (u_{2t} + \gamma_2 u_{1t})$$

$$\leq L_1, \qquad t = 1,2,\ldots,n. \quad \textbf{(2)}$$

Equation (2) yields

$$[(\gamma_2\beta_1')X_{1t} + \beta_2' X_{2t} + (u_{2t} + \gamma_2 u_{1t})] \leq (1 - \gamma_1\gamma_2)L_1, \quad \text{or} \quad \textbf{(3)}$$

$$1/(1 - \gamma_1\gamma_2)[(\gamma_2\beta_1')X_{1t} + \beta_2' X_{2t} + (u_{2t} + \gamma_2 u_{1t})] \leq L_1 \text{ iff } (1 - \gamma_1, \gamma_2) > 0. \quad \textbf{(4)}$$

2. $L_1 < y_{2t}^* < L_2$:

$$y_{1t} = 1/(1 - \gamma_1\gamma_2)[\beta_1' X_{1t} + (\gamma_1\beta_2')X_{2t} + (u_{1t} + \gamma_1 u_{2t})] \quad \textbf{(5)}$$

$$y_{2t}^* = 1/(1 - \gamma_1\gamma_2)[(\gamma_2\beta_1')X_{1t} + \beta_2' X_{2t} + (u_{2t} + \gamma_2 u_{1t})],$$

$$t = 1,2,\ldots,n. \quad \textbf{(6)}$$

Equation (6) yields

$$1/(1 - \gamma_1\gamma_2)[(\gamma_2\beta_1')X_{1t} + \beta_2' X_{2t} + (u_{2t} + \gamma_2 u_{1t})] > L_1, \quad \text{and} \quad \textbf{(7)}$$

$$1/(1 - \gamma_1\gamma_2)[(\gamma_2\beta_1')X_{1t} + \beta_2' X_{2t} + (u_{2t} + \gamma_2 u_{1t})] < L_2. \quad \textbf{(8)}$$

https://doi.org/10.1017/S0266466698143098 Published online by Cambridge University Press

3. $y_{2t}^{*} \geq L_2$:

$$y_{1t} = \beta_1' X_{1t} + \gamma_1 L_2 + u_{1t}, \tag{9}$$

$$y_{2t}^{*} = (\gamma_2 \beta_1')X_{1t} + \beta_2' X_{2t} + (\gamma_1 \gamma_2)L_2 + (u_{2t} + \gamma_2 u_{1t}),$$
$$\geq L_2, \qquad t = 1,2,\ldots,n. \tag{10}$$

Equation (10) yields

$$[(\gamma_2 \beta_1')X_{1t} + \beta_2' X_{2t} + (u_{2t} + \gamma_2 u_{1t})] \geq (1 - \gamma_1 \gamma_2)L_2, \quad \text{or} \tag{11}$$

$$1/(1 - \gamma_1 \gamma_2)[(\gamma_2 \beta_1')X_{1t} + \beta_2' X_{2t} + (u_{2t} + \gamma_2 u_{1t})] \geq L_2 \text{ iff } (1 - \gamma_1 \gamma_2) > 0. \tag{12}$$

A comparison between the inequalities in (4) and (7) reveals that they are mutually exclusive iff $(1 - \gamma_1 \gamma_2) > 0$. Similarly, a comparison between the inequalities in (8) and (12) reveals that they are mutually exclusive iff $(1 - \gamma_1 \gamma_2) > 0$.

Hence, the reduced form of model (1)–(2) in the problem exists iff $(1 - \gamma_1 \gamma_2) > 0$ because this condition guarantees that for any X_{1t}, X_{2t}, u_{1t}, and u_{2t}, the model produces one and only one y_{2t} for $t = 1,2,\ldots,n$.

97.3.3. *Properties of Idempotent Matrix*—Solution,[1] proposed by Victoria Zinde-Walsh. Consider the characteristic equation

$$\det(\gamma I - C) = 0 \tag{1}$$

for a complex $n \times n$ matrix C. By the fundamental theorem of algebra, the polynomial in (1) has exactly n linear factors $\gamma - \gamma_i$, with γ_i, $i = 1,\ldots n$, denoting the ith characteristic root (of course, the roots do not have to be distinct). Note that the coefficient of γ^{n-1} in (1) equals $-\mathrm{tr}(C)$. If $C = A$, denote γ_i by λ_i; for $C = I - A$, denote γ_i by μ_i; note the one-to-one relation $\lambda_i = 1 - \mu_i$. Rearrange the roots so that $\lambda_i = 1$ ($\mu_i = 0$) for $i = 1,\ldots l$, if $l \geq 1$; $\lambda_i \neq 0$, $\lambda_i \neq 1$ ($\mu_i \neq 0$, or 1) for $i = l + 1,\ldots,k$, if $k \geq 1$; $\lambda_i = 0$ ($\mu_i = 1$) for $i = k + 1,\ldots,n$.

Proof of necessity. The characteristic roots of $C = A^2$ are λ_i^2 (Lipschutz, 1991, Example 11.1); thus, if $A = A^2$, $l = k$. Because similar matrices have the same characteristic polynomial (Lipschutz, 1991, Theorem 8.3) and a complex matrix C is similar to a triangular matrix C_Δ, with characteristic roots on the diagonal (by Lipschutz, 1991, Theorem 11.1), it follows that here $\Sigma\lambda_i = \mathrm{tr}(A) = l$ and $\Sigma\mu_i = \mathrm{tr}(I - A) = n - l$. Moreover, $rk(A) = rk(A_\Delta) = l$ and $rk(I - A) = rk(I - A_\Delta) = n - l$.

Proof of sufficiency. Note that $\mathrm{tr}(A) = rk(A) - \sum_{i=l+1}^{k} \mu_i$ and $\mathrm{tr}(I - A) = rk(I - A) - \sum_{i=l+1}^{k} \lambda_i = rk(I - A) - k + l + \sum_{i=l+1}^{k} \mu_i$. By the condition of equality between ranks and traces, $k = l$. Now consider $C = A^2 - A$. For this C, the characteristic polynomial can be factored as

$$-\det\left[\left(\tfrac{1}{2} + \sqrt{\tfrac{1}{4} + \gamma}\right)I - A\right]\det\left[\left(\tfrac{1}{2} - \sqrt{\tfrac{1}{4} + \gamma}\right)I - A\right].$$

https://doi.org/10.1017/S0266466698143098 Published online by Cambridge University Press

Thus, because each root of A is either 1 or 0, each root γ_i of C has to satisfy either $\frac{1}{2} + \sqrt{\frac{1}{4} + \gamma} = 0$ or 1 or $\frac{1}{2} - \sqrt{\frac{1}{4} + \gamma} = 0$ or 1. It is easy to see that this is possible only if $\gamma_i = 0$ for all i. A complex matrix with all its characteristic roots equal to zero is a zero matrix.

NOTE

1. Excellent solutions have been proposed independently by H. Neudecker and M. van de Velden, by H. Quian, by H.J. Werner, by V. Zinde-Walsh, and by R.E. Hartwig and G. Trenkler, the posers of the problem.

REFERENCE

Lipschutz, S. (1991) *Theory and Problems of Linear Algebra.* New York: McGraw-Hill.

97.3.4. *Order Invariability of Idempotent Matrix*—Solutions.[1] The following solutions have been proposed independently by Karim M. Abadir and Walter Distaso and by Naorayex K. Dastoor. These solutions are based on different types of interesting arguments.

1. Solution—proposed by Karim M. Abadir and Walter Distaso. We will prove the following generalization.

LEMMA. *Let P_X be a projector onto the space spanned by the columns of the $n \times k$ matrix X of rank k. Then, for any invertible $k \times k$ matrix R, the projectors P_{XR} and P_X coincide.*

Proof. Denote the complex-conjugate transpose of a matrix A by A^* and the g-inverse of A by A^- (which satisfies $AA^-A = A$). From Rao and Mitra (1971, p. 112), the projector (idempotent) P_X has the representation

$$P_X \equiv X(X^*\Lambda X)^- X^*\Lambda,$$

where Λ is a nonsingular matrix such that $\text{rank}(X^*\Lambda X) = \text{rank}(X)$. Here, we have $\text{rank}(X) = k$, so that $(X^*\Lambda X)$ is invertible and the formula reduces to

$$P_X = X(X^*\Lambda X)^{-1} X^*\Lambda.$$

Applying this simplification to P_{XR} also, we have

$$P_{XR} = XR(R^*X^*\Lambda XR)^{-1} R^*X^*\Lambda = X(X^*\Lambda X)^{-1} X^*\Lambda = P_X,$$

by the invertibility of R.

When $X = [X_a \;\; X_b]$, the special case given in the original problem yields the explicit form for R as the permutation matrix

$$R = \begin{bmatrix} 0 & I_a \\ I_b & 0 \end{bmatrix},$$

https://doi.org/10.1017/S0266466698143098 Published online by Cambridge University Press

leading to $XR = [X_b \ X_a] \equiv X_r$. The original problem also requires $\Lambda = I_n$ and X real, so that $P_X = X(X'X)^{-1}X'$. It follows from our lemma that $M = I_n - P_X = I_n - P_{XR}$. More generally, our lemma shows that matrices like M satisfy invariances relating to more than just the column-ordering of X, when X is of full column rank. Any nonsingular linear homogeneous transformation of the columns of X leaves its column span unchanged, and the projector (idempotent) P_X is unaffected. A group-theoretic proof of this property is also possible but less transparent.

2. Solution—proposed by Naorayex K. Dastoor. The solution is obtained as a special case of the following theorem, which considers more general forms of idempotent matrices.

THEOREM. *Let Σ be an $n \times n$-positive definite matrix and $\mathcal{R}(A)$ denote the range space of A. Given $n \times k$ matrices X and Z, each with full column rank, the following statements are equivalent*:

(a) $I_n - X(X'\Sigma X)^{-1}X'\Sigma = I_n - Z(Z'\Sigma Z)^{-1}Z'\Sigma$;
(b) $X(X'\Sigma X)^{-1}X'\Sigma = Z(Z'\Sigma Z)^{-1}Z'\Sigma$;
(c) $\mathcal{R}(X) = \mathcal{R}(Z)$;
(d) *there exists a $k \times k$ nonsingular matrix C such that $XC = Z$.*

Proof. The equivalence of (a) and (b) is obvious; (b) $\Leftrightarrow$ (c) is obtained by noting that $A(A'\Sigma A)^{-1}A'\Sigma$ is the orthogonal projector onto $\mathcal{R}(A)$ for $A = X, Z$, with the inner product defined by $\langle x, y\rangle = x'\Sigma y$; and (c) $\Leftrightarrow$ (d) is Theorem 5.4.5 in Graybill (1983, p. 89).

Now let

$$\Sigma = I_n, X = [X_A, X_B], C = \begin{bmatrix} 0_{a\times b} & I_a \\ I_b & 0_{b\times a} \end{bmatrix} \quad \text{and} \quad Z = XC = [X_B, X_A] = X_R;$$

then the equivalence of statements (a) and (d) of the theorem provides

$$I_n - X(X'X)^{-1}X' = I_n - X_R(X_R'X_R)^{-1}X_R',$$

as required.

NOTE

1. Excellent solutions have been proposed independently by F. Goerlich, by S.A. El-Sheneity and A.S. Hadi, by J.H. Kim, by D. Li, by W. Mikhail, by G. Trenkler, J. Gross, and S.O. Troschke, by M. Van de Velden, by H.J. Werner, and by K. Ham, E. Im, and M.S. Snow, the posers of the problem.

REFERENCES

Graybill, F.A. (1983) *Matrices with Applications in Statistics*, 2nd ed. Belmont: Wadsworth.
Rao, C.R. & S.K. Mitra (1971) *Generalized Inverse of Matrices and its Applications.* New York: John Wiley & Sons.

https://doi.org/10.1017/S0266466698143098 Published online by Cambridge University Press