The maximum number of omitted variables

Problem 00.2.2

Dmitri L. Danilov and Jan R. Magnus

Econometric Theory, 16, 287–288 (2000)

Econometric Theory, **16**, 2000, 287–299. Printed in the United States of America.

PROBLEMS AND SOLUTIONS

PROBLEMS

00.2.1. *Degeneration of Feasible GLS to 2SLS in a Limited Information Simultaneous Equations Model*, proposed by Chuanming Gao and Kajal Lahiri. Consider a simple limited information simultaneous equations model,

$$y_1 = \gamma y_2 + u, \tag{1}$$

$$y_2 = X\beta + v, \tag{2}$$

where y_1, y_2 are $N \times 1$ vectors of observations on two endogenous variables. X is $N \times K$ matrix of predetermined variables of the system, and $K \geq 1$ such that (1) is identified. Each row of (u,v) is assumed to be i.i.d. $(0,\Sigma)$, and Σ is p.d.s.

Following Lahiri and Schmidt (1978), feasible GLS for (1) and (2) based on a consistent estimate of Σ yields a consistent estimate for γ. Pagan (1979) showed that an iterated Aitken estimator will generate LIML estimate of γ.

Denote $\hat{\gamma}_{2SLS} = (y_2' P y_2)^{-1} y_2' P y_1$, where $P = X(X'X)^{-1}X'$. The residuals $\hat{u} = y - \hat{\gamma}_{2SLS} y_2$ and $\hat{v} = My_2$, where $M = I_N - P$, may be used to generate a consistent estimate for Σ, e.g.,

$$\hat{\Sigma} = \frac{1}{N}\begin{bmatrix} \hat{u}'\hat{u} & \hat{u}'\hat{v} \\ \hat{v}'\hat{u} & \hat{v}'\hat{v} \end{bmatrix}.$$

Show that a feasible GLS estimate of γ using $\hat{\Sigma}$ (i.e., the first iterate of iterated Aitken) degenerates to $\hat{\gamma}_{2SLS}$.

REFERENCES

Lahiri, K. & P. Schmidt (1978) On the estimation of triangular structural systems. *Econometrica* 46, 1217–1222.

Pagan, A. (1979) Some consequences of viewing LIML as an iterated Aitken estimator. *Economics Letters* 3, 369–372.

00.2.2. *The Maximum Number of Omitted Variables*, proposed by Dmitri L. Danilov and Jan R. Magnus. Consider the standard partitioned regression model $y = X_1\beta_1 + X_2\beta_2 + u$, where $X \equiv (X_1 : X_2)$ is a nonstochastic $n \times k$ matrix with full column rank $k = k_1 + k_2$. We are interested in estimating β_1 and consider β_2 as a nuisance parameter. Let $r = \text{rank}(X_1'X_2)$. Show that we may assume, without loss of generality, that $k_2 = r$ and, hence, in particular that $k_2 \leq k_1$. Can we still make this simplifying assumption when drawing inferences about β_1?

In the special case where $r = 0$ and where consequently X_2 is orthogonal to X_1, we may delete X_2 altogether, a well-known result.

 0266-4666/00 $9.50

In another special case where $k_1 = 1$ (one "focus" parameter and the rest nuisance parameters), it is sufficient to consider just *one* nuisance parameter.

00.2.3. *Effects of Transforming the Duration Variable in Accelerated Failure Time (AFT) Models*, proposed by S.K. Sapra. Consider the following AFT model,

$$\ln t = \beta' x + \varepsilon, \tag{1}$$

where x is a $(m \times 1)$ vector of known constants, β is a $(m \times 1)$ vector of unknown parameters, $\varepsilon = \ln t_0 - E(\ln t_0)$, and t_0 is a random variable with a density function not involving x or β.

(a) Show that the following transformations of t lead to AFT models: (i) $y = kt, k > 0$, and (ii) $y = t^k$, where k is a constant.
(b) Show that the following transformations of t do not lead to AFT models: (i) $y = a + bt$, $a > 0, b > 0$, and (ii) $y = \exp(a + bt)$, where a and b are constants.
(c) Derive the hazard functions for the density functions of y in parts (a) and (b) by using the transformations of t defined therein.

00.2.4. *Conflict Among Criteria for Testing Hypotheses: Examples from Non-Normal Distributions*, proposed by Badi H. Baltagi. Berndt and Savin (1977) showed that $W \geq LR \geqq LM$ for the case of a multivariate regression model with normal disturbances. Ullah and Zinde-Walsh (1984) showed that this inequality is not robust to non-normality of the disturbances. In the spirit of the latter article, this problem considers simple examples from non-normal distributions and illustrates how this conflict among criteria is affected.

(a) Consider a random sample $x_1, x_2, \ldots, x_n$ from a Poisson distribution with parameter λ. Show that for testing $\lambda = 3$ versus $\lambda \neq 3$ yields $W \geq LM$ for $\bar{x} \leq 3$ and $W \leq LM$ for $\bar{x} \geq 3$.
(b) Consider a random sample $x_1, x_2, \ldots, x_n$ from an exponential distribution with parameter θ. Show that for testing $\theta = 3$ versus $\theta \neq 3$ yields $W \geq LM$ for $0 < \bar{x} \leq 3$ and $W \leq LM$ for $\bar{x} \geq 3$.
(c) Consider a random sample $x_1, x_2, \ldots, x_n$ from a Bernoulli distribution with parameter θ. Show that for testing $\theta = 0.5$ versus $\theta \neq 0.5$, we will always get $W \geq$ LM. Show also, that for testing $\theta = \frac{2}{3}$ versus $\theta \neq \frac{2}{3}$ we get $W \leq LM$ for $\frac{1}{3} \leq \bar{x} \leq \frac{2}{3}$ and $W \geq LM$ for $\frac{2}{3} \leq \bar{x} \leq 1$ or $0 < \bar{x} \leq \frac{1}{3}$.

REFERENCES

Berndt, E.R. & N.E. Savin (1977) Conflict among criteria for testing hypotheses in the multivariate linear regression model. *Econometrica* 45, 1263–1278.

Ullah, A. & V. Zinde-Walsh (1984) On the robustness of LM, LR and W tests in regression models. *Econometrica* 52, 1055–1065.

Solution

Econometric Theory, 17, 485 (2001)

00.2.2. *The Maximum Number of Omitted Variables*—Solution, proposed by Dmitri L. Danilov and Jan R. Magnus. If $r = 0$, the result is simple and well known. We assume that $r \geq 1$. Let $(S:T)$ be an orthogonal $k_2 \times k_2$ matrix such that

$$X_2' X_1 X_1' X_2 S = S\Lambda, \qquad X_1' X_2 T = 0,$$

where Λ denotes an $r \times r$ diagonal matrix with positive diagonal elements. Notice that the dimensions of S and T are $k_2 \times r$ and $k_2 \times (k_2 - r)$, respectively. Because $X_2' X_2$ has full rank k_2, we obtain

$$r(X_2 T) = r(X_2' X_2 T) = r(T) = k_2 - r,$$

so that the $n \times (k_2 - r)$ matrix $W_2 \equiv X_2 T$ has full column rank. Hence, we may define the idempotent matrix $M_2 = I_n - W_2(W_2' W_2)^{-1} W_2'$.

Now, let $W_1 \equiv M_2 X_2 S$, an $n \times r$ matrix. Because

$$W_1 = M_2 X_2 S = X_2 S - W_2(W_2' W_2)^{-1} W_2' X_2 S,$$

we obtain $X_1' W_1 = X_1' X_2 S$ and, hence,

$$X_2' X_1 X_1' W_1 = X_2' X_1 X_1' X_2 S = S\Lambda,$$

so that $r = r(X_2' X_1 X_1' W_1) \leq r(W_1) \leq r$ and, hence, $r(W_1) = r$.

Next, let $W \equiv (W_1 : W_2)$. We already know that $r(W_1) = r$ and $r(W_2) = k_2 - r$. Because $M_2 W_2 = 0$, it follows that $W_1' W_2 = 0$ and, hence, that $r(W) = r(W_1) + r(W_2) = r + k_2 - r = k_2$.

Finally, we observe that

$$M_2 X_2 = X_2 - X_2 T(W_2' W_2)^{-1} W_2' X_2 = X_2 P$$

for some matrix P and, hence,

$$W = (W_1 : W_2) = (M_2 X_2 S : X_2 T) = (X_2 PS : X_2 T) = X_2 Q$$

for some $k_2 \times k_2$ matrix Q. Because $r(W) = k_2$, Q is non-singular.

It is now easy to see that W_2 is orthogonal to both W_1 and X_1. Also, the space spanned by the k_2 columns of W is identical to the space spanned by the k_2 columns of X_2, so that $X_2\beta_2 = W\delta$ for some choice of δ (namely $\delta = Q^{-1}\beta_2$). Hence, the estimator $\hat{\beta}_1$ obtained from a regression of y on X_1 and X_2 will be identical to the estimator obtained from a regression of y on X_1 and W_1, and W_1 only has r columns.

When drawing inferences about β_1, we assume that $u \sim N(0, \sigma^2 I_n)$. The estimator of σ^2 will be biased upward if we delete W_2 from our regression, even though W_2 is orthogonal to both X_1 and W_1, just as in the standard textbook case.