Optimal structural estimation of triangular systems: I. The stationary case

Boswijk, H.P.

Published in:
Econometric Theory

DOI:
10.1017/S0266466600004667

Link to publication

Citation for published version (APA):
Boswijk, H. P. (1991). Optimal structural estimation of triangular systems: I. The stationary case. Econometric Theory, 7, 428-430. DOI: 10.1017/S0266466600004667

Let $N = I_n - n^{-1}ss'$ and $Y = NX$. Then

$$B = -\tfrac{1}{2}NDN = NXX'N' = YY'$$

is clearly a positive semidefinite matrix as required.

NOTE

1. It has been pointed out by Ali S. Hadi and Martin Wells (Cornell University, USA), that a solution to this problem is given in Theorem 14.2.1(a) of [1]. The solution published here, which is similar to the one proposed by H.P. Boswijk and H. Neudecker, the posers of the problem, is quite elegant, being a solution in matrix terms. A very good solution has been proposed by Klaus Pötzelberg (University of Basel, Switzerland), and a good solution by Bernd Schipp (University of Dortmund, Germany).

REFERENCE

1. Mardia, K.V., J.T. Kent & J.M. Bibby. *Multivariate Analysis.* New York: Academic Press, 1979.

90.2.5. *Optimal Structural Estimation of Triangular Systems: I. The Stationary Case*—Solution,[1] proposed by H. Peter Boswijk. Express the model in matrix notation as

$$y_1 = y_2\beta + u_1, \tag{1}$$

$$y_2 = X\gamma + u_2, \tag{2}$$

with $(u_1, u_2) \sim N(0, \sigma^2\Sigma_0 \otimes I_n)$, where y_1, y_2, u_1, and u_2 are $n \times 1$ vectors, X is an $n \times k$ matrix, β and σ^2 are scalars, γ is a k-vector, and

$$\Sigma_0 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

The conditional model for y_{1t} given y_{2t} reads as

$$y_1 = y_2(1 + \beta) - X\gamma + \nu_1, \tag{1$'$}$$

with $\nu_1 = u_1 - u_2$. Let $P_X = X(X'X)^{-1}X'$, and let the symbol "$\Rightarrow$" denote convergence of probability measures (hence convergence in distribution if the limit is an element of a Euclidean space, and convergence in probability if the limit is nonstochastic).

(A) The 2SLS estimator is

$$\hat{\beta}_{2SLS} = \beta + (y_2'P_Xy_2)^{-1}y_2'P_Xu_1. \tag{3}$$

Because $n^{-1}X'X \to M$ as $n \to \infty$, and X is exogenous, we have $X'u_2 = O_p(n^{1/2})$, so that

$$n^{-1}y_2'P_Xy_2 = n^{-1}\gamma'X'X\gamma + o_p(1) \Rightarrow \gamma'M\gamma, \tag{4}$$

$$n^{-1/2}y_2'P_Xu_1 = n^{-1/2}\gamma'X'u_1 + o_p(1) \Rightarrow N(0, 2\sigma^2\gamma'M\gamma). \tag{5}$$

Hence we have, provided $\gamma \neq 0$,

$$n^{1/2}(\hat{\beta}_{2SLS} - \beta) \Rightarrow N(0, 2\sigma^2(\gamma' M \gamma)^{-1}). \tag{6}$$

(B) The OLS estimator in (1)′ can be expressed as

$$\hat{\beta}_{OLS} = \beta + (y_2'(I_n - P_X)y_2)^{-1} y_2'(I_n - P_X)v_1. \tag{7}$$

We now have

$$n^{-1} y_2'(I_n - P_X)y_2 = n^{-1} u_2' u_2 + o_p(1) \Rightarrow \sigma^2, \tag{8}$$

$$n^{-1/2} y_2'(I_n - P_X)v_1 = n^{-1/2} u_2' v_1 + o_p(1) \Rightarrow N(0, \sigma^4), \tag{9}$$

so that

$$n^{1/2}(\hat{\beta}_{OLS} - \beta) \Rightarrow N(0,1). \tag{10}$$

(C) The loglikelihood equals

$$\begin{aligned}\mathcal{L}(\beta, \gamma, \sigma^2) = {} & -T \ln 2\pi - T \ln \sigma^2 \\ & - (y_1 - y_2(1+\beta) + X\gamma)'(y_1 - y_2(1+\beta) + X\gamma)/(2\sigma^2) \\ & - (y_2 - X\gamma)'(y_2 - X\gamma)/(2\sigma^2).\end{aligned} \tag{11}$$

Differentiating (11) yields the following estimator generating equations:

$$y_2' y_2 \hat{\beta} - y_2' X \hat{\gamma} = y_2'(y_1 - y_2),$$

$$X' y_2 \hat{\beta} - 2X'X\hat{\gamma} = X'(y_1 - 2y_2).$$

After elimination of $\hat{\gamma}$, this results in

$$\begin{aligned}[y_2'(I_n - P_X)y_2 + \tfrac{1}{2} y_2' P_X y_2]\hat{\beta}_{ML} &= y_2'(I_n - P_X)(y_1 - y_2) + \tfrac{1}{2} y_2' P_X y_1 \\ &= y_2'(I_n - P_X)y_2 \hat{\beta}_{OLS} + \tfrac{1}{2} y_2' P_X y_2 \hat{\beta}_{2SLS}.\end{aligned} \tag{12}$$

Thus the maximum likelihood estimator is a weighted average of the OLS and 2SLS estimators, with the weights proportional to their inverted variances. Consider (5) and (9). Because

$$E x_t u_{1t} u_{2t} v_{1t} = E x_t (u_{2t} + v_{1t}) u_{2t} v_{1t} = 0,$$

these two results hold jointly, with zero asymptotic covariance. Therefore, the OLS and 2SLS estimators are asymptotically independent, and, from (12), we have

$$n^{1/2}(\hat{\beta}_{ML} - \beta) \Rightarrow N\left(0, \sigma^2\left(\sigma^2 + \frac{1}{2}\gamma' M \gamma\right)^{-1}\right) \tag{13}$$

It is easily checked that the ML estimator has smaller variance than the OLS and 2SLS estimators. Hence, both econometricians are wrong. Econometrician (A) is wrong because the covariance matrix restrictions *overidentify* the system. Therefore, 2SLS estimation is not efficient. His colleague (B) is wrong because y_{2t}, though predetermined in (1)′, is not *weakly exogenous* for β in the sense of Engle et al. [1]. This is because the parameters of the conditional model (1)′ and the marginal model (2) are not "variation free," as γ enters (unrestrictedly) in both equations.

The relative efficiency of the 2SLS and OLS estimators depends on

$$\rho^2 = \frac{\gamma' M \gamma}{\sigma^2 + \gamma' M \gamma},$$

which may be interpreted as an asymptotic coefficient of determination of equation (2). If $\rho^2 > \frac{2}{3}$, the 2SLS estimator has smaller variance; otherwise the OLS estimator is more efficient.

NOTE

1. Excellent solutions have been proposed, independently, by Juan J. Dolado (Bank of Spain) and Peter C.B. Phillips, the poser of the problem. The remarks below are extracted from P.C.B. Phillips' solution.

REFERENCE

1. Engle, R.F., D.F. Hendry & J.-F. Richard. Exogeneity. *Econometrica* 51 (1983): 277–304.

Remarks (by P.C.B. Phillips): (i) Observe that when $\gamma = 0$ the limit distributions (10) and (13) are identical and $\hat{\beta}_{\mathrm{OLS}}$ and $\hat{\beta}_{\mathrm{ML}}$ are asymptotically equivalent. In this case the first equation (1) would be unidentified were it not for the identifying information in the knowledge of the error covariance matrix structure $\Sigma = \sigma^2 \Sigma_0$. Both $\hat{\beta}_{\mathrm{ML}}$ and $\hat{\beta}_{\mathrm{OLS}}$ utilize this identifying information. $\hat{\beta}_{\mathrm{2SLS}}$ does not, and the variance of the limit distribution (σ) is undefined. In fact, $\hat{\beta}_{\mathrm{2SLS}}$ converges in distribution to a random variable, reflecting the uncertainty about β that persists asymptotically when the identifying information in the error covariance matrix is ignored—see Phillips [1]. But since Σ_0 is known in this case we may regard $\hat{\beta}_{\mathrm{2SLS}}$ as being infinitely deficient, that is, the ratio of the variances of the limit distributions (13) and (σ) is zero.

(ii) For large $\gamma' M \gamma$, the limit distributions (σ) and (13) are close and both are degenerate as $\gamma' M \gamma \to \infty$. On the other hand, the limit distribution (10) is invariant to $\gamma' M \gamma$. Thus, for large $\gamma' M \gamma$, 2SLS is nearly as asymptotically efficient as the full systems MLE. In effect, as $\gamma' M \gamma$ increases, the signal from the regressors in the reduced form equation (2) is stronger. This increases the precision in the estimation of the reduced form and thereby improves the effectiveness of the instruments in the 2SLS regression. The