ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 17 — 1990

Comment on "Instrumental Variables and Maximum Likelihood"

Theo DIJKSTRA, Tom WANSBEEK *

ABSTRACT. – HOLLY and MAGNUS [1988] show that the IV estimator in a linear equation is asymptotically as efficient as the ML estimator in the model that is obtained by "completing" this equation to a complete equations system. Their proof is long and involves many matrix manipulations. A simpler approach is provided here.

Commentaire sur « Variables instrumentales et maximum de vraisemblance »

RÉSUMÉ. – HOLLY et MAGNUS [1988] montrent que l'estimateur IV dans une équation linéaire est asymptotiquement aussi efficace que l'estimateur ML dans le modèle obtenu en « complétant » cette équation par un système d'équations complet. Leur démonstration est longue et comprend de nombreuses manipulations sur des matrices. Nous présentons ici une approche plus simple.

* T. DIJKSTRA, T. WANSBEEK: Department of Economics, Groningen University, P.O. Box 800, 9700 AV Groningen, The Netherlands. We wish to thank the editor and the referees for their useful comments.

1 Introduction

HOLLY and MAGNUS [1988] (hereafter denoted by HM) consider instrumental variables (IV) estimation of the model

(1) $$y = Y\alpha_0 + X_1\beta_0 + u = Z\delta_0 + u$$

where y $(T \times 1)$ is the dependent variable, $Y (T \times N)$ is a matrix of endogenous variables, X_1 $(K_1 \times N)$ a matrix of exogenous variables, and u $(T \times 1)$ a disturbance vector orthogonal to X_1 but not to Y. The existence is assumed of K_2 instrumental variables grouped in the $T \times K_2$-matrix X_2, with the property that the matrix $X'Z/T$ $(K \times (N + K_1))$ has a probability limit of rank $N + K_1$, where $X \equiv (X_1, X_2)$ and $K \equiv K_1 + K_2$. (We use the notation of HM as far as possible.)

HM investigate several issues that arise when "completing" (1) with the N equations

(2) $$Y = X_1\Pi_{10} + X_2\Pi_{20} + V = X\Pi_0 + V$$

where V $(T \times N)$ is a disturbance matrix orthogonal to X and Π_{10} $(K_1 \times N)$ and Π_{20} $(K_2 \times N)$ are coefficient matrices with rank $(\Pi_{20}) = N$. Evidently, (1) and (2) together form a complete system of simultaneous equations. The main result in HM is that the limiting distribution of the IV estimator of δ_0 in (1) is the same as that of its FIML estimator in (1) and (2), with normality and the mutual independence of the rows of (u, V) and X as the working hypothesis. To prove this, HM derive the asymptotic information matrix for the complete model and invert it to obtain the variance of the limiting distribution of the FIML estimator of all model parameters, including that of δ_0. The latter appears to be equal to the result from IV. The HM approach is quite long, and in view of the intuitive plausibility of the result one is tempted to raise the question whether a shorter route is available. The answer is affirmative if the general result of the asymptotic equivalence between FIML and 3SLS is used. Of course, the ancillary results that are spawned by the HM approach may be of independent interest, providing it with a justification. On the other hand, we will prove a stronger result, viz. that the difference between the estimators is $o_p(T^{-1/2})$, in other words, they are asymptotically equivalent. Moreover, this result does not depend on normality.

2 The Proof

Our argument is as follows. First, as is well known, FIML and 3SLS are asymptotically equivalent under very general conditions not requiring normality. For this result, see e.g. ROTHENBERG and LEENDERS [1964], MALINVAUD [1970], BROWNE [1974], HENDRY [1976]; it is already implicit in FERGUSON [1958]. Second, applying the IV method to (1) amounts to applying 2SLS. Consequently we are done when we have shown that 2SLS and 3SLS give identical estimates for δ_0. That this indeed is the case follows from result (3) in section 4.2 of ZELLNER and THEIL [1962]. For the sake of clarity we supply an independent short proof for the special case we are concerned with here.

Let $\hat{\Psi}$ $((N+1)\times(N+1))$ be the positive definite 2SLS estimate of the variance matrix of the disturbance vector (u_t, v_t') and partition it conformably as

$$\hat{\Psi}=\begin{pmatrix} \hat{\sigma}^2 & \hat{\theta}' \\ \hat{\theta} & \hat{\Omega} \end{pmatrix}. \tag{3}$$

Let $P\equiv X(X'X)^{-1}X'$ be the projector onto the X-space. The 3SLS estimator of δ_0 and Π_0 follows from minimizing

$$q\equiv\begin{bmatrix} y-Z\delta \\ \text{vec}\,(Y-X\Pi) \end{bmatrix}'(\hat{\Psi}^{-1}\otimes P)\begin{bmatrix} y-Z\delta \\ \text{vec}\,(Y-X\Pi) \end{bmatrix} \tag{4}$$

with respect to δ and Π. This criterium can be elaborated by using the formula for the inverse of a partitioned matrix, which yields

$$\hat{\Psi}^{-1}=\begin{bmatrix} \hat{\sigma}^{-2} & 0 \\ 0 & 0 \end{bmatrix}+\begin{bmatrix} \hat{\sigma}^{-2}\hat{\theta}' \\ -I_N \end{bmatrix} S\,(\hat{\sigma}^{-2}\hat{\theta},\ -I_N), \tag{5}$$

where S $(N\times N)$ is some positive definite matrix, whose particular form is immaterial. Using (5), q becomes

$$q=\hat{\sigma}^{-2}(y-Z\delta)'\,P\,(y-Z\delta)+r'(S\otimes P)\,r \tag{6}$$

with

$$r\equiv\hat{\sigma}^{-2}\hat{\theta}\otimes P\,(y-Z\delta)-\text{vec}\,(Y-X\Pi). \tag{7}$$

We proceed in two stages. We first fix δ and minimize q in (6) with respect to Π, and then minimize the concentrated criterium with respect to δ. Now, Π only occurs in the second term of (6). Since

$$r'(S\otimes P)\,r=[(I_N\otimes X^+)\,r]'\,(S\otimes X'X)\,[(I_N\otimes X^+)\,r] \tag{8}$$

[with $X^+\equiv(X'X)^{-1}X'$] and

$$(I_N\otimes X^+)\,r=\hat{\sigma}^{-2}\hat{\theta}\otimes X^+(y-Z\delta)-\text{vec}\,(X^+Y-\Pi), \tag{9}$$

the second term of (6) equals 0 for

$$\hat{\Pi}_0 = X^+ (Y - \hat{\sigma}^{-2} (y - Z\delta)\,\hat{\theta}'). \tag{10}$$

Hence the second stage simply is to minimize the first term of (6). This is exactly the criterium that leads to the 2 SLS estimator of δ.

3 Conclusion

Showing the truth of the HM result becomes quite simple when the equality of 2 SLS and 3 SLS of δ_0 is established and if the asymptotic equivalence between FIML and 3 SLS is used. Note that we are not in a situation where 2 SLS and 3 SLS give identical results for *all* coefficients. Such situations have been characterized by Kapteyn and Fiebig [1981], elaborating for the simultaneous equations system the well-known result about the condition n where OLS and GLS coincide. That the situation here is different is also apparent from (10), where $\hat{\Pi}$ differs from the 2 SLS estimator of Π_0, which is $X^+ Y$. The discrepancy is a rank one matrix that is linear in $\hat{\sigma}^{-2}$, in $\hat{\theta}$, and in the "residual" $y - Z\delta$, and vanishes whenever these do. Further, note that Ω plays no role in the 3 SLS estimation of Π_0. In fact, minimization of (4) with Ψ replaced by any conformable positive definite matrix leads to the same estimate for δ_0. This reflects the fact that in the completing part (2) of the model the same regressors enter into every equation or, stated differently, that no restrictions are imposed Π_0. As is apparent from (9) and (10), whenever Π_0 is restricted there is no way to find a $\hat{\Pi}$ such that r vanishes.

• References

Browne, M. W. (1974). – "Generalized Least-Squares Estimators in the Analysis of Covariance Structures", *South African Statistical Journal*, 8, pp. 1-24.

Ferguson, T. S. (1958). – "A Method for Generating Best Asymptotically Normal Estimators with Application to the Estimation of Bacterial Densities", *Annals of Mathematical Statistics*, 29, pp. 1046-1062.

Hendry, D. F. (1976). – "The Structure of Simultaneous Equations Estimators", *Journal of Econometrics*, 4, pp. 87-98.

Holly, A. and Magnus, J. R. (1988). – "A Note on Instrumental Variables and Maximum Likelihood Estimation Procedures, *Annales d'Économie et de Statistique*, 10, pp. 121-138.

Kapteyn, A. and Fiebig, D. (1981). – "When are Two-Stage and Three-Stage Least Squares Estimators Identical?", *Economics Letters*, 8, pp. 53-57.

Malinvaud, E. (1970). – *Statistical Methods of Econometrices*, North-Holland, Amsterdam.

Rothenberg, T. J. and Leenders, C. T. (1964). – "Efficient Estimation of Simultaneous Equations Systems", *Econometrica*, 32, pp. 57-76.

Zellner, A. and Theil, H. (1962). – "Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations, *Econometrica* 30, pp. 54-78.