Economics Letters 8 (1981) 53–57
North-Holland Publishing Company 

# WHEN ARE TWO-STAGE AND THREE-STAGE LEAST SQUARES ESTIMATORS IDENTICAL?

Arie KAPTEYN and Denzil G. FIEBIG

*University of Southern California, Los Angeles, CA 90007, USA*

Received 27 July 1981

Necessary and sufficient conditions are derived for the numerical equivalence of the two-stage and three-stage least squares estimators in a linear simultaneous equations model. The conditions are easy to verify in any practical application.

## 1. Introduction

In a recent note, Srivastava and Tiwari (1978) generalize conditions given by Zellner and Theil (1962), under which two-stage least squares (2SLS) and three-stage least squares (3SLS) give identical estimates of the structural coefficients in a linear simultaneous equation model. In their derivation, Srivastava and Tiwari employ a necessary and sufficient condition for ordinary least squares (OLS) and generalized least squares (GLS) to be identical in a linear regression model, derived by Rao (1968).

Rao's condition is somewhat difficult to apply. Consequently Srivastava and Tiwari derive only sufficient conditions for 2SLS and 3SLS to be identical. It does not become clear from their analysis if it is possible to generalize their conditions any further, nor is the application of their conditions in practice all that obvious, although they provide a number of important special cases.

In this note we use a condition derived by Kruskal (1968) to provide necessary and sufficient conditions for the equivalence of 2SLS and 3SLS. The conditions are easy to check in practice.

Our conditions also follow from the analysis by Gourieroux and Monfort (1980). They do not state the conditions explicitly, however. Since, moreover, their article is rather technical and hence less accessible to applied econometricians, it appears worthwhile to provide an explicit statement of the conditions.

## 2. The model

To facilitate comparison, we follow Srivastava and Tiwari's notation. Thus, we consider a linear system of $M$ structural equations with $M$ jointly dependent and $\Lambda$ predetermined variables:

$$\begin{aligned} y_i &= Y_i\gamma_i + X_i\beta_i + u_i, \qquad i = 1, 2, \ldots, M, \\ &= Z_i\delta_i + u_i, \qquad Z_i = (Y_i X_i), \qquad \delta_i = \begin{pmatrix} \gamma_i \\ \beta_i \end{pmatrix}, \end{aligned} \tag{1}$$

where the $T$-vector $y_i$ contains the observations on the $i$th dependent variable to be explained by the $i$th structural equation, $Y_i$ $(T \times m_i, m_i < M)$ contains observations on jointly dependent variables included as explanatory variables in the $i$th equation, $X_i(T \times l_i, l_i \leqslant \Lambda)$ is the matrix of predetermined variables included in the $i$th equation, $\gamma_i$ and $\beta_i$ are corresponding vectors of unknown parameters, $u_i$ is a $T$-vector of disturbances satisfying

$$\begin{aligned} &E(u_i) = 0, \\ &E(u_i u_j') = \sigma_{ij} I_T, \end{aligned} \qquad i, j = 1, 2, \ldots, M. \tag{2}$$

The distribution of the disturbances is supposed to be independent of the predetermined variables in the system, the reduced form is assumed to exist and the equations are either just identified or over identified.

All equations taken together can be written as

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} Z_1 & 0 & \ldots & 0 \\ 0 & Z_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & Z_M \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_M \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_M \end{bmatrix}, \quad \text{or}$$

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

Let $X(T \times \Lambda)$ be the matrix of all predetermined variables. Premultiplying (1) by $X'$ yields

$$X'y_i = X'Z_i\delta_i + X'u_i, \qquad i = 1, \ldots, M. \tag{4}$$

The error variance–covariance matrix of this equation is $\sigma_{ii}\ X'X$. It is well-known that GLS estimation of $\delta_i$ in this equation yields the 2SLS estimator for $\delta_i$ [e.g., Theil (1971, sect. 5.9)]. Let $R$ be a non-singular $\Lambda \times \Lambda$-matrix such that $X'X = RR'$. Then GLS in (4) is identical to OLS in

$$R^{-1}X'y_i = R^{-1}X'Z_i\,\delta_i + R^{-1}X'u_i, \qquad i = 1,\dots,M. \tag{5}$$

Stacking these $M$ equations in one system we get

$$\left(I_M \otimes R^{-1}X'\right)y = \left(I_M \otimes R^{-1}X'\right)Z\,\delta + \left(I_M \otimes R^{-1}X'\right)u \equiv A\,\delta + w, \tag{6}$$

with $A$ and $w$ implicitly defined. The variance–covariance matrix of the error vector $w$ is $\Sigma \otimes I_\Lambda$. It is easy to check that GLS estimation in (6) yields the 3SLS estimator of $\delta$, $\delta_{3\mathrm{SLS}}$ say. Obviously, OLS in (6) would yield the 2SLS estimator of $\delta$, $\delta_{2\mathrm{SLS}}$ say.

Starting from this fact, Srivastava and Tiwari exploit a result from regression theory to investigate under what conditions 2SLS and 3SLS give identical results. Adapting a result by Rao (1968) to the present situation we have that $\hat{\delta}_{2\mathrm{SLS}}$ and $\hat{\delta}_{3\mathrm{SLS}}$ are identical if and only if there exists a $M\Lambda \times (M\Lambda - \Sigma_{i=1}^{M}\ (l_i + m_i))$ – matrix $B$ with column rank, orthogonal to $A$ such that

$$A'(\Sigma \otimes I_\Lambda)B = 0. \tag{7}$$

To simplify, Srivastava and Tiwari postulate $B$ to be block-diagonal and derive sufficient conditions for the equivalence of $\hat{\delta}_{2\mathrm{SLS}}$ and $\hat{\delta}_{3\mathrm{SLS}}$. They observe that more general conditions can be obtained if $B$ is not a priori taken to be block-diagonal, but they conjecture that the resulting conditions will be difficult to apply since they would depend on the unknown $\sigma_{ij}$.

We will see that it is possible to obtain conditions that are both general and simple, by employing a result due to Kruskal (1968). Kruskal's result states that OLS = GLS (so $\hat{\delta}_{2\mathrm{SLS}} = \hat{\delta}_{3\mathrm{SLS}}$) if there exists some matrix $C$ such that

$$(\Sigma \otimes I_\Lambda)A = AC. \tag{8}$$

Obviously, $C$ has to be a square matrix of order $\Sigma_{i=1}^{M}(l_i + m_i)$. It is easy to prove the equivalence of (7) and (8), cf. Rao and Mitra (1971, sect. 8.2).

## 3. When are 2SLS and 3SLS identical?

Define $A_i \equiv R^{-1}X'Z_i(\Lambda \times (l_i + m_i))$. Then (8) can be written as

$$\begin{bmatrix} \sigma_{11}A_1 & \sigma_{12}A_2 & \cdots & \sigma_{1M}A_M \\ \vdots & & & \vdots \\ \sigma_{M1}A_1 & \sigma_{M2}A_2 & \cdots & \sigma_{MM}A_M \end{bmatrix} = \begin{bmatrix} A_1C_{11} & A_1C_{12} & \cdots & A_1C_{1M} \\ \vdots & & & \\ A_MC_{M1} & A_MC_{M2} & \cdots & A_MC_{MM} \end{bmatrix}, \tag{9}$$

where the typical block $C_{ij}$ of $C$ is of order $(l_i + m_i) \times (l_j + m_j)(i, j = 1,\ldots, M)$ consequently, (8) holds if and only if

$$\sigma_{ii}A_i = A_iC_{ii}, \qquad i = 1,\ldots,M, \quad \text{and} \tag{10}$$

$$\sigma_{ij}A_j = A_iC_{ij}, \qquad i \neq j, \quad i, j = 1,\ldots, M. \tag{11}$$

Equality (10) can easily be satisfied by choosing $C_{ii}$ to be a scalar matrix, i.e., $C_{ii} = \sigma_{ii}I_{(l_i + m_i)}$.

To study the implications of equality (11), we have to distinguish two cases. First, if $\sigma_{ij} = 0$, $C_{ij} = 0$ satisfies the equation. Second, if $\sigma_{ij} \neq 0$ we can derive a number of restrictions on $C_{ij}$. Because of the symmetry of $\Sigma$ there has to hold simultaneously:

$$\sigma_{ij}A_j = A_iC_{ij}, \qquad \sigma_{ij}A_i = A_jC_{ji}. \tag{12}$$

Since it has been assumed that all structural parameters are identified, $A_i$ and $A_j$ have to be of full column rank. From (12) it follows, moreover, that $A_i$ and $A_j$ have to have equal rank, so $C_{ij}$ and $C_{ji}$ are square and non-singular. To be precise:

$$C_{ji} = \sigma_{ij}^2 C_{ij}^{-1}. \tag{13}$$

In view of the definition of $A_i$, we have that

$$\sigma_{ij}R^{-1}X'Z_i = R^{-1}X'Z_jC_{ji}, \quad \text{hence} \tag{14}$$

$$X'Z_i = \sigma_{ij}^{-1} X'Z_j C_{ji}, \quad \text{and} \tag{15}$$

$$X'Z_j = \sigma_{ij}^{-1} X'Z_i C_{ij}, \tag{16}$$

so $X'Z_i$ has to lie in the column space of $X'Z_j$ and vice versa.

These results can be summarized in the following:

*Proposition.* *2SLS and 3SLS give numerically identical results if and only if for all equations $i, j$ with non-zero error covariance $X'Z_i$ and $X'Z_j$ span the same column space.*

A number of special cases are readily derived from this. If the errors of all equations are uncorrelated, 2SLS is identical to 3SLS. If all equations are just identified, and hence $X'Z_i$ is square and non-singular for each $i$ (so that the $X'Z_i$ span the $\Lambda$-dimensional space for all $i$), 2SLS equals 3SLS. A third example would be that the first $p$ equations are just identified with freely covarying errors, whereas the last $M-p$ equations may be overidentified but their errors would have to be mutually uncorrelated and uncorrelated with the errors of the first $p$ equations.

As noticed by Srivastava and Tiwari, a seemingly unrelated regression system is a special case of a simultaneous system, with only purely exogenous variables as explanatory variables. For that special case the Proposition implies that GLS will be identical to OLS applied equation by equation if and only if for all equations $i$, $j$ with non-zero error covariance the matrices of explanatory variables $X_i$ and $X_j$ span the same column space. This was also noted by Gourieroux and Monfort (1980). Special cases are that all errors are uncorrelated or that all equations have the same matrix of explanatory variables.

## References

Gourieroux, C. and A. Monfort, 1980, Sufficient linear structures: Econometric Applications, Econometrica 48, 1083–1097.

Kruskal, W., 1968, When are Gauss—Markoff and least squares estimates identical? A coordinate-free approach, Annals of Mathematical Statistics 39, 70–75.

Rao, C.R., 1968, A note on previous lemma in the theory of least squares and some future results, Sankhya A 30, 259–266.

Rao, C.R. and S.K. Mitra, 1971, Generalized inverses of matrices and its applications (Wiley, New York).

Srivastava, V.K. and R. Tiwari, 1978, Efficiency of two-stage and three-stage least squares estimators, Econometrica 46, 1495–1498.

Theil, H., 1971, Principles of econometrics (Wiley, New York).

Zellner, A. and H. Theil, 1962, Three-stage least squares: Simultaneous estimation of simultaneous equations, Econometrica 30, 54–78.