Reprinted from Econometrics and Economic Theory
Essays in Honour of Jan Tinbergen
Edited by Willy Sellekaerts

4 Specification Bias in Seemingly Unrelated Regressions

BY POTLURI RAO*

MULTIPLE regression analysis specifies a linear relation between a dependent variable and a set of independent variables. When the independent variables are non-stochastic, and the error terms are homoscedastic and serially independent, the ordinary least squares estimation of the parameters yields the best linear unbiased estimates. But when there is a set of linear regression equations whose error terms are contemporaneously correlated, then the ordinary least squares estimation of each of the equations separately is not the 'best' estimation procedure. When the parameters of contemporaneous correlation are known then it is possible to obtain unbiased estimates with smaller variance than the corresponding ordinary least squares estimates by estimating all the regression equations jointly using the Aitken's generalised least squares.[1] In the absence of information on these parameters Professor Zellner [5] suggested the use of estimates of these parameters from residuals of the ordinary least squares. This procedure is called the 'seemingly unrelated regression equations' (SURE) procedure.

The theoretical properties of the SURE estimation procedure have been investigated extensively under the assumption that the estimations are the true relations.[2] In econometric research, however, one can never be certain that the estimated equations are the true specifications. Often we leave out some variables

* This paper was prepared when the author was visiting the University of Washington, Seattle. The author is grateful to Professors Zvi Griliches, Jan Kmenta and Lester Telser for their helpful comments.

[1] For a discussion on the Aitken's generalised least squares procedure in the context of a set of linear regression equations, see [4].

[2] A brief survey of the literature may be found in [1].

in estimation either because we are not aware of their relevance in the true model, or that data are not available on some variables. In such situations we may want to choose an estimator that is less sensitive to mis-specification even though it may not be the 'best' estimator under the ideal set of assumptions. With this objective in mind we investigate the consequences of left out variables on the ordinary least squares and SURE estimation procedures. The consequences of mis-specification on the ordinary least squares estimates were analysed in detail elsewhere [2]. In this paper we concentrate mainly on the consequences of a left-out variable in the SURE estimation procedure and compare these results with that of the ordinary least squares.

The Model

For our analysis we shall concentrate on a simple case with two linear regression equations. Let the true model be

$$\begin{aligned} x_{4t} &= \alpha_1 x_{1t} + x_{3t} + \eta_{1t} \\ x_{5t} &= \alpha_2 x_{2t} + \beta x_{3t} + \eta_{2t} \end{aligned} \qquad (t = 1, \ldots, T) \qquad (1)$$

In this model all the variables are measured from their respective means, hence the constant terms are implicit. The independent variables x_1, x_2 and x_3 are assumed to be non-stochastic; that is, held constant in repeated samples. Since we are interested in isolating the consequences of mis-specification we shall assume that in the true model, error terms (ηs) are contemporaneously uncorrelated. As in the classical linear regression model we shall assume that the error terms (ηs) are homoscedastic and serially independent. The independent variable x_3 is measured in units such that its coefficient in the first equation of (1) is unity. This simplifies discussion without loss of generality. Whenever there is no ambiguity we shall delete the subscript t and use $\mathbf{x}_i$ to denote a vector of T observations on the variable x_i.

Let us suppose that instead of the true model (1) the following mis-specified model is used for estimation

$$\begin{aligned} \mathbf{x}_4 &= \alpha_1 \mathbf{x}_1 + \mathbf{u}_1 \\ \mathbf{x}_5 &= \alpha_2 \mathbf{x}_2 + \mathbf{u} \end{aligned} \qquad (2)$$

The estimated model may also be written in matrix notation as

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\alpha} + \mathbf{U} \qquad (3)$$

where

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{x}_2 \end{bmatrix}$$

$$\mathbf{Y} = \begin{bmatrix} \mathbf{x}_4 \\ \mathbf{x}_5 \end{bmatrix}$$

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$$

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix}$$

Even though the true error terms (ηs) are contemporaneously uncorrelated, error terms in the estimated model (us) are because the same variable is left out in both equations. This kind of problem is common in empirical research. For example, in an investment study, error terms in the investment functions of the General Electric and Westinghouse [5] would be contemporaneously correlated when a common variable, namely the business fluctuation, is left out from both the equations.

The variance-covariance matrix of error terms in the estimated model is

$$E(\mathbf{UU}') = \boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2_{x_3} + \sigma^2_{\eta_1} & \beta\sigma^2_{x_3} \\ \beta\sigma^2_{x_3} & \beta^2\sigma^2_{x_3} + \sigma^2_{\eta_2} \end{bmatrix} \otimes \mathbf{I}_T \qquad (4)$$

where $\mathbf{I}_T$ is an identity matrix of order T.

Ordinary least squares estimation of the equations of model (2) separately yields the following estimates for the αs

$$\begin{aligned} \hat{\alpha}_1 &= (\mathbf{x}_1'\mathbf{x}_1)^{-1}\mathbf{x}_1'\mathbf{x}_4 \\ \hat{\alpha}_2 &= (\mathbf{x}_2'\mathbf{x}_2)^{-1}\mathbf{x}_2'\mathbf{x}_5 \end{aligned} \qquad (5)$$

These estimates may also be expressed in matrix form as

$$\hat{\boldsymbol{\alpha}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \qquad (6)$$

When the variance-covariance matrix of error terms in the estimated model ($\boldsymbol{\Sigma}$) is known, then the generalised least squares estimates of the αs may be obtained as

$$\tilde{\alpha} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y \tag{7}$$

The generalised least squares estimates are not attainable because the matrix Σ is a function of the unknown parameters. However, this variance-covariance matrix may be estimated from residuals of the ordinary least squares defined as

$$\begin{aligned} e_1 &= x_4 - \hat{\alpha}_1 x_1 \\ e_2 &= x_5 - \hat{\alpha}_2 x_2 \end{aligned} \tag{8}$$

An estimate of the variance-covariance matrix is

$$\hat{\Sigma} = T^{-1}\begin{bmatrix} e_1'e_1 & e_1'e_2 \\ e_2'e_1 & e_2'e_2 \end{bmatrix} \otimes I_T \tag{9}$$

SURE estimates of the parameters α may be obtained by using the estimate $\hat{\Sigma}$ instead of the true matrix Σ as

$$\alpha^* = (X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}Y \tag{10}$$

Now we have two sets of estimates for the αs, the ordinary least squares estimates $\hat{\alpha}$ and the SURE estimates α^*. Both these sets are obtained from the mis-specified model (2) and hence may have a certain amount of 'sin'. The question is now: Which of these two estimation procedures has the least sin? *If* we knew that the set of equations of model (1) were the true relations, and *If* we had data for the variable x_3, then of course, there would be no need to resort to these sinful estimation procedures. But these two *Ifs* have a capital I. We shall measure the extent of sin of these estimation procedures in the units of bias and variance of the estimators. In our study we shall concentrate on variance and asymptotic bias of these two alternative estimation procedures and see how they compare.

In order to simplify the algebra involved we shall adopt the following notation

$$\begin{aligned} x_{ij} &= x_i'x_j \\ e_{ij} &= e_i'e_j \\ b_{ij} &= (x_j'x_j)^{-1}x_j'x_i \end{aligned} \tag{11}$$

These *b*s are used only as a shorthand notation and need not have any causal interpretation. Our analysis also needs a statistic γ defined as

$$\gamma = (e_{22})^{-1}e_{21} \tag{12}$$

By substituting equations (1), (5) and (8) in equation (12) it may be seen that

$$\text{plim}(\gamma) = \gamma^* = \frac{\beta\sigma^2_{x_3}}{\beta^2\sigma^2_{x_3} + \beta^2 b_{32}^2 + \sigma^2_{\eta_2}} \tag{13}$$

In our analysis we do not use this complete expression for γ^*. We only use the property that γ^* and β have the same sign.

The Nature of Asymptotic Bias

The ordinary least squares estimates of the parameters α are given by Equation (5). Since the true relations are given by Equation (1), by substituting for the variables x_4 and x_5 we obtain

$$\begin{aligned} \hat{\alpha}_1 &= b_{41} = \alpha_1 + b_{31} + (x_1'x_1)^{-1}x_1'\eta_1 \\ \hat{\alpha}_2 &= b_{52} = \alpha_2 + \beta b_{32} + (x_2'x_2)^{-1}x_2'\eta_2 \end{aligned} \tag{14}$$

Since the results of this investigation are symmetric we shall concentrate on only one of the estimates, namely the $\hat{\alpha}_1$. The expected value of the $\hat{\alpha}_1$ is

$$E(\hat{\alpha}_1) = \alpha_1 + b_{31} \tag{15}$$

The ordinary least squares estimate of the parameter α_1 is a biased estimate, unless the left-out variable x_3 is orthogonal to the included variable x_1.

SURE estimates of the αs are obtained as

$$\alpha^* = (X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}Y \tag{16}$$

where $\hat{\Sigma}$ is an estimate of the variance-covariance matrix Σ obtained from the residuals as given in Equation (9). The expressions in Equation (16) may be evaluated as follows[3]

[3] In the following section we take this direct approach to evaluate these expressions because the compact matrix notation simplifies the algebra at the expense of insight into the problem.

$$\hat{\Sigma} = T^{-1}\begin{bmatrix} e_{11} & e_{12} \\ e_{12} & e_{22} \end{bmatrix} \otimes \mathbf{I}_T \qquad (17)$$

$$\hat{\Sigma}^{-1} = \frac{T}{e_{11}e_{22} - e_{12}e_{12}}\begin{bmatrix} e_{22} & -e_{12} \\ -e_{12} & e_{11} \end{bmatrix} \otimes \mathbf{I}_T \qquad (18)$$

$$\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X} = \begin{bmatrix} \mathbf{x}_1' & 0 \\ 0 & \mathbf{x}_2' \end{bmatrix}\hat{\Sigma}^{-1}\begin{bmatrix} \mathbf{x}_1 & 0 \\ 0 & \mathbf{x}_2 \end{bmatrix}$$

$$= \frac{T}{e_{11}e_{22} - e_{12}e_{12}}\begin{bmatrix} e_{22}x_{11} & -e_{12}x_{12} \\ -e_{12}x_{12} & e_{11}x_{22} \end{bmatrix} \qquad (19)$$

$$(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1} = \frac{e_{11}e_{22} - e_{12}e_{12}}{T(e_{11}e_{22}x_{11}x_{22} - e_{12}e_{12}x_{12}x_{12})}\begin{bmatrix} e_{11}x_{22} & e_{12}x_{12} \\ e_{12}x_{12} & e_{22}x_{11} \end{bmatrix} \qquad (20)$$

$$\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{Y} = \begin{bmatrix} \mathbf{x}_1' & 0 \\ 0 & \mathbf{x}_2' \end{bmatrix}\hat{\Sigma}^{-1}\begin{bmatrix} \mathbf{x}_4 \\ \mathbf{x}_5 \end{bmatrix}$$

$$= \frac{T}{e_{11}e_{22} - e_{12}e_{12}}\begin{bmatrix} e_{22}x_{14} - e_{12}x_{15} \\ e_{11}x_{25} - e_{12}x_{24} \end{bmatrix} \qquad (21)$$

Hence

$$\alpha_1^* = \frac{e_{11}x_{22}e_{22}x_{14} - e_{11}x_{22}e_{12}x_{15} + e_{12}x_{12}e_{11}x_{25} - e_{12}e_{12}x_{12}x_{24}}{e_{11}e_{22}x_{11}x_{22} - e_{12}e_{12}x_{12}x_{12}} \qquad (22)$$

By taking out the expression $(e_{11}e_{22}x_{11}x_{22})$ as a common factor we obtain

$$\alpha_1^* = \frac{b_{41} - \gamma b_{51} + \gamma b_{21}b_{52} - r^2_{e_1e_2}b_{21}b_{42}}{1 - r^2_{x_1x_2}r^2_{e_1e_2}} \qquad (23)$$

where r is the correlation coefficient.

Since the true model is given by Equation (1), by substituting for the variables x_4 and x_5 in Equation (23) we obtain

$$\alpha_1^* = \alpha_1 + \frac{b_{31} - \gamma\beta(b_{32} - b_{31}) - r^2_{e_1e_2}b_{21}b_{32}}{1 - r^2_{x_1x_2}r^2_{e_1e_2}} + f(X'\eta) \qquad (24)$$

where $f(X'\eta)$ stands for the terms containing $X'\eta$, and $\text{plim}\,[f(X'\eta)] = 0$.

For large samples we obtain

$$\text{plim}\,(\alpha_1^*) = \alpha_1 + \frac{b_{31} - \gamma^*\beta(b_{32} - b_{31}) - P^2b_{21}b_{32}}{1 - P^2r^2_{x_1x_2}} \qquad (25)$$

where $P^2 = \text{plim}\,(r^2_{e_1e_2})$.

In the case of mis-specification in the estimated model, the SURE estimates are also biased estimates, unless the left out variable is orthogonal to all the independent variables of all the equations in the model.

We have shown that when estimated equations have left out variables specified by the true relation, both the estimation procedures, the ordinary least squares and SURE, in general yield biased estimates of the parameters of the model. In order to compare which of these two estimates has larger bias let us consider the following cases.

Case 1

The left-out variable $\mathbf{x}_3$ is orthogonal to the independent variables in both equations. In this case b_{31} and b_{32} are zero. Therefore, both estimates, $\hat{\alpha}_1$ and α_1^* are unbiased. Note that in our analysis we are investigating bias in only one of the estimates, namely $\hat{\alpha}_1$. Similar results follow for the other estimate, except that the ordering of these cases with respect to it would be different.

$$E(\hat{\alpha}_1) = \alpha_1$$
$$\text{plim}\,(\alpha_1^*) = \alpha_1 \qquad (26)$$

Case 2

The left-out variable $\mathbf{x}_3$ is orthogonal to the independent variable $\mathbf{x}_1$ of the first equation (the equation with the parameter we are interested in), but not with that of the second equation. In this case b_{31} is zero, but not b_{32}. Hence the SURE estimate is biased whereas the ordinary least squares estimate is not.

$$E(\hat{\alpha}_1) = \alpha_1$$

$$\text{plim}\,(\alpha_1^*) = \alpha_1 - \frac{\gamma^*\beta b_{32} + P^2b_{21}b_{32}}{1 - P^2r^2_{x_1x_2}} \qquad (27)$$

Case 3

The left-out variable $\mathbf{x}_3$ is orthogonal to the independent variable of the second equation, $\mathbf{x}_2$, but not to that of the first equation. In this case both estimates are biased, but the SURE estimate has larger bias than the ordinary least squares estimate.

$$E(\hat{\alpha}_1) = \alpha_1 + b_{31}$$

$$\text{plim}\,(\alpha_1^*) = \alpha_1 + b_{31}\frac{1+\gamma^*\beta}{1-P^2r^2_{x_1x_3}} \tag{28}$$

As we have already shown in Equation (13), γ^* and β are of the same sign. Since the correlation coefficients are fractions, it follows that the bias in the SURE estimate α_1^* is larger in magnitude than the bias in the ordinary least squares estimate $\hat{\alpha}_1$.

Case 4

The left-out variable x_3 is not orthogonal to either of the independent variables x_1 or x_2. In this case both estimates are biased. It is not obvious which one has the smaller bias. When b_{31} and b_{32} are of the same order of magnitude, the bias in the ordinary least squares and SURE procedures are approximately equal. When x_1 and x_2 are the same variable these estimates, α_1^* and $\hat{\alpha}_1$, are identical[4] and hence have the same bias. Bias in the SURE procedure depends on the magnitude of b_{31} relative to b_{32}. One should keep in mind that we are also estimating the parameter α_2 in the SURE procedure. If one were to choose values of b_{31} and b_{32} so as to reduce bias in the estimate α_1^* one would be increasing bias in the estimate α_2^*.

Our analysis leads to the conclusion that when the estimated model is mis-specified both the ordinary least squares and SURE estimates are asymptotically biased and the bias of the SURE estimator is generally larger than the bias of the ordinary least squares estimator. Asymptotic bias of these estimators is a consequence of systematic relation between the left-out variable and independent variables of the estimated model. In many practical situations, however, it may be reasonable to believe that the left-out variable is orthogonal to the included variables (or nearly so), in which case asymptotic bias of these estimators is zero (or nearly zero). In this case the choice of estimation procedure depends crucially on the relative variances.

[4] This point may be seen by substituting x_1 for x_2 in equation (23).

The Relative Efficiency

In order to isolate the consequences of mis-specification on the relative efficiency of the estimators we shall consider the case where both estimators are asymptotically unbiased, that is the left-out variable x_3 is orthogonal to the variables x_1 and x_2. Note that in our model the independent variables are non-stochastic and only the error terms (ηs) change in each of the repeated samples.

As shown in Equation (14) the ordinary least squares estimate $\hat{\alpha}_1$ may be expressed as

$$\hat{\alpha}_1 = \alpha_1 + b_{31} + (x_1'x_1)^{-1}x_1'\eta_1 \tag{29}$$

Since x_3 is orthogonal to x_1 we have

$$\hat{\alpha}_1 = \alpha_1 + (x_1'x_1)^{-1}x_1'\eta_1 \tag{30}$$

Hence the variance of the estimate $\hat{\alpha}_1$ is[5]

$$V(\hat{\alpha}_1) = E[(x_1'x_1)^{-1}x_1'\eta_1\eta_1'x_1(x_1'x_1)^{-1}] \tag{31}$$

Since the xs are non-stochastic the variance of the estimate $\hat{\alpha}_1$ reduces to

$$V(\hat{\alpha}_1) = (x_1'x_1)^{-1}\sigma^2_{\eta_1} \tag{32}$$

The SURE estimates of α may be expressed as

$$\begin{aligned}\alpha^* &= (X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}Y\\ &= \alpha + (X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}\eta\end{aligned} \tag{33}$$

because x_3 is orthogonal to x_1 and x_2 by assumption.

The variance-covariance matrix of the SURE estimates is

$$\begin{aligned}V(\alpha^*/\hat{\Sigma}) &= E[(X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}\eta\eta'\hat{\Sigma}^{-1}X(X'\hat{\Sigma}^{-1}X)^{-1}]\\ &= (X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}E(\eta\eta')\hat{\Sigma}^{-1}X(X'\hat{\Sigma}^{-1}X)^{-1}\end{aligned} \tag{34}$$

To be able to compare the variance of the SURE estimate with that of the ordinary least squares estimate let us express the variance of the estimates of the αs from the ordinary least

[5] The estimate $\hat{\alpha}_1$ has the same variance, even when x_3 is not orthogonal to x_1.

squares as

$$V(\hat{\alpha}) = \begin{bmatrix} x_{11} & 0 \\ 0 & x_{22} \end{bmatrix}^{-1} \begin{bmatrix} \sigma^2_{\eta_1} & 0 \\ 0 & \sigma^2_{\eta_2} \end{bmatrix} \otimes \mathbf{I}$$

$$= (\mathbf{X}'\mathbf{X})^{-1}E(\eta\eta')$$

$$= (\mathbf{X}'[E(\eta\eta')]^{-1}\mathbf{X})^{-1} \qquad (35)$$

Now let us consider a hypothetical seemingly unrelated regression equations model with no mis-specification in the estimated model. Let the true variance-covariance matrix of the errors be φ. Instead of φ another matrix $\hat{\Sigma}(\hat{\Sigma} \neq \varphi)$ is used in its place in computing the seemingly unrelated estimates. The variance matrix of the resulting estimates is

$$\mathbf{V} = (\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\Sigma}^{-1}\varphi\hat{\Sigma}^{-1}\mathbf{X}(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1} \qquad (36)$$

The variance matrix of the generalised least squares estimates using the true variance matrix φ is

$$\mathbf{V}^* = (\mathbf{X}'\varphi^{-1}\mathbf{X})^{-1} \qquad (37)$$

But the generalised least squares estimates are the minimum variance unbiased estimates. Therefore it follows that

$$(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\Sigma}^{-1}\varphi\hat{\Sigma}^{-1}\mathbf{X}(\mathbf{X}'\hat{\Sigma}^{-1})^{-1} \geq (\mathbf{X}'\varphi^{-1}\mathbf{X})^{-1} \qquad (38)$$

This inequality becomes an equality only when $\hat{\Sigma} = \varphi$.

Since the inequality (38) is true for all φ, it is true for $E(\eta\eta')$ as well. Hence by substituting $E(\eta\eta')$ for φ in the inequality (38) *and noting that the inequality is valid for all* $\hat{\Sigma}$ we obtain[6]

$$V(\alpha^*) \geq V(\hat{\alpha}) \qquad (39)$$

The SURE estimation procedure leads to less efficient estimates than the ordinary least squares when the contemporaneous correlation in the error terms is caused by mis-specification of the model. When there is no mis-specification in the model, then of course, these results do not hold. Even in the case where the estimated model is the true relation, for small sample sizes the SURE estimates are less efficient than the ordinary least squares. This result is due to Professor Zellner [6].

[6] This inequality holds for any number of equations in a set, and also for any number of independent variables in each equation.

The Nature of Bias in Standard Errors

The results presented in the above section may come as a surprise to researchers who worked with ordinary least squares and SURE estimation. Often standard errors of SURE estimates are smaller than standard errors of corresponding least squares estimates. Many researchers attributed this to the 'true' efficiency of SURE procedure. The true efficiency of an estimator is reflected by its variance and *not* by an estimate of the variance. Even though the ordinary least squares estimator has smaller variance than SURE whenever there is a left-out variable, the estimated variance would be larger for ordinary least squares than for SURE simply because of the way these estimates are computed from the data. As we shall prove below standard errors of the estimates do not reflect the true relative efficiency of these estimators.

The variance of the estimate $\hat{\alpha}_1$ is computed as[7]

$$\hat{V}(\hat{\alpha}_1) = T^{-1}(x_{11})^{-1}e_{11} \qquad (40)$$

In Equation (40) the expression $(T^{-1}e_{11})$ is an estimate of the variance u_1 $(= x_3 + \eta_1)$ and not that of η_1. The appropriate expression to compute $\hat{V}(\hat{\alpha}_1)$ is the variance of η_1 and not that of u_1. In the context of left-out variables the standard errors of ordinary least squares estimates are always upward biased [3, pp. 136–8].

On the other hand, the SURE procedure estimates the variance of α_1^* as

$$\hat{V}(\alpha^*) = (\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1} \qquad (41)$$

By using Equation (20) we may obtain the estimate of $\mathbf{V}(\alpha_1^*)$ as

$$\hat{V}(\alpha_1^*) = \frac{e_{11}x_{22}(e_{11}e_{22} - e_{12}e_{12})}{T(e_{11}e_{22}x_{11}x_{22} - e_{12}e_{12}x_{12}x_{12})} \qquad (42)$$

which simplifies to

$$\hat{V}(\alpha_1^*) = T^{-1}(x_{11})^{-1}e_{11}\frac{1 - r^2_{e_1e_2}}{1 - r^2_{e_1e_2}r^2_{x_1x_2}} \qquad (43)$$

[7] In this study we are dividing the residual sum of squares by T instead of the number of degrees of freedom merely for simplicity of notation.

A comparison of Equations (40) and (43) reveals that

$$\hat{V}(\alpha_1^*) \leq \hat{V}(\hat{\alpha}_1) \quad (44)$$

Whether the estimated model is a true model or a misspecified one, standard errors of SURE estimates will be smaller than the corresponding ordinary least squares estimates because of the way they are computed. Even in cases where $V(\alpha_1^*) > V(\hat{\alpha}_1)$ the estimates yield the inequality $\hat{V}(\alpha_1^*) \leq \hat{V}(\hat{\alpha}_1)$, in every sample. Therefore the computed standard errors of these estimators do not reflect the true relative efficiency.

Conclusions

Our analysis points out that bias and variance of the ordinary least squares and SURE estimators are sensitive to misspecification caused by a left-out variable, and the SURE procedure is more sensitive than the ordinary least squares. When contemporaneous correlation in a set of regression equations is the result of a common left-out variable, the SURE estimation procedure results in less efficient estimates than ordinary least squares. The true relative efficiency of these estimators is not reflected by the computed standard errors of the estimates.

References

[1] KMENTA, JAN and GILBERT, R., Small Sample Properties of Alternative Estimators of Seemingly Unrelated Regressions, *Journal of the American Statistical Association*, LXIII (December 1968) 1180–1200.

[2] RAO, POTLURI, Some Notes on Misspecification in Multiple Regressions, *The American Statistician*, XXV (December 1971) 37–9.

[3] RAO, POTLURI and MILLER, R. L., *Applied Econometrics*, Belmont: Wadsworth Publishing Co., 1971.

[4] TELSER, LESTER G., Iterative Estimation of a Set of Linear Regression Equations, *Journal of the American Statistical Association*, LIX (September 1964) 845–62.

[5] ZELLNER, ARNOLD, An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation

Bias, *Journal of the American Statistical Association*, LVII (June 1962) 348–68.

[6] ZELLNER, ARNOLD, Estimators for Seemingly Unrelated Regression Equations: Some Exact Finite Sample Results, *Journal of the American Statistical Association*, LVIII (December 1963) 977–92.