



AgEcon SEARCH

RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

UCD

Working paper series

89-17

OMISSION OF RELEVANT EXPLANATORY
VARIABLES IN SUR MODELS:
OBTAINING THE BIAS USING A TRANSFORMATION

Richard Green
University of California, Davis
Dept. of agricultural economics

ABSTRACT

This paper obtains an expression of the bias when relevant explanatory variables are omitted in SUR models. A simple demand system is provided to illustrate the direction of the bias terms when a variable is omitted.

GIANNINI FOUNDATION OF
AGRICULTURAL ECONOMICS
LIBRARY
WITHDRAWN
MAY 30 1990

Professor Richard Green
Department of Agricultural Economics
University of California
Davis, CA 95616

OMISSION OF RELEVANT EXPLANATORY VARIABLES IN SUR MODELS: OBTAINING THE BIAS USING A TRANSFORMATION

The purpose of this paper is to obtain an expression of the bias when relevant explanatory variables are omitted in SUR models. The approach uses a suitable transformation for the seemingly unrelated regressions in order to derive the bias terms (Bacon, Kmenta). In particular, the transformation allows the use of ordinary least squares estimation techniques. This simplifies the derivations considerably. In addition, a simple SUR model is used to illustrate the direction of the bias terms when certain variables are omitted in demand subsystems.

The Model

Suppose that the true model is given by

$$y_1 = X_1\beta_1 + X_3\beta_3 + \varepsilon_1 \quad (1)$$

$$y_2 = X_2\beta_2 + X_4\beta_4 + \varepsilon_2 \quad (2)$$

where y_i is a $T \times 1$ vector of observations on the dependent variable, X_i is a $T \times K_i$ matrix of values of the explanatory variables, β_i is a $K_i \times 1$ vector of parameters and ε_i is a $T \times 1$ vector of disturbances. The stochastic assumptions are

$$E(\varepsilon_i) = 0, \quad i = 1,2 \quad (3)$$

and

$$E(\varepsilon_i\varepsilon_j') = \sigma_{ij}I, \quad i,j = 1,2. \quad (4)$$

That is, the error terms across equations are contemporaneously correlated and there does not exist serial correlation over time.

Now suppose that X_3 and X_4 are erroneously omitted from the "true" model in (1) and (2). The disturbance terms become

$$\varepsilon_1^* = X_3\beta_3 + \varepsilon_1 \quad (5)$$

and

$$\varepsilon_2^* = X_4\beta_4 + \varepsilon_2. \quad (6)$$

The SUR estimators of β_1 and β_2 can be obtained by applying a transformation due to Bacon and presented in Kmenta, pp. 640-641. The transformed SUR regressions are

$$\begin{bmatrix} I & 0 \\ a_1I & a_2I \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ a_1I & a_2I \end{bmatrix} \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} I & 0 \\ a_1I & a_2I \end{bmatrix} \begin{bmatrix} \varepsilon_1^* \\ \varepsilon_2^* \end{bmatrix} \quad (7)$$

where $a_1 = \pm \sqrt{\frac{\rho^2}{1-\rho^2}}$ and $a_2 = \pm \sqrt{\frac{\sigma_{11}}{\sigma_{22}(1-\rho^2)}}$ and $\rho = \sigma_{11}/\sqrt{\sigma_{11}\sigma_{22}}$. The

transformation converts the disturbance terms such that the variance-covariance matrix of the original error terms in (1) and (2) is σ^2I . Also assume that a_1 and a_2 are known, i.e., assume that the covariances of the original disturbances are known.

Then the SUR estimators of β_1 and β_2 from (7) are

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \left\{ \begin{bmatrix} X_1 & 0 \\ a_1X_1 & a_2X_2 \end{bmatrix}' \begin{bmatrix} X_1 & 0 \\ a_1X_1 & a_2X_2 \end{bmatrix} \right\}^{-1} \begin{bmatrix} X_1 & 0 \\ a_1X_1 & a_2X_2 \end{bmatrix}' \begin{pmatrix} y_1 \\ a_1y_1 + a_2y_2 \end{pmatrix}. \quad (8)$$

Taking expectations of both sides of (8) yields

$$E(b) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' E \begin{pmatrix} y_1 \\ a_1y_1 + a_2y_2 \end{pmatrix} \quad (9)$$

where $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ and $\mathbf{X} = \begin{bmatrix} X_1 & 0 \\ a_1X_1 & a_2X_2 \end{bmatrix}$.

Since, the "true" values of y_1 and y_2 are given in (1) and (2),

$$E(y_1) = X_1\beta_1 + X_3\beta_3 \quad (10)$$

and

$$E(y_2) = X_2\beta_2 + X_4\beta_4. \quad (11)$$

Thus,

$$\begin{aligned} E(b) &= (\underline{X}' \underline{X})^{-1} \underline{X}' \begin{pmatrix} X_1\beta_1 + X_3\beta_3 \\ a_1(X_1\beta_1 + X_3\beta_3) + a_2(X_2\beta_2 + X_4\beta_4) \end{pmatrix} \quad (12) \\ &= (\underline{X}' \underline{X})^{-1} \underline{X}' \left\{ \underline{X} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{bmatrix} X_3 & 0 \\ a_1X_3 & a_2X_4 \end{bmatrix} \begin{pmatrix} \beta_3 \\ \beta_4 \end{pmatrix} \right\} \\ &= \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + (\underline{X}' \underline{X})^{-1} \underline{X}' \begin{pmatrix} X_3 & 0 \\ a_1X_3 & a_2X_4 \end{pmatrix} \begin{pmatrix} \beta_3 \\ \beta_4 \end{pmatrix}. \end{aligned}$$

The bias term is given by the second term on the right hand side of equation (12).

After simplifying, the bias term becomes

$$\begin{bmatrix} (1 + a_1^2)X_1'X_1 & a_1a_2X_1'X_2 \\ a_1a_2X_2'X_1 & a_2^2X_2'X_2 \end{bmatrix}^{-1} \begin{bmatrix} (1 + a_1^2)X_1'X_3 & a_1a_2X_1'X_4 \\ a_1a_2X_2'X_3 & a_2^2X_2'X_4 \end{bmatrix} \begin{pmatrix} \beta_3 \\ \beta_4 \end{pmatrix}. \quad (13)$$

In general, it is obvious that the direction of the bias when relevant variables are omitted is difficult to determine. However, in special cases the sign of the bias term can be obtained.

Special Cases

Let X_1, X_2, X_3 and X_4 be $T \times 1$ vectors of observations on the explanatory variables. Furthermore, let $X_3 = X_4$, i.e., assume that the same explanatory variable is omitted from each equation. With these assumptions, $X_1'X_1 = \sum X_{1t}^2$, $X_1'X_2 = \sum X_{1t}X_{2t}$, etc. Thus, the bias term becomes, after simplifications

$$\frac{1}{D} \begin{bmatrix} (1 + a_1^2)a_2^2 \sum X_{2t}^2 (\sum X_{1t}X_{3t}) - (a_1a_2)^2 (\sum X_{1t}X_{2t}) (\sum X_{2t}X_{3t}) \\ - (1 + a_1^2)(a_1a_2) (\sum X_{1t}X_{2t}) (\sum X_{1t}X_{3t}) + a_1a_2(1 + a_1^2) (\sum X_{1t}^2) (\sum X_{2t}X_{3t}) \\ a_2^2 a_1 a_2 (\sum X_{2t}^2) (\sum X_{1t}X_{3t}) - a_1 a_2 \cdot a_2^2 (\sum X_{1t}X_{2t}) (\sum X_{2t}X_{3t}) \\ - (a_1a_2)^2 (\sum X_{1t}X_{2t}) (\sum X_{1t}X_{3t}) + a_2^2 (1 + a_1^2) \sum X_{1t}^2 (\sum X_{2t}X_{3t}) \end{bmatrix} \begin{pmatrix} \beta_3 \\ \beta_4 \end{pmatrix} \quad (14)$$

where $D = a_2^2(1 + a_1^2) \sum X_{1t}^2 \sum X_{2t}^2 - (a_1a_2)^2 (\sum X_{2t}X_{1t})^2$, the determinant of product-cross product matrix. First, consider the sign of D . An alternative expression for D is

$$a_1^2 a_2^2 \left[\sum X_{1t}^2 \sum X_{2t}^2 - (\sum X_{2t}X_{1t})^2 \right] + a_2^2 \sum X_{1t}^2 \sum X_{2t}^2. \quad (15)$$

The term in brackets in (15) is positive by the Cauchy-Schwartz Inequality. The second term is obviously positive since it contains all squared terms. Hence, D is positive.

Next, consider the first row of (14) excluding $\frac{1}{D}$ which we have already shown is positive. The first row contains the bias term for β_1 . The first row is

$$\begin{bmatrix} (1 + a_1^2)a_2^2 (\sum X_{2t}^2) (\sum X_{1t}X_{2t}) - (a_1a_2)^2 (\sum X_{1t}X_{2t}) (\sum X_{2t}X_{3t}) \\ a_2^2 a_1 a_2 (\sum X_{2t}^2) (\sum X_{1t}X_{3t}) - a_1 a_2 a_2^2 (\sum X_{1t}X_{2t}) (\sum X_{2t}X_{3t}) \end{bmatrix} \begin{pmatrix} \beta_3 \\ \beta_4 \end{pmatrix} \quad (16)$$

Suppose, for example, that the omitted variable, say income is omitted from equations (1) and (2). Then both β_3 and β_4 are positive if the commodity is a superior good. If, in addition, the omitted variable is positively correlated with the included variables than each of the sums of the cross-products are positive. For instance, if income were omitted then it is usually positively correlated with prices. The coefficient of β_3 can be rewritten as

$$(a_1 a_2)^2 \left[(\sum X_{2t}^2)(\sum X_{1t} X_{2t}) - (\sum X_{1t} X_{2t})(\sum X_{2t} X_{3t}) \right] + a_2^2 (\sum X_{2t}^2)(\sum X_{1t} X_{2t}). \quad (17)$$

The second term in (17) is positive given the above assumptions. The term in brackets in (17) is just the numerator of the least squares estimator of α_3 (coefficient of X_3) obtained by the auxiliary regression of X_1 on X_2 and X_3 . Thus, the term in brackets is positive if X_1 (say the price of a commodity) is positively correlated with X_3 (say the omitted income variable or another price).

The sign of the coefficient on β_4 can be determined as follows. The coefficient of β_4 can be rewritten as

$$a_1 a_2 a_2^2 \left[(\sum X_{2t}^2)(\sum X_{1t} X_{3t}) - (\sum X_{1t} X_{2t})(\sum X_{2t} X_{3t}) \right]. \quad (18)$$

The term in brackets can be interpreted as the least squares coefficient on X_3 obtained from the auxiliary regression of X_1 on X_2 and X_3 . Thus, the least squares coefficient of X_3 is expected to be positive if X_1 is price and X_3 is another price or income. The sign of the only remaining term to consider is $a_1 a_2 a_2^2$. Now $a_1 a_2$ must satisfy the restriction

$$a_1 \sigma_{11} + a_2 \sigma_{12} = 0 \quad (\text{Kmenta, p. 641}) \quad (19)$$

where σ_{11} is the variance of ε_{1t} and σ_{12} is covariance between ε_{1t} and ε_{2t} . Clearly σ_{11} is positive. If σ_{12} is negative, then (19) implies that a_1 and a_2 are the same sign. Consequently, the bias term for β_1 , the SUR estimator when X_3 and X_4 are omitted, must be positive. If $\sigma_{12} > 0$, as would be expected if the two equations were demand functions, then (19) requires that a_1 and a_2 be of opposite signs. In this case the term involving β_4 in the bias expression is negative given the above assumptions; see (18). Thus, the sign of the overall bias term associated with the SUR estimator of β_1 depends upon the relative magnitudes of the two terms in (16). A priori it is difficult to state what direction the bias would be.

Consider the second row of (14). This row contains the bias associated with estimating β_2 . The bias is given by

$$\frac{1}{D} \left\{ \left[- (1 + a_1^2) (a_1 a_2) (\sum X_{1t} X_{2t}) (\sum X_{1t} X_{3t}) + a_1 a_2 (1 + a_1^2) (\sum X_{1t}^2) (\sum X_{2t} X_{3t}) \right] \beta_3 \right. \\ \left. + \left[- (a_1 a_2)^2 (\sum X_{1t} X_{2t}) (\sum X_{1t} X_{3t}) + a_2^2 (1 + a_1^2) (\sum X_{1t}^2) (\sum X_{2t} X_{3t}) \right] \beta_4 \right\}. \quad (20)$$

The coefficient of β_3 , excluding $\frac{1}{D}$ can be rewritten as

$$a_1 a_2 a_1^2 \left[(\sum X_{1t}^2) (\sum X_{2t} X_{3t}) - (\sum X_{1t} X_{2t}) (\sum X_{1t} X_{3t}) \right] - a_1 a_2 (\sum X_{1t} X_{2t}) (\sum X_{1t} X_{3t}). \quad (21)$$

The term in brackets is the numerator of the least squares estimator of X_2 obtained by regressing X_3 on X_1 and X_2 , e.g., by regressing the omitted income variable on included prices. The term in brackets is expected to be positive in most demand contexts. If the other variables are positively correlated, then the sign of the β_3 coefficients depends upon the sign of $a_1 a_2$. As before, if $\sigma_{12} > 0$, then a_1 and a_2 have opposite signs and the coefficient is positive. On the other hand if $\sigma_{12} < 0$, then a_1 and a_2 will have the same signs and the coefficient of β_3 will be negative given the positive correlation of the explanatory variables.

Finally, the coefficient of β_4 can be rewritten as

$$(a_1 a_2)^2 \left[(\sum X_{1t}^2)(\sum X_{2t} X_{3t}) - (\sum X_{1t} X_{2t})(\sum X_{1t} X_{3t}) \right] + a_2^2 (\sum X_{1t}^2)(\sum X_{2t} X_{3t}). \quad (22)$$

By using similar arguments as those above, this term will be positive; again making the same assumptions concerning the correlation between the explanatory variables.

The bias associated with the SUR estimator of β_2 can be either positive or negative; however, in special cases the sign can be determined to be positive.

Conclusions

This paper has developed an expression for the bias in a two equation SUR model with relevant explanatory variables omitted. In addition, by using the transformation of Bacon, the bias can be signed in special cases. An illustration based on a demand subsystem was employed as the special case, although other cases could easily be demonstrated.

REFERENCES

- Bacon, R., 1974, A Simplified Exposition of Seemingly Unrelated Regressions and the Three Stage Least Squares, *Oxford Bulletin of Economics and Statistics* 36, 229-233.
- Kmenta, J., 1986, *Elements of Econometrics*, 2nd ed., (Macmillan Publishing Co., New York).