# THE COVARIANCE TRANSFORMATION AND THE INSTRUMENTAL VARIABLES ESTIMATOR OF THE FIXED EFFECTS MODEL

by
Jeffrey Pliskin*

Uorking Paper No. 28

July 1989

Submitted to
The Jerome Levy Economics Institute
Bard College

*Hamilton College, Clinton, New York

This paper was written while I was a resident scholar at the Jerome Levy Economics Institute at Bard College. I am grateful to the Institute for its hospitality. I am indebted to Betsy Jensen, Bob Turner, and Jim White for comments on an earlier draft. I am responsible for all errors.

## ABSTRACT

The covariance transformation is a useful and often necessary procedure to estimate the fixed effects model. When some explanatory variables are contemporaneously correlated with the disturbance term, the covariance transformation can be used in conjunction with an instrumental variables procedure to obtain a consistent estimator. This paper describes how to correctly compute the IV estimator as a two stage least squares estimator. In addition, I show that if the IV estimator is incorrectly computed using a two stage least squares approach where the covariance transformation is not applied until the second stage, the resulting estimator is not in general consistent.

The use of panel data in economics has become more widespread in recent years (Hsiao (1986)). One model that is often adopted to take into account the heterogeneity of the cross-sectional units in the panel is the fixed effects model. When the number of cross-sectional units is large, it is computationally difficult, if not impossible, to compute the ordinary least squares (OLS) estimates using the standard formula. This computational problem has been overcome by the covariance transformation, which enables one to obtain the OLS estimates using standard computer packages. Although descriptions of how to use the covariance transformation to estimate the fixed effects model by OLS are widely available, I am not aware of a corresponding treatment of how to use the transformation in conjunction with an instrumental variables (IV) procedure. The purpose of this note is to describe how to use the covariance transformation to obtain the IV estimates for the fixed effects model when the model contains explanatory variables that are contemporaneously correlated with the disturbance term. In particular, I show that when the IV estimator is computed using a two stage least squares approach, the covariance transformation should be used in the first stage as well as in the second. If the transformation is used only in the second stage, the estimator will not be in general consistent.

I consider the following fixed effects model

$$y_{it} = \alpha_i + \Sigma x_{itk}\beta_k + \epsilon_{it} \quad i = 1,2,...,H; t = 1,2,...,T \qquad (1)$$

where the $\alpha_i$'s are the fixed effects which are intended to capture the (time-invariant) heterogeneity of the H cross-sectional units in the sample, the $\epsilon_{it}$'s are assumed to be identically and independently distributed random variables with zero mean and constant variance $\sigma^2$. In addition, the first K, of the K explanatory variables are assumed to be contemporaneously correlated with the disturbance term, i.e., $Ex_{itk}\epsilon_{it} \neq 0$ for $k = 1,2,\ldots,K_1$; the remaining $K_2$ of the explanatory variables are exogenous. For concreteness, I will assume that the sample consists of a panel of H households observed over T periods.

If the observations are ordered first by household and then by time, we can rewrite (1) in matrix notation as

$$Y = D\alpha + X\beta + \epsilon \tag{2}$$

where D is a (HT x H) matrix of household dummy variables which is given by D = I, $\otimes J_T$, with I, being a (H x H) identity matrix and J, being a (T x 1) vector of ones. The dimensions of Y and $\epsilon$ are (HT x 1) and the dimension of X is (HT x K). Let $X = [X_1 \; X_2]$, W = [D X], and N = HT, where X, and $X_2$ are (N x $K_1$) and (N x $K_2$) respectively.

I begin by considering the OLS estimator to illustrate the covariance transformation. The OLS estimator of $(\alpha' \; \beta')'$ is given by $(W'W)^{-1}W'Y$. When the number of households is large, it is difficult, if not impossible, to invert $(W'W)$, a matrix of order (H + K). In these instances, the OLS estimate of $\beta$ is obtained by using the covariance transformation to transform the model and then estimating the transformed model by least squares. Specifically, let $M_D = I_n - D(D'D)^{-1}D'$ be an idempotent matrix (i.e., $M_D = M_D M_D$) which transforms observations on Y and X into deviations from their respective household means. For example, the vector $Y^* = M_D Y$ has a typical element $y^*_{it} = y_{it} - \bar{y}_i$, where $\bar{y}_i = (1/T)\Sigma y_{it}$. One can transform the model (2) by premultiplying by $M_D$ to obtain

$$M_D Y = M_D D\alpha + M_D X\beta + M_D \epsilon \qquad (3)$$

Since D is orthogonal to $M_D$, i.e. $M_D D = 0$, (3) simplifies to

$$Y^* = X^*\beta + \epsilon^* \qquad (4)$$

The OLS estimate of $\beta$, b, can be obtained by regressing $Y^*$ on $X^*$, i.e.,

$$b = (X^{*\prime}X^*)^{-1}X^{*\prime}Y^* \qquad (5)$$

$$= (X'M_D X)^{-1}(X'M_D Y) \qquad (6)$$

It is well known that this estimate is identical to the last K elements of $(W'W)^{-1}W'Y$. Thus, the covariance transformation simply enables one to obtain the OLS estimate of $\beta$ by a computationally convenient method.

Given the assumption that $X_t$ is contemporaneously correlated with $\epsilon$, the OLS estimator of $\beta$ is not consistent. A common remedy to this problem is to use instrumental variables. Thus, assume that there exists a (N x R) matrix (R ≥ K) of instruments, $Z = [Z_1\ X_2]$, for X such that' $\text{plim}(1/H)(Z'M_D\epsilon) = 0$, $\text{plim}(1/H)(Z'M_DZ)$ is a non-singular matrix of finite constants, and $\text{plim}(1/H)(Z'M_DX)$ is a matrix of finite constants of full column rank. Let $\delta' = (\alpha'\ \beta')'$. The IV estimator of 6, $d^{IV}$, is given by (see Bowden and Turkington (1984))

$$d^{IV} = (W'P_QW)^{-1}W'P_QY \tag{7}$$

where Q = [D Z] is a matrix of full column rank and $P_Q = Q(Q'Q)^{-1}Q'$ is an idempotent matrix. Let $\hat{W} = P_QW$. Since $\hat{W}'\hat{W} = W'P_QP_QW = W'P_QW$, we have

$$d^{IV} = (\hat{W}'\hat{W})^{-1}\hat{W}'Y \tag{8}$$

The version of the IV estimator given by (8) shows that it can be obtained as a two stage least squares estimator: $\hat{W}$ is obtained in stage one as the fitted values of W based on a regression of W on Q; Y is regressed on $\hat{W}$ in stage two, thereby yielding $d^{IV}$.

When H is large, it will be computationally difficult to obtain the IV estimates from either (7) or. (8) because each involves inverting a matrix of order (H + K). Thus, one might want to use the covariance transformation to reduce the dimensions of the matrices that need to be inverted. I will consider two approaches to applying the covariance transformation. The first yields an estimate of $\beta$ that is identical to the one given in (7) or (8). The second yields an estimator that differs from $b^{IV}$, and moreover, is an inconsistent estimator of $\beta$. I consider this second approach because, as I will discuss below, it corresponds to an error that is often made in obtaining the two stage least squares estimates of the parameters of simultaneous equation models.

The correct way of using the covariance transformation to compute the IV estimator as a two stage least squares estimator is to apply it at both stages. To see this, I begin by examining the regression equations for the first stage. Since $W = [D\ X,\ X_2]$ and $Q = [D\ Z_1 X_2]$, one obtains $\hat{W}$ by running the following regressions:

$$D = D\pi_{00} + Z_1\pi_{01} + X_2\pi_{02} + \epsilon_0 \qquad (9)$$

$$X_1 = D\pi_{10} + Z_1\pi_{11} + X_2\pi_{12} + \epsilon_1 \qquad (10)$$

$$X_2 = D\pi_{20} + Z_1\pi_{21} + X_2\pi_{22} + \epsilon_2 \qquad (11)$$

It is clear from (9) and (11) that $\hat{D} = D$ and $\hat{X}_2 = X_2$. Applying the results for the inverse of a partitioned matrix to (8), one can show that the IV estimator of $\beta$ is given by

$$b^{IV} = (\hat{X}'M_D\hat{X})^{-1}(\hat{X}'M_DY) \qquad (12)$$

This implies that $b^{IV}$ is obtained by regressing $Y^*$ on $\hat{X}^* = M_D\hat{X}$ $= [M_D\hat{X}_1 \; M_DX_2]$, where $M_D\hat{X}_1$ is obtained as follows. Since D is orthogonal to $M_D$, (10) implies that

$$M_D\hat{X}_1 = M_DZ_1\hat{\pi}_{11} + M_DX_2\hat{\pi}_{12} \tag{13}$$

where $\hat{\pi}_{11}$ and $\hat{\pi}_{12}$ are the OLS estimates of $\pi_{11}$ and $\pi_{12}$, which can be easily computed by applying the covariance transformation to (10) and estimating the transformed model by least squares. Given. the assumptions made above, $b^{IV}$ is a consistent estimator of $\beta$. (See Proposition 1 in the Appendix.)

In contrast, an inconsistant estimator results when the following two stage least squares approach is used. In the first stage regression, the fitted values of X, are obtained by regressing X, on Z = $[Z_1 \; X,]$, yielding $\tilde{X}_1 = Z\tilde{\phi}$, where $\tilde{\phi} = (Z'Z)^{-1}Z'X_1$. In the second stage, $Y^*$ is regressed on $\tilde{X}^* = M_D\tilde{X} = [M_D\tilde{X}_1 \; M_DX_2]$. This is equivalent to obtaining an estimator of $\beta$ by using X in place of $\hat{X}$ in (12), i.e.,

$$\tilde{b} = (\tilde{X}'M_D\tilde{X})^{-1}(\tilde{X}'M_DY) \tag{14}$$

In the Appendix (see Proposition 2), I show that the resulting estimator is not in general consistent.*

The inconsistency of $\tilde{b}$ arises because X, is not computed using all exogenous variables (i.e., Q), while the "fitted values" of D and $X_2$ (which are equal to their actual values) are implicitly based on all components of Q including D. This error is essentially the same as one that is often made in estimating simultaneous equation models, where many researchers omit some of the predeterminedvariables in the first stage regressions (Hausman (1983) and Bowden and Turkington (1984)).

In summary, I have described how to use the covariance transformation to correctly compute the IV estimator of the fixed effects model as a two stage least squares estimator. In particular, I showed that the transformation should be used in both stages.

APPENDIX

Proposition 1. If $\text{plim}(1/H)(Z'M_D\epsilon) = 0$; $\text{plim}(1/H)(Z'M_DZ) = B$, a non-singular matrix of finite constants: and $\text{plim}(1/H)(Z'M_DX) = C$, a matrix of finite constants of full column rank, then $\text{plim}\ b^{IV} = \beta$.

Proof. Using equation (12) in the text and the definition of $\hat{X}$, we have

$$b^{IV} = \{(X'M_DZ)(Z'M_DZ)^{-1}(Z'M_DX)\}^{-1} \times\{(X'M_DZ)(Z'M_DZ)^{-1}(Z'M_DX)\beta + (X'M_DZ)(Z'M_DZ)^{-1}(Z'M_D\epsilon)\} \quad (A1)$$

$$= \beta + \{(X'M_DZ)(Z'M_DZ)^{-1}(Z'M_DX)\}^{-1}(X'M_DZ)(Z'M_DZ)^{-1}Z'M_D\epsilon \quad (A2)$$

The assumptions made above imply that

$$\text{plim}\ b^{IV} = \beta + (C'B^{-1}C)^{-1}C'B^{-1}0 \quad (A3)$$

$$= \beta \quad (A4)$$

Proposition 2. In addition to the assumptions of Proposition 1, suppose that $\text{plim}(1/H)(Z'Z) = F$, a non-singular matrix of finite constants and $\text{plim}(1/H)(Z'X) = G$, a matrix of finite constants of full column rank. Then $\text{plim}\ \tilde{b} \neq \beta$.

Proof. Using equation (14) of the text and the definition of $\tilde{X}$, we have

$$\tilde{b} = \{X'Z(Z'Z)^{-1}(Z'M_DZ)^{-1}(Z'Z)^{-1}(Z'X)\}^{-1}$$
$$\times\{(X'Z)(Z'Z)^{-1}(Z'M_DX)\beta + (X'Z)(Z'Z)^{-1}(Z'M_D\epsilon)\} \qquad \text{(A5)}$$

The assumptions made above imply that

$$\text{plim } \tilde{b} = \{G'F^{-1}B^{-1}F^{-1}G\}^{-1}\{G'F^{-1}C\beta + G'F^{-1}0\} \qquad \text{(A6)}$$

$$= \{G'F^{-1}B^{-1}F^{-1}G\}^{-1}(G'F^{-1}C)\beta \qquad \text{(A7)}$$

$$\neq \beta \qquad \text{(A8)}$$

# FOOTNOTES

[1]Since most panel data sets are characterized by many cross-sectional units observed over a small number of time periods, I state consistency properties for fixed T and $H \to \infty$. Consequently, all probability limits (plims) are defined for $H \to \infty$.

[2]In addition to the assumptions made in the text, a set of sufficient conditions for the inconsistency of $\tilde{b}$ are (i) plim(1/H)(Z'Z) is a nonsingular matrix of finite constants and (ii) plim(1/H)(Z'X) is a matrix of finite constants of full column rank.

## REFERENCES


Bowden, R. J. and Turkington D. A. (1984). Instrumental Variables, Cambridge, Cambridge University Press.

Hausman, J. A. (1983). "Specification and Estimation of Simultaneous Equation Models," in Griliches, Z. and Intrilligator, M. (eds.), Handbook of Econometrics, Vol. 1, Amsterdam, North-Holland.

Hsiao, C. (1986). Analysis of Panel Data, Cambridge, Cambridge University Press.