Econometric Reviews
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597248

Alternative transformations to eliminate fixed effects

Marno Verbeek [a]
[a] Department of Econometrics and CentER, Tilburg University, Tilburg, LE, The Netherlands

Online Publication Date: 01 January 1995

To cite this Article: Verbeek, Marno (1995) 'Alternative transformations to eliminate fixed effects', Econometric Reviews, 14:2, 205 — 211

To link to this article: DOI: 10.1080/07474939508800315
URL: http://dx.doi.org/10.1080/07474939508800315

ECONOMETRIC REVIEWS, 14(2), 205-211 (1995)

ALTERNATIVE TRANSFORMATIONS TO ELIMINATE FIXED EFFECTS

Marno Verbeek

Department of Econometrics and CentER
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands

Keywords: panel data, fixed effects models
JEL classification: C23

ABSTRACT

In a panel data model with fixed individual effects, a number of alternative transformations are available to eliminate these effects such that the slope parameters can be estimated from ordinary least squares on transformed data. In this note we show that each transformation leads to algebraically the same estimator if the transformed data are used efficiently (i.e. if GLS is applied). If OLS is used, however, differences may occur and the routinely computed variances, even after degrees of freedom correction, are incorrect. In addition, it may matter whether "redundant" observations are used or not.

1. INTRODUCTION

It is common practice in panel data modeling to allow each unit to have its own intercept term to account for heterogeneity in individual behavior. Consider the linear model

$$y_{it} = x_{it}\beta + \alpha_i + \varepsilon_{it}, \quad i = 1, \dots, N \ ; \ t = 1, \dots, T \tag{1}$$

where i denotes individuals, t denotes time and β is a k dimensional parameter vector of interest. The α_i's are treated as fixed unknown parameters, while ε_{it} is an i.i.d. zero mean error term, with variance σ_ε^2 and uncorrelated with x_{is}, i.e. the x's are strictly exogenous.

It is well known (cf. Hsiao (1986, p. 30)) that under these assumptions, ordinary least squares in (1) is best linear unbiased. The implied estimator for β is identical to the least squares estimator obtained from a transformed regression, where the data on y_{it} and x_{it} are taken in deviation from their individual means. Stacking observations by individual, (1) can be written as

$$y_i = X_i\beta + \alpha_i\iota_T + \varepsilon_i, \quad i = 1, \dots, N \tag{2}$$

where $\iota_T = (1, 1, \dots, 1)'$ of dimension T. The least squares dummy variable estimator for β is obtained through OLS on transformed data with transformation matrix $Q = I_T - \frac{1}{T}\iota_T\iota_T'$. Thus, noting that $QQ = Q$ and $Q = Q'$, one obtains

$$\hat{\beta}_Q = \left(\sum_{i=1}^{N} X_i'QX_i\right)^{-1} \left(\sum_{i=1}^{N} X_i'Qy_i\right) \tag{3}$$

with variance

$$V\{\hat{\beta}_Q\} = \sigma_\varepsilon^2 \left(\sum_{i=1}^{N} X_i'QX_i\right)^{-1} . \tag{4}$$

If $\hat{\beta}_Q$ is computed from transformed data, the standard estimate for its variance is

$$\hat{V}\{\hat{\beta}_Q\} = \hat{\sigma}_\varepsilon^2 \left(\sum_{i=1}^{N} X_i'QX_i\right)^{-1} \tag{5}$$

where

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{NT-k}\sum_{i=1}^{N}(y_i - X_i\hat{\beta}_Q)'Q'Q(y_i - X_i\hat{\beta}_Q). \tag{6}$$

Since this estimator ignores the singularity in the transformed data, it is biased. In particular, it underestimates the true variance by $(NT-k)/(NT-N-k)$, which, for large N, approximately equals $T/(T-1)$. It is straightforward to adjust for this underestimation.

Instead of transforming the data into deviations from individual means, it is possible to eliminate the fixed effects in (1) by using any transformation matrix A satisfying $A\iota_T = 0$. When the exogenous variables in the model are affected by measurement error, for example, it is argued that using different transformations may lead to different biases, so that the corresponding (inconsistent) estimators can be combined to produce a consistent one (see Griliches and Hausman (1986)). In the next section, a general discussion of the properties of estimators based on alternative transformations will be given. Section 3 discusses the implications of these results and concludes.

2. ALTERNATIVE TRANSFORMATIONS

Suppose fixed effects are eliminated by an $S \times T$ transformation matrix A of rank $T-1$, satisfying $A\iota_T = 0$, where S equals either T or $T-1$. The transformed model is given by

$$Ay_i = AX_i\beta + A\varepsilon_i, \quad i = 1,\ldots,N. \tag{7}$$

The OLS-estimator on this newly transformed data is

$$\hat{\beta}_A = \left(\sum_{i=1}^{N} X_i'A'AX_i\right)^{-1}\left(\sum_{i=1}^{N} X_i'A'Ay_i\right), \tag{8}$$

and, consequently, $\hat{\beta}_A = \hat{\beta}_Q$ whenever

$$A'A = cQ \tag{9}$$

for some (nonzero) constant c. The variance of this estimator is

$$V\{\hat\beta_A\} = \sigma_\varepsilon^2 \left(\sum_{i=1}^N X_i'A'AX_i\right)^{-1} \sum_{i=1}^N X_i'A'AA'AX_i \left(\sum_{i=1}^N X_i'A'AX_i\right)^{-1}. \quad (10)$$

The standard least squares estimator $\hat\sigma_{\varepsilon A}^2$ for σ_ε^2 obtained from (7) will, in general, not be unbiased. In particular, its expectation can be shown to equal

$$E\{\hat\sigma_{\varepsilon A}^2\} = \sigma_\varepsilon^2 \frac{N \operatorname{tr} A'A - \operatorname{tr}\left(\sum_{i=1}^N X_i'A'AX_i\right)^{-1} \sum_{i=1}^N X_i'A'AA'AX_i}{NS-k}. \quad (11)$$

By combining (10) and (11), it follows that the routinely computed variance given by

$$\hat V\{\hat\beta_A\} = \hat\sigma_{\varepsilon A}^2 \left(\sum_{i=1}^N X_i'A'AX_i\right)^{-1} \quad (12)$$

is an unbiased estimator for the true variance, apart from a known factor $(NT-N-k)/(NS-k)$, only if

$$A'AA'A = cA'A \quad (13)$$

for some (nonzero) constant c. Although this condition seems weaker than (9), it is not. The symmetry of $A'A$, together with (13) requires that $A'A$ is a constant times a linear projection matrix. As the row space of A is characterized by $A\iota_T = 0$, the appropriate projection matrix is Q. Consequently, condition (13) corresponds with condition (9). Thus, if (9) holds, the routinely computed variance is unbiased for $S = T-1$. If $S = T$, the variance is, similar to our earlier results, underestimated by $(NT-k)/(NT-N-k)$. The estimator for σ_ε^2 is unbiased only if (9) holds for $c = 1$ and $S = T-1$.

As the transformed error term in (7) is no longer i.i.d., it is a reasonable alternative to estimate β using generalized least squares. Note that $V\{A\varepsilon_i\} = \sigma_\varepsilon^2 AA'$ is singular if $S = T$. Whether or not this is the case does not affect the GLS estimator because the same singularity also holds for the transformed model. Restricting attention, for the moment, to the case where AA' is of full rank, i.e. $S = T-1$, the GLS estimator is given by

$$\hat\beta_A(GLS) = \left(\sum_{i=1}^N X_i'A'(AA')^{-1}AX_i\right)^{-1} \left(\sum_{i=1}^N X_i'A'(AA')^{-1}Ay_i\right). \quad (14)$$

Since $A'(AA')^{-1}A$ is the projection matrix onto the row space of A characterized by $A\iota_T = 0$, it is generally true that $A'(AA')^{-1}A = Q$. Consequently, for any choice of the transformation matrix A, the GLS estimator for β is identical to the within estimator $\hat{\beta}_Q$. Using the results of Amemiya (1985, p. 185), it can be shown that this also holds for the case where AA' is singular.

As $\hat{\beta}_A(GLS) = \hat{\beta}_Q$ for any choice of A, both estimators have the same variance matrix. Moreover, it can be easily shown that the estimated variance of $\hat{\beta}_A(GLS)$ equals the estimated variance of $\hat{\beta}_Q$ and thus underestimates the true variance.

We can summarize the results of this section as follows. Independent of the transformation that is used to eliminate the fixed individual effects, efficient use of the available information leads to the same estimator with the same properties, namely the within estimator $\hat{\beta}_Q$. However, if ordinary least squares is applied to transformed data, differences may arise if the condition in (9) is not fulfilled. Only when this condition is fulfilled is an unbiased estimator of the least squares variance readily obtained from standard regression output, given a degrees of freedom correction if $S = T$. We shall illustrate these points in the next section.

3. IMPLICATIONS

A common alternative to the within transformation is the transformation in first differences. The $T \times T$ matrix A in this case is given by

$$A_1 = \begin{bmatrix} 1 & 0 & & \dots & 0 & -1 \\ -1 & 1 & & & & 0 \\ \vdots & -1 & 1 & \dots & & \vdots \\ 0 & \dots & & & -1 & 1 \end{bmatrix} \tag{15}$$

such that the elements in the transformed data set are $y_{i1} - y_{iT}$ and $y_{it} - y_{i,t-1}$ $(t = 2, \dots, T)$. It is easily verified that $A_1'A_1 = TQ$ if $T = 2, 3$, from which it immediately follows that the first difference estimator and the within estimator are identical for $T = 2$ and $T = 3$. For the more common case where $T > 3$

the estimators are different. Moreover, the routinely computed variance, even after the degrees of freedom correction, will be biased. An unbiased estimate for the variance of the estimator based on the first difference transformation given in (15) cannot be easily derived.

Now suppose we drop the first observation for each individual because it is an exact linear combination of the last $T-1$ observations. Then the corresponding transformation matrix is A_2, say, which is a $(T-1)\times T$ matrix corresponding to A_1 with its first row deleted. It is easily verified that there is no constant c such that $A_2'A_2 = cA_1'A_1$ unless $T=2$. Thus, for $T>2$ the first difference estimator based on A_2 does *not* equal the first difference estimator based on A_1 although the only difference is that a "redundant" observation is dropped. In addition, when the number of time periods T exceeds 2, the variance of $\hat{\beta}_{A_1}$ is incorrectly estimated, and it is not straightforward to adjust for the bias in the estimated variance.

In the case of three time periods ($T=3$), we have seen that the first difference estimator based on OLS using A_1 is best linear unbiased, as it equalled the within estimator. From the discussion above, it follows that the first difference estimator based on A_2 is not best linear unbiased. Of course, the GLS estimator based on A_2, as was shown in the previous section, is identical to the within estimator and thus best linear unbiased.

Note that this conclusion implies that, when applying OLS to differenced observations, the routinely computed OLS variances are only valid for $T=2$, and for $T=3$ provided the transformation matrix includes the transformation $y_{i1}-y_{iT}$, i.e. if A_1 is used (assuming that a degrees of freedom correction is applied).

The general conclusion from the results above is the following. Instead of the within transformation to eliminate individual effects in panel data models, alternative transformations can be applied. If the transformed models are estimated with a generalized least squares procedure, the resulting estimator is equal to the within estimator which is BLUE. However, if the transformed model is estimated with ordinary least squares, different estimators result. In the special case where the transformation matrix A is such that $A'A$ equals a constant times the projection matrix Q, the OLS estimator equals the within estimator. Moreover, this is the only situation in which the routinely computed OLS variance is an unbiased estimator for the true variance (given a degrees of freedom correction, if needed). In general, obtaining an unbiased (or a consistent) estimator for the true variance is not straightforward.

ACKNOWLEDGMENTS

I have benefited from useful discussions with Theo Nijman and Geert Thijssen.

REFERENCES

Amemiya, T., 1985, *Advanced Econometrics,* Basis Blackwell, Oxford.

Griliches, Z. and J.A. Hausman, 1986, Errors in Variables in Panel Data, *Journal of Econometrics*, 31, 93-118.

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