

Fixed Effects Models with Time Invariant Variables: A Theoretical Note

Ronald L. Oaxaca (University of Arizona) and
Iris Geisler (Dun & Bradstreet)*

Revised January 2003

Abstract

This paper demonstrates the equivalence between a consistent two-stage *GLS* estimator and the pooled *OLS* estimator of the coefficients on time invariant covariates in an unbalanced *FE* panel. In general the estimated standard errors differ between these two procedures.

JEL: C2

Key Words: panel data, time invariant regressors

It is not uncommon to find explanatory variables of interest in panel data sets that are time invariant, e.g. race, sex, regional location. In a fixed effects model these variables are “swept away” by the within estimator of the coefficients on the time varying covariates. Nevertheless, it is possible to identify and consistently estimate the effects of the time invariant regressors through two-stage procedures. Hausman and

*Corresponding author: Ronald L. Oaxaca, Department of Economics, 401 McClelland Hall, University of Arizona, PO Box 210108, Tucson, AZ 85721-0108. Tel. no. (520) 621-4135, Fax no. (520) 621-8450. email: rlo@email.arizona.edu.

Taylor (1981) analyzes models in which some of the variables (both time varying and time invariant) are endogenous. Baltagi (1995) provides a comprehensive treatment of panel data models in the contexts of both single equation and systems methods. Polachek and Kim (1994) examines a single equation model in which the slope parameters of time-invariant regressors vary across individuals. In this paper we consider the case of a single equation model with an unbalanced design and time invariant regressors. We develop a two-stage *GLS* estimation procedure for consistent estimation of the coefficients on the time invariant regressors, and demonstrate the equivalence between these *GLS* coefficient estimates and the *OLS* coefficient estimates on the time invariant covariates from a pooled cross-section, time-series model. However, in general the estimated standard errors differ between these two procedures.

The basic model may be expressed as

$$Y_{it} = \alpha_i + X_{it}\beta + \varepsilon_{it}, \quad i = 1, \dots, n \quad t = 1, \dots, T_i, \quad (1)$$

where α_i is an individual specific intercept term, X_{it} is a $1 \times k_2$ vector of observations on the time varying covariates, and β is a $k_2 \times 1$ parameter vector. Consider the restriction $\alpha_i = \alpha + Z_{1i}\gamma^1$ where α is a constant, Z_{1i} is a $1 \times k_1$ vector of observations on time invariant covariates, γ^1 is a $k_1 \times 1$ parameter vector. Under the restrictions on the fixed effects, pooled *OLS* would be the estimator of choice given the usual assumptions on ε_{it} . If the restrictions are not true, the fixed effects within estimator can be used to estimate β from the model:

$$y_{it} = x_{it} \beta + \varepsilon_{it} - \varepsilon_i. \quad (2)$$

where $y_{it} = Y_{it} - Y_i$, $x_{it} = X_{it} - X_i$, and $i.$ denotes the group mean average for the i th cross-sectional unit. The *FE* estimator is accordingly

$$\tilde{\beta}^{fe} = (X' M_D X)^{-1} X' M_D Y \quad (3)$$

where X is a $n\bar{T} \times k_2$ observation matrix for the X 's, $M_D = I_{n\bar{T}} - D(D'D)^{-1}D'$, D is a $n\bar{T} \times n$ observation matrix for the individual dummy variables, $\bar{T} = n^{-1} \sum_{i=1}^n T_i$, and Y is a $n\bar{T} \times 1$ observation vector on the Y 's.

Note that the group means representation of the restricted model is given by

$$Y_i = \alpha + Z_{1i}\gamma^1 + X_i\beta + \varepsilon_i. \quad (4)$$

Upon subtracting $X_i\tilde{\beta}^{fe}$ from both sides of (4) we obtain

$$\begin{aligned} Y_i - X_i\tilde{\beta}^{fe} &= \alpha + Z_{1i}\gamma^1 + \eta_i, \quad i = 1, \dots, n \\ &= Z_i\gamma + \eta_i, \end{aligned} \quad (5)$$

where $\eta_i = X_i(\beta - \tilde{\beta}^{fe}) + \varepsilon_i$, $Z_i = (1, Z_{1i})$ and $\gamma = (\alpha, \gamma^1)'$. Note that $E(\eta_i) = 0$, and $\text{Var}(\eta) = \sigma_\varepsilon^2 \Omega$, where $\eta = (\eta_1, \dots, \eta_n)'$. A typical diagonal element of Ω is given by $\left(\frac{1}{T_i} + X_i(X'M_D X)^{-1}X_i'\right)$, and a typical off-diagonal element is given by $X_i(X'M_D X)^{-1}X_j'$. Therefore, $\Omega = (D'D)^{-1} + \bar{X}(X'M_D X)^{-1}\bar{X}'$, where $\bar{X} = (D'D)^{-1}D'X$ is a $n \times k_2$ observation matrix on the X_i 's. Polachek and Kim (1994) estimate (5) by *OLS* and a *GLS* method that corrects only for the heteroscedasticity inherent in the diagonal elements of Ω . Our two-stage *GLS* estimator of γ also takes account of the non zero off diagonal elements of Ω :

$$\tilde{\gamma}^{fe} = (\bar{Z}'\Omega^{-1}\bar{Z})^{-1}\bar{Z}'\Omega^{-1}(\bar{Y} - \bar{X}\tilde{\beta}^{fe}), \quad (6)$$

where \bar{Z} is a $n \times (k_1 + 1)$ observation matrix on the constant term and the Z_i 's, and \bar{Y} is a $n \times 1$ observation vector on the Y_i 's. The variance/covariance matrix for $\tilde{\gamma}^{fe}$ is given by $\sigma_\varepsilon^2 (\bar{Z}'\Omega^{-1}\bar{Z})^{-1}$.

Now consider the restricted pooled cross-section, time series model specified by

$$Y_{it} = Z_i\gamma + X_{it}\beta + \varepsilon_{it}, \quad i = 1, \dots, n \quad t = 1, \dots, T_i. \quad (7)$$

Let $\hat{\gamma}^{ols}$ denote the *OLS* estimator of γ from equation (7):

$$\hat{\gamma}^{ols} = (Z'M_x Z)^{-1}Z'M_x Y,$$

where Z is the $n\bar{T} \times (k_1 + 1)$ observation matrix on the time invariant covariates and constant term, and $M_x = I_{n\bar{T}} - X(X'X)^{-1}X'$. It is easily shown that $\text{Var}(\hat{\gamma}^{ols}) = \sigma_\varepsilon^2 (Z'M_x Z)^{-1}$. It turns out that $\hat{\gamma}^{ols} = \tilde{\gamma}^{fe}$, or more formally

Theorem: $(Z'M_x Z)^{-1} Z'M_x Y = (\bar{Z}'\Omega^{-1}\bar{Z})^{-1} \bar{Z}'\Omega^{-1} (\bar{Y} - \bar{X} \tilde{\beta}^{fe})$.

The following results will be used in the proof of the theorem:

$$\bar{Z} = (D'D)^{-1} D'Z$$

$$\bar{X} = (D'D)^{-1} D'X$$

$$\bar{Y} = (D'D)^{-1} D'Y$$

Proof:

Part 1

We first show that $(Z'M_x Z)^{-1} = (\bar{Z}'\Omega^{-1}\bar{Z})^{-1}$. This is equivalent to showing $Z'M_x Z = \bar{Z}'\Omega^{-1}\bar{Z}$.

Upon substitution for \bar{Z} , we have $Z'M_x Z = Z' [D(D'D)^{-1}\Omega^{-1}(D'D)^{-1}D'] Z$.

Since $Z \neq \begin{matrix} 0 \\ (n\bar{T} \times (k_1+1)) \end{matrix}$, it suffices to show $M_x = D(D'D)^{-1}\Omega^{-1}(D'D)^{-1}D'$.

Premultiplying by D' and postmultiplying by D , yields

$$D'M_x D = \Omega^{-1} \Rightarrow (D'M_x D)^{-1} = \Omega. \text{ Substitution for } M_x \text{ yields}$$

$$[D'D - D'X(X'X)^{-1}X'D]^{-1} = \Omega$$

We make use of the matrix algebra result

$$[A + BCB']^{-1} = A^{-1} - A^{-1}B[C^{-1} + B'A^{-1}B]^{-1}B'A^{-1} \text{ (see Greene, 4th ed.)}$$

where $A = D'D$, $B = D'X$, and $C = (-X'X)^{-1}$. Upon making these substitutions, we obtain

$$\begin{aligned}
& (D'D)^{-1} - (D'D)^{-1} D'X [-X'X + X'D(D'D)^{-1} D'X]^{-1} X'D (D'D)^{-1} = \Omega \\
\Rightarrow & (D'D)^{-1} + (D'D)^{-1} D'X [X'X - X'D(D'D)^{-1} D'X]^{-1} X'D (D'D)^{-1} = \Omega \\
\Rightarrow & (D'D)^{-1} + (D'D)^{-1} D'X [X'M_D X]^{-1} X'D (D'D)^{-1} = \Omega
\end{aligned}$$

Substituting \bar{X} for $(D'D)^{-1} D'X$ yields

$$(D'D)^{-1} + \bar{X} (X'M_D X)^{-1} \bar{X}' = \Omega, \text{ which completes the first part of the proof.}$$

This also proves that

$$\text{Var}(\hat{\gamma}^{ols}) = \sigma_\varepsilon^2 (Z'M_x Z)^{-1} = \sigma_\varepsilon^2 (\bar{Z}'\Omega^{-1}\bar{Z})^{-1} = \text{Var}(\tilde{\gamma}^{fe}).$$

Part 2

The second half of the proof is to show $Z'M_x Y = \bar{Z}'\Omega^{-1}(\bar{Y} - \bar{X}\tilde{\beta}^{fe})$.

Upon substituting for \bar{Z} , \bar{Y} and \bar{X} and collecting terms, we have

$$Z'M_x Y = Z'D (D'D)^{-1} \Omega^{-1} (D'D)^{-1} D' (Y - X\tilde{\beta}^{fe}).$$

As shown above in Part 1, $M_x = D (D'D)^{-1} \Omega^{-1} (D'D)^{-1} D'$; therefore, we can write

$$Z'M_x Y = Z'M_x (Y - X\tilde{\beta}^{fe}) \Rightarrow$$

$$Z'M_x Y = Z'M_x Y - Z'M_x X\tilde{\beta}^{fe} = Z'M_x Y \text{ since } M_x X = \begin{matrix} 0 \\ (n\bar{T} \times k_2) \end{matrix}.$$

This proves the second half of the theorem: $Z'M_x Y = \bar{Z}'\Omega^{-1}(\bar{Y} - \bar{X}\tilde{\beta}^{fe})$

Thus, $(Z'M_x Z)^{-1} Z'M_x Y = (\bar{Z}'\Omega^{-1}\bar{Z})^{-1} \bar{Z}'\Omega^{-1}(\bar{Y} - \bar{X}\tilde{\beta}^{fe})$ or $\hat{\gamma}^{ols} = \tilde{\gamma}^{fe}$.

Although the true standard errors for $\hat{\gamma}^{ols}$ and $\tilde{\gamma}^{fe}$ are the same, the estimated standard errors are in general different for the two estimation procedures. This is due to differences in estimating the error variance. For model (7)

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^n \sum_{t=1}^{T_i} (Y_{it} - Z_i \hat{\gamma}^{ols} - X_{it} \hat{\beta}^{ols})^2}{n\bar{T} - (k_1 + k_2 + 1)},$$

and for *GLS* estimation of model (5) the error variance for ε is estimated with the residuals from the within (fixed effects) estimator of model (2):

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^n \sum_{t=1}^{T_i} \left[(Y_{it} - Y_{i.}) - (X_{it} - X_{i.}) \tilde{\beta}^{fe} \right]^2}{n\bar{T} - (n + k_2)}.$$

Consistency of the two-stage *GLS* estimator is shown next. From (5) and (6) we can express the estimator as

$$\begin{aligned} \tilde{\gamma}^{fe} &= (\bar{Z}'\Omega^{-1}\bar{Z})^{-1} \bar{Z}'\Omega^{-1} (\bar{Z}\gamma + \eta) \\ &= \gamma + (\bar{Z}'\Omega^{-1}\bar{Z})^{-1} \bar{Z}'\Omega^{-1}\eta \\ &= \gamma + (\bar{Z}'\Omega^{-1}\bar{Z})^{-1} \bar{Z}'\Omega^{-1} \left[\bar{X} (\beta - \tilde{\beta}^{fe}) + \varepsilon \right]. \end{aligned}$$

Consistency depends on the time series observations approaching infinity so that $T_i \rightarrow \infty, \forall i \Rightarrow \bar{T} \rightarrow \infty$. Therefore,

$$\begin{aligned} p\lim \tilde{\gamma}^{fe} &= \gamma + \left[\lim_{\bar{T} \rightarrow \infty} \bar{T}^{-1} (\bar{Z}'\Omega^{-1}\bar{Z}) \right]^{-1} \lim_{\bar{T} \rightarrow \infty} \bar{T}^{-1} (\bar{Z}'\Omega^{-1}\bar{X}) p\lim_{\bar{T} \rightarrow \infty} (\beta - \tilde{\beta}^{fe}) \\ &\quad + \left[\lim_{\bar{T} \rightarrow \infty} \bar{T}^{-1} (\bar{Z}'\Omega^{-1}\bar{Z}) \right]^{-1} p\lim_{\bar{T} \rightarrow \infty} \bar{T}^{-1} (\bar{Z}'\Omega^{-1}\varepsilon) \\ &= \gamma \end{aligned}$$

assuming $\left[\lim_{\bar{T} \rightarrow \infty} \bar{T}^{-1} (\bar{Z}'\Omega^{-1}\bar{Z}) \right]^{-1}$ is finite and positive definite, $\lim_{\bar{T} \rightarrow \infty} \bar{T}^{-1} (\bar{Z}'\Omega^{-1}\bar{X})$ is finite, $\tilde{\beta}^{fe}$ is consistent, and $p\lim_{\bar{T} \rightarrow \infty} \bar{T}^{-1} (\bar{Z}'\Omega^{-1}\varepsilon) = 0$.

A straight forward F test can be used to test $\mathbf{H}_0: \alpha_i = \alpha + Z_{1i}\gamma^1$, $\mathbf{H}_1: \alpha_i \neq \alpha + Z_{1i}\gamma^1$:
(pooled OLS) (FE)

$$\frac{(\hat{\varepsilon}'_{ols} \hat{\varepsilon}_{ols} - \hat{\varepsilon}'_{fe} \hat{\varepsilon}_{fe}) / (n - k_1 - 1)}{\hat{\varepsilon}'_{fe} \hat{\varepsilon}_{fe} / (n\bar{T} - n - k_2)} \sim F_{(n-k_1-1), (n\bar{T}-n-k_2)}.$$

Endnotes

We gratefully acknowledge the helpful comments of Badi Baltagi, Greg Crawford, Alfonso Flores-Lagunes, Daniel Houser, and William Horrace.

REFERENCES

- Baltagi, B.H., 1995. *Econometric Analysis of Panel Data*, (John Wiley and Sons, Chichester)
- Greene, W. H, 2000. *Econometric Analysis*, 4th ed. (Prentice Hall, Upper Saddle River, NJ)
- Hausman, J. A., and W.E. Taylor, 1981. "Panel Data and Unobservable Individual Effects," *Econometrica*, 49, 1377-1398.
- Polacheck, S.W. and Kim, M., 1994. "Panel Estimates of the Gender Earnings Gap: Individual-Specific Intercept and Individual-Specific Slope Models," in *The Econometrics of Labor Market Segregation and Discrimination*, Shoshana Neuman and Jacques Silber, eds., *Journal of Econometrics*, 61, 23-42.