10.1 Motivation: The Omitted Variables Problem

We are interested in the partial effects of the observable explanatory variables $x_j$ in the population regression function

\[ E(y|x_1, x_2, \ldots, x_k, c) \]  \hspace{1cm} (10.1)

- An unobserved, time-constant variable is called an unobserved effect in panel data analysis.

- Such a data set is usually called a balanced panel because the same time periods are available for all cross section units.
10.2 Assumptions about the Unobserved Effects and Explanatory Variables

10.2.1 Random or Fixed Effects?

- The basic unobserved effects model (UEM) can be written, for a randomly drawn cross section observation \( i \), as

\[
y_{it} = x_{it}\beta + c_i + u_{it}, \quad t = 1, 2, ..., T
\]

- In addition to unobserved effect, there are many other names given to \( c_i \) in applications: unobserved component, latent variable, and unobserved heterogeneity are common.

- If \( i \) indexes individuals, then \( c_i \) is sometimes called an individual effect or individual heterogeneity; analogous terms apply to families, firms, cities, and other cross-sectional units.

- The \( u_{it} \) are called the idiosyncratic errors or idiosyncratic disturbances because these change across \( t \) as well as across \( i \).

- Nevertheless, later we will label two different estimation methods random effects estimation and fixed effects estimation.
10.2.2 Strict Exogeneity Assumptions on the Explanatory Variables

With an unobserved effect the most revealing form of the strict exogeneity assumption is

\[ E(y_{it}|x_{i1}, x_{i2}, ..., x_{iT}, c_i) = E(y_{it}|x_{it}, c_i) = x_{it} \beta + c_i \]  \hspace{1cm} (10.12)

\[ \text{• When assumption (10.12) holds, we say that the } x_{it} : t = 1, 2, ..., T \text{ are strictly exogenous conditional on the unobserved effect } c_i. \]

10.2.3 Some Examples of Unobserved Effects Panel Data Models
• Example 10.1 (Program Evaluation):

\[
\log(wage_{it}) = \theta_t + z_{it} \gamma + \delta_1 prog_{it} + c_i + u_{it} \tag{10.16}
\]

• Example 10.2 (Distributed Lag Model):

\[
patents_{it} = \theta_t + z_{it} \gamma + \delta_0 RD_{it} + \delta_1 RD_{i,t-1} + \cdots + \delta_5 RD_{i,t-5} + c_i + u_{it} \tag{10.17}
\]
• Example 10.3 (Lagged Dependent Variable):

\[
\log(wage_{it}) = \beta_1 \log(wage_{i,t-1}) + c_i + u_{it}, \ t = 1,2,\ldots, T
\]  

(10.18)

10.3 Estimating Unobserved Effects Models by Pooled OLS

Under certain assumptions, the pooled OLS estimator can be used to obtain a consistent estimator of \( \beta \) in model (10.11). Write the model as

\[
y_{it} = x_{it} \beta + v_{it}, \ t = 1,2,\ldots, T
\]

(10.21)

where \( v_{it} = c_i + u_{it} \), \( t = 1,2,\ldots T \) are the composite errors.
10.4 Random Effects Methods

10.4.1 Estimation and Inference under the Basic Random Effects Assumptions

• As with pooled OLS, a random effects analysis puts \( c_i \) into the error term.

• Assumption RE.1:
  1. \( E(u_{it}|x_i, c_i) = 0, \ t = 1, \ldots, T. \)
  2. \( E(c_i|x_i) = E(c_i) = 0 \)

where \( x_i \equiv (x_{i1}, x_{i2}, \ldots, x_{iT}) \).

• Assumption RE.2: \( \text{rank} \ E(X_i'\Omega^{-1}X_i) = K. \)

• When \( \Omega \) has the form \( (10.31) \), we say it has the random effects structure.

• Assumption RE.3:
  1. \( E(u_i'u'_i|x_i, c_i) = \sigma_u^2 I_T. \)
  2. \( E(c_i'^2|x_i) = \sigma_c^2 \)

• In a panel data context, the FGLS estimator that uses the variance matrix \( (10.33) \) is what is known as the random effects estimator:

\[
\hat{\beta}_{RE} = \left( \sum_{i=1}^{N} X_i'\hat{\Omega}^{-1}X_i \right)^{-1} \left( \sum_{i=1}^{N} X_i'\hat{\Omega}^{-1}y_i \right)
\]  

(10.34)

• Example 10.4(RE Estimation of the Effects of Job Training Grants):

\[
\log(\hat{\text{scrap}}) = .415 - .093 \ d88 - .270 \ d89 + .548 \ union - .215 \ grant - .377 \ grant_{-1}
\]

\[
(.243) \ (.109) \ (.132) \ (.411) \ (.148) \ (.205)
\]
10.4.3 A General FGLS Analysis

If the idiosyncratic errors $u_{it}: t = 1,2,\ldots,T$ are generally heteroskedastic and serially correlated across $t$, a more general estimator of $\Omega$ can be used in FGLS:

$$\hat{\Omega} = N^{-1} \sum_{i=1}^{N} \hat{\delta}_i \hat{\delta}_i'$$  \hspace{1cm} (10.38)
10.4.4 Testing for the Presence of an Unobserved Effect

From equation (10.37), we base a test of $H_0 : \sigma^2_c = 0$ on the null asymptotic distribution of

$$N^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} \hat{u}_{it} \hat{u}_{is}$$

which is essentially the estimator $\hat{\sigma}_c^2$ scaled up by $\sqrt{N}$.

10.5 Fixed Effects Methods

10.5.1 Consistency of the Fixed Effects Estimator

Again consider the linear unobserved effects model for $T$ time periods:

$$y_{it} = x_{it} \beta + c_i + u_{it}, \ t = 1, \ldots, T$$

(10.41)
• Assumption FE.1: $E(u_{it}|x_i, c_i) = 0, \ t = 1, 2, \ldots, T$.

• In this section we study the fixed effects transformation, also called the within transformation.

• The fixed effects (FE) estimator, denoted by $\hat{\beta}_{FE}$, is the pooled OLS estimator from the regression

$$\bar{y}_{it} \text{ on } \bar{x}_{it}, \ t = 1, 2, \ldots, T; i = 1, 2, \ldots, N$$

(10.48)

• This set of equations can be obtained by premultiplying equation (10.42) by a time-demeaning matrix.

• Assumption FE.2: $\text{rank} \left( \sum_{t=1}^{T} E(\bar{x}_{it}' \bar{x}_{it}) \right) = \text{rank} \left[ E(\bar{X}_i' \bar{X}_i) \right] = K$.

• It is also called the within estimator because it uses the time variation within each cross section.

• The between estimator, which uses only variation between the cross section observations, is the OLS estimator applied to the time-averaged equation (10.45).
10.5.2 Asymptotic Inference with Fixed Effects

- Assumption FE.3: \( E(u_i u_i' | x_i, c_i) = \sigma_u^2 I_T \).

- To see how to estimate \( \sigma_u^2 \), we use equation (10.51) summed across \( t: \sum_{t=1}^{T} E(\ddot{u}_{it}^2) = (T - 1)\sigma_u^2 \), and so \( [N(T - 1)]^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} E(\ddot{u}_{it}^2) = \sigma_u^2 \). Now, define the fixed effects residuals as

\[
\ddot{u}_{it} = \ddot{y}_{it} - \ddot{x}_{it} \hat{\beta}_{FE}, \quad t = 1, 2, \ldots, T; i = 1, 2, \ldots, N
\]

which are simply the OLS residuals from the pooled regression (10.48).

- Example 10.5(FE Estimation of the Effects of Job Training Grants):

\[
\log(\ddot{scrap}) = -0.80 \ dd88 - 0.247 \ dd89 - 0.252 \ grant - 0.422 \ grant_{-1} \\
\quad (0.109) \quad (0.133) \quad (0.151) \quad (0.210)
\]
10.5.3 The Dummy Variable Regression

- The estimator of $\beta$ obtained from regression (10.57) is, in fact, the fixed effects estimator. This is why $\hat{\beta}_{FE}$ is sometimes referred to as the dummy variable estimator.

10.5.4 Serial Correlation and the Robust Variance Matrix Estimator

Applying equation (7.26), the robust variance matrix estimator of $\hat{\beta}_{FE}$ is

$$\text{Var}(\hat{\beta}_{FE}) = (\bar{X}'\bar{X})^{-1} \left( \sum_{i=1}^{N} \bar{X}'_{i} \tilde{u}_{i} \tilde{u}_{i}' \bar{X}_{i} \right) (\bar{X}'\bar{X})^{-1}$$

(10.59)

which was suggested by Arellano (1987) and follows from the general results of White (1984, Chapter 6).
Example 10.5 (continued):

\[
\log(\hat{scrap}) = -.080 \, d88 - .247 \, d89 - .252 \, grant - .422 \, grant_{-1}
\]

\[
(\hat{.109}) \quad (\hat{.133}) \quad (\hat{.151}) \quad (\hat{.210})
\]

\[
[.096] \quad [.193] \quad [.140] \quad [.276]
\]

10.5.5 Fixed Effects GLS

- Assumption FEGLS.3: \( E(u_i u_i' | x_i, c_i) = \Lambda \), a \( T \times T \) positive definite matrix.

- The fixed effects GLS (FEGLS) estimator is defined by

\[
\hat{\beta}_{FEGLS} = (\sum_{i=1}^{N} \hat{X}_i' \hat{\Omega}^{-1} \hat{X}_i)^{-1} (\sum_{i=1}^{N} \hat{X}_i' \hat{\Omega}^{-1} \hat{y}_i)
\]

where \( \hat{X}_i \) and \( \hat{y}_i \) are defined with the last time period dropped.
• Assumption FEGLS.2: \( \text{rank}E(\hat{X}_i'\hat{\Omega}^{-1}\hat{X}_i) = K \).

10.5.6 Using Fixed Effects Estimation for Policy Analysis

Consider the model

\[ y_{it} = x_{it} \beta + u_{it} = z_{it} \gamma + \delta w_{it} + v_{it} \]

where \( v_{it} \) may or may not contain an unobserved effect.
10.6 First Differencing Methods

10.6.1 Inference

- Assumption FD.1: Same as Assumption FE.1.

Lagging the model (10.41) one period and subtracting gives

\[ \Delta y_{it} = \Delta x_{it} \beta + \Delta u_{it}, \quad t = 2, 3, \ldots, T \quad (10.63) \]

where \( \Delta y_{it} = y_{it} - y_{i,t-1}, \Delta x_{it} = x_{it} - x_{i,t-1}, \) and \( \Delta u_{it} = u_{it} - u_{i,t-1}. \)

- As with the FE transformation, this first-differencing transformation eliminates the unobserved effect \( c_i. \)

- The first-difference (FD) estimator, \( \hat{\beta}_{FD} \), is the pooled OLS estimator from the regression

\[ \Delta y_{it} on \Delta x_{it}, \quad t = 2, \ldots, T; i = 1, 2, \ldots, N \quad (10.65) \]

- Assumption FD.2: \( \text{rank} \left( \sum_{t=2}^{T} E(\Delta x'_{it} \Delta x_{it}) \right) = K. \)

- Assumption FD.3: \( E(e_i e'_i | x_{i1}, \ldots, x_{iT}, c_i) = \sigma_{e_i}^2 I_{T-1}, \) where \( e_i \) is the \((T - 1) \times 1\) vector containing \( e_{it}, t = 2, \ldots, T. \)
10.6.2 Robust Variance Matrix

If Assumption FD.3 is violated, then, as usual, we can compute a robust variance matrix. The estimator in equation (7.26) applied in this context is

\[
\text{Avâr}(\hat{\beta}_{FD}) = (\Delta X' \Delta X)^{-1} \left( \sum_{i=1}^{N} \Delta X_i' \hat{\epsilon}_i \hat{\epsilon}_i' \Delta X_i \right) (\Delta X' \Delta X)^{-1}
\] (10.70)

where \(\Delta X\) denotes the \(N(T-1) \times K\) matrix of stacked first differences of \(x_{it}\).

- Example 10.6 (FD Estimation of the Effects of Job Training Grants):

\[
\Delta \log(\text{scrap}_{it}) = \delta_1 + \delta_2 d89_t + \beta_1 \Delta \text{grant}_{it} + \beta_2 \Delta \text{grant}_{it-1} + \Delta u_{it}
\]
10.6.3 Testing for Serial Correlation

The regression is based on \( T - 2 \) time periods:

\[
\hat{e}_{it} = \hat{p}_1 \hat{e}_{i,t-1} + \text{error}_{it}, \; t = 3, 4, \ldots, T; i = 1, 2, \ldots, N \quad (10.71)
\]

- Example 10.6 (continued):

10.6.4 Policy Analysis Using First Differencing

In this one case, \( \Delta \text{prog}_i = \Delta \text{prog}_{i2} \), and the first-differenced equation can be written as

\[
\Delta y_{i2} = \theta_2 + \Delta z_{i2} \gamma + \delta_1 \text{prog}_{i2} + \Delta u_{i2} \quad (10.72)
\]
• When $\Delta z_{i2}$ is omitted, the estimate of $\delta_1$ from equation (10.72) is the difference-in-differences (DID) estimator (see Problem 10.4): $\hat{\Delta}_1 = \overline{\Delta y}_{treat} - \overline{\Delta y}_{control}$.

10.7 Comparison of Estimators

10.7.1 Fixed Effects versus First Differencing

With more than two time periods, a test of strict exogeneity is a test of $H_0: \gamma = 0$ in the expanded equation

$$\Delta y_t = \Delta x_t \beta + w_t \gamma + \Delta u_t, \ t = 2, \ldots, T$$

where $w_t$ is a subset of $x_t$ (that would exclude time dummies).
10.7.2 The Relationship between the Random Effects and Fixed Effects Estimators

Using the fact that $j_T'j_T = T$, we can write $\Omega$ under the random effects structure as

$$\Omega = \sigma_u^2 I_T + \sigma_c^2 j_T j_T' = \sigma_u^2 I_T + T \sigma_c^2 j_T (j_T'j_T)^{-1} j_T' = \sigma_u^2 I_T + T \sigma_c^2 P_T = (\sigma_u^2 + \sigma_c^2)(P_T + \eta Q_T)$$

where $P_T \equiv I_T - Q_T = j_T (j_T'j_T)^{-1} j_T'$ and $\eta \equiv \sigma_u^2 / (\sigma_u^2 + T \sigma_c^2)$.

- Equation (10.76) shows that the random effects estimator is obtained by a quasitime demeaning: rather than removing the time average from the explanatory and dependent variables at each $t$, random effects removes a fraction of the time average.

- Example 10.7 (Job Training Grants):

10.7.3 The Hausman Test Comparing the RE and FE Estimators