Chapter 1. Linear Panel Models and Heterogeneity

School of Economics and Management - University of Geneva

Christophe Hurlin, Université of Orléans

University of Orléans

February 2018

Introduction

The outline of this chapter is the following:

Section 1

Specification tests and analysis of covariance

1. Specification tests and analysis of covariance

Objectives

1. Define the concept of **homogeneous panel data model**.
2. Define the concept of **heterogeneous panel data model**.
3. Define the concept of **individual (unobserved) effects**.
4. Introduce the **specification tests** (Hsiao, 2003).
5. Propose an **empirical application** for the strike days in OECD.

1. Specification tests and analysis of covariance

Notations

Let us consider the following linear model

$$y_{it} = \alpha_{it} + \beta'_{it} x_{it} + \varepsilon_{it}$$

- $\forall\, i = 1, .., n$, $\forall\, t = 1, .., T$
- α_{it} is a scalar that varies across i and t.
- $\beta_{it} = (\beta_{1it}, \beta_{2it}, ..., \beta_{Kit})'$ is a $K \times 1$ vector of parameters that vary across i and t,
- $x_{it} = (x_{1it}, ..., x_{Kit})'$ is a $K \times 1$ vector of exogenous variables,
- ε_{it} is an error term.

1. Specification tests and analysis of covariance

Differentrestrictions on the regression coefficients can be tested:

1. the homogeneity of **regression slope coefficients**
2. the homogeneity of **regression intercept coefficients**
3. the **time stability of parameters** (slopes and constants). We will not consider this issue here (since it is not specific to panel data models).

1. Specification tests and analysis of covariance

Fact (Time stability)

We assume that the parameters are **constant over time** *(no structural break, no regime switching, etc.), but can vary across individuals.*

$$y_{it} = \alpha_i + \beta_i' x_{it} + \varepsilon_{it}$$

1. Specification tests and analysis of covariance

Three types of restrictions can be imposed on this model.

1. Regression slope coefficients are identical, and intercepts are not (**model with individual / unobserved effects**).

$$y_{it} = \alpha_i + \beta' x_{it} + \varepsilon_{it}$$

2. Regression intercepts are the same, and slope coefficients are not (**unusual)**.

$$y_{it} = \alpha + \beta_i' x_{it} + \varepsilon_{it}$$

3. Both slope and intercept coefficients are the same (**homogeneous / pooled panel)**.

$$y_{it} = \alpha + \beta' x_{it} + \varepsilon_{it}$$

1. Specification tests and analysis of covariance

Definition (Heterogeneous panel data model)

An **heterogeneous panel data model** is a model in which all parameters (constant and slope coefficients) vary accross individuals.

1. Specification tests and analysis of covariance

Definition (Homogeneous panel data model)

An **homogeneous panel data model** (or pooled model) is a model in which all parameters (constant and slope coefficients) are common

1. Specification tests and analysis of covariance

Definition (individual effects)

In a panel data model, the **individual (unobserved) effects** are captured by the constant terms α_i.

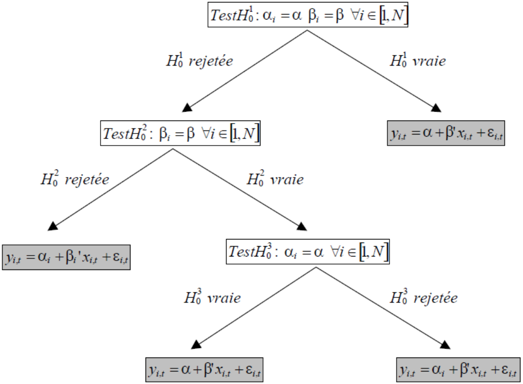
1. Specification tests and analysis of covariance

How to choose the appropriate specification of the panel data model?

- **Economic interpretation**: is it plausible to assume the homogeneity of the parameters across individuals?
- **Specification tests**: testing strategy proposed by Hsiao (2003) for instance.

Specification Tests

1. Specification tests and analysis of covariance



1. Specification tests and analysis of covariance

Lemma (Normality assumption)

Under the assumption that the ε_{it} are independently normally distributed over i and t with mean zero and variance σ_{ε}^2:

$$\varepsilon_{it} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_{\varepsilon}^2\right)$$

different Fisher F-tests can be used to test the restrictions on β and α.

1. Specification tests and analysis of covariance

First step (homogeneous/ pooled assumption)

Let us consider the general model

$$y_{it} = \alpha_i + \beta' x_{it} + \varepsilon_{it}$$

The hypothesis of common intercept and slope can be viewed as a set of $(K+1)(n-1)$ linear restrictions:

$$H_0^1 : \beta_i = \beta \quad \alpha_i = \alpha \quad \forall\, i \in \{1, ..., n\}$$

$$H_a^1 : \exists\, (i, j) \in \{1, ..., n\}^2 \; / \; \beta_i \neq \beta_j \text{ or } \alpha_i \neq \alpha_j$$

1. Specification tests and analysis of covariance

Consider the model

$$y_{it} = \alpha_i + \beta' x_{it} + \varepsilon_{it}$$

$$H_0^1 : \beta_i = \beta \;\; \alpha_i = \alpha \;\; \forall\, i \in \{1, ..., n\}$$

- Under the alternative H_1, there are nK estimated slope coefficients for the n vectors β_i $(K \times 1)$ and n estimated constants.
- Under H_1, the unrestricted residual sum of squares S_1 divided by σ_ε^2 has a chi-square distribution with $nT - n(K+1)$ degrees of freedom.

1. Specification tests and analysis of covariance

Definition (Homogeneity test)

Under the homogeneous assumption H_0^1,

$$H_0^1 : \underset{(K,1)}{\beta_i} = \underset{(K,1)}{\beta} \quad \alpha_i = \alpha \;\; \forall\, i \in \{1, ..., n\}$$

the F statistic, denoted F_1,and defined by:

$$F_1 = \frac{(RSS_{1,c} - RSS_1) \,/\, [(n-1)(K+1)]}{RSS_1 \,/\, [nT - n(K+1)]}$$

has a Fischer distribution with $(n-1)(K+1)$ and $nT - n(K+1)$ degrees of freedom. RSS_1 denotes the residual sum of squares of the model and $RSS_{1,c}$ the residual sum of squares of the constrained model

$$y_{it} = \alpha + \beta' x_{it} + \varepsilon_{it}$$

Remark 1

Under H_1, the residual sum of squares is equal to the sum of the n residual sum of squares associated to the n individual regressions:

$$RSS_1 = \sum_{i=1}^{n} RSS_{1,i} = \sum_{i=1}^{n} \widehat{\varepsilon}_{it}^2 = \sum_{i=1}^{n} \left[S_{yy,i} - S'_{xy,i} S_{xx,i}^{-1} S_{xy,i} \right]$$

$$S_{yy,i} = \sum_{t=1}^{T} \left(y_{it} - \overline{y}_i \right)^2 \quad \text{with} \quad \overline{x}_i = \frac{1}{T} \sum_{t=1}^{T} x_{it} \quad \text{and} \quad \overline{y}_i = \frac{1}{T} \sum_{t=1}^{T} y_{it}$$

$$S_{xx,i} = \sum_{t=1}^{T} \left(x_{it} - \overline{x}_i \right) \left(x_{it} - \overline{x}_i \right)' \qquad S_{xy,i} = \sum_{t=1}^{T} \left(x_{it} - \overline{x}_i \right) \left(y_{it} - \overline{y}_i \right)$$

1. Specification tests and analysis of covariance

Remark 2

Under H_0^1, the model becomes:

$$y_{it} = \alpha + \beta' x_{it} + \varepsilon_{it}$$

The least-squares regression of the pooled model yields parameter estimates

$$\widehat{\beta} = S_{xx}^{-1} S_{xy}$$

$$S_{xx} = \sum_{i=1}^{n} \sum_{t=1}^{T} \left(x_{it} - \overline{x}\right) \left(x_{it} - \overline{x}\right)' \quad \text{with} \quad \overline{x} = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} x_{it}$$

$$S_{xy} = \sum_{i=1}^{n} \sum_{t=1}^{T} \left(x_{it} - \overline{x}\right) \left(y_{it} - \overline{y}\right) \quad \text{with} \quad \overline{y} = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} y_{it}$$

1. Specification tests and analysis of covariance

Under H_0^1, the overall RSS is defined by

$$SCR_{1,c} = S_{yy} - S'_{xy} S_{xx}^{-1} S_{xy}$$

with

$$S_{yy} = \sum_{i=1}^{n} \sum_{t=1}^{T} \left(y_{it} - \overline{y}_i \right)^2$$

$$S_{xx,i} = \sum_{i=1}^{n} \sum_{t=1}^{T} \left(x_{it} - \overline{x}_i \right) \left(x_{it} - \overline{x}_i \right)'$$

$$S_{xy,i} = \sum_{i=1}^{n} \sum_{t=1}^{T} \left(x_{it} - \overline{x}_i \right) \left(y_{it} - \overline{y}_i \right)$$

1. Specification tests and analysis of covariance

Second step (individual/unobserved effects)

Let us consider the general model

$$y_{it} = \alpha_i + \beta_i' x_{it} + \varepsilon_{it}$$

The hypothesis of heterogeneous intercepts but homogeneous slopes can be reformulated as subject to $(n-1)K$ linear restrictions (no restrictions on α_i).

$$H_0^2 : \beta_i = \beta \quad \forall \, i = 1, ..n$$

1. Specification tests and analysis of covariance

Definition (Test for common slope parameters)

Under the assumption H_0^2,

$$H_0^2 : \beta_i = \beta \ \ \forall \, i = 1, ..n$$

the F statistic, denoted F_2, and defined by:

$$F_2 = \frac{\left(RSS_{1,c'} - RSS_1\right) / \left[(n-1)\, K\right]}{RSS_1 / \left[nT - n\,(K+1)\right]}$$

has a Fischer distribution with $(n-1)\,K$ et $nT - n\,(K+1)$ degrees of freedom under H_0^2. RSS_1 denotes the residual sum of squares of the model and $RSS_{1,c'}$ the residual sum of squares of the constrained model (model with individual effects):

$$y_{it} = \alpha_i + \beta' x_{it} + \varepsilon_{it}$$

1. Specification tests and analysis of covariance

Under H_0^2, the residual sum of squares is:

$$RSS_{1,c'} = \sum_{i=1}^{n} S_{yy,i} - \left(\sum_{i=1}^{n} S_{xy,i} \right)' \left(\sum_{i=1}^{n} S_{xx,i} \right)^{-1} \left(\sum_{i=1}^{n} S_{xy,i} \right)$$

$$S_{yy,i} = \sum_{t=1}^{T} \left(y_{it} - \overline{y}_i \right)^2 \quad \text{with} \quad \overline{x}_i = \frac{1}{T} \sum_{t=1}^{T} x_{it} \quad \overline{y}_i = \frac{1}{T} \sum_{t=1}^{T} y_{it}$$

$$S_{xx,i} = \sum_{t=1}^{T} \left(x_{it} - \overline{x}_i \right) \left(x_{it} - \overline{x}_i \right)'$$

$$S_{xy,i} = \sum_{t=1}^{T} \left(x_{it} - \overline{x}_i \right) \left(y_{it} - \overline{y}_i \right)$$

Third step: homogeneous constants

If H_0^2 is not rejected, one can also apply a conditional test for homogeneous intercepts ($n-1$ linear restrictions).

$$H_0^3 : \alpha_i = \alpha \quad \forall\, i = 1, .., n \text{ given } \beta_i = \beta$$

- Under the null, the model is homogeneous (pooled) and the restricted residual sum of squares is $SCR_{1,c}$.
- Under the alternative, the model is $y_{it} = \alpha_i + \beta' x_{it} + \varepsilon_{it}$, and there is $nT - K - n$ degrees of freedom

1. Specification tests and analysis of covariance

Definition (Test for homogeneous constant terms)

Under the assumption H_0^3,

$$H_0^3 : \alpha_i = \alpha \ \ \forall\, i = 1, .., n \quad \text{given} \quad \beta_i = \beta$$

the F statistic, denoted F_3, and defined by:

$$F_3 = \frac{\left(RSS_{1,c} - RSS_{1,c'}\right) / \left(n - 1\right)}{RSS_{1,c'} / \left[n\left(T - 1\right) - K\right]} \tag{1}$$

has a Fischer distribution with $n - 1$ and $n\left(T - 1\right) - K$ degrees of freedom under H_0^2. $RSS_{1,c'}$ denotes the residual sum of squares of the model with individual effects and $SCR_{1,c}$ the residual sum of squares of the pooled model previously defined.

Application: Strikes in OECD

1. Specification tests and analysis of covariance

Example (strikes in OECD countries)

Let us consider a simple panel regression model for the total number of strike days in OECD countries. We consider a balanced panel data set for 17 countries ($n = 17$) and annual data form 1951 to 1985 ($T = 35$).

General idea: evaluate the link between strikes and some macroeconomic factors (inflation, unemployment, etc..).

1. Specification tests and analysis of covariance

We consider the following model

$$s_{it} = \alpha_i + \beta_i u_{it} + \gamma_i p_{it} + \varepsilon_{it}$$

- s_{it} the number of strike days for 1,000 workers for the country i at time t.
- u_{it} the unemployment rate
- p_{it} the inflation rate

1. Specification tests and analysis of covariance

```
                          PANEL DATA ESTIMATION
                         =====================

Balanced data:  NI=   17, T=   35, NOB=   595

TOTAL (plain OLS) Estimates:

Dependent variable: SRT

 Mean of dependent variable = 305.076        Std. error of regression = 557.258
Std. dev. of dependent var. = 571.637                       R-squared = .052874
   Sum of squared residuals = .183838E+09           Adjusted R-squared = .049674
      Variance of residuals = 310536.

           Estimated    Standard
Variable  Coefficient    Error       t-statistic
U          27.4379      7.53997      3.63899
P          18.6136      4.98653      3.73277
C          95.0787      43.1414      2.20389

F test of A,B=Ai,Bi:  F(48,544) = 3.8320,  P-value = [.0000]
Critical F value for diffuse prior (Leamer, p.114) =   7.6418
```

Recall that, we have:

$$y_{it} = \alpha_i + \beta' x_{it} + \varepsilon_{it}$$

$$H_0^1 : \beta_i = \beta \;\; \alpha_i = \alpha \;\; \forall\, i \in \{1, ..., n\}$$

The test statistic satisfies

$$F_1 \underset{H_0^1}{\sim} F\,(48, 544)$$

since

$$(n-1)\,(K+1) = (17-1) \times (2+1) = 48$$

$$nT - n\,(K+1) = 595 - 17 \times (2+1) = 544$$

1. Specification tests and analysis of covariance

```
BETWEEN (OLS on means) Estimates:

Dependent variable: SRT

 Mean of dependent variable = 305.076       Std. error of regression = 189.663
Std. dev. of dependent var. = 278.196                      R-squared = .593303
   Sum of squared residuals = 503607.             Adjusted R-squared = .535203
      Variance of residuals = 35972.0

           Estimated    Standard
Variable   Coefficient   Error       t-statistic
U          80.9542      23.1650      3.49467
P          59.3882      32.6068      1.82134
C          -341.546     197.643      -1.72810

WITHIN (fixed effects) Estimates:

Dependent variable: SRT

Sum of squared residuals = .146958E+09                     R-squared = .242875
   Variance of residuals = 255136.                Adjusted R-squared = .219215
Std. error of regression = 505.110

           Estimated    Standard
Variable   Coefficient   Error       t-statistic
U          -21.5968     9.19158      -2.34963
P          16.2729      4.75658      3.42113

F test of Ai,B=Ai,Bi:  F(32,544) = 1.1845,  P-value = [.2266]
Critical F value for diffuse prior (Leamer, p.114) =   6.9699

F test of A,B=Ai,B:  F(16,576) = 9.0342,  P-value = [.0000]
Critical F value for diffuse prior (Leamer, p.114) =   6.7476
```

1. Specification tests and analysis of covariance

For the second test:

$$H_0^2 : \beta_i = \beta \;\; \forall\, i = 1, ..n$$

the F statistic, denoted F_2, has the following distribution

$$F_2 \underset{H_0^1}{\sim} F\,(32, 544)$$

since

$$(n-1)\,K = (17-1) \times 2 = 32$$

$$nT - n\,(K+1) = 595 - 17 \times (2+1) = 544$$

1. Specification tests and analysis of covariance

For the third test:

$$H_0^3 : \alpha_i = \alpha \ \ \forall\, i = 1, .., n \quad \text{given} \quad \beta_i = \beta$$

the F statistic, denoted F_3, satisfies:

$$F_3 \underset{H_0^1}{\sim} F\left(16, 576\right)$$

since

$$\left(n - 1\right) = 17 - 1 = 16$$

$$n\left(T - 1\right) - K = 17 \times \left(35 - 1\right) - 2 = 576$$

1. Specification tests and analysis of covariance

Should we use these specification tests?

- These heterogeneity / homogeneity tests of the parameters are valid under **specific assumptions** (normality of residuals).
- More generally, the assumption of heterogeneity / homogeneity of the parameters (slope coefficients and constants) has to be evaluated through an **economic reasoning**.

1. Specification tests and analysis of covariance

Example

It is reasonnable to assume that the slope parameters of the production function are the same accros countries? what does it imply? Should I impose a common mean for the TFP for France and Germany? The answer is probably no.

Key Concepts Section 1

1. Heterogeneous panel data model
2. Homogeneous panel data model
3. Individual (unobserved) effects
4. Specification tests (Hsiao, 2003)

Section 2

Linear Unobserved Effects Panel Data Models

Objectives

1. Define the concept of **linear unobserved effects** panel data model.
2. Define the concept of **individual effect**.
3. Write the linear model in a **vectorial form**.
4. Define the notion of **fixed effects**.
5. Define the notion of **random effects**.

2. Linear unobserved effects panel data models

Definition (linear unobserved effects panel data model)

A linear unobserved (individual) effects panel data model is defined as:

$$y_{it} = \alpha_i + \beta' x_{it} + \varepsilon_{it}$$

where α_i is a scalar, $\beta = (\beta_1, \beta_2, ..., \beta_K)'$ denotes a $K \times 1$ vector of parameters, $x_{it} = (x_{1it}, ..., x_{Kit})'$ is a $K \times 1$ vector of exogenous variables, and ε_{it} is an error term, assumed to be $i.i.d.$, with $\forall\, i = 1, .., n$, $\forall\, t = 1, .., T$

$$\mathbb{E}(\varepsilon_{it}) = 0 \quad \mathbb{E}(\varepsilon_{it}^2) = \sigma_\varepsilon^2$$

2. Linear unobserved effects panel data models

Definition (individual effects)

There are many names for the scalars α_i, $i = 1, ..., n$: (1) **unobserved effects**, (2) **individual effects**, (3) **unobserved components**, and (4) **latent variables** (for random effects models).

2. Linear unobserved effects panel data models

Definition (error terms)

The errors ε_{it} are called the **idiosyncratic errors** or **idiosyncratic disturbances**. They change accross t as well as accross i.

Vectorial form (1)

Let us denote

$$\underset{(T,1)}{y_i} = \begin{pmatrix} y_{i,1} \\ y_{i,2} \\ \dots \\ y_{it} \end{pmatrix} \qquad \underset{(T,K)}{X_i} = \begin{pmatrix} x_{1,i,1} & x_{2,i,1} & \dots & x_{K,i,1} \\ x_{1,i,2} & x_{2,i,2} & \dots & x_{K,i,2} \\ \dots & \dots & \dots & \dots \\ x_{1,it} & x_{2,it} & \dots & x_{K,it} \end{pmatrix}$$

Let us denote e a unit vector and ε_i the vector of errors:

$$\underset{(T,1)}{e} = \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} \qquad \underset{(T,1)}{\varepsilon_i} = \begin{pmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \\ \dots \\ \varepsilon_{it} \end{pmatrix} \qquad \underset{(K,1)}{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_K \end{pmatrix}$$

2. Linear unobserved effects panel data models

Definition (vectorial form)

For any individual $\forall\, i = 1, ..., n$, the **linear unobserved effects panel data model** can be defined as follows:

$$y_i = e\alpha_i + X_i\beta + \varepsilon_i$$

$$\mathbb{E}\left(\varepsilon_i\right) = 0$$

$$\mathbb{E}\left(\varepsilon_i \varepsilon_i'\right) = \sigma_\varepsilon^2 I_T$$

$$\mathbb{E}\left(\varepsilon_i \varepsilon_j'\right) = \underset{(T,T)}{0} \quad \text{if } i \neq j$$

Example (production function)

Let us consider the case of a Cobb Douglas production function in log, as defined previously, for the case $T = 3$ and $K = 2$. We have:

$$y_{it} = \alpha_i + \beta_k k_{it} + \beta_n n_{it} + \varepsilon_{it} \quad \forall\, i, \forall\, t \in \{1, 2, 3\}$$

or in a vectorial form for a country i as:

$$\underset{(3,1)}{y_i} = \underset{(3,1)}{e}\,\underset{(1,1)}{\alpha_i} + \underset{(3,2)}{X_i}\,\underset{(2,1)}{\beta} + \underset{(3,1)}{\varepsilon_i}$$

$$\begin{pmatrix} y_{i,1} \\ y_{i,2} \\ y_{i,3} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \alpha_i + \begin{pmatrix} k_{i,1} & n_{i,1} \\ k_{i,2} & n_{i,2} \\ k_{i,3} & n_{i,3} \end{pmatrix} \begin{pmatrix} \beta_k \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \\ \varepsilon_{i,3} \end{pmatrix}$$

Vectorial form (2)

It is also possible to stackle all these vectors/matrices as follows

$$Y = \widetilde{e}\widetilde{\alpha} + X\beta + \varepsilon$$

$$\underset{(Tn,1)}{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} \qquad \underset{(Tn,K)}{X} = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_n \end{pmatrix} \qquad \underset{(Tn,1)}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{pmatrix}$$

where 0_T is the null vector $(T, 1)$.

$$\underset{(Tn,n)}{\widetilde{e}} = I_n \otimes e = \begin{pmatrix} e & 0_T & \dots & 0_T \\ 0_T & e & \dots & 0_T \\ \dots & \dots & \dots & 0_T \\ 0_T & 0_T & \dots & e \end{pmatrix} \qquad \underset{(n,1)}{\widetilde{\alpha}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{pmatrix}$$

2. Linear unobserved effects panel data models

Example (production function)

Consider the case of the production function with $T = 3$ and $n = 2$

$$Y = \widetilde{e}\widetilde{\alpha} + X\beta + \varepsilon$$

$$\begin{pmatrix} y_{1,1} \\ y_{1,2} \\ y_{1,3} \\ y_{2,1} \\ y_{2,2} \\ y_{2,3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} k_{11} & n_{11} \\ k_{12} & n_{12} \\ k_{13} & n_{13} \\ k_{21} & n_{21} \\ k_{22} & n_{22} \\ k_{,3} & n_{23} \end{pmatrix} \begin{pmatrix} \beta_k \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{1,3} \\ \varepsilon_{2,1} \\ \varepsilon_{2,2} \\ \varepsilon_{2,3} \end{pmatrix}$$

2. Linear unobserved effects panel data models

Especially in methodological papers, but also in applications, one often sees a discussion about whether the individual effects α_i have to be treated as a **random effect** or a **fixed effect**.

2. Linear unobserved effects panel data models

Definition (Traditional approach)

In the traditional approach to panel data models, α_i is called a "**random effect"** when it is treated as a random variable and a **"fixed effect"** when it is treated as a parameter to be estimated for each cross section observation i.

Discussion

- For Wooldridge (2010), these discussions about whether the α_i should be treated as random variables or as parameters to be estimated are wrongheaded for microeconomic panels.

- With a large number of random draws from the cross section, **it almost always makes sense to treat the unobserved effects, α_i, as random draws from the population**, along with y_{it} and x_{it}.

- This approach is certainly appropriate from an omitted variables or neglected heterogeneity perspective.

2. Linear unobserved effects panel data models

Fact (Mundlak's approach)

As suggested by Mundlak (1978), the key issue involving α_i is whether or not it is uncorrelated with the observed explanatory variables x_{it}.

Mundlak Y. (1978), "On the Pooling of Time Series and Cross Section Data", *Econometrica*, 46, 69-85

2. Linear unobserved effects panel data models

Definition (a "modern" approach)

In a modern approach, "random effect" is synonymous with **zero correlation** between the observed explanatory variables and the unobserved (random) effect α_i:

$$cov\left(x_{it}, \alpha_i\right) = 0, \quad \forall t, \forall i$$

Remarks

- Actually, a stronger conditional mean **independence** assumption,

$$\mathbb{E}\left(\alpha_i \middle| x_{i1}, ..., x_{iT}\right) = 0$$

  is needed to fully justify statistical inference.

- In applied papers, when α_i is referred to an "individual random effect," then α_i is probably being assumed to be uncorrelated with the x_{it}.

2. Linear unobserved effects panel data models

Definition (Fixed effects)

In microeconometric applications, the term "**fixed effect**" does not usually mean that α_i is being treated as nonrandom; rather, it means that one is allowing for arbitrary correlation between the unobserved effect α_i and the observed explanatory variables x_{it}.

2. Linear unobserved effects panel data models

Remarks

- Wooldridge (2010) avoids referring to α_i as a random effect or a fixed effect. Instead, he refers to α_i as unobserved effect, **unobserved heterogeneity**, and so on.

- Nevertheless, later we will label two different estimation methods as **random effects estimation** and **fixed effects estimation** methods.

- This terminology is so ingrained that it is pointless to try to change it now.

Fixed or random effects?

- The economic interpretation of the individual effects generally allows to show that they are **probably correlated** to the explanatory variables.
- But, in case of doubt, it is possible to use a specification test (**Hausman's test, 1978**)

2. Linear unobserved effects panel data models

Example (Production function)

Let us consider the simple example of the Cobb Douglass production function.

$$y_{it} = \beta_i k_{it} + \gamma_i n_{it} + \alpha_i + v_{it}$$

In this case, α_i corresponds to the unobserved effect on TFP due to scountry specific omitted factor (climate, institutions, organization, etc..). In this case, we might expect that the the level of factors are positivily correlated with this component of TFP: the more a country is productive, the more it invests in capital for instance.

$$cov\left(\alpha_i, k_{it}\right) > 0 \qquad cov\left(\alpha_i, n_{it}\right) > 0$$

2. Linear unobserved effects panel data models

Example (Patents and R&D)

Hausman, Hall, and Griliches (1984) estimate (nonlinear) distributed lag models to study the relationship between patents awarded to a firm and current and past levels of R&D spending. A linear version of their model is:

$$patents_{it} = \theta_t + z_{it}\gamma + \delta_0 RD_{it} + \delta_1 RD_{it-1} + .. + \delta_5 RD_{it-5} + \alpha_i + v_{it}$$

where RD_{it} is spending on R&D for firm i at time t and z_{it} contains other explanatory variables. α_i is a firm heterogeneity term that may influence $patents_{it}$ and that may be correlated with current, past, and future R&D expenditures.

$$cov\left(\alpha_i, RD_{it-k}\right) \neq 0 \quad \forall k$$

2. Linear unobserved effects panel data models

Definition (Hausman's test)

The Hausman test (1978), is a test of the null hypothesis

$$cov\left(x_{it},\alpha_i\right)=0,\quad \forall\left(it\right)$$

and is generally presented as **a specification test** (fixed or random) for the **unobserved effects**.

Key Concepts Section 2

1. Linear unobserved effects panel data model.
2. Vectorial form of the linear panel data model
3. Individual effects.
4. Unobserved effects
5. Random effects.
6. Fixed effects.

Section 3

Fixed Effects Estimation Methods

3. Fixed effects estimation methods

Objectives

1. Specify the linear regression model with **fixed effects**.
2. Introduce the **LSDV (within)** estimator.
3. Define the **within** transformation.
4. Estimate the **slope parameters**.
5. Estimate the **fixed effects**.

3. Fixed effects estimation methods

Notations

Let us denote

$$\underset{(T,1)}{y_i} = \begin{pmatrix} y_{i,1} \\ y_{i,2} \\ \dots \\ y_{it} \end{pmatrix} \qquad \underset{(T,K)}{X_i} = \begin{pmatrix} x_{1,i,1} & x_{2,i,1} & \dots & x_{K,i,1} \\ x_{1,i,2} & x_{2,i,2} & \dots & x_{K,i,2} \\ \dots & \dots & \dots & \dots \\ x_{1,it} & x_{2,it} & \dots & x_{K,it} \end{pmatrix}$$

Let us denote e a unit vector and ε_i the vector of errors:

$$\underset{(T,1)}{e} = \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} \qquad \underset{(T,1)}{\varepsilon_i} = \begin{pmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \\ \dots \\ \varepsilon_{it} \end{pmatrix} \qquad \underset{(K,1)}{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_K \end{pmatrix}$$

3. Fixed effects estimation methods

We consider the fixed effects model:

$$y_i = e\alpha_i + X_i\beta + \varepsilon_i \quad \forall\, i = 1, .., n$$

where α_i is assumed to be a **constant term** or a random variable satisfying $\mathbb{E}\left(\alpha_i \mid x_{i1}, ..., x_{iT}\right) = 0$.

3. Fixed effects estimation methods

Assumptions (H1) *The errors terms* ε_{it} *are* $i.i.d.$ $\forall$ (it) with:

- $\mathbb{E}\left(\varepsilon_{it}\right) = 0$
- $\mathbb{E}\left(\varepsilon_{it}\varepsilon_{i,s}\right) = \begin{cases} \sigma_{\varepsilon}^{2} & t = s \\ 0 & \forall t \neq s \end{cases}$, *or* $\mathbb{E}\left(\varepsilon_{i}\varepsilon_{i}'\right) = \sigma_{\varepsilon}^{2} I_{T}$ *where* I_t *denotes the identity matrix* (T, T).
- $\mathbb{E}\left(\varepsilon_{it}\varepsilon_{j,s}\right) = 0, \forall j \neq i, \forall (t, s)$, or $\mathbb{E}\left(\varepsilon_{i}\varepsilon_{j}'\right) = 0$ *where* 0 *denotes the null matrix* (T, T).

3. Fixed effects estimation methods

Theorem

Under assumptions H_1, the ordinary-least-squares (OLS) estimator of β is the best linear unbiased estimator (BLUE).

3. Fixed effects estimation methods

Definition (LSDV estimator)

In this context, the OLS estimator $\widehat{\beta}$ is called the **least-squares dummy-variable (LSDV)** or **Fixed Effect (FE)** estimator, because the observed values of the variable for the coefficient α_i takes the form of dummy variables.

3. Fixed effects estimation methods

The OLS estimators of α_i and β and are obtained by minimizing

$$\begin{aligned}\left\{\widehat{\alpha}_i, \widehat{\beta}_{LSDV}\right\} &= \underset{\{\alpha_i,\beta\}_{i=1}^{n}}{\arg\min} \sum_{i=1}^{n} \varepsilon_i' \varepsilon_i \\ &= \sum_{i=1}^{n} \left(y_i - e\alpha_i - X_i\beta\right)' \left(y_i - e\alpha_i - X_i\beta\right)\end{aligned}$$

3. Fixed effects estimation methods

FOC1 (with respect to α_i) gives:

$$\widehat{\alpha}_i = \overline{y}_i - \widehat{\beta}'_{LSDV}\overline{x}_i$$

with

$$\overline{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it} \quad \overline{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}$$

Given the second FOC (with respect to β) and the previous result, we can derive the formula for $\widehat{\beta}_{LSDV}$.

Definition (LSDV estimator)

Under assumption H_1, the fixed effect estimator or **LSDV estimator** of β is defined by:

$$\begin{aligned}\widehat{\beta}_{LSDV} &= \left(\sum_{i=1}^{n} \sum_{t=1}^{T} (x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)' \right)^{-1} \\ &\quad \left(\sum_{i=1}^{n} \sum_{t=1}^{T} (x_{it} - \overline{x}_i)(y_{it} - \overline{y}_i) \right)\end{aligned}$$

3. Fixed effects estimation methods

Remarks

1. The computational procedure for estimating the slope parameters in this model does not require that the dummy variables for the individual (and/or time) effects actually be included in the matrix of explanatory variables.

2. We only need (1) the empirical **means of time-series observations separately for each cross-sectional unit**, (2) transform the observed variables by subtracting out these means, and (3) then apply the least squares method to the transformed data.

3. Fixed effects estimation methods

The foregoing procedure is equivalent to premultiplying the i^{th} equation

$$y_i = e\alpha_i + X_i\beta + \varepsilon_i$$

by a $T \times T$ idempotent (covariance) transformation matrix (**within operator**)

$$Q = I_T - \frac{1}{T}ee'$$

to "sweep out" the individual effect α_i so that individual observations are measured as deviations from individual means (over time).

3. Fixed effects estimation methods

$$\underset{(T,T)}{Q} = I_T - \frac{1}{T}ee' = \begin{pmatrix} 1-\frac{1}{T} & -\frac{1}{T} & ... & -\frac{1}{T} & -\frac{1}{T} \\ -\frac{1}{T} & 1-\frac{1}{T} & ... & -\frac{1}{T} & -\frac{1}{T} \\ .... & .... & ... & ... & ... \\ -\frac{1}{T} & -\frac{1}{T} & ... & 1-\frac{1}{T} & -\frac{1}{T} \\ -\frac{1}{T} & -\frac{1}{T} & ... & -\frac{1}{T} & 1-\frac{1}{T} \end{pmatrix}$$

3. Fixed effects estimation methods

Qy_i and QX_i correspond to the observations are measured as deviations from individual means :

$$
\begin{aligned}
Qy_i &= \left(I_T - \frac{1}{T} ee' \right) y_i \\
&= y_i - e \left(\frac{1}{T} e' y_i \right) \\
&= \begin{pmatrix} y_{i,1} \\ y_{i,2} \\ \dots \\ y_{it} \end{pmatrix} - \left(\frac{1}{T} \sum_{t=1}^{T} y_{it} \right) \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} \\
&= y_i - \overline{y}_i e
\end{aligned}
$$

3. Fixed effects estimation methods

$$
\begin{aligned}
QX_i &= X_i - \frac{1}{T} ee' X_i \\
&= \begin{pmatrix} x_{1,i,1} & x_{2,i,1} & \cdots & x_{K,i,1} \\ x_{1,i,2} & x_{2,i,2} & \cdots & x_{K,i,2} \\ \ldots & \ldots & \ldots & \ldots \\ x_{1,it} & x_{2,it} & \cdots & x_{K,it} \end{pmatrix} \\
& \quad - \frac{1}{T} \begin{pmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{pmatrix} \begin{pmatrix} \sum_{t=1}^{T} x_{1,it} & \sum_{t=1}^{T} x_{2,it} & \ldots & \sum_{t=1}^{T} x_{K,it} \end{pmatrix}
\end{aligned}
$$

3. Fixed effects estimation methods

Finally, when the transformation Q is applied to a vector of constant (or a **time invariant variable**), it lead to a null vector.

$$\begin{aligned} Qe &= \left(I_T - \frac{1}{T} ee' \right) e \\ &= e - \frac{1}{T} ee'e \\ &= e - e = 0 \end{aligned}$$

since

$$e'e = \begin{pmatrix} 1 & .. & 1 \end{pmatrix} \begin{pmatrix} 1 \\ .. \\ 1 \end{pmatrix} = T$$

3. Fixed effects estimation methods

So, we have:

$$y_i = e\alpha_i + X_i\beta + \varepsilon_i$$

$$\iff Qy_i = Qe\alpha_i + QX_i\beta + Q\varepsilon_i$$

$$\iff Qy_i = QX_i\beta + Q\varepsilon_i$$

3. Fixed effects estimation methods

Definition (Within - LSDV estimator)

Under assumption H_1, the fixed effect estimator or **LSDV estimator** or **Within estimator** of parameter β is defined by:

$$\widehat{\beta}_{LSDV} = \left(\sum_{i=1}^{n} X_i' Q X_i \right)^{-1} \left(\sum_{i=1}^{n} X_i' Q y_i \right)$$

where

$$Q = I_T - \frac{1}{T} e e'$$

3. Fixed effects estimation methods

Fact (Time-invariant regressors)

If the explanatory variables contain some time-invariant variables z_i, their coefficients cannot be estimated by LSDV, because the covariance transformation eliminates z_i.

$$Qz_i = \left(I_T - \frac{1}{T} ee' \right) z_i = z_i - \frac{1}{T} ee' z_i = z_i - \overline{z}_i e = 0_T$$

3. Fixed effects estimation methods

Example

Let us consider a simple panel regression model for the total number of strike days in OECD countries. We have a balanced panel data set for 17 countries ($n = 17$) and annual data form 1951 to 1985 ($T = 35$). General idea: evaluate the link between strikes and some macroeconomic factors (inflation, unemployment etc..)

3. Fixed effects estimation methods

We consider the following model

$$s_{it} = \alpha_i + \beta_i u_{it} + \gamma_i p_{it} + \varepsilon_{it}$$

- s_{it} the number of strike days for 1000 workers for the country i at time t.
- u_{it} the unemployement rate
- p_{it} the inflation rate

3. Fixed effects estimation methods

```
                    PANEL DATA ESTIMATION
                    =====================

Balanced data:  NI=   17, T=   35, NOB=   595

WITHIN (fixed effects) Estimates:

Dependent variable: SRT

Sum of squared residuals = .146958E+09                  R-squared = .242875
   Variance of residuals = 255136.             Adjusted R-squared = .219215
Std. error of regression = 505.110

          Estimated    Standard
Variable  Coefficient   Error        t-statistic
U         -21.5968     9.19158       -2.34963
P         16.2729      4.75658       3.42113
```

3. Fixed effects estimation methods

```
                                   Equation   1
                                   ============

                       Method of estimation = Ordinary Least Squares


Dependent variable: SRTC
Number of observations:  595

 Mean of dependent variable = -.159906E-07
Std. dev. of dependent var. = 503.791
   Sum of squared residuals = .146958E+09
      Variance of residuals = 247822.
   Std. error of regression = 497.817
                  R-squared = .025220
         Adjusted R-squared = .023576
    Durbin-Watson statistic = 1.98516
  F-statistic (zero slopes) = 15.3421
 Schwarz Bayes. Info. Crit. = 12.4386
 Log of likelihood function = -4538.36

           Estimated    Standard
Variable  Coefficient    Error        t-statistic
UC        -21.5968      9.05887       -2.38405
PC        16.2729       4.68791       3.47125


                       1             2             3             4
@FIXED             355.49000     -2.58225     801.98611     211.05529


                       5             6             7             8
@FIXED             387.83512     339.06634     57.54147     582.61959


                       9            10            11            12
@FIXED            1012.66380     114.22273     27.49515     164.23965


                      13            14            15            16
@FIXED              14.43655      14.93684     -44.85353     284.45232


                      17
@FIXED             495.87113
```

3. Fixed effects estimation methods

Theorem

The LSDV estimator $\widehat{\beta}$ is unbiased and consistent when either n, or T, or both tend to infinity.

$$\widehat{\beta}_{LSDV} \underset{nT \rightarrow \infty}{\overset{p}{\longrightarrow}} \beta$$

3. Fixed effects estimation methods

Theorem

The estimator for the unobserved effects $\widehat{\alpha}_i$, although unbiased, is consistent only when $T \rightarrow \infty$.

$$\widehat{\alpha}_i \xrightarrow[T \rightarrow \infty]{p} \alpha_i$$

3. Fixed effects estimation methods

Theorem

The asymptotic variance–covariance matrix of the LSDV estimator $\widehat{\beta}$ is given by:

$$\mathbb{V}\left(\widehat{\beta}_{LSDV}\right) = \sigma_{\varepsilon}^{2}\left(\sum_{i=1}^{n} X_i' Q X_i\right)^{-1}$$

3. Fixed effects estimation methods

Estimator of the asymptotic covariance matrix

$$\widehat{\mathbb{V}}\left(\widehat{\beta}_{LSDV}\right) = \widehat{\sigma}_{\varepsilon}^{2}\left(\sum_{i=1}^{n} X_i' Q X_i\right)^{-1}$$

with

$$\widehat{\sigma}_{\varepsilon}^{2} = \frac{1}{nT - K - n}\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{\varepsilon}_{it}^{2}$$

3. Fixed effects estimation methods

```
                    PANEL DATA ESTIMATION
                    =====================

Balanced data:  NI=    17, T=    35, NOB=    595

WITHIN (fixed effects) Estimates:

Dependent variable: SRT

Sum of squared residuals = .146958E+09                  R-squared = .242875
   Variance of residuals = 255136.             Adjusted R-squared = .219215
Std. error of regression = 505.110

           Estimated    Standard
Variable  Coefficient    Error        t-statistic
U         -21.5968      9.19158       -2.34963
P         16.2729       4.75658       3.42113
```

$$
\begin{aligned}
\widehat{\sigma}_{\varepsilon}^{2} &= \frac{1}{nT-K-n}\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{\varepsilon}_{it}^{2} \\
&= \frac{1}{595-2-17}\times 0.146958e^{09} \\
&= 505.1093
\end{aligned}
$$

3. Fixed effects estimation methods

```
                                 Equation   1
                                 ============

                Method of estimation = Ordinary Least Squares

Dependent variable: SRTC
Number of observations:   595

 Mean of dependent variable = -.159906E-07
Std. dev. of dependent var. = 503.791
   Sum of squared residuals = .146958E+09
      Variance of residuals = 247822.
   Std. error of regression = 497.817
                  R-squared = .025220
         Adjusted R-squared = .023576
     Durbin-Watson statistic = 1.98516
  F-statistic (zero slopes) = 15.3421
 Schwarz Bayes. Info. Crit. = 12.4386
 Log of likelihood function = -4538.36

          Estimated     Standard
Variable  Coefficient    Error         t-statistic
UC        -21.5968      9.05887        -2.38405
PC        16.2729       4.68791        3.47125
```

Be careful with a simple OLS method!

$$
\begin{aligned}
\widehat{\sigma}_{\varepsilon}^{2} &= \frac{1}{nT-K}\sum_{i=1}^{n}\sum_{t=1}^{T}\widehat{\varepsilon}_{it}^{2} \\
&= \frac{1}{595-2}\times 0.146958e^{09} \\
&= 497.8165
\end{aligned}
$$

as it does not take into account the correct number of constant terms.

Key Concepts Section 3

1. Linear unobserved effects panel data model.
2. Fixed effects and assumptions H_1.
3. LSDV or within estimator.
4. Within transformation
5. Properties of the LSDV estimator.
6. Asymptotic variance-covariance of the LSDV estimator.

Section 4

Random Effects Estimation Methods

4. Random effects estimation methods

Objectives

1. Specify the **error-component** model.
2. Define the **Generalized Least Squares (GLS)** estimator.
3. Define the **between** and **pooled** estimators.
4. Write the GLS estimator as a **weigthed average** of the LSDV and between estimators.
5. Study the **properties** of the GLS estimator..
6. Define the **feasible GLS** estimator.

4. Random effects estimation methods

Definition (error-component model)

The random specification of unobserved effects corresponds to a particular case of variance-component or **error-component model**, in which the error is assumed to consist of three components

$$y_{it} = \beta' x_{it} + \varepsilon_{it} \quad \forall \ (it)$$

$$\varepsilon_{it} = \alpha_i + \lambda_t + v_{it}$$

4. Random effects estimation methods

Terminology

$$\varepsilon_{it} = \alpha_i + \lambda_t + v_{it}$$

- α_i : **individual** (random) effect
- λ_t : **time** (random) effect
- v_{it} : **idiosyncratic error** term

4. Random effects estimation methods

Assumptions (H2) *The errors terms $\varepsilon_{it} = \alpha_i + \lambda_t + v_{it}$ are i.i.d.* $\forall$ (it) with:

- $\mathbb{E}(\alpha_i) = \mathbb{E}(\lambda_t) = \mathbb{E}(v_{it}) = 0$
- $\mathbb{E}(\alpha_i \lambda_t) = \mathbb{E}(\lambda_t v_{it}) = \mathbb{E}(\alpha_i v_{it}) = 0$
- $\mathbb{E}(\alpha_i \alpha_j) = \begin{cases} \sigma_\alpha^2 & i = j \\ 0 & \forall i \neq j \end{cases}$
- $\mathbb{E}(\lambda_t \lambda_s) = \begin{cases} \sigma_\lambda^2 & t = s \\ 0 & \forall t \neq s \end{cases}$
- $\mathbb{E}(v_{it} v_{j,s}) = \begin{cases} \sigma_v^2 & t = s,\ i = j \\ 0 & \forall t \neq s,\ \forall i \neq j \end{cases}$
- $\mathbb{E}(\alpha_i x_{it}') = \mathbb{E}(\lambda_t x_{it}') = \mathbb{E}(v_{it} x_{it}') = 0$

4. Random effects estimation methods

Remark

As suggested by Wooldridge (2001), the "fixed effect" specification can be viewed as a case in which α_i is a random parameter with

$$cov\left(\alpha_i, x_{it}'\right) \neq 0$$

whereas the "random effect model" correspond to the situation in which

$$cov\left(\alpha_i, x_{it}'\right) = 0$$

4. Random effects estimation methods

Definition (error-component model)

Under H2, the variance of y_{it} conditional on x_{it} is equal to:

$$\sigma^2_{y|x} = \sigma^2_{\varepsilon} = \sigma^2_{\alpha} + \sigma^2_{\lambda} + \sigma^2_{v}$$

4. Random effects estimation methods

Definition (centered individual effects)

If the individual effects α_i^* are supposed to have a **non zero mean**, with

$$\mathbb{E}(\alpha_i) = \mu$$

then we can defined individual effects $\alpha_i = \mu + \alpha_i^*$ with zero mean. The error-component model is then defined as:

$$y_{it} = \mu + \beta' x_{it} + \varepsilon_{it}$$

$$\varepsilon_{it} = \alpha_i^* + \lambda_t + v_{it}$$

4. Random effects estimation methods

Random coefficient model

In the sequel, for simplicity we do not introduce any **time effects** and consider a simple random effect model with

$$\varepsilon_{it} = \alpha_i + v_{it}$$

4. Random effects estimation methods

Vectorial form

The vectorial expression of the individual effects model is then defined as:

$$\underset{(T,1)}{y_i} = \underset{(T,K+1)}{\widetilde{X}_i}\underset{(K+1,1)}{\gamma} + \underset{(T,1)}{\varepsilon_i}$$

$$\underset{(T,1)}{\varepsilon_i} = \underset{(T,1)}{e}\underset{(1,1)}{\alpha_i} + \underset{(T,1)}{v_i}$$

$$\widetilde{X}_i = (e : X_i) \quad \text{and} \quad \gamma' = \left(\mu : \beta'\right)$$

4. Random effects estimation methods

Definition (variance-covariance matrix of errors)

Under assumptions H_2, the **variance-covariance matrix** of ε_i is equal to:

$$V = \mathbb{E}\left(\varepsilon_i \varepsilon_i'\right) = \mathbb{E}\left(\left(\alpha_i e + v_i\right)\left(\alpha_i e + v_i\right)'\right) = \sigma_\alpha^2 ee' + \sigma_v^2 I_T$$

Its inverse is:

$$V^{-1} = \frac{1}{\sigma_v^2}\left(I_T - \left(\frac{\sigma_\alpha^2}{\sigma_v^2 + T\sigma_\alpha^2}\right) ee'\right)$$

Remark

The presence of α_i produces a correlation among errors of the same cross-sectional unit (**autocorrelation**) as

$$V = \mathbb{E}\left(\varepsilon_i \varepsilon_i'\right) = \sigma_\alpha^2 ee' + \sigma_v^2 I_T$$

$$\underset{(T,T)}{V} = \begin{pmatrix} \sigma_\alpha^2 + \sigma_v^2 & \sigma_\alpha^2 & \ldots & \sigma_\alpha^2 \\ & \sigma_\alpha^2 + \sigma_v^2 & \ldots & \sigma_\alpha^2 \\ & & \ldots & \sigma_\alpha^2 \\ & & & \sigma_\alpha^2 + \sigma_v^2 \end{pmatrix}$$

Remark

If we consider the $nT \times 1$ vector of errors $\varepsilon = \left(\varepsilon_1', ..., \varepsilon_n'\right)'$, we have

$$\mathbb{V}\left(\varepsilon\right) = \mathbb{E}\left(\varepsilon\varepsilon'\right) = V \otimes I_n$$

$$\underset{(nT,nT)}{\mathbb{V}\left(\varepsilon\right)} = \begin{pmatrix} \underset{(T,T)}{V} & \underset{(T,T)}{0} & \dots & \underset{(T,T)}{0} \\ & \underset{(T,T)}{V} & \dots & \underset{(T,T)}{0} \\ & & \dots & \underset{(T,T)}{0} \\ & & & \underset{(T,T)}{V} \end{pmatrix}$$

4. Random effects estimation methods

Within transformation

Regardless of whether the α_i are treated as fixed or as random, the individual-specific effects for a given sample can be swept out by the idempotent (covariance) transformation matrix Q

$$Qy_i = Qe\mu + QX_i\beta + Qe\alpha_i + Qv_i$$

Since $Qe = \left(I_T - T^{-1}ee'\right)e = 0$, we have

$$Qy_i = QX_i\beta + Qv_i$$

4. Random effects estimation methods

Theorem

Under assumptions H_2, when α_i are treated as random, the LSDV estimator is unbiased and consistent either n, or T, or both tend to infinity. However, whereas the LSDV is the BLUE under the assumption that α_i are fixed constants, it is not the BLUE when α_i are assumed random. The BLUE in the latter case is the **Generalized-Least-Squares (GLS)** *estimator.*

Summary

Assumptions	LSDV	GLS
$H_2 + \mathbb{E}\left(\alpha_i \mid x_{i1}, .., x_{iK}\right) = 0$	Unbiased	BLUE
$H_2 + \mathbb{E}\left(\alpha_i \mid x_{i1}, .., x_{iK}\right) \neq 0$	BLUE	Biased

4. Random effects estimation methods

Notations

- Let us consider the model

$$y_i = \widetilde{X}_i \gamma + \varepsilon_i \qquad \forall\, i = 1, .., n$$

  where $\varepsilon_i = \alpha_i e + v_i$, $\widetilde{X}_i = (e : X_i)$ and $\gamma' = \left(\mu : \beta'\right)$.

- We assume that the **variance covariance matrix** $V = \mathbb{E}\left(\varepsilon_i \varepsilon_i'\right)$ is **known**.

4. Random effects estimation methods

Definition (GLS estimator)

If the variance covariance matrix V is known, the **GLS estimator** *of the* γ vector, denoted $\widehat{\gamma}_{GLS}$, *is defined by:*

$$\widehat{\gamma}_{GLS} = \left(\sum_{i=1}^{n} \widetilde{X}_i' V^{-1} \widetilde{X}_i \right)^{-1} \left(\sum_{i=1}^{n} \widetilde{X}_i' V^{-1} y_i \right)$$

Under assumptions H_2, this estimaor is BLUE.

Definition (inverse of the variance-covariance matrix)

Following Maddala (1971), we can write V^{-1} as:

$$V^{-1} = \frac{1}{\sigma_v^2}\left(Q + \psi \frac{1}{T} ee'\right)$$

where $Q = \left(I_T - ee' / T\right)$ *and where the parameter* ψ *is defined by:*

$$\psi = \left(\frac{\sigma_v^2}{\sigma_v^2 + T\sigma_\alpha^2}\right)$$

4. Random effects estimation methods

Proof

$$
\begin{aligned}
V^{-1}V &= \frac{1}{\sigma_v^2}\left(Q+\psi\frac{1}{T}ee'\right)\left(\sigma_\alpha^2 ee' + \sigma_v^2 I_T\right) \\
&= \frac{1}{\sigma_v^2}\left(\sigma_\alpha^2 Qee' + \sigma_v^2 Q + \psi\frac{\sigma_\alpha^2}{T}ee'ee' + \psi\frac{\sigma_v^2}{T}ee'\right) \\
&= \frac{1}{\sigma_v^2}\left(\sigma_v^2 Q + \psi\sigma_\alpha^2 ee' + \psi\frac{\sigma_v^2}{T}ee'\right) \quad \text{as } e'e = T \\
&= \frac{1}{\sigma_v^2}\left(\sigma_v^2 Q + \frac{1}{T}ee'\psi\left(T\sigma_\alpha^2+\sigma_v^2\right)\right) \\
&= I_T - \frac{1}{T}ee' + \frac{1}{T}ee'\psi\left(\frac{T\sigma_\alpha^2+\sigma_v^2}{\sigma_v^2}\right)
\end{aligned}
$$

4. Random effects estimation methods

Proof (ct'd)

$$V^{-1}V = I_T - \frac{1}{T}ee' + \frac{1}{T}ee'\psi\left(\frac{T\sigma_\alpha^2 + \sigma_v^2}{\sigma_v^2}\right)$$

Since

$$\psi = \left(\frac{\sigma_v^2}{\sigma_v^2 + T\sigma_\alpha^2}\right)$$

we have

$$V^{-1}V = I_T - \frac{1}{T}ee' + \frac{1}{T}ee' = I_T \quad \square$$

4. Random effects estimation methods

Given this definition of V^{-1}, we have:

$$\begin{aligned}\widehat{\gamma}_{GLS} &= \left(\sum_{i=1}^{n}\widetilde{X}_i'\left(Q+\psi\frac{1}{T}ee'\right)\widetilde{X}_i\right)^{-1}\left(\sum_{i=1}^{n}\widetilde{X}_i'\left(Q+\psi\frac{1}{T}ee'\right)y_i\right)\\ &= \left(\sum_{i=1}^{n}\widetilde{X}_i'Q\widetilde{X}_i+\psi\frac{1}{T}\sum_{i=1}^{n}\widetilde{X}_i'ee'\widetilde{X}_i\right)^{-1}\left(\sum_{i=1}^{n}\widetilde{X}_i'Qy_i+\psi\frac{1}{T}\sum_{i=1}^{n}\widetilde{X}_i'ee'y_i\right)\end{aligned}$$

with $\widetilde{X}_i = (e\ X_i)$ and $\gamma' = (\mu\ \beta')$

4. Random effects estimation methods

It is possible to show that

$$\begin{pmatrix} \widehat{\mu}_{GLS} \\ \widehat{\beta}_{GLS} \end{pmatrix} = \begin{pmatrix} \psi nT & \psi T \sum\limits_{i=1}^{n} \overline{x}_i' \\ \psi T \sum\limits_{i=1}^{n} \overline{x}_i & \sum\limits_{i=1}^{n} X_i' Q X_i + \psi T \sum\limits_{i=1}^{n} \overline{x}_i \overline{x}_i' \end{pmatrix}^{-1} \begin{pmatrix} \psi nT\overline{y} \\ \sum\limits_{i=1}^{n} X_i' Q y_i + \psi T \sum\limits_{i=1}^{n} \overline{x}_i \overline{y}_i \end{pmatrix}$$

Using the formula of the partitioned inverse, we can derive $\widehat{\beta}_{GLS}$.

Definition (GLS estimator)

If the variance covariance matrix V is known, the **GLS estimator** of β is:

$$\widehat{\beta}_{GLS} = \left(\frac{1}{T} \sum_{i=1}^{n} X_i' Q X_i + \psi \sum_{i=1}^{n} (\overline{x}_i - \overline{x}) (\overline{x}_i - \overline{x})' \right)^{-1} \left(\frac{1}{T} \sum_{i=1}^{n} X_i' Q y_i + \psi \sum_{i=1}^{n} (\overline{x}_i - \overline{x}) (\overline{y}_i - \overline{y}) \right)$$

with $\psi = \sigma_v^2 \left(\sigma_v^2 + T \sigma_\alpha^2 \right)^{-1}$

The GLS estimator can be expressed as a weigthed average of the **LSDV (OLS) estimator** and the **between estimator**:

$$\begin{aligned}\widehat{\beta}_{GLS} &= \left(\frac{1}{T}\sum_{i=1}^{n} X_i' Q X_i + \psi \sum_{i=1}^{n} \left(\overline{x}_i - \overline{x}\right)\left(\overline{x}_i - \overline{x}\right)'\right)^{-1} \\ &\qquad \left(\frac{1}{T}\sum_{i=1}^{n} X_i' Q y_i + \psi \sum_{i=1}^{n} \left(\overline{x}_i - \overline{x}\right)\left(\overline{y}_i - \overline{y}\right)\right)\end{aligned}$$

4. Random effects estimation methods

Definition (between group estimator)

The between-group estimator or **between estimator** $\widehat{\beta}_{BE}$ corresponds to the *OLS estimator obtained in the model:*

$$\overline{y}_i = c + \beta' \overline{x}_i + \varepsilon_i \quad \forall\, i = 1, .., n$$

$$\widehat{\beta}_{BE} = \left(\sum_{i=1}^{n} \left(\overline{x}_i - \overline{x} \right) \left(\overline{x}_i - \overline{x} \right)' \right)^{-1} \left(\sum_{i=1}^{n} \left(\overline{x}_i - \overline{x} \right) \left(\overline{y}_i - \overline{y} \right) \right)$$

The estimator $\widehat{\beta}_{BE}$ is called the between-group estimator because it ignores variation within the group

4. Random effects estimation methods

Definition (pooled estimator)

The **pooled estimator** $\widehat{\beta}_{pooled}$ corresponds to the *OLS estimator obtained in the pooled model:*

$$y_{it} = \alpha + \beta' x_{it} + \varepsilon_{it} \quad \forall\, i = 1, .., n \ \ \forall\, t = 1, .., T$$

$$\widehat{\beta}_{pooled} = \left(\sum_{t=1}^{T} \sum_{i=1}^{n} \left(x_{it} - \overline{x} \right) \left(\overline{x}_i - \overline{x} \right)' \right)^{-1} \left(\sum_{t=1}^{T} \sum_{i=1}^{n} \left(x_{it} - \overline{x} \right) \left(y_{it} - \overline{y} \right) \right)$$

4. Random effects estimation methods

Theorem

Under assumptions H_2, the GLS estimator $\widehat{\beta}_{GLS}$ is a weighted average of the between-group $\widehat{\beta}_{BE}$ and the within-group (LSDV) estimators $\widehat{\beta}_{LSDV}$.

$$\widehat{\beta}_{GLS} = \Delta \widehat{\beta}_{BE} + (I_K - \Delta)\, \widehat{\beta}_{LSDV}$$

where Δ denotes a weigth matrix defined by:

$$\Delta = \psi T \left(\sum_{i=1}^{n} X_i' Q X_i + \psi T \sum_{i=1}^{n} (\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})' \right)^{-1} \left(\sum_{i=1}^{n} (\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})' \right)$$

4. Random effects estimation methods

GLS estimator properties

1. If $\psi \rightarrow 0$, the GLS estimator converges to LSDV estimator.

$$\widehat{\beta}_{GLS} \xrightarrow[\psi \rightarrow 0]{p} \widehat{\beta}_{LSDV}$$

2. If $\psi \rightarrow 1$, then GLS converges to the OLS pooled estimator.

$$\widehat{\beta}_{GLS} \xrightarrow[\psi \rightarrow 1]{p} \widehat{\beta}_{pooled}$$

4. Random effects estimation methods

Proof: case 1

$$\widehat{\beta}_{GLS} = \Delta \widehat{\beta}_{BE} + (I_K - \Delta) \widehat{\beta}_{LSDV}$$

$$\Delta = \psi T \left(\sum_{i=1}^{n} X_i' Q X_i + \psi T \sum_{i=1}^{n} (\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})' \right)^{-1} \left(\sum_{i=1}^{n} (\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})' \right)$$

Consider the case $\psi = 0$ then

$$\Delta = 0 \qquad \widehat{\beta}_{GLS} = \widehat{\beta}_{LSDV}$$

So, if $\psi \rightarrow 0$, the GLS estimator converges to LSDV estimator.

$$\widehat{\beta}_{GLS} \underset{\psi \rightarrow 0}{\overset{p}{\longrightarrow}} \widehat{\beta}_{LSDV}$$

4. Random effects estimation methods

Proof: case 2

$$\widehat{\beta}_{GLS} = \Delta\widehat{\beta}_{BE} + (I_K - \Delta)\,\widehat{\beta}_{LSDV}$$

Consider the case $\psi = 1$ we have

$$\Delta = T\left(\sum_{i=1}^{n} X_i' Q X_i + T\sum_{i=1}^{n}(\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})'\right)^{-1}\left(\sum_{i=1}^{n}(\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})'\right)$$

$$\widehat{\beta}_{BE} = \left(\sum_{i=1}^{n}(\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})'\right)^{-1}\left(\sum_{i=1}^{n}(\overline{x}_i - \overline{x})(\overline{y}_i - \overline{y})\right)$$

$$\widehat{\beta}_{LSDV} = \left(\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)'\right)^{-1}\left(\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \overline{x}_i)(y_{it} - \overline{y}_i)\right)$$

4. Random effects estimation methods

Proof: case 2

$$
\begin{aligned}
&\widehat{\beta}_{GLS} = \Delta\widehat{\beta}_{BE} + (I_K - \Delta)\,\widehat{\beta}_{LSDV} \\
= \quad & T\left(\sum_{i=1}^{n} X_i' Q X_i + T\sum_{i=1}^{n}(\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})'\right)^{-1}\left(\sum_{i=1}^{n}(\overline{x}_i - \overline{x})(\overline{y}_i - \overline{y})\right) \\
& + \left(\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)'\right)^{-1}\left(\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \overline{x}_i)(y_{it} - \overline{y}_i)\right) \\
& - T\left(\sum_{i=1}^{n} X_i' Q X_i + T\sum_{i=1}^{n}(\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})'\right)^{-1}\left(\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \overline{x}_i)(y_{it} -
\end{aligned}
$$

4. Random effects estimation methods

Proof: case 2

So, if $\psi = 1$ we have

$$\begin{aligned} \widehat{\beta}_{GLS} &= \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \left(x_{it} - \overline{x} \right) \left(x_{it} - \overline{x} \right)' \right)^{-1} \\ & \qquad \left(\sum_{i=1}^{n} \sum_{t=1}^{T} \left(x_{it} - \overline{x} \right) \left(y_{it} - \overline{y} \right) \right) \\ &= \widehat{\beta}_{pooled} \end{aligned}$$

So, if $\psi \rightarrow 1$, the GLS estimator converges to the OLS pooled estimator.

$$\widehat{\beta}_{GLS} \xrightarrow[\psi \rightarrow 1]{p} \widehat{\beta}_{pooled}$$

4. Random effects estimation methods

GLS estimator properties

The parameter $\psi = \sigma_v^2 \left(\sigma_v^2 + T\sigma_\alpha^2\right)^{-1}$ measures the weight given to the between-group variation.

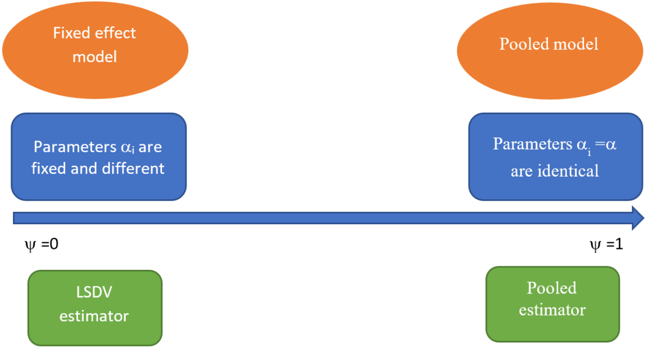
- In the LSDV (or fixed-effects model) procedure, this source of variation is completely ignored $(\psi = 0)$.
- The OLS procedure (pooled model) corresponds to $\psi = 1$. The between-group and within-group variations are just added up.

4. Random effects estimation methods

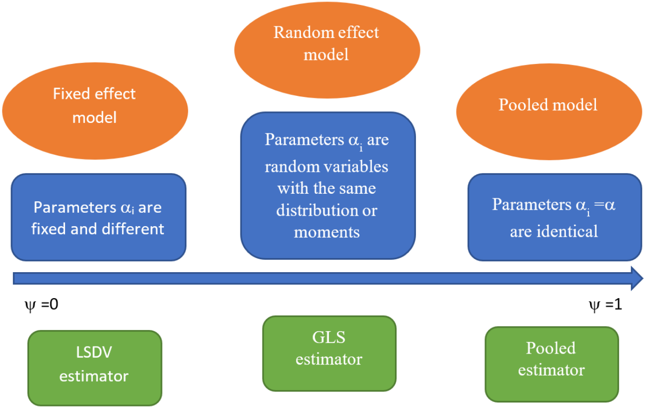
Fact

The procedure of treating α_i as random coefficients provides an **intermediate solution** *between treating them all as different (fixed effects, LSDV) and treating them all as equal (pooled model).*

4. Random effects estimation methods



4. Random effects estimation methods



GLS estimator properties

Given the definition of ψ, we have:

$$\lim_{T \to \infty} \psi = \lim_{T \to \infty} \left(\frac{\sigma_v^2}{\sigma_v^2 + T\sigma_\alpha^2} \right) = 0$$

4. Random effects estimation methods

Theorem (GLS and LSDV)

When T tends to infinity, the GLS estimator converges to the LSDV estimator:

$$\widehat{\beta}_{GLS} \underset{T \to \infty}{\longrightarrow} \widehat{\beta}_{LSDV}$$

4. Random effects estimation methods

Interpretation

$$\widehat{\beta}_{GLS} \underset{T\rightarrow\infty}{\longrightarrow} \widehat{\beta}_{LSDV}$$

- When $T \rightarrow \infty$, we have an infinite number of observations for each i.
- Therefore, we can consider each α_i as a random variable which has been drawn once and forever
- For each i we assume that they are just like fixed parameters.

4. Random effects estimation methods

Definition (Transformation matrix)

Computation of the GLS estimator can be simplified by introducing a **transformation matrix** P such that

$$P = \left(I_T - \left(1 - \psi^{1/2} \right) (1/T)\, ee' \right)$$

We have

$$V^{-1} = \frac{1}{\sigma_v^2} P' P$$

Premultiplying the model by the transformation matrix P, we obtain the GLS estimator by applying the least-squares method to the transformed model (Theil (1971, Chapter 6)).

4. Random effects estimation methods

Transformation matrix

The GLS estimator is equivalent to

1. Transforming the data by subtracting a fraction $\left(1-\psi^{1/2}\right)$ of individual means $\overline{y}_i$ and $\overline{x}_i$ from their corresponding y_{it} and x_{it}
2. Regressing $y_{it}-\left(1-\psi^{1/2}\right)\overline{y}_i$ on a constant and $x_{it}-\left(1-\psi^{1/2}\right)\overline{x}_i$ using simple OLS.

4. Random effects estimation methods

Definition (Asymptotic variance covariance matrix)

Under assumptions H2, the asymptotic **variance covariance matrix** of the GLS estimator is given by:

$$\mathbb{V}\left(\widehat{\beta}_{GLS}\right)=\sigma_v^2\left(\sum_{i=1}^{n}X_i'QX_i+\psi T\sum_{i=1}^{n}\left(\overline{x}_i-\overline{x}\right)\left(\overline{x}_i-\overline{x}\right)'\right)^{-1}$$

4. Random effects estimation methods

Remark

$$\mathbb{V}\left(\widehat{\beta}_{GLS}\right) = \sigma_v^2 \left(\sum_{i=1}^{n} X_i' Q X_i + \psi T \sum_{i=1}^{n} \left(\overline{x}_i - \overline{x}\right)\left(\overline{x}_i - \overline{x}\right)'\right)^{-1}$$

$$\mathbb{V}\left(\widehat{\beta}_{LSDV}\right) = \sigma_\varepsilon^2 \left(\sum_{i=1}^{n} X_i' Q X_i\right)^{-1}$$

As $\psi > 0$, the difference between the covariance matrices of $\widehat{\beta}_{LSDV}$ and $\widehat{\beta}_{GLS}$ is a positive semidefinite matrix. For $K = 1$, we have:

$$\mathbb{V}\left(\widehat{\beta}_{GLS}\right) \leq \mathbb{V}\left(\widehat{\beta}_{LSDV}\right)$$

=> the **LSDV is not BLUE**

4. Random effects estimation methods

Definition (feasible GLS)

If the variance components σ_{ε}^2 and σ_{α}^2 are unknown, we can use a two-step GLS estimation procedure, called as **feasible GLS**.

1. In the first step, we estimate the variance components using some consistent estimators.
2. In the second step, we substitute their estimated values into

$$\widehat{\gamma}_{GLS} = \left(\sum_{i=1}^{n} \widetilde{X}_i' \widehat{V}^{-1} \widetilde{X}_i \right)^{-1} \left(\sum_{i=1}^{n} \widetilde{X}_i' \widehat{V}^{-1} y_i \right)$$

or its equivalent form.

Two-step GLS estimator

Define $\overline{y}_i = \alpha_i + \beta' \overline{x}_i + \overline{\varepsilon}_i$ and $(y_{it} - \overline{y}_i) = (x_{it} - \overline{x}_i) + (v_{it} - \overline{v}_i)$, we can use the within and between-group residuals to estimate σ_{ε}^2 and σ_{α}^2 by

$$\widehat{\sigma}_v^2 = \frac{\sum_{i=1}^{n} \sum_{t=1}^{T} \left((y_{it} - \overline{y}_i) - \widehat{\beta}'_{LSDV} (x_{it} - \overline{x}_i) \right)^2}{n(T-1) - K}$$

$$\widehat{\sigma}_{\alpha}^2 = \frac{\sum_{i=1}^{n} \left(\overline{y}_i - \widehat{\beta}'_{LSDV} \overline{x}_i \right)^2}{n - K - 1} - \widehat{\sigma}_v^2$$

4. Random effects estimation methods

Then, we have an estimate of ψ and V^{-1}

$$\widehat{\psi} = \frac{\widehat{\sigma}_v^2}{\widehat{\sigma}_v^2 + T\widehat{\sigma}_\alpha^2}$$

$$\widehat{V}^{-1} = \frac{1}{\widehat{\sigma}_v^2}\left(Q + \widehat{\psi}\frac{1}{T}ee'\right)$$

4. Random effects estimation methods

Lemma

When the sample size is large (in the sense of either $n \longrightarrow \infty$,or $T \longrightarrow \infty$), the two-step GLS estimator will have the same asymptotic efficiency as the GLS procedure with known variance components.

4. Random effects estimation methods

Lemma

Even for moderate sample size (for $T \geq 3$, $n - (K + 1) \geq 9$; for $T \geq 2$, $n - (K + 1) \geq 10$), the two-step procedure is still more efficient than the LSDV estimator in the sense that the difference between the covariance matrices of the covariance estimator and the two-step estimator is nonnegative definite.

4. Random effects estimation methods

Example

Let us consider a simple panel regression model for the total number of strikes days in OECD countries. We have a balanced panel data set for 17 countries ($n = 17$) and annual data form 1951 to 1985 ($T = 35$).

$$s_{it} = \alpha_i + \beta u_{it} + \gamma p_{it} + \varepsilon_{it}$$

4. Random effects estimation methods

Figure: Random effects method

```
              PANEL DATA ESTIMATION
              =====================

Balanced data:  NI=   17, T=   35, NOB=   595

Variance Components (random effects) Estimates:

VWITH (variance of Uit)   =  0.25514E+06
VBET  (variance of Ai)    =  55401.
(computed from small sample formula)
THETA (0=WITHIN, 1=TOTAL) =  0.11628

Dependent variable: SRT

Sum of squared residuals = .152560E+09               R-squared = .214013
  Variance of residuals = 264862.            Adjusted R-squared = .189450
Std. error of regression = 514.647

         Estimated    Standard
Variable Coefficient    Error     t-statistic
U         -12.0814     8.78623    -1.37504
P          16.4247     4.72670     3.47489
C          248.622     71.9106     3.45737
```

4. Random effects estimation methods

Here we have

$$\widehat{\sigma}_v^2 = 0.25514e^{06} = 255,140$$

$$\widehat{\sigma}_\alpha^2 = 55,401$$

$$\widehat{\psi} = \frac{\widehat{\sigma}_v^2}{\widehat{\sigma}_v^2 + T\widehat{\sigma}_\alpha^2} = \frac{255,140}{255,140 + 35 \times 55,401} = 0.1163$$

4. Random effects estimation methods

Key Concepts Section 4

1. Error-component model.
2. GLS, Between and pooled estimators.
3. Write the GLS estimator as a weighted average.
4. Feasible GLS estimator.
5. Properties of the GLS estimator.
6. Asymptotic variance-covariance matrix of the GLS estimator.

Section 5

Specification tests: Fixed or Random effects?

5. Specifications tests

Objectives

1. Define the **Mundlak's specification**.
2. Discuss the **independence assumption** between random effects and explanatory variables.
3. Show that the GLS estimator may be **not consistent** when T is fixed.
4. Show that the GLS is **always consistent** when T tends to infinity.
5. Introduce the **Hausman's lemma**.
6. Define the **Hausman's specification test**.

5. Specifications tests

Fact (large T sample)

Whether to treat the effects as fixed or random makes no difference when T **is large**, *because both the LSDV estimator and the generalized least-squares estimator become the same estimator:*

$$\widehat{\beta}_{GLS} \underset{T \to \infty}{\longrightarrow} \widehat{\beta}_{LSDV}$$

Fixed or random effects

- When T is finite and n is large, whether to treat the effects as fixed or random is not an easy question to answer.
- It can make a surprising amount of difference in the estimates of the parameters.

5. Specifications tests

Example (Hausman, 1978)

Hausman (1978) estimates a wage equation using a sample of 629 high school graduates followed over six years by the Michigan income dynamics study. The explanatory variables include a piecewise-linear representation of age, the presence of unemployment or poor health in the previous year, and dummy variables for self-employment, living in the South, or living in a rural area.

5. Specifications tests

Table 3.3. *Wage equations (dependent variable: log wage[a])*

Variable	Fixed effects	Random effects
1. Age 1 (20–35)	0.0557	0.0393
	(0.0042)	(0.0033)
2. Age 2 (35–45)	0.0351	0.0092
	(0.0051)	(0.0036)
3. Age 3 (45–55)	0.0209	−0.0007
	(0.0055)	(0.0042)
4. Age 4 (55–65)	0.0209	−0.0097
	(0.0078)	(0.0060)
5. Age 5 (65–)	−0.0171	−0.0423
	(0.0155)	(0.0121)
6. Unemployed previous year	−0.0042	−0.0277
	(0.0153)	(0.0151)
7. Poor health previous year	−0.0204	−0.0250
	(0.0221)	(0.0215)
8. Self-employment	−0.2190	−0.2670
	(0.0297)	(0.0263)
9. South	−0.1569	−0.0324
	(0.0656)	(0.0333)
10. Rural	−0.0101	−0.1215
	(0.0317)	(0.0237)
11. Constant	—	0.8499
	—	(0.0433)
s^2	0.0567	0.0694
Degrees of freedom	3,135	3,763

[a]3,774 observations; standard errors are in parentheses.
Source: Hausman (1978).

5. Specifications tests

In the random-effects framework, there are two fundamental assumptions.

1. One is that the unobserved individual effects α_i are random draws from a common population.

2. The explanatory variables are strictly exogenous: it implies that **all the components of the error terms** are orthogonal to the regressors:

$$\mathbb{E}\left(\varepsilon_{it} \middle| x_{i1}, .., x_{iK} \right) = \mathbb{E}\left(\alpha_i \middle| x_{i1}, .., x_{iK} \right) = \mathbb{E}\left(v_{it} \middle| x_{i1}, .., x_{iK} \right) = 0$$

5. Specifications tests

What happens when this condition is violated?

$$\mathbb{E}\left(\alpha_i \middle| x_{i1}, .., x_{iK}\right) \neq 0 \quad \text{or} \quad \mathbb{E}\left(\alpha_i x_{it}'\right) \neq 0$$

1. **The Mundlak's specification (1978)**
2. **The Hausman's specification test**

Subsection 5.1

The Mundlak's Specification

5.1. Mundlak's specification

Mundlak's specification

- Mundlak (1978) criticized the random-effects formulation on the grounds that it neglects the correlation that may exist between the effects α_i and the explanatory variables x_{it}.
- There are reasons to believe that in many circumstances α_i and x_{it} are indeed correlated.

Mundlak Y. (1978), "On the Pooling of Time Series and Cross Section Data", *Econometrica*, 46, 69-85.

Mundlak's specification

- The properties of various estimators we have discussed thus far depend on the existence and extent of the relations between the X's and the effects α_i.
- Therefore, we have to consider the joint distribution of these variables. However, α_i are unobservable.
- **Mundlak (1978) suggests to approximate $\mathbb{E}\left(\alpha_i x_{it}\right)$ by a linear function.**

5.1. Mundlak's specification

Definition (Mundlak's specification)

Let us assume that the individual effects satisfy:

$$\alpha_i = \underbrace{\overline{x}_i' a}_{\text{component proportional to } x} + \underbrace{\alpha_i^*}_{\text{component orthogonal to } x}$$

with $a \in \mathbb{R}^K$, $\overline{x}_i = T^{-1} \sum_{t=1}^{T} x_{it}$ the $K \times 1$ vector of individual means of the explanatory variables and

$$\mathbb{E}\left(\alpha_i^* x_{it}'\right) = 0$$

5.1. Mundlak's specification

Definition (Mundlak's specification)

With the Mundlak's specification, the unobserved effects model becomes*:*

$$y_{it} = \mu + \beta' x_{it} + \overline{x}_i' a + \varepsilon_{it}$$

$$\varepsilon_{it} = \alpha_i^* + v_{it}$$

5.1. Mundlak's specification

Assumption H3: The error term $\varepsilon_{it} = \alpha_i^* + v_{it}$ *are i.i.d.* $\forall\ (it)$ with:

- $\mathbb{E}\left(\alpha_i^*\right) = \mathbb{E}\left(v_{it}\right) = 0$
- $\mathbb{E}\left(\alpha_i^* v_{it}\right) = 0$
- $\mathbb{E}\left(\alpha_i^* \alpha_j^*\right) = \begin{cases} \sigma_{\alpha^*}^2 & i = j \\ 0 & \forall i \neq j \end{cases}$
- $\mathbb{E}\left(v_{it} v_{j,s}\right) = \begin{cases} \sigma_v^2 & t = s,\ i = j \\ 0 & \forall t \neq s,\ \forall i \neq j \end{cases}$
- $\mathbb{E}\left(v_{it} x_{it}'\right) = \mathbb{E}\left(\alpha_i^* x_{it}'\right) = 0$

5.1. Mundlak's specification

Mundlak's specification

The model can be rewritten as follows:

$$\underset{(T,1)}{y_i} = \underset{(T,K+1)}{\widetilde{X}_i^*} \underset{(K+1,1)}{\gamma} + \underset{(T,1)}{\varepsilon_i} \qquad \forall\, i = 1, .., n$$

with

$$\varepsilon_i = \alpha_i^* e + v_i$$

$$\widetilde{X}_i^* = \left(e\overline{x}_i' : e : X_i \right)$$

$$\gamma' = \left(a' : \mu : \beta' \right)$$

Mundlak's specification

The variance-covariance matrix of the error term is defined as:

$$\begin{aligned} \mathbb{E}\left(\varepsilon_i \varepsilon_j'\right) &= \mathbb{E}\left(\left(\alpha_i^* e + v_i\right)\left(\alpha_j^* e + v_j\right)'\right) \\ &= \begin{cases} \sigma_{\alpha^*}^2 ee' + \sigma_v^2 I_T = V^* & i = j \\ 0 & i \neq j \end{cases} \end{aligned}$$

GLS estimator

Utilizing the expression for the inverse of a partitioned matrix, we obtain the GLS estimator of μ, β, and a as:

$$\widehat{\mu}^*_{GLS} = \overline{y} - \overline{x}'\widehat{\beta}_{BE}$$

$$\widehat{\beta}^*_{GLS} = \Delta\widehat{\beta}_{BE} + (I_K - \Delta)\,\widehat{\beta}_{LSDV}$$

$$\widehat{a}^*_{GLS} = \widehat{\beta}_{BE} - \widehat{\beta}_{LSDV}$$

Between estimator

The **between estimator** $\widehat{\beta}_{BE}$ corresponds to the *OLS estimator obtained in the model:*

$$\overline{y}_i = c + (\beta + a)' \, \overline{x}_i + \varepsilon_i = c + \theta' \overline{x}_i + \varepsilon_i \quad \forall \, i = 1, .., n$$

$$\widehat{\theta}_{BE} = \left(\sum_{i=1}^{n} \left(\overline{x}_i - \overline{x} \right) \left(\overline{x}_i - \overline{x} \right)' \right)^{-1} \left(\sum_{i=1}^{n} \left(\overline{x}_i - \overline{x} \right) \left(\overline{y}_i - \overline{y} \right) \right)$$

5.1. Mundlak's specification

GLS estimator properties

Under H3, the GLS estimator $\widehat{\beta}^{*}_{GLS}$ is consistent (cf. section 4):

$$\widehat{\beta}^{*}_{GLS} \underset{nT\to\infty}{\longrightarrow} \beta$$

Besides, we have

$$\widehat{\beta}^{*}_{GLS} \underset{T\to\infty}{\longrightarrow} \widehat{\beta}_{LSDV}$$

5.1. Mundlak's specification

Question: What is the consequence to neglect the dependence between α_i and x_{it} and to wrongly consider the following model?

$$y_{it} = \mu + \beta' x_{it} + \varepsilon_{it}$$

$$\varepsilon_{it} = \alpha_i + v_{it}$$

$$\alpha_i = \overline{x}_i' a + \alpha_i^*$$

5.1. Mundlak's specification

Let us assume that the DGP corresponds to the Mundlak's model

$$\alpha_i = \overline{x}_i' a + \alpha_i^*$$

and we apply GLS to the initial model:

$$y_{it} = \mu + \beta' x_{it} + \varepsilon_{it}$$

$$\varepsilon_{it} = \alpha_i + v_{it}$$

5.1. Mundlak's specification

In general, we have:

$$\widehat{\beta}_{GLS} = \Delta\widehat{\beta}_{BE} + (I_K - \Delta)\,\widehat{\beta}_{LSDV}$$

It is easy to show that:

$$\widehat{\beta}_{BE} \xrightarrow[n\rightarrow\infty]{p} \beta + a$$

$$\widehat{\beta}_{LSDV} \xrightarrow[n\rightarrow\infty]{p} \beta$$

5.1. Mundlak's specification

Let us assume that

$$\Delta \xrightarrow[n\rightarrow\infty]{p} \overline{\Delta}$$

with

$$\Delta = \psi T \left(\sum_{i=1}^{n} X_i' Q X_i + \psi T \sum_{i=1}^{n} \left(\overline{x}_i - \overline{x}\right)\left(\overline{x}_i - \overline{x}\right)' \right)^{-1} \left(\sum_{i=1}^{n} \left(\overline{x}_i - \overline{x}\right)\left(\overline{x}_i - \overline{x}\right)' \right)$$

5.1. Mundlak's specification

When T is fixed and n tends to infinity, the **GLS is not consistent** if there is a correlation between individual effects and the expanatory variables:

$$\begin{aligned}
\underset{n\rightarrow\infty}{\text{plim}}\, \widehat{\beta}_{GLS} &= \overline{\Delta} \times \underset{n\rightarrow\infty}{\text{plim}}\, \widehat{\beta}_{BE} + \left(I_K - \overline{\Delta}\right) \times \underset{n\rightarrow\infty}{\text{plim}}\, \widehat{\beta}_{LSDV} \\
&= \overline{\Delta} \times (\beta + a) + \left(I_K - \overline{\Delta}\right) \times \beta \qquad (2) \\
&= \beta + \overline{\Delta} a
\end{aligned}$$

with

$$\overline{\Delta} = \underset{n\rightarrow\infty}{\text{plim}} \Delta.$$

5.1. Mundlak's specification

Theorem (GLS bias)

If $\alpha_i = \overline{x}_i' a + \alpha_i^*$ *with* $a \neq 0$*, the GLS is* **not consistent** *when* T *is fixed and* n *tends to infinity:*

$$\widehat{\beta}_{GLS} \underset{n \to \infty}{\overset{p}{\longrightarrow}} \beta + \overline{\Delta} a$$

As usual, the GLS is consistent with T*:*

$$\widehat{\beta}_{GLS} \underset{T \to \infty}{\overset{p}{\longrightarrow}} \beta$$

Summary

	$\mathbb{E}\left(\alpha_i \mid x_{i1},..,x_{iK}\right) = 0$		$\mathbb{E}\left(\alpha_i \mid x_{i1},..,x_{iK}\right) \neq 0$	
	LSDV	GLS	LSDV	GLS
T fixed, $n \rightarrow \infty$	Consistent	—	Consistent	Not Consistent
$T \rightarrow \infty$ and $n \rightarrow \infty$	Consistent	BLUE	Consistent	Consistent

Subsection 5.2

The Hausman's Specification Test

5.2. Hausman's specification test

Hausman (1978) proposes a **general specification test**, that can be applied in the specific context of linear panel models to the issue of specification of individual effects (fixed or random).

Hausman J.A., (1978) "Specification Tests in Econometrics", *Econometrica*, 46, 1251-1271

5.2. Hausman's specification test

General idea of the Hausman's lemma

- Let us consider a general model

$$y = f(x; \beta) + \varepsilon$$

  and particular hypothesis H_0 on the parameters, error terms, etc.

- Let us consider two estimators of the K-vector β, denoted $\widehat{\beta}_1$ and $\widehat{\beta}_2$, **both consistent** under H_0 and asymptotically normally distributed.

- Under H_0, the estimator $\widehat{\beta}_1$ reaches the asymptotic Cramer–Rao bound.

- Under H_1, the estimator $\widehat{\beta}_2$ is biased and not consistent.

General idea of the Hausman's lemma

By examing the **distance** between $\widehat{\beta}_1$ and $\widehat{\beta}_2$, it is possible to conclude about H_0 :

1. If the distance is small, H_0 can not be rejected.
2. If the distance is large, H_0 can be rejected.

5.2. Hausman's specification test

Distance measure

- This distance is naturally defined as follows:

$$H = \left(\widehat{\beta}_2 - \widehat{\beta}_1\right)' \left(\mathbb{V}\left(\widehat{\beta}_2 - \widehat{\beta}_1\right)\right)^{-1} \left(\widehat{\beta}_2 - \widehat{\beta}_1\right)$$

- **However, the issue is to compute the variance-covariance matrix $\mathbb{V}\left(\widehat{\beta}_2 - \widehat{\beta}_1\right)$ of the difference between both estimators.**

- Generaly we know $\mathbb{V}\left(\widehat{\beta}_2\right)$ and $\mathbb{V}\left(\widehat{\beta}_1\right)$, but not $\mathbb{V}\left(\widehat{\beta}_2 - \widehat{\beta}_1\right)$.

5.2. Hausman's specification test

Lemma (Hausman, 1978)

Based on a sample of n observations, consider two estimates $\widehat{\beta}_1$ and $\widehat{\beta}_2$ that are both consistent and asymptotically normally distributed, with $\widehat{\beta}_1$ attaining the asymptotic Cramer–Rao bound so that $\sqrt{n}\left(\widehat{\beta}_1 - \beta\right)$ is asymptotically normally distributed with variance–covariance matrix V_1. Suppose $\sqrt{n}\left(\widehat{\beta}_2 - \beta\right)$ is asymptotically normally distributed, with mean zero and variance–covariance matrix V_2. Let $\widehat{q} = \widehat{\beta}_2 - \widehat{\beta}_1$. Then the limiting distributions [under the null] of $\sqrt{n}\left(\widehat{\beta}_1 - \beta\right)$ and $\sqrt{n}\widehat{q}$ have zero covariance:

$$\mathbb{E}(\widehat{\beta}_1 \widehat{q}') = 0_K$$

5.2. Hausman's specification test

Theorem

From this lemma, it follows that

$$\mathbb{V}\left(\widehat{\beta}_2-\widehat{\beta}_1\right)=\mathbb{V}\left(\widehat{\beta}_2\right)-\mathbb{V}\left(\widehat{\beta}_1\right)$$

Thus, Hausman suggests using the test statistic

$$H=\left(\widehat{\beta}_2-\widehat{\beta}_1\right)'\left(\mathbb{V}\left(\widehat{\beta}_2\right)-\mathbb{V}\left(\widehat{\beta}_1\right)\right)^{-1}\left(\widehat{\beta}_2-\widehat{\beta}_1\right)$$

or equivalently

$$H=\widehat{q}'\left(\mathbb{V}\left(\widehat{q}\right)\right)^{-1}\widehat{q}$$

5.2. Hausman's specification test

Under the null hypothesis, the test statistic H has an asymptotic chi-square distribution with K degrees of freedom.

$$H \underset{n \to \infty}{\overset{H_O}{\longrightarrow}} \chi^2 (K)$$

Under the alternative, it has a noncentral chi-square distribution with noncentrality parameter $\widetilde{q}' \left(\mathbb{V} \left(\widehat{q} \right) \right)^{-1} \widetilde{q}$, where $\widetilde{q}$ is defined as follows:

$$\widetilde{q} = \underset{H_1 / n \to \infty}{\text{plim}} \left(\widehat{\beta}_2 - \widehat{\beta}_1 \right)$$

5.2. Hausman's specification test

Specification test for fixed versus random effects

- Let us apply the Hausman's test to discriminate between fixed effects methods and random effects methods.
- We assume that α_i are random variable and the key assumption tested is here defined as:

$$H_0 : \mathbb{E}\left(\alpha_i \mid X_i\right) = 0$$

$$H_1 : \mathbb{E}\left(\alpha_i \mid X_i\right) \neq 0$$

5.2. Hausman's specification test

Definition (Hausman's specification test)

The Hausman specification test is a test of the null of no dependence between the (random) individual effects and the explanatory variables.

$$H_0 : \mathbb{E}\left(\alpha_i \middle| X_i\right) = 0$$

$$H_1 : \mathbb{E}\left(\alpha_i \middle| X_i\right) \neq 0$$

Specification test for fixed versus random effects

The Hausman's test can also be interpreted as a specification test between "fixed effect methods" and "random effect methods".

1. If the null is rejected, the correlation between individual effects and the explicative variables induces a bias in the GLS estimates. So, a standard LSDV approach (**fixed effects** method) has to be privilegiated.

2. If the null is not rejected, we can use a GLS estimator (random effect method) and specify the individual effects as random variables (**random effects** model).

5.2. Hausman's specification test

Hausman's specification test

How to implement this test? Let us consider the standard model with random effects ($\mu = 0$):

$$y_i = X_i \beta + e\alpha_i + v_i$$

1. Under H_0 (and assumptions A_2) we know that $\widehat{\beta}_{LSDV}$ and $\widehat{\beta}_{GLS}$ are consistent and asymptotically normally distributed.

2. Under H_0, $\widehat{\beta}_{GLS}$ is BLUE and attains asymptotic Cramer–Rao bound.

3. Under H_1, $\widehat{\beta}_{GLS}$ is not consistent.

5.2. Hausman's specification test

According to the Hausman's lemma, we have (for $K = 1$):

$$cov\left(\widehat{\beta}_{GLS}, \left(\widehat{\beta}_{LSDV} - \widehat{\beta}_{GLS}\right)\right) = 0 \iff cov\left(\widehat{\beta}_{LSDV}, \widehat{\beta}_{GLS}\right) = \mathbb{V}\left(\widehat{\beta}_{GLS}\right)$$

Since,

$$\mathbb{V}\left(\widehat{\beta}_{LSDV} - \widehat{\beta}_{GLS}\right) = \mathbb{V}\left(\widehat{\beta}_{LSDV}\right) + \mathbb{V}\left(\widehat{\beta}_{GLS}\right) - 2cov\left(\widehat{\beta}_{LSDV}, \widehat{\beta}_{GLS}\right)$$

We have:

$$\mathbb{V}\left(\widehat{\beta}_{LSDV} - \widehat{\beta}_{GLS}\right) = \mathbb{V}\left(\widehat{\beta}_{LSDV}\right) - \mathbb{V}\left(\widehat{\beta}_{GLS}\right)$$

Definition (Hausman's specification test)

The Hausman specification test statistic of individual effect can be defined as follows:

$$H = \left(\widehat{\beta}_{LSDV} - \widehat{\beta}_{GLS}\right)' \left(\mathbb{V}\left(\widehat{\beta}_{LSDV}\right) - \mathbb{V}\left(\widehat{\beta}_{GLS}\right)\right)^{-1} \left(\widehat{\beta}_{LSDV} - \widehat{\beta}_{GLS}\right)$$

Under $H_0 : \mathbb{E}\left(\alpha_i \middle| X_i\right) = 0$, we have:

$$H \underset{nT \to \infty}{\overset{H_O}{\longrightarrow}} \chi^2\left(K\right)$$

Remarks

$$H=\left(\widehat{\beta}_{LSDV}-\widehat{\beta}_{GLS}\right)'\left(\mathbb{V}\left(\widehat{\beta}_{LSDV}\right)-\mathbb{V}\left(\widehat{\beta}_{GLS}\right)\right)^{-1}\left(\widehat{\beta}_{LSDV}-\widehat{\beta}_{GLS}\right)$$

Let us assume that $K=1$, then under the null $H_0:\mathbb{E}\left(\alpha_i\middle|X_i\right)=0$

$$\mathbb{V}\left(\widehat{\beta}_{LSDV}\right)-\mathbb{V}\left(\widehat{\beta}_{GLS}\right)>0$$

since $\widehat{\beta}_{GLS}$ is the BLUE.

5.2. Hausman's specification test

Remarks

1. When n is fixed and T tends to infinity, $\widehat{\beta}_{GLS}$ and $\widehat{\beta}_{MCG}$ become identical. However, it was shown by Ahn and Moon (2001) that the numerator and denominator of H approach zero at the same speed. Therefore the ratio remains chi-square distributed. However, in this situation the fixed-effects and random-effects models become indistinguishable for all practical purposes.

2. The more typical case in practice is that n is large relative to T, so that differences between the two estimators or two approaches are important problems.

5.2. Hausman's specification test

Example

Let us consider a simple panel regression model for the total number of strikes days in OECD countries. We have a balanced panel data set for 17 countries ($n = 17$) and annual data form 1951 to 1985 ($T = 35$).

$$s_{it} = \alpha_i + \beta_i u_{it} + \gamma_i p_{it} + \varepsilon_{it}$$

5.2. Hausman's specification test

```
            PANEL DATA ESTIMATION
            ====================

Balanced data:  NI=   17, T=   35, NOB=   595

WITHIN (fixed effects) Estimates:

Dependent variable: SRT

Sum of squared residuals = .146958E+09              R-squared = .242875
  Variance of residuals = 255136.             Adjusted R-squared = .219215
Std. error of regression = 505.110

        Estimated   Standard
Variable  Coefficient    Error      t-statistic
U        -21.5968     9.19158     -2.34963
P        16.2729      4.75658      3.42113

Variance Components (random effects) Estimates:

VWITH (variance of Uit)   =  0.25514E+06
VBET  (variance of Ai)    =   55401.
(computed from small sample formula)
THETA (0=WITHIN, 1=TOTAL) =  0.11628

Dependent variable: SRT

Sum of squared residuals = .152560E+09              R-squared = .214013
  Variance of residuals = 264862.             Adjusted R-squared = .189450
Std. error of regression = 514.647

        Estimated   Standard
Variable  Coefficient    Error      t-statistic
U        -12.0814     8.78623     -1.37504
P        16.4247      4.72670      3.47489
C        248.622      71.9106      3.45737

Hausman test of H0:RE vs. FE:  CHISQ(2) = 13.924,  P-value = [.0009]
```

Key Concepts Section 5

1. Mundlak's specification.
2. Dependence between random effects and explanatory variables.
3. The GLS estimator may be not consistent (fixed T, n tends to infinity).
4. The GLS is always consistent when T tends to infinity.
5. Hausman's lemma.
6. Hausman's specification test for fixed or random effects models.

Section 6

Heterogeneous Panel Data Models

6. Heterogeneous panel data models

Objectives

1. Define the **heterogeneous** panel data model.
2. Introduce the **random coefficient model**.
3. Introduce the **Swamy's** model.
4. Define the **GLS** estimator

6. Heterogeneous panel data models

There are cases in which there are changing economic structures or different socioeconomic and demographic background factors that imply that the **slope parameters** β may be varying over time and/or may be different for different crosssectional units.

6. Heterogeneous panel data models

Heterogeneous panel data model

The most general form of an heterogeneous and time-varying coefficient model is:

$$y_{it} = \sum_{k=1}^{K} \beta_{kit} x_{kit} + v_{it} \quad i = 1, .., n \text{ and } t = 1, .., T$$

In contrast to previous sections, we no longer treat the intercept differently than other explanatory variables and let $x_{1it} = 1$.

6. Heterogeneous panel data models

Assumption H4: *We assume that parameters do not vary with time. Then, we have:*

$$\beta_{kit} = \beta_{ki}$$

$$y_{it} = \sum_{k=1}^{K} \beta_{ki} x_{kit} + v_{it} \quad i = 1, .., n \text{ and } t = 1, .., T$$

6. Heterogeneous panel data models

Panel or not panel?

This model is equivalent to postulating a separate regression for each cross-sectional unit

$$y_{it} = \beta_i' x_{it} + v_{it} \quad i = 1, .., n$$

where $\beta_i = (\beta_{1i}, \beta_{2i}, ..., \beta_{Ki})'$ is a $K \times 1$ vector of parameters, and $x_{it} = (x_{1it}, ..., x_{Kit})'$ is a $K \times 1$ vector of exogenous variables.

6. Heterogeneous panel data models

Panel or not panel?

$$y_{it} = \beta_i' x_{it} + v_{it} \quad i = 1, .., n$$

But some "links" between the individuals may require a panel regression model :

- The error terms v_{it} are cross-correlated among cross-units.
- The slope parameters β_i are considered as random variable with a **common probability distribution** or at least **common moments**

6. Heterogeneous panel data models

Definition (heterogeneous slope parameters)

The vectors of **slope parameters** β_i are assumed to satisfy

$$\underset{(K,1)}{\beta_i} = \underset{(K,1)}{\beta} + \underset{(K,1)}{\zeta_i}$$

for $i = 1, .., n$, where β is a $K \times\ 1$ vector of constants, and ζ_i denotes a $K \times\ 1$ vector of constant or random variables.

6. Heterogeneous panel data models

The **heterogeneous coefficient model** becomes

$$y_{it} = \sum_{k=1}^{K} \left(\beta_k + \xi_{ki}\right) x_{kit} + v_{it} \qquad i = 1, .., n \text{ and } t = 1, .., T$$

- $\beta = (\beta_1, \beta_2, ..., \beta_K)'$ denotes the common mean coefficient $K \times 1$ vector.
- $\xi_i = (\xi_{1i}, \xi_{2i}, ..., \xi_{Ki})'$ is the vector of individual deviation from the common mean.
- The errors terms may be cross-correlated or not, i.e. $cov\left(v_{jit}, v_{it}\right) \neq 0$ or $cov\left(v_{jit}, v_{it}\right) = 0$.

6. Heterogeneous panel data models

For this type of model we are interested in

1. Estimating the mean coefficient vector β,
2. Predicting each individual component β_i,
3. Estimating the dispersion of the individual-parameter vectors β_i,

Heterogeneous panel models with fixed or random coefficients

1. If individual observations are heterogeneous, then ξ_i can be treated as **fixed constants**.

2. If conditional on x_{kit}, individual units can be viewed as random draws from a common population, then ξ_i are generally treated as **random variables** having for instance, zero means and constant variances and covariances.

$$\mathbb{E}\left(\zeta_i\right) = 0 \quad \text{and} \quad \mathbb{V}\left(\zeta_i\right) = \Delta$$

6. Heterogeneous panel data models

Definition (Fixed-coefficient model)

When β_i are treated as **fixed constants**, we can stack the nT observations in the form of the Zellner (1962) seemingly unrelated regression (SURE) model

$$\begin{pmatrix} y_1 \\ . \\ y_n \end{pmatrix} = \begin{pmatrix} X_1 & 0 & 0 \\ 0 & .. & .. \\ 0 & .. & X_n \end{pmatrix} \begin{pmatrix} \beta_1 \\ . \\ \beta_n \end{pmatrix} + \begin{pmatrix} v_1 \\ . \\ v_n \end{pmatrix}$$

where y_i and v_i are $T \times 1$ vectors $(y_{it}, ..., y_{iT})$ and $(v_{it}, ..., v_{iT})$, and X_i is the $T \times K$ matrix of the time-series observations of the i^{th} individual's explanatory variables with the t^{th} row equal to x_{it}.

Heterogeneous panel models with fixed coefficients

1. If the covariances between different cross-sectional units are not zero, *e.g.* $\mathbb{E}\left(v_i v_{j'}\right) \neq 0$, the GLS estimator of $\left(\beta_1', ..., \beta_n'\right)$ is more efficient than the single-equation estimator of i for each cross-sectional unit. Panel data is **useful**.

2. If X_i are identical for all i or $\mathbb{E}\left(v_i v_{i'}\right) = \sigma_i^2 I_T$ and $\mathbb{E}\left(v_i v_{j'}\right) = 0$ for $i \neq j$, the GLS estimator for $\left(\beta_1', ..., \beta_n'\right)$ is the same as applying least squares separately to the time-series observations of each cross-sectional unit. Panel data is **useless**.

6. Heterogeneous panel data models

Definition (random coefficient model)

Alternatively, each regression coefficient can be viewed as **a random variable** with a common probability distribution:

$$\beta_i \overset{i.i.d.}{\sim} \text{Common Distribution}$$

or at least common moments:

$$\mathbb{E}(\beta_i) = \beta \quad \mathbb{V}(\beta_i) = \Delta \quad \forall i = 1, ..., n$$

Random coefficient model

1. The random-coefficient specification reduces the number of parameters to be estimated substantially, while still allowing the coefficients to differ from unit to unit and/or from time to time.

2. Depending on the type of assumption about the parameter variation, it can be further classified into one of two categories: stationary and nonstationary random-coefficient models.

3. For more details, see Hurwicz (1950), Klein (1953), Theil and Mennes (1959), or Zellner (1966).

Subsection 6.1

Random Coefficient Models

6.1. Random coefficient model

Definition (random coefficient model)

The vectors of slope parameters β_i are **randomly distributed** with a common mean $\mathbb{E}(\beta_i) = \beta$, and

$$y_{it} = \sum_{k=1}^{K} (\beta_k + \xi_{ki}) x_{kit} + v_{it} \qquad i = 1, .., n$$

with $\beta = (\beta_1, \beta_2, ..., \beta_K)'$ and $\xi_i = (\xi_{1i}, \xi_{2i}, ..., \xi_{Ki})$. Let us denote

$$\beta_{ki} = \beta_k + \xi_{ki}$$

6.1. Random coefficient model

Remarks

1. The vector $x_i = (x_{1i}..x_{Ki})$ includes a constant term. The parameter β_{ki} associated to this constant term corresponds to an **individual (random) effect**.

2. An alternative notation is:

$$y_{it} = \alpha + \sum_{k=2}^{K} (\beta_k + \xi_{ki}) x_{kit} + \alpha_i + v_{it}$$

$$\mathbb{E}(\alpha_i) = 0$$

6.1. Random coefficient model

Consider the set of assumptions used in the seminal paper of Swamy (1970).

Swamy P.A. (1970), "Efficient Inference in a Random Coefficient Regression Model", *Econometrica*, 38, 311-323

6.1. Random coefficient model

Assumption H5 (Swamy's model): *Let us assume that*

- $\mathbb{E}\left(\xi_i\right) = 0, \quad \mathbb{E}\left(v_i\right) = 0$
- $\mathbb{E}\left(\xi_i \xi_j'\right) = \begin{cases} \Delta & i = j \\ 0 & \forall i \neq j \end{cases}$
- $\mathbb{E}\left(x_{it} \xi_j'\right) = 0, \; E\left(\xi_i v_j'\right) = 0, \quad \forall \left(i, j\right)$
- $\mathbb{E}\left(v_i v_{j'}\right) = \begin{cases} \sigma_i^2 I_T & t = s, \; i = j \\ 0 & \forall t \neq s, \forall i \neq j \end{cases}$

6.1. Random coefficient model

Remark: we assume that the error term v_i is heteroskedastic:

$$\mathbb{E}\left(v_i v_i'\right) = \sigma_i^2 I_T$$

6.1. Random coefficient model

Definition (moments of slopes parameters)

The two first moments of the vector of random parameters $\beta_i = \beta + \xi_i$ are defined by, $\forall\, i = 1, ..n$:

$$\underset{(K,1)}{\mathbb{E}\left(\beta_i\right)} = \underset{(K,1)}{\beta}$$

$$\underset{(K,K)}{\mathbb{V}\left(\beta_i\right)} = \mathbb{E}\left(\xi_i \xi_i'\right) = \underset{(K,K)}{\Delta}$$

6.1. Random coefficient model

Homegeneous moments

The variance covariance matrix Δ of the random parameters $\beta_i = (\beta_{1i}, \beta_{2i}, .., \beta_{Ki})'$ is assumed to be **common** to all cross section units:

$$\underset{(K,K)}{\Delta} = \mathbb{E}\left((\beta_i - \beta)(\beta_i - \beta)'\right) = \begin{pmatrix} \sigma^2_{\beta_1} & \sigma_{\beta_1,\beta_2} & ... & \sigma_{\beta_1,\beta_K} \\ \sigma_{\beta_2,\beta_1} & \sigma^2_{\beta_2} & ... & \sigma_{\beta_2,\beta_K} \\ ... & ... & ... & ... \\ \sigma_{\beta_K,\beta_1} & \sigma_{\beta_K,\beta_2} & ... & \sigma^2_{\beta_K} \end{pmatrix}$$

6.1. Random coefficient model

Vectorial form

For each cross section unit, we have:

$$y_i = X_i \beta + X_i \xi_i + v_i$$

$$\beta_i = \beta + \xi_i$$

where the vector X_i include a constant term (i.e. the average of random individual effects, α).

6.1. Random coefficient model

Definition (random coefficient model)

The **random coefficient model** can be rewriten as follows :

$$y_i = X_i \beta + \varepsilon_i$$

$$\varepsilon_i = X_i \xi_i + v_i = X_i \left(\beta_i - \beta \right) + v_i$$

6.1. Random coefficient model

Covariance matrix

For a given cross unit, the covariance matrix for the composite disturbance term $\varepsilon_i = X_i \xi_i + v_i$ is defined by:

$$\begin{aligned}
\Phi_i &= \mathbb{E}\left(\varepsilon_i \varepsilon_i'\right) \\
&= \mathbb{E}\left[\left(X_i \xi_i + v_i\right)\left(X_i \xi_i + v_i\right)'\right] \\
&= X_i \mathbb{E}\left(\xi_i \xi_i'\right) X_i' + \mathbb{E}\left(v_i v_i'\right) \\
&= X_i \Delta X_i' + \sigma_i^2 I_T
\end{aligned}$$

6.1. Random coefficient model

Definition

For a given cross unit, the covariance matrix for the composite disturbance term $\varepsilon_i = X_i \xi_i + v_i$ is defined by:

$$\Phi_i = X_i \Delta X_i' + \sigma_i^2 I_T$$

Stacking all nT observations, the covariance matrix for the composite disturbance term is block-diagonal and heteroskedastic.

6.1. Random coefficient model

Remarks

1. Under Swamy's assumption, the simple regression of y on X will yield an unbiased and consistent estimator of β if $(1/nT)X'X$ converges to a nonzero constant matrix.

2. But the estimator is inefficient, and the usual least-squares formula for computing the variance–covariance matrix of the estimator is incorrect, often leading to misleading statistical inferences.

Definition (GLS estimator)

The best linear unbiased estimator of β is the GLS estimator

$$\widehat{\beta}_{GLS} = \left(\sum_{i=1}^{n} X_i' \Phi_i^{-1} X_i \right)^{-1} \left(\sum_{i=1}^{n} X_i' \Phi_i^{-1} y_i \right)$$

6.1. Random coefficient model

Definition (GLS estimator)

The GLS estimator $\widehat{\beta}_{GLS}$ is a matrix-weighted average of the least-squares estimator $\widehat{\beta}_i$ for each cross-sectional unit, with the weights inversely proportional to their covariance matrices:

$$\widehat{\beta}_{GLS} = \sum_{i=1}^{n} \omega_i \widehat{\beta}_i$$

$$\omega_i = \left(\sum_{i=1}^{n} \left(\Delta + \sigma_i^2 \left(X_i' X_i \right)^{-1} \right)^{-1} \right)^{-1} \left[\Delta + \sigma_i^2 \left(X_i' X_i \right)^{-1} \right]^{-1}$$

$$\widehat{\beta}_i = \left(X_i' X_i \right)^{-1} X_i' y_i$$

6.1. Random coefficient model

The covariance matrix for the GLS estimator is:

$$\begin{aligned}\mathbb{V}\left(\widehat{\beta}_{GLS}\right) &= \left(\sum_{i=1}^{n} X_i'\Phi_i^{-1}X_i\right)^{-1} \\ &= \left(\sum_{i=1}^{n}\left(\Delta+\sigma_i^2\left(X_i'X_i\right)^{-1}\right)^{-1}\right)^{-1}\end{aligned}$$

6.1. Random coefficient model

Feasible GLS estimator

Swamy proposes to use the OLS estimators $\widehat{\beta}_i = \left(X_i' X_i\right)^{-1} X_i' y_i$ and their residuals $\widehat{v}_i = y_i - X_i \widehat{\beta}_i$ to obtain unbiased estimators of σ_i^2 and Δ

$$\begin{aligned}
\widehat{\sigma}_i^2 &= \frac{1}{T-K} y_i' \left(I_T - X_i \left(X_i' X_i\right)^{-1} X_i' \right) y_i \\
&= \frac{1}{T-K} \sum_{t=1}^{T} \widehat{v}_{it}
\end{aligned}$$

with

$$y_{it} = \beta_i' x_{it} + v_{it}$$

6.1. Random coefficient model

Feasible GLS estimator

For the Δ matrix, we have:

$$\underset{(K,K)}{\widehat{\Delta}} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\left(\widehat{\beta}_i - \frac{1}{n}\sum_{i=1}^{n}\widehat{\beta}_i\right)\left(\widehat{\beta}_i - \frac{1}{n}\sum_{i=1}^{n}\widehat{\beta}_i\right)'\right) - \frac{1}{n}\sum_{i=1}^{n}\widehat{\sigma}_i^2\left(X_i'X_i\right)^{-1}$$

6.1. Random coefficient model

Definition (estimator for Δ)

However, the previous estimator $\widehat{\Delta}$ is not necessarily nonnegative definite. In this situation, Swamy (1970) has suggested replacing this estimator by:

$$\underset{(K,K)}{\widehat{\Delta}} = \frac{1}{n-1} \sum_{i=1}^{n} \left(\left(\widehat{\beta}_i - \frac{1}{n} \sum_{i=1}^{n} \widehat{\beta}_i \right) \left(\widehat{\beta}_i - \frac{1}{n} \sum_{i=1}^{n} \widehat{\beta}_i \right)' \right)$$

This estimator, although not unbiased, is nonnegative definite and is consistent when T tends to infinity.

6.1. Random coefficient model

Remarks

1. Swamy proved that substituting $\widehat{\sigma}_i^2$ and $\widehat{\Delta}$ for σ_i^2 and Δ yields an asymptotically normal and efficient estimator of β.

2. The speed of convergence of the GLS estimator is $n^{1/2}$.

6.1. Random coefficient model

Summary: how to estimate a random coefficient model?

1. Run the n individual regressions $y_{it} = \beta_i' x_{it} + v_{it}$.

2. Compute $\widehat{\sigma}_i^2$ and the Swamy's estimator $\widehat{\Delta}$ as follows

$$\underset{(K,K)}{\widehat{\Delta}} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\left(\widehat{\beta}_i - \frac{1}{n}\sum_{i=1}^{n}\widehat{\beta}_i\right)\left(\widehat{\beta}_i - \frac{1}{n}\sum_{i=1}^{n}\widehat{\beta}_i\right)'\right)$$

$$\widehat{\sigma}_i^2 = \frac{1}{T-K}\sum_{t=1}^{T}\widehat{v}_{it}$$

3. Compute the GLS estimate of the mean of the parameters β_i

$$\widehat{\beta}_{GLS} = \left(\sum_{i=1}^{n} X_i' \widehat{\Phi}_i^{-1} X_i \right)^{-1} \left(\sum_{i=1}^{n} X_i' \widehat{\Phi}_i^{-1} y_i \right)$$

with

$$\widehat{\Phi}_i = X_i \widehat{\Delta} X_i' + \widehat{\sigma}_i^2 I_T$$

6.1. Random coefficient model

Example (Swamy, 1970)

Swamy (1970) used this model to reestimate the Grunfeld investment function with the annual data of 11 U.S. corporations. His GLS estimates of the common-mean coefficients of the firms' beginning-of-year value of outstanding shares and capital stock are 0.0843 and 0.1961, with asymptotic standard errors 0.014 and 0.0412, respectively. The estimated dispersion measure of these coefficients is

$$\widehat{\Delta} = \begin{pmatrix} 0.0011 & -0.0002 \\ & 0.0187 \end{pmatrix}$$

Predicting Individual Coefficients

1. Sometimes one may wish to predict the individual component β_i, because it provides information on the behavior of each individual and also because it provides a basis for predicting future values of the dependent variable for a given individual.

2. Swamy (1970, 1971) has shown that the best linear unbiased predictor, conditional on given x_i, is the least-squares estimator $\widehat{\beta}_i$.

6.1. Random coefficient model

Definition (individual predictors for β_i)

Lee and Griffiths (1979) suggest predicting β_i by

$$\beta_i^* = \widehat{\beta}_{GLS} + \Delta X_i' \left(X_i \Delta X_i' + \sigma_i^2 I_T \right)^{-1} \left(y_i - X_i \widehat{\beta}_{GLS} \right)$$

This predictor is the best linear unbiased estimator in the sense that $\mathbb{E}\left(\beta_i^* - \beta_i\right) = 0$, where the expectation is an unconditional one.

Predicting Individual Coefficients : Lindley and Smith (1972)

1. It is also possible to consider a Bayesian approach: the random coefficient model is also called the **hierarchical model**
2. In this case, the prior distribution for the β_i parameters is specified with the value for β, Δ and σ_i^2.
3. From a Bayesian perspective, the likelihood is combined with priors to generate posterior distributions of the parameters.

Subsection 6.2

Other Heterogeneous Panel Data Models

6.2. Other hetorogeneous panel data models

Heterogeneous panel data models

There are many other way to model the heterogeneity of slope parameters:

1. Mixed fixed and random (MFR) coefficients model.
2. Mean group estimation.
3. Panel threshold regression models.
4. Grouped Patterns of Heterogeneity.
5. etc.

6.2. Other hetorogeneous panel data models

Heterogeneous panel data models

There are many other way to model the heterogeneity of slope parameters:

1. **Mixed fixed and random (MFR) coefficients model.**
2. Mean group estimation.
3. Panel threshold regression models.
4. Grouped Patterns of Heterogeneity.
5. etc.

6.2. Other hetorogeneous panel data models

Definition (mixed fixed and random coefficient model)

We assume that each cross section unit is different

$$y_{it} = \sum_{k=1}^{K} \beta_{ki} x_{ki} + \sum_{l=1}^{m} \gamma_{li} w_{li} + v_{it} \qquad i = 1, .., n$$

where x_{it} and w_{it} are each a $K \times 1$ and an $m \times 1$ vector of explanatory variables that are independent of the error of the equation v_{it}.

Hsiao C. (1989), "Modelling Ontrario Regional Electricity System Demand Using a Mixed Fixed and Random Coeffcient Approaoch", *Regional Science and Urban Economics*, 19, 565-587

6.2. Other hetorogeneous panel data models

Assumptions

- The parameters $\underset{(nK,1)}{\beta} = (\beta_1', \beta_2', ..., \beta_K')'$ are assumed to be randomly distributed
- The parameters $\underset{(nm,1)}{\gamma} = (\gamma_1', \gamma_2', ..., \gamma_m')'$ are fixed.

6.2. Other hetorogeneous panel data models

In a vectorial form, we have

$$\underset{(nT,1)}{y} = \underset{(nT,nK)}{X}\underset{(nK,1)}{\beta} + \underset{(nT,nm)}{W}\underset{(nm,1)}{\gamma} + \underset{(nT,1)}{v}$$

$$X = \begin{pmatrix} X_1 & 0 & 0 \\ 0 & X_2 & \\ & & \\ 0 & & X_n \end{pmatrix} \quad W = \begin{pmatrix} W_1 & 0 & 0 \\ 0 & W_2 & \\ & & \\ 0 & & W_n \end{pmatrix}$$

6.2. Other hetorogeneous panel data models

Assumptions on the random coefficients

The coefficients of x_{it} are assumed to be subject to stochastic restrictions of the form

$$\beta = A_1 \overline{\beta} + \zeta$$

- A_1 is an $nK \times L$ matrix with known elements,
- $\overline{\beta}$ is an $L \times 1$ vector of constants,
- ζ is assumed to be (normally distributed) random variables with mean 0 and nonsingular constant covariance matrix C and is independent of x_i.

6.2. Other hetorogeneous panel data models

Assumptions on the fixed coefficients

The coefficients of w_{it} are assumed to be subject to

$$\gamma = A_2 \overline{\gamma}$$

- A_2 is an $nm \times n$ matrix with known elements,
- $\overline{\gamma}$ is an $n \times 1$ vector of constants.

MFR model

Since A_2 is known, we can rewrite the model as

$$\underset{(nT,1)}{y} = \underset{(nT,nK)}{X}\underset{(K,1)}{\beta} + \underset{(nT,n)}{\widetilde{W}}\underset{(n,1)}{\overline{\gamma}} + \underset{(nT,1)}{v}$$

where $\widetilde{W} = WA_2$.

6.2. Other hetorogeneous panel data models

Many of the linear panel data models with unobserved individual specific but time-invariant heterogeneity can be treated as special cases of this model.

6.2. Other hetorogeneous panel data models

Example

A common model for all cross-sectional units. If there is no interindividual difference in behavioral patterns, we may let $X = 0$, $A_2 = e_n \otimes I_m$, so model becomes

$$y_{it} = w_{it}\overline{\gamma} + v_{it}$$

6.2. Other hetorogeneous panel data models

Example

When each individual is considered different, then $X = 0$, $A_2 = I_n \otimes I_m$, and the model becomes

$$y_{it} = w_{it}\gamma_i + v_{it}$$

6.2. Other hetorogeneous panel data models

Example

When the effects of the individual specific, time-invariant omitted variables are treated as random variables just as in the assumption on the effects of other omitted variables, we can let $X_i = e_T$, $\zeta' = (\zeta_1, ..., \zeta_n)$, $A_1 = e_n$, $C = I_n \sigma_\alpha^2$, $\overline{\beta}$ be an unknown constant, and w_{it} not contain an intercept term. Then the model becomes:

$$y_{it} = \overline{\beta} + \overline{\gamma}' w_{it} + \zeta_i + v_{it}$$

6.2. Other hetorogeneous panel data models

Heterogeneous panel data models

There are many other way to model the heterogeneity of slope parameters:

1. Mixed fixed and random (MFR) coefficients model.
2. Mean group estimation.
3. Panel threshold regression models.
4. Grouped Patterns of Heterogeneity.
5. etc.

6.2. Other hetorogeneous panel data models

Heterogeneous panel data models

There are many other way to model the heterogeneity of slope parameters:

1. Mixed fixed and random (MFR) coefficients model.
2. **Mean group estimation**.
3. Panel threshold regression models.
4. Grouped Patterns of Heterogeneity.
5. etc.

6.2. Other hetorogeneous panel data models

Consider an heterogeneous panel data model

$$y_{it} = \beta_i' x_{it} + v_{it} \quad i = 1, .., n$$

- A consistent estimator of $\beta = \mathbb{E}\left(\beta_i\right)$ can be obtained under more general assumptions concerning β_i and the regressors.
- One such possible estimator is the **Mean Group (MG)** estimator proposed by Pesaran and Smith (1995) for estimation of dynamic random coefficient models.

6.2. Other hetorogeneous panel data models

Definition (mean group estimator)

The mean group (MG) estimator is defined as the simple average of the OLS estimators $\widehat{\beta}_i$

$$\widehat{\beta}_{MG} = \frac{1}{n}\sum_{i=1}^{n}\widehat{\beta}_i$$

6.2. Other hetorogeneous panel data models

Mean Group (MG) estimator

When the regressors are strictly exogeneous and the errors are i.i.d, an unbiased estimator of the covariance matrix is given by

$$\mathbb{V}\left(\widehat{\beta}_{MG}\right) = \frac{1}{n}\widehat{\Delta}$$

$$\widehat{\Delta} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\left(\widehat{\beta}_i - \frac{1}{n}\sum_{i=1}^{n}\widehat{\beta}_i\right)\left(\widehat{\beta}_i - \frac{1}{n}\sum_{i=1}^{n}\widehat{\beta}_i\right)'\right)$$

6.2. Other hetorogeneous panel data models

For more details

Pesaran, M.H. and R. Smith (1995), "Estimation of Long-Run Relationships from Dynamic Heterogeneous Panels", *Journal of Econometrics*, 68, 79-114.

Pesaran, M.H.,Y. Shin and R.P. Smith, (1999), "PooledMean Group Estimation of Dynamic Heterogeneous Panels", *Journal of the American Statistical Association*, 94, 621-634.

6.2. Other hetorogeneous panel data models

Heterogeneous panel data models

There are many other way to model the heterogeneity of slope parameters:

1. Mixed fixed and random (MFR) coefficients model.
2. Mean group estimation.
3. Panel threshold regression models.
4. Grouped Patterns of Heterogeneity.
5. etc.

6.2. Other hetorogeneous panel data models

Heterogeneous panel data models

There are many other way to model the heterogeneity of slope parameters:

1. Mixed fixed and random (MFR) coefficients model.
2. Mean group estimation.
3. Panel threshold regression models.
4. **Grouped Patterns of Heterogeneity**.
5. etc.

6.2. Other hetorogeneous panel data models

Bonhomme S. and E. Manresa (2015), Grouped Patterns of Heterogeneity in Panel Data, *Econometrica*, 83(3), 1147-1184

- This paper introduces time-varying **Grouped Patterns of Heterogeneity** in linear panel data models.
- A distinctive feature of this approach is that group membership is left unrestricted.
- The authors estimate the parameters of the model using a "grouped fixed-effects" estimator that minimizes a least-squares criterion with respect to all possible groupings of the cross-sectional units.

6.2. Other hetorogeneous panel data models

A simple linear model with grouped patterns of heterogeneity takes the following form

$$y_{it} = x_{it}'\theta + \alpha_{g_i t} + v_{it}$$

- $\alpha_{g_i,t} \in \mathcal{A} \subset \mathbb{R}$ denotes a group-specific unobservable variable
- The group membership variables $g_i \in \{1, ..., G\}$ and the group-specific time effects $\alpha_{g_i t}$ are unrestricted.
- Units in the same group share the same time profile α_{gt}
- The number of groups G is to be set or estimated by the researcher.

Definition

The grouped fixed-effects estimator in this model is defined as the solution of the following minimization problem:

$$\left(\widehat{\theta},\widehat{\alpha},\widehat{\gamma}\right) = \arg\min \sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it} - x_{it}'\theta - \alpha_{g_i t}\right)^2$$

where the minimum is taken over all possible groupings $\gamma = \{g_1, ..., g_n\}$ of the n units into G groups, common parameters θ, and group-specific time effects α.

Definition

For given values of θ and α, the optimal group assignment for each individual unit is given by:

$$\widehat{g}_i\left(\theta,\alpha\right) = \underset{g\in\{1,\ldots,G\}}{\arg\min} \sum_{t=1}^{T} \left(y_{it} - x_{it}'\theta - \alpha_{gt}\right)^2$$

6.2. Other hetorogeneous panel data models

This model can be extended to allow for group-specific effects of covariates (heterogeneous slope coefficients):

$$y_{it} = x_{it}'\theta_{g_i} + \alpha_{g_i t} + v_{it}$$

6.2. Other hetorogeneous panel data models

The authors propose a very intuitive iterative algorithm to:

- Estimate the parameters
- Determine the group membership

For more details, see

Bonhomme S. and E. Manresa (2015), Grouped Patterns of Heterogeneity in Panel Data, *Econometrica*, 83(3), 1147-1184

6. Heterogeneous panel data models

Key Concepts Section 6

1. Heterogeneous panel data model.
2. Random coefficient model.
3. GLS estimator.
4. Hierarchical model.
5. Mixed fixed and random coefficients model.
6. Mean group estimator.
7. Grouped Patterns of Heterogeneity.

End of Chapter 1

Christophe Hurlin (University of Orléans)