Kurt Schmidheiny
Unversität Basel

Panel Data: Fixed and Random Effects

matrix-free

1 Introduction

In panel data, individuals (persons, firms, cities, ...) are observed at several points in time (days, years, before and after treatment, ...). This handout focuses on panels with relatively few time periods (small T) and many individuals (large N).

This handout introduces the two basic models for the analysis of panel data, the fixed effects model and the random effects model, and presents consistent estimators for these two models. The handout does not cover so-called dynamic panel data models.

Panel data are most useful when we suspect that the outcome variable depends on explanatory variables which are not observable but correlated with the observed explanatory variables. If such omitted variables are constant over time, panel data estimators allow to consistently estimate the effect of the observed explanatory variables.

2 The Econometric Model

Consider the multiple linear regression model for individual $i = 1, ..., N$ who is observed at several time periods $t = 1, ..., T$

$$y_{it} = \alpha + \beta_1 x_{it1} + ... + \beta_K x_{itK} + c_i + u_{it}$$

where y_{it} is the dependent variable, $x_{it1}, ..., x_{itK}$ are K explanatory variables, α and β_k are $K + 1$ parameters, c_i is an *individual-specific effect* and u_{it} is an *idiosyncratic* error term.

We will assume throughout this handout that each individual i is observed in all time periods t. This is a so-called *balanced panel*. The treatment of unbalanced panels is straightforward but tedious.

Version: 7-5-2012, 13:59

The data generation process (dgp) is described by:

PL1: Linearity

$y_{it} = \alpha + \beta_1 x_{it1} + ... + \beta_K x_{itK} + c_i + u_{it}$
where $E(u_{it}) = 0$ and $E(c_i) = 0$

The model is linear in parameters α, β, individual effect c_i and error u_{it}.

PL2: Independence

$\{x_{i11}...x_{iTK}, y_{i1}...y_{iT}\}_{i=1}^{N}$
i.i.d. (independent and identically distributed)

The observations are independent across individuals but not necessarily across time. This is guaranteed by random sampling of individuals.

PL3: Strict Exogeneity

$E(u_{it}|x_{i11}...x_{iTK}, c_i) = 0$ (mean independent)

The idiosyncratic error term u_{it} is assumed uncorrelated with the explanatory variables of all past, current and future time periods of the same individual. This is a strong assumption which e.g. rules out lagged dependent variables. *PL3* also assumes that the idiosyncratic error is uncorrelated with the individual specific effect.

PL4: Error Variance

a) $V(u_{it}|x_{i11}...x_{iTK}, c_i) = \sigma_u^2 > 0$ and $< \infty$
$Corr(u_{it}, u_{is}|x_{i11}...x_{iTK}, c_i) = 0$ for all $s \neq t$
(homoscedastic and no serial correlation)

b) $V(u_{it}|x_{i11}...x_{iTK}, c_i) = \sigma_{u,it}^2 > 0$ and $< \infty$
$Corr(u_{it}, u_{is}|x_{i11}...x_{iTK}, c_i) = 0$ for all $s \neq t$
(no serial correlation)

c) $V(u_{it}|x_{i11}...x_{iTK}, c_i) = \sigma_{u,it}^2 > 0$ and $< \infty$
$Corr(u_{it}, u_{is}|x_{i11}...x_{iTK}, c_i) < 1$ and > -1 for all $s \neq t$

The remaining assumptions are divided into two sets of assumptions: the random effects model and the fixed effects model.

2.1 The Random Effects Model

In the random effects model, the individual-specific effect is a random variable that is uncorrelated with the explanatory variables.

RE1: Unrelated effects

$a)$ $E(c_i|x_{i11}...x_{iTK}) = 0$ and
$V(c_i|x_{i11}...x_{iTK}) = \sigma_c^2 < \infty$

$b)$ $E(c_i|x_{i11}...x_{iTK}) = 0$ and
$V(c_i|x_{i11}...x_{iTK}) = \sigma_{c,i}^2(x_{i11}...x_{iTK}) < \infty$

RE1 assumes that the individual-specific effect is uncorrelated with the explanatory variables of all past, current and future time periods of the same individual. Version *RE1a* assumes constant variance.

RE2: Identifiability

$(1, x_{it1}, \cdots, x_{itK})$ are not linearly dependent and
$V(x_{itk}) > 0$ and $< \infty$

RE2 assumes that the regressors including a constant are not perfectly collinear, that all regressors (but the constant) have non-zero variance and not too many extreme values.

The random effects model can be written as

$$y_{it} = \alpha + \beta_1 x_{it1} + ... + \beta_K x_{itK} + v_{it}$$

where $v_{it} = c_i + u_{it}$. Assuming *PL2*, *PL4* and *RE1* in the special versions *PL4a* and *RE1a* leads to

$$V(v_{it}|x_{i11}...x_{iTK}) = \sigma_v^2 = \sigma_c^2 + \sigma_u^2 \text{ for all } i, t$$
$$Cov(v_{it}, v_{is}|x_{i11}...x_{iTK}) = \sigma_c^2 \text{ for all } i \text{ and } s \neq t$$
$$Cov(v_{it}, v_{js}|x_{i11}...x_{iTK}, x_{j11}...x_{jTK}) = 0 \text{ for all } s, t \text{ and } i \neq j\,.$$

This special case under the a) versions of *PL4* and *RE1* is therefore called the *equicorrelated random effects model.*

2.2 The Fixed Effects Model

In the fixed effects model, the individual-specific effect is a random variable that is allowed to be correlated with the explanatory variables.

FE1: related effects

–

FE1 explicitly states the absence of the independence assumption in *RE1*.

FE2: Identifiability

$(\ddot{x}_{it1}, \cdots, \ddot{x}_{itK})$ are not linearly dependent and
$V(\ddot{x}_{itk}) > 0$ and $< \infty$ for all k
where $\ddot{x}_{itk} = x_{itk} - \bar{x}_{ik}$ and $\bar{x}_{ik} = 1/T \sum_i x_{itk}$

FE2 assumes that the explanatory variables are not perfectly collinear, that all regressors have non-zero within-variance (i.e. variation over time for a given individual) and not too many extreme values. Hence x_{it} can not include a constant or any other time-invariant variables.

3 Estimation with Pooled OLS

The *pooled OLS estimator* ignores the panel structure of the data and simply estimates α and β by regressing y_{it} on a constant and on $x_{it1}, \cdots, x_{itK}$. In the special case with one regressor x_{it}, the resulting pooled OLS estimators of α and β are:

$$\widehat{\beta}^{POLS} = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x})(y_{it} - \bar{y})}{\sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x})^2}$$

$$\widehat{\alpha}^{POLS} = \bar{y} - \hat{\beta}_1 \bar{x}$$

where $\bar{y} = 1/NT \sum_i \sum_t y_{it}$ and $\bar{x} = 1/NT \sum_i \sum_t x_{it}$.

Random effects model: The pooled OLS estimator of α and β is unbiased under *PL1*, *PL2*, *PL3*, *RE1*, and *RE2* in small samples. Additionally assuming *PL4* and normally distributed idiosyncratic and individual-specific errors, it is normally distributed in small samples. It is consistent and approximately normally distributed under *PL1*, *PL2*, *PL3*, *PL4*, *RE1*, and *RE2* in samples with a large number of individuals ($N \to \infty$). However, the pooled OLS estimator is not efficient. More importantly, the usual standard errors of the pooled OLS estimator is incorrect and tests (t-, F-, z-, Wald-) based on them are not valid. Correct standard errors can be estimated with the so-called cluster-robust covariance estimator treating each individual as a cluster (see the handout on "Clustering in the Linear Model").

Fixed effects model: The pooled OLS estimators of α and β are biased and inconsistent, because the variable c_i is omitted and potentially correlated with the other regressors.

4 Random Effects Estimation

The *random effects estimator* is the feasible generalized least squares (GLS) estimator. GLS transforms the data (dependent and explanatory variables) such that the error terms in the transformed model are uncorrelated across all N individuals *and* all time periods T. The GLS is similar to the weighted least squares (WLS) estimator but not easily expressed without matrix notation (see the handout on "Heteroscedasticity in the linear model").

The transformation of the model depends on the two unknown parameters σ_v^2 and σ_c^2 only. There are many different ways to estimate these two parameters. For example,

$$\widehat{\sigma}_v^2 = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\widehat{v}_{it}^2 \quad , \quad \widehat{\sigma}_c^2 = \widehat{\sigma}_v^2 - \widehat{\sigma}_u^2$$

where

$$\widehat{\sigma}_u^2 = \frac{1}{NT-N}\sum_{t=1}^{T}\sum_{i=1}^{N}(\widehat{v}_{it} - \overline{\widehat{v}}_i)^2$$

and $\widehat{v}_{it} = y_{it} - \alpha^{POLS} - \widehat{\beta}_1^{POLS} x_{it1} - \ldots - \widehat{\beta}_K^{POLS} x_{itK}$ and $\overline{\widehat{v}}_i = 1/T\sum_{t=1}^{T}\widehat{v}_{it}$. The degree of freedom correction in $\widehat{\sigma}_u^2$ is also asymptotically important when $N \to \infty$.

Random effects model: The RE estimator of α and β is unbiased under *PL1 - PL3*, *RE1a* and *RE2* in small samples. However, we cannot unbiasedly estimate its variance. Small samples t- and F-tests are only asymptotically valid.

The RE estimator is consistent and asymptotically normally distributed under *PL1 - PL4*, *RE1* and *RE2* when the number of individuals $N \to \infty$ even if T is fixed. It can therefore be approximated in samples with many individual observations N as

$$\widehat{\alpha}^{RE} \overset{A}{\sim} N\left(\alpha, Avar(\widehat{\alpha}^{RE})\right)$$

and for all k

$$\widehat{\beta}_k^{RE} \overset{A}{\sim} N\left(\beta_k, Avar(\widehat{\beta}_k^{RE})\right)$$

Assuming the equicorrelated model (*PL4a* and *RE1a*), $\widehat{\sigma}_v^2$ and $\widehat{\sigma}_c^2$ are consistent estimators of σ_v^2 and σ_c^2, respectively. Then $\widehat{\alpha}_{RE}$ and $\widehat{\beta}_{RE}$ are asymptotically efficient and the asymptotic variance can be consistently estimated. Allowing for arbitrary conditional variances and for serial correlation of the combined error v_{it} (*PL4b* and *RE1b*), the asymptotic variance can be consistently estimated with the so-called cluster-robust covariance estimator treating each individual as a cluster (see the handout on "Clustering in the Linear Model"). In both cases, the usual tests (z-, Wald-) for large samples can be performed.

In practice, we can rarely be sure about equicorrelated errors and better always use cluster-robust standard errors for the RE estimator.

Fixed effects model: Under the assumptions of the fixed effects model (*FE1*, i.e. *RE1* violated), the random effects estimators of α and β are biased and inconsistent, because the variable c_i is omitted and potentially correlated with the other regressors.

5 Fixed Effects Estimation

Subtracting time averages $\bar{y}_i = 1/T \sum_t y_{it}$ from the initial model

$$y_{it} = \alpha + \beta_1 x_{it1} + ... + \beta_K x_{itK} + c_i + u_{it}$$

yields the *within model*

$$\ddot{y}_{it} = \beta_1 \ddot{x}_{it1} + ... + \beta_K \ddot{x}_{itK} + \ddot{u}_{it}$$

where $\ddot{y}_{it} = y_{it} - \bar{y}_i$, $\ddot{x}_{itk} = x_{itk} - \bar{x}_{ik}$ and $\ddot{u}_{it} = u_{it} - \bar{u}_i$. Note that the individual-specific effect c_i and the intercept α cancel. Also note that time-invariant regressors (e.g. the constant) where $x_{itk} = \bar{x}_{ik}$ cancel as $\ddot{x}_{itk} = x_{itk} - \bar{x}_{ik} = 0$ and their effect cannot be estimated by the within estimator (see *FE2*).

The *fixed effects estimator* or *within estimator* of the slope coefficient β estimates the within model by OLS In the special case with one regressor, the resulting FE estimators of β_0 and β_1 are

$$\widehat{\beta}_1^{FE} = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} \ddot{x}_{it} \, \ddot{y}_{it}}{\sum_{i=1}^{N} \sum_{t=1}^{T} \ddot{x}_{it}^2}$$

The general case with K explanatory variables is derived likewise but not easily expressed without matrix notation.

Random effects model and fixed effects model: The fixed effects estimator of β is unbiased under *PL1*, *PL2*, *PL3*, and *FE2* in small samples. Additionally assuming *PL4*, *RE1* or *FE1* and normally distributed idiosyncratic errors, it is normally distributed in small samples. Assuming homoscedastic errors with no serial correlation (*PL4a*), the variance $V(\widehat{\beta}_k^{FE}|X)$ can be unbiasedly estimated with the usual OLS estimator in the transformed model. In the special case with one regressor, it is estimated as

$$\widehat{V}(\widehat{\beta}_1^{FE}|x_{111}, ..., x_{NTK}) = \frac{\widehat{\sigma}_u^2}{\sum_{i=1}^{N} \sum_{t=1}^{T} \ddot{x}_{it}^2}$$

where $\widehat{\sigma}_u^2 = \widehat{\ddot{u}}' \widehat{\ddot{u}} / (NT - N - K)$ and $\widehat{\ddot{u}}_{it} = \ddot{y}_{it} - \widehat{\beta}_1^{FE} \ddot{x}_{it1} - ... - \widehat{\beta}_K^{FE} \ddot{x}_{itK}$. Note the non-usual degrees of freedom correction. The usual z- and F-tests can be performed.

The FE estimator is consistent and asymptotically normally distributed under *PL1 - PL4*, *FE1* or *RE1* and *FE2* when the number of individuals

$N \to \infty$ even if T is fixed. It can therefore be approximated in samples with many individual observations N as

$$\widehat{\beta}_k^{FE} \overset{A}{\sim} N\left(\beta_k, Avar(\widehat{\beta}_k^{FE})\right)$$

Assuming homoscedastic errors with no serial correlation ($PL4a$), the asymptotic variance can be consistently estimated as the usual OLS estimator in the transformed model. In the special case with one regressor, it is estimated as

$$\widehat{Avar}(\widehat{\beta}_1^{FE}) = \frac{\widehat{\sigma}_u^2}{\sum_{i=1}^N \sum_{t=1}^T \ddot{x}_{it}^2}$$

where $\widehat{\sigma}_u^2 = \widehat{\ddot{u}}'\widehat{\ddot{u}}/(NT-N)$. Allowing for heteroscedasticity and serial correlation of unknown form ($PL4b$), the asymptotic variance $Avar(\widehat{\beta})$ can be consistently estimated with the so-called cluster-robust covariance estimator treating each individual as a cluster (see the handout on "Clustering in the Linear Model"). In both cases, the usual tests (z-, Wald-) for large samples can be performed.

In practice, the idiosyncratic errors are often likely serially correlated (violating $PL4a$) when $T > 2$. Bertrand, Duflo and Mullainathan (2004) show that the usual standard errors of the fixed effects estimator are drastically understated in the presence of serial correlation. It is therefore advisable to always use cluster-robust standard errors for the fixed effects estimator.

6 Least Squares Dummy Variables Estimator (LSDV)

The FE estimator is numerically identical to pooled OLS including a set of $N-1$ dummy variables which identify the individuals and hence an additional $N-1$ parameters $\gamma_2, ..., \gamma_N$. Note that one of the individual dummies is dropped because we include a constant. The estimators $\widehat{\beta}_1^{LSDV}, ..., \widehat{\beta}_K^{LSDV}, \widehat{\gamma}_2^{LSDV}, ..., \widehat{\gamma}_N^{LSDV}$ are generally not consistent as the number of parameters goes to infinity as $N \to \infty$. From the numerical identity with the FE estimator we know that $\widehat{\beta}_{LSDV}$ is consistent while $\widehat{\gamma}_{LSDV}$ is inconsistent. However, this so-called *incidental parameters* problem generally causes a bias in *non-linear* fixed effects models like the probit.

7 First Difference Estimator

Subtracting the lagged value $y_{i,t-1}$ from the initial model

$$y_{it} = \alpha + \beta_1 x_{it1} + ... + \beta_K x_{itK} + c_i + u_{it}$$

yields the *first-difference model*

$$\dot{y}_{it} = \beta_1 \dot{x}_{it1} + ... + \beta_K \dot{x}_{itK} + \dot{u}_{it}$$

where $\dot{y}_{it} = y_{it} - y_{i,t-1}$, $\dot{x}_{it} = x_{it} - x_{i,t-1}$ and $\dot{u}_{it} = u_{it} - u_{i,t-1}$. Note that the individual-specific effect c_i and the intercept α cancel. Also note that time-invariant regressors (e.g. the constant) cancel as $x_{it} - x_{i,t-1} = 0$ and their effect cannot be estimated by the first-difference estimator (see *FE2*). The *first-difference estimator* of the slope coefficient β estimates the first-difference model by OLS In the special case with one regressor, the resulting FE estimators of β_0 and β_1 are

$$\widehat{\beta}_1^{FD} = \frac{\sum_{i=1}^{N} \sum_{t=2}^{T} \dot{x}_{it}\, \dot{y}_{it}}{\sum_{i=1}^{N} \sum_{t=2}^{T} \dot{x}_{it}^2}$$

The general case with K explanatory variables is derived likewise but not easily expressed without matrix notation. In the special case $T = 2$, the FD estimator is numerically identical to the FE estimator.

Random effects model and fixed effects model: The FD estimator is a consistent estimator of β under the same assumptions as the FE estimator. It is less efficient than the FE estimator if u_{it} is not serially correlated (*PL4a*).

8 Random Effects vs. Fixed Effects Estimation

The random effects model can be consistently estimated by both the RE estimator or the FE estimator. The former is efficient while the latter is inefficient. So we would prefer the RE estimator if we can be sure that the individual-specific effect really is an unrelated effect (*RE1*).

This can be tested by a (Durbin-Wu-) Hausmann test. The test is performed by comparing $\widehat{\beta}_{RE}$ and $\widehat{\beta}_{FE}$ for the subset of coefficients of time-varying variables:

$$H \overset{A}{\sim} \chi^2_J$$

follows a χ^2 distribution with J degrees of freedom where J is the number of time-varying regressors. The null hypothesis is that the individual-specific effect is uncorrelated with the regressors and the errors are equicorrelated. Under H_0, $\widehat{\beta}_{RE}$ is consistent and efficient and $\widehat{\beta}_{FE}$ is consistent but inefficient; under H_A, $\widehat{\beta}_{RE}$ is inconsistent but $\widehat{\beta}_{FE}$ remains consistent. See e.g. Wooldridge (2002) for robust alternatives that do not require full efficiency of the RE estimator.

Note: Assumption *RE1* is an extremely strong assumption and the FE estimator is almost always much more convincing than the RE estimator. Interest in the effect of a time-invariant variable is no sufficient reason to use the RE estimator.

9 Time Fixed Effects

We often also suspect that there are time-specific effects δ_t which affect all individuals in the same way

$$y_{it} = \alpha + \beta_1 x_{it1} + ... + \beta_K x_{itK} + \delta_t + c_i + u_{it} \, .$$

We can estimate this extended model by including a dummy variable for each time period. Assuming a fixed number of time periods T and the number of individuals $N \to \infty$, both the RE estimator and the FE estimator are consistent using time dummy variables under the same conditions as above.

10 Implementation in Stata 12.0

Stata provides a series of commands that are especially designed for panel data. See `help xt` for an overview.

Stata requires panel data in the so-called *long form*: there is one line for every individual and every time observation. The very powerful Stata command `reshape` helps transforming data into this format. Before working with panel data commands, we have to tell Stata the variables that identify the individual and the time period. For example, load data

```
webuse nlswork.dta
```

and define individuals (variable *idcode*) and time periods (variable *year*)

```
xtset idcode year
```

The fixed effects estimator is calculated by the Stata command `xtreg` with the option `fe`:

```
generate ttl_exp2 = ttl_exp^2
xtreg ln_wage ttl_exp ttl_exp2, fe
```

Note that the effect of time-constant variables like *grade* is not identified by the fixed effects estimator. The parameter reported as *_cons* in the Stata output is the average fixed effect $1/N \sum_i c_i$. Stata uses $NT - N - K$ degrees of freedom for small sample tests. Cluster-robust Huber/White standard errors are reported with the `vce` option:

```
xtreg ln_wage ttl_exp ttl_exp2, fe vce(cluster idcode)
```

Since version 10, Stata automatically assumes clustering with robust standard errors in fixed effects estimations. So we could also just use

```
xtreg ln_wage ttl_exp ttl_exp2, fe vce(robust)
```

Stata reports small sample t- and F-tests with $N - 1$ degrees of freedom. Furthermore, Stata multiplies the cluster-robust covariance by $N/(N-1)$ to correct for degrees of freedom in small samples. This is ad-hoc in the cross-section context with large N but useful in the time series context with large T (see Stock and Watson, 2008).

The random effects estimator is calculated by the Stata command `xtreg` with the option `re`:

```
xtreg ln_wage grade ttl_exp ttl_exp2, re
```

Stata reports asymptotic z- and Wald tests with random effects estimation. Cluster-robust Huber/White standard errors are reported with the `vce` option:

```
xtreg ln_wage grade ttl_exp ttl_exp2, re vce(cluster idcode)
```

The Hausman test is calculated by estimating RE and FE and then comparing the estimates:

```
xtreg ln_wage grade ttl_exp ttl_exp2, re
estimates store b_re
xtreg ln_wage ttl_exp ttl_exp2, fe
estimates store b_fe
hausman b_fe b_re
```

The pooled OLS estimator with corrected standard errors is calculated with the standard ols command `regress`:

```
reg ln_wage grade ttl_exp ttl_exp2, vce(cluster idcode)
```

where the `vce` option was used to report correct cluster-robust Huber/White standard errors.

The least squares dummy variables estimator is calculated by including dummy variables for individuals in pooled OLS. It is numerically only feasibly with relatively few individuals:

```
drop if idcode > 50
xi: regress ln_wage ttl_exp ttl_exp2 i.idcode
```

where only the first 43 individuals are used in the estimation. The long list of estimated fixed effects can suppressed by using the `areg` command:

```
areg ln_wage ttl_exp ttl_exp2, absorb(idcode)
```

11 Implementation in R 2.13

The R package plm provides a series of functions and data structures that are especially designed for panel data.

The plm package works with data stored in a structured dataframe format. The function plm.data transforms data from the so-called *long form* into the plm structure. Long form data means that there is one line for every individual and every time observation. For example, load data

```
> library(foreign)
> nlswork <- read.dta("http://www.stata-press.com/data/r11/nlswork.dta")
```

and define individuals (variable *idcode*) and time periods (variable *year*)

```
> library(plm)
> pnlswork <- plm.data(nlswork, c("idcode","year"))
```

The fixed effects estimator is calculated by the R function plm and its model option within:

```
> ffe<-plm(ln_wage~ttl_exp+I(ttl_exp^2), model="within", data = pnlswork)
> summary(ffe)
```

Note that the effect of time-constant variables like *grade* is not identified by the fixed effects estimator. Cluster-robust Huber/White standard errors are reported with the lmtest package:

```
> library(lmtest)
> coeftest(ffe, vcov=vcovHC(ffe, cluster="group"))
```

where the option cluster="group" defines the clusters by the individual identifier in the plm.data dataframe.

The random effects estimator is calculated by the R function plm and its model option random:

```
> fre <- plm(ln_wage~grade+ttl_exp+I(ttl_exp^2), model="random",
   data = pnlswork)
> summary(fre)
```

Cluster-robust Huber/White standard errors are reported with the lmtest package:

```
> library(lmtest)
> coeftest(fre, vcov=vcovHC(fre, cluster="group"))
```

The Hausman test is calculated by estimating RE and FE and then comparing the estimates:

```
> phtest(ffe, fre)
```

The pooled OLS estimator with corrected standard errors is calculated by the R function `plm` and its model option `pooled`:

```
> fpo <- plm(ln_wage~grade+ttl_exp+I(ttl_exp^2), model="pooled",
   data = pnlswork)
> library(lmtest)
> coeftest(fpo, vcov=vcovHC(fpo, cluster="group"))
```

where the `lmtest` package was used to report correct cluster-robust Huber/White standard errors.

The least squares dummy variables estimator is calculated by including dummy variables for individuals in pooled OLS. It is numerically only feasibly with relatively few individuals:

```
> lsdv <- lm(ln_wage~ttl_exp+I(ttl_exp^2)+factor(idcode),
   data = nlswork, subset=(idcode<=50))
> summary(lsdv)
```

where only the first 43 individuals are used in the estimation.

References

Introductory textbooks

Stock, James H. and Mark W. Watson (2011), Introduction to Econometrics, 3rd ed., Pearson Addison-Wesley. Chapter 10.

Wooldridge, Jeffrey M. (2009), Introductory Econometrics: A Modern Approach, 4th ed., South-Western Cengage Learning. Chapters 13 and 14.

Advanced textbooks

Cameron, A. Colin and Pravin K. Trivedi (2005), Microeconometrics: Methods and Applications, Cambridge University Press. Chapter 21.

Wooldridge, Jeffrey M. (2002), Econometric Analysis of Cross Section and Panel Data, MIT Press. Chapter 10.

Companion textbooks

Angrist, Joshua D. and Jörn-Steffen Pischke (2009), Mostly Harmless Econometrics: An Empiricist's Companion, Princeton University Press. Chapter 5.

Articles

Manuel Arellano (1987), Computing Robust Standard Errors for Within-Group Estimators, Oxford Bulletin of Economics and Statistics, 49, 431-434.

Bertrand, M., E. Duflo and S. Mullainathan (2004), How Much Should We Trust Differences-in-Differences Estimates?, Quarterly Journal of Economics, 119(1), 249-275.

Stock, James H. and Mark W. Watson (2008), Heteroskedasticity-Robust Standard Errors for Fixed Effects Panel Data Regression, Econometrica, 76(1), 155-174. [advanced]