ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Department of Statistics

ECONOMETRICS
Panel Data Models

Ioannis Vrontos
Associate Professor in Statistics
Department of Statistics
School of Information Sciences and Technology
Athens University of Economics & Business
Email: vrontos@aueb.gr
Office: Troias & Kimolou, 4^{th} floor, 412
Tel.: 210 8203927

Panel Data Models

- Introduction and motivation
- Panel data models
  - General dynamic panel model
  - Basic linear panel model
  - Simple panel model
  - Fixed effects panel model
  - Random effects panel model
- Examples and applications using R

Panel data models

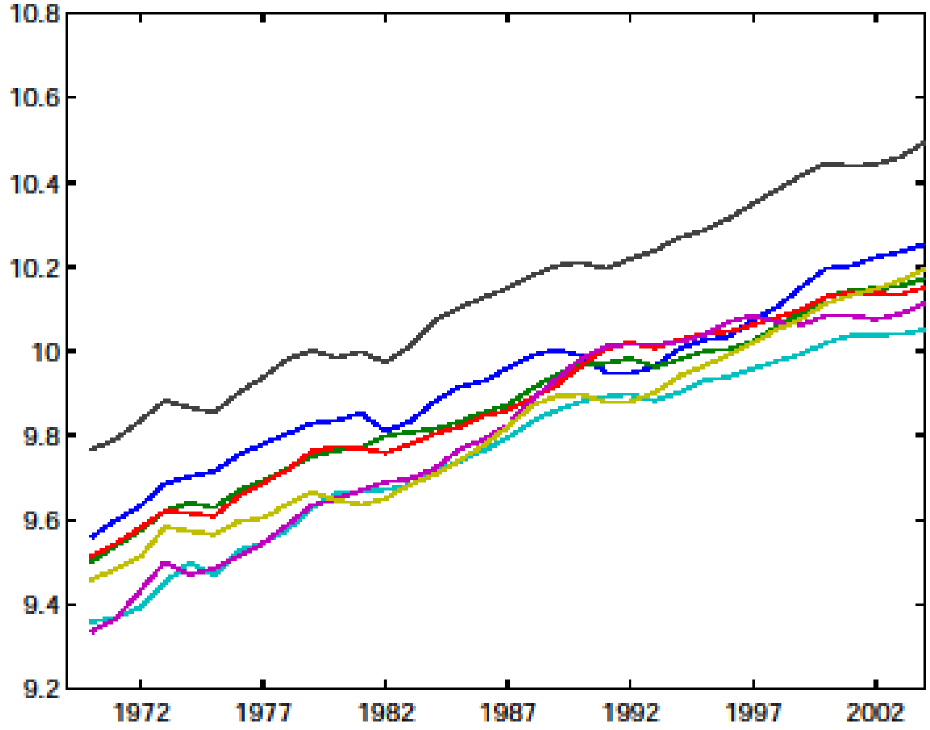
Introduction and motivation

Advantages:

- Larger number of data points (more degrees of freedom)
- Combine cross section and time series data
- Limitation of the omitted variable problem
- Takes into account common and cross sectional explanatory variables
- May account for cross sectional dependence

Panel data

GDP series for the G7 countries



Panel data

Dependent variables and Specific cross sectional explanatory variables

- **Panel data:**

  Dependent variable: Y_{it}, $i = 1, \dots, N$ $(cross\ section\ unit)$
  $t = 1, \dots, T$ $(time\ period)$

  Explanatory variables (specific cross sectional unit variables, i.e. different for each unit)

  X_{ijt}, $i = 1, \dots, N$ $(cross\ section\ unit)$
  $j = 1, \dots, k$ $(explanatory\ variables)$
  $t = 1, \dots, T$ $(time\ period)$

  E.g.: Y_{it}: GDP for different countries across time, X_{ijt}: investments, industrial production, unemployment for different countries across time

Y1	Y2	…	YN	X11	X12	…	X1k	…	XN1	XN2	…	XNk
Y_{11}	Y_{21}	…	Y_{N1}	X_{111}	X_{121}	…	X_{1k1}	…	X_{N11}	X_{N21}	…	X_{Nk1}
Y_{12}	Y_{22}	…	Y_{N2}	X_{112}	X_{122}	…	X_{1k2}	…	X_{N12}	X_{N22}	…	X_{Nk2}
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	…	$\vdots$	$\vdots$	$\ddots$	$\vdots$
Y_{1T}	Y_{2T}	…	Y_{NT}	X_{11T}	X_{12T}	…	X_{1kT}	…	X_{N1T}	X_{N2T}	…	X_{NkT}

Panel data

Dependent variables and Common explanatory variables

- **Panel data:**

  Dependent variable: $Y_{it}, \quad i = 1, \dots, N \ (cross\ section\ unit)$
  $t = 1, \dots, T \ (time\ period)$

  Explanatory variables (common variables, i.e. common for each unit)

  $X_{jt}, \quad j = 1, \dots, k \ (explanatory\ variables)$
  $t = 1, \dots, T \ (time\ period)$

  E.g.: Y_{it}: GDP for different countries across time, X_{jt}: global volatility index, global panic index across time (common explanatory variables for all units)

Y1	Y2	…	YN
Y_{11}	Y_{21}	…	Y_{N1}
Y_{12}	Y_{22}	…	Y_{N2}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
Y_{1T}	Y_{2T}	…	Y_{NT}

X1	X2	…	Xk
X_{11}	X_{21}	…	X_{k1}
X_{12}	X_{22}	…	X_{k2}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
X_{1T}	X_{2T}	…	X_{kT}

Panel data models

A general dynamic panel data model

- Consider the following dynamic panel data model:

$$Y_{it} = \delta_i + \beta_i t + \sum_{k=1}^{K} \alpha_{ik} X_{kt} + \sum_{q=1}^{Q} \gamma_{iq} X_{iqt} + \varepsilon_{it},$$
$$\varepsilon_{it} = \varphi_i \varepsilon_{i,t-1} + u_{it},$$

where

- $i = 1, \dots, N$ denotes the cross-sectional units of the panel (e.g. countries)
- $t = 1, \dots, T$ denotes the time period
- Y_{it} denotes the economic or financial dependent variable (e.g. gross domestic product)
- X_{kt} is a set of K exogenous explanatory factors or covariates (e.g. a global volatility index, a market or a commodity index) which are common for all cross-sectional units of the panel, and affect the dependent variables of the panel through separate/different coefficients α_{ik}
- X_{iqt} is a set of Q of cross-sectional specific factors (explanatory variables) which are different for each cross-sectional unit with separate coefficients γ_{iq}
- δ_i are the intercept coefficients of the model, which are different for each cross sectional unit
- β_i are the trend coefficients, which are different for each cross sectional unit
- ε_{it} is a zero-mean autoregressive one, AR(1), process which captures the dynamics of the panel data model. The error term u_{it} is assumed not to be serially correlated, i.e. $E(u_{is}u_{it}) = 0$, for all $t \neq s$ and for all i, but it is heterogeneous and correlated across i, i.e. $E(u_{it}u_{jt}) \neq 0$ for all i and j, and therefore, allows for cross sectional dependence across units

Panel data models

Basic linear panel data model

- The **basic linear panel data model** used in econometric literature can be described through suitable restrictions of the following general model:

$$Y_{it} = \alpha_{it} + \beta_{it} X_{it} + u_{it},$$

where

- $i = 1, \dots, N$ denotes the cross-sectional units of the panel
- $t = 1, \dots, T$ denotes the time period
- Y_{it} denotes the economic or financial dependent variable
- X_{it} is the explanatory factors or covariates
- α_{it} is the intercept, which is different for each cross sectional unit, and across time
- β_{it} are the explanatory variables coefficients, which are different for each cross sectional unit, and across time
- u_{it} is the error term

- Number of observations: NT
- Number of parameters to be estimated: 2NT
- Can not be estimated !! Thus, **suitable restrictions are imposed** on α_{it} and β_{it}

Panel data models

A simple panel data model (model I)

- The simplest panel data model assumes parameter homogeneity. In the general linear panel model $\mathrm{Y}_{it} = \alpha_{it} + \beta_{it}\mathrm{X}_{it} + u_{it}$ impose the following restrictions: $\alpha_{it} = \alpha$ and $\beta_{it} = \beta$ for all i, t. The resulting model can be written in the form:

$$\mathrm{Y}_{it} = a + \beta\mathrm{X}_{it} + u_{it},$$

where

- $i = 1, \dots, N$ denotes the cross-sectional units of the panel
- $t = 1, \dots, T$ denotes the time period
- Y_{it} denotes the economic or financial dependent variable
- X_{it} is the explanatory factors or covariate
- a is the intercept of the model, which is common for all cross sectional units and across time
- β is the explanatory variable coefficient, which is common for all cross sectional units across time
- u_{it} is the error term

- This simple panel model does not account for the heterogeneity of the cross sectional units (common a and β parameter)
- Number of observations: NT
- Number of parameters to be estimated: 2

Panel data models

Fixed effects panel data model (model II)

- In the general linear panel data model $\mathrm{Y}_{it} = \alpha_{it} + \beta_{it}\mathrm{X}_{it} + u_{it}$ impose the following restrictions: $\alpha_{it} = \alpha_i$ for all t, and $\beta_{it} = \beta$ for all i and t. The resulting model is called the **fixed effects model**:

$$\mathrm{Y}_{it} = \alpha_i + \beta\mathrm{X}_{it} + u_{it},$$

where

- $i = 1, \dots, N$ denotes the cross-sectional units of the panel
- $t = 1, \dots, T$ denotes the time period
- Y_{it} denotes the economic or financial dependent variable
- X_{it} is the explanatory factors or covariate
- α_i is the intercept of the model, which is different for each cross sectional unit
- β is the slope, which is common for all cross sectional units
- u_{it} is the error term

- This panel model accounts for the heterogeneity of the cross sectional units (different intercept coefficients α_i, common β parameter)
- Number of observations: NT
- Number of parameters to be estimated: N+1

Panel data models

Fixed effects panel data model (model III)

- In the general linear panel data model $\mathrm{Y}_{it} = \alpha_{it} + \beta_{it}\mathrm{X}_{it} + u_{it}$ impose the following restrictions: $\alpha_{it} = \alpha_i$, and $\beta_{it} = \beta_i$ for all t. The resulting model is:

$$\mathrm{Y}_{it} = \alpha_i + \beta_i \mathrm{X}_{it} + u_{it},$$

- $i = 1, \dots, N$ denotes the cross-sectional units of the panel
- $t = 1, \dots, T$ denotes the time period
- Y_{it} denotes the economic or financial dependent variable
- X_{it} is the explanatory factors or covariate
- α_i is the intercept of the model, which is different for each cross sectional unit
- β_i is the slope, which is different for each cross sectional unit
- u_{it} is the error term

- This panel model accounts for the heterogeneity of the cross sectional units (different intercept coefficients α_i, and different β_i coefficients)
- Different parameters across units, but constant across time
- Number of observations: NT
- Number of parameters to be estimated: 2N
- To be able to estimate this model: T>2

Panel data models: Fixed Effects

Least Squares Dummy variables (LSDV)

- Consider the following fixed effects panel data model:

$$\mathrm{Y}_{it} = \alpha_i + \beta \mathrm{X}_{it} + u_{it}$$

- Fixed parameters α_i (non stochastic), different for each sectional unit, common β
- We are interested in:
  - Estimate the model parameters
  - Perform hypothesis testing (for the heterogeneity of intercept parameters)
- Re-write the model by using/constructing appropriate dummy variables

$$\mathrm{Y}_{it} = \alpha + \beta \mathrm{X}_{it} + \gamma_2 d_{2t} + \gamma_3 d_{3t} + \cdots + \gamma_N d_{Nt} + u_{it}$$

- Number of observations: NT, number of parameters: $2 + (N - 1) = N + 1$

- The model implies that:

$i = 1$: $\mathrm{Y}_{1t} = \alpha + \beta \mathrm{X}_{1t} + u_{1t}$

$i = 2$: $Y_{2t} = \alpha + \beta \mathrm{X}_{2t} + \gamma_2 d_{2t} + u_{2t} \Rightarrow \mathrm{Y}_{2t} = (\alpha + \gamma_2) + \beta \mathrm{X}_{2t} + u_{2t}$

$i = 3$: $Y_{3t} = \alpha + \beta \mathrm{X}_{3t} + \gamma_3 d_{3t} + u_{3t} \Rightarrow \mathrm{Y}_{3t} = (\alpha + \gamma_3) + \beta \mathrm{X}_{3t} + u_{3t}$

$\vdots$

$i = N$: $Y_{Nt} = \alpha + \beta \mathrm{X}_{Nt} + \gamma_N d_{Nt} + u_{Nt} \Rightarrow \mathrm{Y}_{Nt} = (\alpha + \gamma_N) + \beta \mathrm{X}_{Nt} + u_{Nt}$

Panel data models: Fixed Effects

Construct the dummy variables

i	t	Y	X	D2	D3
1	1	Y_{11}	X_{11}	0	0
	2	Y_{12}	X_{12}	0	0
	⋮	⋮	⋮	⋮	⋮
	T	Y_{1T}	X_{1T}	0	0
2	1	Y_{21}	X_{21}	1	0
	2	Y_{22}	X_{22}	1	0
	⋮	⋮	⋮	⋮	⋮
	T	Y_{2T}	X_{2T}	1	0
3	1	Y_{31}	X_{31}	0	1
	2	Y_{32}	X_{32}	0	1
	⋮	⋮	⋮	⋮	⋮
	T	Y_{2T}	X_{3T}	0	1

Example: number of units $\mathrm{N} = 3$ (i.e. construct $N - 1 = 2$ dummy variables), time period T, dependent variable Y_{it}, explanatory variable X_{it}

Panel data models: Fixed Effects

Hypothesis testing (t-test)

- For the fixed effects panel data model of the form:

$$Y_{it} = \alpha + \beta X_{it} + \gamma_2 d_{2t} + \gamma_3 d_{3t} + \cdots + \gamma_N d_{Nt} + u_{it}$$

we can perform two types of hypothesis tests:

- **(i) t-test:**

$$H_0: \gamma_i = 0$$
$$H_1: \gamma_i \neq 0$$

by using the following test statistic:

$$T = \frac{\hat{\gamma}_i}{se(\hat{\gamma}_i)}$$

The hypothesis testing about parameters γ_i is very important, since if we reject the null hypothesis (H_0: $\gamma_i = 0$), this implies that parameter γ_i is statistically significant at level α, and the corresponding intercept parameter for unit i, is $\alpha + \gamma_i$, and is statistically different than the intercept (α) of the baseline cross sectional unit 1.

Panel data models: Fixed Effects

Hypothesis testing (F-test)

- **(ii) F-test:**
  $H_0: \gamma_2 = \gamma_3 = \cdots = \gamma_N = 0$
  $H_1: not\ H_0$
  The null hypothesis implies the following restricted model:
  $$H_0:\ Y_{it} = \alpha + \beta X_{it} + u_{it}$$
  while the alternative hypothesis implies the unrestricted model:
  $$H_1:\ Y_{it} = \alpha + \beta X_{it} + \gamma_2 d_{2t} + \gamma_3 d_{3t} + \cdots + \gamma_N d_{Nt} + u_{it}$$

- The F-test statistic can be used:
  $$F = \frac{(RSS_R - RSS_{Unr})/(df_R - df_{Unr})}{RSS_{Unr}/df_{Unr}} \sim F_{\mathrm{N-1,NT-N-1}}$$

- RSS_R and RSS_{Unr} are the residual sum of squares of the restricted and the unrestricted model, respectively

- $df_R = NT - 2$ and $df_{Unr} = NT - 2 - (N-1) = NT - N - 1$ are the degrees of freedom of the restricted and the unrestricted model, respectively, and $df_R - df_{Unr} = [NT-2] - [NT - N - 1] = N - 1$

Panel data models: Fixed Effects

Within Estimator

- Consider the following fixed effects panel data model:

$$\mathrm{Y}_{it} = \alpha_i + \beta \mathrm{X}_{it} + u_{it} \quad (1)$$

- Define

$$\bar{Y}_i = \frac{1}{T}\sum_{t=1}^{T} \mathrm{Y}_{it}, \qquad \bar{X}_i = \frac{1}{T}\sum_{t=1}^{T} X_{it}, \qquad \bar{u}_i = \frac{1}{T}\sum_{t=1}^{T} u_{it}$$

- Then

$$\mathrm{Y}_{it} = \alpha_i + \beta \mathrm{X}_{it} + u_{it} \Rightarrow$$

$$\Rightarrow \frac{1}{T}\sum_{t=1}^{T} \mathrm{Y}_{it} = \frac{1}{T}\sum_{t=1}^{T} \alpha_i + \frac{1}{T}\beta \sum_{t=1}^{T} \mathrm{X}_{it} + \frac{1}{T}\sum_{t=1}^{T} u_{it} \Rightarrow$$

$$\Rightarrow \bar{Y}_i = \frac{1}{T}\mathrm{T}\alpha_i + \beta \bar{\mathrm{X}}_i + \bar{u}_i \Rightarrow$$

$$\Rightarrow \bar{Y}_i = \alpha_i + \beta \bar{\mathrm{X}}_i + \bar{u}_i \quad (2)$$

Panel data models: Fixed Effects

Within Estimator

- By taking the difference (1)-(2) in equations (1) and (2)

$$\mathrm{Y}_{it} = \alpha_i + \beta \mathrm{X}_{it} + u_{it} \quad (1)$$

and

$$\bar{Y}_i = \alpha_i + \beta \bar{\mathrm{X}}_i + \bar{u}_i \quad (2)$$

we obtain:

$$\mathrm{Y}_{it} - \bar{Y}_i = \beta(\mathrm{X}_{it} - \bar{\mathrm{X}}_i) + (u_{it} - \bar{u}_i)$$

$$Y_{it}^* = \beta X_{it}^* + u_{it}^* \quad (3)$$

- Within Estimator steps:
  - Apply OLS in equation (3) and estimate β, i.e. obtain $\hat{\beta}$
  - Using equation (2), estimate α_i, i.e. $\hat{\alpha_i} = \bar{Y}_i - \hat{\beta}\bar{\mathrm{X}}_i$
  - Computationally easy
  - Can not conduct hypothesis testing

Panel data models: Fixed Effects

First-difference Estimator

- Another way of estimating panel data models is by first-differencing the data: lagging the model and subtracting, the time-invariant components are eliminated, and the model

$$\Delta \mathrm{Y}_{i,t} = \beta \Delta \mathrm{X}_{i,t} + \Delta u_{i,t}$$

can be consistently estimated by pooled OLS. This is called the first-difference estimator

- The differences are defined as follows $\Delta \mathrm{Y}_{i,t} = \mathrm{Y}_{i,t} - \mathrm{Y}_{i,t-1}$, $\Delta \mathrm{X}_{i,t} = \mathrm{X}_{i,t} - \mathrm{X}_{i,t-1}$, $\Delta u_{i,t} = u_{i,t} - u_{i,t-1}$, for $t = 2, \dots, T$

- Its relative efficiency, and so reasons for choosing it against other consistent alternatives, depends on the properties of the error term. The first-difference estimator is usually preferred if the errors u_{it} are strongly persistent in time, because then the $\Delta u_{i,t}$ will tend to be serially uncorrelated.

Panel data models: Random Effects

The Random effects model

- Consider the following panel data model:

$$\mathrm{Y}_{it} = \alpha_i + \beta \mathrm{X}_{it} + u_{it} \quad (1),$$

$$u_{it} \sim iid(0, \sigma^2)$$

- Suppose that $a_i \sim D(a, \omega^2)$, that is: $E(a_i) = a$ and $V(a_i) = \omega^2$
- Then, that $a_i = a + \mu_i$ (2), $\mu_i \sim iid\ D(0, \omega^2)$, and $\mathrm{E}(\mu_i) = 0$, and $\mathrm{V}(\mu_i) = \omega^2$
- In this model, we have two parameters (a, ω^2) instead of N parameters ($a, \mathrm{N} - 1$ dummy variables) with respect to the intercept
- The model can be written:

$$\mathrm{Y}_{it} = \alpha_i + \beta \mathrm{X}_{it} + u_{it} \Rightarrow$$

$$\Rightarrow \mathrm{Y}_{it} = a + \mu_i + \beta \mathrm{X}_{it} + u_{it} \Rightarrow$$

$$\Rightarrow \mathrm{Y}_{it} = a + \beta \mathrm{X}_{it} + (\mu_i + u_{it}) \Rightarrow$$

$$\Rightarrow \mathrm{Y}_{it} = a + \beta \mathrm{X}_{it} + v_{it}, \qquad \text{where} \qquad v_{it} = \mu_i + u_{it}$$

Panel data models: Random Effects

The Random effects model

- For the process

$$v_{it} = \mu_i + u_{it}$$

we observe that: $v_{i1} = \mu_i + u_{i1}, v_{i2} = \mu_i + u_{i2}, \ \dots, \ v_{iT} = \mu_i + u_{iT}$. Therefore, v_{it} contains a common component, μ_i, across time, and thus v_{it} is auto-correlated.

- Its mean is: $E(v_{it}) = E(\mu_i + u_{it}) = E(\mu_i) + E(u_{it}) = 0$

- The variance is:

$$V(v_{it}) = V(\mu_i + u_{it}) = V(\mu_i) + V(u_{it}) = \omega^2 + \sigma^2$$

- The covariance of v_{it} with v_{is}, $t \neq s$, is:

$$Cov(v_{it}, v_{is}) = E(v_{it} v_{is}) = E[(\mu_i + u_{it})(\mu_i + u_{is})] =$$

$$= E[\mu_i^2 + \mu_i u_{is} + u_{it}\mu_i + u_{it} u_{is}] =$$

$$= E\left[\mu_i^2\right] = V(\mu_i) = \omega^2 \neq 0,$$ there is auto-correlation at v_{it}

Panel data models: Random Effects

The Random effects model

- Therefore, the covariance matrix of $\boldsymbol{v_i} = (v_{i1}, v_{i2}, \dots, v_{iT})'$ is

$$\boldsymbol{\Omega} = \begin{bmatrix} \omega^2+\sigma^2 & \omega^2 & \cdots & \omega^2 \\ \omega^2 & \omega^2+\sigma^2 & \cdots & \omega^2 \\ \vdots & \vdots & \ddots & \vdots \\ \omega^2 & \omega^2 & \cdots & \omega^2+\sigma^2 \end{bmatrix}$$

- However, there is not heteroskedasticity. The covariances between the elements of $\boldsymbol{v_i} = (v_{i1}, v_{i2}, \dots, v_{iT})'$ and $\boldsymbol{v_j} = \left(v_{j1}, v_{j2}, \dots, v_{jT}\right)'$ are given by:

$$Cov\left(v_{it}, v_{jt}\right) = E\left(v_{it}v_{jt}\right) = E\left[(\mu_i + u_{it})(\mu_j + u_{jt})\right] =$$

$$= E[\mu_i\mu_j + \mu_i u_{jt} + u_{it}\mu_j + u_{it}u_{jt}] = 0$$

- Thus,

$$\text{Cov}\begin{pmatrix} \boldsymbol{v_1} \\ \boldsymbol{v_2} \\ \vdots \\ \boldsymbol{v_N} \end{pmatrix} = \begin{bmatrix} \boldsymbol{\Omega} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Omega} \end{bmatrix}$$

Panel data models:

Fixed or Random Effects Model (Hausman test)

- **Decide between fixed or random effects model: we can run the Hausman test**
- Null hypothesis: the preferred model is random effects
  vs. the alternative the fixed effects (see Green, 2008, chapter 9).
- Steps:
  - Run a fixed effects model and save the estimates
  - Run a random model and save the estimates
  - Perform Hausman test. If the p-value is significant (for example a=0.05 > p-value) then use fixed effects, if not use random effects

- R command:
  - `fixed <- plm(y ~ sf1, data=data, index=c("country"), model="within")`
  - `random <- plm(y ~ sf1, data=data, index=c("country"), model="random")`
  - `phtest(fixed, random)`

Panel data models:

Tests - Diagnostics

- **Breusch-Pagan Lagrange Multiplier for random effects**
- The LM test helps you decide between a random effects regression and a simple OLS regression
- The null hypothesis in the LM test: no panel effect (i.e. OLS better). That is, no significant difference across units (i.e. no panel effect)

- R command:
  - pool <- plm(y ~ sf1, data = data, model="pooling")
  - plmtest(pool, type=c("bp"))

Panel data models:

Tests - Diagnostics

- **Breusch-Pagan Lagrange Multiplier test for cross-sectional dependence in panels**
- Cross-sectional dependence is a problem in panel data (especially with long time series)
- The null hypothesis in the Breusch-Pagan/LM Cross-sectional dependence tests is that residuals across units are not correlated
- Breusch-Pagan/LM (cross-sectional dependence) tests are used to test whether the residuals are correlated across units

- R commands:
  - Breusch-Pagan Lagrange Multiplier test for cross-sectional dependence in panels
  - `fixed <-plm(y ~ sf1, data=data, index=c("country"), model="within")`
  - `pcdtest(fixed, test = c("lm"))`

  - Pesaran test for cross-sectional dependence in panels
  - `pcdtest(fixed, test = c("cd"))`

Panel data models:

Testing for serial correlation

- **Breusch-Godfrey/Wooldridge test for serial correlation in panel models**
- Serial correlation tests apply to panel data. The null is that there is not serial correlation

- R commands: (required package: lmtest)
  - fixed <-plm(y ~ sf1, data=data, index=c("country"), model="within")
  - pbgtest(fixed)

Panel data models:

Testing for heteroskedasticity

- **Breusch-Pagan test for heteroskedasticity**
- The null hypothesis for the Breusch-Pagan test is homoskedasticity
- If hetersokedaticity is detected we can use robust covariance matrix to account for it, or model the conditional variances (using, for example, ARCH/GARCH type models)

- R commands: (required package: lmtest)
  - bptest(y ~ sf1 + factor(country), data = data, studentize=F)

Panel data models:

Controlling for heteroskedasticity

- Robust covariance matrix estimation (Sandwich estimator)
- The ‘vcovHC’ function estimates three heteroskedasticity-consistent covariance estimators:
  - "white1": for general heteroskedasticity but no serial correlation (recommended for random effects)
  - "white2“: is "white1" restricted to a common variance within groups (recommended for random effects)
  - "arellano“: both heteroskedasticity and serial correlation (recommended for fixed effects)
- The following options can be applied:
  - HC0 : heteroskedasticity consistent (default)
  - HC1,HC2, HC3 : Recommended for small samples. HC3 gives less weight to influential observations
  - HC4 : small samples with influential observations
  - HAC : heteroskedasticity and autocorrelation consistent (type ?vcovHAC for more details)
- R commands: (required package: lmtest)
  - fixed <-plm(y ~ sf1, data=data, index=c("country"), model="within")
  - coeftest(fixed) # Original coefficients
  - coeftest(fixed, vcovHC) # Heteroskedasticityconsistent coefficients
  - coeftest(fixed, vcovHC(fixed, method = "arellano")) # Heteroskedasticity consistent coefficients (Arellano)
  - random <-plm(y ~ sf1, data=data, index=c("country"), model="random")
  - coeftest(random) # Original coefficients
  - coeftest(random, vcovHC) # Heteroskedasticity consistent coefficients

Panel data models: Application to R

- Several panel data models will be implemented in R
- See corresponding R-file

Thank you