

# List 01. Panel analysis

Nikita V. Artamonov

September 27, 2024

Necessary Python package: `pandas`, `numpy`, `linearmodels`<sup>1</sup>

## Contents

<b>1 Baseline model</b>	<b>1</b>
<b>2 Models with lags &amp; differences</b>	<b>5</b>

## 1 Baseline model

#1. For the panel `Grunfeld` consider a regression

inv on value & capital

**Remark:** here `firm` is an individual index, `year` is a time index.

1. write down the specification of a baseline model
2. import the dataset and set a MultiIndex over a pandas DataFrame (please, mind individual and time indices).
3. Is the panel balanced or not?
4. Evaluate matrices of regression design  $\mathbf{y}$ ,  $\mathbf{X}$
5. Fit Pooling, RE and FE (within-estimators) models

---

<sup>1</sup>`conda install -c conda-forge linearmodels pip install linearmodels`

- via matrices of regression design
- via regression specification

and report the fitting result.

6. For each model test the significance of coefficients. Perform both non-robust and robust test.
7. Test overall significance of each model. Perform both non-robust and robust test
8. Perform Hausman test

**#2** (Gun Laws and the Effect on Crime). For the panel **Guns** consider a regression

log(violent) on law (+ another regressors)

**Remark:** here **state** is an individual index, **year** is a time index.

1. write down the specification of a baseline model
2. import the dataset and set a MultiIndex over a pandas DataFrame (please, mind individual and time indices).
3. Is the panel balanced or not?
4. Evaluate matrices of regression design  $\mathbf{y}, \mathbf{X}$  (law is categorical!)
5. Fit Pooling, RE and FE (within-estimators) models
  - via matrices of regression design
  - via regression specification

and report the fitting result.

6. For each model test the significance of coefficients. Perform both non-robust and robust test.
7. Test overall significance of each model. Perform both non-robust and robust test
8. Perform Hausman test

#3. For the panel `Emp1UK` consider a regression

$\log(\text{emp})$  on  $\log(\text{wage})$ ,  $\log(\text{capital})$ ,  $\log(\text{output})$

**Remark:** here `firm` is an individual index, `year` is a time index.

1. write down the specification of a baseline model
2. import the dataset and set a MultiIndex over a pandas DataFrame (please, mind individual and time indices).
3. Is the panel balanced or not?
4. Evaluate matrices of regression design  $\mathbf{y}$ ,  $\mathbf{X}$
5. Fit Pooling, RE and FE (within-estimators) models
  - via matrices of regression design
  - via regression specification

and report the fitting result.

6. For each model test the significance of coefficients. Perform both non-robust and robust test.
7. Test overall significance of each model. Perform both non-robust and robust test
8. Perform Hausman test

#4. For the panel `Wages`<sup>2</sup> consider a regression

$\text{lwage}$  on  $\text{ed}$ ,  $\text{exp}$ ,  $\text{exp}^2$ ,  $\text{south}$ ,  $\text{smsa}$ ,  $\text{married}$ ,  $\text{bluecol}$

**Remark:** here `id` is an individual index, `time` is a time index.

1. write down the specification of a baseline model
2. import the dataset and set a MultiIndex over a pandas DataFrame (please, mind individual and time indices).
3. Is the panel balanced or not?

---

<sup>2</sup>In the panel `ed` is time-invariant!

4. Evaluate matrices of regression design  $\mathbf{y}, \mathbf{X}$
5. Fit Pooling, RE and FE (within-estimators) models
  - via matrices of regression design
  - via regression specification
 and report the fitting result.
6. For each model test the significance of coefficients. Perform both non-robust and robust test.
7. Test overall significance of each model. Perform both non-robust and robust test
8. Perform Hausman test

**#5** (Production costs of Turkish banks). For the panel `TurkishBanks` consider a regression

$\log(\text{cost})$  on  $\log(\text{output})$  (+ another regressors)

**Remark:** here `id` is an individual index, `year` is a time index.

1. write down the specification of a baseline model
2. import the dataset and set a MultiIndex over a pandas DataFrame (please, mind individual and time indices).
3. Is the panel balanced or not?
4. Evaluate matrices of regression design  $\mathbf{y}, \mathbf{X}$
5. Fit Pooling, RE and FE (within-estimators) models
  - via matrices of regression design
  - via regression specification
 and report the fitting result.
6. For each model test the significance of coefficients. Perform both non-robust and robust test.
7. Test overall significance of each model. Perform both non-robust and robust test
8. Perform Hausman test

## 2 Models with lags & differences

#1. For the panel `Grunfeld` consider a regression

$$\Delta \text{inv} \text{ on } \Delta \text{value} \ \& \ \Delta \text{capital}$$

**Remark:** here `firm` is an individual index, `year` is a time index.

1. Fit Pooling, RE and FE (within-estimators) models and report the fitting result.
2. For each model test the significance of coefficients. Perform both non-robust and robust test.
3. Test overall significance of each model. Perform both non-robust and robust test
4. Perform Hausman test

#2. For the panel `Grunfeld` consider a regression

$$\text{inv}_{it} \text{ on } \text{value}_{it}, \text{value}_{i,t-1}, \text{capital}_{it}, \text{capital}_{i,t-1}$$

**Remark:** here `firm` is an individual index, `year` is a time index.

1. Fit Pooling, RE and FE (within-estimators) models and report the fitting result.
2. For each model test the significance of coefficients. Perform both non-robust and robust test.
3. Test overall significance of each model. Perform both non-robust and robust test
4. Perform Hausman test

#3. For the panel `EmplUK` consider a regression

$$\Delta \log(\text{emp}) \text{ on } \Delta \log(\text{wage}), \Delta \log(\text{capital}), \Delta \log(\text{output})$$

**Remark:** here `firm` is an individual index, `year` is a time index.

1. Fit Pooling, RE and FE (within-estimators) models and report the fitting result.

2. For each model test the significance of coefficients. Perform both non-robust and robust test.
3. Test overall significance of each model. Perform both non-robust and robust test
4. Perform Hausman test

#4. For the panel **EmplUK** consider a regression

$$\log(\text{emp})_{it} \text{ on } \log(\text{wage})_{it}, \log(\text{wage})_{i,t-1}, \log(\text{capital})_{it}, \log(\text{capital})_{i,t-1}, \\ \log(\text{output})_{it}, \log(\text{output})_{i,t-1}$$

**Remark:** here **firm** is an individual index, **year** is a time index.

1. Fit Pooling, RE and FE (within-estimators) models and report the fitting result.
2. For each model test the significance of coefficients. Perform both non-robust and robust test.
3. Test overall significance of each model. Perform both non-robust and robust test
4. Perform Hausman test

#5 (Production costs of Turkish banks). For the panel **TurkishBanks** consider a regression

$$\Delta \log(\text{cost}) \text{ on } \Delta \log(\text{output}) (+ \text{ another regressors})$$

**Remark:** here **id** is an individual index, **year** is a time index.

1. Fit Pooling, RE and FE (within-estimators) models and report the fitting result.
2. For each model test the significance of coefficients. Perform both non-robust and robust test.
3. Test overall significance of each model. Perform both non-robust and robust test
4. Perform Hausman test

#6 (Production costs of Turkish banks). For the panel `TurkishBanks` consider a regression

$\log(\text{cost})_{it}$  on  $\log(\text{output})_{it}$ ,  $\log(\text{output})_{i,t-1}$ , (+ another regressors)

**Remark:** here `id` is an individual index, `year` is a time index.

1. Fit Pooling, RE and FE (within-estimators) models and report the fitting result.
2. For each model test the significance of coefficients. Perform both non-robust and robust test.
3. Test overall significance of each model. Perform both non-robust and robust test
4. Perform Hausman test