



Microeconometrics

Aneta Dzik-Walczak

Kateryna Zabarina

2022/2023

Microeconometrics

□ Classes:

- OLS revision
- Initial Data Analysis, OLS estimation, diagnostic tests
- Panel Data Analysis (RE, FE)
- Advanced Panel Models
- LPM, Logit, Probit, Ordered Models
- Test
- Students' Presentations

Microeconometrics

- 1 absence is allowed
- Class grade (each part has to be passed)
 - Model (50%)
 - Presentation (20%)
 - Test (30%)
- Final grade (Jerzy Mycielski)
 - Based on class grade and exam grade

Model

- Introduction
- Economic theory
- Literature review (5 empirical articles)
 - Hypothesis
- Dataset description
- Initial data analysis
 - Variables description
 - Descriptive statistics
 - Graphical analysis (histogram, box-plot)
 - Relations between variables (correlation, scatterplot)
 - Tests (e.g. Equality of Means)
- Estimation
- Parameters interpretation
- Verification of hypothesis
- Summary (literature context)

Parametric vs. non-parametric tests

	Parametric tests		Non-parametric tests
2 independent samples	Small sample, normal distribution, T test	Large sample, T test	Mann-Withney
2 dependent samples	Small sample, normal distribution, T test	Large sample, T test	Wilcoxon
>2 independent samples	Small sample, normal distribution, ANOVA	Large sample, ANOVA	Kruskal-Wallis
	H0 rejection		
	Tuckey, Bonnferoni, Scheffe		Mann-Withney
>2 dependent samples	Small sample, normal distribution, ANOVA	Large sample, ANOVA	Friedman
	H0 rejection		
	Tuckey, Bonnferoni, Scheffe		Wilcoxon

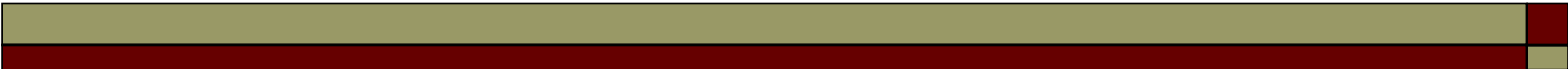
Presentation

□ Introduction

- Main goal

- Motivation

} like in Bachelor Thesis



Presentation

- Economic Theory
 - Main idea
 - Definitions
 - Assumptions



Presentation

- Hypotheses
 - Statements that can be verified
 - Women earn less than men

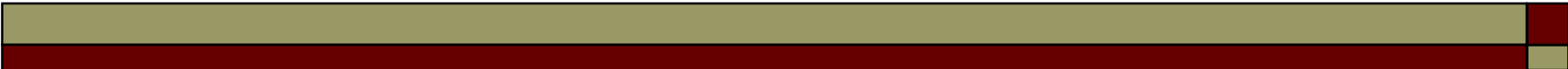
Presentation

□ Results

- Presented in a friendly and transparent way
 - No screenshots from STATA

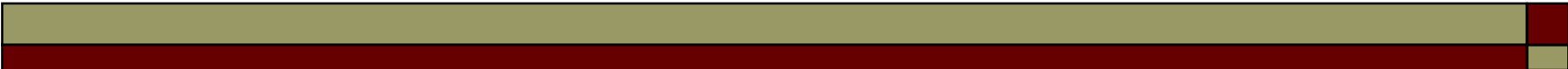
Variable	Hypothesis	Parameter estimate
Sex 0-man 1-womam	-	-100 **
Experience	+	30 ***

- Significance level: * 10%, **5%, *** 1%
- Coefficient interpretation
- Summary in the context of the hypotheses and literature



Deadlines

- Investigated problem
 - 09.01.2023 (email)
- Presentation
 - 23.01.2023 (email)
- Model (email: report, dataset, do-file)
 - 23.01.2023



Models from previous years

- ❑ The Effect of Human Capital on the Performance of Enterprises in an Industrial Cluster in Northern Vietnam
- ❑ Why FDI impacts on economic growth in Sub-Saharan African countries?
- ❑ The effective elements on women participation in the labor market-Poland
- ❑ Short- term impact of public debt and fiscal deficit on growth
- ❑ The impact of the WTO on its members' trade. Panel data analysis
- ❑ WHO GETS THE JOB? ANALYZING FACTORS DETERMINING PROBABILITY OF BEING EMPLOYED
- ❑ Economic status of Spanish-speaking Hispanic US citizens – the case of Puerto Rico

[illegible]

Ordinary Least Squares (OLS)

$$Y = X\beta + \varepsilon$$

- Wooldridge, Jeffrey M, *Econometric analysis of cross section and panel data*, The MIT Press, Cambridge 2002

OLS - coefficients interpretation

```
xi: reg dochg i.sex i.gredu
```

```
i.sex          _Isex_1-2          (naturally coded; _Isex_1 omitted)
i.gredu        _Igredu_1-4        (naturally coded; _Igredu_1 omitted)
```

Source	SS	df	MS	Number of obs = 20884		
Model	1.3302e+10	4	3.3255e+09	F(4, 20879) = 577.49		
Residual	1.2023e+11	20879	5758596.66	Prob > F = 0.0000		
Total	1.3354e+11	20883	6394479.9	R-squared = 0.0996		
				Adj R-squared = 0.0994		
				Root MSE = 2399.7		

dochg	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
_Isex_2	-695.714	36.61821	-19.00	0.000	-767.4885	-623.9394
_Igredu_2	-1391.613	47.21045	-29.48	0.000	-1484.149	-1299.077
_Igredu_3	-2005.016	47.66322	-42.07	0.000	-2098.44	-1911.593
_Igredu_4	-2431.217	64.16855	-37.89	0.000	-2556.992	-2305.441
_cons	5361.137	41.00044	130.76	0.000	5280.773	5441.501

OLS - coefficients interpretation

dochg	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
_Isex_2	-695.714	36.61821	-19.00	0.000	-767.4885	-623.9394
_Igredu_2	-1391.613	47.21045	-29.48	0.000	-1484.149	-1299.077
_Igredu_3	-2005.016	47.66322	-42.07	0.000	-2098.44	-1911.593
_Igredu_4	-2431.217	64.16855	-37.89	0.000	-2556.992	-2305.441
_cons	5361.137	41.00044	130.76	0.000	5280.773	5441.501

- Explained variable:
 - Income (dochg)- continuous variable (unit = 1 euro)
- Explanatory variables:
 - Sex (sex) – binary variable (1-man, 2 -woman)
 - Education level (gredu) – discrete variable (1- higher, 2- secondary, 3- lower secondary, 4- primary)
- Women have a lower income than men by 695 Euros (on average)
- People with secondary education have income lower by 1391 Euros than those with higher education (on average)
- People with lower secondary education have income lower by 2005 Euros than those with higher education (on average)
- People with primary education have income lower by 2431 Euros than those with higher education (on average)

OLS - coefficients interpretation

```
reg so dochg sex
```

	so	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
	+						
	dochg	.0000332	6.78e-07	48.90	0.000	.0000319	.0000345
	sex	-.0053723	.0036936	-1.45	0.146	-.0126121	.0018674
	_cons	.0577133	.0058828	9.81	0.000	.0461825	.0692441

- Explained variable:
 - Savings rate (so) - continuous variable (unit= 1 percentage point)
- Explanatory variables:
 - Income (dochg) – continuous variable (unit= 1 euro)
 - Sex (sex) – binary variable (1-man, 2-woman)
- We interpret coefficients as **partial effect**:
- The increase in income by one unit increases the savings rate by 0.00003 units. More specifically: increase in income by 1 Euro increases the savings rate by 0.00003 percentage points.
- Knowing the properties of OLS model, we know that the scaling of variable leads to an appropriate scaling of parameter. Thus, the interpretation may be as follows: Increase in income by 1000 Euros increases the savings rate by 0.03 percentage points.

OLS - coefficients interpretation

```
reg lnso lndochg sex
```

lnso	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lndochg	.0425764	.0139643	31.69	0.000	.4152048	.4699479
sex	-.0038458	.0166148	-0.23	0.817	-.0364127	.0287211
_cons	-5.162073	.1176933	-43.86	0.000	-5.392765	-4.93138

- Explained variable
 - Logarithm of savings rate (lnso)- continuous variable
- Explanatory variables
 - Logarithm of income (lndochg) –continuous variable
 - Sex (sex) – binary variable (1-man, 2-woman)
- We interpret lndochg coefficient as **elasticity**
 - The increase in income by 1% increases the savings rate by 0,04%
- We interpret sex coefficient as **semi-elasticity**
 - Women have a lower savings rate than men on average by 0,3%

OLS - coefficients interpretation

```
reg lnso dochg
```

lnso	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
+						
dochg	.0000825	2.85e-06	28.99	0.000	.0000769	.0000881
_cons	-1.898676	.0135798	-139.82	0.000	-1.925294	-1.872058

- Explained variable
 - Logarithm of savings rate (lnso) - continuous variable (unit= 1 Euro)
- Explanatory variables
 - Income (dochg) – continuous variable
- We interpret dochg coefficient as **semi-elasticity**
 - The increase in income by 1 Euro increases the savings rate by 0,008%

Example

Source	SS	df	MS	Number of obs = 20884		
Model	153.615074	5	30.7230149	F(5, 20878) = ...		
Residual	1270.62287	20878	.060859415	Prob > F = 0.0000		
Total	1424.23795	20883	.068200831	R-squared = ...		
				Adj R-squared =		
				Root MSE = .2467		

so	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
dochg	.0000353	7.11e-07	49.62	0.078	.0000339	.0000367
_Isex_2	.0014056	.0037969	0.37	0.000	-.0060366	.0088477
_Igredu_2	.031791	.0049533	6.42	0.320	.0220821	.0414999
_Igredu_3	.0423637	.0051033	8.30	0.000	.0323607	.0523666
_Igredu_4	.0594209	.0068197	8.71	0.000	.0460537	.0727881
_cons	.0101066	.0056846	1.78	0.075	-.0010356	.0212488

- The model for savings rate was estimated. Explanatory variables: respondent's income (*dochg*); sex (*sex*): woman *sex*=0, man *sex*=1; education (*gredu*): primary *gredu*=1, gymnasium *gredu*=2, secondary *gredu*=3, higher *gredu*=4.

■ Assume **significance level $\alpha = 0.05$**

1. Compute R^2 and F statistics.
2. Comment fit of the model.
3. Are variables jointly significant?
4. Which variables are significant?
5. Interpret model coefficients.

Hint: Remember to give names of used tests, test statistics, null hypothesis.