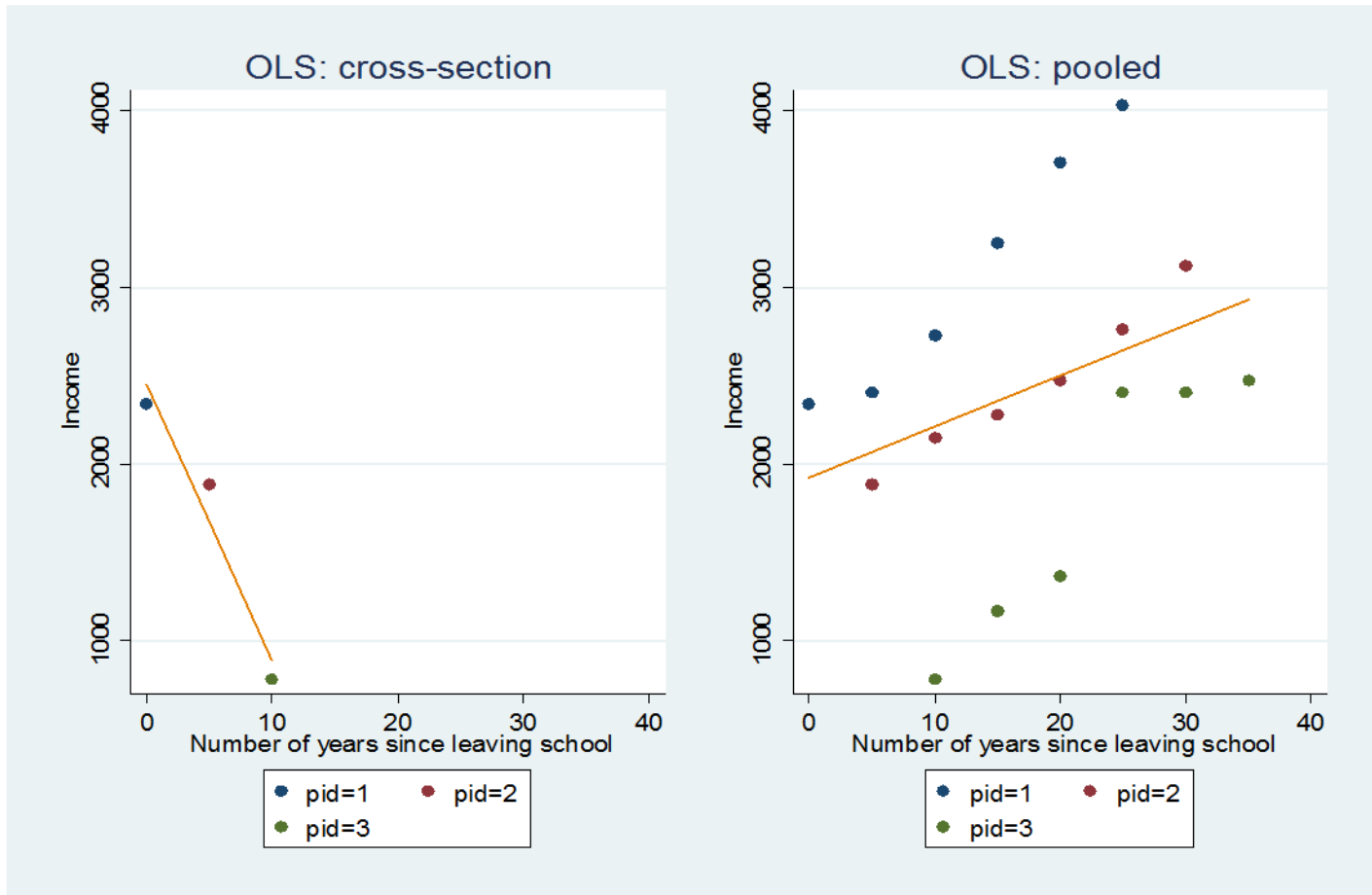


# Analiza danych panelowych

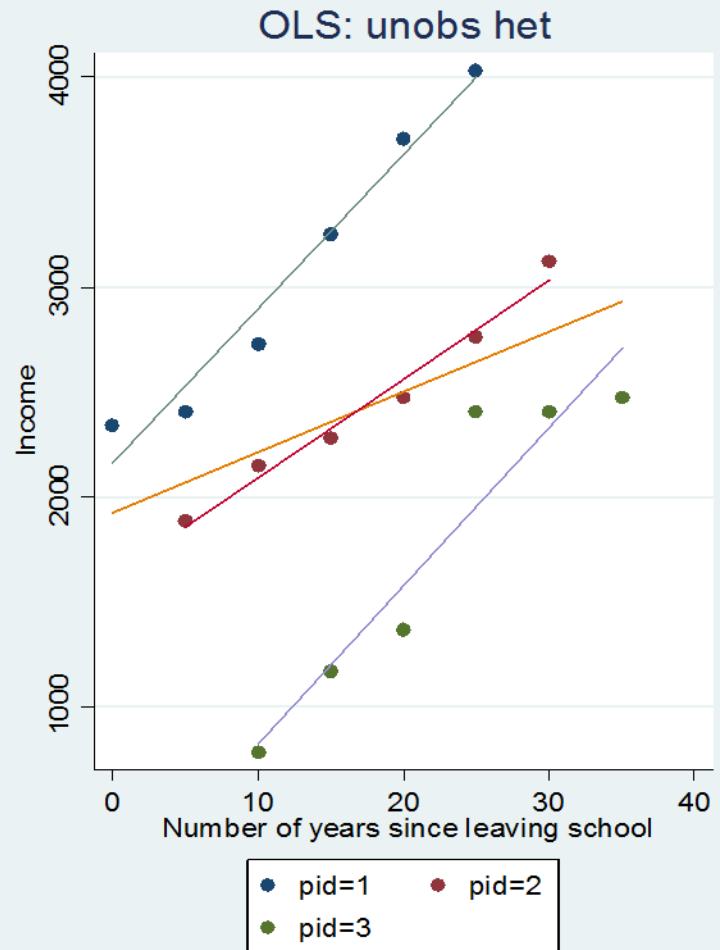
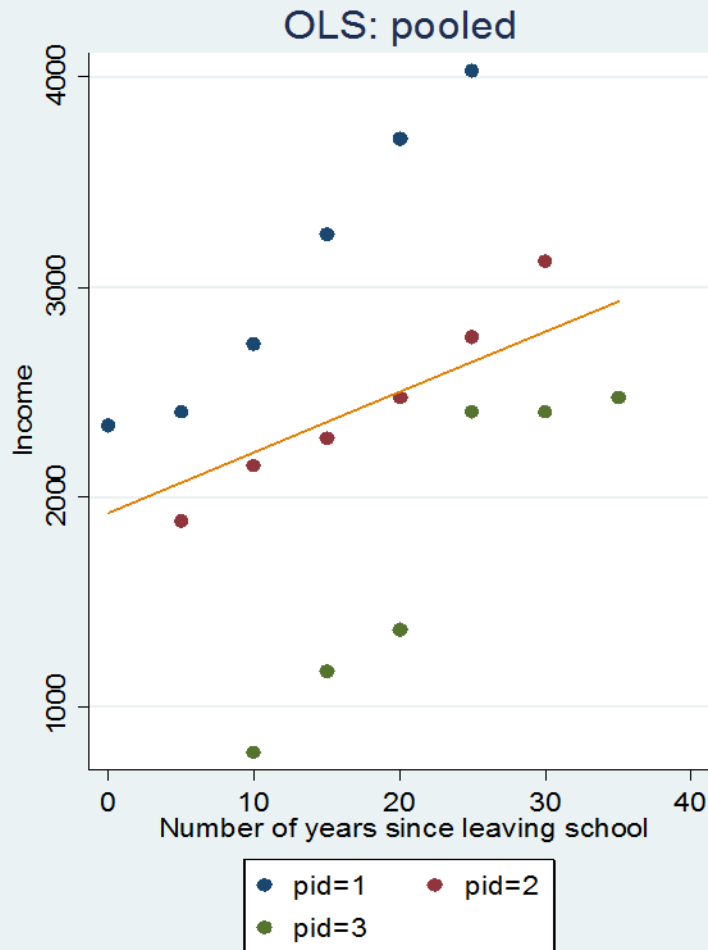
**Stanisław Cichocki**

**Natalia Nehrebecka**

# Plan zajęć



# Plan zajęć



# Plan zajęć

## 1. Wprowadzenie do danych panelowych

- a) Charakterystyka danych panelowych
- b) Zalety i ograniczenia

## 2. Modele ekonometryczne danych panelowych

- a) Model efektów nieobserwowalnych
- b) Model efektów losowych
- c) Model efektów stałych
- d) Modele z dwukierunkowym komponentem błędu

## 3. Testy

- a) Test istotności efektów stałych i losowych
- b) Test Hausmana

# Plan zajęć

1. Wprowadzenie do danych panelowych
  - a) Charakterystyka danych panelowych
  - b) Zalety i ograniczenia
  
2. Modele ekonometryczne danych panelowych
  - a) Model efektów nieobserwowalnych
  - b) Model efektów losowych
  - c) Model efektów stałych
  - d) Modele z dwukierunkowym komponentem błędu
  
3. Testy
  - a) Test istotności efektów stałych i losowych
  - b) Test Hausmana

# Dane statystyczne

- ▶ Dane statystyczne dzieli się na trzy główne grupy:
- ▶ Szeregi czasowe
- ▶ Dane przekrojowe
- ▶ Dane panelowe
  - Dzielone są na dwie kategorie:
  - Dane „połączone” (*pooled data*)
  - Dane panelowe (właściwe)
    - Panel / longitudinal data

# Próby przekrojowo-czasowe

- ▶ Dane „połączone” (*pooled data*)
- ▶ Dane z kilku okresów czasu dotyczące **różnych** prób przekrojowych
- ▶ Dane „połączone” to taki rodzaj danych, w którym nie rozróżniamy kolejnych jednostek badania.
- ▶ **Korzyści:** zwiększenie liczby stopni swobody z  $N$  na  $NT$  ( $N$  – liczba obserwacji,  $T$  – liczba okresów)

# Próby przekrojowo-czasowe

<i>obsno</i>	<i>year</i>	<i>hprice</i>	<i>proptax</i>	<i>sqrft</i>	<i>bdrms</i>	<i>bthrms</i>
1	1993	85500	42	1600	3	2.0
2	1993	67300	36	1440	3	2.5
3	1993	134000	38	2000	4	2.5
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
250	1993	243600	41	2600	4	3.0
251	1995	65000	16	1250	2	1.0
252	1995	182400	20	2200	4	2.0
253	1995	97500	15	1540	3	2.0
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
520	1995	57200	16	1100	2	1.5



# Dane panelowe

## ▶ Dane panelowe właściwe

- Dane panelowe mają cechy zarówno próby przekrojowej, jak i szeregu czasowego.
- Zawierają one szereg obserwacji dla danej próby przekrojowej.
- Każda obserwacja przypisana jest jednej konkretnie jednostce badania za pomocą zmiennej, która identyfikuje tę jednostkę.
- Jednostki badania mogą być różnych rozmiarów: np. kraj, firma, rodzina, poszczególne osoby.
- W danych panelowych występuje także zmienna wskazująca na okres, w którym dana obserwacja była zebrana (kwartał, rok, itd.).

# Dane panelowe

A Two-Year Panel Data Set on City Crime Statistics

<i>obsno</i>	<i>city</i>	<i>year</i>	<i>murders</i>	<i>population</i>	<i>unem</i>	<i>police</i>
1	1	1986	5	350000	8.7	440
2	1	1990	8	359200	7.2	471
3	2	1986	2	64300	5.4	75
4	2	1990	1	65100	5.5	75
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
297	149	1986	10	260700	9.6	286
298	149	1990	6	245000	9.8	334
299	150	1986	25	543000	4.3	520
300	150	1990	32	546200	5.2	493

# Dane panelowe - zbilansowane

. *xtset country year*

panel variable: country (strongly balanced)

time variable: year, 1990 to 1999

delta: 1 unit

# Dane panelowe - zbilansowane

## *xtdes*

```
country: 1, 2, ..., 7                n = 7
year: 1990, 1991, ..., 1999         T = 10
Delta(year) = 1 unit
Span(year) = 10 periods
(country*year uniquely identifies each observation)
```

```
Distribution of T_i:  min    5%    25%    50%    75%    95%    max
                   10     10     10     10     10     10     10
```

Freq.	Percent	Cum.	Pattern
7	100.00	100.00	1111111111
7	100.00		XXXXXXXXXXXX

# Dane panelowe - niezbilansowane

. *xtset idcode year*

panel variable: idcode (unbalanced)

time variable: year, 68 to 88, but with gaps

delta: 1 unit

# Dane panelowe - niezbilansowane

## *xtdes*

```

idcode: 1, 2, ..., 5159          n =          4709
year:   68, 69, ..., 88          T =           15
Delta(year) = 1 unit
Span(year)  = 21 periods
(idcode*year uniquely identifies each observation)
Distribution of T_i:  min      5%      25%      50%      75%      95%      max
                   1        1        3        5        9       13       15

```

Freq.	Percent	Cum.	Pattern
136	2.89	2.89	1.....
114	2.42	5.31	.....1
89	1.89	7.20	.....1.11
87	1.85	9.05	.....11
86	1.83	10.87	<b>111111.1.11.1.11.1.11</b>
61	1.30	12.17	.....11.1.11
56	1.19	13.36	11.....
54	1.15	14.50	.....1.1.11
54	1.15	15.65	.....1.11.1.11.1.11
3972	84.35	100.00	(other patterns)
4709	100.00		XXXXXX.X.XX.X.XX.X.XX

# Pytania teoretyczne

1. Czym różni się panel od próby przekrojowo-czasowej?

**Dziękuję za uwagę**