

STATISTICAL MODELING AND CAUSAL INFERENCE WITH R

Week 9: Panel Data

Manuel Bosancianu

November 9, 2020

Hertie School of Governance

Max Schaub

Applied example

Example taken from Brüderl and Ludwig (2015), based on SOEP.

Standard setup for panel data:

- ✓ sample of households, with everyone in HH interviewed (including foreigners)
- ✓ ongoing since 1984 (expanded in 1990 to cover former GDR)
- ✓ annual interviews (occasionally, top-up households added)
- ✓ expanded again in 2016 to cover refugees

Extensive potential

- ✓ social inequalities
 1. perceptions on incomes
 2. social mobility
 3. changes in income distribution

- ✓ well-being and health
 1. impact of Covid-19
 2. impact of aging

- ✓ migrant integration
 1. language skills and job market success
 2. requalification in host country
 3. success of second generation

Impact of marriage

Impact of first marriage on life satisfaction?

Problem: happier people are more likely to get married (Frey & Stutzer, 2005).

Similarities re. self-selection: impact of institutions on development.

Variables:

- ✓ **id**: person ID
- ✓ **year**: wave
- ✓ **marry**: married
- ✓ **yrsmarried**: marriage length
- ✓ **age**: R's age
- ✓ **loghhinc**: HH income (log)
- ✓ **woman**: gender
- ✓ **happy**: life satisfaction

Who to include?

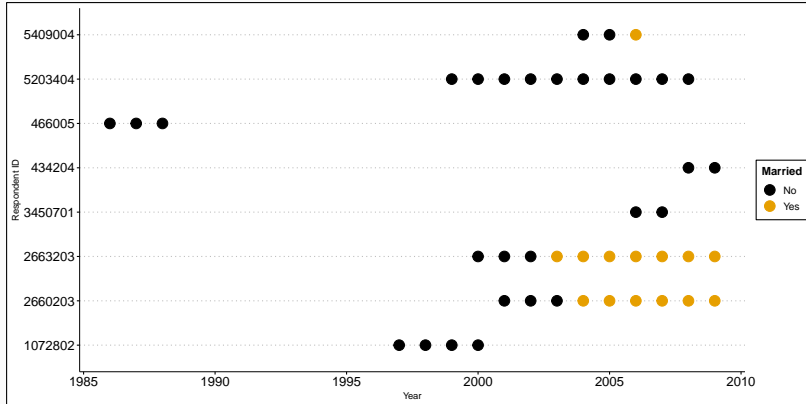
Any unit who can potentially change from not-treated to treated (Brüderl & Ludwig, 2015; Sobel, 2012).

If unit's treatment status changed during period, they get included.

If unit never got treatment, they get included (control group).

What about already treated? If treatment effect varies over time, including them might bias estimation (older vs. younger).

Trajectory of 8 units



Explore variation

How is the ratio “between-unit / within-unit” in terms of variation?

	Gender	Satisfaction	Marriage
Mean	0.467	7.147	0.230
σ_{within}	0	1.260	0.267
$\sigma_{between}$	0.499	1.282	0.269

Gender doesn't have any within-unit variation for us.

Fixed effects specifications

We can't use gender in this specification.

	DV: satisfaction
Marriage	0.167*** (0.017)
HH income (log)	0.125*** (0.009)
Age	-0.041*** (0.001)
R ²	0.016
Adj. R ²	-0.119
Num. obs.	121919

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Can't add either gender or age in this model.

	DV: satisfaction
Marriage	0.163*** (0.017)
HH income (log)	0.126*** (0.009)
R ²	0.003
Adj. R ²	−0.134
Num. obs.	121919

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Should we trust SE?

Assumption for errors is $v_{it} \sim iid(0, \sigma_v^2)$.

With longer panels, serial correlation between errors is a problem: $Cor(v_{it}, v_{i(t-1)}) \neq 0$.

```
pbgtest(model2, order = 2)
```

^^IBreusch-Godfrey/Wooldridge test for serial correlation in panel models

data: happy ~ marry + loghhinc

chisq = 473.01, df = 2, p-value < 2.2e-16

alternative hypothesis: serial correlation in idiosyncratic errors

Manual correction needed.

Robust standard errors

Correction done on SEs to factor in serial correlation.

```
coeftest(model2,  
         vcov. = function(x) vcovHC(x, type = "sss"))
```

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
marry	0.163145	0.022726	7.1789	7.072e-13	***
loghhinc	0.125558	0.012250	10.2497	< 2.2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Random effects specification

Random effects logic

$$\begin{cases} \text{Happy}_{it} = \beta_{i0} + \beta_1 * \text{Marriage}_{it} + \beta_2 * \text{Age}_{it} + \beta_3 * \text{Income}_{it} + \epsilon_{it} \\ \beta_{i0} = \gamma_{00} + \gamma_{01} * \text{Gender}_i + v_i \end{cases} \quad (1)$$

The unit fixed effects (actually, the intercepts) are modeled with 2nd equation.

Goal is to explain their variance (instead of treating them as nuisance).

REs results

```
model3 <- plm(happy ~ marry + loghhinc + age + woman, index = c("id", "year"), effect = "twoways",  
model = "random", data = df.happy)
```

Function specifies random effects, as well as shape of correlation.

Serial correlation of errors still needs to be addressed.

REs results

DV: satisfaction	
(Intercept)	5.736*** (0.069)
Marriage	0.134*** (0.011)
HH income (log)	0.193*** (0.006)
Age	-0.020*** (0.001)
Woman	0.058*** (0.015)
s_idios	1.331
s_id	1.048
s_time	0.087
R ²	0.023
Adj. R ²	0.023
Num. obs.	121919

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Other questions

Same design, different question

What is the impact of social media use on mental health, e.g. depression?

Original Investigation

FREE

July 15, 2019

Association of Screen Time and Depression in Adolescence

Elroy Boers, PhD^{1,2}; Mohammad H. Afzali, PhD^{1,2}; Nicola Newton, PhD³; [et al](#)

[» Author Affiliations](#) | [Article Information](#)

JAMA Pediatr. 2019;173(9):853-859. doi:10.1001/jamapediatrics.2019.1759

The screenshot shows the top portion of a Scientific American article. At the top left is a 'Latest Issues' button. The center features the 'SCIENTIFIC AMERICAN' logo with a '175' anniversary badge. To the right are links for 'Cart 0', 'Sign In', and 'Store'. Below the header is a 'MENTAL HEALTH' category bar. The main title 'Does Social Media Cause Depression?' is prominently displayed in a large, bold, serif font. Below it, the subtitle 'The answer's not black-and-white' is in a smaller, regular serif font. A thin horizontal line separates the subtitle from the byline, which reads 'By Jade Wu Savvy Psychologist on January 16, 2020'.

Latest Issues

SCIENTIFIC AMERICAN 175

Cart 0 Sign In | Store

MENTAL HEALTH

Does Social Media Cause Depression?

The answer's not black-and-white

By Jade Wu Savvy Psychologist on January 16, 2020

Discussion topics

- ✓ How would you design a study to assess this?
- ✓ What variables would you need to study this?
- ✓ How often do you need to collect data for this design?
- ✓ What is treatment and comparison group?

Thank **you** for the kind attention!

- Brüderl, J., & Ludwig, V. (2015). Fixed-effects panel regression. In H. Best & C. Wolf (Eds.), *The SAGE Handbook of Regression Analysis and Causal Inference* (pp. 327–357). London: SAGE Publications.
- Frey, B. S., & Stutzer, A. (2005). Happiness Research: State and Prospects. *Review of Social Economy*, 63(2), 207-228.
- Sobel, M. E. (2012). Does Marriage Boost Men's Wages? Identification of Treatment Effects in Fixed Effects Regression Models for Panel Data. *Journal of the American Statistical Association*, 107(498), 521–529.