Health Econometrics

Health Economics and Management

Tutorials

Ylenia Brilli
ylenia.brilli@univr.it

A.Y. 2019-2020

Info

- My email address: ylenia.brilli@univr.it
- **Office hours:** before/after the tutorials (contact me before)
- **Problem set:** out on October 14th, due on **October 25th**
- **By October 7th**, email me your group information

Tutorials:

1. Sept 26th, 13-16, Lab Scaravilli
2. Sept 30th, 12-13.30, Lab Ranzani
3. Oct 3rd, 13-16, Lab Scaravilli (lab booked until 17.00)
4. Oct 14th, 11-13.30, Lab Ranzani

Lab material is available at this Dropbox link [protected with password]

Lab 1 - Panel data

Aim of the analysis: study the effect of **health status** on **income**

1. Explore a panel dataset
2. Pooled OLS
3. First Difference
4. STATA command for fixed effects

Data

- **German Socio-Economic Panel** (GSOEP), available on the *Journal of Applied Econometrics Data Archive*
- Used in the paper by Riphahn, Wambach, and Million, "Incentive Effects in the Demand for Health Care: A Bivariate Panel Count Data Estimation", Journal of Applied Econometrics, Vol. 18, No. 4, 2003, pp. 387-405.
- Individuals observed across several years;
- information about socio-demographic characteristics, economic status and health-related behaviours
- Further information on the data/variable here

Estimation model - Pooled OLS

$$hhninc_{it} = \beta_0 + \beta_1 hsat_{it} + X_{it}^{'}\beta_k + D_{year}^{'}\delta + a_i + v_{it}$$

- *hhninc* is yearly income measured in 10.000 German marks
- *hsat* is an index of health status (0-10 scale)
- *X* is a vector of explanatory variables of our choice
- $D_{year}^{'}$ is a vector of yearly dummies (to control for the *time trends*)
- a_i is the **unobserved fixed heterogeneity**

What is the expected sign of $\hat{\beta}_1$?

Estimation model - First difference

$$\Delta hhninc_{it} = \beta_0 + \beta_1 \Delta hsat_{it} +_{it}^{'} \beta_k + \Delta D_{year}^{'} \delta + a_i + v_{it}$$

- create manually each new variable
- time-invariant variables (e.g. *female*) will not enter the equation
- the constant becomes a time trend

Lab 2 - Research on Health Policy

- Overview of Regression Discontinuity designs
- Analysis of a health policy using cross-sectional data

Sharp RD

- Rules to assign treatments may be arbitrary and may be used to analyze the effects of policies
- Sharp Regression Discontinuity (RD) is used when the treatment status is a *deterministic* and *discontinuous* function of an observable variable
- Examples:
  - students obtain scholarship awards if the score they obtain at the end of high school is above a certain level
  - youths are allowed to drink/drive whenever their age is above 16/18
  - families are entitled to cash-transfers if their income is below a certain threshold

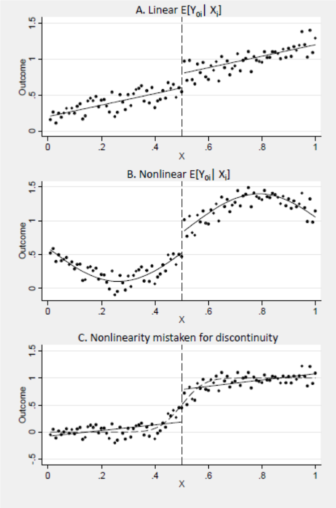


Figure: RD and non-RD examples. Source: Angrist and Pischke 2008

The RD estimation model

- $y_i = \alpha + \beta x_i + \rho D_i + \eta_i$
- ρ is the effect of the treatment on the outcome(s)
- Here, D_i is not only correlated with x_i, but also a deterministic function of it:

$$D_i = \begin{cases} 1 & \text{if} \quad x_i \geq 0.5 \\ 0 & \text{if} \quad x_i < 0.5 \end{cases}$$

- The RD method captures the coeff. of interests by distinguishing the discontinuous fnc $1(x_i \geq 0.5)$ from the smooth fnc x_i

Baseline assumptions in RD designs

1. The treatment changes discontinuously at the cutoff
2. Individuals are not able to manipulate x_i in order to be treated
    - Check the distribution of x_i: there should not be heapings
    - Other predetermined characteristics must be continuous at the cutoff
3. There should not be other relevant changes occurring at the cutoff
    - The only change should be the change in treatment status

$\Rightarrow$ The assumptions above are more likely to hold in *the vicinity* of the cutoff (**bandwidth selection**)

The paper:

Vaccination take-up and health:
evidence from a flu vaccination program for the elderly

joint work with Claudio Lucifora (Catholic Univ. Milan), Marco Tonello (Bank of Italy) and Antonio Russo (Agency for Health Protection-Milan)

- Seasonal influenza is a serious public health issue
- It causes severe illness and death, especially for the elderly population
- Main strategy to protect at-risk categories, such as the elderly, has been to implement *vaccination* programs and policies
- EU Council Recommendation (2009): EU member states shall implement policies aimed at reaching a vaccination coverage among the older age groups of 75%
- Many countries provides free flu vaccination for individuals above a certain age:
  - US and majority of EU countries use 65 years as age-threshold
  - Others use younger ages (60, 59, 55)

This paper

1. We assess the effects of a flu vaccination program for the elderly implemented in Italy on vaccination take-up and health
2. According to the program, the flu vaccine is provided for free to individuals aged 65 or more
3. We adopt a RD strategy around the age cutoff
4. We use data from the metropolitan area of Milan, North-Italy for the flu vaccination campaign 2013

The influenza vaccination program in Italy

- Vaccination against influenza regulated by the National Plan for Preventive Vaccination (NPPV)
- Flu vaccination campaign between late Oct and end of Dec (while the virus circulates between Nov and April in the following year)
- NPPV establishes that vaccination shall be free for:
    1. the **elderly**, from the year they turn 65
    2. individuals affected by a chronic disease
    3. individuals institutionalized in nursing homes
    4. Women in 2nd/3rd trimester of pregnancy
- For the categories above, the GP provides the vaccine and makes the injection
- The GP shall also actively reach the targeted categories above
- The others shall buy the vaccine at the pharmacy and pay the doctor/a nurse for the injection

Data and sample

- Data from **Agency for Health Protection of Milan** (North-West of Italy):
  - Metropolitan area (MA) of Milan: ~ 3 millions inhabitants
- **General Health register**, flu season 2013-2014:
  - demographic characteristics (gender, municipality of residence, date of birth)
  - date of flu vaccination (if any)
  - whether individual is exempted from cost-sharing because of chronic disease, low income or both
  - physician's characteristics: age, experience, number of patients
- **Hospitalizations records**, years 2013-2014:
  - occurrence & duration of hospitalization
  - type of hospitalization (planned or emergency)
- **Sample:** individuals aged 64 or 65 in the year 2013, excluding:
  1. those with a disability
  2. those institutionalized in nursing homes

Variables of interests

- Running variable: date of birth (dob)
- An individual is *treated* if aged 65 in 2013, i.e. born before Jan 1, 1949
- Vaccination take-up:
  $V_i = 1$ if individual took the flu vaccine in Oct-Dec 2013
- Health outcomes:
  - Probability of hospitalization (if occurred at least once)
  - Number of days at hospital
  - The health outcomes are defined over the period when the influenza virus was active (Mid Oct, 2013 - Mid April, 2014)

See descriptive stats

Identification strategy

$$y_i = \alpha + \beta ET_i + ET_i(dob - 1Jan1949) + (1 - ET_i)(dob - 1Jan1949) + \epsilon_i$$

- y_i indicates any outcome variable (e.g., vaccination or health)
- $ET_i = 1$ if the individual is aged 65 in 2013, i.e., eligible for free vaccination

Identifying assumptions

1. The age threshold determines a jump in the probability of getting the vaccination
2. Absence of manipulation in the running variable
   $\Rightarrow$ For the baseline analysis, we drop individuals born at the cutoff date (Jan 1st, 1949) and on the day before/after the cutoff date (*donut1* specification)
3. Continuity of other characteristics

Universal eligibility to free vaccination and take up

Parametric estimates

	(1)	(2)	(3)	(4)	(5)	(6)
(A) BW=1 month						
Treated	0.068***	0.068***	0.066***	0.068***	0.074**	0.072**
	(0.017)	(0.018)	(0.017)	(0.017)	(0.030)	(0.030)
N	5618	5618	5618	5618	5618	5618
(B) BW=3 months						
Treated	0.060***	0.059***	0.059***	0.060***	0.054***	0.051***
	(0.009)	(0.010)	(0.009)	(0.009)	(0.015)	(0.015)
N	17276	17276	17276	17276	17276	17276
(C) BW=6 months						
Treated	0.062***	0.062***	0.061***	0.062***	0.055***	0.055***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.010)	(0.010)
N	35398	35398	35398	35398	35398	35398
(D) BW=12 months						
Treated	0.080***	0.079***	0.079***	0.080***	0.058***	0.058***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.008)	(0.008)
N	68962	68962	68962	68962	68962	68962
Covariates		✓		✓		✓
Specification:						
(i) Linear						
(ii) Linear with interaction	✓	✓				
(iii) Quadratic			✓	✓		
(iv) Quadratic with interaction					✓	✓

Universal eligibility for free vaccination and health

Parametric estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Bandwidth size:	BW=1m				BW=3m				BW=6m				BW=12			
Panel A. Prob. Hospitalization																
Treated	-0.013	-0.014	-0.013	-0.013	-0.003	-0.003	-0.003	-0.017*	0.004	0.005	0.004	-0.003	0.002	0.002	0.002	0.005
	(0.011)	(0.011)	(0.011)	(0.018)	(0.006)	(0.006)	(0.006)	(0.009)	(0.004)	(0.004)	(0.004)	(0.006)	(0.003)	(0.003)	(0.003)	(0.005)
Panel B. No. Of days at hospital																
Treated	-0.102	-0.104	-0.102	0.041	-0.085	-0.084	-0.085	-0.196	0.008	0.010	0.009	-0.082	-0.010	-0.010	-0.010	0.015
	(0.178)	(0.176)	(0.177)	(0.279)	(0.097)	(0.096)	(0.097)	(0.151)	(0.063)	(0.063)	(0.063)	(0.101)	(0.048)	(0.048)	(0.048)	(0.073)
N	5618	5618	5618	5618	17276	17276	17276	17276	35398	35398	35398	35398	68962	68962	68962	68962
Covariates	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Specification:																
(i) Linear	✓				✓				✓				✓			
(ii) Linear with interac		✓				✓				✓				✓		
(iii) Quadratic			✓				✓				✓				✓	
(iv) Quadratic with interac				✓				✓				✓				✓

Lab 3 - Difference in Differences (DID)

Aim of the analysis: study the effect of a *vaccination program* on vaccination take-up

1. Explore the data
2. Compute the DID estimator from sample means
3. Estimate DID specification
4. Show time trends in a graph

Background

- The materials for this tutorial have been taken from the paper by C. Carpenter and E. Lawler (2019), **Direct and spillover effects of Middle School Vaccination Requirements**, *American Economic Journal: Economic Policy*, 2019, 11(1), 95-125
- You can find the paper & replication files here

Background for the analysis

- The paper considers the introduction of Middle School vaccination mandates (for several diseases) between 2005 and 2015 and analyzes their effects on take-up and diseases incidence
- In our exercise:
  - we focus on the vaccination against *Tetanus*, *Diphtheria* and *Pertussis*;
  - we consider the introduction of a mandate at Middle School entry for a booster (the *Tdap* vaccine) in 2005 in selected US states;
  - we look at the Tdap vaccination take-up among 13-year-olds

Diseases covered by the Tdap vaccine

- Tetanus, Diphtheria and Pertussis are bacterial diseases, for which vaccination, especially among young children, has been strongly recommended since the 1940s
- Pertussis (or *whooping cough*) is highly contagious, and can have serious consequences especially among infants and young children (high rate of hospitalizations and of severe complications).

The Tdap vaccine mandate

- In the US it is very common to have school-based mandatory vaccination laws as tool to increase the take-up.
- For tetanus, diphtheria and pertussis, the first Tdap vaccine mandate at Middle School entry was introduced in 2005 by 5 states (Florida, Kentucky, Massachusetts, Montana and Wyoming)
- However, exemptions could be obtained (for religious or philosophical reasons)

Data

- Data on timing of Tdap vaccine mandates are linked to data on adolescent vaccination from the **National Immunization Survey-Teen**
- This is a repeated cross-section, sampling adolescents between 13 and 17 years of age
- The data provides info on whether the teen received the Tdap vaccine between 10 and 13 years of age
- The data also gives demographic information on the teen and the household

Timing

- We observe the Tdap vaccine take-up of 13-year-olds in 2004, 2005 and 2006
- Starting from 2005, the Treated states introduced the Tdap vaccine mandate at Middle School entry
- The other states introduced similar mandates later, but we do not observe these later periods in the data
- Thus, they act as Control states.

Treatment and Control groups

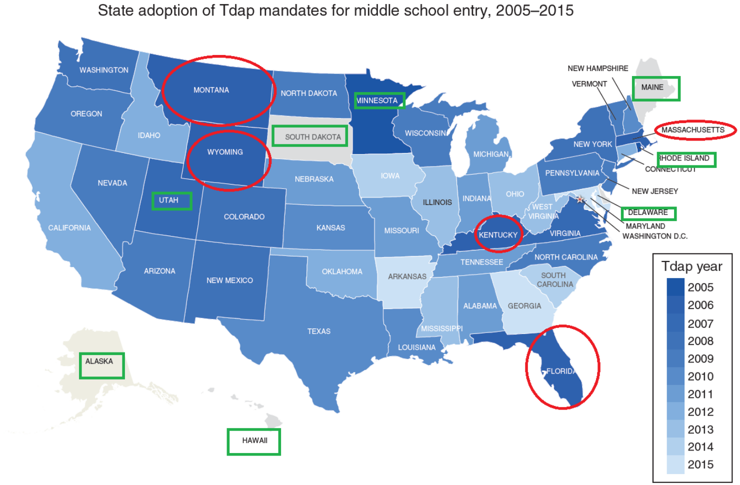
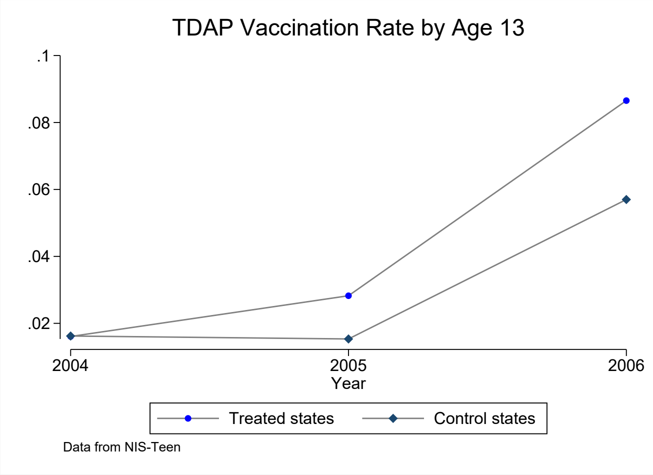


Figure: Timing of Tdap Mandate adoption. Source: Carpenter and Lawler 2019 (circles and squares added)

Trends in Treated and Control states



DID from comparison of sample means

	Before	*After*
Treated	0.01607	0.06292
Control	0.01621	0.04041
DID	0.02265	

DID estimation model

$$y_{ist} = \beta_0 + \delta_0 post_{it} + \beta_1 Treated_{si} + \delta_1 post_{it} * Treated_{si} + \gamma x_{ist} + v_{it}$$

- i individual, s state, t year
- y_{ist} is the probability that 13-year-old i has received the Tdap vaccine
- $post_{it} = 1$ for years 2005 and 2006
- $Treated_{si} = 1$ identifies the Treated states
- x_{ist} is a vector of controls at the individual i and/or state s levels

Some remarks...

- Since we only have 3 periods, all the time trends are absorbed by the $post$ dummy
- However, in general, you should control for year dummies as well

- When using DID it is important to check the trends in Treated and Control states *before* the policy change
- Here, we only have one pre-policy period
- Ideally, you should have more