Lecture 13: Panel Data (Part III)

Jaakko Meriläinen

5304 Econometrics @ Stockholm School of Economics

Extending Panel Data Models

- In the previous lectures, we have seen the following models looking at variation over time:
  - Difference-in-differences
  - First-differences
  - Fixed effects

- We will conclude the discussion on panel models by discussing extensions: non-panel difference in differences, difference-in-differences with IV, triple-differenced models, difference-in-discontinuities, and synthetic controls

Plan for This Lecture: Extensions

1. Introduction
2. Non-panel difference-in-differences
3. Difference-in-differences with IV
4. Triple-differenced models
5. Difference-in-discontinuities
6. Synthetic controls
7. References

Non-Panel Difference-in-Differences

- Sometimes you see difference-in-differences applications that are not using panel data per se
- Let us take an example from my on-going research
- Meriläinen and Mitrunen study the impacts of exposure to Marxist propaganda in elementary school on labor-market outcomes in adulthood
  - Meriläinen and Mitrunen use a rogue school experiment in a Finnish municipality in the 1970s
  - Kids on the 5th grade were taught material taken from the Soviet Union
  - The experiment generates variation across cohorts (so no time variation per se) and geographical variation
  - Meriläinen and Mitrunen use a generalized difference-in-differences approach and document negative effects on adulthood incomes
  - These negative effects seem to run through occupational choice—exposed individuals became more likely to choose "left-wing occupations" that also have lower wages

The Pirkkala Experiment

Source: Archives of Yle.

Figure 1. A demonstration against the Finnish society.

Table 1. Word counts in the Pirkkala handout and in a core curriculum text book.

Word	Detected word in Finnish	Count in Pirkkala	Count in core curriculum
(1)	(2)	(3)	(4)
Work	työ	339	57
Worker	työläis / työvä	111	11
Peasant	talonpoik	91	31
Class	luok	33	2
Ruling class	hallitsev luok	7	0
Corporation	yritys / yhtiö	20	0
Revolution	vallankum	22	0
Socialism	sosialis	8	0
Capitalism	kapital	25	0
Soviet Union	neuvostol	11	1
United States	yhdysval	10	0
Total word count		18622	22715

Notes: The table reports certain word frequencies counted in the 5th grade history materials. The words in the column (3) are counted from the socialist Pirkkala handout Historia: 5. luokka: ihmiskunnan kehityksen yleispiirteet vanhimmista ajoista nykypäiviin saakka. The word counts in column (4) are drawn from the replaced core curriculum 5th-grade textbook *Peruskoulun historia. 1, Viidettä kouluvuotta varten* by authors Lehtonen and Huttunen. We scanned the original material and used OCR on Python to convert the text into data. From these data, we count words using the roots of the Finnish words to account for inflection.

The Pirkkala Experiment: Sources of Identification

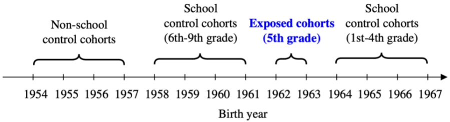


Figure 3. Cohort variation.

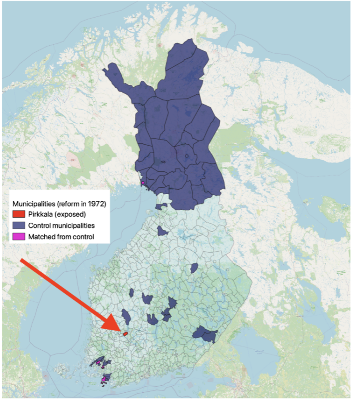


Figure 4. Geographical variation.

Notes: The map shows the municipality exposed to the curriculum experiment (Pirkkala), the other municipalities that had the comprehensive school reform at the same time (plausible control municipalities), as well as the municipalities that are chosen by propensity score matching as the most alike municipalities to Pirkkala.

The Pirkkala Experiment: Empirical Approach

- Meriläinen and Mitrunen estimate

$$Y_{imc} = \beta_c \mathbf{1}[Cohort]_c \times \mathbf{1}[Pirkkala]_m + \lambda_m + \delta_c + \varepsilon_{icm}$$

  where
  - β_c are the cohort-specific effects
  - $\mathbf{1}[5th\ grader\ 1973-1975]_c$ is an indicator for the individual being a 5th-grader in 1973-1975
  - $\mathbf{1}[Pirkkala]_m$ is an indicator for the individual living in Pirkkala in 1970
  - λ_m present the municipality (or finer urban area) fixed effects
  - δ_c present cohort fixed effects

- Other controls include municipality characteristics (1970) interacted with cohort fixed effects and gender fixed effects and household 1970 controls

Results: Income Effects

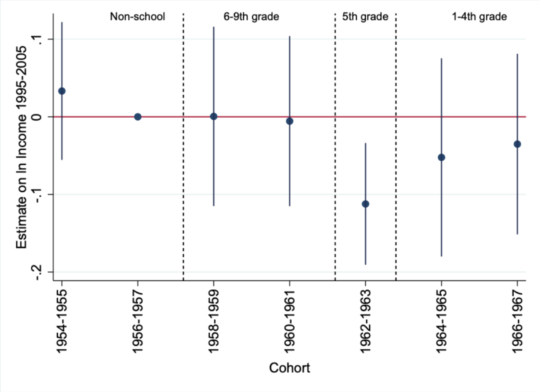


Figure 5. Estimated income effects by cohort.

Notes: The figure plots point estimates from an event-study specification and their 95% confidence intervals that are constructed using standard errors clustered at the municipality-cohort level. We use 1956-1957 cohorts as the base group. Here we use the full set of municipalities. Estimations include municipality fixed effects, cohort fixed effects. The vertical line marks the cohorts that were in school and the curriculum experiment exposed children (cohorts born 1962-1963).

Panel Models and IVs

- Fixed effects and first-differenced models help deal with **time-invariant** unobservables
- But in many settings, we may still be concerned with bias from **time-varying** unobservables
- In these settings, we may need to combine our methods
- It is straightforward to combine the panel methods we have studied (FE, FD) with IV models
- We will now look at a particular example
    - Again, trying to identify the causal returns to schooling in the labor market
    - But this time, in developing countries

Returns to Education in Indonesia (Duflo 2001)

- Duflo (2001) is the seminal paper in developing countries which estimates the causal effect of education on earnings; a classic question in labor economics (which we have now seen a few times!)

- **Context:** School building program in Indonesia
  - Between 1973-1974 and 1978-1979, 61,000 schools were built
  - Enrollment rate increased from 69% to 83% between 1973 and 1978
  - Number of schools built in each region was inversely related to number of children not enrolled in 1972

- **Data:**
  - Men born between 1950-1972 from 1995 census: birth date, salary level and level of education in 1995
  - Intensity of the building program (based on the program) in the birth region of each person

Identification and Estimation: Program Effects on Educational Attainment

- Exploit sources of variation in program intensity for individuals (number of schools built in each region + cohort of child) and estimate

$$S_{ijk} = c_1 + \alpha_{1j} + \beta_{1k} + \sum_{l=2}^{23}(P_j d_{il})\gamma_{1l} + \sum_{l=2}^{23}(\mathbf{C}_j d_{il})\delta_{1l} + \epsilon_{ijk}$$

  where

  S_{ijk} = education level of person i in region j in cohort k
  α_{1j}= district-of-birth fixed effects
  β_{1k}= birth cohort fixed effects
  P_j = 1 if there person was born in a high intensity region
  d_{il} = dummy variable for belonging to cohort l
  C_j = vector of regional characteristics

- Each γ_{1l} can be interpreted as an estimate of program impact for the given cohort

Program effects on educational attainment

Figure 1. Coefficients of the Interactions Age in 1974* Program Intensity in the Region of Birth in the Education Equation

Estimating the Returns to Education: Specifications and Identification

- Use variation in education predicted by program to estimate returns to schooling

- Second stage: $y_{ijk} = d + a_j + b_k + S_{ijk}b + \eta_{ijk}$

- Because S_{ijk} may be correlated with η_{ijk}, instrument it with
  - The vector of interaction terms $(P_j.d_{il})$
  - Or a single interaction between being in "young" cohort and program intensity in region of birth

- Identifying assumption(s):
  - The program is as good as randomly assigned (conditional on cohort and district dummies)
  - The program affected wages only through increasing attainment (and not directly).

Estimating the Returns to Education

TABLE 7—EFFECT OF EDUCATION ON LABOR MARKET OUTCOMES: OLS AND 2SLS ESTIMATES

Method	Instrument	(1)	(2)	(3)	(4)
Panel A: Sample of Wage Earners					
Panel A1: Dependent variable: log(hourly wage)					
OLS		0.0776	0.0777	0.0767	
		(0.000620)	(0.000621)	(0.000646)	
2SLS	Year of birth dummies*program intensity in region of birth	0.0675	0.0809	0.106	0.0908
		(0.0280)	(0.0272)	(0.0222)	(0.0541)
		[0.96]	[0.9]	[0.93]	[0.9]
2SLS	(Aged 2–6 in 1974)*program intensity in region of birth	0.0752	0.0862	0.104	
		(0.0338)	(0.0336)	(0.0304)	
		(0.0338)	(0.0336)	(0.0304)	

Triple-Differenced models Through Examples

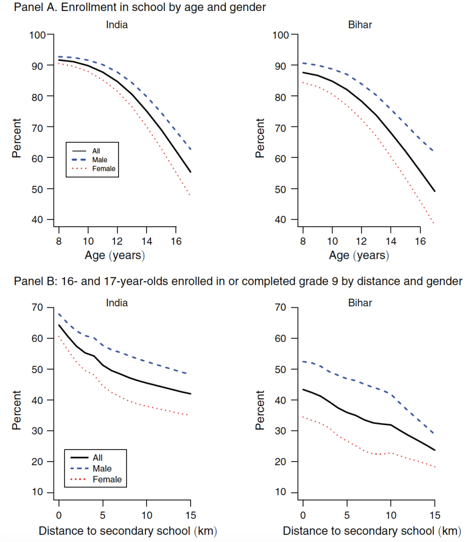
- Parallel trends assumption might sometimes be rejected, which would make a difference-in-differences model not suitable

- There may still be something else you could do... and the best way to look at this is through an empirical example

- Context of the "Cycling to School in India" paper by Muralidharan and Prakash (2017):
  - Girls drop out of secondary schooling before boys do in India
  - A particular barrier is just getting to school
  - There are time costs and safety costs, both of which increase in distance

- In 2006, a newly-elected state government in Bihar announced a program which provided money to buy a bicycle for all girls who enrolled in Grade 9
  - Akin to a Conditional Cash Transfer
  - The aim of the policy was to encourage girls to stay in school and finish secondary education

The Policy Issue



Thinking About Sources of Variation

- The main challenge for evaluation is that the policy was implemented in the entire state at once

- Within Bihar, there are two sources of variation:
  - Across ages—the policy applies only to incoming 9th grade students (~14 years old)
  - Across sex—the policy applies only to girls

- Unfortunately for the authors, the enrollment trends for boys and girls **are not parallel** (which they test for)

- So they use a third source of variation
  - Look at Bihar and bordering districts in the neighboring state
  - This neighboring state was a part of Bihar until 2001
  - Under the assumption that the trends for boys/girls are identical across these two neighboring states, can come up with a **triple-differences** approach (difference-in-difference-in-differences or DDD)

Trends in Enrollment for Boys and Girls Are Not Parallel in Bihar...

Table 1—Testing the Parallel Trends Assumption

	Dependent variable: log (9th grade enrollment by school, gender, and year)
Panel A. Testing parallel trends for the difference-in-differences (DD)	
Female × year	**0.0518**
	(0.00)
Female	–0.870
	(0.06)
Year (coded as 1 to 4)	0.0852
	(0.01)
Constant	4.235
	(0.05)
Observations	20,266
R^2	0.167

But Trends in Enrollment for Boys and Girls Are Parallel Across the Two States!

Panel B. Testing parallel trends for the triple differences (DDD)

Female × year × Bihar	**−0.0100**
	(0.01)
Female × year	0.0618
	(0.01)
Female × Bihar	0.175
	(0.11)
Bihar × year	0.0290
	(0.01)
Female	−1.045
	(0.09)
Year (coded as 1 to 4)	0.0562
	(0.01)
Bihar	−0.123
	(0.12)
Constant	4.358
	(0.11)
Observations	22,279
R^2	0.171

Empirical Specification

$$(1)\ y_{ihv} = \beta_0 + \beta_1 \cdot F_{ihv} \cdot T_{ihv} \cdot BH_{ihv} + \beta_2 \cdot F_{ihv} \cdot BH_{ihv} + \beta_3 \cdot T_{ihv} \times BH_{ihv} + \beta_4 \cdot F_{ihv} \cdot T_{ihv} + \beta_5 \cdot F_{ihv} + \beta_6 \cdot T_{ihv} + \beta_7 \cdot BH_{ihv} + \varepsilon_{ihv},$$

- Here

  y_{ihv} denotes outcomes for child i in household h and village v
  F_{ihv} is a female dummy
  T_{ihv} is an indicator for being in the treated cohort (14-15 years)
  BH_{ihv} is an indicator for being in Bihar

- Sample: all 14-17 year-olds in a detailed cross-sectional panel survey from 2007-2008 (˜18 months after the policy roll-out)

Results

Table 2—Triple Difference (DDD) Estimate of the Impact of Being Exposed to the Cycle Program on Girl's Secondary School Enrollment

Dependent variable: Enrolled in or completed grade 9				
Treatment group = age 14 and 15				
Control group = age 16 and 17	(1)	(2)	(3)	(4)
Treat × female × Bihar	**0.103**	**0.091**	**0.052**	**0.052**
	(0.030)	**(0.029)**	**(0.025)**	**(0.025)**
Treat × female	0.020	0.024	0.038	0.039
	(0.026)	(0.026)	(0.021)	(0.021)
Treat × Bihar	−0.044	−0.042	−0.029	−0.028
	(0.018)	(0.018)	(0.016)	(0.016)
Female × Bihar	−0.094	−0.091	−0.067	−0.066
	(0.023)	(0.023)	(0.020)	(0.020)
Treat	−0.148	−0.143	−0.138	−0.138
	(0.014)	(0.014)	(0.013)	(0.013)
Female	−0.092	−0.088	−0.100	−0.101
	(0.020)	(0.020)	(0.017)	(0.017)
Bihar	0.011	−0.044	−0.032	−0.044
	(0.016)	(0.016)	(0.015)	(0.015)
Constant	0.464	0.771	0.593	0.562
	(0.013)	(0.024)	(0.027)	(0.040)
Demographic controls	No	Yes	Yes	Yes
HH socioeconomic controls	No	No	Yes	Yes
Village level controls	No	No	No	Yes
Observations	30,295	30,295	30,147	30,112
R^2	0.035	0.088	0.207	0.208

Results: Mechanism

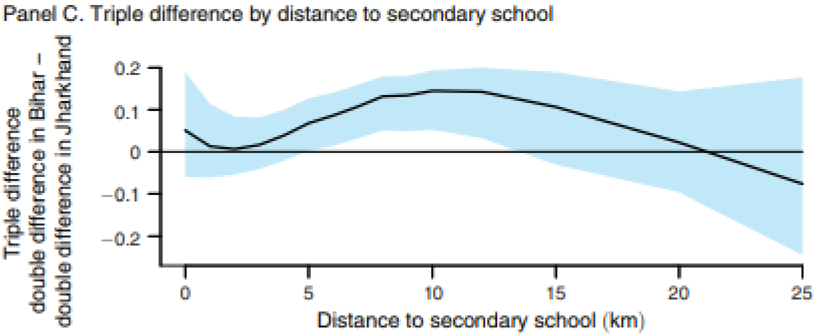


FIGURE 2. NON-PARAMETRIC DOUBLE AND TRIPLE DIFFERENCE ESTIMATES OF IMPACT OF THE CYCLE PROGRAM

(by distance to nearest secondary school)

Effects of EU Enlargement onf Labor-Market Outcomes

- In recent years, few things have shaped Western countries economically, politically, and socially the way that globalization and open borders have

- A prominent example is the European Union, the expansion of which has been speculated to be one of the key reasons for phenomena such as Brexit

- Kuosmanen and Meriläinen (forthcoming) study the labor-market implications of EU enlargement to
  - Many Estonians in particular came to Finland as "posted workers" after 2004
  - Many of these worked in construction-sector occupations that had no certification requirements ("vulnerable occupations") especially in Southern Finland (the Helsinki region)
  - This creates three sources of variation that Kuosmanen and Meriläinen can exploit
  - Main findings: "[...] robust evidence that the entry of new EU countries decreased the annual earnings of vulnerable workers relative to less vulnerable workers in the construction sector. This decrease was persistent [...] and economically meaningful, about 9 percent (roughly a month's salary per year) [...] a small increase in unemployment and provide evidence on different adjustment mechanisms."

A First Glance at Differences

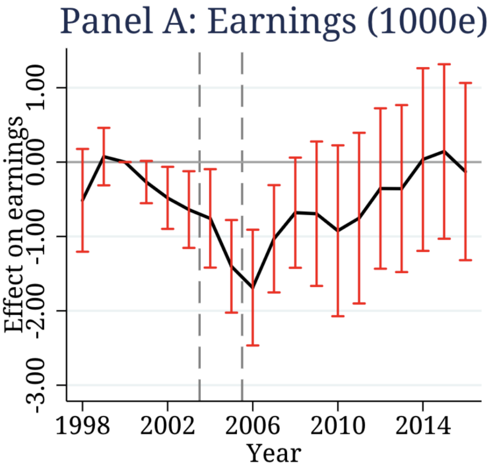
Table 1. EU enlargement and labor market outcomes: differences in means.

	Helsinki CZ			Control CZs		
	Vulnerable (1)	Control (2)	Difference (3)	Vulnerable (4)	Control (5)	Difference (6)
Panel A: Earnings						
Before	24.325	31.432	-7.106***	22.652	29.647	-6.995***
	(15.740)	(15.144)	(1.175)	(14.626)	(14.119)	(0.800)
	[74,809]	[16,154]	[90,963]	[72,543]	[15,823]	[88,366]
After	23.976	32.015	-8.039***	22.564	28.984	-6.420***
	(21.705)	(21.330)	(0.736)	(19.865)	(20.524)	(0.701)
	[151,752]	[33,352]	[185,104]	[148,540]	[32,518]	[181058]
Difference	-0.349	0.584	*DD*	-0.088	-0.663	*DD*
	(0.227)	(0.653)	-0.933	(0.246)	(0.554)	0.575
	[226,561]	[49,506]	(0.575)	[221,083]	[48,341]	(0.522)

DDD = -1.508* (0.766)

Notes: The table reports region- and occupation-level averages before and after the enlargement of the EU in 2004. Vulnerable occupations are builders, painters, carpenters, and plumbers. Control refers to electricians. Earnings are measured in thousands of 2015 euros. Standard errors clustered at the municipality level and reported in parentheses. Number of observations are shown in brackets. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Problem with Difference-in-Differences Comparing Treated and Non-Treated Regions



Triple-Differences Specification

- Kuosmanen and Meriläinen estimate

$$Y_{ioct} = \beta \text{Vulnerable}_o \times \text{Helsinki}_c \times \text{After}_t \\ + \lambda_i + \lambda_{ct} + \lambda_{ot} + \epsilon_{ioct}.$$

- Y_{ioct} refers to the worker-level outcome of an individual i in occupation o, commuting zone c, and year t
- Treated individuals are those who were builders, carpenters, painters, and plumbers (vulnerable occupations) in the year 2000
- Compare workers in these occupations with electricians for whom $\text{Vulnerable}_o = 0$
- The treatment dummy Helsinki_c takes the value 1 if the individual worked in the Helsinki region in the year 2000 (high-exposure commuting zone), and 0 otherwise
- The indicator variable After_t takes the value 1 after the enlargement of the EU in 2004, and is equal to zero before that
- β is the coefficient of interest; it captures the effect of EU enlargement on individuals in vulnerable occupations in the treated region
- Also include individual fixed effects (λ_i) to capture time-invariant differences between individuals, and net out commuting zone-time (λ_{ct}), and occupation-time (λ_{ot}) fixed effects

Triple Differences Results: Also Use Variation Across Occupations

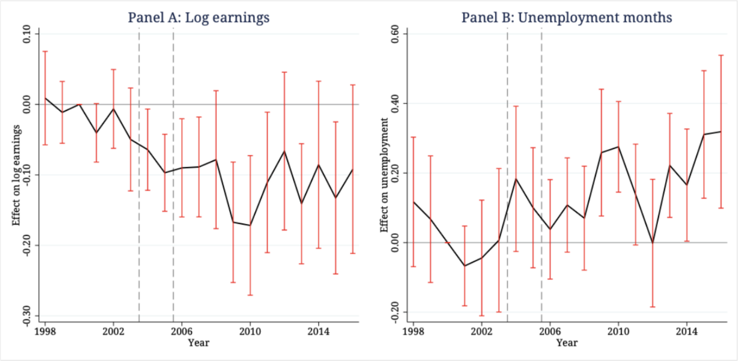


Figure 6. Effects of EU enlargement on annual earnings and unemployment months: event study specification.

Notes: Figure shows point estimates and their 95% confidence intervals constructed using standard errors clustered at the municipality level. Gray dashed lines mark the transition period.

Difference-in-Discontinuities

- Suppose our RDD and DiD (a dream couple!) had a baby
- That baby is called difference-in-discontinuities
- Sometimes, RDDs are problematic because several treatments happen at the cutoff—how do we identify the treatment effect of interest?
  - A typical example are population-based cutoffs: many government policies may change at the same cutoff
- Idea of difference-in-discontinuities: exploit over-time variation in the presence of different rules
- Comparing the discontinuities in outcomes at the cutoff at two points of time gives the difference-in-discontinuities estimator

Difference-in-Discontinuities: An Empirical Example

Table 1: Rules of the Domestic Stability Pact (DSP)

Year	Target of the DSP rules	Covered municipalities
1997	None	All
1998	None	All
1999	Fiscal gap: zero growth	All
2000	Fiscal gap: zero growth	All
2001	Fiscal gap: max 3% growth	Above 5,000
2002	Fiscal gap: max 2.5% growth	Above 5,000
2003	Fiscal gap: zero growth	Above 5,000
2004	Fiscal gap: zero growth	Above 5,000

Notes. The *Domestic Stability Pact* is a set of fiscal rules imposed by the central government to discipline the fiscal management of local governments. The main target is the *Fiscal gap* (see Appendix Table A1 for details). The growth of the fiscal gap with respect to its value two years before is constrained to be either zero or below 2.5%/3% depending on the year of the DSP. Legislative sources: annual national budget law (*Legge Finanziaria*) from 1999 to 2004.

Table 2: Legislative thresholds for Italian municipalities, 1997–2004

Population	Wage of mayor	Wage of executive committee	Size of executive committee	Size of city council	Electoral rule
Below 1,000	1,291	15%	4	12	single
1,000-3,000	1,446	20%	4	12	single
3,000-5,000	2,169	20%	4	16	single
5,000-10,000	2,789	50%	4	16	single
10,000-15,000	3,099	55%	6	20	single
15,000-30,000	3,099	55%	6	20	runoff
30,000-50,000	3,460	55%	6	30	runoff
50,000-100,000	4,132	75%	6	30	runoff
100,000-250,000	5,010	75%	10	40	runoff
250,000-500,000	5,784	75%	12	46	runoff
Above 500,000	7,798	75%	14-16	50-60	runoff

Notes. Policies varying at different legislative thresholds in the period 1999–2004. *Population* is the number of resident inhabitants as measured by the last available Census. *Wage of mayor* and *Wage of executive committee* refer to the monthly gross wage of the mayor and the members of the executive committee, respectively; the latter is expressed as a percentage of the former, which refers to 2000 and is measured in Euros. *Size of executive committee* is the maximum allowed number of executives appointed by the mayor. *Size of city council* is the number of seats in the city council. The wage thresholds at 1,000 and 10,000 were introduced in 2000; all of the other thresholds date back to 1960. Since 1993, the *Electoral rule* for the mayor is plurality with either single round or runoff.

- Grembi et al. (2016) consider the impact of abolishing fiscal rules on government spending policies in Italy

Estimation: Compare Differences at the 5,000 Cutoff Before and After 2001

- An RDD estimation using the 5,000 inhabitants cutoff would be confounded by the change in politicians remuneration (which could also mean higher incentives to perform better/attract better politicians in power)

- But the DSP rules are only present for municipalities with less than 5,000 inhabitants before the year 2001!

- We can quantify the discontinuities before and after 2001 and compare them with each other:

$$y_{it} = \delta_0 + \delta_1 P_{it} + S_i(\gamma_0 + \gamma_1 P_i t) + T_t[(\alpha_0 + \alpha_1 P_{it}) + S_i(\beta_0 + \beta_1 P_{it})] + \xi_{it}$$

  where P_{it} is the (normalized) population in municipality i at time t, S_i is an indicator for the municipality having less than 5,000 inhabitants, and T_t is an indicator for the year being 2001 or later

Difference-in-Discontinuities Estimator

- $\hat{\beta}_0$ is the difference-in-discontinuities estimator
- It identifies the causal effect of <u>not having</u> fiscal sustainability rules if three conditions hold:
  1. All covariates are continuous at the 5,000 inhabitants threshold before and after the treatment
  2. The effect of the confounding policy (wage change) is constant over time
  3. The effect of the treatment does not depend on the confounding policy
- It is better to limit our attention within some small window around the 5,000 inhabitants cutoff just like in RDD
- Note: Optimal bandwidth selection methods that we saw in the RDD lectures could be applied after some modifications

Main Results

Table 4: Effect of relaxing fiscal rules, diff-in-disc estimates

Panel A: Fiscal Discipline and Expenditures

	Deficit	**Fiscal Gap**	**Current Outlays**	**Capital Outlays**	**Debt Service**
Calonico et al. (2014)	17.495**	59.468*	-47.698	102.557	-2.607
	(7.737)	(32.079)	(59.522)	(101.152)	(8.057)
h	600	513	443	427	404
Obs.	2,414	2,136	1,828	1,724	1,646
Cross Validation	9.454**	48.469**	-10.665	-4.221	-2.096
	(4.343)	(23.315)	(32.756)	(83.336)	(3.587)
h	1,498	833	979	944	1,202
Obs.	5,858	3,438	4,112	3,974	4,908
Mean	13.393	190.757	489.515	475.815	29.651

Panel B: Revenues and Tax Instruments

	Taxes	**Fees& tariffs**	**Central Transfers**	**Other Revenues**	**Real estate tax rate**	**Income tax surcharge**
Calonico et al. (2014)	-76.083**	-2.879	35.001	-21.900	-0.050*	-0.070*
	(32.597)	(10.140)	(27.634)	(120.248)	(0.026)	(0.039)
h	378	505	564	399	435	441
Obs.	1,536	2,104	2,286	1,622	1,782	1,310
Cross Validation	-34.748*	1.413	32.938	-81.308	-0.027*	-0.044*
	(20.166)	(7.199)	(21.721)	(62.926)	(0.016)	(0.026)
h	684	795	833	1,498	907	871
Obs.	2,810	3,238	3,438	5,858	3,806	2,594
Mean	184.811	57.836	131.026	531.925	0.581	0.309

Notes. Municipalities between 3,500 and 7,000 inhabitants; budget years between 1999 and 2004. Diff-in-disc estimates of the impact of relaxing fiscal rules on policy outcomes and tax instruments below 5,000 after 2001. Estimation method: Local Linear Regression with two optimal bandwidth h, as in equation (1). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2014a, 2014b), or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007). All policy outcomes are per capita and in 2009 Euros. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

A Quantitative Approach to Case Studies

- The last panel estimator we will see is the **synthetic control method**
- Developed for the case where:
  - Treatment happens at an aggregate level **for one cluster**
  - Data are available at the aggregate level for lots of periods
  - There are many other clusters that are untreated
- The method has since been extended to multiple treated units but we will stay with the one-cluster case in this course

A Quantitative Approach to Case Studies

- To motivate the idea further, let us consider the first-ever study using a synthetic control approach
- Abadie and Gardeazabal (2003) are interested in the economic costs of the conflict/terrorism in the Basque Country from 1970s
- Difficult to see how to evaluate this
  - Just looking at changes over time in the Basque country confounds secular changes over time
  - Spain had a general downturn in late-70s to mid-80s
  - Basque Country differs from other regions in Spain on factors that matter for growth
- We do not expect parallel trends to hold if we compared the Basque Country with other parts of Spain—who should we be comparing them to, then!?

Abadie and Gardeazabal (2003): The Basic Idea

- Use the other 16 regions in Spain to construct a "synthetic" Basque country using pre-treatment (conflict) data
  - Generate weights for the other regions which sum to one
  - The synthetic Basque is a weighted average of the other regions
  - In the pre-treatment period "synthetic" Basque should look like "true" Basque on covariates
- Then compare the trends for the synthetic Basque Country to the actual Basque Country over time
- The difference between the real Basque Country and the the synthetic Basque Country in any given year is the treatment effect of terrorism in that year

Main Result Graphically

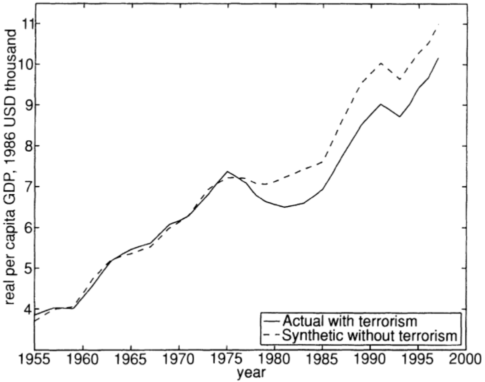


FIGURE 1. PER CAPITA GDP FOR THE BASQUE COUNTRY

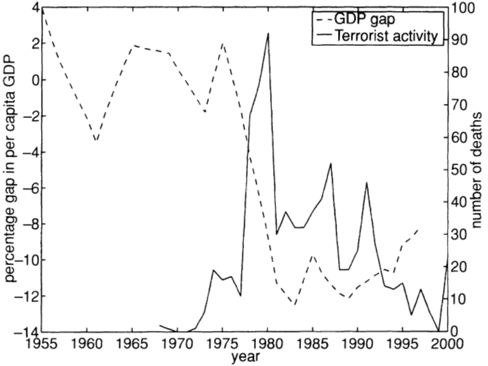


FIGURE 2. TERRORIST ACTIVITY AND ESTIMATED GAP

Advantages

- Precludes extrapolation
- Does not require access to post-treatment outcomes in the "design" phase of the study, when synthetic controls are calculated
- Makes explicit the contribution of each comparison unit to the counterfactual of interest
- Allows researchers to use quantitative and qualitative techniques to analyze the similarities and differences between the units representing the case of interest and the synthetic control
- Formalizing the way comparison units are chosen not only represents a way of systemizing comparative case studies, it also has direct implications for inference

Set Up

- Suppose that we observe $J+1$ units in periods 1,2,..., T
- Unit "one" is exposed to the intervention of interest (that is, "treated") during periods T_0+1,..., T
- The remaining J are an untreated reservoir of potential controls (a "donor pool")
- Let Y_{it}^N be the outcome that would be observed for unit i at time t in the **absence of the intervention**
- Let Y_{it}^I be the outcome that would be observed for unit i at time t if unit i is **exposed to the intervention** in periods T_0+1 to T
- We aim to estimate the effect of the intervention on the treated unit $(\alpha_{1T_0+1}, ..., \alpha_{1T})$ where $\alpha_{1t} = Y_{1t}^I - Y_{1t}^N = Y_{1t} - Y_{1t}^N$ for $t > T_0$ and Y_{1t} is the outcome for unit one at time t

How to Implement This in Practice?

- Let $W = (w_2, ..., w_{J+1})'$ with $w_j \geq 0$ for $j = 2, ..., J+1$ and $w_2 + \cdots + w_{j+1} = 1$
- Each value of W represents a potential synthetic control
- Let X_1 be a $(k \times 1)$ vector of pre-intervention characteristics for the treated unit
- Similarly, let X_0 be a $(k \times J)$ matrix which contains the same variables for the unaffected units
- The vector $W^* = (w_2^*, ..., w_{J+1}^*)'$ is chosen to minimize $\|X_1 - X_0 W\|$, subject to our weight constraints
- Let Y_{jt} be the value of the outcome for unit j at time t; then, for a post-intervention period t (with $t \geq T_0$), the synthetic control estimator is:

$$\hat{\alpha_{1t}} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}$$

How to Implement This in Practice?

- Abadie et al. consider $\|X_1 - X_0 W\| = \sqrt{(X_1 - X_0 W)' V (X_1 - X_0 W)}$
- Let X_{jm} be the value of the m-th covariate for unit j
- Typically, V is a $(k \times k)$ diagonal matrix with main diagonal $v_1, ..., v_k$
- Then, the synthetic control weights $W^* = (w_2^*, ..., w_{J+1}^*)'$ minimize:

$$\sum_{m=1}^{k} v_m (X_{1m} - \sum_{j=2}^{J+1} w_j X_{jm})^2$$

where v_m is a weight that reflects the relative importance that we assign to the m-th variable when we measure the discrepancy between the treated unit and the synthetic controls

How to Implement This in Practice?

- The choice of V is important
- W^* depends on the choice of V
- The synthetic control $W^*(V)$ is meant to reproduce the behavior of the outcome variable for the treated unit in the absence of the treatment
- Therefore, the weights $v_1, \ldots, v_k$ should reflect the predictive value of the covariates

How to Implement This in Practice?

- Choice of $v_1, ..., v_k$ can be based on different things...
  - Subjective assessment of the predictive power of each of the covariates, or calibration inspecting how different values for $v_1, ..., v_k$ affect the discrepancies between the treated unit and the synthetic control
  - Use regression to assess the predictive power of the covariates; minimize mean square prediction error (MSPE):

$$\sum_{t=1}^{T_0} \left(Y_{1t} - \sum w_{j=2\ j}^{J+1\ *}(V) Y_{jt} \right)^2$$

  - **Cross-validation:** Divide the pre-treatment period into an initial **training period** and a subsequent **validation period**
  - For any given V, calculate $W^*(V)$ in the training period; minimize the MSPE of $W^*(V)$ in the validation period

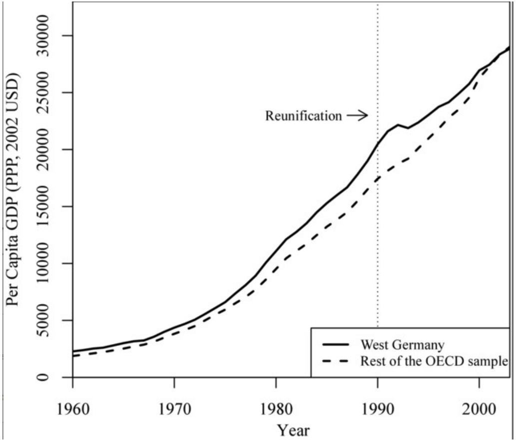
What about unobserved factors?

- Comparative case studies are complicated by unmeasured factors affecting the outcome variables as well as heterogeneity in the effect of observed and unobserved factors
- However, f the number of pre-intervention periods in the data is large, matching on pre-intervention outcomes allows us to control for heterogeneous responses to multiple unobserved factors
- Intuition: only units that are alike in observed and unobserved determinants of the outcome variable as well as in the effect of those determinants on the outcome variable should produce similar trajectories of the outcome variable over extended periods of time

Time for More Cool Applications! The Unification of West and East Germany (Abadie, Diamond and Hainmuller 2014)

- Cross-country regressions are often criticized because they put side-by-side countries of very different characteristics
- "What do Thailand, the Dominican Republic, Zimbabwe, Greece and Bolivia have in common that merits their being put in the same regression analysis? Answer: For most purposes, nothing at all." (Harberger 1987)
- The synthetic control method provides a data-driven procedure to select a comparison unit
- Application: the economic impact of the 1990 German reunification in West Germany
- Donor pool is restricted to 16 OECD countries

The Unification of West and East Germany



The Unification of West and East Germany

Country Weights in the Synthetic West Germany

Country	Weight	Country	Weight
Australia	0	Netherlands	0.10
Austria	0.42	New Zealand	0
Belgium	0	Norway	0
Denmark	0	Portugal	0
France	0	Spain	0
Greece	0	Switzerland	0.11
Italy	0	United Kingdom	0
Japan	0.16	United States	0.22

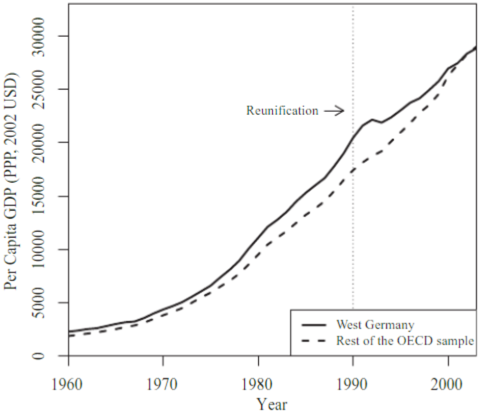
The Unification of West and East Germany

TABLE 2 Economic Growth Predictor Means before German Reunification

	West Germany	Synthetic West Germany	OECD Sample
GDP per capita	15808.9	15802.2	8021.1
Trade openness	56.8	56.9	31.9
Inflation rate	2.6	3.5	7.4
Industry share	34.5	34.4	34.2
Schooling	55.5	55.2	44.1
Investment rate	27.0	27.0	25.9

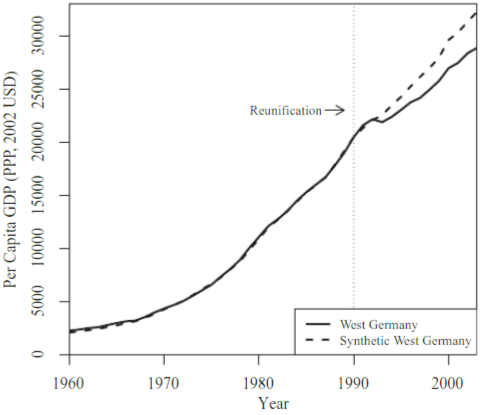
The Unification of West and East Germany

FIGURE 1 Trends in per Capita GDP: West Germany versus Rest of the OECD Sample



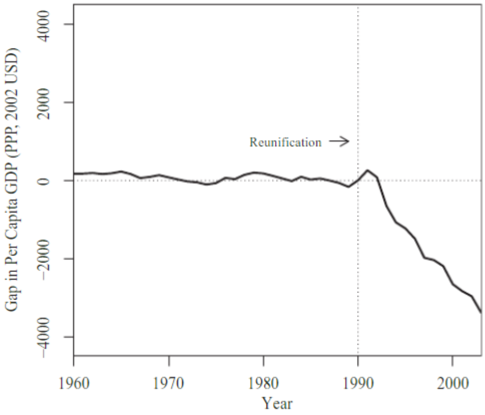
The Unification of West and East Germany

FIGURE 2 Trends in per Capita GDP: West Germany versus Synthetic West Germany



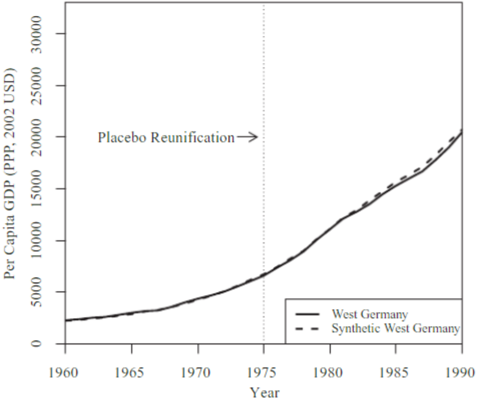
The Unification of West and East Germany

FIGURE 3 **Per Capita GDP Gap between West Germany and Synthetic West Germany**



The Unification of West and East Germany

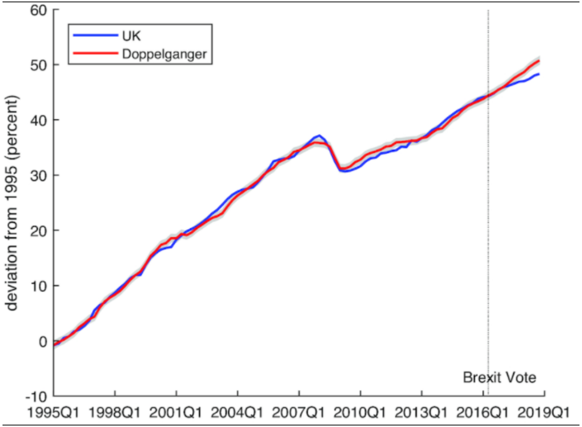
FIGURE 4 Placebo Reunification 1975–Trends in per Capita GDP: West Germany versus Synthetic West Germany



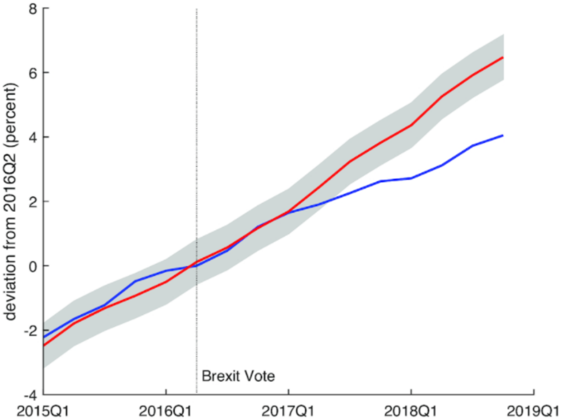
Another Application: The Brexit Referendum (Born et al. 2019)

- Born et al (2019) study the effects of the Brexit referendum on UK real GDP
- Huge political interest in the question, very difficult to study
- Data: quarterly data for 23 countries from 1995Q1 to 2016Q2
- Create a doppelganger UK and compare actual post-referendum UK to it!

The Brexit Referendum



The Brexit Referendum

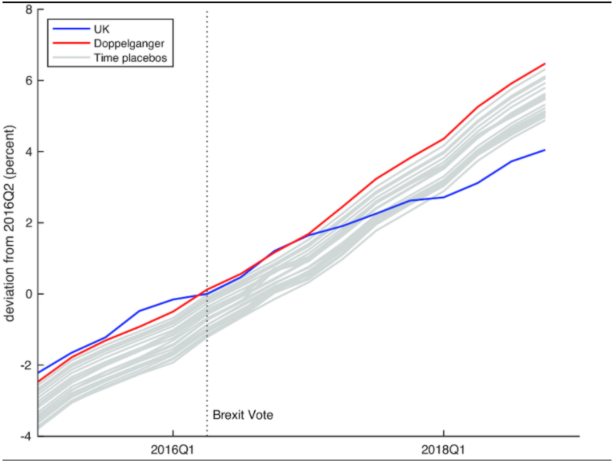


The Brexit Referendum

Table 2. Composition of the Doppelganger: Country Weights.

Australia	<0.01	Austria	<0.01	Belgium	<0.01	Canada	<0.01
Finland	<0.01	France	<0.01	Germany	0.05	Hungary	0.11
Iceland	0.01	Ireland	0.01	Italy	0.17	Japan	<0.01
Korea	<0.01	Luxembourg	<0.01	Netherlands	<0.01	New Zealand	0.14
Norway	<0.01	Portugal	<0.01	Slovak Republic	<0.01	Spain	<0.01
Sweden	<0.01	Switzerland	<0.01	United States	0.51		

The Brexit Referendum



Contextual Requirements (Abadie 2021)

- **Size and volatility of the effect:** The effect needs to be large relative to ordinary noise
- **Availability of a comparison group ("donor pool")**
- **No anticipation:** You can define treatment as being before a particular date when you expect no anticipation
- **No interference/spillovers**
- **Convex hull condition:** A weighted average of the donors can approximate the doppelganger
- **Time horizon:** You have enough periods after to see the full effect
- **Data:** Sufficient pre-treatment data on aggregate units

Recap of the Methods Discussed in This Lecture

1. Introduction
2. Non-panel difference-in-differences
3. Difference-in-differences with IV
4. Triple-differenced models
5. Difference-in-discontinuities
6. Synthetic controls
7. References

Readings

Essential

- Cunningham (2019): Causal Inference: The Mixtape–the chapter on synthetic controls

Recommended

- Abadie, A. (2021). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects, Journal of Economic Literature

Papers Mentioned in This Lecture

- Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. American Economic Review, 91(4), 795-813.
- Muralidharan, K., & Prakash, N. (2017). Cycling to school: increasing secondary school enrollment for girls in India. American Economic Journal: Applied Economics, 9(3), 321-350.
- Kuosmanen, I., & Meriläinen, J. 2022. Labor Market Effects of Open Borders: Evidence from the Finnish Construction Sector after EU Enlargement. Forthcoming in Journal of Human Resources.
- Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. American Journal of Political Science, 59(2), 495-510
- Born, B., Muller, G. J., Schularick, M., & Sedláček, P. (2019). The costs of economic nationalism: evidence from the Brexit experiment. The Economic Journal, 129(623), 2722-2744.