Principles of Empirical Analysis

Lecture 9a: Difference-in-differences

Spring 2022
Tuukka Saarimaa

Outline

- **Basic idea of difference-in-difference (DID, DD, diff-in-diff) designs**
  - DID with two groups and two time periods
  - More general case with many time periods
- **Applications**
  - Card & Krueger (1994): classic paper on minimum wage
  - Saarimaa & Tukiainen (2015): common pool
  - Other examples

DID

- **We have talked about the idea of using group differences to estimate causal effects**
  - We would like to find treatment and control groups who can be assumed to be similar in every way except receipt of treatment
  - Without randomization this is very difficult/implausible
- **A weaker assumption is that in the absence of treatment, the difference between treatment and control groups is constant over time (parallel or common trends)**
- **With this assumption (and some others) we can use observations in treatment and control groups before and after the treatment to estimate a causal effect**

DID

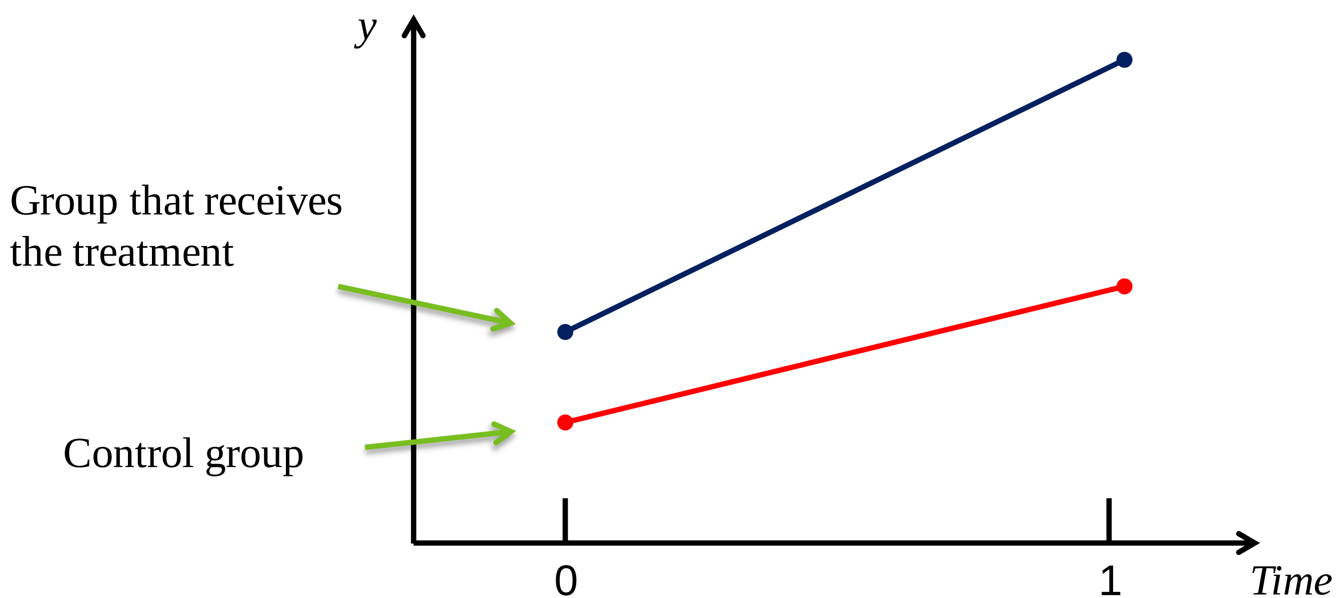
- **Idea:**
  - Pre-treatment difference is ‘normal’ difference
  - Post-treatment difference is ‘normal’ difference + causal effect of treatment
  - Difference-in-differences is the causal effect
- **DID relies heavily on common or parallel time trends so visual inspection of the data is a very important part of any DID analysis**

Difference-in-differences – two groups, two time periods

DID: two groups, two time periods

- **The canonical DID design contains two time periods and two groups where the timing of treatment is the same for all treated**
  - Most current DID applications, however, use data from more than two time periods and often the treatment occurs at different times
  - There is currently a lively discussion going on regarding what to do when dealing with these more complicated designs!
- **The 2x2 design is still an excellent pedagogical point of departure**

Difference-in-differences graphically

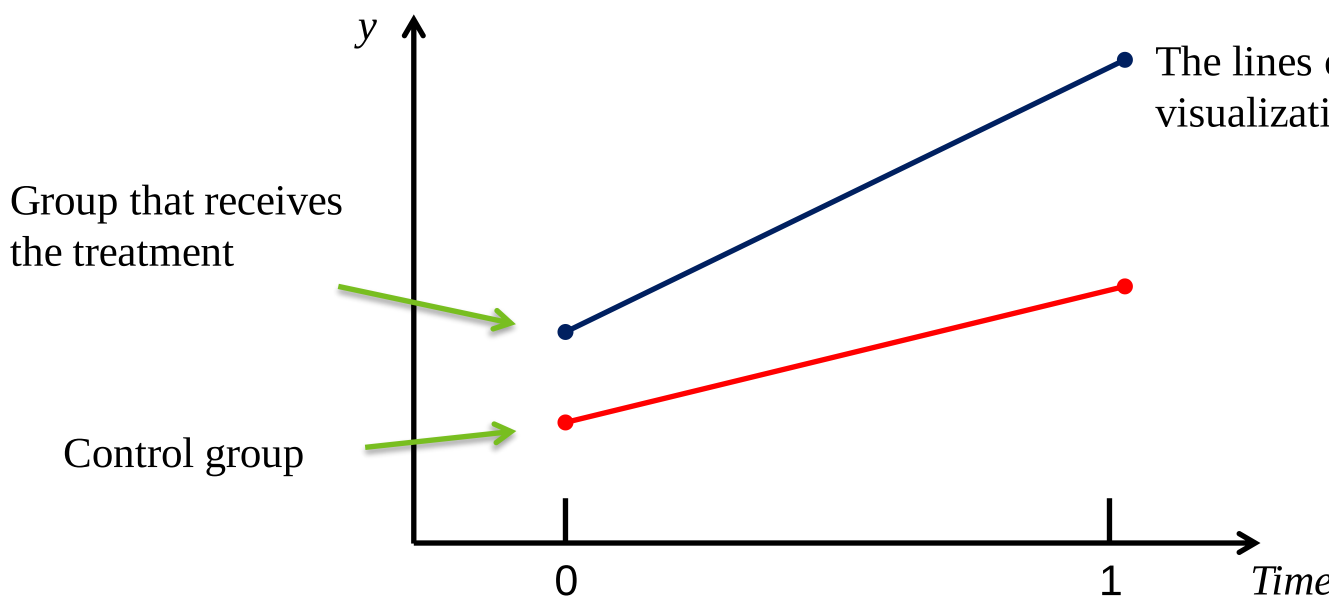


Difference-in-differences graphically

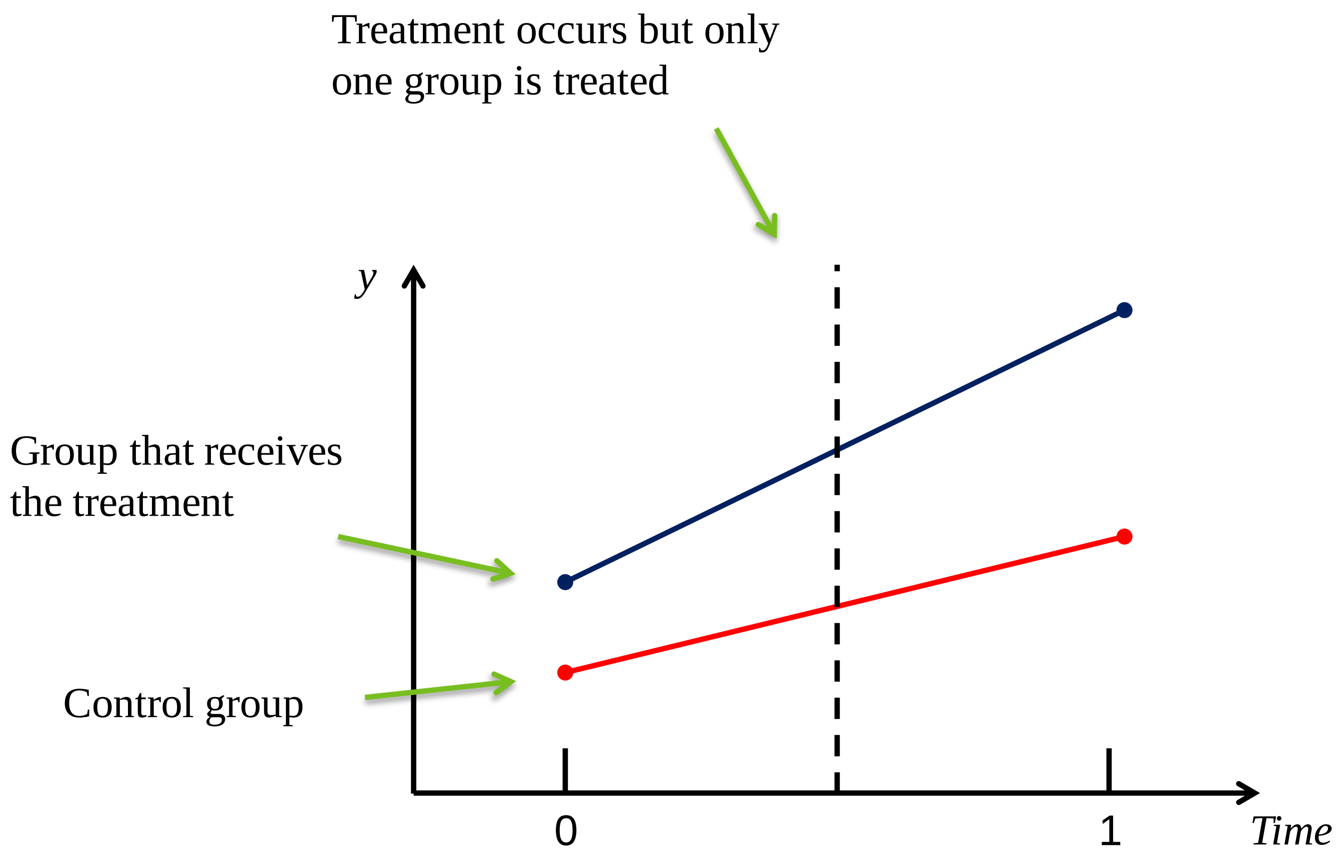
We are measuring the outcome of interest (y) in two time periods

The dots are means of the outcome for each group in each time period

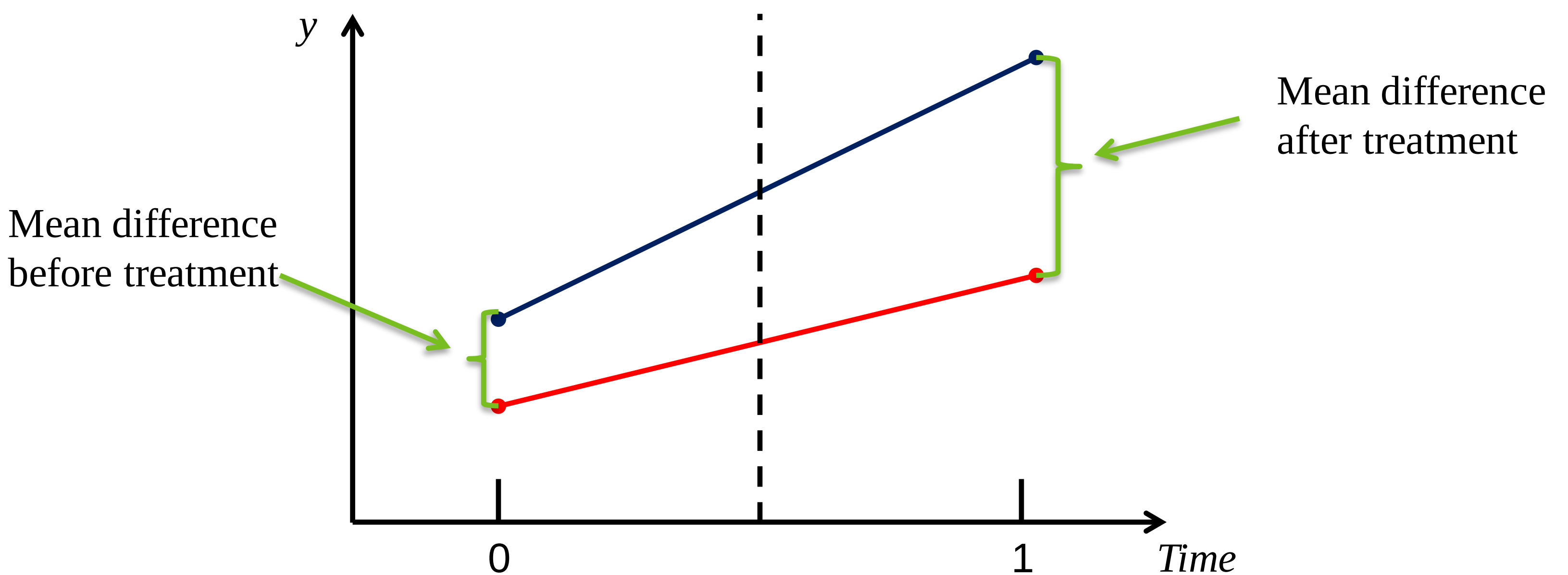
The lines connecting the dots are just for visualization purposes



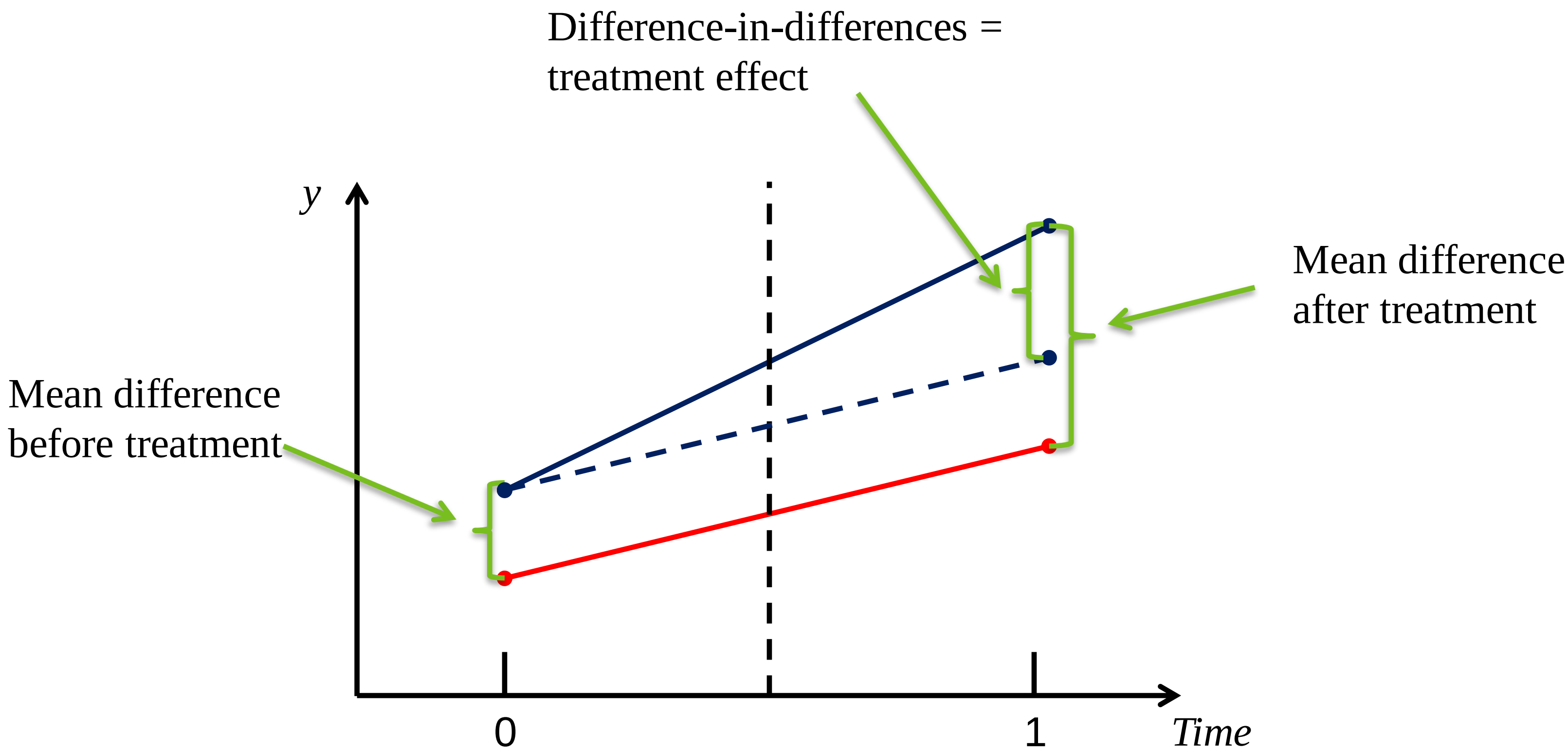
Difference-in-differences graphically



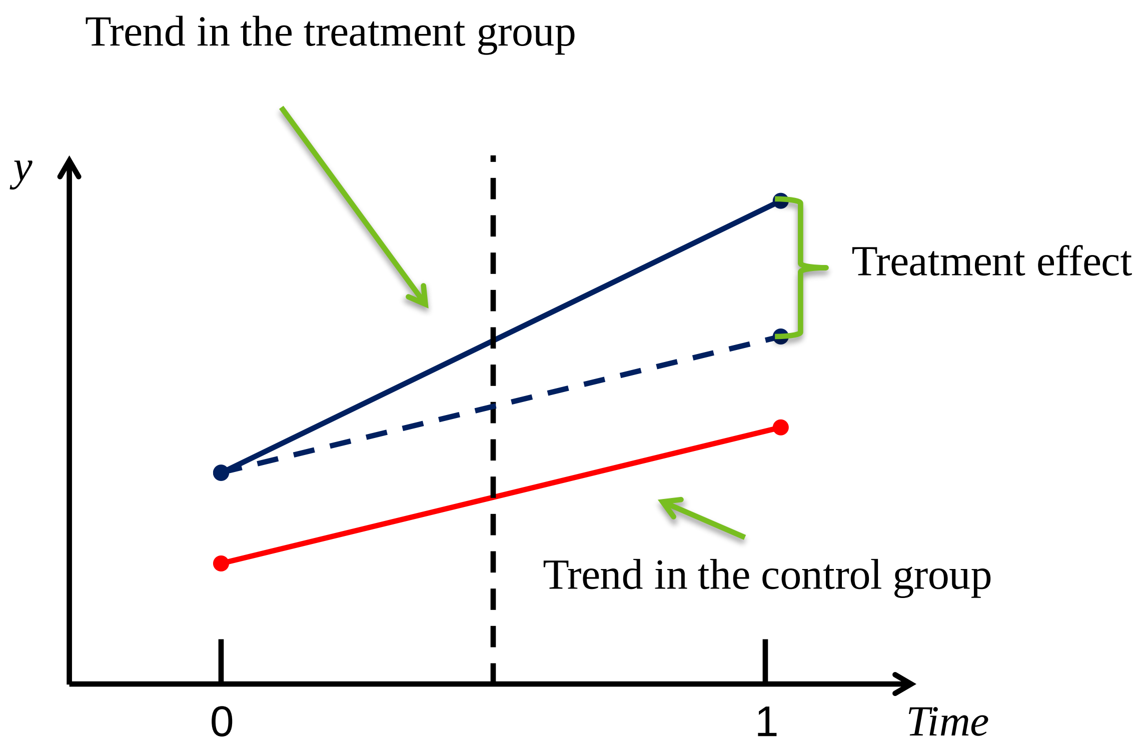
Difference-in-differences graphically



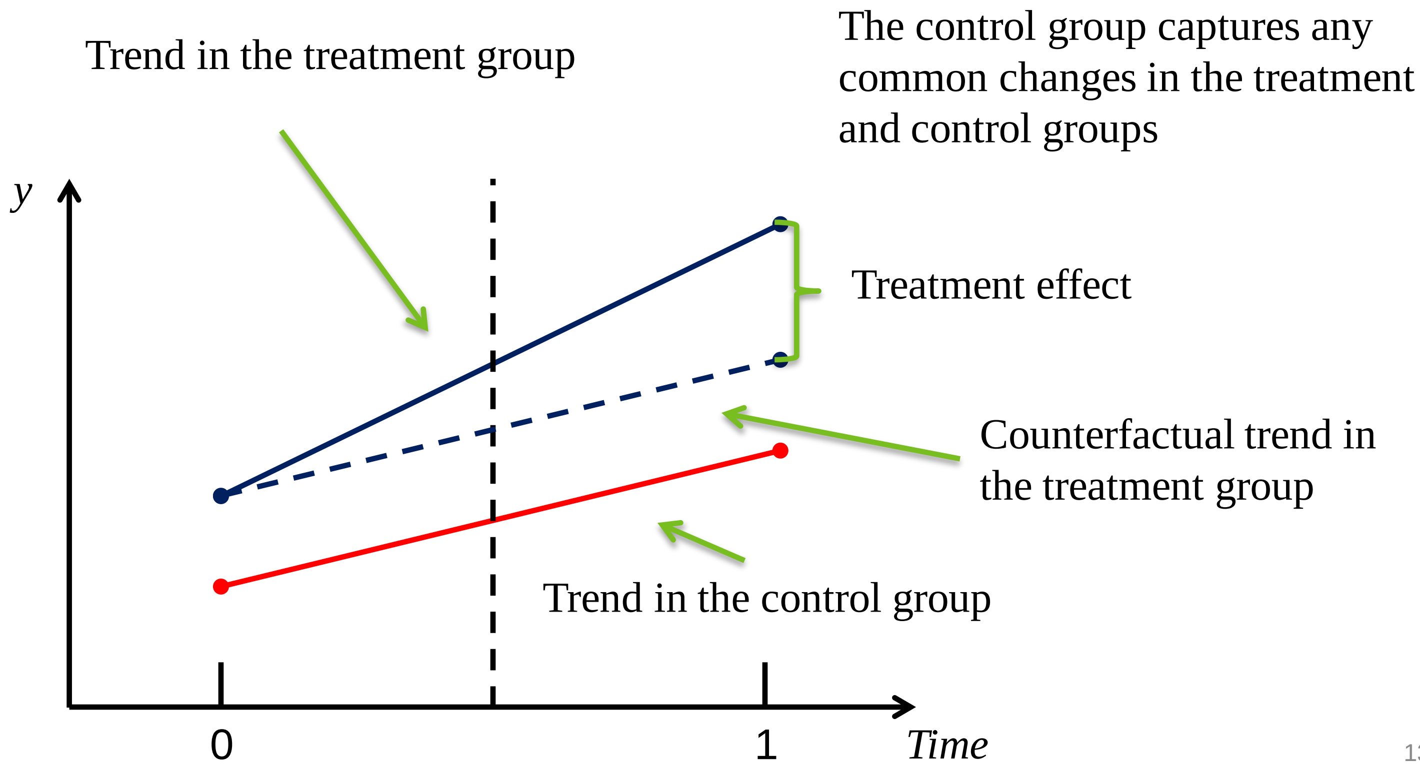
Difference-in-differences graphically



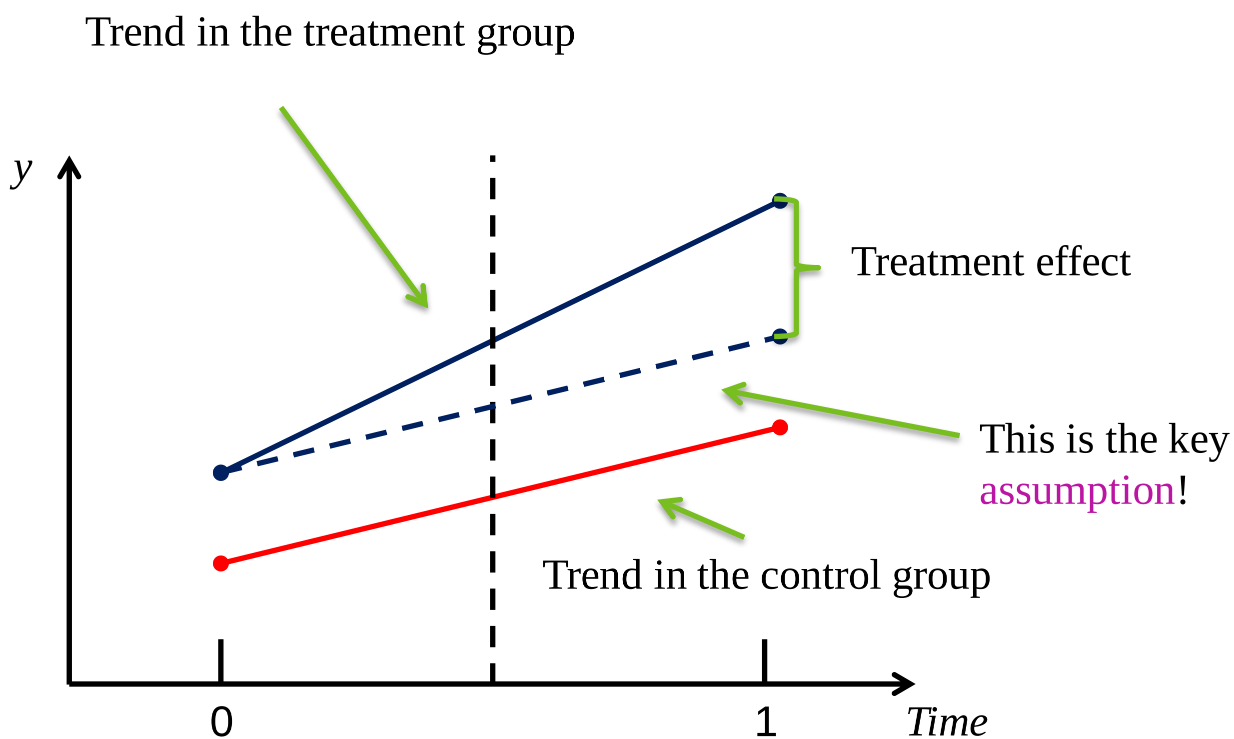
Difference-in-differences graphically



Difference-in-differences graphically



Difference-in-differences graphically

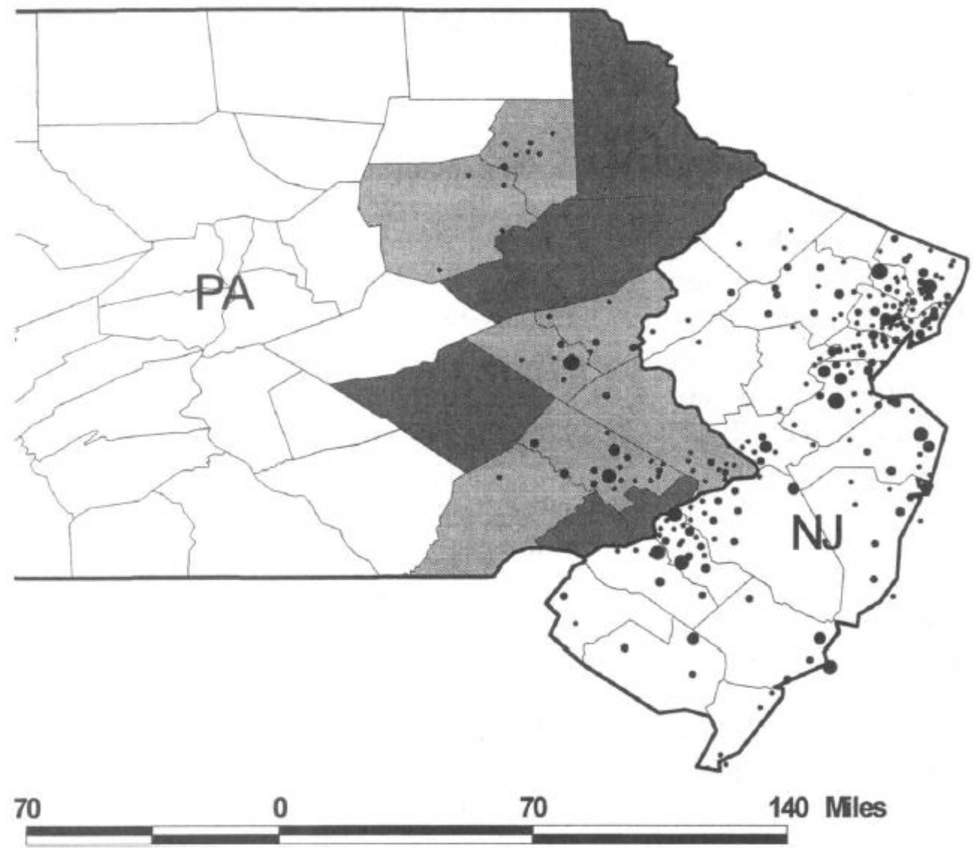


Example: New Jersey minimum wage increase

On April 1, 1992, NJ increased the state minimum wage from $4.25 to $5.05; PA's minimum wage stayed at $4.25

Card & Krueger (1994) surveyed about 400 fast food stores both in NJ and in PA before (February) and after (November) the minimum wage increase

Any common macroeconomic trends captured by using the control group



Card & Krueger (1994)

- **DID more formally**

  y_{ist}: employment at restaurant i, state s, time t

- **In DID, we need the following means**

  $E[y_{ist}|s = NJ, t = Feb]$

  $E[y_{ist}|s = NJ, t = Nov]$

  $E[y_{ist}|s = PA, t = Feb]$

  $E[y_{ist}|s = PA, t = Nov]$

Card & Krueger (1994)

- **In Jersey:**

$E[y_{ist}|s = NJ, t = Feb] = \text{mean employment in February}$

$E[y_{ist}|s = NJ, t = Nov] = \text{mean employment in November}$

$E[y_{ist}|s = NJ, t = Nov] - E[y_{ist}|s = NJ, t = Feb]$

$= \text{difference in employment}$

- **In Pennsylvania:**

$E[y_{ist}|s = PA, t = Feb] = \text{mean employment in February}$

$E[y_{ist}|s = PA, t = Nov] = \text{mean employment in November}$

$E[y_{ist}|s = PA, t = Nov] - E[y_{ist}|s = PA, t = Feb]$

$= \text{difference in employment}$

Card & Krueger (1994)

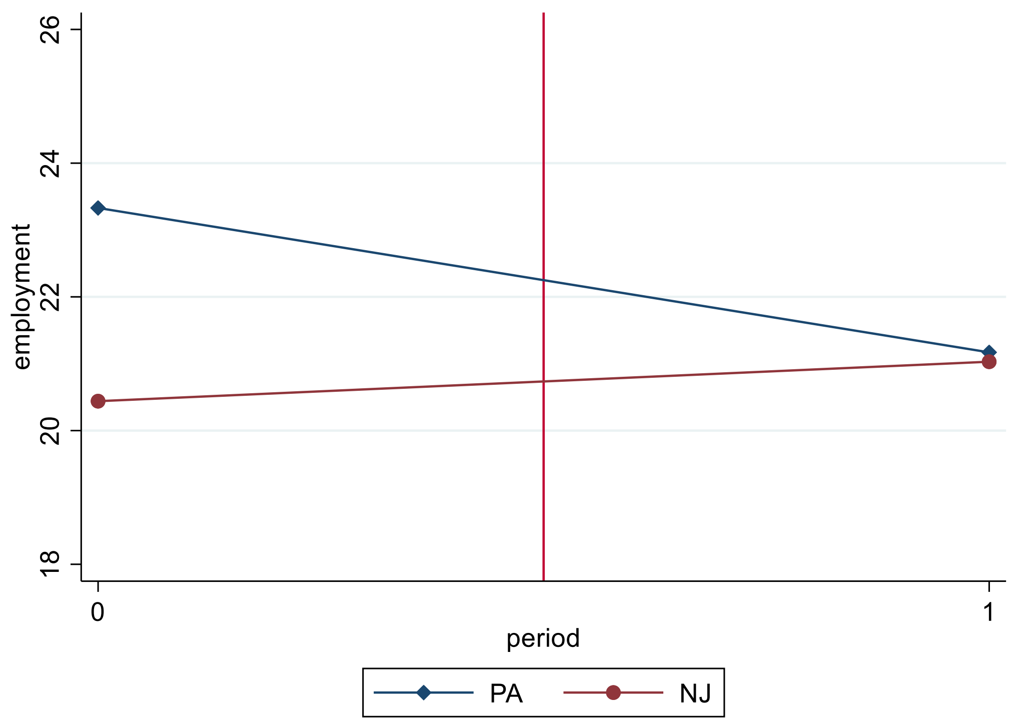
- **The population DID is the treatment effect we are after**

$$\delta = (E[y_{ist}|s = NJ, t = Nov] - E[y_{ist}|s = NJ, t = Feb]) - (E[y_{ist}|s = PA, t = Nov] - E[y_{ist}|s = PA, t = Feb])$$

- **The DID estimator is the sample analog:**

$$\hat{\delta} = \left(\overline{y}_{NJ,Nov} - \overline{y}_{NJ,Feb}\right) - \left(\overline{y}_{PA,Nov} - \overline{y}_{PA,Feb}\right)$$

Card & Krueger (1994)



	Stores by state		
Variable	PA (i)	NJ (ii)	Difference, NJ – PA (iii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (0.51)	−2.89 (1.44)
2. FTE employment after, all available observations	21.17 (0.94)	21.03 (0.52)	−0.14 (1.07)
3. Change in mean FTE employment	−2.16 (1.25)	0.59 (0.54)	$\boxed{2.76\ (1.36)}$

$(21.03 - 20.44) - (21.17 - 23.33) = 2.76$

- Surprisingly, if anything employment increased in New Jersey!

DID using regression

In the 2x2 case, the regression model would look like this

$$y_{it} = \alpha + \beta treated_i + \gamma after_t + \delta treated_i \cdot after_t + u_{it}$$

treated = 1 if observation is in the treatment group, 0 otherwise

after = 1 if observation is from the after period, 0 otherwise

*treated*after* = 1 for if observation is in the treatment group observed after the treatment

In econometrics jargon, *treated* and *after* are dummy variables and their product is called an interaction term

α is referred to as the intercept or the constant term

DID using regression

In Card & Krueger minimum wage study this would be

$$y_{ist} = \alpha + \beta NJ_s + \gamma Nov_t + \delta NJ_s \cdot Nov_t + u_{ist}$$

NJ = 1 if observation is in New Jersey the treatment group, 0 otherwise (regardless of the time period)

Nov = 1 if observation is from the after period, 0 otherwise (regardless of the state)

*NJ*Nov* = 1 for if observation is in New Jersey observed after the treatment

DID using regression

In Card & Krueger minimum wage study this would be

$$y_{ist} = \alpha + \beta NJ_s + \gamma Nov_t + \delta NJ_s \cdot Nov_t + u_{ist}$$

- *NJ before:* $E[y_{ist} | NJ = 1, Nov = 0] = \alpha + \beta$
- *NJ after:* $E[y_{ist} | NJ = 1, Nov = 1] = \alpha + \beta + \gamma + \delta$
- *PA before:* $E[y_{ist} | NJ = 0, Nov = 0] = \alpha$
- *PA after:* $E[y_{ist} | NJ = 0, Nov = 1] = \alpha + \gamma$

- Assuming that $E[u_{ist} | NJ, Nov] = 0$

DID using regression

$$DID = (NJ\ after - NJ\ before) - (PA\ after - PA\ before)$$

- $NJ\ after - NJ\ before = (\alpha + \beta + \gamma + \delta) - (\alpha + \beta) = \gamma + \delta$
- $PA\ after - PA\ before = (\alpha + \gamma) - \alpha = \gamma$

So, we have:

- $DID = (NJ\ after - NJ\ before) - (PA\ after - PA\ before) = \delta$

Estimating the regression model using OLS produces the DID estimate and standard errors which is very convenient

General case

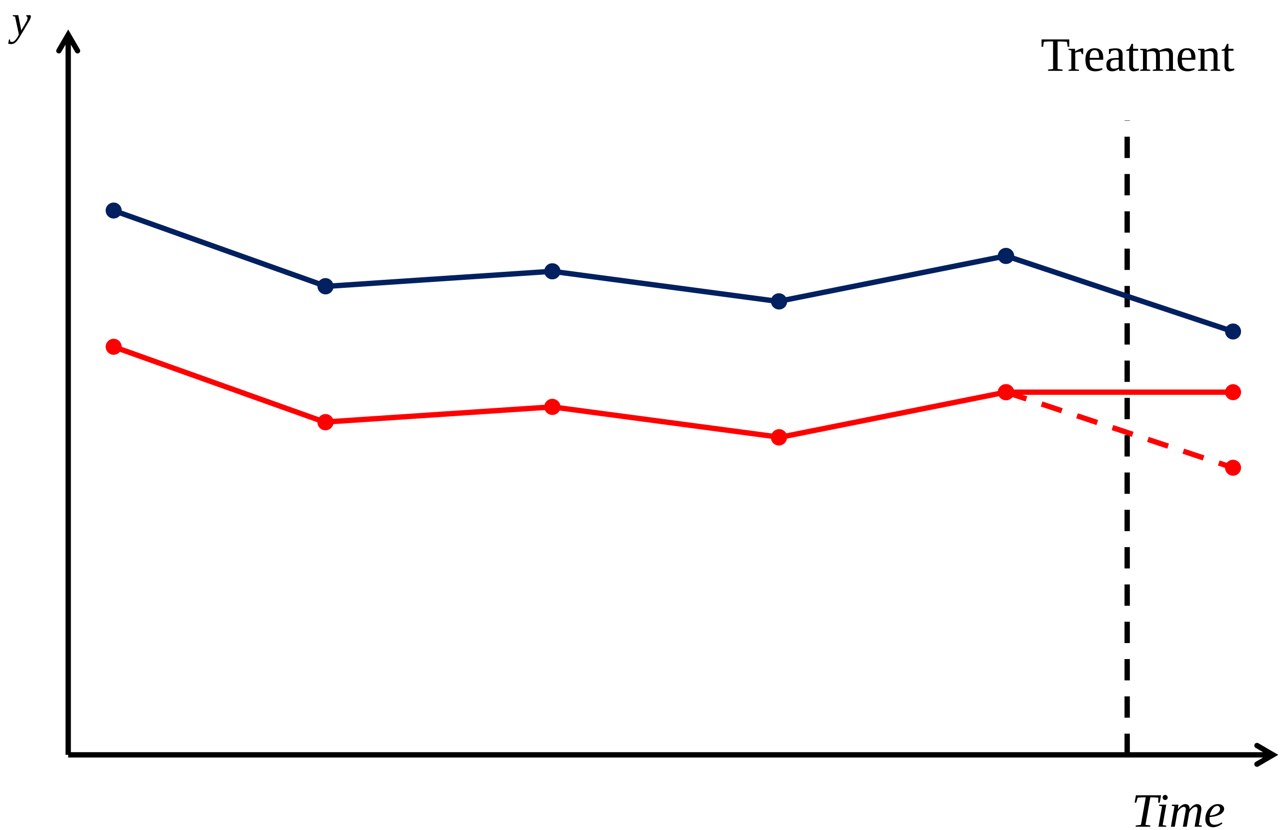
Key assumption I

- **The key assumption for any DID strategy is that the outcome in the treatment and control groups would follow the same time trend in the absence of treatment**
  - This does not mean that they must have the same mean (or level) of the outcome
- **This common or parallel trend assumption is impossible to test because you never observe the counterfactual for the treatment group**
  - But you can test it indirectly using pre-treatment data to show that the trends have been the same in past (only indirect evidence!)

Key assumption I

- **Even if pre-trends are the same one still must worry about other policies or changes coinciding with the treatment**
  - Nothing else happens at the same time as the treatment takes place that would affect the control and treatment groups differently
- **It is very important to be familiar with the institutional details of the reform/policy change that you are analyzing**
  - This applies to all empirical research, but is particularly important in quasi-experimental settings
  - It is not unusual that when presenting this type of work in scientific seminars, most of the discussion concerns institutional details

Common pre-treatment trends



Key assumption II

- **The second key assumption is that there are no spillover effects or that group compositions do not change because of treatment (if using repeated cross-sections)**
  - In the minimum wage example, this would mean that New Jersey's minimum wage increase does not directly affect employment in Pennsylvania
  - How plausible is this?

Applications

ELSEVIER

European Journal of Political Economy
Volume 38, June 2015, Pages 140-152

Common pool problems in voluntary municipal mergers

Tuukka Saarimaa, Janne Tukiainen [1]

⊞ Show more

https://doi.org/10.1016/j.ejpoleco.2015.02.006

Get rights and content

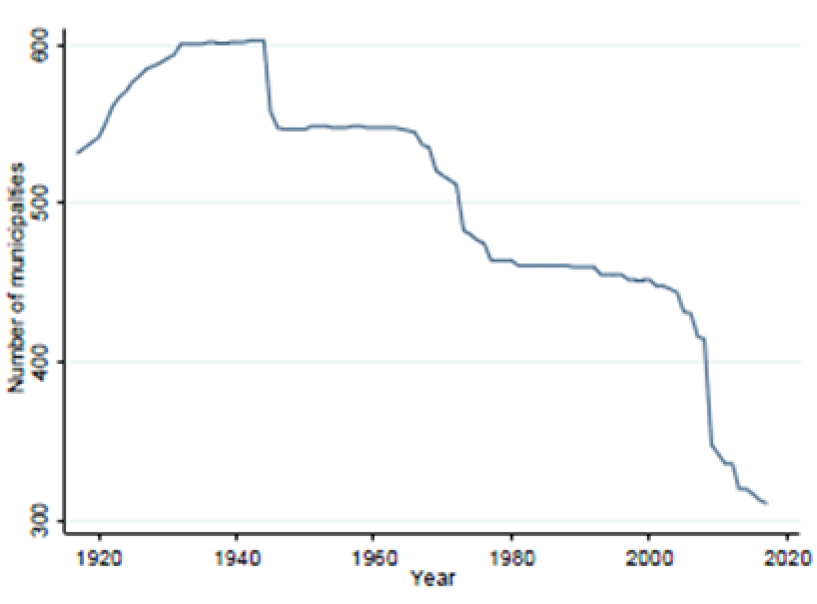
Link to paper: https://www.sciencedirect.com/science/article/abs/pii/S017626801500021X

Saarimaa & Tukiainen (2015)

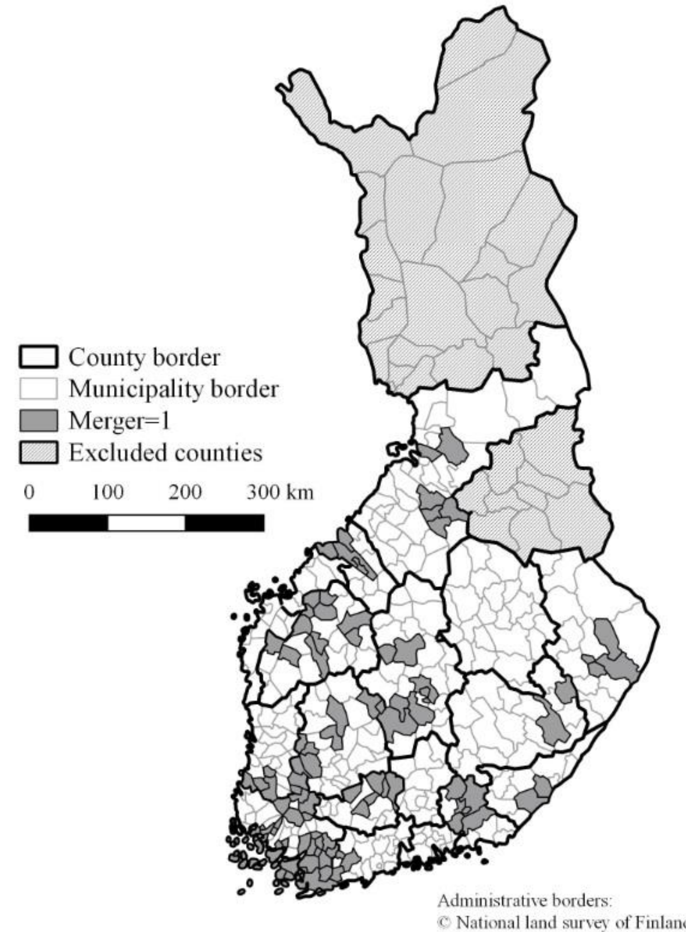
Highlights

- We analyze free-riding behavior of Finnish municipalities prior to municipal mergers.
- A time lag between the initial decision and the actual merger creates a common pool.
- Municipalities exploit the common pool by substantially increasing municipal debt.
- The results are consistent with the "law of $1/n$".

The mergers



Number of municipalities
300
400
500
600
1920
1940
1960
1980
2000
2020
Year



County border
Municipality border
Merger=1
Excluded counties
0 100 200 300 km
Administrative borders:
© National land survey of Finland

Motivation I

- **Very important policy question**
  - This was the first attempt to reform publicly provided social and health services in a major way (so-called PARAS-hanke)
  - Larger municipalities would be better able to produce and fund these services (median municipality population size was 6000)
  - Central government incentivized the mergers with subsidies, but merging was voluntary
  - It’s important to know all the effects of major policy reforms
- **Recently, a new tier of government (counties) was established to provide these services**
  - We just had the county elections

Motivation II

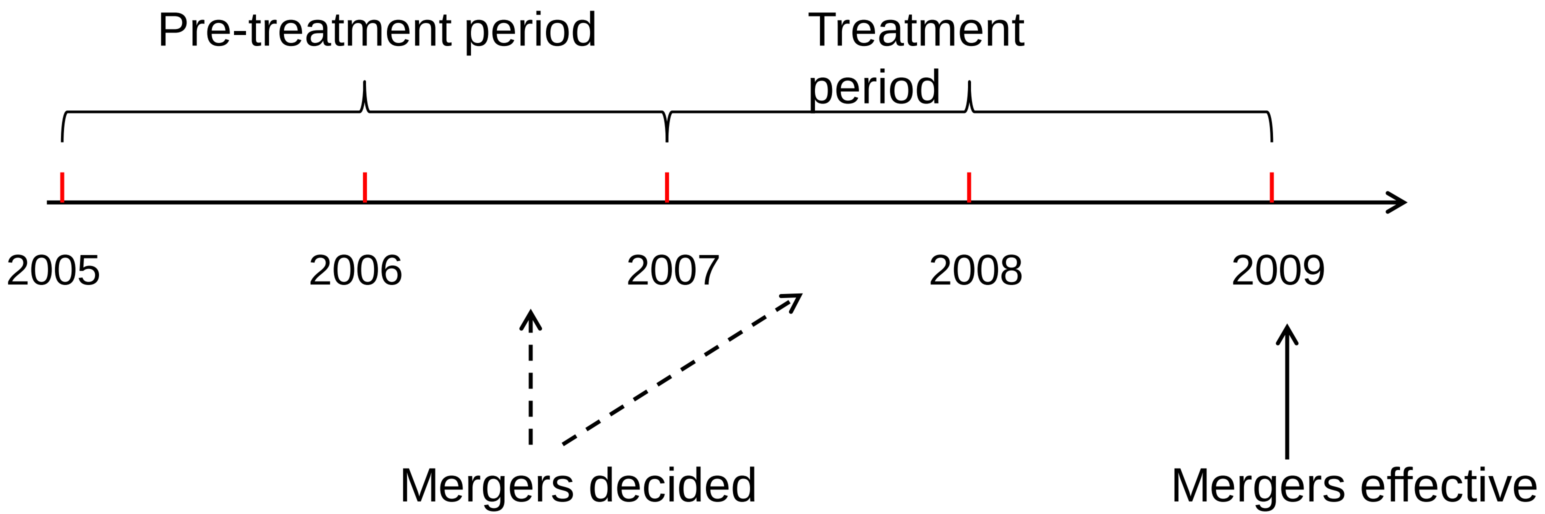
- **A common pool problem arises when the costs of an activity that benefits a small group of people are shared among a larger group**
  - "Law of $1/n$" (Weingast et al. 1981, JPE): Due to common pool funding and universalism, spending increases as the number of decision makers increase => free-riding
  - Think about having dinner with friends and splitting the check

Motivation II

- **A common pool problem arises when the costs of an activity that benefits a small group of people are shared among a larger group**
  - "Law of 1/*n*" (Weingast et al. 1981, JPE): Due to common pool funding and universalism, spending increases as the number of decision makers increase => free-riding
  - Think about having dinner with friends and splitting the check
- **Municipality mergers open an opportunity to study the common pool problem and the law of 1/n**
  - Municipalities are autonomous before the merger takes place
  - Free-riding incentives related to relative size of merging municipalities, not the number of municipalities in the merger
  - Electoral punishment unlikely

The mergers

- **In 2007 a provisional law introduced: merger subsidies and other merger incentives**
- **At the start of 2009, 32 mergers (involving 99 municipalities) took place; decided in 2006–07**



Common pool incentives

- **Who has incentives to free-ride?**
  - Some incentives for all municipalities that merge, but stronger for relatively small municipalities
- **We define a measure of free-riding incentives for municipality *i* in merger *j* as**

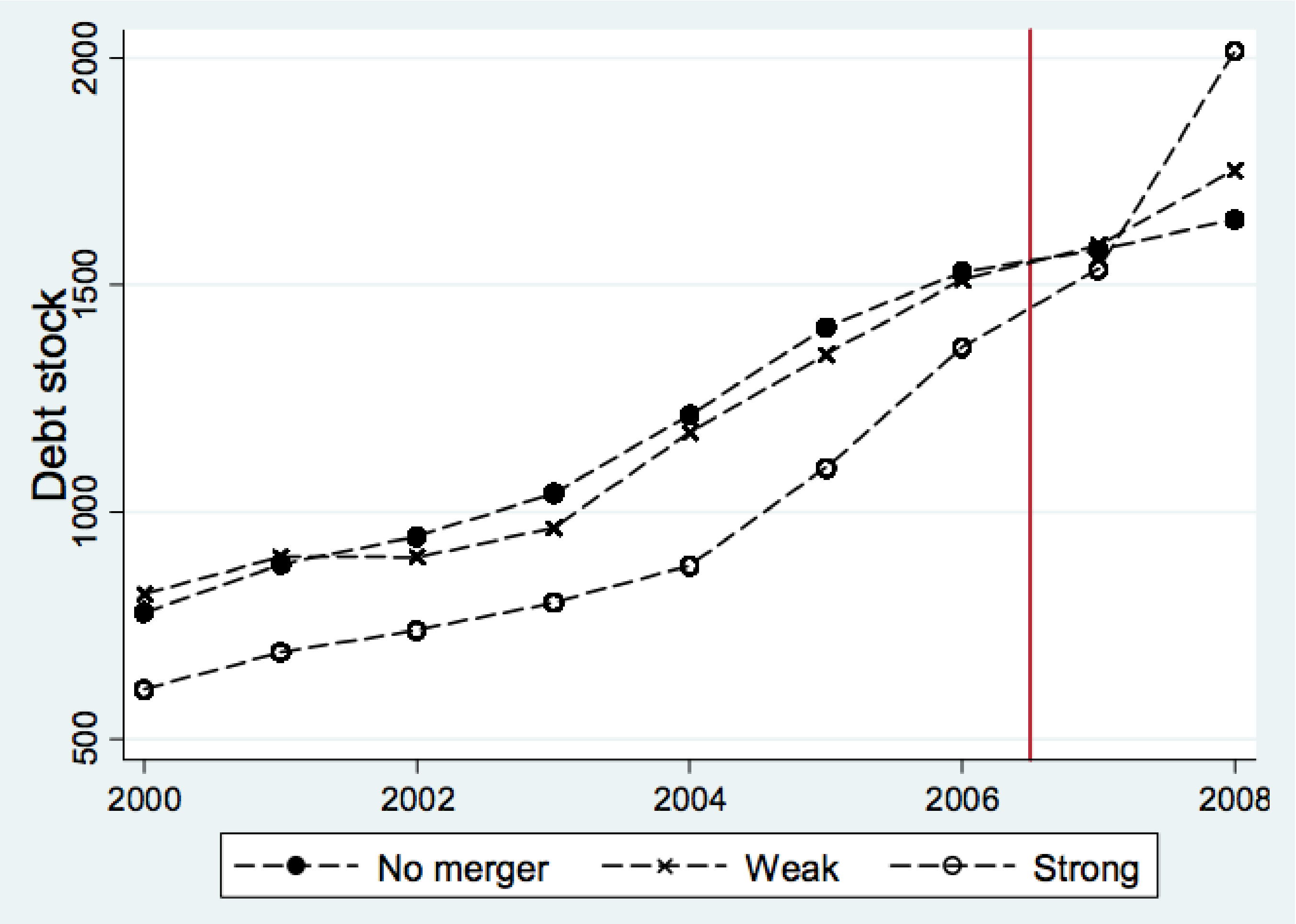
$$freeride_i = 1 - taxbase_i/taxbase_j$$

  - Idea: municipality *i* internalizes $taxbase_i/taxbase_j$ of the social marginal cost of borrowing
  - Higher values of *freeride* imply stronger incentives to freeride

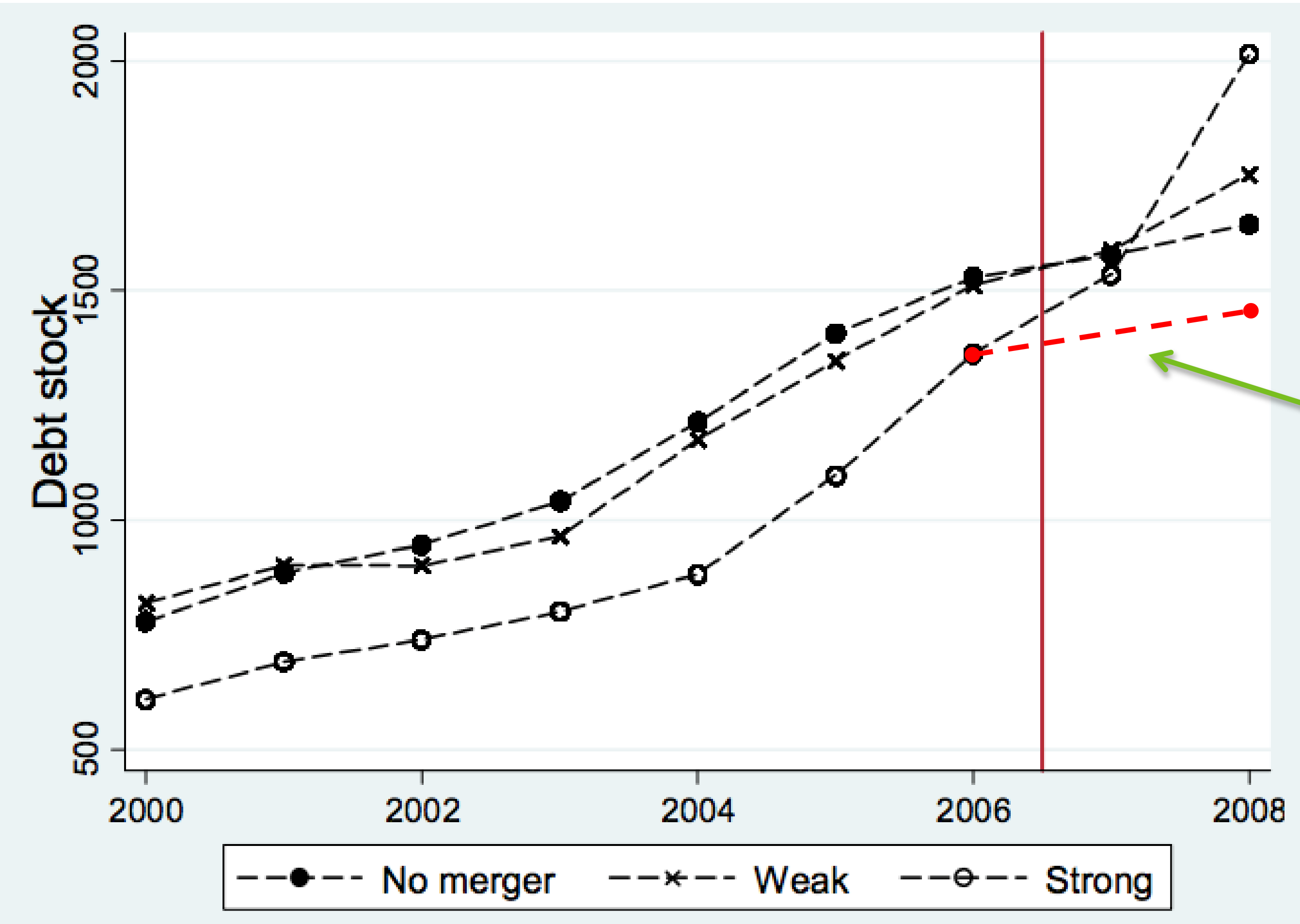
Analysis

- **The incentives measure is continuous, so for the graphical analysis we use groups**
  - Divide the merger group into municipalities with weak and strong incentives to free-ride (according to median value of *freeride*)
  - Compare these groups to the no-merger group and each other
  - So, we have two treatment groups and a control group, but the DID idea is the same

Main results – debt stock per capita

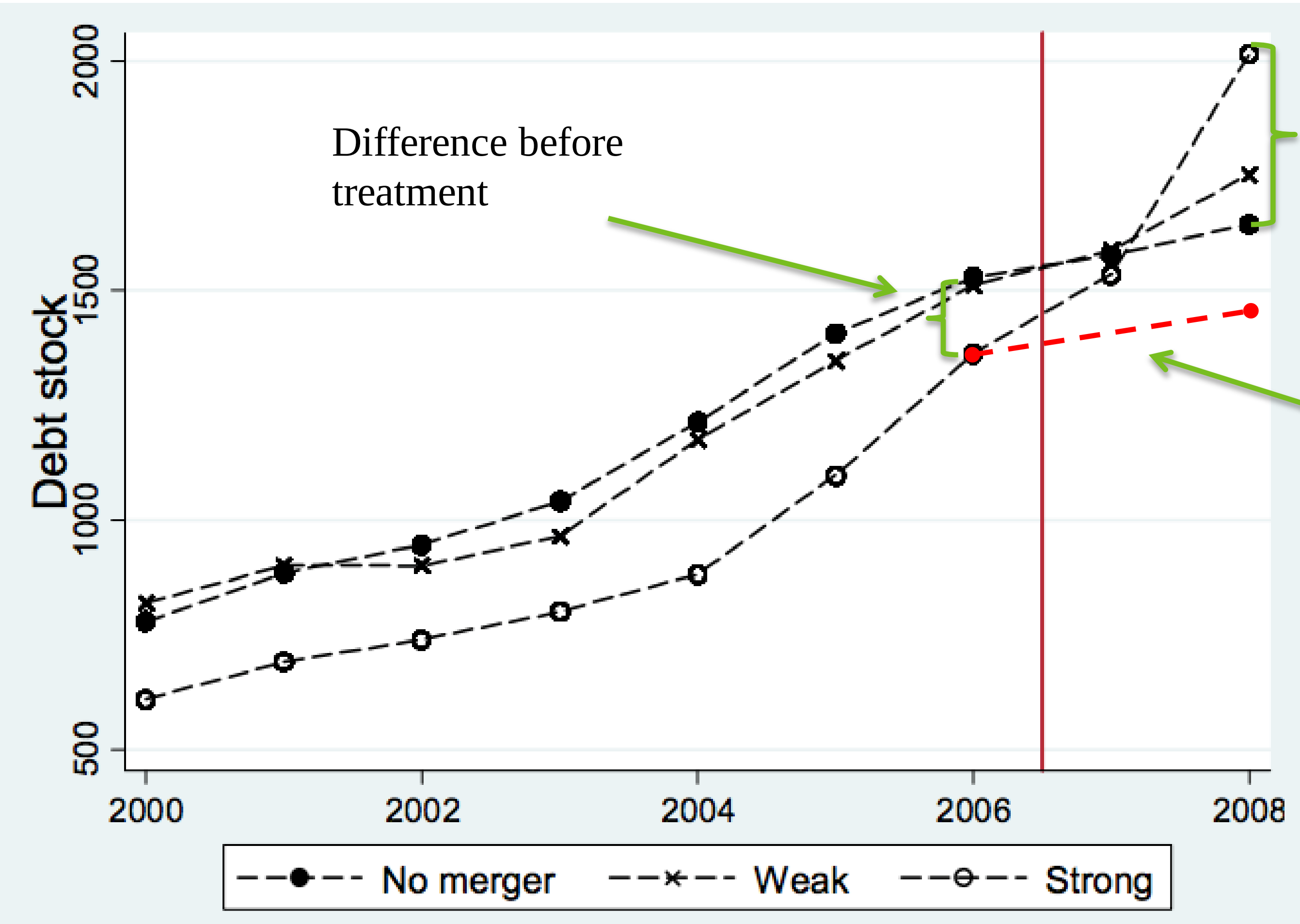


Main results – debt stock per capita



Counterfactual trend in the treatment groups

Main results – debt stock per capita

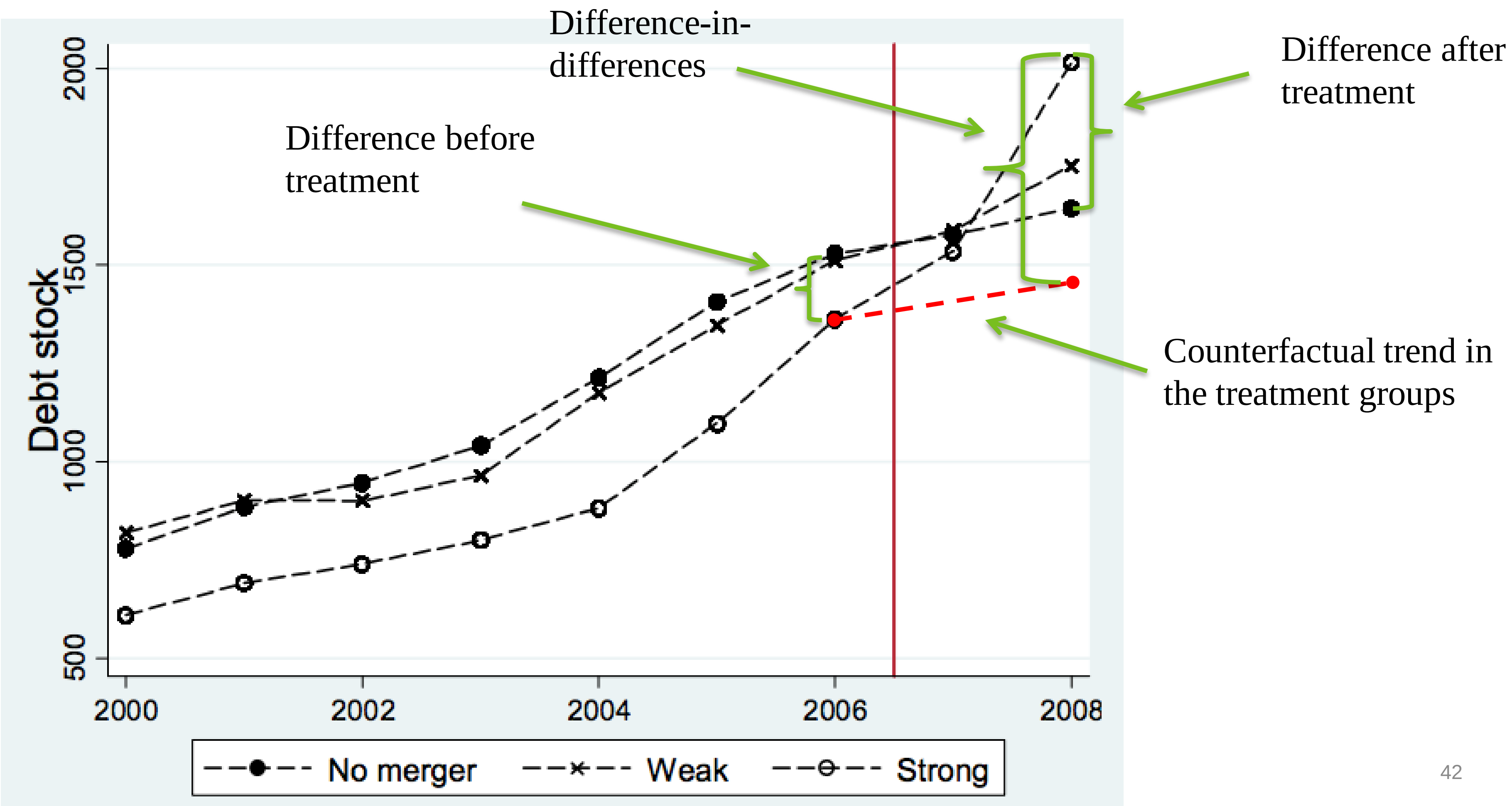


Difference before treatment

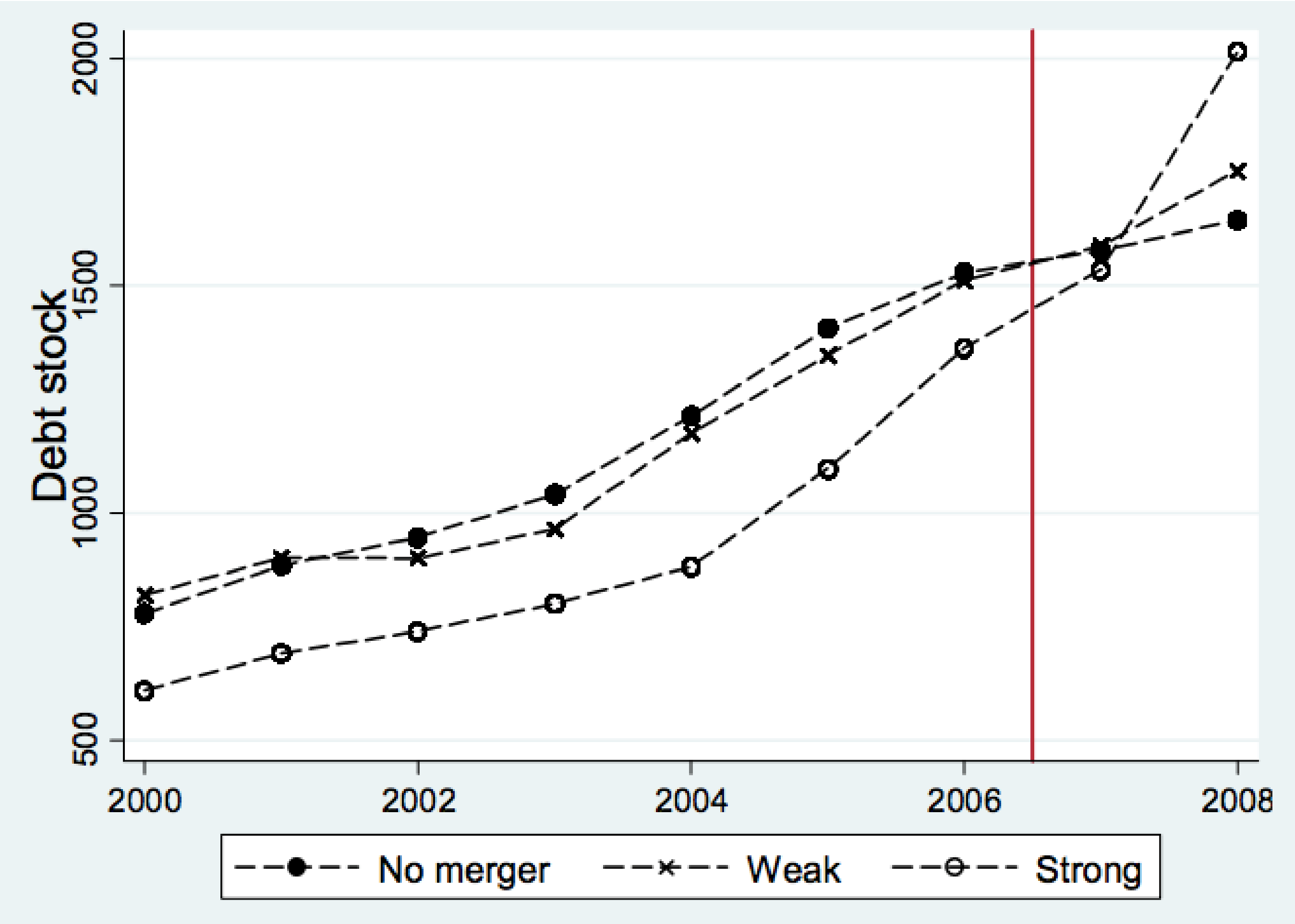
Difference after treatment

Counterfactual trend in the treatment groups

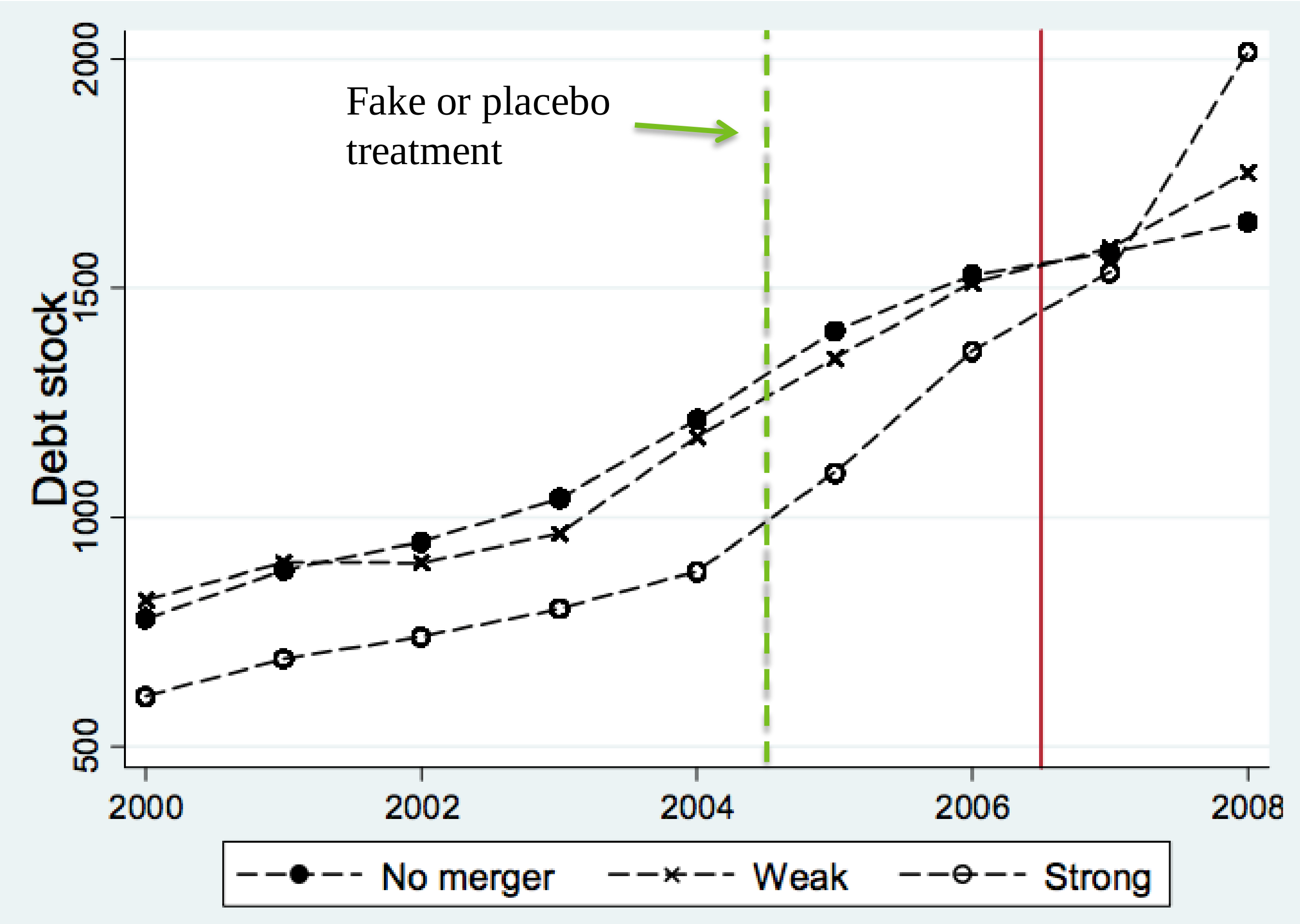
Main results – debt stock per capita



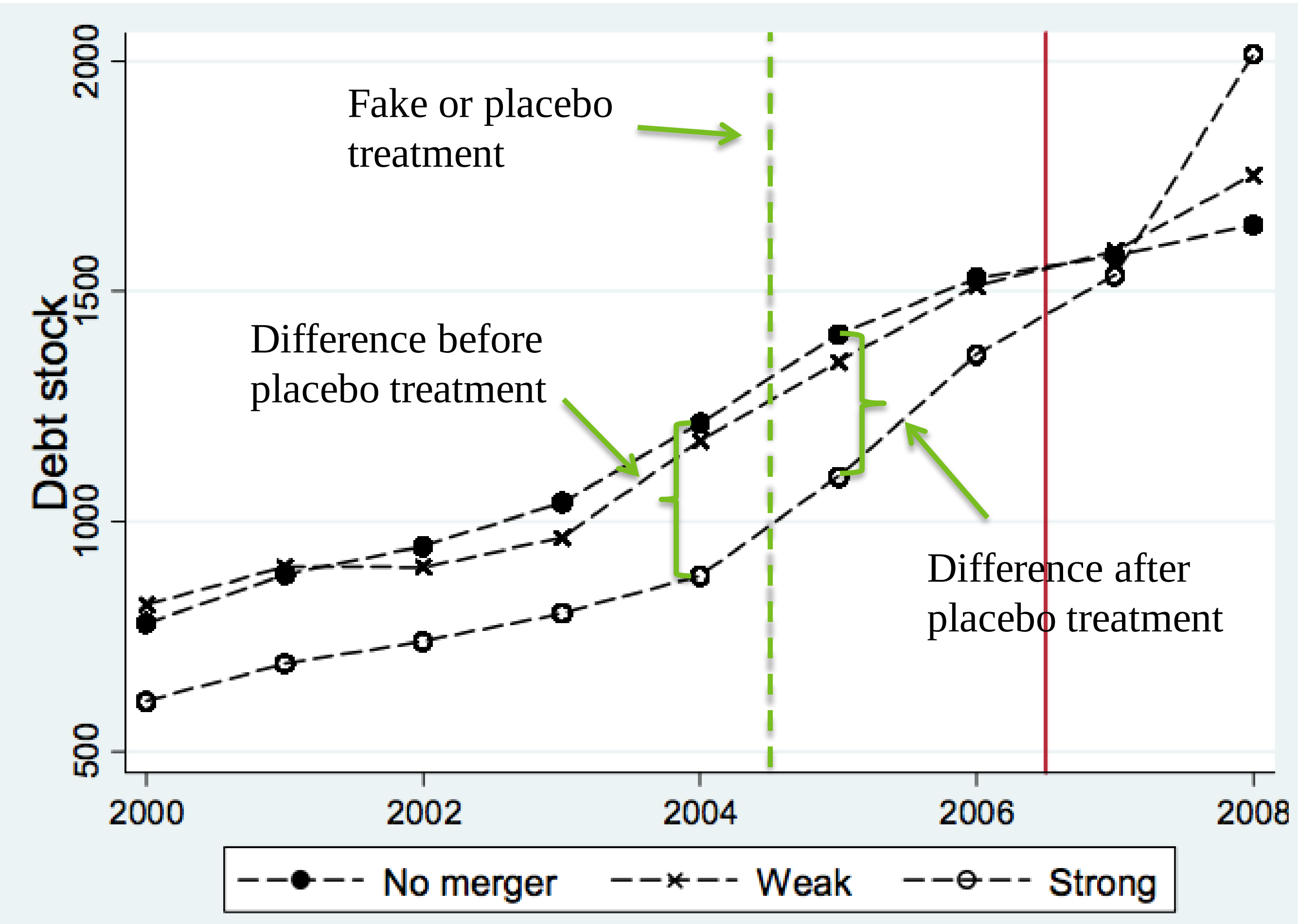
Testing for common trends (indirectly)



Testing for common trends (indirectly)



Testing for common trends (indirectly)

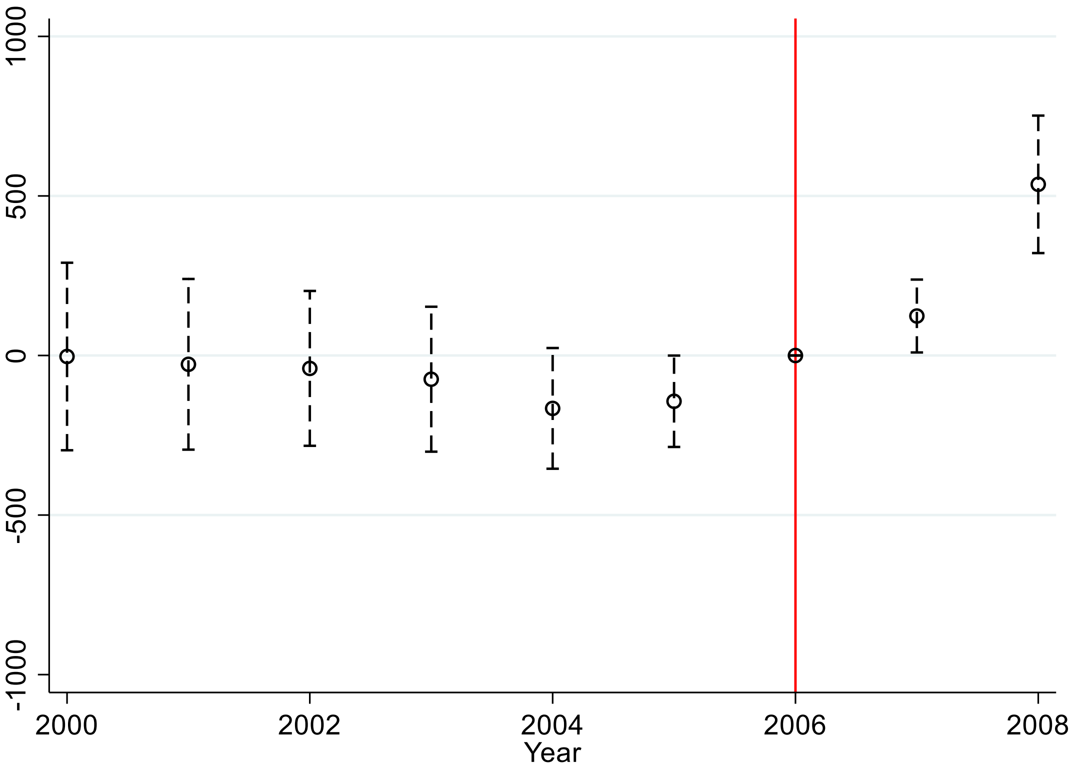


Estimate a DID as if the treatment took place at the placebo cutoff

Repeat with all the pre-treatment years

If you have a lot of pre-treatment data, placebo tests are a convincing way to argue that the common trends assumption is valid

Event study graph



Often DID results are presented like this

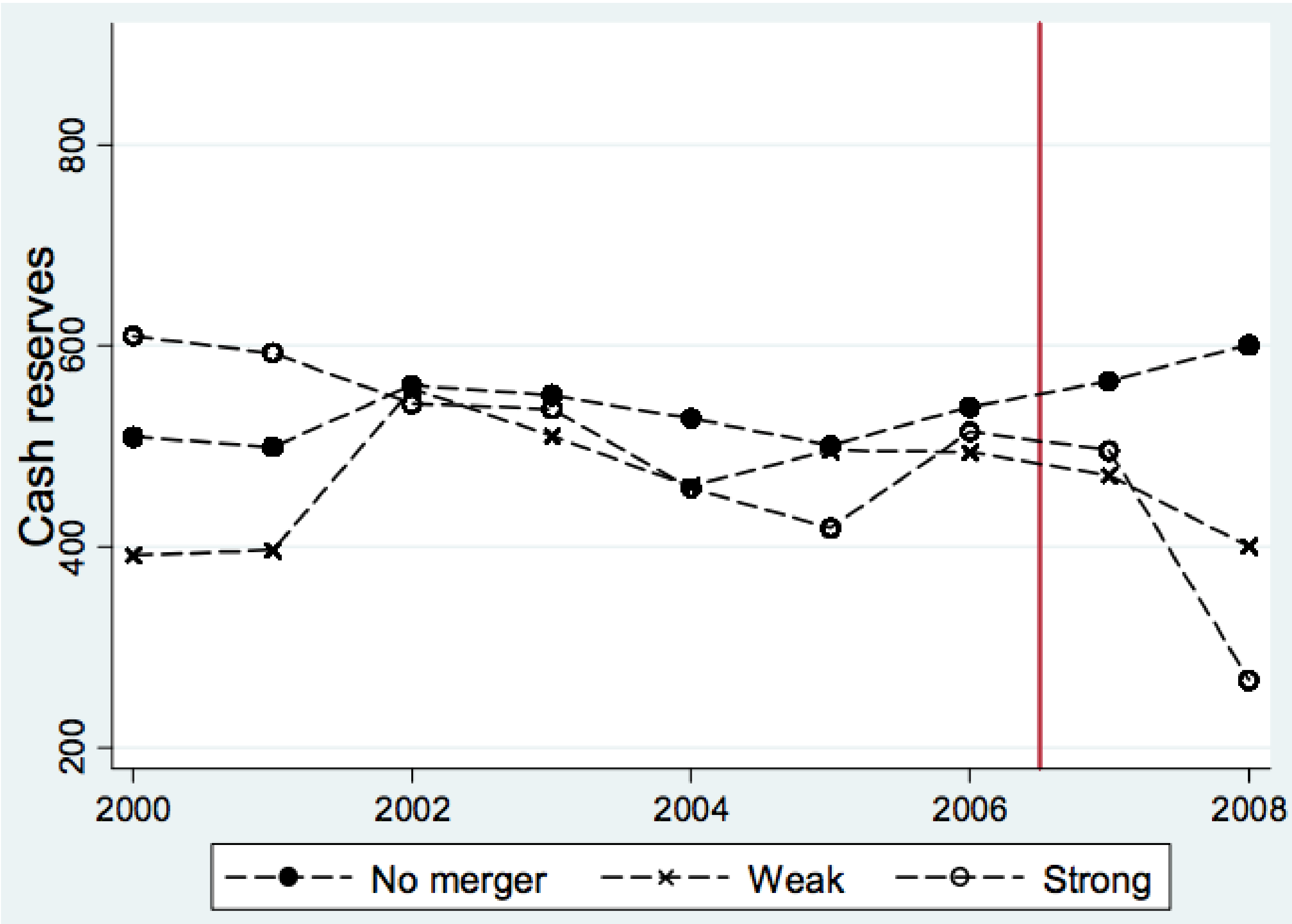
Here we are illustrating how the mean difference between the control group and the strong incentive group behaves through time

The comparison year is 2006

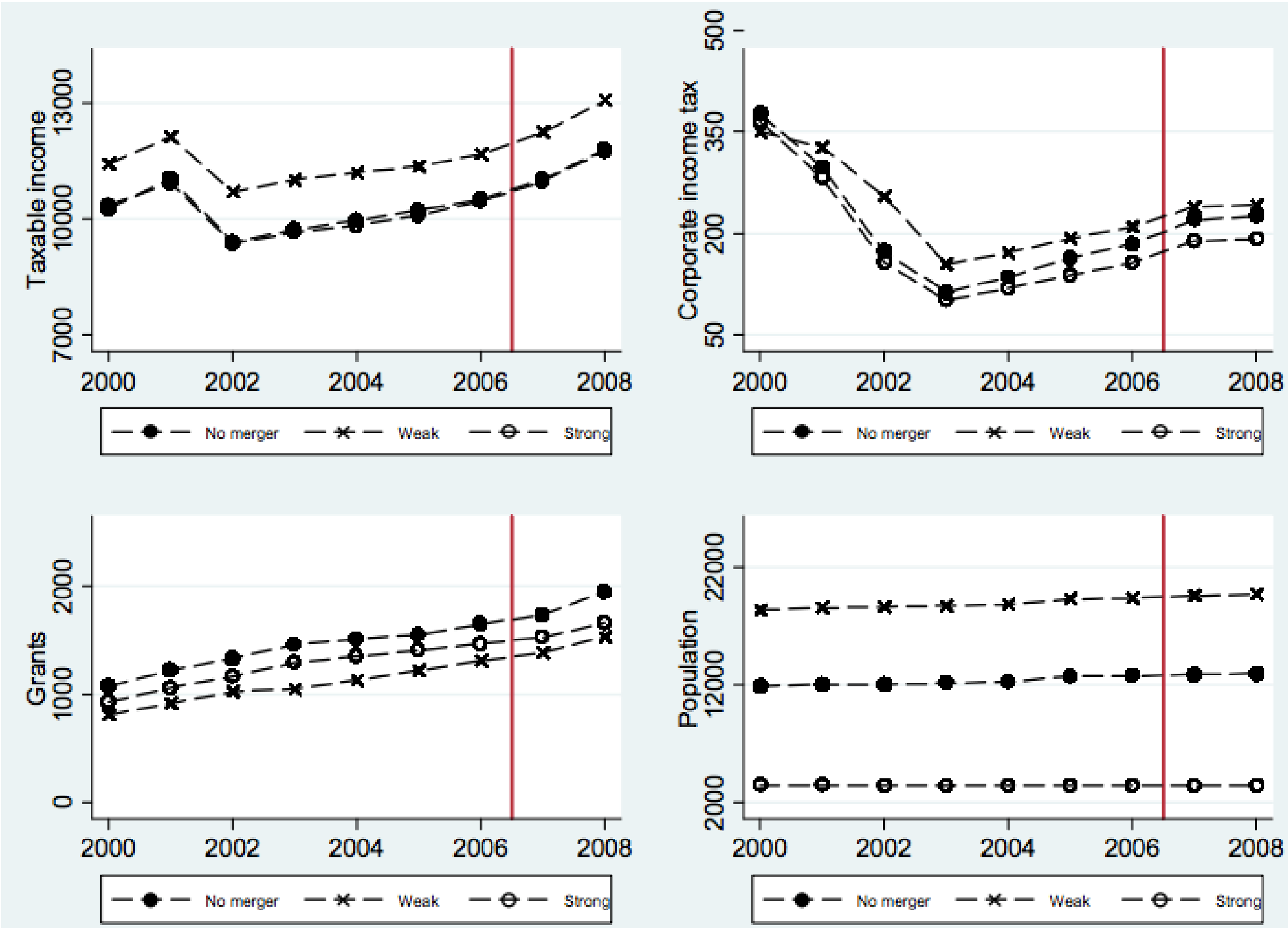
From this type of figure, you can see both the main result and its confidence interval and the pre-treatment placebo tests

These "event study graphs" are very useful when treatment timing varies across units

Main results – cash reserves per capita

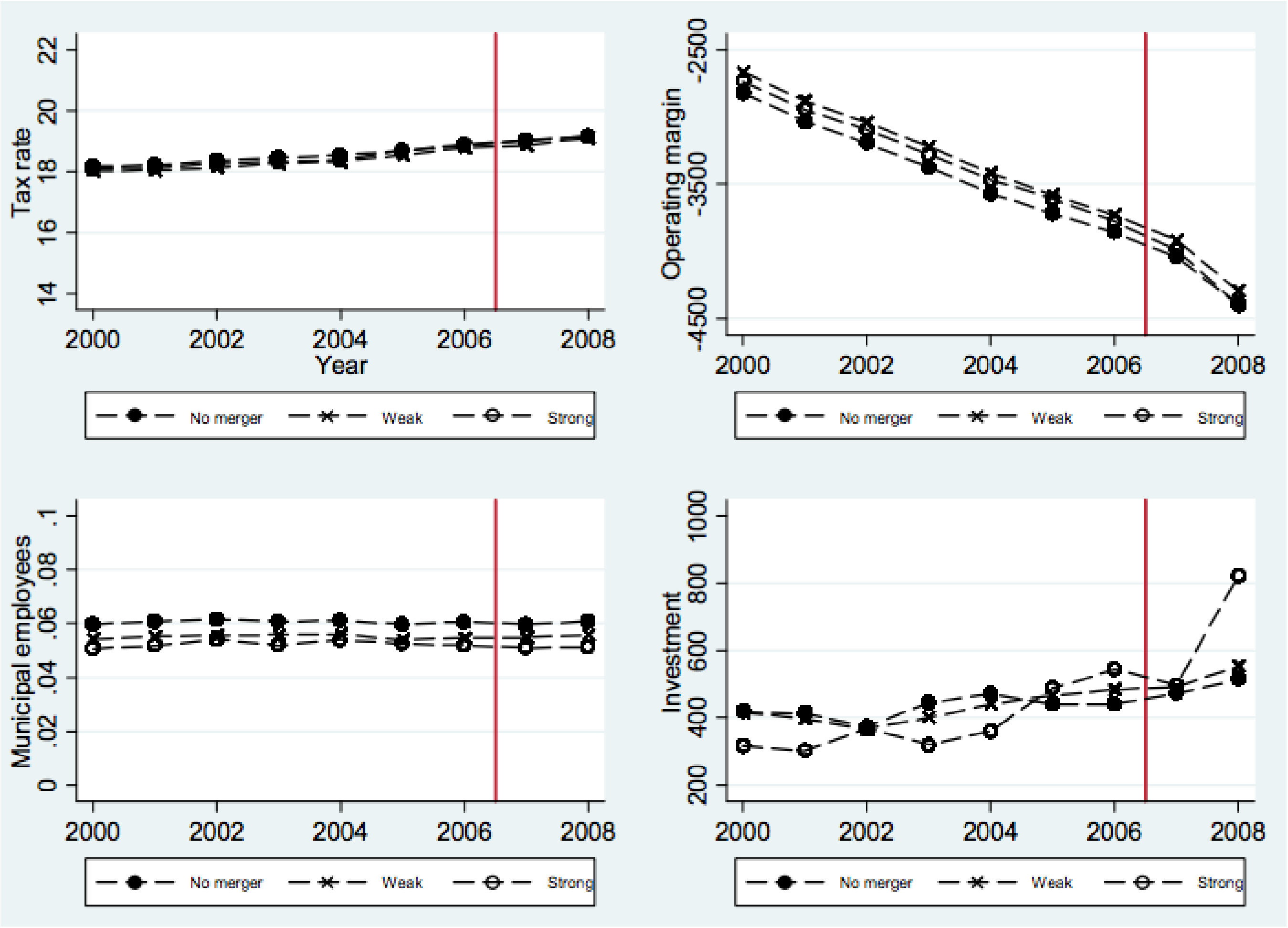


Other explanations



Did something else change at the same time that could explain the increase in debt?

Where did the money go?



Conclusions – Saarimaa & Tukiainen (2015)

- **Consistent with the law of 1/*n*, find clear evidence of free-riding among merging Finnish municipalities**
  - Debt-financed investments
- **Some policy advice:**
  - Policy 1: Politicians are may exploit a common pool if given the opportunity to do so (similar evidence from other countries)
  - Policy 2: During a merger process, some financial constraints on the local level may be a good idea (as in Denmark)
  - Policy 3: For the mergers to be beneficial overall, benefits need to be substantial

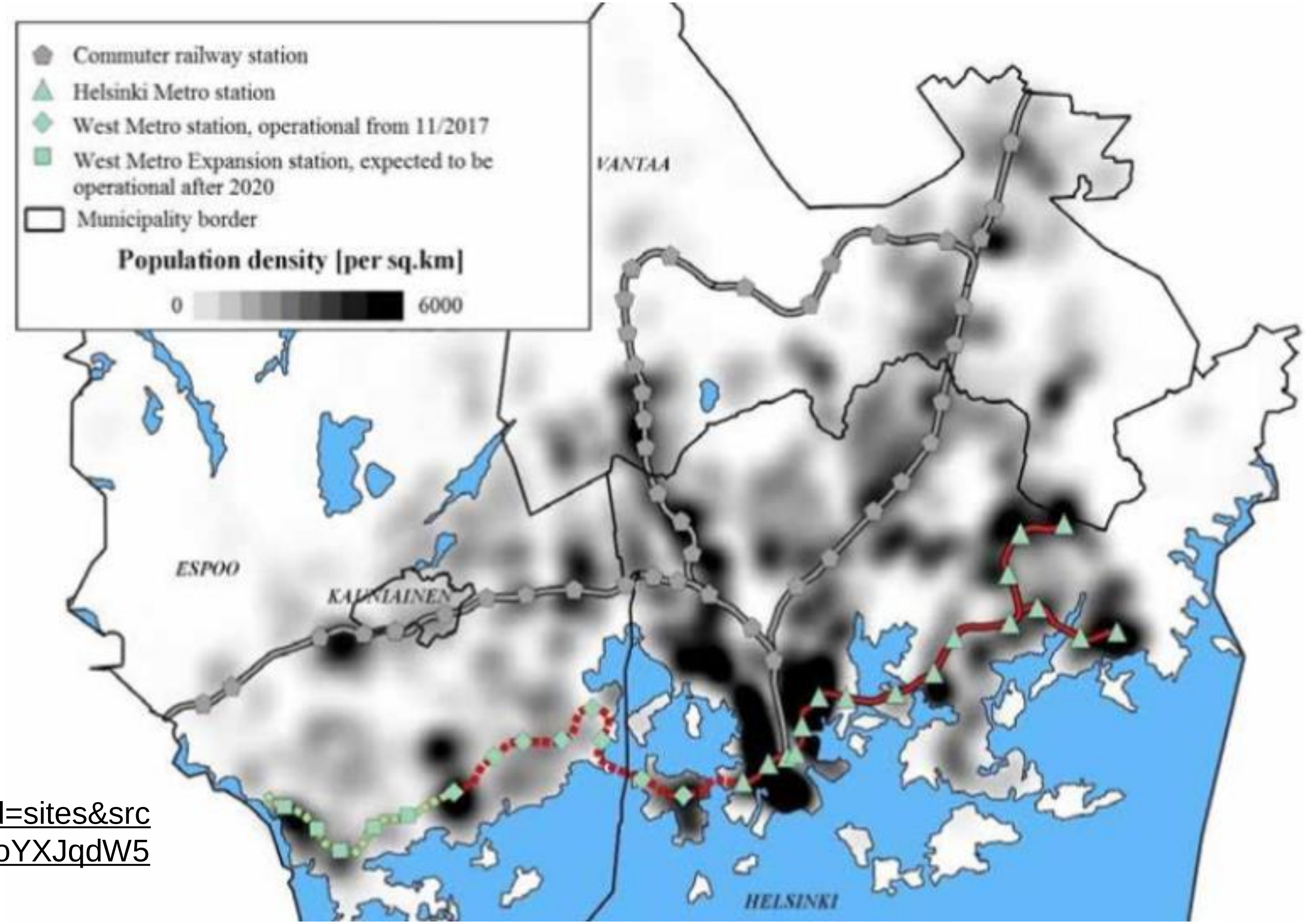
Other examples

Harjunen (2018): West Metro extension and house prices

Fig. A2. A route map of the metro in Helsinki in 2016 (© HSL 2016)

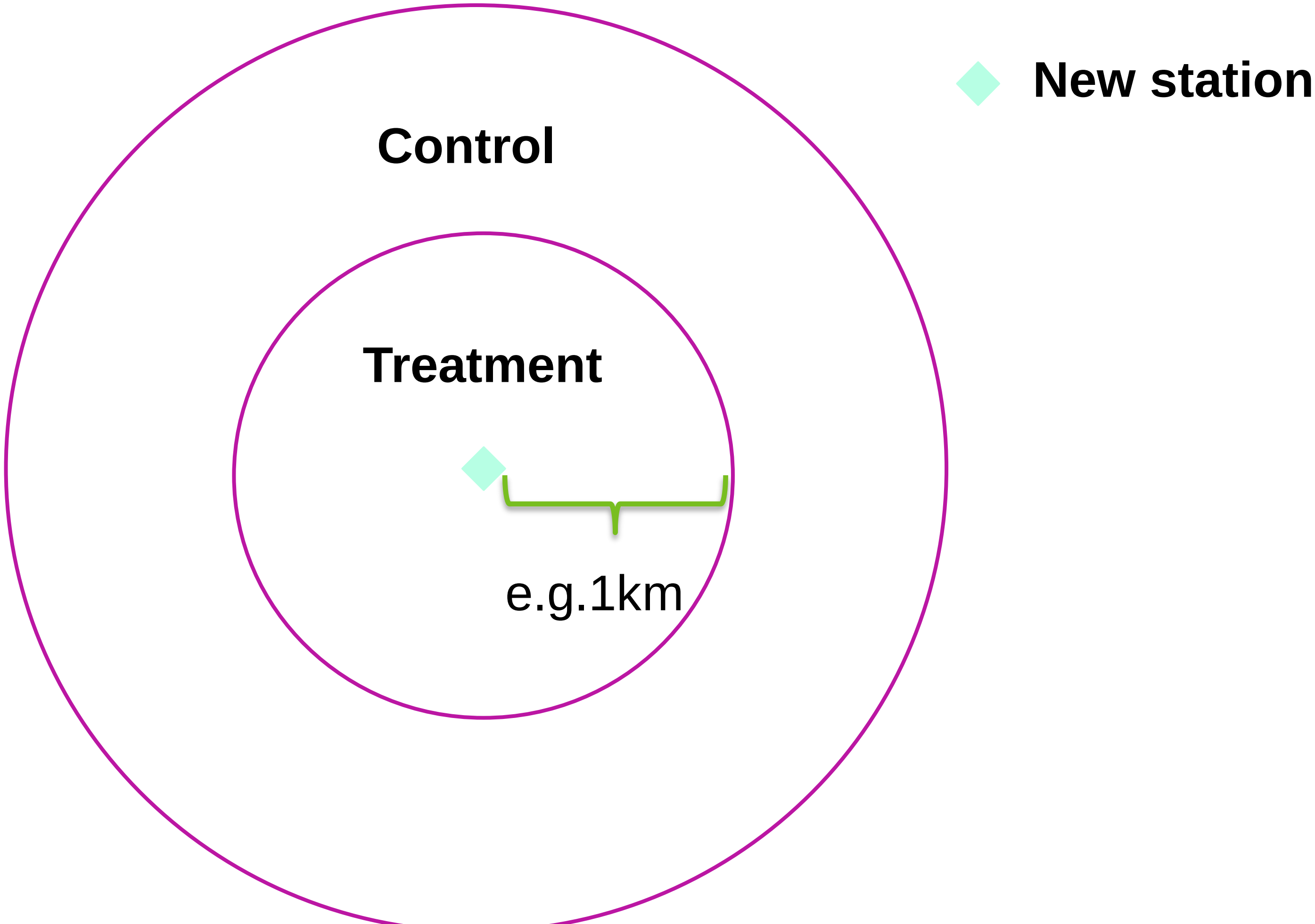
Fig. A3. A route map of the metro in Helsinki and Espoo after west metro is operational in 2017 (© HSL)

West Metro extension in the HMA



https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxvc2thcmloYXJqdW5lbnxneDoxY2JkNmZiMGM4ZWJmMjcx

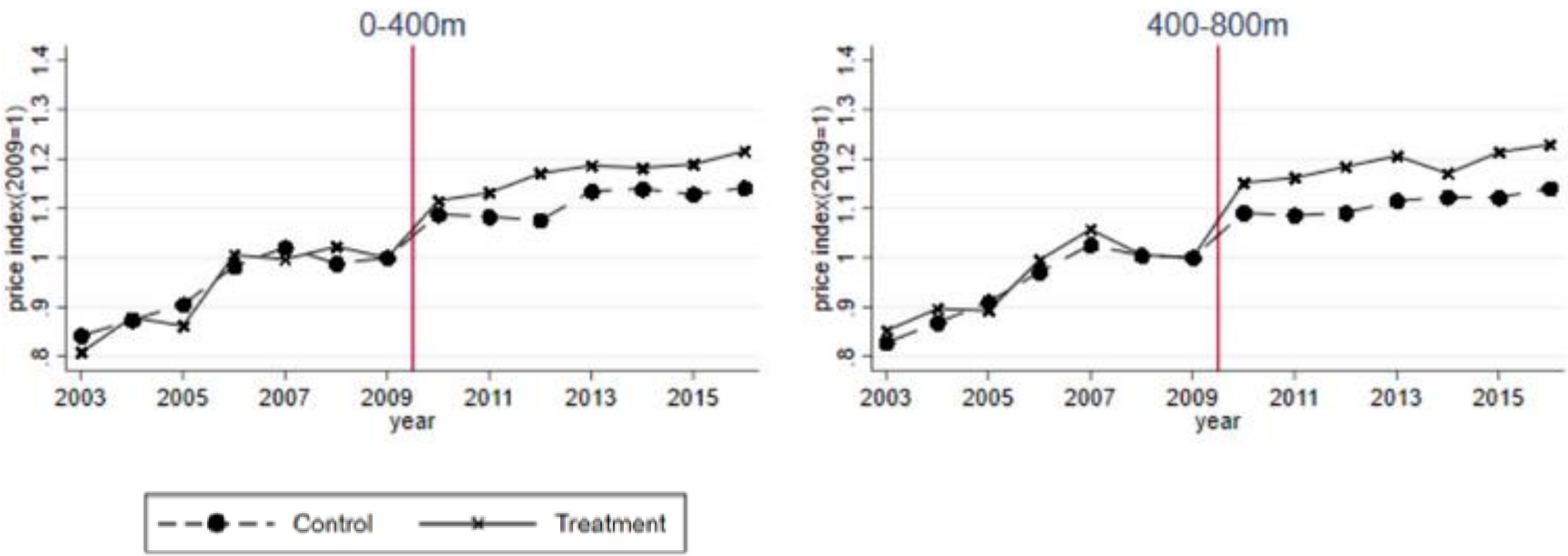
Is this a good DID design?



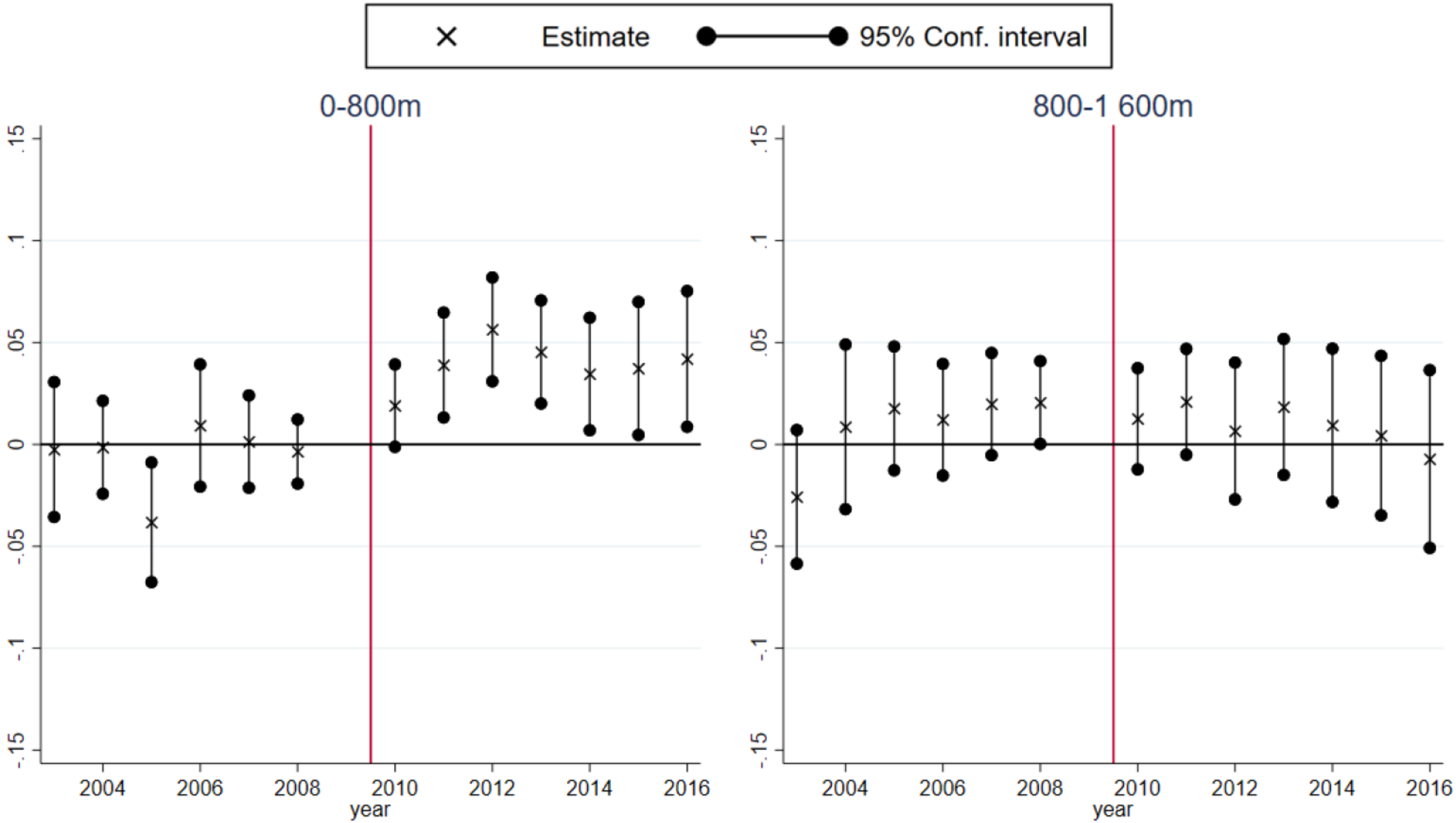
Data

Sample	Whole data (Helsinki and Espoo)	0 to 800m		800 to 1 600m	
Status		Treated	Control	Treated	Control
N	43 025	6 868	15 640	4 429	11 267
Sale price	223 668	252 024	196 154	311 661	199 122
	[110 007]	[119 458]	[78 980]	[156 343]	[82 107]
Square price	3 506	4 181	3 325	3 877	3 242
	[918]	[951]	[805]	[919]	[805]
Area	66	62	61	82	64
	[29]	[27]	[25]	[38]	[27]
Age	37	43	32	32	39
	[17]	[17]	[17]	[13]	[18]
Maint. Charge (€/m2)	3,5	3,8	3,5	3,5	3,5
	[1.2]	[1.1]	[1.2]	[1.2]	[1.3]
Floor number	2,4	2,7	2,5	2,3	2,3
	[1.6]	[1.7]	[1.5]	[1.5]	[1.4]
Floors in building	3,8	4,4	3,8	3,6	3,4
	[3.0]	[2.2]	[2.1]	[2.3]	[1.9]
Dist. to nearest station (m)	869	482	484	1 168	1 134
	[489]	[190]	[185]	[239]	[239]
Dist to CBD (km)	12	9	13	11,2	12,5
	[4.6]	[3.6]	[4.8]	[3.2]	[4.6]

Results

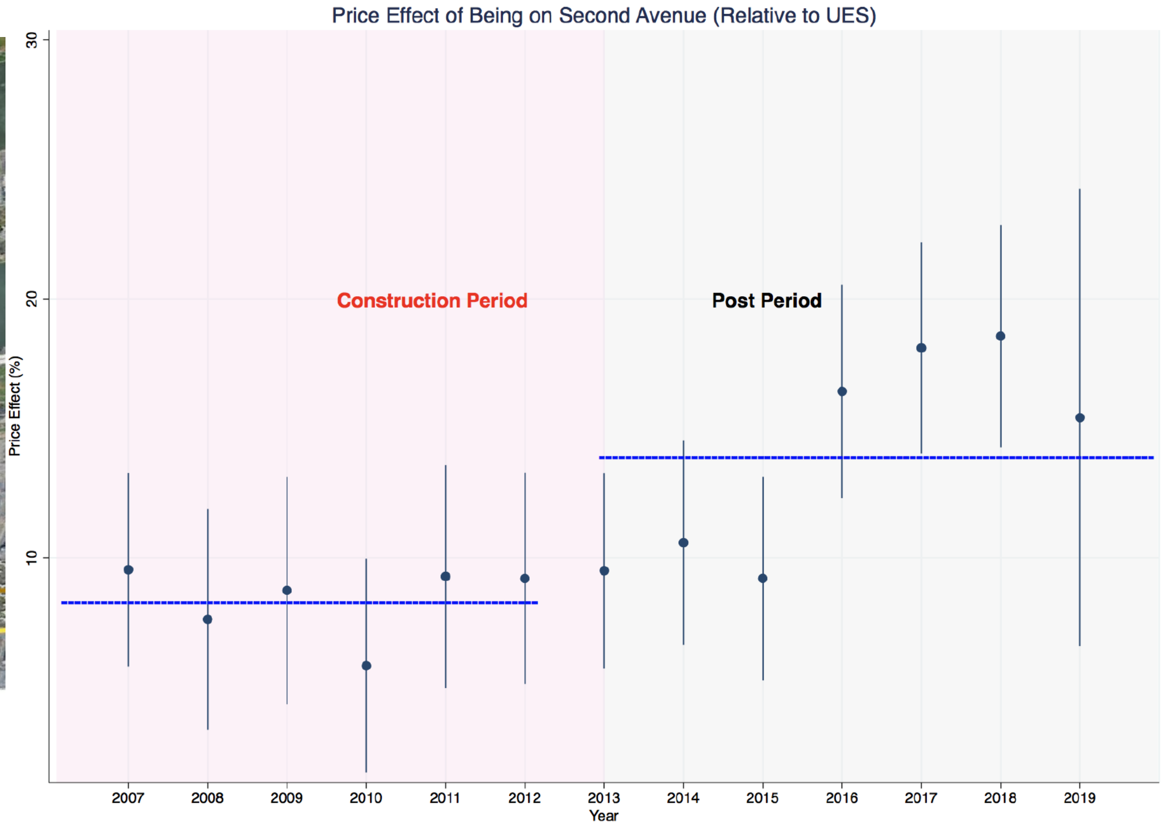

0-400m
400-800m
price index(2009=1)
.8
.9
1
1.1
1.2
1.3
1.4
2003
2005
2007
2009
2011
2013
2015
year
Control
Treatment

Results



Estimate
95% Conf. interval
0-800m
800-1 600m
.15
.1
.05
0
-.05
-.1
-.15
2004
2006
2008
2010
2012
2014
2016
year

Gupta, Van Nieuwerburgh and Kontokosta (2022): New subway line in NYC



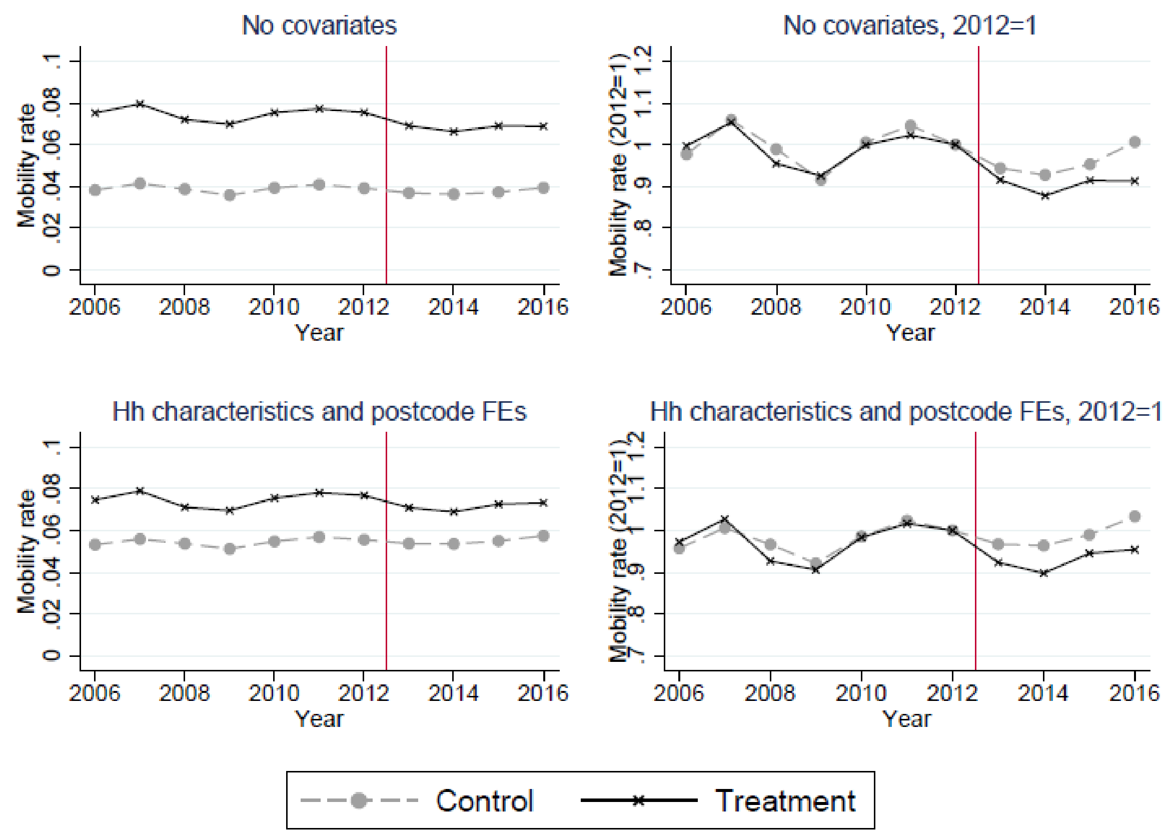
Link to paper: https://www.sciencedirect.com/science/article/pii/S0094119021001042

Eerola, Harjunen, Saarimaa & Lyytikäinen (2021): transfer tax

- **Exploit a tax reform implemented in March 2013**
- **Raised the transfer tax rate of apartments in multi-unit buildings without affecting the tax rate of single-family detached houses => a DID design**
  - Treatment group = homeowners living in housing units subject to the tax increase (tax rate 1.6% -> 2%)
  - Control group = homeowners whose housing units were unaffected by the reform (tax rate constant at 4%)
  - Outcome: mobility, i.e. probability that the household moves
  - Data: all Finns 2006–2016

Link to paper: https://www.sciencedirect.com/science/article/pii/S0094119021000498

DID results


No covariates
No covariates, 2012=1
Hh characteristics and postcode FEs
Hh characteristics and postcode FEs, 2012=1
Mobility rate
Mobility rate (2012=1)
Year
0
.02
.04
.06
.08
.1
.7
.8
.9
1
1.1
1.2
2006
2008
2010
2012
2014
2016
Control
Treatment

Is this a good design?

Table 2: Mobility rates before and after reform by origin and destination house type.

	House		Apartment	
	Pre-reform	Post-reform	Pre-reform	Post-reform
Moved to house	0.0144	0.0130	0.0193	0.0167
Moved to apartment	0.1012	0.0091	0.0334	0.0283
Moved to rental	0.1454	0.0153	0.0222	0.0233
Total	0.0391	0.0375	0.0749	0.0683

Notes: Table reports mobility rates of home-owners by origin and destination housing type.

Spillover to control group

- **In a housing market setting the design may be flawed due to spillovers between the treatment and control groups**
  - If homeowners in the treatment group move less often because of the tax increase, the homeowners in the control group may also be indirectly affected as now they have less trading partners
- **Complement empirical analysis with a model with two homeowner segments, apartments and single-family houses**
  - Combining the empirical and theoretical analyses, we find a roughly 7.2% reduction in treatment group mobility due to a 0.5 percentage point increase in the transfer tax rate
  - Our DID estimate of the effect is roughly 5.6%, suggesting a 22% downward bias in the estimate. The bias arises because mobility decreases by 1.6% also in the control group.

DID recap

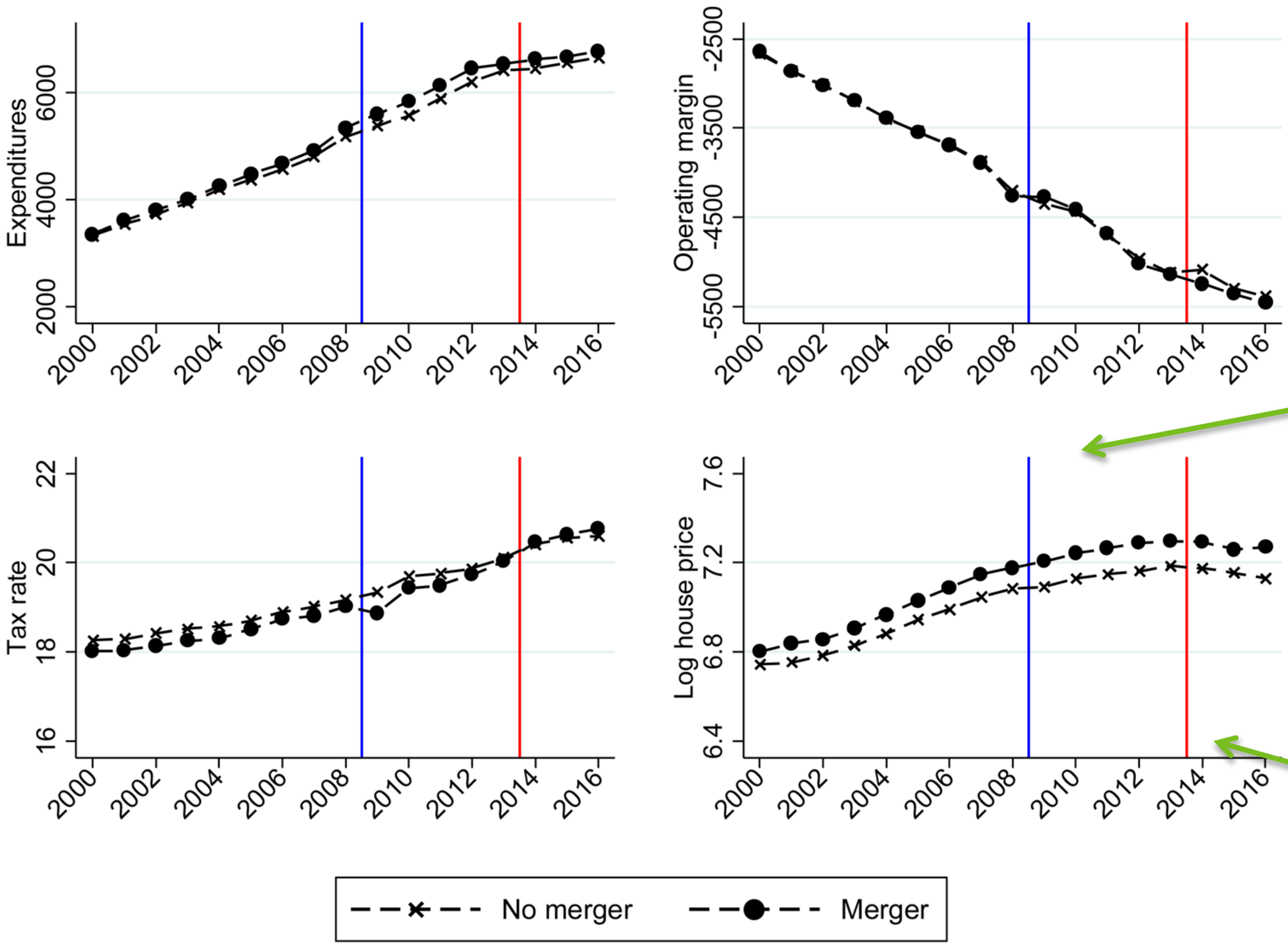
- **Idea:**
  - Even if treated and control groups differ in baseline characteristics, we can use observations on treatment and control groups before and after the treatment to estimate a causal effect
- **Assumptions:**
  - The potential outcomes (not observed) would have developed in a parallel manner for both groups in the absence of treatment
  - No spillovers
- **Testing for design validity:**
  - Visualization and testing: are trends in outcomes parallel before treatment? (indirect test)
  - Is there anything else that could have happened to one group but not the other? (know your institutional setting!)

Implementation

- **In practice, estimation of the treatment effect is implemented using regression models**
  - Produces "automatically" the estimate of the treatment effect and the standard error and we can add control variables
- **Data can be either**
  - Panel data: data where you observe the same individuals (units) in multiple time periods
  - Repeated cross-sectional data: e.g. repeated random samples from a population where you observe different individuals in different time periods
- **There are complicated issues concerning staggered designs and the literature is moving forward on this**

Extra

Harjunen, Saarimaa & Tukiainen (2021)



Mergers took place

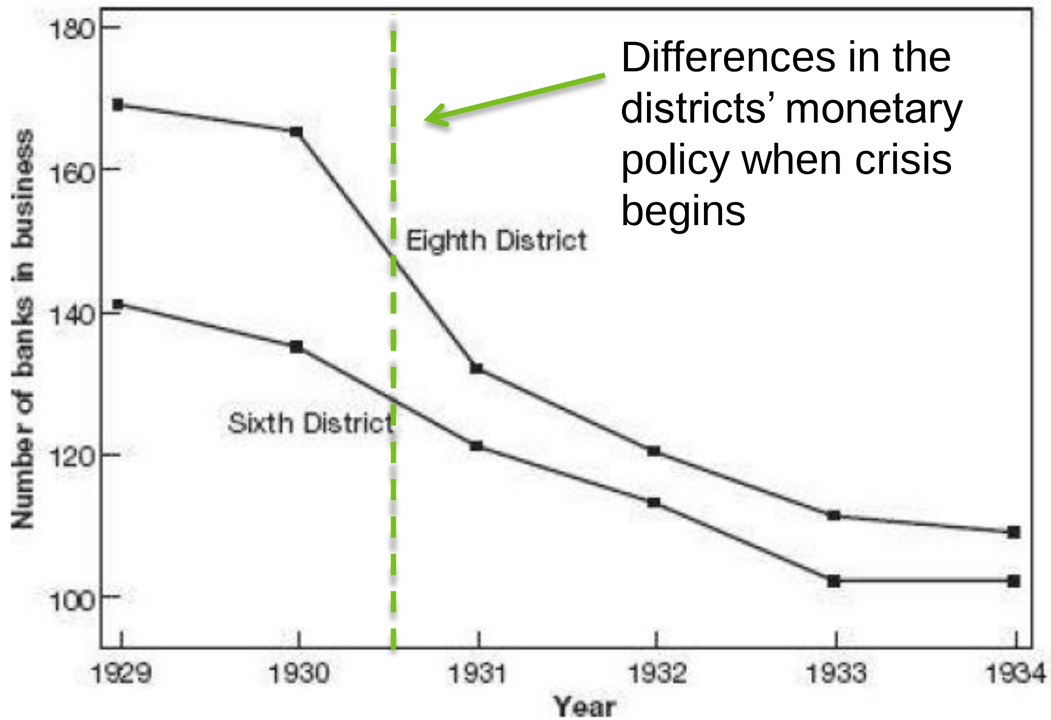
A five year lay-off protection ended

https://www.cambridge.org/core/journals/political-science-research-and-methods/article/political-representation-and-effects-of-municipal-mergers/1DC538037E1E3DC260EA276CD845318D#article

Richardson & Troots (2009)

FIGURE 5.2

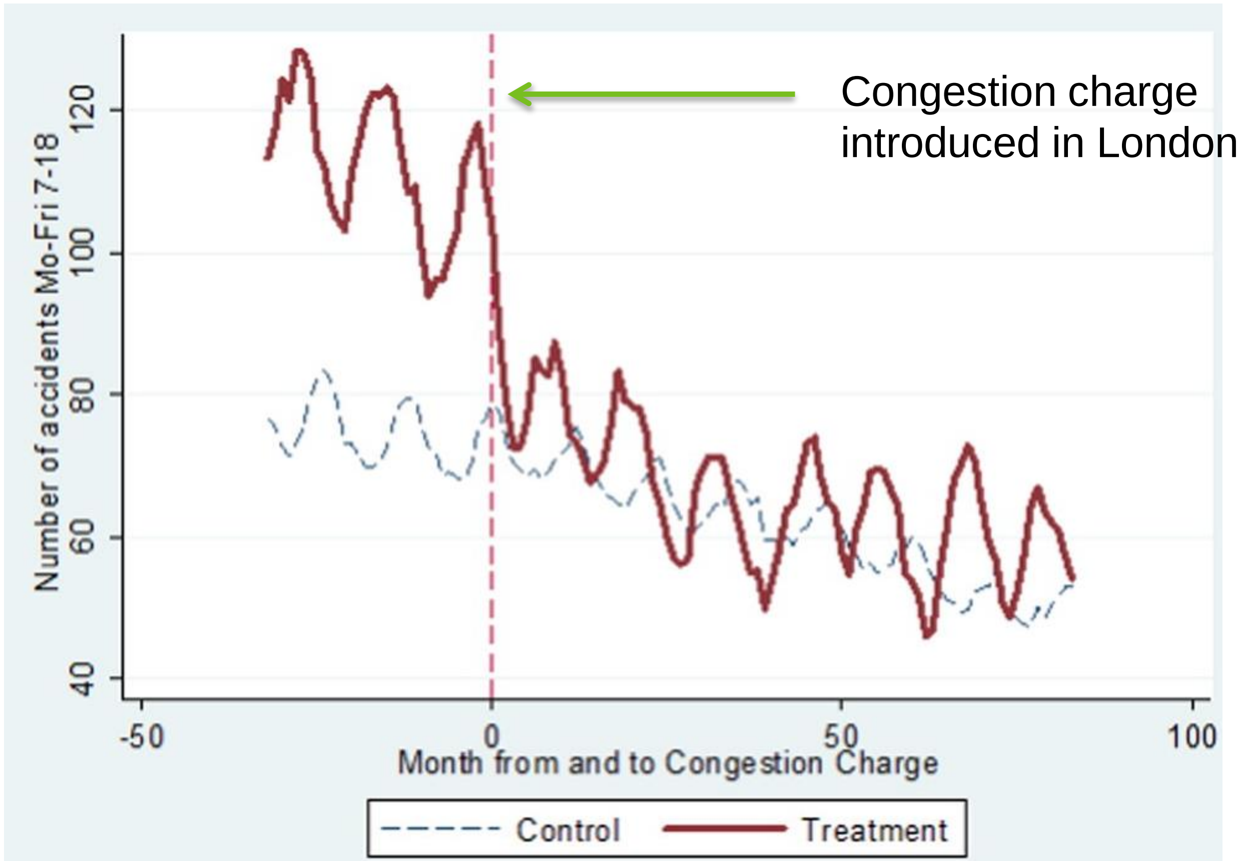
Trends in bank failures in the Sixth and Eighth Federal Reserve Districts



Note: This figure shows the number of banks in operation in Mississippi in the Sixth and Eighth Federal Reserve Districts between 1929 and 1934.

Link to paper: https://www.journals.uchicago.edu/doi/abs/10.1086/649603

Green, Haywood and Navarro (2016)



https://www.sciencedirect.com/science/article/pii/S0047272715001929