Applied Microeconometrics I

Lecture 9: Differences-in-differences (continued)

Tuomas Pekkarinen

Aalto University

October 6, 2021

Lecture Slides

Nobel Prize in Economics 2021

The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2021

Ill. Niklas Elmehed © Nobel Prize Outreach.

David Card

Prize share: 1/2

Ill. Niklas Elmehed © Nobel Prize Outreach.

Joshua D. Angrist

Prize share: 1/4

Ill. Niklas Elmehed © Nobel Prize Outreach.

Guido W. Imbens

Prize share: 1/4

Nobel Prize in Economics 2021

- This course is very much based on the contributions of Angrist, Card, and Imbens
  - The textbook of course
  - Angrist and Krueger (1991): Quarter of birth as an IV for schooling
  - Angrist, Imbens, and Rubin (1996): LATE interpretation of IV
  - Angrist et al (2011): IV as an RCT with imperfect compliance
  - Angrist (2006): LATE and ATT
  - Card and Krueger (1994): Dif-in-dif analysis of minimum wages

What did we do last time?

- Differences-in-differences
- Two groups: Treatment and control
- Two periods: t and $t+1$

$$Y_{it} = \alpha_i + \lambda_t + \rho D_{it} + \epsilon_{it}$$

  where $i = \{T, C\}$, $t = \{t, t+1\}$, $D_{T,t} = D_{C,t} = D_{C,t+1} = 0$, and $D_{T,t+1} = 1$
- OLS estimation of ρ gives the treatment effect

Difference-in-differences (DID)

- Fixed effects vs. Differences-in-differences
- Example: Card and Krueger (1994)
- Improving traditional DID set up
- Other examples

Fixed effects vs. Differences-in-differences

- Differences-in-differences is an application of the familiar individual fixed-effects model with panel data:
$$Y_{it} = \alpha_i + \lambda_t + \rho D_{it} + X_{it}^{'}\beta + \epsilon_{i,t}$$
where t denotes time (or something else, we return to this later) and i individuals
- α_i varies across i but not across t whereas λ_t varies across t but not across i
- The key to the identification of ρ is that we have repeated observations on i over t

Fixed effects vs. Differences-in-differences

- Then we can "eliminate" α_i and identify ρ either by converting the data into deviations from i-specific means:

$$Y_{it} - \bar{Y_i} = \lambda_t - \bar{\lambda} + \rho(D_{it} - \bar{D_i}) + (X_{it} - \bar{X_i})^{'}\beta + (\epsilon_{it} - \bar{\epsilon_i})$$

- or by differencing over t

$$Y_{it} - Y_{it-1} =$$
$$\lambda_t - \lambda_{t-1} + \rho(D_{it} - D_{it-1}) + (X_{it} - X_{it-1})^{'}\beta + \epsilon_{it} - \epsilon_{it-1}$$

- These transformations will provide more or less the same results

Fixed effects vs. Differences-in-differences

- Differences-in-differences is an application of the fixed effects model where:
  - i often refers to more aggregate groups
  - Units in the treatment group start being exposed to the treatment at time t (i.e.: a new a law is implemented in a certain region, but not in the control regions)
- The differences-in-differences framework helps us to think much more carefully about identification issues.

Famous example: Card and Krueger (1994)

Effect of Minimum wages on employment

- Theory:
  - In a competitive model the result of increasing the minimum wage is to reduce employment.
  - However, in a monopsonistic model an increase in minimum wages can actually increase employment.
- On April 1, 1992, New Jersey raised the state minimum wage from $4.25 to $5.05, whereas in the bordering state of Pennsylvania the minimum wage stayed at $4.25 throughout this period.
- Card and Krueger (1994) evaluated the effect of this change on the employment of low wage workers.
- They conducted a survey to some 400 fast food restaurants from the two states just before the NJ reform, and a second survey to the same outlets 7-8 months after.

Treatment and Control Locations

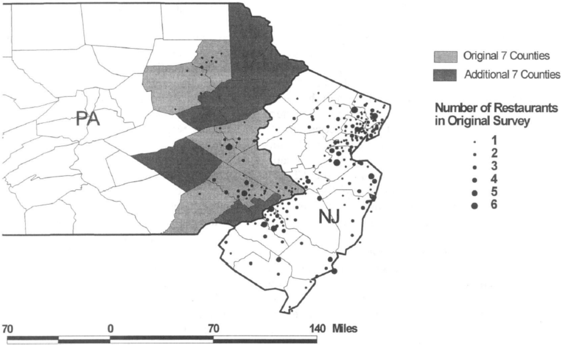
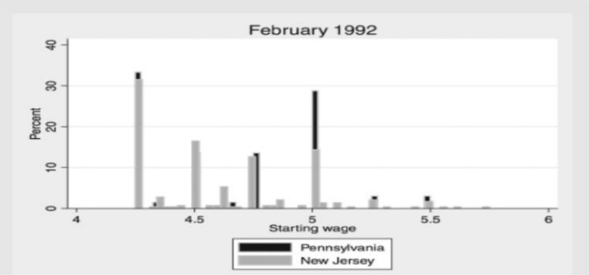


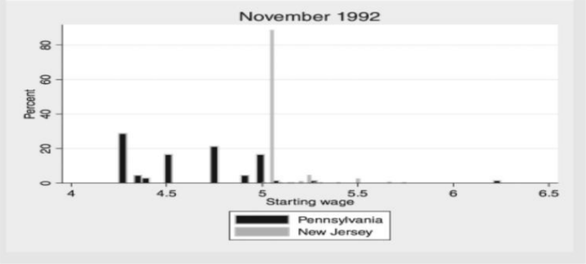
FIGURE 1. AREAS OF NEW JERSEY AND PENNSYLVANIA COVERED BY ORIGINAL SURVEY AND BLS DATA

Famous example: Card and Krueger (1994)

- Characteristics of fast food restaurants:
  1. A large source of employment for low-wage workers.
  2. They comply with minimum wage regulations (especially franchised restaurants).
  3. Fairly homogeneous job, so good measures of employment and wages can be obtained.
  4. Easy to get a sample frame of franchised restaurants (yellow pages) with high response rates.
  5. Response rates 87% and 73% (less in Penn, because the interviewer was less persistent).

Distribution of wage rates, before and after





Famous example: Card and Krueger (1994)

- Treatment group: Fast-food restaurants in New Jersey ($i = NJ$)
- Control group: Fast-food restaurants in Pennsylvania ($i = PA$)
- Denote the period before April 1992 with $t = 0$ and period after April 1992 with $t = 1$
- At period $t = 0$ the minimum wage is $w_{PA,0} = w_{NJ,0} = w$
- At period $t + 1$ the minimum wages differ so that $w_{PA,1} = w$ and $w_{NJ,1} = w + \Delta$

Famous example: Card and Krueger (1994)

- Write employment is state i at period t as:

$$L_{it} = \alpha_i + \lambda_t + \rho w_{it} + \epsilon_{it}$$

- Now:

$$\begin{aligned} E(L_{PA,0}) &= \alpha_{PA} + \lambda_0 + \rho w \\ E(L_{NJ,0}) &= \alpha_{NJ} + \lambda_0 + \rho w \\ E(L_{PA,1}) &= \alpha_{PA} + \lambda_1 + \rho w \\ E(L_{NJ,1}) &= \alpha_{NJ} + \lambda_1 + \rho(w + \Delta) \end{aligned}$$

- The differences-in-differences estimator of ρ is:

$$\rho = [E(L_{NJ,1}) - E(L_{NJ,0})] - [E(L_{PA,1}) - E(L_{PA,0})]$$

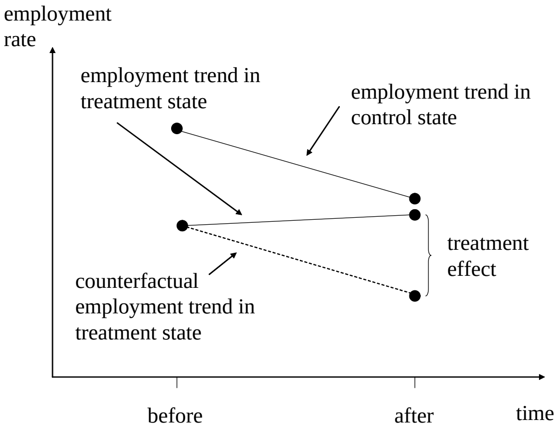


Figure 5.2.1: Causal effects in the differences-in-differences model

Table 5.2.1: Average employment per store before and after the New Jersey minimum wage increase

Variable		PA (i)	NJ (ii)	Difference, NJ-PA (iii)
1.	FTE employment before,	23.33	20.44	-2.89
	all available observations	(1.35)	(0.51)	(1.44)
2.	FTE employment after,	21.17	21.03	-0.14
	all available observations	(0.94)	(0.52)	(1.07)
3.	Change in mean FTE	-2.16	0.59	2.76
	employment	(1.25)	(0.54)	(1.36)

Notes: Adapted from Card and Krueger (1994), Table 3. The table reports average full-time equivalent (FTE) employment at restaurants in Pennsylvania and New Jersey before and after a minimum wage increase in New Jersey. The sample consists of all stores with data on employment. Employment at six closed stores is set to zero. Employment at four temporarily closed stores is treated as missing. Standard errors are reported in parentheses

Simple Regression Dif-in-Dif

We can also use the following regression to estimates the differences-in-differences coefficient:

$$Y_{ist} = \alpha + \gamma TREAT_s + \lambda AFTER_t + \rho(AFTER_t * TREAT_s) + \varepsilon_{ist}$$

- Y_{ist} is the number of full-time employees working in establishment i, located in state $s \in$\{NJ,PA\}, in period $t \in$\{Feb 1992, Nov 1992\}
- $TREAT_s$: dummy variable equal to 1 when s=\{NJ\}
- $AFTER_t$: dummy variable equal to 1 when t=\{Nov 1992\}
- $TREAT_s * AFTER_t$ interaction term that takes value one when s=\{NJ\} & t=\{Nov 1992\}

Simple Regression DD: interpreting coefficients

$$Y_{ist} = \alpha + \gamma TREAT_s + \lambda AFTER_t + \rho(AFTER_t * TREAT_s) + \varepsilon_{ist}$$

- α: average Y in non-treated group (PA) in the pre-treatment period
- γ: difference in Y between treatment group (NJ) and control group (PA) in the pre-treatment period
- λ: ΔY in the control group between the pre-treatment and the treatment period
- ρ: ΔY in the treatment group between the pre-treatment and the treatment period, relative to the ΔY in the control group
  $\rightarrow$ captures effect of the policy!

```
.
. xi: reg EMPTOT i.NEWJERSEY*i.AFTER, cluster(ID)
i.NEWJERSEY       _INEWJERSEY_0-1     (naturally coded; _INEWJERSEY_0 omitted)
i.AFTER           _IAFTER_0-1         (naturally coded; _IAFTER_0 omitted)
i.NEW~Y*i.AFTER   _INEWXAFT_#_#       (coded as above)

Linear regression                                      Number of obs =     794
                                                       F(  3,   409) =    1.80
                                                       Prob > F      =  0.1462
                                                       R-squared     =  0.0074
                                                       Root MSE      =  9.4056

                                   (Std. Err. adjusted for 410 clusters in ID)
------------------------------------------------------------------------------
             |               Robust
      EMPTOT |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
_INEWJERSEY_1 |  -2.891761   1.439546    -2.01   0.045    -5.721593   -.0619281
   _IAFTER_1 |  -2.165584   1.218025    -1.78   0.076    -4.559954    .2287855
_INEWXAFT_1_1 |   2.753606   1.306607     2.11   0.036     .1851025    5.322109
       _cons |   23.33117   1.346536    17.33   0.000     20.68417    25.97816
------------------------------------------------------------------------------
```

Regression DD: Including controls

$Y_{ist} =$
$\alpha + \gamma TREAT_s + \lambda AFTER_t + \rho(AFTER_t * TREAT_s) + X'_{ist}\beta + \varepsilon_{ist}$

- Sometimes you may want to control for certain time-varying covariates (X_{ist}).
- Including controls may help to obtain more precise estimates, but make sure not to include *bad controls*

Results of the CK Study

- Wages increased by 10% in NJ, remained constant PA
- ... but employment rose in NJ and decreased in PA
- The dif-in-dif estimate suggests that **the rise of minimum wage *increased* employment**
- Result robust to alternative specifications and to an alternative control group (workers with salaries above the minimum salary)

Reactions to the CK Study

- Angus Deaton: "*The reception accorded to Princeton faculty by their colleagues in other institutions is what might be expected by the friends and defenders of child-molesters*"
- James Buchanan in the Wall Street Journal:
  "*no self-respecting economist would claim that increases in the minimum wage increase employment. Such a claim, if seriously advanced, becomes equivalent to a denial that there is even minimum scientific content in economics, and that, in consequence, economists can do nothing but write as advocates for ideological interests. Fortunately, only a handful of economists are willing to throw over the teaching of two centuries; we have not yet become a bevy of camp-following whores*"

See Angus Deaton's "Letters from America" for more:
www.princeton.edu/~deaton/downloads/letterfromamerica_oct1996.html

Reactions to the CK Study

- Neumark and Wascher (2000, AER)
  - CK data has a lot of measurement error
  - data provided by Employment Policies Institute reveal that the minimum wage rise did decrease employment
- Card and Krueger (2000, AER)
  - administrative data from Bureau of Labor Statistics confirm the key findings of the 1994 paper
  - "*calls into question the representativeness of the sample assembled by Berman, Neumark and Wascher*"

See John Schimtt's "Cooked to Order" for more:
www.prospect.org/cs/articles?article=cooked_to_order

Treatment and Control Locations (Card and Krueger, 2000)

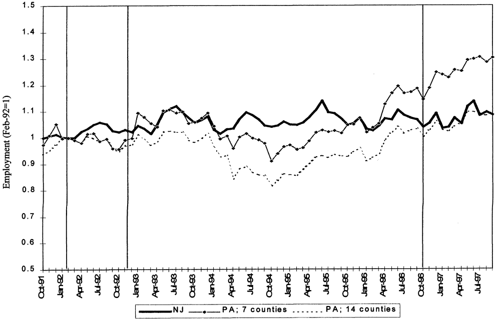


FIGURE 2. EMPLOYMENT IN NEW JERSEY AND PENNSYLVANIA FAST-FOOD RESTAURANTS, OCTOBER 1991 TO SEPTEMBER 1997

Note: Vertical lines indicate dates of original Card-Krueger survey and the October 1996 federal minimum-wage increase.
Source: Authors' calculations based on BLS ES-202 data.

- Some potential methodological concerns:
  1. The authors do not examine how the trends evolved in the past. Information from future trends suggests that they are not parallel.
  2. At the end of the day, we only have two observations. Possible common shocks may affect the treatment or the control group.
  3. Other policies?
  4. Note also the tension between having observations that are geographically close and the potential existence of an impact of the treatment on the control group.

Improving Traditional Dif-and-Dif Set Up

- The crucial assumption in DD set up is that the control group provides information about how the treatment group would have evolved in the absence of treatment (parallel trends)
- With more than two periods this can be investigated in several ways...
  1. Illustrate graphically that the outcomes evolved similarly in the years before the policy was implemented
  2. Estimate placebo models: does the placebo policy introduced in t-1, t-2, etc. have any significant impact?
  3. Include group-specific trends

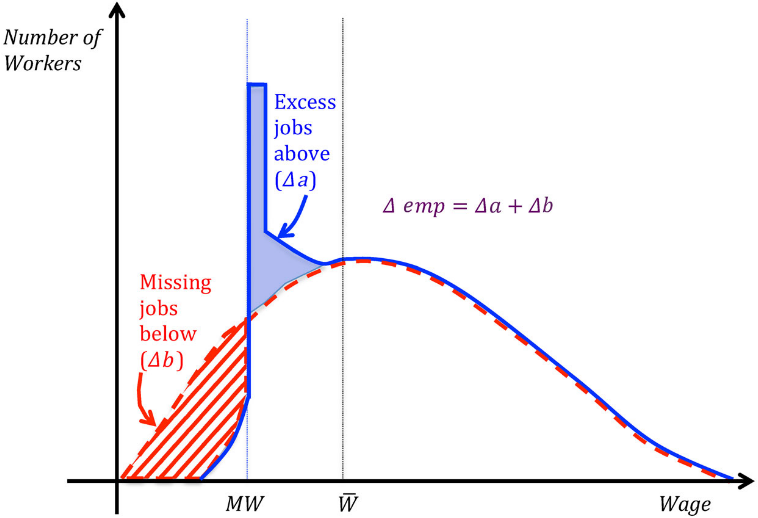
Example:Cengiz et al (2019): The effect of minimum wage on low-wage jobs

- Cengiz et al (2019) is the latest addition to the literature on the effects of minimum wages
- Like Card and Krueger(1994) it is a dif-in-dif paper but modern in its approach
- Exploit 138 state-level changes in minimum wages between 1979 and 2016 to identify the effect
- Instead of a particular sector, focus on the effect on the total employment

Example: Cengiz et al (2019): The effect of minimum wage on low-wage jobs

- Dif-in-dif design to estimate the impact of minimum wage increases on the entire distribution of wages
- Zoom on the bottom of the distribution to estimate the impact on employment and wages of affected workers
- Idea:
  - An increase in minimum wages will directly affect jobs that were paying less than the minimum wage before the raise
  - Some of these jobs will disappear as a result of the raise
  - Some of these jobs will increase wages and show up as "excess jobs" at and above the minimum wages
  - At the upper tail of the wage distribution we shouldn't see any effects of minimum wages

The impact of minimum wages on the frequency distribution of wages



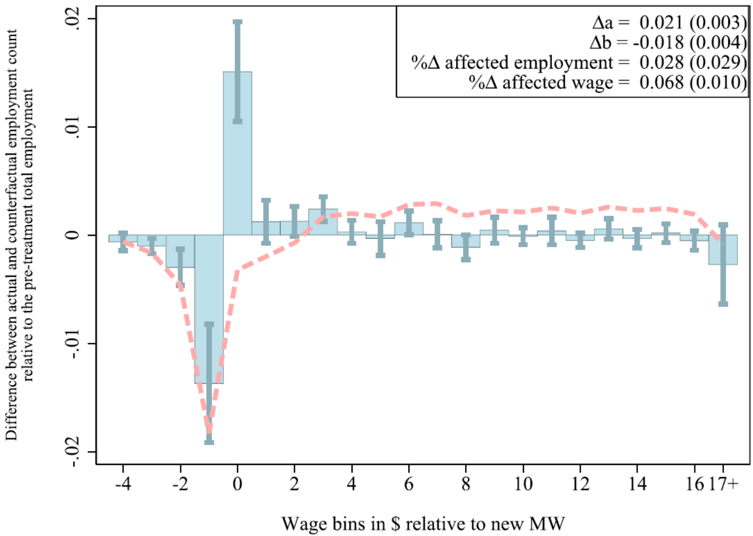
Example: Cengiz et al (2019): The effect of minimum wage on low-wage jobs

- Estimate employment changes in bins of the wage distribution relative to the minimum wage for three years prior to and five years following an event

$$\frac{E_{sjt}}{N_{st}} = \sum_{\tau=-3}^{4} \sum_{k=-4}^{17} \alpha_{\tau k} I_{sjt}^{\tau k} + \mu_{sj} + \rho_{jt} + \Omega_{sjt} + u_{sjt}$$

  where E_{sjt} is the employment in wage bin j in state s at a quarter t, N_{st} is the population in s at t
- The treatment dummy $I_{sjt}^{\tau k}$ is equal to one if the minimum wage was increased τ years from date t for bin j that falls between k and $k+1$ dollars of the new minimum wages
- State-bin effects μ_{sj}, period-bin effects ρ_{jt}

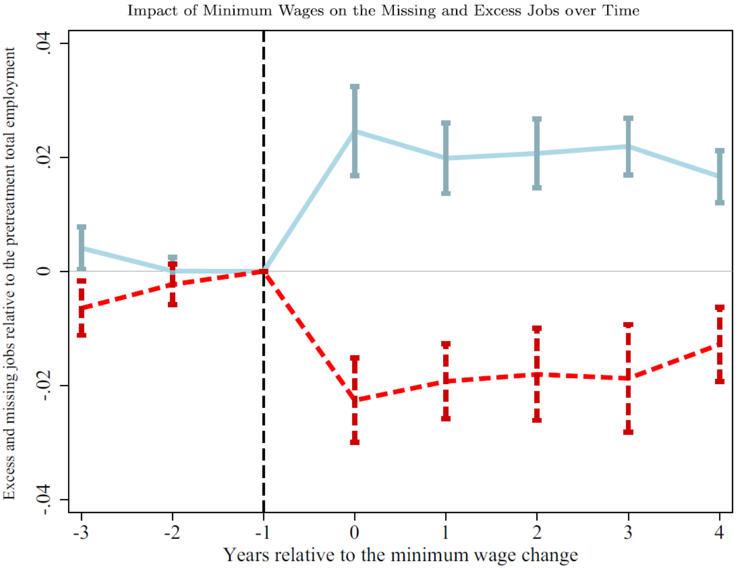
The impact of minimum wages on the wage distribution



Example:Cengiz et al (2019): The effect of minimum wage on low-wage jobs

- The autors have data on states several years before and after the increase in minimum wage change
- This allows the to estimate the "lead"and "lag" effects to assess pre-existing trends
- If employment below and above the new minimum wage diverge already before the raise, parallel trends assumption could be violated
- No evidence of diverging trends before the raise

The impact of minimum wages on the wage distribution



Example: Ashenfelter and Rouse (1998): Income, schooling and identical twins

- So far we have used differences-in-differences to identify causal effects using data on units over time
- However, we can have multiple observations of a unit in other dimensions as well
- Examples:
  - Plants within firms
  - Family member within families
  - Pupils within classes or schools
- We can also exploit this kind of variation in a differences-in-differences style strategy

Example: Ashenfelter and Rouse (1998): Income, schooling and identical twins

- Let's return to the problem of estimating the effect of schooling S_{if} on earnings Y_{if}:
  $$Y_{if} = \alpha + \rho S_{if} + \gamma A_f + \epsilon_{if}$$
  where f denotes family and A_f is unobserved
- A_f is fixed within families now captures all the unobserved determinants of earnings that are fixed within families
- It is highly unlikely that $Cov(S_{if}, A_f) = 0$
- Hence, omitting A_f would lead to biased estimates of ρ

Example: Ashenfelter and Rouse (1998): Income, schooling and identical twins

- Ashenfelter and Rouse have data on identical twins $i = 1, 2$

$$\text{Twin 1: } Y_{1f} = \alpha + \rho S_{1f} + \gamma A_f + \epsilon_{1f}$$
$$\text{Twin 2: } Y_{2f} = \alpha + \rho S_{2f} + \gamma A_f + \epsilon_{2f}$$

- If A_f is common to the pair of twins, then differencing yields:

$$Y_{1f} - Y_{2f} = \rho(S_{1f} - S_{2f}) + (\epsilon_{1f} - \epsilon_{2f})$$

- Under these assumptions estimaing ρ with the differenced equation gives us the causal effect of schooling on earnings

OLS estimates in the population and in the twin sample

TABLE II
OLS ESTIMATES OF THE (MEAN) RETURN TO SCHOOLING USING THE CPS AND TWINS DATA

	CPS[a]	Identical twins
	OLS (1)	OLS (2)
Own education	0.085	0.110
	(0.0003)	(0.009)
Age	0.071	0.104
	(0.0004)	(0.010)
Age^2 (÷100)	−0.074	−0.106
	(0.0005)	(0.013)
Female	−0.253	−0.318
	(0.001)	(0.040)
White	0.087	−0.100
	(0.002)	(0.072)
Sample size	476,851	680
R^2	0.332	0.339

Standard errors are in parentheses. All regressions include a constant.

a. The Current Population Survey (CPS) sample is drawn from the 1991–1993 Outgoing Rotation Group files. The sample includes workers age 18–65 with an hourly wage greater than $1 per hour in 1993 dollars; the regression is weighted using the earnings weight. We converted the 1992 and 1993 education categories into a continuous measure according to the categorization suggested by Park [1994].

First difference estimates

	GLS (1)	GLS (2)	3SLS (3)	First-difference (4)	First-difference by IV (5)
Own education	0.102	0.066	0.091	0.070	0.088
	(0.010)	(0.018)	(0.024)	(0.019)	(0.025)
Avg. education		0.051	0.033		
$[(S_1 + S_2)/2]$		(0.022)	(0.028)		
Age	0.104	0.103	0.103		
	(0.013)	(0.013)	(0.013)		
Age^2 (÷100)	−0.107	−0.104	−0.104		
	(0.015)	(0.015)	(0.015)		
Female	−0.315	−0.309	−0.306		
	(0.049)	(0.049)	(0.049)		
White	−0.106	−0.105	−0.101		
	(0.090)	(0.091)	(0.091)		
Covered by a union					
Married					
Tenure (years)					
Sample size	680	680	680	340	340
R^2	0.262	0.264	0.267	0.039	

Example: Ashenfelter and Rouse (1998): Income, schooling and identical twins

- Why do identical twin have different levels of schooling?
- What if it is only mis-reporting?
- More general problem: First differencing (or conversion to deviations from unit-specific means) exacerbates measurement error problems
- Intuition: If identical twins always have the same level of schooling in reality, then all the observed variation will be just measurement error

Example: Ashenfelter and Rouse (1998): Income, schooling and identical twins

- Ashenfelter and Rouse solution: Assume that twins report each other's schooling with independent measurement errors
- Then we can use one's twins reporting of one's own schooling as an instrument for one's own reporting
- Intuition: Both my recollection and my twin sibling's recollection are mismeasured assessments of my real level of schooling. Instrumenting my own recollection with my sibling's recollection will clean away the measurement error

Final comments on diff-and-diff:

- Identification again relies on assumptions that cannot be tested
- But at least we can show that in the past trends were parallel
- Discuss explicitly why it is a good assumption to believe that the timing of the treatment/policy was as good as random
- Discuss explicitly the existence of alternative policies that might contemporaneously affect the treatment or the control group
- Discuss the possibility that the control group is affected by the treatment.

Clustering standard errors

- Imagine that we want to estimate the impact of taking a certain pill on individual happiness. Individuals in the control group will receive a placebo
- Sample size: 1000 individuals from Kuopio and 1000 from Tampere
- How do we assign individuals to treatment and control? Two proposals
    1. Flip a coin once: tail, individuals from Kuopio are treated, heads, individuals from Tampere are treated
    2. Flip a coin 2000 times, once for each individual: tail, the individual is assigned to treatment; heads she is assigned to control
- Which of the two implementations would be more informative about the impact of the treatment? Why?

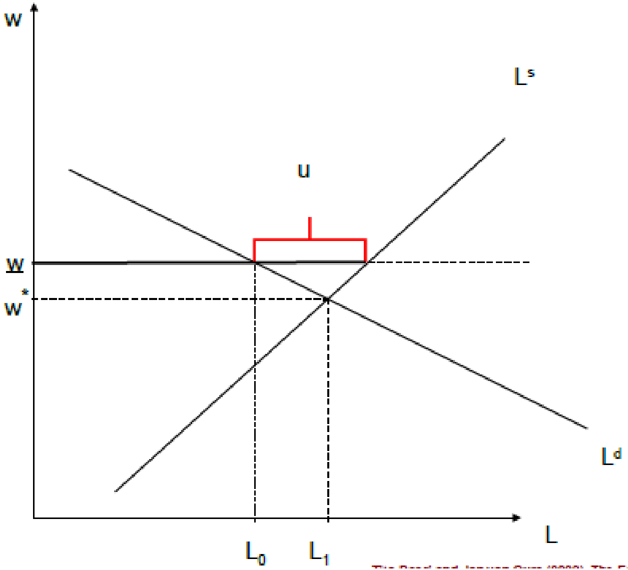
Clustering standard errors

- However, the OLS standard errors are similar in both cases. What's wrong?
- The potential presence of a common shock:
  - There might be some common shock affecting all individuals in the treatment group or in the control group (Moulton 1990).
  - OLS standard errors assume that all observations are independent realizations. Standard errors have to be corrected to account for the presence of common shocks
- In a differences-in-differences common shocks problem also may lead to serial correlation in the standard errors
- Shocks are now common to time*group cells

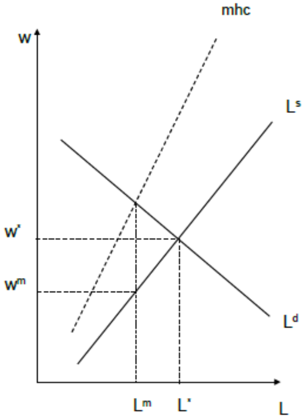
Solution

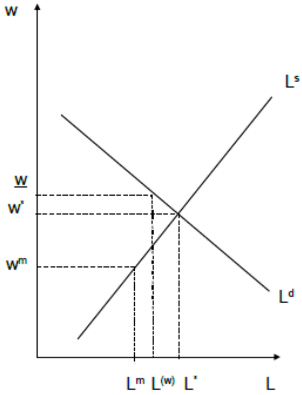
- **Cluster standard errors!**
  - In some sense, it implies acknowledging how many independent sources of information there are in the data.
- When the number of groups is large enough (rule of thumb: N>50), use the 'sandwich formula' (stata command *cluster(group)*
- When the number of groups is small, the corresponding asymptotic properties do not hold. There are some alternatives:
  - Block-bootstrap
  - Collapse the time series information into a "pre" and "post" period

Minimum wages in a competitive labour market



Minimum wages and monopsony





Example

State	Wage group	t=1	t=2
	$wg_{1,low}$	$empl_{1,low,1}$	$empl_{1,low,2}$
s_1	$wg_{1,medium}$	$empl_{1,medium,1}$	$empl_{1,medium,2}$
	$wg_{1,high}$	$empl_{1,high,1}$	$empl_{1,high,2}$
	$wg_{2,low}$	$empl_{2,low,1}$	$empl_{2,low,2}$
s_2	$wg_{2,medium}$	$empl_{2,medium,1}$	$empl_{2,medium,2}$
	$wg_{2,high}$	$empl_{2,high,1}$	$empl_{2,high,2}$

Minimum wage increases in s=1 at t=2

Increases wages in $wg_{1,low}$ at t=2,

Employment may decrease in $wg_{1,low}$ and increase in $wg_{1,medium}$

No effect: $(empl_{1,high,2}-empl_{1,high,1})-(empl_{2,high,2}-empl_{2,high,1})$

Net effect: $[(empl_{1,medium,2}-empl_{1,medium,1})-(empl_{2,medium,2}-empl_{2,medium,1})]-[(empl_{1,low,2}-empl_{1,low,1})-(empl_{2,low,2}-empl_{2,low,1})]$

Measurement error and IV

- We want to estimate
$$y = \beta s + e$$
where for convenience we assume that $Cov(s, e) = 0$
- Suppose that s is measured with error by both siblings i and j:
$$s_i = s + u$$
$$s_j = s + v$$
so that $Cov(s, u) = 0$, $Cov(s, v) = 0$, $Cov(e, u) = 0$, and $Cov(e, v) = 0$
- If we only had one measure s_i of s, we would encounter the familiar measurement error problem

Measurement error and IV

- But note that $Cov(s_j, e) = 0$ and that $Cov(s_i, s_j) = Var(s) \neq 0$
- We can use s_j as an instrument for s_i
- Instrument s_i with s_j to get:

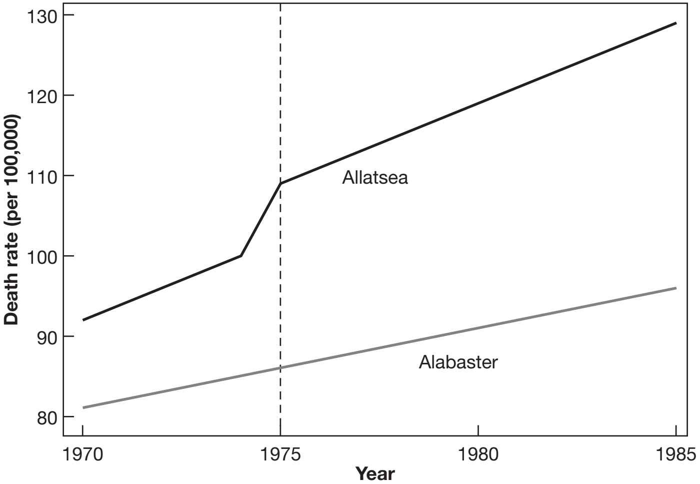
$$\frac{Cov(s_j, y)}{Cov(s_j, s_i)} = \frac{Cov(s + v, \beta s + e)}{Cov(s + u, s + v)} = \beta$$

What did we do last time?

- Analogy between dif-in-dif and fixed effects regression
- Card and Krueger
- Cengiz et al
- Question about group-specific trends: Example Autor (2003)

Dif-in-dif with non-parallel linear trends

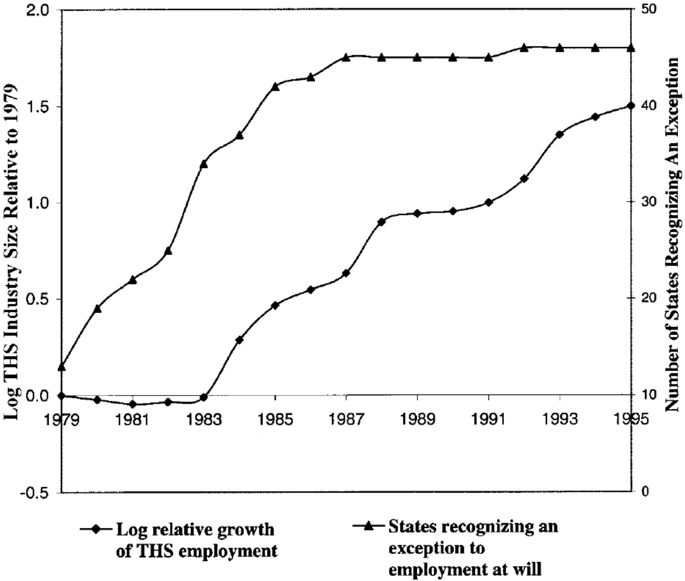
FIGURE 5.6
A real MLDA effect, visible even though trends are not parallel



Example: Autor (2003): Outsourcing at will

- Autor (2003) studies the effect of stricter employment protection (wrongful discharge laws) on outsourcing of employment to temporary help firms (THS)
- Between 1973 and 1995 45 state courts limited employers' discretion to terminate workers
- Between 1979 and 1995 employment in THS grew at 11% annually
- Are these phenomena causally linked?

Wrongful discharge laws and THS employment



Example: Autor (2003): Outsourcing at will

- Some states change employment protection and some don't
- Autor uses data on sectoral employment in U.S states over 1979-95 to regress THS employment on state and time dummies, controls and D_{st}:

  $$Y_{st} = \alpha_s + \lambda_t + \rho D_{st} + X_{st}^{'}\gamma + \epsilon_{st}$$

  where $D_{st} = 1$ if state s adopts a wrongful discharge law at t and zero otherwise
- This is a standard differences-in-differences setting

The impact of wrongful discharge laws on THS employment

Table 3
The Estimated Impact of Common Law Exceptions to Employment at Will on THS Employment, 1979–95

Exceptions Recognized	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Implied contract	.112	.136					.096	.137
	(.099)	(.063)					(.099)	(.062)
Public policy			.135	−.026			.126	−.023
			(.092)	(.060)			(.094)	(.058)
Good faith					.106	−.071	.100	−.079
					(.113)	(.095)	(.113)	(.093)
State and year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × time trends	No	Yes	No	Yes	No	Yes	No	Yes
R^2	.969	.988	.969	.988	.968	.988	.969	.988

SOURCE.—For dependent variable, see County Business Patterns, various years.
NOTE.—THS = temporary help services. Dependent variable: log state THS employment; n = 850. Ordinary least squares estimates given. Huber-White robust SEs in parentheses allow for arbitrary correlation of residuals within each state. For state common law information, see table A1.

Example: Autor (2003): Outsourcing at will

- Autor has data on many states that change legislation and on several years before and after the law change
- This makes it possible to "test" for causality in a placebo sense. Autor uses two strategies
- First, include state-specific time trends to the regression:

$$Y_{st} = \alpha_s + \gamma_s t + \lambda_t + \rho D_{st} + X_{st}^{'} \beta + \epsilon_{st}$$

  where $\gamma_s t$ is a state-specific linear trend
- Results are robust to inclusion of state-specific trends

Autor (2003): Outsourcing at will

- Second, include leads and lags of D_{st}:
  $Y_{st} = \alpha_s + \lambda_t + \sum_{\tau=0}^{m} \rho_{-\tau} D_{st-\tau} + \sum_{\tau=1}^{k} \rho_{\tau} D_{st+\tau} + X'_{ist}\beta + \epsilon_{ist}$
  where the sums allow for m postrteatment and q anticipatory effects
- If the dif-in-dif assumption holds there shouldn't be any anticipatory effects
- No evidence of anticipatory effects
- Posttreatment effects also seem to increase with time

The effect of wrongful discharge laws before, during, and after the adoption

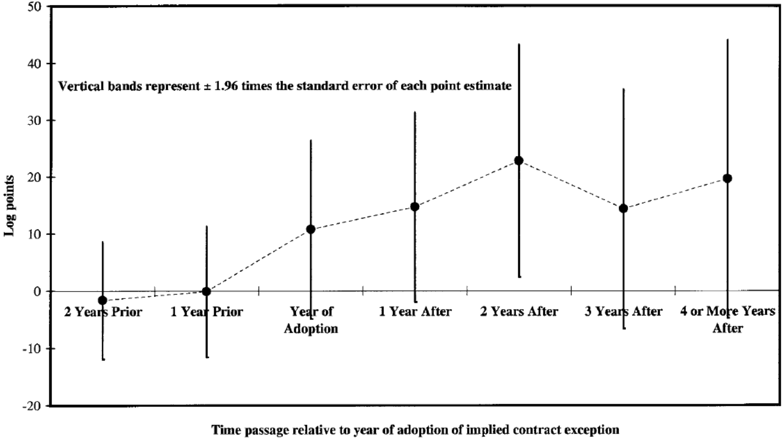


FIG. 3.—Estimated impact of implied contract exception on log state temporary help supply industry employment for years before, during, and after adoption, 1979–95.