CONDITIONING ON UNOBSERVABLES: DIFFERENCES-IN-DIFFERENCES AND INSTRUMENTAL VARIABLES

- DD Principle
- Control group in DD models
- Validity of the identifying assumption in DD models
- Statistical inference in DD models
- Instrumental variables

Introduction

- In the previous chapter, we considered settings where treatment assignment is confounded but where there exists a set of observed covariates X such that treatment assignment becomes unconfounded after conditioning on X.
- In practice however, we have the problem of possible **unobserved confounders**
- In this case, there is no universal solution but identification can be based on specific features of the data and further identifying assumptions
  - Restriction in which unobserved confounders affect the outcome of interest over time: difference in differences (this chapter)
  - Existence of an instrument: instrumental variables (this chapter)
  - Existence of thresholds: regression discontinuity (next chapter)

DD Principle

- Configuration: data available for each individual before and after the treatment
  - Is outcome before treatment a good counterfactual?
    - *i.e.* does $E(Y^{t0} | T^{t1} = 1) = E(Y_0^{t1} | T^{t1} = 1)$?
  - Problems:
    - An estimator that omits the effects of outcome dynamics (economic cycle effect, enhanced experience, etc.)
    - Bias related to anticipation effects

DD Principle

- **Idea:** combine the two naïve estimators
  - Data on the treated group and a control group; before (date t_0) and after (date t_1) treatment
  - Only one group is treated in the second period
  - Estimator:

$$E(Y^{t1} - Y^{t0}|T = 1) - E(Y^{t1} - Y^{t0}|T = 0)$$

    - The first difference removes individual effects
    - The second difference removes time effects

  - We can also write:

$$(E(Y^{t1}\,|T = 1) - E(Y^{t1}\,|T = 0)) – (E(Y^{t0}|T = 1) - E(Y^{t0}|T = 0))$$

DD Principle

- Conditions for using the method
  - When we have repeated observations for participants **and** non-participants
  - When we suspect unobserved characters of individuals are present explaining both their participation in the programme and the value of the variable of interest (independently of participation)
  - When this unobserved heterogeneity is fixed and additive

QUIZ 1

- To reduce tobacco consumption, several regions organized an information campaign in 2015 on how smoking is harmful to health. You are tasked with evaluating the impact of the campaign on effective tobacco consumption in the regions concerned.
- You have data on smoking in 2014 and 2016 in the treated regions and in other non-treated regions.
- Which of the following are difference-in-difference methods.
  - *The difference in average tobacco consumption in 2016 between treated and non-treated regions.*
  - *The change between 2014 and 2016 in average tobacco consumption in the treated regions.*
  - *The difference between treated and non-treated regions in the change in average tobacco consumption between 2014 and 2016.*
  - *The change between 2014 and 2016 in the difference in average tobacco consumption between treated and non-treated regions.*

QUIZ 2

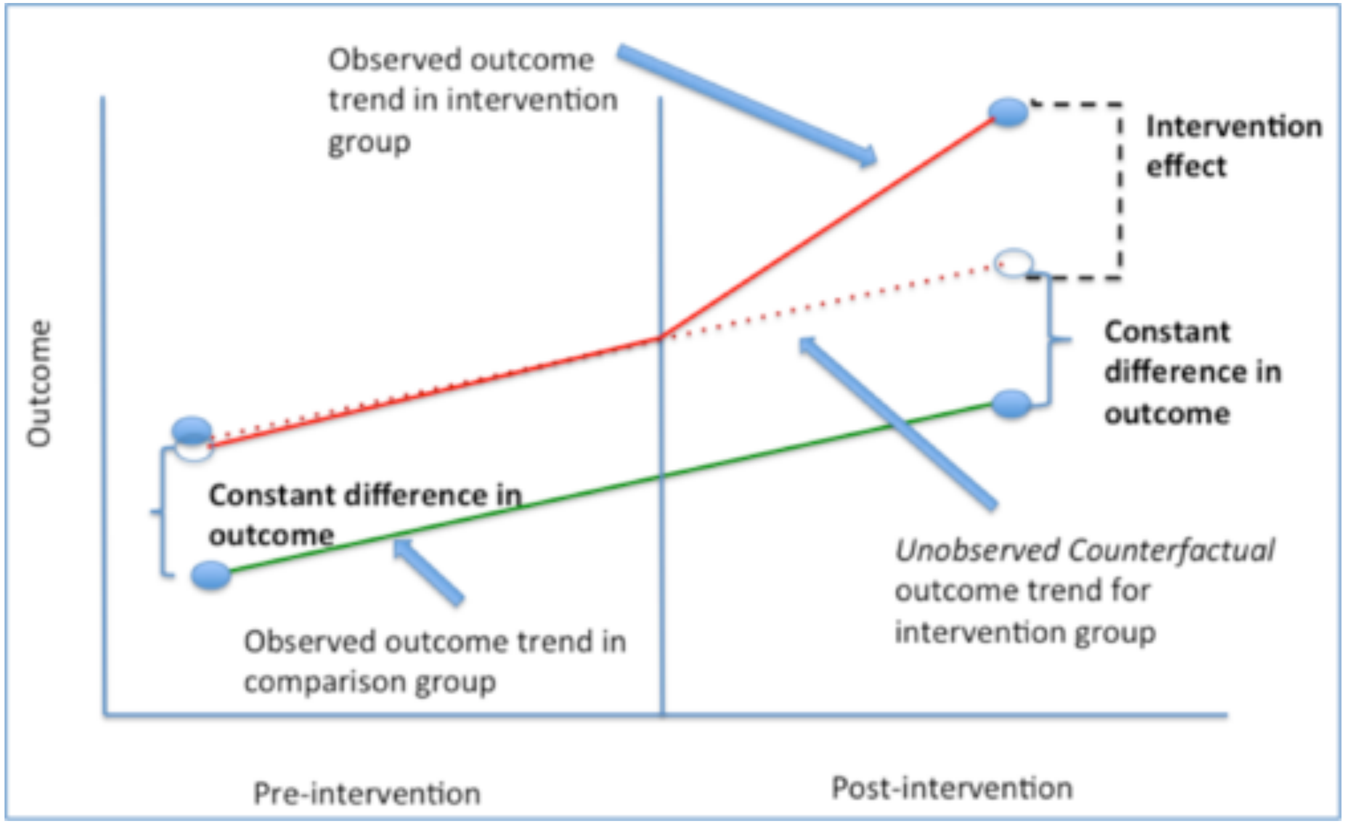
- Situation: In 2012, the government of Gondwana introduced major tax cuts on corporate profits. You are tasked with evaluating whether this tax break has increased net job creation by companies.
- You have data on net job creation by companies in the country in question (***G* = 1**), from 2012 to 2016. You also have the same data for neighbouring countries (***G* = 0**). You also know that a macroeconomic shock hit all of the countries in 2014 cutting net job creation in the same proportions everywhere.
- Can you evaluate the impact of the tax break by differences-in-differences?
  - *Yes, by comparing the change in job creation by companies in the treated country between 2012 and 2016 with that observed in the neighbouring countries.*
  - *Yes, by comparing job creation by companies in the treated country in 2016 with that observed in the neighbouring countries.*
  - *No, because you do not have data on job creation by the companies of the different countries before 2012.*
  - *No, because the macroeconomic shock prevents you from attributing the change in job creation in the treated country to the tax break.*

QUIZ 3

- Situation: To improve HIV screening the government launched a major scheme in 2006 to open free HIV testing centres in each major city in the country. The centres opened progressively between 2006 and 2010 by when all the major cities had a centre. You are tasked with evaluating the impact of these centres on the number of people detected positive.
- You have annual data from 2000 to 2015 showing the number of people detected positive for each major city. You also know the date each centre opened.
- Can you evaluate the impact of the screening centres by differences-in-differences?
  - *Yes, because the major cities are not all treated at the same time.*
  - *No, because not all the major cities are treated.*
  - *Yes, because you have data for several years before 2006 and after 2010.*
  - *No, because all the major cities are not treated at the same time.*

DD Principle

- **Identifying assumption:** in the absence of treatment, the average outcome of the treated and control groups would have changed in the same way → common trend assumption (or parallel trend assumption)



QUIZ 4

- The differences-in-differences method can be used to estimate the impact of treatment under the ‘parallel trends’ assumption.
- A first way to formulate this assumption is related to the non-observed characteristics that affect membership of the treated group and the value of the outcome.
- Under the parallel trends assumption, these characteristics should:
  - *be stable over time.*
  - *be identical on average in the two groups at the beginning of the period.*
  - *be similar from one individual to the next.*

QUIZ 5

- Situation: To reduce tobacco consumption, several regions organized an information campaign in 2013 on how smoking is harmful to health. You are tasked with evaluating the impact of the campaign on effective tobacco consumption in the regions concerned. You have data on tobacco consumption in (A) the treated regions and (B) the non-treated regions over the period 2010–2017.
- To evaluate the credibility of the parallel trends assumption, it is important to study the data on the variable of interest.
- Which of the following reduce(s) the credibility of the parallel trends assumption.
  - *Tobacco consumption is historically lower in the B regions than the A regions.*
  - *Tobacco consumption has been falling continually for 15 years in the B regions and stable in the A regions.*
  - *Tobacco consumption has risen in the B regions since 2013 but remained stable in the A regions.*

QUIZ 6

- The parallel trends assumption can also be formulated in terms of potential outcomes.
- Which of the following corresponds to the parallel trends assumption?
  - *In the absence of treatment, the mean outcome of the treated group would have the same value as in the non-treated group for each period.*
  - *In the absence of treatment, the mean outcome of the treated group would change in the same way as that of the non-treated group.*
  - *In the absence of treatment, the mean outcome of the treated group would be stable over time and like that of the non-treated group.*

QUIZ 7

- Situation: To reduce tobacco consumption, several regions organized an information campaign in 2015 on how smoking is harmful to health. You are tasked with evaluating the impact of the campaign on effective tobacco consumption in the regions concerned.
- You have data on tobacco consumption, ***Y***, in 2014 and in 2016 in the treated regions (***G = 1***) and in other non-treated regions (***G = 0***)
- You decide to use a differences-in-difference method, by estimating the impact of the information campaign in treated regions by the difference between the change in average tobacco consumption between 2014 and 2016 in the treated regions with that observed in the non-treated regions.
- Which of the following hypotheses is (are) necessary for your comparison to be valid?
  - *If the information campaign had not take place, average tobacco consumption in treated regions would be the same in 2016 as in 2014.*
  - *If the information campaign had not taken place, average tobacco consumption in the treated regions would be the same in 2016 as in the non-treated regions.*
  - *If the information campaign had not taken place, average tobacco consumption in the treated regions would have changed between 2014 and 2016 in the same way as in the non-treated regions.*

DD Principle

- **Implementation:**
  - Classical case: A is the control group and B the treatment group

$$y = \beta_0 + \beta_1 1_{t>t_0} + \beta_2 1_{i \in B} + \tau 1_{t>t_0} 1_{i \in B} + u$$

  - $1_{t>t0}$: captures conjuncture effects
  - $1_{i \in B}$: captures individual effects
  - τ : coefficient of interest

$$\hat{\tau} = (\bar{y}_{B,t_1} - \bar{y}_{B,t_0}) - (\bar{y}_{A,t_1} - \bar{y}_{A,t_0})$$

DD Principle

- **Generalization:**
  - G groups (states, age group) and T periods
  - Some but not all groups are treated at t:
    - T_{gt} = dummy for group g treated at time t
  - Implementation for estimation: set of dummies for each group, each date and one for T_{gt}

DD Principle

- **Comments**
  - There is no need for panel data…
    - Average data per group can be used
    - … but group composition must be stable
  - There can be **neither attrition nor endogenous selection** between the two time periods.
  - If additional information is available, conditioning by observables:

$$y = \beta_0 + \beta_1 1_{t>t_0} + \beta_2 1_{i\in B} + \tau 1_{t>t_0} 1_{i\in B} + X\beta + u$$

Control Group in DD models

- **Control group:**
  - The major issue is to define a credible control group
    - How can we be sure that once we allow for systematic differences by group and period there is no selection because of anticipation of unobservable change in potential income? To do this we can use a shock to the economic environment or public policy that affects population categories differently.
  - We are then dealing with ‘natural experiments’ or ‘quasi-experiments’
  - Card and Krueger (AER, 1994): effect of an increase in the minimum wage on the employment of young people

Control Group in DD models

- **Card and Krueger (1995)**
  - Increase in minimum wage from $4.25 to $5.05 in April 1992 in New Jersey (decided on in 1990)
  - What was the effect on employment of unskilled labour?
  - CK made surveys before and after in fast-food outlets in NJ and the neighbouring state of Pennsylvania:
    - Mostly unskilled jobs paid minimum wage
    - Homogeneous technology and products
    - Easy access to data (phone book), good response rate, no attrition

TABLE 1—SAMPLE DESIGN AND RESPONSE RATES

	All	Stores in: NJ	PA
Wave 1, February 15 – March 4, 1992:			
Number of stores in sample frame:[a]	473	364	109
Number of refusals:	63	33	30
Number interviewed:	410	331	79
Response rate (percentage):	86.7	90.9	72.5
Wave 2, November 5 – December 31, 1992:			
Number of stores in sample frame:	410	331	79
Number closed:	6	5	1
Number under rennovation:	2	2	0
Number temporarily closed:[b]	2	2	0
Number of refusals:	1	1	0
Number interviewed:[c]	399	321	78

[a]Stores with working phone numbers only; 29 stores in original sample frame had disconnected phone numbers.

[b]Includes one store closed because of highway construction and one store closed because of a fire.

[c]Includes 371 phone interviews and 28 personal interviews of stores that refused an initial request for a phone interview.

Différence de Différence :

Variable	Stores by state 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)	2.76 (1.36)

Control Group in DD models

- **Control group (continued):**
  - More ‘agnostic’ definition from data using a synthetic control estimator
    - Synthetic control : weighted average of untreated units chosen to reproduce characteristics of the treated unit before the intervention
    - Idea : a combination of untreated units may provide a better comparison to the treatment unit than any untreated unit alone.

QUIZ 8

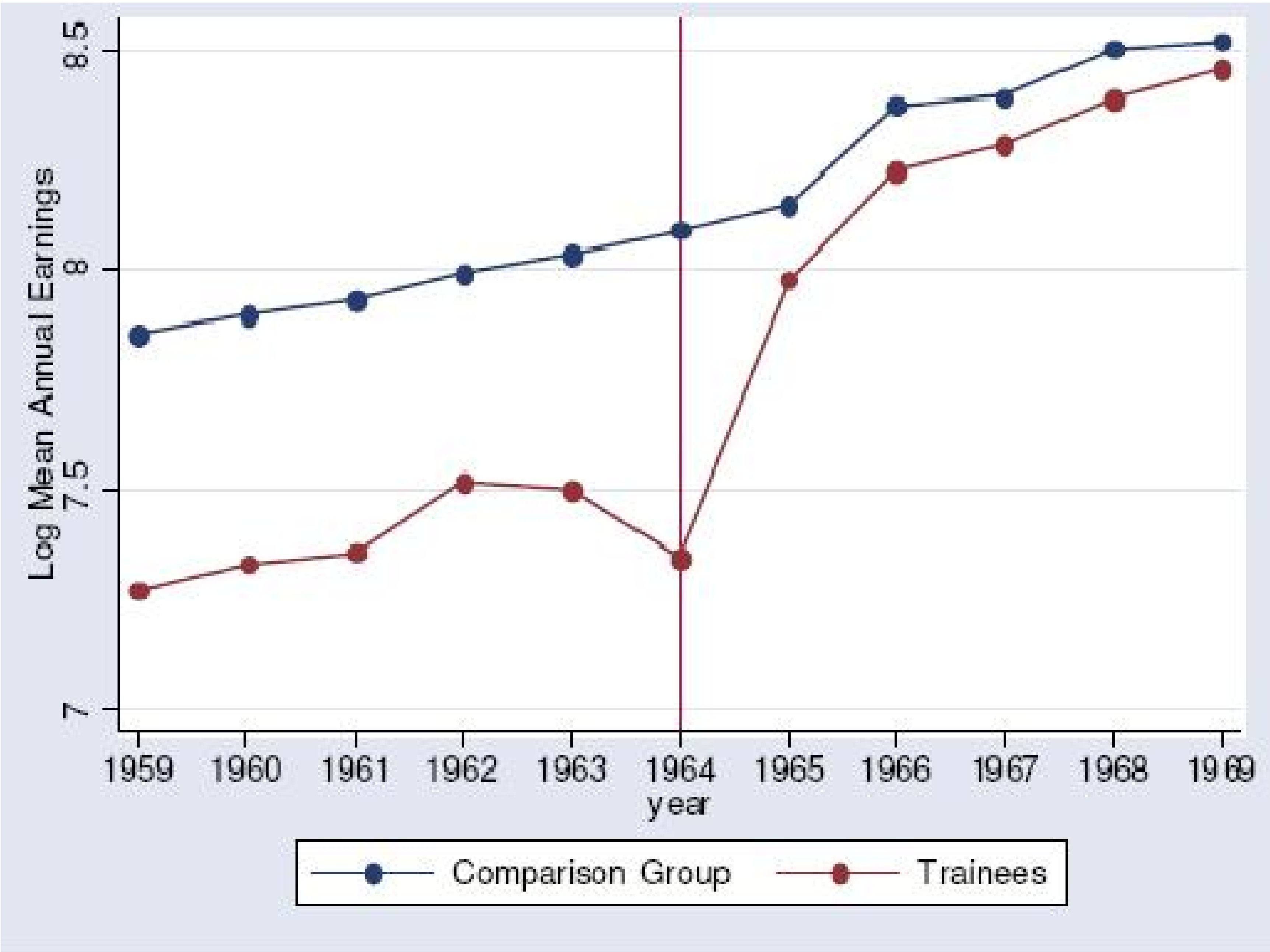
- Which of the following enable you in practice to estimate the impact of treatment by differences-in-differences?
  - *Find just the difference in means of* ***Y*** *at each date and for each group (treated and non-treated)*
  - *A linear regression of* ***Y*** *on indicative variables covering each date of observation and indicating the position with respect to treatment (treated and non-treated).*

QUIZ 9

- With the differences-in-differences method, what can be done with linear regression that cannot be done with a straightforward comparison of means of ***Y*** at each date and in each group?
  - *Easily calculate the standard deviation of the estimator*
  - *Conduct the analysis on more than two years*
  - *Control for the effect of other factors*
  - *Obtain an unbiased estimator of the impact*

Validity of the Identifying Assumption in DD models

- The validity of the DD estimator is based on a 'common trend' assumption: the income trend is the same for treated and non-treated groups
- This assumption is never testable
  - Otherwise it would mean we could observe the counterfactual change in the treated group in the absence of treatment
- With more than two groups and/or two periods we can, however, get some idea of whether it is plausible:
  - When several periods are available, we can check the changes are the same in the different groups
  - 'Ashenfelter's dip': beware of catch-up phenomena!

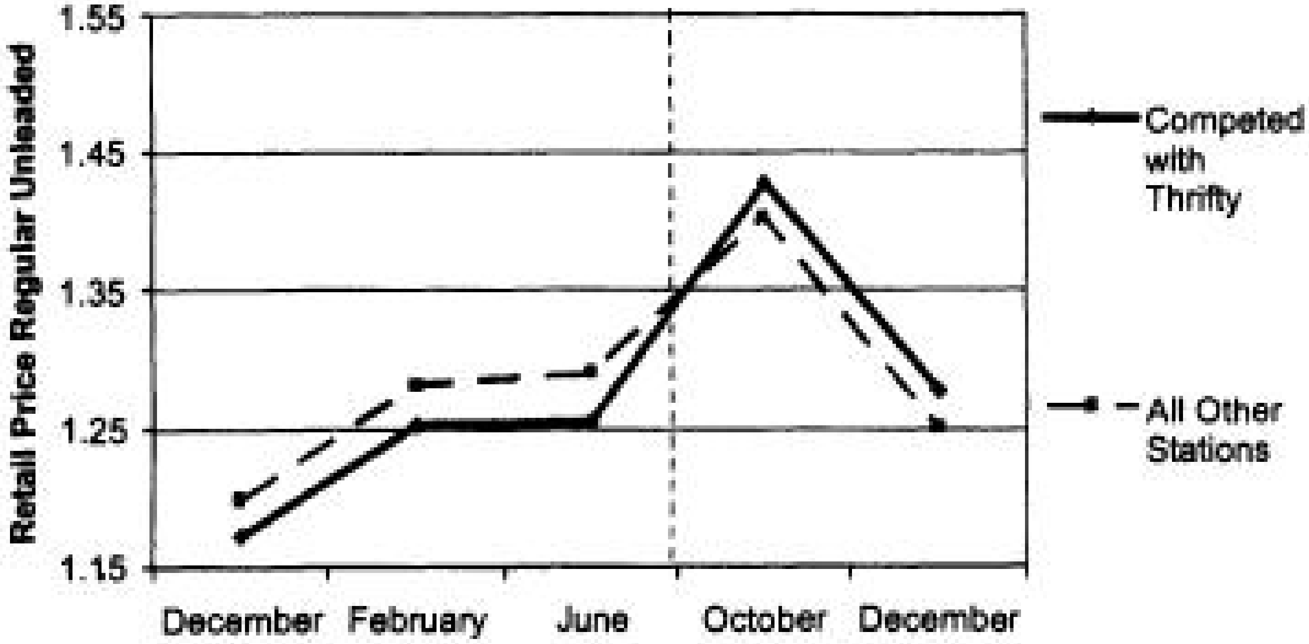


Log Mean Annual Earnings
8.5
8
7.5
7
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
year
Comparison Group
Trainees

Validity of the Identifying Assumption in DD models

- **Example:** Hastings, AER (2004) ‘Vertical relationship and competition in retail gasoline markets’
  - Higher prices at the pump concomitant with a wave of vertical concentration (buy-out of distributor by producer to reduce competition)?
  - Necessity to legislate on concentration to ensure competition?
  - Use of natural experiment: ARCO’s buy-out of an independent distributor ‘Thrifty’
  - Control group: nearby (1 mile) stations in competition with Thrifty stations

Evolution of gasoline price according to the distance from a "Thrifty" station:



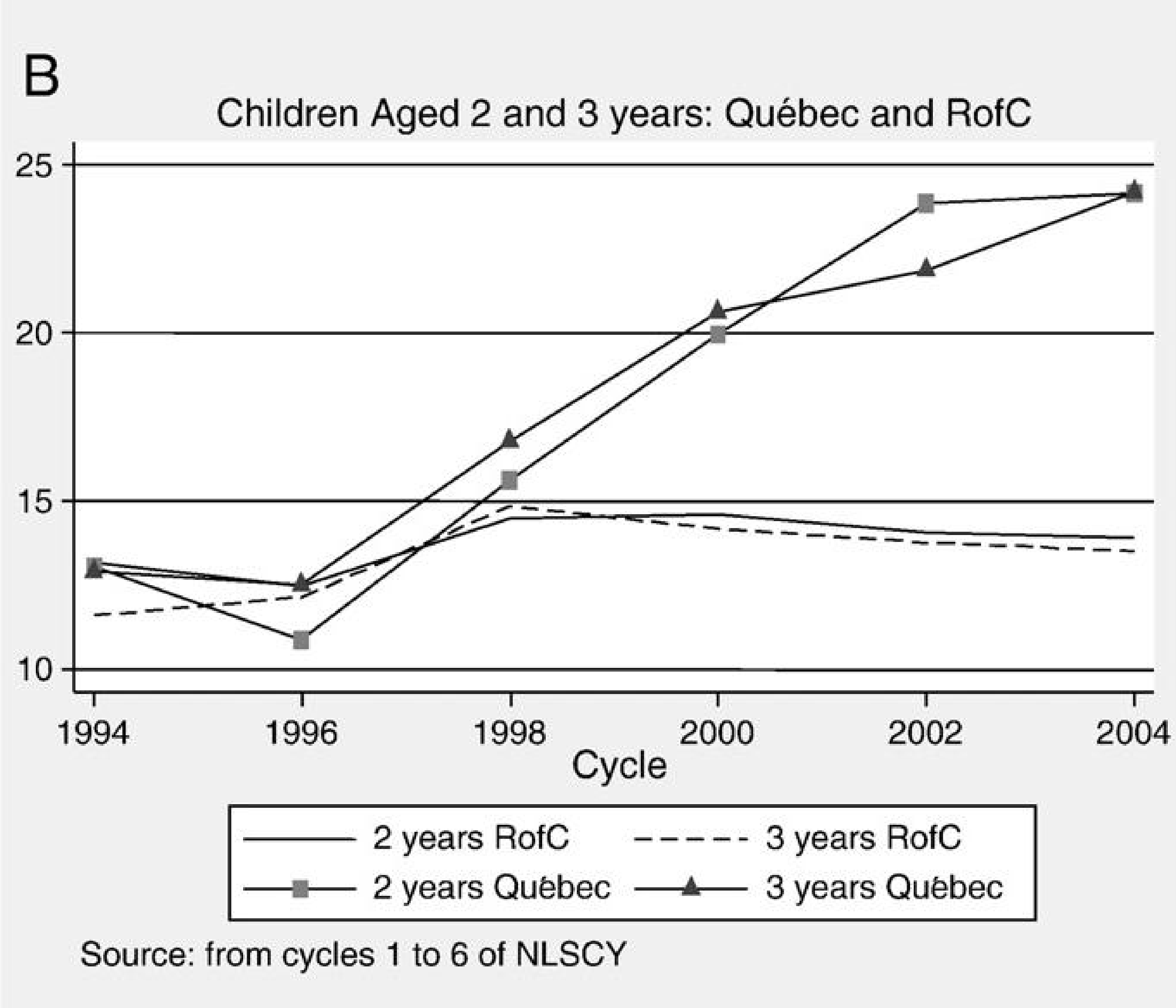
Validity of the Identifying Assumption in DD models

- **Indirect tests**
  - Placebo (Angrist Krueger 1999)
    - When we apply the same method at a date when nothing happened, we should find a non-significant effect
  - Use of another control group (e.g. based on a different criterion) Card & Krueger: comparison with high wages in NJ
  - May lead to differences-in-differences-in-differences…

Validity of the Identifying Assumption in DD models

- **Differences-in-differences-in-differences**
  - Example: Lefèvre et al. (2009) ‘Dynamic labour supply effects of childcare subsidies: Evidence from a Canadian natural experiment on low-fee universal child care’, Labor Economics
  - Evaluation of long-run on the income of working mothers further to the introduction in 1997 of a programme for building spaces for children under 4 in Quebec
    - The $5/day/child subsidy positively changes the life cycle profile for mothers of young children for all the years the child is of pre-school age
    - If the number of years (experience) in work rises, the policy may have positive effects on earnings later in the life cycle, which may further increase the labour supply
  - No similar programme in the other provinces of Canada

Mean hours per week in primary care by age



Validity of the Identifying Assumption in DD models

- **Differences-in-difference-in-differences (continued)**
  - Treatment group: mothers in Quebec with at least one child aged 6 to 11 years and no children under 6 years old
  - Control groups:
    - Mothers in rest of Canada with at least one child aged 6 to 11 years and no children under 6 years old… but the provinces may be affected differently by business cycle shocks, and policy effects may be confused
    - Mothers in Quebec with at least one child aged 12 to 17 years and no children under 12 years old… but they may have different outcomes on the labour market

Validity of the Identifying Assumption in DD models

- **Differences-in-difference-in-differences (continued)**
  - Solution: comparison of the difference on the change in outcomes on the labour market of mothers with a child aged 6 to 11 years and those with a child aged 12 to 17 years in Quebec with the same difference for other mothers in rest of Canada.
  - Formally:

$$
\begin{aligned}
Y_{it} = \quad & \beta_0 + \beta_1 Q_{it} + \beta_2 I_{t \geq 1997} + \beta_3 C_{it} + \beta_4 Q_{it} I_{t \geq 1997} + \beta_5 Q_{it} C_{it} \\
& + \beta_6 C_{it} I_{t \geq 1997} + \delta Q_{it} C_{it} I_{t \geq 1997} + e_{it}
\end{aligned}
$$

$Q_i t$ = 1 if Quebec, C_{it} = 1 if child aged 6 to 11 years.

Parameter of interest: δ

QUIZ 10

- Situation : To reduce tobacco consumption, several regions organized an information campaign in 2013 on how smoking is harmful to health. You are tasked with evaluating the impact of the campaign on effective tobacco consumption in the regions concerned. You have data on tobacco consumption in (A) the treated regions and (B) the non-treated regions over the period 2010–2017.
- You observe that the change in tobacco consumption is very similar in regions A and B between 2010 and 2013. By studying the institutional context, you learn that in 2013 the A regions also introduced subsidies to reduce the price of treatment for nicotine addiction. Can you evaluate the impact of the information campaign by differences-in-differences?

  - *Yes, because the subsidies were only introduced in 2013.*
  - *Yes, because the change in tobacco consumption was very similar in the A and B regions*
  - *No, because the change in tobacco consumption after 2013 in the A and B region would have diverged even in the absence of the information campaign between 2010 and 2013.*

QUIZ 11

- Situation: In 2012, the government introduced major tax cuts on corporate profits. You are tasked with evaluating by differences-in-differences whether this tax break has increased net job creation by companies. To do so you have data on net job creation by companies in the country from 2011 to 2013 and by companies in neighbouring countries over the same period.
- In 2011 an economic shock caused job creation to fall in the country in question, but only temporarily. Does that entail a problem for evaluating the impact of the tax break by differences-in-differences?
  - *Yes, there is a risk of overestimating the impact of the tax break.*
  - *Yes, there is a risk of underestimating the impact of the tax break.*
  - *No, because the shock was temporary.*

QUIZ 12

- Situation: In 2012, the government introduced major tax cuts on corporate profits. You are tasked with evaluating by differences-in-differences whether this tax break has increased net job creation by companies. To do so you have data on net job creation by companies in the country from 2011 to 2013 and by companies in neighbouring countries over the same period.
- Which of the following proposals for improving the data available are most likely to improve the credibility of the parallel trends assumption?
  - *More years of observations before 2011*
  - *More years of observations after 2013*
  - *Data on a larger number of countries*
  - *Data on the economic characteristics of the countries and companies.*

Statistical inference in DD models

- **Problem**
  - Classical inference: we attempt to control for bias arising from observing a sample from a population
  - In the case of DD estimators
    - At least one additional problem: quality of the control group
    - → More problems on the variance-covariance matrix

Statistical inference in DD models

- **Consequences:**
  - Bertrand, Duflo and Mullainathan (QJE, 2004): ‘How much should we trust in DD estimators?’
    - The assumption of an identical change between groups *a priori* more credible for shorter periods
    - However, policy impact is often interesting in the medium or even long term
    - In practice, estimation over fairly long time spans
    - In this case, the assumption of no group*period cross effects is unlikely
    - Ignoring it biases inference

Statistical inference in DD models

- **BDM**
  - Model:

$$y_{igt} = e_t + e_g + \delta D_{gt} + X_{igt}\beta + u_{igt}$$

e_t and e_e : time/period fixed effects

D_{gt} : indicator group g was treated at t

  - Two sources of problems
    - Individual data: generally shocks common to a group → group heteroscedasticity
    - Temporal autocorrelation

Statistical inference in DD models

- **Placebo laws generated**
  - Data: panel CPS (wage, long-term employment in US)
  - 200 random draws
    - One year
    - Half of states treated
  - DD method is applied to evaluate the impact of these virtual laws
    - Real effect = 0
    - Classical test of significance at 5%: Student's comparison at 1.96. By definition, we should have 5% erroneous rejection of null hypothesis

Statistical inference in DD models

- **Outcome of simulation**
  - BDM rejects null hypothesis in two out of three cases!
  - Aggregation of data at state level (simultaneous shocks): again one-in-two chance of being mistaken
  - Correction of autocorrelation by *AR(k)*: little improvement (not enough points to properly estimate parameters)
  - Solutions proposed:
    - Aggregate data before/after reform (if single date)
    - Use White's variance-covariance estimator if enough groups
    - Inference by block-bootstrapping (re-sampling method)

TABLE I
Survey of DD Papers[a]

Number of DD papers	92	
Number with more than 2 periods of data	69	
Number which collapse data into before-after	4	
Number with potential serial correlation problem	65	
Number with some serial correlation correction	5	
GLS	4	
Arbitrary variance-covariance matrix	1	
Distribution of time span for papers with more than 2 periods	**Average**	16.5
Informal techniques used to assess endogeneity	**Number**	
Graph dynamics of effect	15	
See if effect is persistent	2	
DDD	11	
Include time trend specific to treated states	7	
Look for effect prior to intervention	3	
Include lagged dependent variable	3	
Number with potential clustering problem	80	
Number which deal with it	36	

Example

- Bauer T.K., Braun S.T., Kvasnicka M. (2017) Nuclear power plant closures and local housing values: Evidence from Fukushima and the German housing market, *Journal of Urban Economics*, 99, 94-106
- Problem
  - What effects of proximity to a negative amenity, here nuclear power plants on property prices?
  - Identification from shocks
    - Fukushima nuclear accident of 11 March 2011
    - Closure of 17 nuclear power plants in Germany in 2011 and announced closure of remaining nine by 2022

Example

Our main identifying assumption for a causal interpretation of our results is that conditional on control variables, which include a large set of individual house characteristics, housing prices in treatment and control regions would have followed the same trend in the absence of Fukushima. We corroborate this identifying assumption in various ways. For example, we show that pre-Fukushima trends in prices did not differ statistically between houses close to and further away from a NPP site. We also show that our results do not change when we restrict the estimation sample to a more homogeneous set of regions (e.g., by excluding house offers from urban districts).

Example

- Model
  - Outcome: Housing price
  - Treatment: dummy variable indicating location near nuclear power plant

$$Y_{ijt} = X_{ijt}\beta + \gamma NPP_{ij} + \delta(NPP_{ij} \times Fukushima_t) + D_j + D_t + \epsilon_{ijt}$$

  - Y_{ijt} : asking price for housing *i* in region *j* in month *t*
  - NPP_i *:* dummy for location near nuclear plant (5 km)
  - $Fukushima_t$ *:* dummy taking value 1 if *t* > date of accident
  - X_{ijt} are the housing characteristics
  - D_j and D_t are regional and time dummies
  - Ordinary least squares estimation with robust statistical inference

Example

Table 1
Summary statistics by distance to NPP site, before and after Fukushima.

	Before Fukushima					After Fukushima					DiD
	Distance to next NPP < 5 km			≥ 5 km	≥ 25 km	Distance to next NPP < 5 km			≥ 5 km	≥ 25 km	
	all	operating	non-operating			all	operating	non-operating			[(6) – (1)]– [(9) – (4)]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Price (Euro)	217, 586	233, 615	178, 196	264, 692	264, 597	226, 222	232, 414	206, 790	268, 061	267, 843	5267
	[103.320]	[110.295]	[69.515]	[193.186]	[196.315]	[115.604]	[118.540]	[103.516]	[220.801]	[223.999]	(8498)
Log Price (Euro)	12.188	12.261	12.009	12.316	12.312	12.199	12.231	12.101	12.282	12.277	0.045
	[0.462]	[0.459]	[0.421]	[0.571]	[0.576]	[0.533]	[0.520]	[0.560]	[0.652]	[0.658]	(0.033)
Age (years)	33.024	31.163	37.598	29.215	29.158	40.976	40.256	43.232	36.790	36.745	0.377
	[30.621]	[31.293]	[28.392]	[29.990]	[29.962]	[32.373]	[32.916]	[30.504]	[32.563]	[32.557]	(1.316)
In construction (%)	3.615	4.516	1.399	5.085	5.040	2.280	2.952	0.173	3.903	3.886	–0.152
	[18.666]	[20.767]	[11.747]	[21.968]	[21.876]	[14.929]	[16.927]	[4.152]	[19.366]	[19.327]	(1.315)
Living space (m^2)	146.252	150.314	136.270	152.747	152.653	152.595	155.123	144.662	156.489	156.358	2.602
	[50.392]	[51.288]	[46.645]	[55.572]	[55.845]	[52.723]	[52.724]	[51.942]	[59.350]	[59.524]	(3.531)
Base area (m^2)	700.898	658.047	806.202	729.524	739.245	740.618	680.371	929.679	810.401	822.010	–41.157
	[682.403]	[654.769]	[735.526]	[776.146]	[784.590]	[752.178]	[618.603]	[1045.826]	[871.125]	[881.044]	(36.443)
Detached house (%)	27.846	27.724	28.148	33.367	33.109	25.320	26.109	22.842	29.289	29.147	1.551
	[44.826]	[44.766]	[44.977]	[47.152]	[47.060]	[43.487]	[43.927]	[41.994]	[45.5091]	[45.444]	(2.688)
Observations	14,580	10,363	4217	4,546,480	4,217,770	7192	5454	1738	1,910,446	1,768,295	6,478,698

Note: Columns (1)–(4) and (6)–(9) show the mean of each variable for property located within or outside a 5 km range of a NPP site. Among the property within the 5 km range, we further distinguish between property near sites that were operating before Fukushima and those that were not (NPPs Brunsbüttel, Krümmel). Columns (5) and (10) show the mean of each variable for the control group of property in our baseline specification from column (4) of Table 2. Averages are calculated for the pre-Fukushima period March 2007 to February 2011 and for the post-Fukushima period April 2011 to March 2013. Standard deviations are reported in square brackets (columns (1)–(10)). Standard errors clustered at the level of regional planning units (*Raumordnungsregionen*) are reported in round brackets (column (11)).

Example

Table 2
Main regression results.

	(1)	(2)	(3)	(4)
Treatment effect:				
NPP$<$ 5 km $\times$ Post-Fukushima	-0.025 (0.032)	-0.032^{*} (0.017)		
operat. NPP$<$ 5 km $\times$ Post-Fukushima			-0.048^{***} (0.016)	-0.049^{***} (0.016)
non-operat. NPP$<$ 5 km $\times$ Post-Fukushima			0.014 (0.031)	0.014 (0.031)
5 km $\leq$ operat. NPP $<$ 10 km $\times$ Post-Fukushima				-0.019 (0.015)
10 km $\leq$ operat. NPP $<$ 15 km $\times$ Post-Fukushima				-0.011 (0.016)
15 km $\leq$ operat. NPP $<$ 20 km $\times$ Post-Fukushima				-0.009 (0.012)
20 km $\leq$ operat. NPP $<$ 25 km $\times$ Post-Fukushima				-0.009 (0.008)
Month dummies	yes	yes	yes	yes
Zip code fixed effects	yes	yes	yes	yes
Property characteristics	no	yes	yes	yes

NOTES: The endogenous variable is the log of the nominal house price posted. All regressions include a dummy for the post-Fukushima period. Regressions in columns (1) and (2) include a NPP dummy that indicates whether a house on offer is located within 5 km from a NPP site and an interaction of this indicator with the dummy for the post-Fukushima period. In columns (3) and (4), we instead add separate NPP dummies for houses near NPPs that were operating and non-operating right before the Fukushima accident, along with their respective interactions with the dummy for the post-Fukushima period. In column (4), we add further NPP dummies which indicate whether a house on offer is located between 5 and 10 km, 10 and 15 km, 15 and 20 km or 20 and 25 km from a NPP site that was operating right before the Fukushima accident and the interactions of these additional distance measures with the post-Fukushima dummy. Property characteristics include age (and its square), a dummy for property still under construction, living space (and its square), base area (and its square), and a dummy for detached houses. The estimation sample comprises sales offers for single-unit detached and terraced houses posted on the internet platform ImmobilienScout24 in the months March 2007 to March 2013 (March 2011 offers are excluded). The sample size is 6,478,698 (offer $\times$ month observations). The number of zip code fixed effects is 8116. ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Standard errors are clustered at the level of 96 regional planning units (*Raumordnungsregionen*).

Instrumental variables

- Widely used method
  - In RCT : when there is imperfect compliance
  - With observational data: when one variable is endogenous due to measurement error, omitted variable bias, simultaneity bias.
  - The instrumental variable method relies on the availability of valid instruments
    - In our case, variables that affect the treatment intake but do not directly affect potential outcomes

QUIZ 13

- What conditions must an instrument satisfy to be valid?
  - *It should not have an impact on treatment T*
  - *It should not have an impact on outcome Y*
  - *It should have an impact on outcome Y only via treatment T*
  - *It should not have an impact on other variables other than benefiting from treatment T*

QUIZ 14

- You must evaluate the impact of the number of children *F* on women's earnings *E.* Some studies have shown that couples with two children have a greater probability of having a third children when the first two children are of the same sex *Z* = 1*.*
- The second condition for variable *Z* can be used for instrument to evaluate the impact of *F* on *E* is the exogeneity condition. What does it mean here?
  - Having 2 children of the same sex must have an impact on the number of children
  - Having 2 children of the same sex must have an impact on women's earnings
  - Having 2 children of the same sex must have an effect on women's earnings only via its effect on the number of children
  - Having 2 children of the same sex must be a quasi-random phenomenon and not linked to unobserved characteristics that influence women's earnings.

Instrumental variables

- In RCT with imperfect compliance, the instruments are usually provided by randomized assignment to treatment
  - Z = 1 if assigned to treatment and 0 otherwise
  - The **observed** treatment status is $T = ZT_1 + (1 - Z)T_0$
  - When the assignment is perfect, then $\Pr(T_1 = 1) = \Pr(T_0 = 0) = 1$ and $\Pr(T = Z) = 1$
  - When we have noncompliance, there are 4 cases:
    - **Compliers**: take the treatment only if are assigned to the treatment group $T_1 > T_0$ or ($T_0 = 0$ and $T_1 = 0$)
    - **Always-takers**: always take the treatment if they are assigned to the control group $T_1 = T_0 = 1$
    - **Never-takers**: never take the treatment if they are assigned to the treatment group $T_1 = T_0 = 0$
    - **Defiers**: do the opposite of their treatment status $T_1 < T_0$ or ($T_0 = 1$ and $T_1 = 1$)

QUIZ 15

- You must evaluate the impact of going to a private school on the success to the “baccalauréat”. Pupils are not randomly allocated between public and private schools. Therefore you use a random experiment as an instrument: some families randomly drawn benefit of a reduced cost to enrol in a private school while the families not drawn must pay the normal price.
- What is a complier in this situation?
  - A pupil in a private school if her family has benefited from the reduced cost and in public school if not
  - A pupil that succeeds to the baccalauréat is her family has benefited from the reduced cost and fails otherwise
  - A pupil that succeeds to the baccalauréat if she is in a private school but fails if she is in a public school
  - A pupil enrolled in a private school if she fails and in a public school if she succeeds

Instrumental variables

- Identification and estimation in RCT with imperfect compliance
  - We assume that treatment assignment is not trivial (eg not everybody is assigned to treatment or to control) and that the assignment has an effect on treatment intake:

$$0 < \Pr(Z = 1) < 1 \text{ and } \Pr(T_1 = 1) \neq \Pr(T_0 = 1)$$

  - The identification assumption is the exclusion restriction:

$$(Y_1, Y_0, T_1, T_0) \perp Z$$

  - The IV parameter can the be recovered as:

$$\tau_{IV} = \frac{cov(Y,Z)}{cov(T,Z)} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(T|Z=1) - E(T|Z=0)}$$

  - If the assumptions hold, then

$$\tau_{IV} = \frac{E[(Y_1 - Y_0)(W_1 - W_0)]}{E[W_1 - W_0]}$$

Instrumental variables

- Identification and estimation in RCT with imperfect compliance
  - The sample analog is:

$$\tau_{IV} = \frac{\left(\sum_{i=1}^{n} Z_i Y_i / \sum_{i=1}^{n} Z_i\right) - \left(\sum_{i=1}^{n} (1-Z_i) Y_i / \sum_{i=1}^{n} (1-Z_i)\right)}{\left(\sum_{i=1}^{n} Z_i T_i / \sum_{i=1}^{n} Z_i\right) - \left(\sum_{i=1}^{n} (1-Z_i) T_i / \sum_{i=1}^{n} (1-Z_i)\right)}$$

  Note that this estimator may fail to identify the ATE in case of treatment heterogeneity

Instrumental variables

- Imbens and Angrist (1994) define the local average treatment effect (LATE) as:

$$\tau_{LATE} = E(Y_1 - Y_0 | T_1 > T_0)$$

ie the average effect of the treatment for compliers, ie the average treatment effect for the units that change their treatment status according to their treatment assignment.

In general τ_{LATE} is not the same as ATE or ATT

- Now, such interpretations of IV appear even outside the context of RCT

Example

- Angrist and Imbens (1996) study the impact of being a Vietnam veteran on mortality once gone back to civilian life
  - Does veteran have different mortality rate than the rest of the population?
  - Problem: there might be self selection
  - Who goes to war?
    - Always takers: always go to war
    - Never takers: cautious or pacifists that will do everything possible not to go
    - Compliers: go to war if they are called but don't go otherwise
  - What is the link with mortality?
    - Always takers take more risk on average, and hence might have a higher mortality rate even if gone back to civilian life
    - Cautious people are… cautious and therefore have a lesser mortality rate

Example

- Identification issue:
  - If only always takers go to war and only never takers don’t go, it is not possible to evaluate the impact of veteran status on mortality as in this case, this impact is confounding with attitude toward risk.
  - To identify the impact we need to have compliers, ie individuals the participation to war of which is exogenous
  - Instrument: draft lottery. A number corresponding to the day of birth was used as assignment to enrol or not
  - Estimation with 2SLS