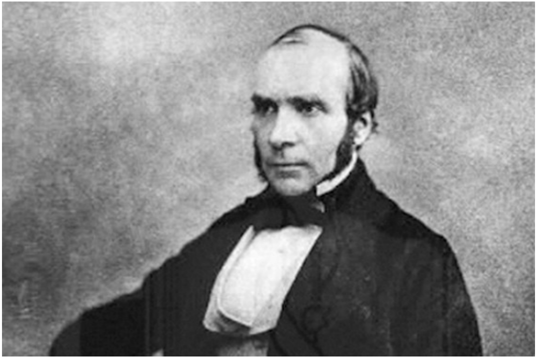
Quantitative Policy Evaluation
Difference in Differences.

Eduardo Fé

Snow on Cholera

Meet John Snow, and anaesthesiologist from 19th century London and a pioneer of **Natural Experiments**...



Snow on Cholera

Cholera epidemics were devastating London

At the time, cholera was thought to be a ‘bad air’ (miasma)

But Snow believed that cholera was a ‘waste’ (water borne disease)

because he observed that

- Epidemics seemed to follow the ‘great track of human intercourse’...
- And that sailors did not become infected until they disembarked (which would provide evidence against the miasma hypothesis)
- He collected data on where deaths occurred, and he mapped these data, observing that deaths were concentrated around the Broad Street water pump...

Snow on Cholera

Snow on Cholera

His observations shaped his views, **but** the strongest evidence to support his hypothesis came when studying the 1853-54 cholera outbreak:

1

Snow on Cholera

His observations shaped his views, **but** the strongest evidence to support his hypothesis came when studying the 1853-54 cholera outbreak:

- Large areas of London were served by two water companies (Lambeth and Southward & Vauxhall).

Snow on Cholera

His observations shaped his views, **but** the strongest evidence to support his hypothesis came when studying the 1853-54 cholera outbreak:

- Large areas of London were served by two water companies (Lambeth and Southward & Vauxhall).
- In 1852 the Lambeth moved its intake pipe further upstream on the Thames thereby "*obtaining a supply of water quite free from the sewage of London*" while Southwark and Vauxhall left its intake pipe in place.

1

Snow on Cholera

His observations shaped his views, **but** the strongest evidence to support his hypothesis came when studying the 1853-54 cholera outbreak:

- Large areas of London were served by two water companies (Lambeth and Southward & Vauxhall).
- In 1852 the Lambeth moved its intake pipe further upstream on the Thames thereby "*obtaining a supply of water quite free from the sewage of London*" while Southwark and Vauxhall left its intake pipe in place.
- Snow recorded cholera deaths throughout London together with information about the company supplying water to each household.

1

Snow on Cholera

His observations shaped his views, **but** the strongest evidence to support his hypothesis came when studying the 1853-54 cholera outbreak:

- Large areas of London were served by two water companies (Lambeth and Southward & Vauxhall).
- In 1852 the Lambeth moved its intake pipe further upstream on the Thames thereby "*obtaining a supply of water quite free from the sewage of London*" while Southwark and Vauxhall left its intake pipe in place.
- Snow recorded cholera deaths throughout London together with information about the company supplying water to each household.
- Then he compared cholera death rates by source of water:

1

Snow on Cholera

His observations shaped his views, **but** the strongest evidence to support his hypothesis came when studying the 1853-54 cholera outbreak:

- Large areas of London were served by two water companies (Lambeth and Southward & Vauxhall).
- In 1852 the Lambeth moved its intake pipe further upstream on the Thames thereby "*obtaining a supply of water quite free from the sewage of London*" while Southwark and Vauxhall left its intake pipe in place.
- Snow recorded cholera deaths throughout London together with information about the company supplying water to each household.
- Then he compared cholera death rates by source of water:

Results[1]:

Southwark & Vauxhall	315
Lambeth	37
Rest of London	56

[1] Deaths per 10,000 population

Snow on Cholera

Why were these results credible? Because those affected by the move and those who were not affected were comparable...

- *The pipes of each company go... into nearly all the courts and alleys*
- *[...]Each company supplies both rich and poor, both large houses and small; there is no difference either in the condition or occupation of the persons receiving the water from different companies*

... and critically,

- The move of the Lambeth company implied that more than 300,000 people were divided into two groups without their choice or even their knowledge.
- One group received sewage from London, the other clean water free from impurity.

Snow on Cholera

In other words, the move by the Lambeth created

- Two groups
- Comparable in all respects
- And which group received the ‘treatment’ (clean water) was allocated as if ‘random’ and independently of individuals’ traits.

Almost as in a randomized experiment. The difference is that:

- In randomized experiments it is a researcher who ultimately controls who gets what
- In the natural experiment the underlying sample is more representative of the full population.

Difference-in-Differences (DiD)

Fast forward to 1994, and[2] Card and Krueger (1994).

Is economic theory right? Do higher minimum wages lead to cuts of employment?
In 1992, the state of New Jersey increased the minimum wage from \$4.25 to \$5.05 per hour

This presented an opportunity to evaluate this question. Specifically, they focus on the **fast-food** industry.

- Leading employer of low-wage workers
- They comply with minimum wage regulations
- Their products are pretty homogeneous
- From a practical point of view, the same chains are available across the US...

[2] A controversial paper. See also Neumark and Wascher (2000), Card and Krueger (2000).

Difference-in-Differences (DiD)

… but where do we find a suitable comparison group for the whole state of New Jersey?

… they chose fast-food restaurants located in neighbouring eastern Pennsylvania, where minimum wage stayed invariant at $4.25.

Difference-in-Differences (DiD)

They constructed a sample of fast food restaurants; the first interviews were done in March 1992 (one month before the scheduled increase in minimum wage) and in November-December 1992, about eight months after the minimum-wage increase.

The sample include 410 responses in the first wave; of these 99.8% provided follow up information in the second wave.

Difference-in-Differences (DiD)

TABLE 2—MEANS OF KEY VARIABLES

Variable	Stores in: NJ	Stores in: PA	t^a
1. *Distribution of Store Types (percentages):*			
a. Burger King	41.1	44.3	−0.5
b. KFC	20.5	15.2	1.2
c. Roy Rogers	24.8	21.5	0.6
d. Wendy's	13.6	19.0	−1.1
e. Company-owned	34.1	35.4	−0.2

Difference significant at 5% if $|t| > 1.96$

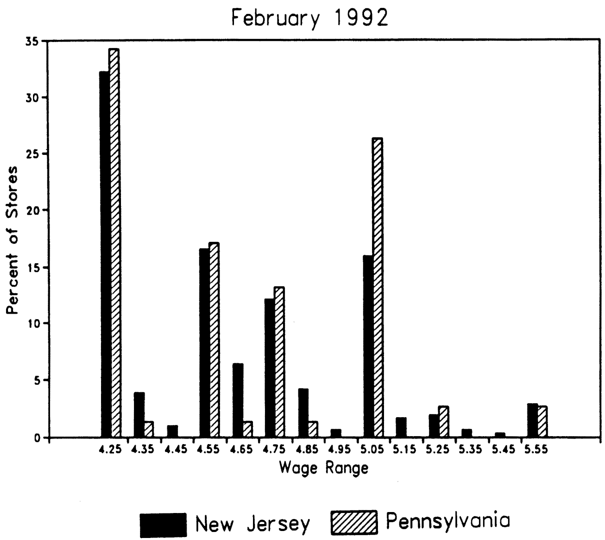
Difference-in-Differences (DiD)

TABLE 2—MEANS OF KEY VARIABLES

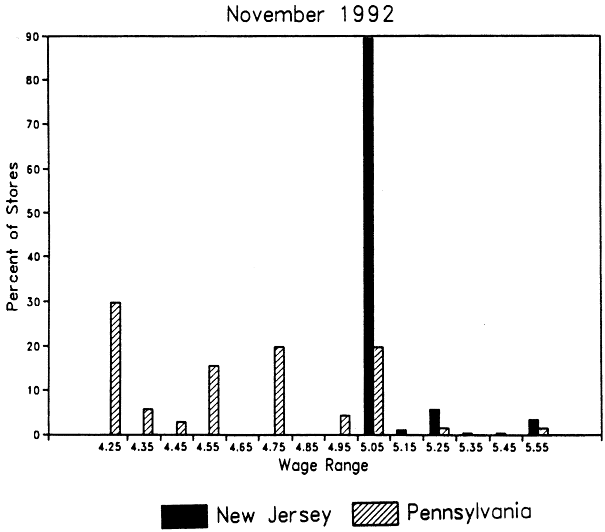
Variable	Stores in: NJ	Stores in: PA	t^a
2. *Means in Wave 1:*			
a. FTE employment	20.4 (0.51)	23.3 (1.35)	−2.0
b. Percentage full-time employees	32.8 (1.3)	35.0 (2.7)	−0.7
c. Starting wage	4.61 (0.02)	4.63 (0.04)	−0.4
d. Wage = \$4.25 (percentage)	30.5 (2.5)	32.9 (5.3)	−0.4
e. Price of full meal	3.35 (0.04)	3.04 (0.07)	4.0
f. Hours open (weekday)	14.4 (0.2)	14.5 (0.3)	−0.3
g. Recruiting bonus	23.6 (2.3)	29.1 (5.1)	−1.0

Difference significant at 5% if $|t| > 1.96$

Difference-in-Differences (DiD)



Difference-in-Differences (DiD)



Difference-in-Differences (DiD)

TABLE 2—MEANS OF KEY VARIABLES

Variable	Stores in: NJ	Stores in: PA	t^a
3. *Means in Wave 2:*			
a. FTE employment	21.0 (0.52)	21.2 (0.94)	−0.2
b. Percentage full-time employees	35.9 (1.4)	30.4 (2.8)	1.8
c. Starting wage	5.08 (0.01)	4.62 (0.04)	10.8
d. Wage = \$4.25 (percentage)	0.0	25.3 (4.9)	—
e. Wage = \$5.05 (percentage)	85.2 (2.0)	1.3 (1.3)	36.1
f. Price of full meal	3.41 (0.04)	3.03 (0.07)	5.0
g. Hours open (weekday)	14.4 (0.2)	14.7 (0.3)	−0.8
h. Recruiting bonus	20.3 (2.3)	23.4 (4.9)	−0.6

Difference significant at 5% if $|t| > 1.96$

Difference-in-Differences (DiD)

The descriptive analysis reveals the change in minimum wages. It also suggests that the characteristics of the restaurants in the sample were comparable in Pennsylvania and New Jersey prior to the increase in minimum wage in New Jersey. But contrary to economic theory the results seem to suggest an increase in employment.

Now, although the units in the sample from Pennsylvania and New Jersey were comparable in January 1992, where they still comparable in November 1992? Specifically, what if Pennsylvania and New Jersey were sitting on different trends of employment?

Difference-in-Differences (DiD)

Then, first, the similarity of traits in January 1992 might have just been a coincidence.

Second, and more critical, the comparisons in the descriptive analysis are meaningless: we just cannot know if the difference in employment in November are due to the differences in trends or to the effect of the minimum-wage rise because we cannot know New Jersey's level of employment in the absence of the increase in minimum wage...

Since data alone are mute, lets **assume** troubles out:

In the absence of any changes to legislation, the relative difference in employment between New Jersey and Pennsylvania had stayed the same between January and November of 1992.

... and come up with some compelling way of arguing in favour of such assumption.

Potential outcomes framework for DiD

We have a collection of units, $i = 1 \ldots, N$ *seen* in two time periods $s = 0, 1$. The units are sampled from a superpopulation of size infinity.

The units come from two strata or groups, identified by the variable $G \in \{0, 1\}$.

Unit i's membership to group G remains constant over time.

Units in strata 1 receive an active treatment in period 1. The binary indicator T equals 1 if a unit receives the treatment, 0 otherwise.

T_{is} is the treatment status of unit i at time s

$G_{is} = S_i$ denotes unit i's group.

Potential outcomes framework for DiD

We maintain SUTVA[3]: unit i's level of outcome under treatment T at time s is $Y_{is}(T_{is})$

Our interest is on the effect of T on Y *for the group affected by the intervention*. Specifically, we want to estimate

$$\begin{aligned}\tau_{\text{ATT}} &= E\big[Y_{i1}(1) - Y_{i1}(0) | G_i = 1\big] \\ &= E\big[Y_{i1}(1) | G_i = 1\big] - E\big[Y_{i1}(0) | G_i = 1\big] \end{aligned} \tag{1}$$

that is the **Average Treatment Effect on the Treated** (ATT).

The selection problem kicks in because $E\big[Y_{i1}(0) | G_i = 1\big]$ is counterfactual.

[3]Note, however, that the no-interference component of SUTVA can be too strong in settings such as the minimum wage study, where geographical spill-overs might be likely; for instance consider two restaurants located at each side of the Penn-New Jersey border. The restaurant in Pennsylvania might have an incentive to increase minimum wages *de facto* if the movement of labour force is fluid. Worse, yet, if panel data are available, non-interference over time periods may be a fanciful assumption.

Potential outcomes framework for DiD

$$\begin{aligned}\tau_{\mathrm{ATT}} &= E\big[Y_{i1}(1) - Y_{i1}(0) | G_i = 1\big] \\ &= E\big[Y_{i1}(1) | G_i = 1\big] - E\big[Y_{i1}(0) | G_i = 1\big]\end{aligned} \tag{2}$$

The selection problem kicks in because $E\big[Y_{i1}(0)|G_i = 1\big]$ is counterfactual.

What if we do a **before-after** comparison of outcomes?

$$E\big[Y_{i1}(1)|G_i = 1\big] - E\big[Y_{i0}(0)|G_i = 1\big] = E\big[Y_{i1}|G_i = 1\big] - E\big[Y_{i0}|G_i = 1\big]$$

For this to identify τ_{ATT} we need to assume that

$$E\big[Y_{i1}(0)|G_i = 1\big] = E\big[Y_{i0}(0)|G_i = 1\big]$$

that is, in the absence of treatment, $Y_{i\bullet}(0)$ had stayed the same. This is too strong for economic variables, for instance, which exhibit trends.

Potential outcomes framework for DiD

$$
\begin{aligned}
\tau_{\text{ATT}} &= E\big[Y_{i1}(1) - Y_{i1}(0) | G_i = 1\big] \\
&= E\big[Y_{i1}(1) | G_i = 1\big] - E\big[Y_{i1}(0) | G_i = 1\big]
\end{aligned} \tag{3}
$$

The selection problem kicks in because $E\big[Y_{i1}(0)|G_i = 1\big]$ is counterfactual.

What if we do a **between-groups** comparison of outcomes?

$$E\big[Y_{i1}(1)|G_i = 1\big] - E\big[Y_{i1}(0)|G_i = 0\big] = E\big[Y_{i1}|G_i = 1\big] - E\big[Y_{i1}|G_i = 0\big]$$

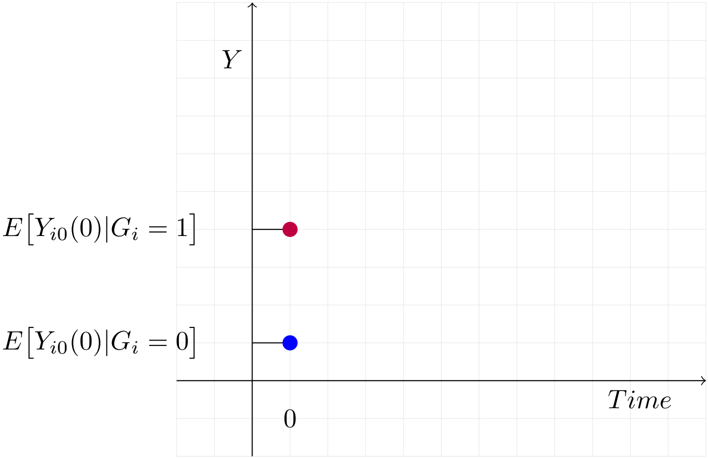
For this to identify τ_{ATT} we need to assume that

$$E\big[Y_{i1}(0)|G_i = 1\big] = E\big[Y_{i1}(0)|G_i = 0\big]$$

that is, $Y_{i\bullet}(0)$ is uncorrelated with group membership; in the absence of treatment both groups share the same mean outcome. This is too strong and implies very strong levels of comparability of units across groups.

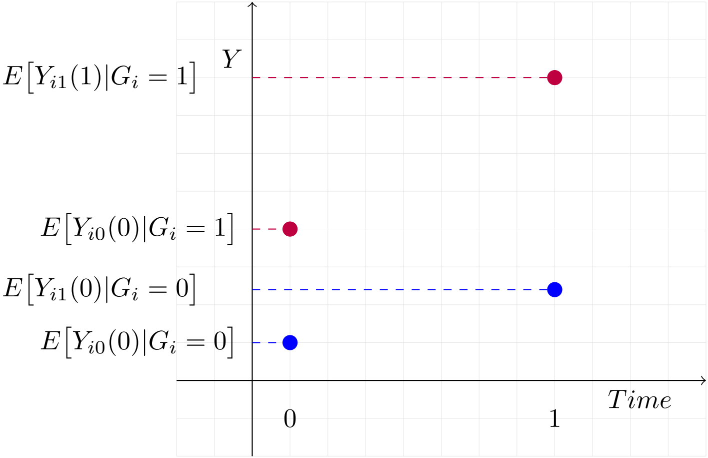
Potential outcomes framework for DiD

In the beginning:



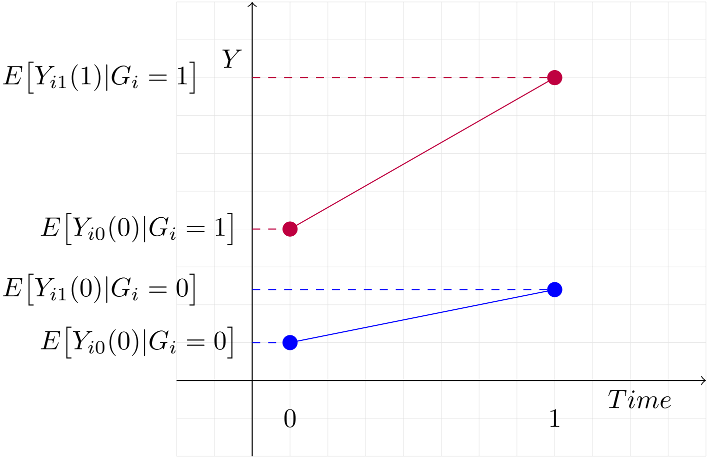
Potential outcomes framework for DiD

...in the end...



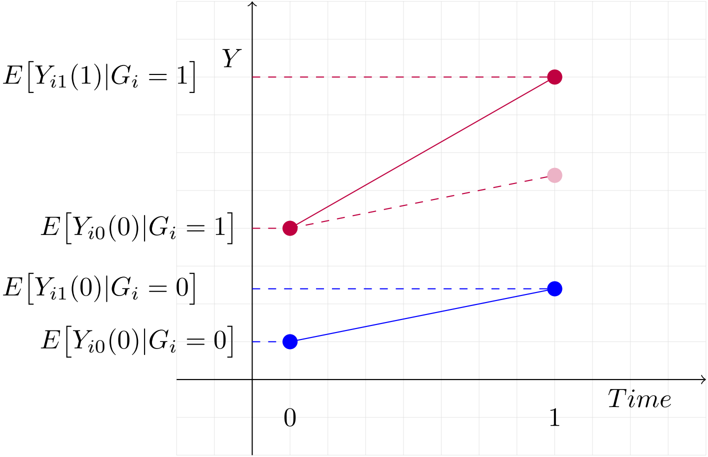
Potential outcomes framework for DiD

...the trends...



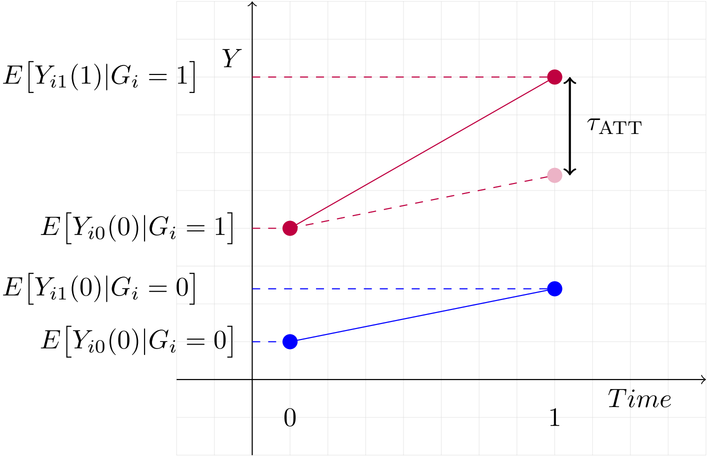
Potential outcomes framework for DiD

...if the trends in $Y_{is}(0)$ are the same across groups ...



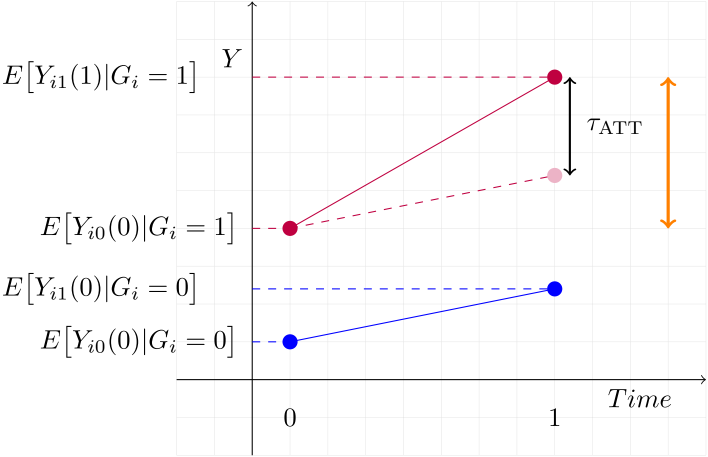
Potential outcomes framework for DiD

...if the trends in $Y_{is}(0)$ are the same across groups ...



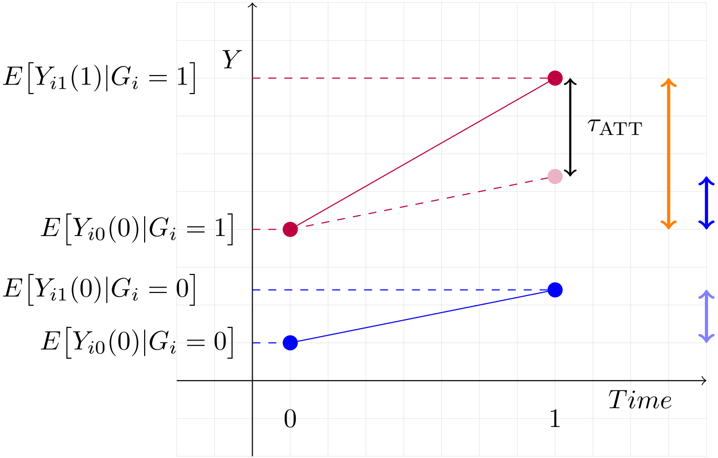
Potential outcomes framework for DiD

compute $E\big[Y_{i1}(1)|G_i=1\big] - E\big[Y_{i0}(0)|G_i=1\big]$



Potential outcomes framework for DiD

subtract trend $E\left[Y_{i1}(0)|G_i = 0\right] - E\left[Y_{i0}(0)|G_i = 0\right]$



Potential outcomes framework for DiD

Identifying assumption

$$E\big[Y_{i1}(0) - Y_{i0}(0)|G_i = 0\big] = E\big[Y_{i1}(0) - Y_{i0}(0)|G_i = 1\big]$$

Then, under SUTVA and the above assumption,

$$\begin{aligned}
\tau_{\text{ATT}} = &\Bigg[E\big[Y_{i1}(1)|G_i = 1\big] - E\big[Y_{i0}(0)|G_i = 1\big]\Bigg] \\
&- \Bigg[E\big[Y_{i1}(0)|G_i = 0\big] - E\big[Y_{i0}(0)|G_i = 0\big]\Bigg] \qquad (4)
\end{aligned}$$

where all the terms can be estimated from data.

Potential outcomes framework for DiD

Often the identifying assumption will be plausible across subgroups defined by a collection of pre-treatment charateristics covariates $\mathbf{X}_{i0} = \mathbf{X}_i$. Then,

$$\begin{aligned}
\tau_{\text{ATT}} &= E\big[Y_{i1}(1) - Y_{i1}(0) | G_i = 1\big] \\
&= \sum_x E\big[Y_{i1}(1) - Y_{i1}(0) | G_i = 1, \mathbf{X}_i = x\big] P\big(\mathbf{X}_i = x | G_i = 1\big)
\end{aligned}$$

where each $E\big[Y_{i1}(1) - Y_{i1}(0) | G_i = 1, \mathbf{X}_i = x\big] = \tau_{\text{ATT}|\mathbf{X}}$ is identified by

$$\begin{aligned}
\tau_{\text{ATT}|\mathbf{X}} = &\Big[E\big[Y_{i1}(1) | G_i = 1, \mathbf{X}_i = x\big] - E\big[Y_{i0}(0) | G_i = 1, \mathbf{X}_i = x\big]\Big] \\
&- \Big[E\big[Y_{i1}(0) | G_i = 0, \mathbf{X}_i = x\big] - E\big[Y_{i0}(0) | G_i = 0, \mathbf{X}_i = x\big]\Big]
\end{aligned}$$

DiD: Estimation

The ATT parameter,

$$\tau_{\text{ATT}} = \Bigg[E\big[Y_{i1}(1)|G_i = 1\big] - E\big[Y_{i0}(0)|G_i = 1\big] \Bigg] - \Bigg[E\big[Y_{i1}(0)|G_i = 0\big] - E\big[Y_{i0}(0)|G_i = 0\big] \Bigg] \tag{5}$$

is nonparametrically identified, since

$$E\big[Y_{i1}(1)|G_i = 1\big] = E\big[Y_{i1}|G_i = 1\big],\; E\big[Y_{i0}(0)|G_i = 1\big] = E\big[Y_{i0}|G_i = 1\big]$$
$$E\big[Y_{i1}(0)|G_i = 0\big] = E\big[Y_{i1}|G_i = 0\big],\; E\big[Y_{i0}(0)|G_i = 0\big] = E\big[Y_{i0}|G_i = 0\big]$$

and all the right-hand-side moments can be directly estimated from data, without further modelling.

To formally discuss estimation notation will differ depending on whether we have panel data or a *repeated cross-section*

DiD: Estimation with panel data

If each unit i in each group is observed at time $s = 0$ and $s = 1$, then

$$\frac{1}{N_1}\sum_{i=1}^{N} Y_{is} \cdot G_i \rightarrow E\big[Y_{is}|G_i = 1\big] \text{ and } \frac{1}{N_0}\sum_{i=1}^{N} Y_{is} \cdot (1 - G_i) \rightarrow E\big[Y_{is}|G_i = 0\big]$$

(where convergence is in probability) and so,

$$\begin{aligned}
\hat{\tau}_{\text{ATT}} &= \left[\frac{1}{N_1}\sum_{i=1}^{N} Y_{i1} \cdot G_i - \frac{1}{N_1}\sum_{i=1}^{N} Y_{i0} \cdot G_i\right] \\
&\quad - \left[\frac{1}{N_0}\sum_{i=1}^{N} Y_{i1} \cdot (1 - G_i) - \frac{1}{N_0}\sum_{i=1}^{N} Y_{i0} \cdot (1 - G_i)\right] \\
&= \left[\frac{1}{N_1}\sum_{i=1}^{N} (Y_{i1} - Y_{i0}) \cdot G_i - \frac{1}{N_0}\sum_{i=1}^{N} (Y_{i1} - Y_{i0}) \cdot (1 - G_i)\right] \qquad (6)
\end{aligned}$$

DiD: Estimation with panel data

To estimate the variance of

$$\hat{\tau}_{\text{ATT}} = \left[\frac{1}{N_1} \sum_{i=1}^{N} (Y_{i1} - Y_{i0}) \cdot G_i - \frac{1}{N_0} \sum_{i=1}^{N} (Y_{i1} - Y_{i0}) \cdot (1 - G_i) \right] \tag{7}$$

note that $\hat{\tau}_{\text{ATT}}$ is a difference in means, similar to that used in Randomized Experiments; therefore, in principle, we could use the sum of the sample variances of each term as an estimator of the variance of $\hat{\tau}_{\text{ATT}}$. Note, however, that this estimator does not take into account variation due to inter-group and inter-individual correlations.

Alternatively, one could use a block-bootstrap estimator, resampling with replacement at group and individual level.

DiD: Estimation with repeated cross-section

If different units are sample in each group at time $s = 0$ and $s = 1$, then we introduce additional notation; specifically, let S_i denote if unit i was sampled at time $s = 0, 1$. Then our data reveals (Y_i, S_i, T_i). The parameter we can estimate needs to be re-defined as,

$$\tau_{\text{ATT}} = \Big[E\big[Y_i(1)|G_i = 1, S_i = 1\big] - E\big[Y_i(0)|G_i = 1, S_i = 1\big]\Big] \tag{8}$$

This parameter is identified by

$$\begin{aligned}\tau_{\text{ATT}} = &\Big[E\big[Y_i(1)|G_i = 1, S_i = 1\big] - E\big[Y_{i0}(0)|G_i = 1, S_i = 0\big]\Big] \\ &- \Big[E\big[Y_i(0)|G_i = 0, S_i = 1\big] - E\big[Y_i(0)|G_i = 0, S_i = 0\big]\Big]\end{aligned} \tag{9}$$

DiD: Estimation with repeated cross-section

Then

$$\frac{\sum_{i=1}^{N} Y_i \cdot S_i \cdot G_i}{\sum_{i=1}^{N} G_i \cdot S_i} \rightarrow E\big[Y_{is} | G_i = 1, S_i = s\big]$$

and so on[4]; therefore

$$\begin{aligned}\hat{\tau}_{\mathrm{ATT}} = & \left[\frac{\sum_{i=1}^{N} Y_i \cdot S_i \cdot G_i}{\sum_{i=1}^{N} G_i \cdot S_i} - \frac{\sum_{i=1}^{N} Y_i \cdot (1 - S_i) \cdot G_i}{\sum_{i=1}^{N} G_i \cdot (1 - S_i)}\right] \\ & - \left[\frac{\sum_{i=1}^{N} Y_i \cdot S_i \cdot (1 - G_i)}{\sum_{i=1}^{N} (1 - G_i) \cdot S_i} - \frac{\sum_{i=1}^{N} Y_i \cdot (1 - S_i) \cdot (1 - G_i)}{\sum_{i=1}^{N} (1 - G_i) \cdot (1 - S_i)}\right] \end{aligned} \tag{10}$$

[4] Convergence is in probability.

Regression framework for DiD

We can provide a identification strategy based on regression. Let's focus first on instances with repeated cross sections. First, define

$$E\left[Y_i(0)|G_i = 0, S_i = 0\right] = \alpha \tag{11}$$

Regression framework for DiD

The identification assumption for DiD implies that, in the absence of treatment

$$\begin{aligned} & E\big[Y_i(0)|G_i = 0, S_i = 1\big] - E\big[Y_i(0)|G_i = 0, S_i = 0\big] \\ & = E\big[Y_i(0)|G_i = 1, S_i = 1\big] - E\big[Y_i(0)|G_i = 1, S_i = 0\big] \end{aligned} \tag{12}$$

This can summarised by assuming a common linear trend for both groups, so that,

$$E\left[Y_i(0)|G_i = 0, S_i = s\right] = \alpha + \delta \cdot s \tag{13}$$

Regression framework for DiD

Next, we may allow some structural differences between groups; we can summarise this via an additive effect

$$E\left[Y_i(0)|G_i = g, S_i = s\right] = \alpha + \delta \cdot s + \gamma \cdot g \tag{14}$$

Regression framework for DiD

Finally, we allow for the differential effect of time and group membership -that is, the treatment effect,

$$\begin{aligned} E\left[Y_i(1)|G_i = g, S_i = s\right] &= E\left[Y_i(0)|G_i = g, S_i = s\right] + \tau_{\mathrm{ATT}} g \cdot s \\ &= \alpha + \delta \cdot s + \gamma \cdot g + \tau_{\mathrm{ATT}} \cdot g \cdot s \end{aligned} \tag{15}$$

which is a standard regression model,

$$Y_i = \alpha + \delta \cdot S_i + \gamma \cdot G_i + \tau_{\mathrm{ATT}} \cdot G_i \cdot S_i + \varepsilon_i \tag{16}$$

estimable via least squares, provided that ε_i is uncorrelated with S_i, G_i.

As before, it is recommended that the standard errors are adjusted to take into account potential correlations of units within the same group.

Regression framework for DiD

The availability of panel data will allow us to work with more flexible specifications. In particular, we can define

$$Y_{is} = \alpha_i + \delta S_t + \tau_{\text{ATT}} \cdot T_{it} + \varepsilon_{it} \tag{17}$$

where now each unit i can have its own intercept (time-invariant unobserved heterogeneity). Furthermore, the error term and α_i need not be uncorrelated.

Taking first differences, we see

$$Y_{i1} - Y_{i0} = \delta + \tau_{\text{ATT}} \cdot (T_{i1} - T_{i0}) + (\varepsilon_{i1} - \varepsilon_{i0}) \tag{18}$$

which can be estimated by least squares; if desired, this latter model can be extended by adding **pre-test** covariates.

Violations of the identifying assumption

The critical assumption for identification implies that, in the absence of treatment, the rate of variation between $s = 0$ and $s = 1$ is identical across the groups.

This will hold if the groups sit on similar trends.

However, the threats to this assumption are substantial...

Violations of the identifying assumption

1. Behind each policy, there is an intentionality. If areas/units targeted by the policy are consistently better/worse, this might present a challenge to the identifying assumption.
2. In repeated cross-sections the composition of the sample might vary over time (due to migration, death, health dynamics, etc). This could potentially carry attached a change in trends.
3. Time span between $s = 0$ and $s = 1$; the larger the temporal gap between periods, the less credible the identifying assumption might be.
4. The magnitude and sign of the estimates in DiD models may be very sensitive to functional form specification and transformations of the dependent variable (specifically, using $\log(Y)$ instead of Y).

You cannot directly test the identifying assumption, but a number of steps are now routine in studies to try to protect the integrity of DiD studies.

Defending your DiD study.

Pre-treatment trends in the outcome
A first check consists on looking at whether before the intervention the groups exhibited parallel trends.

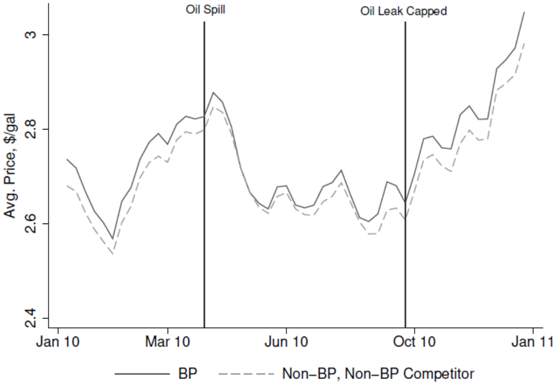
Example: Barrage et al. (2014)

- Did the 2010 *Deepwater Horizon* oil spill damaged BP's retail sales?
- Prior to 2010 BP had heavily invested in re-branding as an evironmentally friendly company (from British Petroleum $\rightarrow$ Beyond Petroleum).
- Between April 2010 and July 2010 an estimated 205.8 million gallons of oil leaked into the gulf of Mexico.
- To study the effect of this oil spill on BP's retail sales, the authors built a sample detailing retail sales in BP petrol stations (treatment group) and carefully selected non-BP petrol stations (control group).

Defending your DiD study.

Pre-treatment trends in the outcome
A first check consists on looking at whether before the intervention the groups exhibited parallel trends.

Figure 1: Average Weekly Retail Price for BP and Comparison Group Stations



Defending your DiD study.

Placebo tests
The intervention should only affect outcomes in the period and area of implementation.

Therefore, apply the same DiD model to compare,...

- ...using the same groups, two consecutive periods where no intervention took place
- ...using the same time periods, replace the treated group by a group that was not affected by the treatment.
- ...using the same groups and time periods, an outcome that should be unaffected by the intervention.

In both instances, the DiD estimate should be statistically insignificant at all levels.

References I

Barrage, L., **Chyn, E.**, and **Hastings, J.** (2014). Advertising and environmental stewardship: Evidence from the bp oil spill. Working Paper 19838, National Bureau of Economic Research

Card, D. and **Krueger, A.B.** (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review*, 84(4): 772–793

Card, D. and **Krueger, A.B.** (2000). Minimum wages and employment: A case study of the fast-food industry in new jersey and pennsylvania: Reply. *The American Economic Review*, 90(5): 1397–1420. ISSN 00028282

Neumark, D. and **Wascher, W.** (2000). Minimum wages and employment: A case study of the fast-food industry in new jersey and pennsylvania: Comment. *The American Economic Review*, 90(5): 1362–1396. ISSN 00028282