UCLA Anderson
SCHOOL of MANAGEMENT

PMDA

Lecture 6: Difference-in-Differences

A few updates and reminders

Classes:

- Class 7: Friday, February 23rd at the regular time and classroom
- **Final Exam:** Thursday, March 21st 8:30 to 11:30 am, classroom TBA

Problem sets:

- If you want to look at how your PS was graded, please attend the TA's office hours or schedule a time to meet with them
- Problem Set 3: Panel Data & Diff-in-diff due Feb 25th
- Problem Set 4: Lasso Regression due March 16th

Last class

- **Interaction terms**
    - Implementation: include product of two existing variables as a separate variable in the regression
    - Interpretation: can be used to analyze whether treatment affects different consumers differently (→ targeting & segmentation)

- **Panel regression with two-way fixed effects**
    - Way to eliminate bias due to omitted variables (observable and unobservable) that do not vary over time / do not vary across units
    - Avoids having to look for controls (unless they vary in both dimensions)
    - Careful: causality not guaranteed!

Last class

- Two-way fixed effects regression
  - Implicitly controls for store size and summer dummy by including unit and time fixed effects (i.e., store and week dummies)
  - In fact, it controls for any variable that varies either only across stores or only over time (e.g. store quality, other seasonal effects such as holidays)
  - → Powerful tool that eliminates bias from a variety of possible (and unobservable) omitted variables that do not vary in both dimensions

Agenda for today

- Difference-in-differences (diff-in-diff) regression
  - Goal: how to estimate treatment effects with observational data
  - Application: Online search advertising (eBay study)
  - Differences-in-differences in a regression framework

Treatment effects with experimental data

- We have seen how to estimate the causal effect of a treatment (e.g., showing an ad) in fully randomized (experimental) data
- Random assignment of the treatment creates a treated group and a control group
- The difference in outcomes (e.g., revenue) between treated and control group measures the treatment effect
    - Estimate this by regressing revenue on treatment
- We then saw that in partially randomized data (e.g., targeted ads) you can still get causal estimates by controlling for the cause of non-random assignment (age)
- Today we will see how and when you can estimate treatment effects if you only have observational data

Treatment effects with observational data

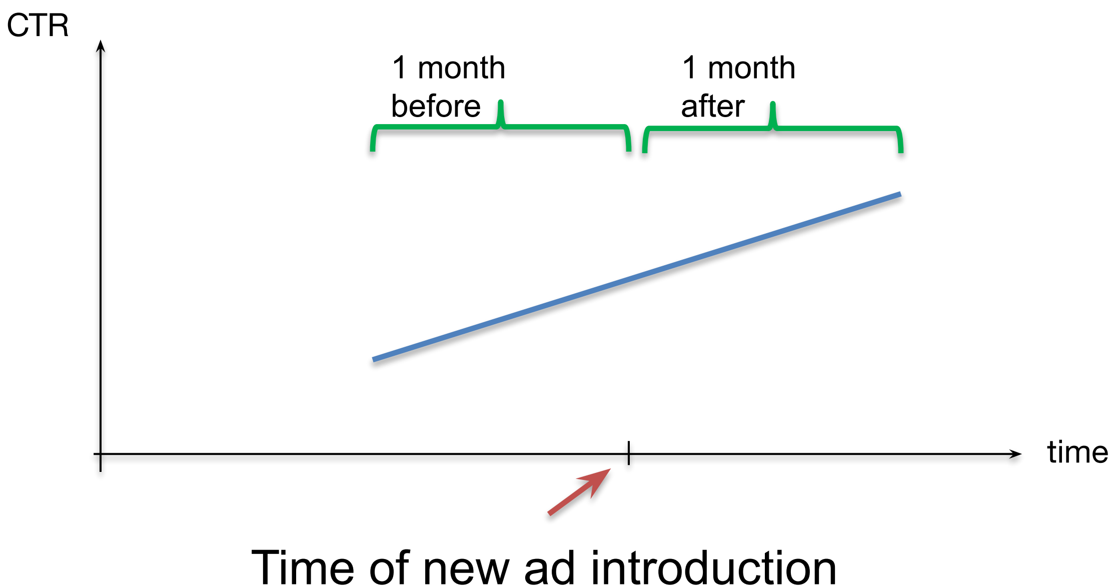
- Quite often companies perform "tests" without running a proper experiment
- For instance a company might re-design their online ad on MSN search engine and show it to everybody, not just a random set of users
- They then compare some measure of performance before and after the change
    - E.g. compute average CTR (click-through rate) or other metric of success over a period before and a period after the change

Q: Suppose we find that average CTR is different the month before and after the introduction of the new ad. Can we be sure that the change was caused by the ad?

A: Maybe something else happened during the period that makes CTR change over time (general trends, seasonal effects) → biased estimate of the treatment effect

Problems with before/after comparison

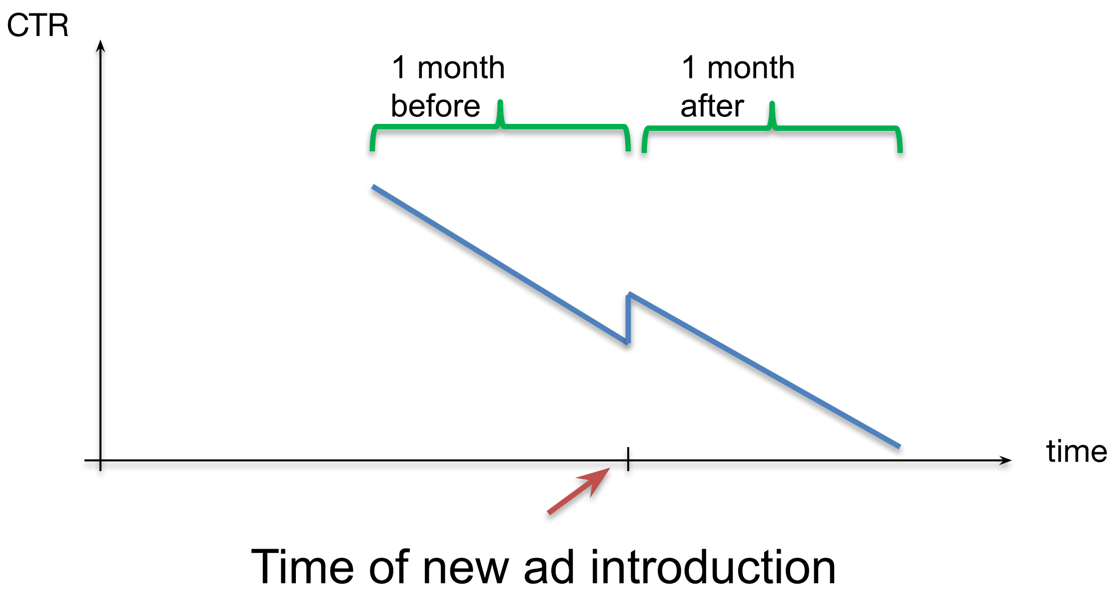
- Say there is a general upward trend in CTRs



- At the time of the ad introduction there is no additional change in CTR besides the trend
- If we compared average CTR over the month before and the month after the ad intro, we would incorrectly conclude that the CTR increases due to the ad

Problems with before/after comparison

- Say there is a general downward trend in CTRs



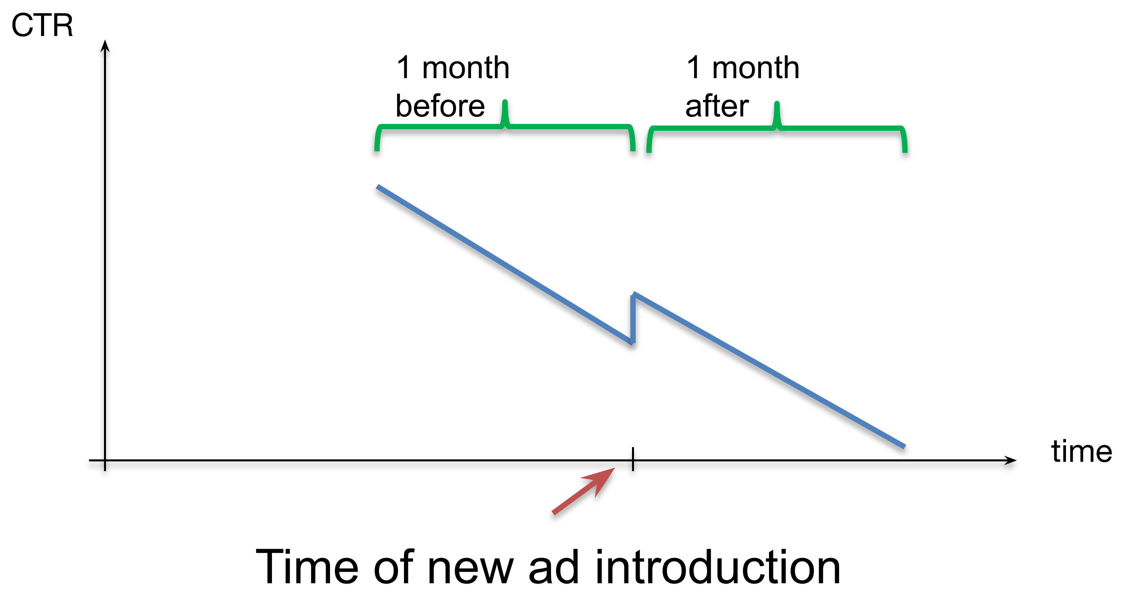
- At the time of the ad introduction there is indeed an increase in CTR (the line shifts upwards) slido

Comparing the average CTR over the month before and the month after leads to concluding that

ⓘ Click **Present with Slido** or install our Chrome extension to activate this poll while presenting.

Problems with before/after comparison

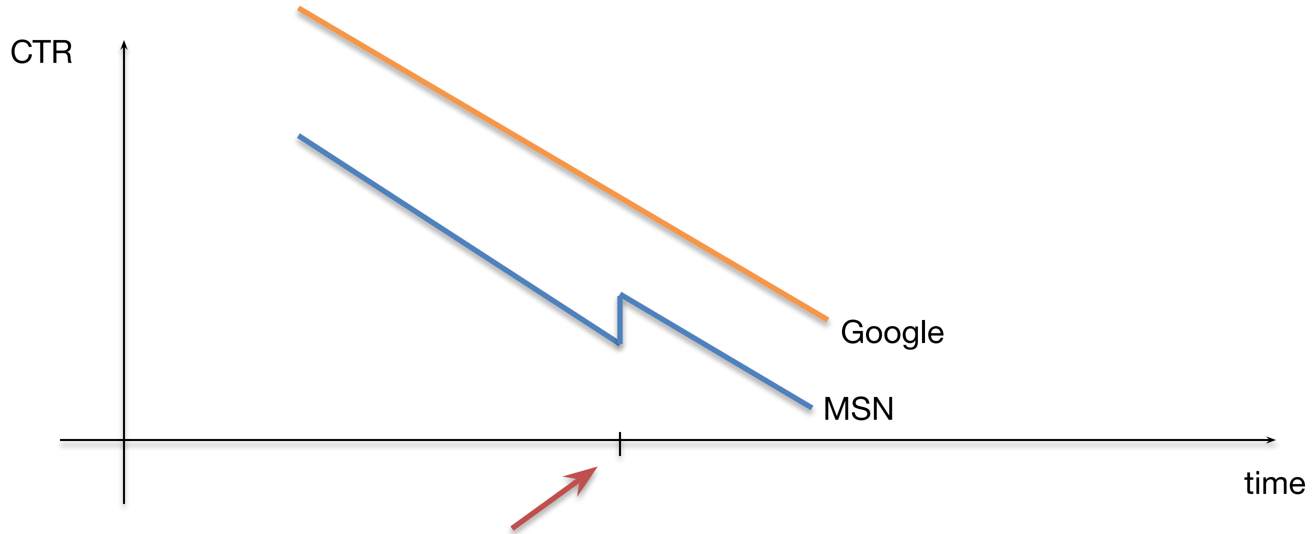
- Say there is a general downward trend in CTRs



- At the time of the ad introduction there is indeed an increase in CTR (the line shifts upwards)
- However, if we compare average CTR over the month before and the month after, the downward trend would “drown out” the positive effect and we would conclude that the CTR decreases

Solution?

- Suppose the company was advertising on both Google and MSN, but only changed the ad on MSN

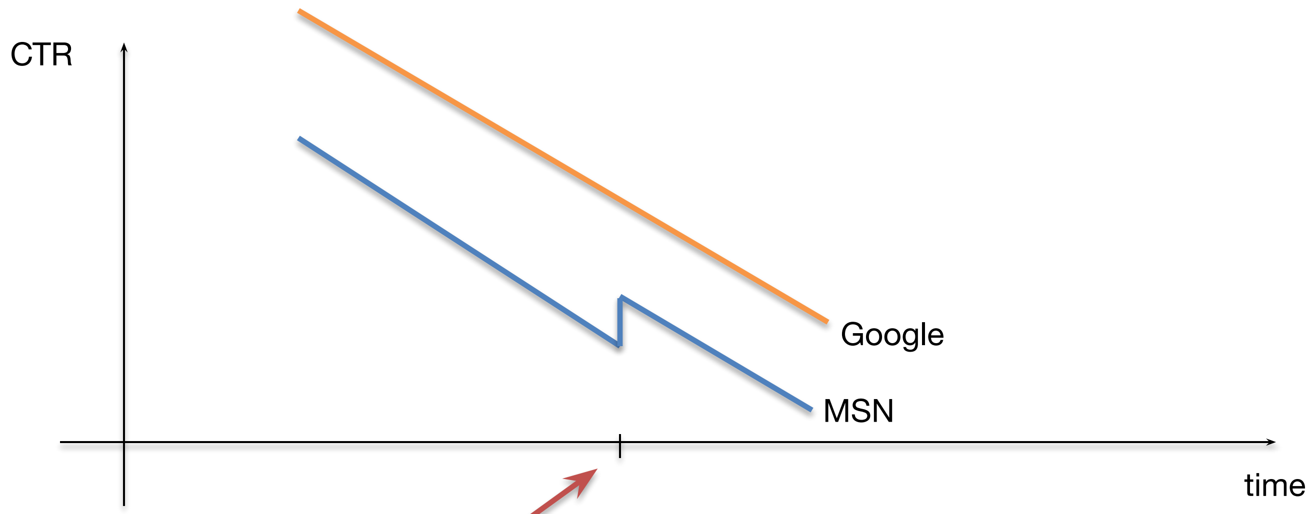


Time of new ad introduction

Q: How does this help?

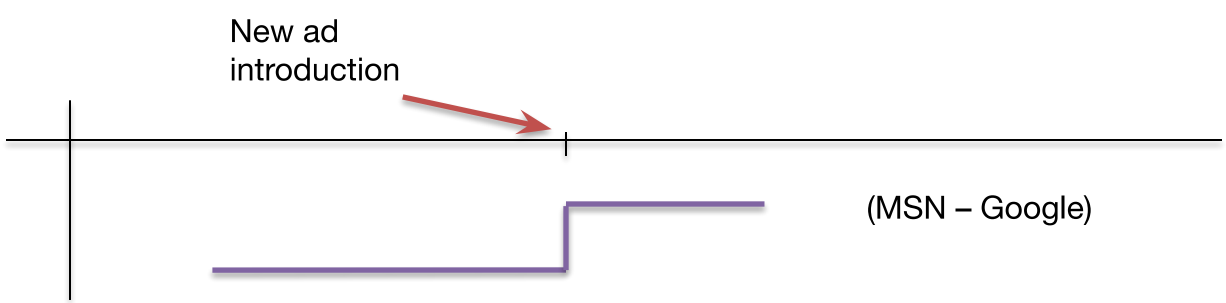
A: We have DIFFERENT LEVELS of CTR on the two platforms but the SAME TIME VARIATION (downward trend), except for the jump in CTR on MSN at the time of the new ad introduction

Finding a control group

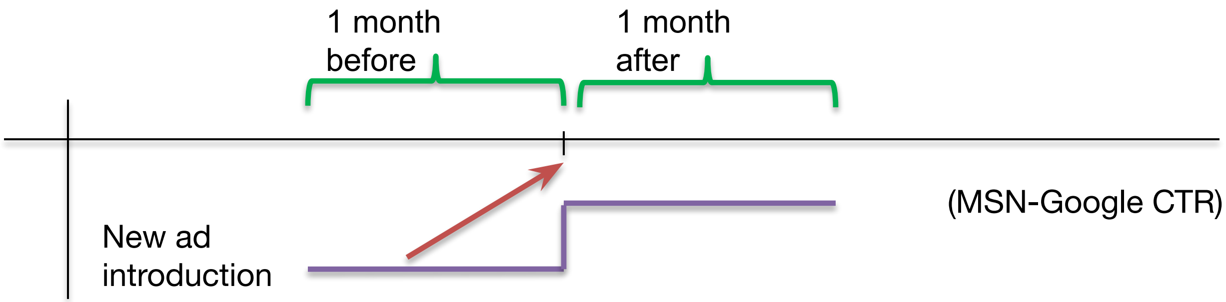


Time of new ad introduction

Q: What does a graph of the difference between MSN and Google CTR look like?



Double differencing …



Q: How can we use the graph above to measure the effect of the new ad?

A: It is simply the increase in the difference (CTR_MSN – CTR_Google) when the new ad is introduced

- We are evaluating whether the difference in CTR between MSN and Google is different before and after the new ad was introduced only on MSN
- Is (CTR_MSN – CTR_Google)_after larger or smaller than (CTR_MSN – CTR_Google)_before ?
- Same as checking if (CTR_MSN – CTR_Google)_after - (CTR_MSN – CTR_Google)_before is positive or negative
- → This is the difference-in-differences estimator of the ad effect on CTR

Diff-in-diffs estimator in formal terms

- A group of people (treated group) was assigned a treatment (saw the ad) at a time t^Treat, but the treatment was **not** randomly assigned
- You have panel data for the outcome of interest (average revenue, CTR) for the treated group before and after t^Treat, let's call them $Y_{before}^{treated}$ and $Y_{after}^{treated}$
- If
  - You also have panel data for a group of people who were not exposed to the treatment (control group), $Y_{before}^{control}$ and $Y_{after}^{control}$
  - The outcome in the treated and control groups would have the same time variation if it weren't for the treatment ("parallel trends" assumption)
- Then you can estimate the treatment effect on observational data using the diff-in-diff estimator:

$$(Y_{after}^{treated} - Y_{after}^{control}) - (Y_{before}^{treated} - Y_{before}^{control})$$

Same as

$$(Y_{after}^{treated} - Y_{before}^{treated}) - (Y_{after}^{control} - Y_{before}^{control})$$

Agenda for today

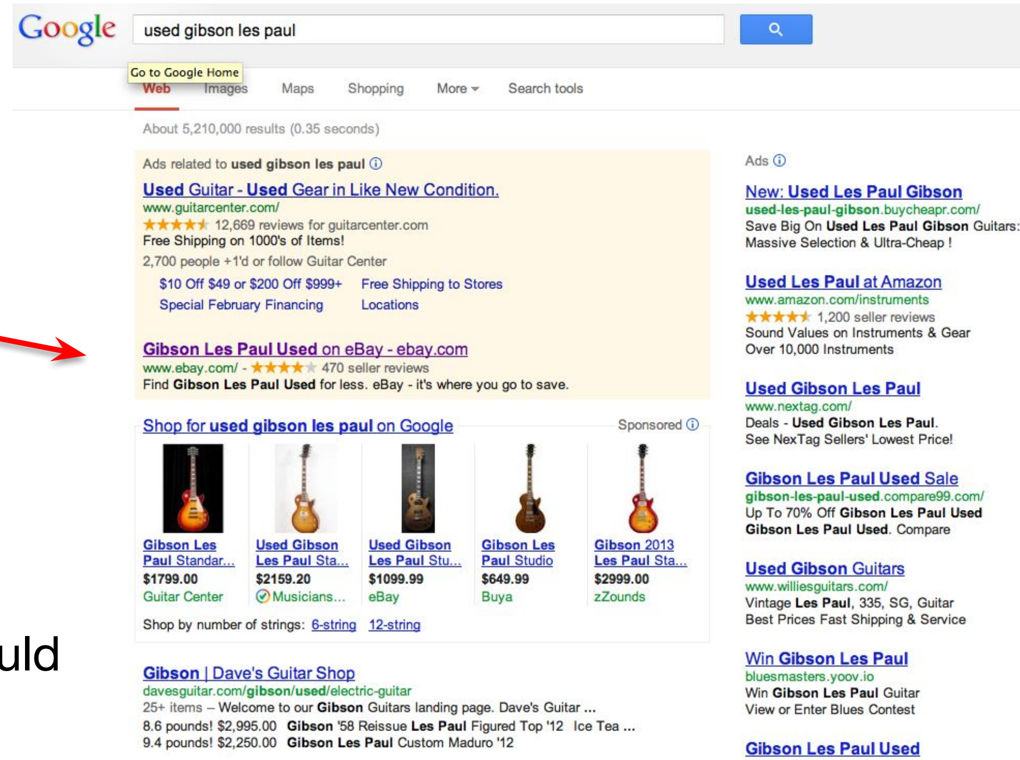
- Difference-in-difference regression
    - Goal: how to estimate treatment effects with observational data
    - Application: online search advertising (eBay study)
    - Differences-in-differences in a regression framework

Search advertising

- Blake, Nosko, Tadelis (see our first lecture) analyze the effect of paid links for eBay

Search Advertising /

Paid Link

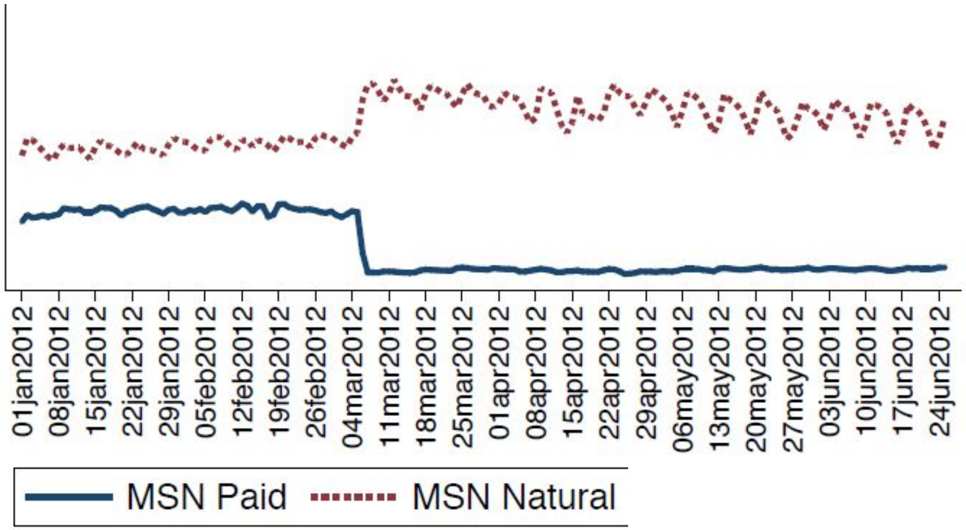


- They are interested in the causal effect of the search ads (what would happen if we switched them off?)

- They conjectured that many consumers might still come to the eBay webpage via organic links

Search advertising

- They convinced eBay to stop their search ads on MSN (they kept it running on other search engines)
- The graph below shows the main results

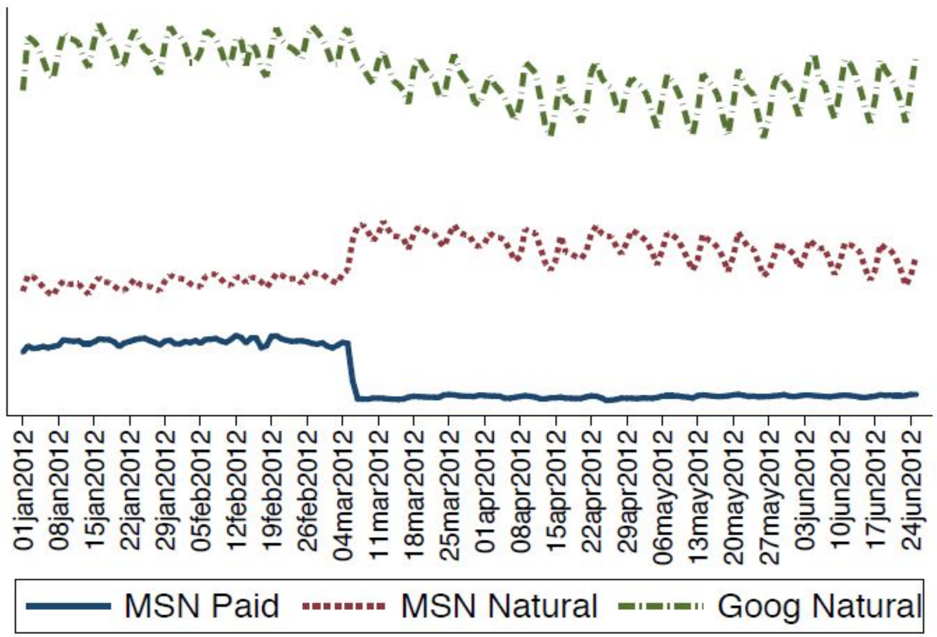


Q: Does it seem like there is a one to one substitution between paid and organic links?

A: The jump up in the red line is a bit smaller than the drop in the blue one. This difference turns out to be about 5%

Search advertising

- They also compared the pattern of time variation with the one from Google (where they still had search ads)



Q: Does the Google pattern change your assessment of the MSN paid versus organic substitution?

A: It seems that the CTR also dropped on Google around the time search ads were removed from MSN

Agenda for today

- Difference-in-difference regression
    - Goal: how to estimate treatment effects with observational data
    - Application: Online search advertising (eBay study)
    - Differences-in-differences in a regression framework

How to estimate diff-in-diffs in a regression

- Consider the following regression:

$$CTR_{it} = \beta_0 + \beta_1 MSNDummy_i + \beta_2 PostDummy_t + \beta_3 MSNDummy_i \times PostDummy_t + e_{it}$$

- Where i denotes the search engine and t the time period (day)
- In other words, we include
  - A dummy for MSN
  - A dummy for whether the observation comes from the time period after the treatment (search ads removed from MSN)
  - An interaction of the two terms above (you can think of the interaction as a “treatment dummy”, because it equals 1 only for MSN after the treatment)
- **NOTE:** this is similar to two-way fixed effects, we will make the comparison more explicit in a bit

Diff-in-diff regression mechanics

$$CTR_{it} = \beta_0 + \beta_1 MSNDummy_i + \beta_2 PostDummy_t + \beta_3 MSNDummy_i \times PostDummy_t + e_{it}$$

Q: What do the following coefficients (/combinations) represent?

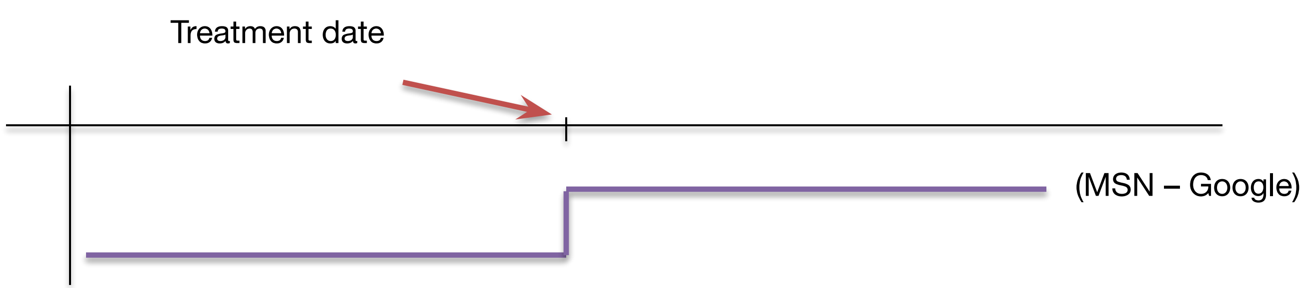
- β_0 E(CTR | Google, before treatment)
- $\beta_0 + \beta_2$ E(CTR | Google, after treatment)
- $\beta_0 + \beta_1$ E(CTR | MSN, before treatment)
- $\beta_0 + \beta_1 + \beta_2 + \beta_3$ E(CTR | MSN, after treatment)

Q: What is the difference in average CTR between MSN and Google before and after treatment?

A: before after

$$[(\beta_0 + \beta_1) - (\beta_0)] = \beta_1 \quad [(\beta_0 + \beta_1 + \beta_2 + \beta_3) - (\beta_0 + \beta_2)] = \beta_1 + \beta_3$$

Diff-in-diff regression mechanics



- The difference in CTR between MSN and Google equals:
  - Before treatment:

$$\Delta CTR_{before} = \beta_1$$

  - After treatment:

$$\Delta CTR_{after} = \beta_1 + \beta_3$$

- The difference in the two expressions above captures the effect of the treatment

$$\Delta CTR_{after} - \Delta CTR_{before} = \beta_3$$

→ The estimator of β_3 is the diff-in-diff estimator!

Two-way fixed effects & diff-in-diff

- Diff-in-diff regression:

$$CTR_{it} = \beta_0 + \beta_1 MSNDummy_i + \beta_2 PostDummy_t + \beta_3 MSNDummy_i \times PostDummy_t + e_{it}$$

- We could extend this to a two-way fixed effects model:

$$CTR_{it} = \delta_t + \beta_1 MSNDummy_i + \beta_3 MSNDummy_i \times PostDummy_t + e_{it}$$

  - This includes a dummy for every day (captured by the time fixed effect δ_t), rather than just PostDummy_t (which drops out when you have δ_t)
  - Since we have only two cross-sectional units, we already have search-engine fixed effects by including MSNDummy_i
  - You can think of MSNDummy_i∗PostDummy_t as a treatment variable that varies across i and t (like price in our earlier example)
  - i.e. treatment is switched on only on MSN and only in later weeks

Two-way fixed effects & diff-in-diff

- Two-way fixed effects model:

$$CTR_{it} = \delta_t + \beta_1 MSNDummy_i + \beta_3 MSNDummy_i \times PostDummy_t + e_{it}$$

- The interpretation of β_3 is the same as before (except that the expectations below depend on t):

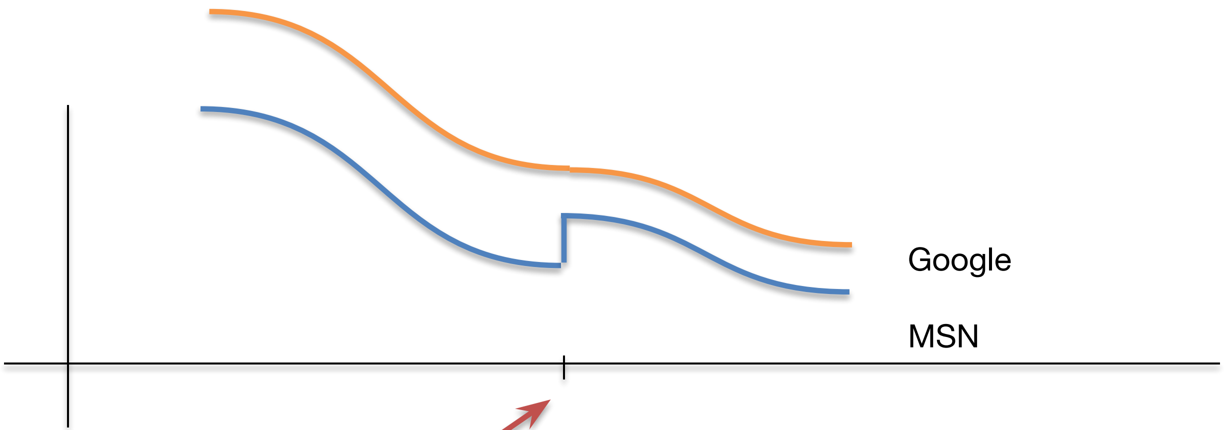
- $E(CTR|Google, before\ \ treatment) = \delta_t$
- $E(CTR|MSN, before\ \ treatment) = \delta_t + \beta_1$
  $\rightarrow \Delta CTR_{before} = \beta_1$
- $E(CTR|Google, after\ \ treatment) = \delta_t$
- $E(CTR|MSN, after\ \ treatment) = \delta_t + \beta_1 + \beta_3$
  $\rightarrow \Delta CTR_{after} = \beta_1 + \beta_3$

$$\Delta CTR_{after} - \Delta CTR_{before} = \beta_3$$

→ The estimator of β_3 is again the diff-in-diff estimator!

Two-way fixed effects & diff-in-diff

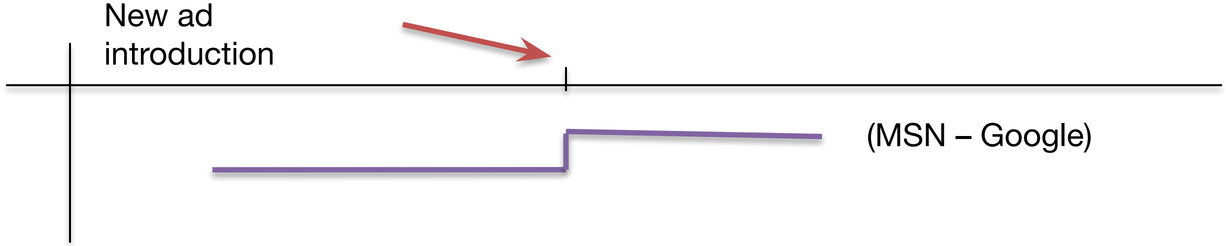
- Two-ways fixed effects model allows for nonlinear trends:



New ad
introduction

- But the parallel trends assumption implies that the difference between MSN and Google CTRs looks the same as before

New ad
introduction



Results for the eBay study

- Columns (1) and (3) regress log clicks on PostDummy_t for MSN and Google
- Estimate of drop in total clicks (paid and organic) on MSN is 5%
- Clicks on Google also dropped by 3.2% (column 3)

	MSN		Google
	(1)	(2)	(3)
	Log Clicks	Log Clicks	Log Clicks
Period	-0.0560***		-0.0321*
	(0.00861)		(0.0124)
Interaction		-0.00529	
		(0.0177)	
Google		5.088	
		(10.06)	
Yahoo		1.375	
		(5.660)	
Constant	12.82***	11.33*	14.34***
	(0.00583)	(5.664)	(0.00630)
Date FE		Yes	

slido

The MSN regression in (1) gives a coefficient of -0.056 while the Google regression in (3) gives a coefficient of -0.0321. These result suggests that in (1)

Results for the eBay study

- Columns (1) and (3) regress log clicks on PostDummy_t for MSN and Google
- Estimate of drop in total clicks (paid and organic) on MSN is 5%
- However, clicks on Google also dropped by 3.2% (column 3)
- Column (2) is a two-ways fixed effects regression (also considering Yahoo data, so must include another dummy)

	MSN		Google
	(1)	(2)	(3)
	Log Clicks	Log Clicks	Log Clicks
Period	$\boxed{-0.0560^{***}}$		$\boxed{-0.0321^{*}}$
	$\boxed{(0.00861)}$		$\boxed{(0.0124)}$
Interaction		$\boxed{-0.00529}$	
		$\boxed{(0.0177)}$	
Google		5.088	
		(10.06)	
Yahoo		1.375	
		(5.660)	
Constant	12.82***	11.33*	14.34***
	(0.00583)	(5.664)	(0.00630)
Date FE		Yes	

Time fixed effects

Search engine fixed effects

- The interaction variable is MSNDummy_i∗PostDummy_t so its coefficient is the diff-in-diff estimator

- The drop in clicks is now only 0.5% (rather than 5%) and not significant!

Summary: diff-in-diff

- We can use diff-in-diff for causal inference
- All we need is some kind of control group which we think experiences a similar evolution over time (even if the level is different) but is not exposed to treatment
- Using diff-in-diff is particularly important in markets with high seasonal variability where simple pre-/post comparison is difficult

Workshop

- Workshop analyzes the impact of the Philadelphia soda tax

- Background:
    - Several US cities have implemented such taxes, others are thinking about it
    - Beverage industry is very worried about this
    - Tax is substantial: 1.5 cents/Oz
    - This is actual data (anonymized)
    - Data records store level sales at stores in Philadelphia and stores outside (where the tax does not apply)

Workshop comments

- Question 3. We want to know
  - How much sales changed in Philly after the tax, relative to stores outside Philly and more than 6 miles away
  - How much sales changed in stores outside Philly and less than 6 miles away, relative to stores more than 6 miles away
- The regression we want is (with stores FE):

$$Sales_{it} = \alpha_i + \beta_1 PostDummy_t + \beta_2 Philly_i \times PostDummy_t$$
$$+ \beta_3 NearPhilly_i \times PostDummy_t + e_{it}$$

where NearPhilly_i is a dummy for stores less than 6 miles from Philly

- We have three dummies for: Philly stores, stores outside and near Philly and stores outside and far from Philly, so we interpret the coefficients as:
  - β_1 is the change for the left out category (stores far from Philly)
  - β_2 is the change for Philly, <u>relative to</u> that for stores far from Philly
  - β_3 is the change for stores near Philly, <u>relative to</u> that for stores far from Philly