Diff in Diff: Applications

Lachlan Deer

Social Media and Web Analytics, Spring 2025

Learning Goals

1. Interpret Difference in Difference results found in the literature
2. Explain advantages and shortcomings of choices made in existing research designs that leverage natural experiments
3. Interpret analysis in search engine advertising markets and on social media platforms

Where are we now?

In the previous class:

- Difference in differences as a research design to analyse data from natural experiments

This class

- Applications of Difference in Differences research design in digital markets

Today's Topics

1. **Search engine advertising effectiveness**

- Randomised Control Trial, but with imperfect randomization into Treatment and Control
- Shuts down search engine ads by eBay in geographic regions of the US
- Examines impact on sales of eBay products
- Discussion below from Blake, Nosko and Tadelis

2. **Effect of social media on product demand**

- Natural experiment leveraging a shutdown of social media in mainland China but not in Hong Kong * To study how social media impacts TV viewership
- Discussion below from Seiler, Yao and Wang

1/ Search Engine Ad Effectiveness

The Business Problem

Seeking answers to the following strategic questions:

- Are Brand based SEM ads effective at bringing traffic to my site?
- Are non-Brand based SEM ads effective at generating sales?
- Are the effects heterogenous across consumers?
- Are the effects heterogenous across companies?

Are Paid Search Ads Effective?

Motivation: Is Search Engine Marketing Effective?

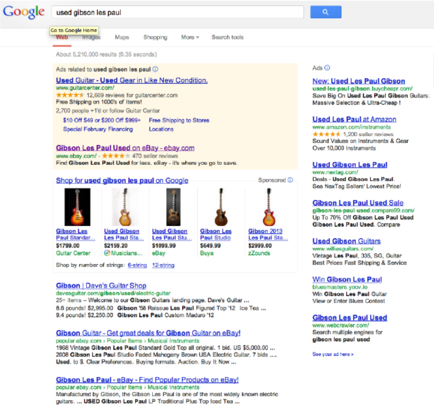
Specific Questions:

- Does SEM generate a positive Return on Investment?
- Is SEM an informative or persuasive form of advertising?

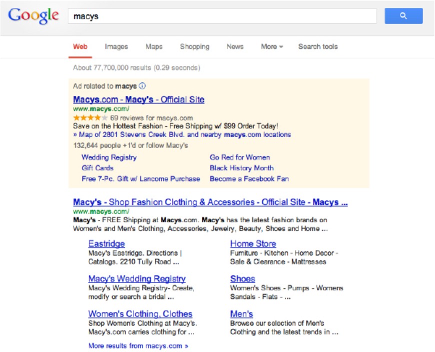
How?

- A series of controlled experiments at eBay
- First, a “proof of concept”
- Then a larger scale experiment

Paid Search in 2012



(a) Used Gibson Les Paul



(b) Macys

Brand Search Terms Experiment

Brand Terms: any queries that include the name of the brand

- Examples: 'ebay shoes', 'de bijenkorf dress',

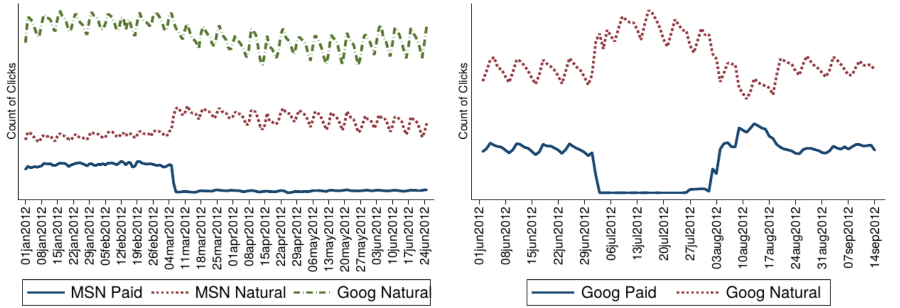
Hypothesis: Users who type the brand name intend to go to that site anyway

$\implies$ brand ads are intercepting what would otherwise be organic clicks

Experiments:

- Experiment 1 (March to June 2012): Shutdown brand ads on MSN and Yahoo!
- Experiment 2 (July 2012): Shutdown brand ads on Google

Brand Search Terms Eyeconometrics



(a) MSN Test

(b) Google Test

- 99.5 % of click traffic is retained!

Non-Brand Search Terms

Non-Brand Search terms: queries that do not include the name of the brand

- Examples: ‘shoes’, ‘long dress’

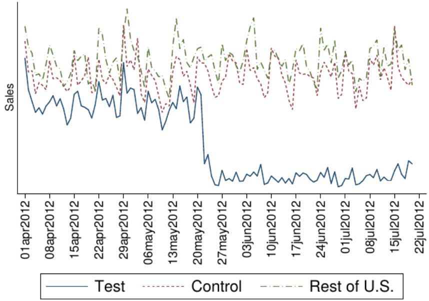
Key difference: Users might not know product is available at a advertiser’s website

Hypothesis: Non brand ads steer consumers to advertiser’s site

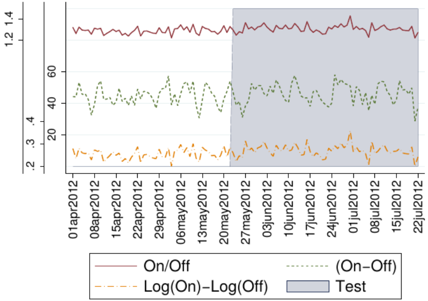
Experiment: Large scale Randomized Control Trial

- Suspend non-brand ads in 30% of all DMAs in USA
- Control vs Test Split chosen via an algorithm
- DMA: region of the US, roughly equivalent to a metro area

Non-Brand Search Terms Eyeconometrics



(a) Attributed Sales by Region



(b) Differences in Total Sales

Non-Brand Search Terms Econometrics

Method: Difference in Differences

$$\begin{aligned}\ln(\text{Sales}_{it}) = \beta_0 + \beta_1 \text{Treatment Group}_i + \beta_2 Post_t \\ + \delta \text{Treatment Group}_i \times Post_t + \text{Fixed Effects} + \varepsilon_{it}\end{aligned}$$

- i is a DMA (region) of the US
- t is time (calendar date)

Coefficient of Interest: δ

Non Brand Search Terms Results

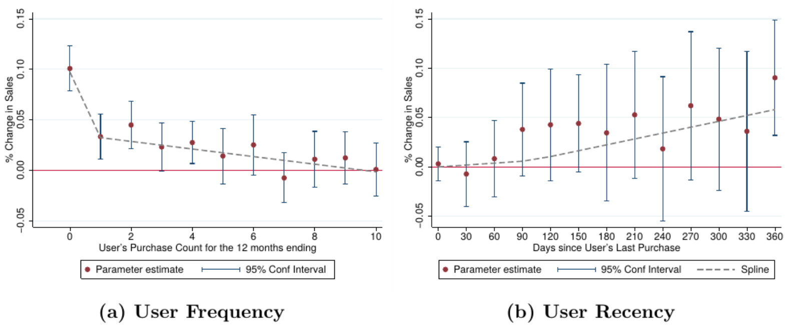
Table 1: Return on Investment

	OLS		IV		DnD	
	(1)	(2)	(3)	(4)	(5)	
Estimated Coefficient	0.88500	0.12600	0.00401	0.00188	0.00659	A
(Std Err)	(0.0143)	(0.0404)	(0.0410)	(0.0016)	(0.0056)	
DMA Fixed Effects		Yes		Yes	Yes	
Date Fixed Effects		Yes		Yes	Yes	
N	10500	10500	23730	23730	23730	
Δln(Spend) Adjustment	3.51	3.51	3.51	3.51	1	B
Δln(Rev) (β)	3.10635	0.44226	0.01408	0.00660	0.00659	C=A*B
Spend (Millions of $)	$ 51.00	$ 51.00	$ 51.00	$ 51.00	$ 51.00	D
Gross Revenue (R')	2,880.64	2,880.64	2,880.64	2,880.64	2,880.64	E
ROI	4173%	1632%	-22%	-63%	-63%	F=A/(1+A)*(E/D)-1
ROI Lower Bound	4139%	697%	-2168%	-124%	-124%	
ROI Upper Bound	4205%	2265%	1191%	-3%	-3%	

The upper panel presents regression estimates of SEM's effect on sales. Columns (1) and (2) naively regress sales on spending in the pre-experiment period. Columns (3) and (4) show estimates of spending's effect on revenue using the difference-in-differences indicators as excluded instruments. Column (5) shows the reduced form difference-in-differences interaction coefficient. The lower panel translates these estimates into a return on investment (ROI) as discussed in Section 4 and shows its 95% confidence interval.

Consumer Heterogeneity

Figure 4: Paid Search Effect by User Segment



Panel (a) shows difference-in-differences estimates and 95% confidence bands of paid search effects on sales for different user segments as defined by how many purchases were made in the previous 12 months. Panel (b) shows similar estimates where users were segmented by the time since last purchase.

Main Takeaways

- Ads served via **Brand Search terms** are**, on average**, **ineffective** at bringing clicks to site
- Ads served via **Non-Brand Search terms** are, **on average**, **do not generate sales**
- **Non-Brand Search terms might be** effective for:
    - Consumers who **do not purchase frequently** on site
    - Consumers who **haven't purchased in a long time**

Results are suggestive of Search Engine Ads being **informative**

Discussion Q:

- Are the consumers for whom ads might be effective usually the type of consumers a firm would advertise to?

Generalizability of Results?

How **generalizable** are the eBay results across different companies?

- Coviello, Gneezy and Goette (2017) run the same experiments for a 'more representative company'
    - Company: Edmunds - a large auto insurer in the US
    - Experiment: Shutdown branded keyword ads on Yahoo and Bing
        - Split markets into 'Treatment' and 'Control'
    - Analysis: Difference in Differences

Generalizability of Results?

Table 2: Difference-in-differences estimates of the treatment effects
WLS Regressions

Dependent variable: change in web-traffic category, normalized by average total web traffic in market during the baseline phase.

Dependent variable:	paid traffic		organic traffic		total traffic	
Treatment Market (=1)	-0.098***	-0.102***	0.042***	0.040***	-0.056***	-0.062***
	(0.008)	(0.003)	(0.012)	(0.011)	(0.017)	(0.012)
Fraction of paid sessions in BL		-0.756***		-0.435*		-1.191***
		(0.100)		(0.255)		(0.321)
Constant	-0.020***	0.092***	-0.077***	-0.012	-0.097***	0.080
	(0.002)	(0.015)	(0.009)	(0.040)	(0.009)	(0.050)
R^2	0.746	0.918	0.163	0.232	0.173	0.473
Obs	210	210	210	210	210	210

Notes: Heteroskedasticity-robust WLS standard errors are in parentheses. Estimates are weighted by the average total web traffic in a market during the baseline (the normalizing variable). *, **, and *** denote significance at the 10, 5, and 1 percent levels, respectively.

Result: **5.6 percentage point reduction in total traffic**

$\Longrightarrow$ search engine ads are not a "zero" effect for all firms

2/ Does Online Word of Mouth Matter?

What is Word of Mouth Marketing?

Consumer's interest in a company's product or service is reflected in their "daily dialogues"

- Why is this *new* in "social media"?
    - It isn't a new idea ...
    - The "social web" with it's increasing connectivity makes it more salient
    - ... and measurable

Types of Word of Mouth

Organic word of mouth:

- People become advocates for a product and have a desire to share their views.
- **This is our focus this week**

Amplified word of mouth:

- Marketers launch campaigns designed to encourage or accelerate WoM in existing or new communities.
- We'll come back to this later in the course – "Social Advertising"

Online versus Offline

- Distinction is always lurking in the background

Social Media Word of Mouth Matters

- Consumers now spend more than 135 mins per day on social media
  - Social media sites contain a treasure-trove of decision relevant information
  - Twitter is the main platform for opinion exchange
- Social Media fostered growing importance of WoM marketing
- Chief Marketing Officers think online WoM matters
  - … Rationalized by consumer’s trust in online info from peers (Nielsen, 2013)
  - 64% of marketing executives believe word of mouth is the most effective form of marketing
  - Only 6% say they have mastered it.

Why Word of Mouth Might Matter

Four potential mechanisms at play:

1. Awareness
2. Buzz
3. Social learning
4. Consumption complementarities

Most often we see:

- Awareness & Buzz $\rightarrow$ volume of tweets
- Social learning $\rightarrow$ sentiment in tweet's text
    - Sentiment often called valence

Online WoM & Causality

Motivation: Causal inference is particularly difficult in the realm of online WOM due to the fact that firms are not directly in control of the amount of WOM.

Specific Business Questions:

- What is the demand elasticity of demand wrt volume of posts?
- What is the mechanism through which online WoM influences choice?

How?: Natural experiment – shutdown if Sina Weibo due to political events in mainland China but not HK

- Sina Weibo $\approx$ Chinese Twitter

Empirical Approach

Industry: TV show viewership – soapies

- Not really **new** products

Data:

- TV ratings (i.e. viewership) at episode/city level in mainland China and HK
- Microblogging activity about each show

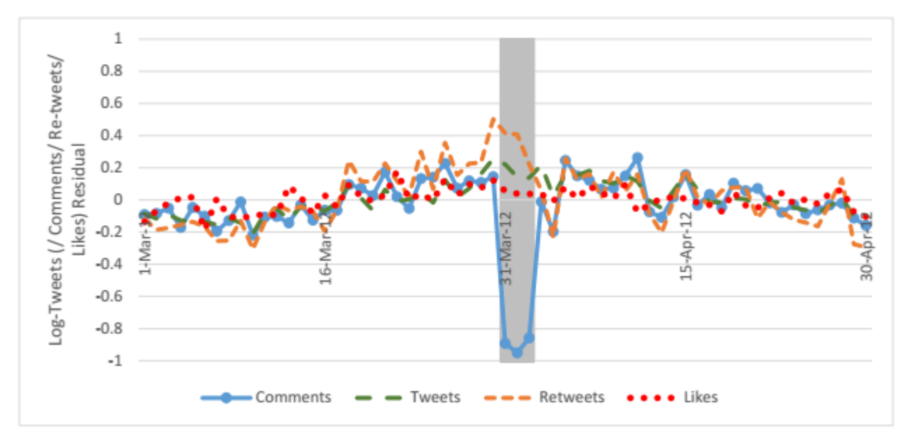
The Natural Experiment: Censorship block on Sina Weibo

- Large, random shock, unrelated to TV
- Block in mainland China, **but not** HK

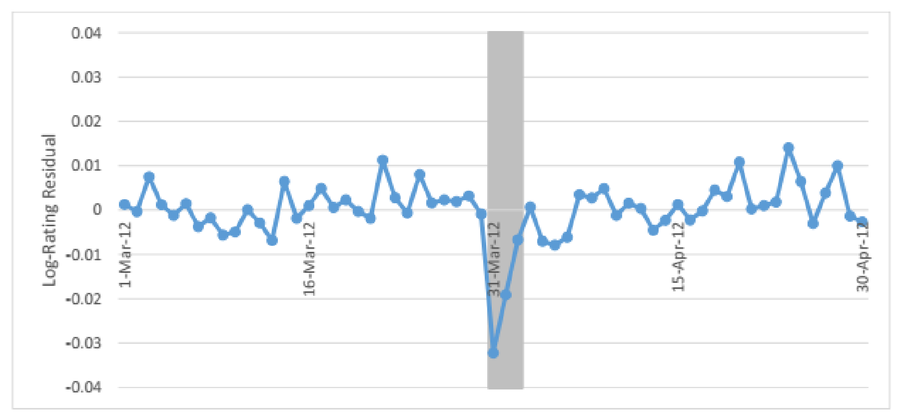
Difference in Differences Regression

$$LogRating_{jt} = \alpha Block_t + \beta Mainland_j + \delta_j Block_t \times Mainland_j + Weekday_t' \gamma + \varepsilon_{jt}$$

Graphical Evidence I



Graphical Evidence II



Diff in Diff Results

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Log Rating	Log Rating	Log Rating	Log Rating	Log Rating	Log Rating
Sample	Mainland China	HK and Mainland China	HK and Shenzhen (respective shows)	HK and Shenzhen (mainland shows)	24 Cities in Mainl. China	24 Cities in Mainl. China
Censor Dummy	-0.017***	0.005	0.002	-0.008***	-0.010	-0.008
	(0.005)	(0.010)	(0.010)	(0.002)	(0.006)	(0.006)
Mainland		-0.026**				
× Censor Dummy		(0.012)				
Shenzhen			-0.035**	-0.017*		
× Censor Dummy			(0.014)	(0.010)		
Sina Weibo Penetration					-0.027*	
× Censor Dummy					(0.014)	
Above Median Penet.						-0.016***
× Censor Dummy						(0.006)
Show FEs	Yes	Yes	Yes	Yes	Yes	Yes
Weekday Dummies	Yes	Yes	Yes	Yes	Yes	Yes
City FEs	n/a	n/a	n/a	n/a	Yes	Yes
Observations	7,899	11,427	11,427	15,798	189,576	189,576
Shows	193	325	325	193	193	193
R^2	0.881	0.964	0.951	0.774	0.479	0.479

Table 2: **Difference-in-Differences Regressions: Geographical Differences.** The unit of observation is an episode in columns (1) to (4) and an episode/city combination in columns (5) and (6). Standard errors are clustered at the show level.

What is the Mechanism?

Dependent Variable	(1) Log Rating	(2) Log Rating	(3) Log Rating	(4) Log Rating
Censor Dummy	-0.005	-0.001	-0.002	-0.002
	(0.005)	(0.007)	(0.007)	(0.007)
Medium **Daily** Activity	-0.008			
× Censor Dummy	(0.011)			
High **Daily** Activity	-0.026**			
× Censor Dummy	(0.012)			
Medium **Pre-Show** Activity		-0.007	-0.007	-0.007
× Censor Dummy		(0.010)	(0.012)	(0.012)
High **Pre-Show** Activity		0.011	0.024	0.028
× Censor Dummy		(0.020)	(0.019)	(0.019)
Medium **Post-Show** Activity		-0.007	-0.007	-0.008
× Censor Dummy		(0.009)	(0.009)	(0.009)
High **Post-Show** Activity		-0.041**	0.001	0.005
× Censor Dummy		(0.020)	(0.018)	(0.019)
Medium **Post-Show (Any) Sentiment** Comments			0.007	
× Censor Dummy			(0.014)	
High **Post-Show (Any) Sentiment** Comments			-0.060***	
× Censor Dummy			(0.016)	
Medium **Post-Show Positive Sentiment** Comments				0.017
× Censor Dummy				(0.014)
High **Post-Show Positive Sentiment** Comments				-0.039**
× Censor Dummy				(0.017)
Medium **Post-Show Negative Sentiment** Comments				-0.017
× Censor Dummy				(0.014)
High **Post-Show Negative Sentiment** Comments				-0.041**
× Censor Dummy				(0.018)
Show FEs	Yes	Yes	Yes	Yes
Day of the Week Dummies	Yes	Yes	Yes	Yes
Observations	7,899	7,899	7,899	7,899
Shows	193	193	193	193
R^2	0.881	0.881	0.881	0.881

Table 5: **Timing and Content: The Differential Impact of Weibo Activity.** The unit of observation is an episode. Standard errors are clustered at the show level.

Takeaways

- Estimated Volume elasticity: between 0.016 and 0.026
- WoM influnces demand via consumption complementarities
    - Can chat about it later online
- Managerial Implications:
    - Fostering post-show discussion
    - Doesn't appear to be sentiment effects
        - (maybe because quality is known?)
        - Does sentiment matter is a big conversation in the literature

An Alternative Approach?

How far to ”believable” numbers can get get without experimental variation?

Can we reconcile the volume vs sentiment debate?

- If we can **control** for (almost) all the omitted variables
- And impose **structure** on the consumer decision making problem
    - Substitute: Clean variation (experiment) for more mathematical modelling and assumptions
- Studied by Deer, Crawford, Chintagunta (2022)

Setting: US Movie Industry & Twitter WoM

Important Distinction for new products:

- Pre- vs Post- release volume and sentiment

Main Result - Demand Elasticities

	Estimate	Std. Error	95% CI
Opening Weekend			
Tweet stock	0.055**	0.034	[0.014, 0.144]
Pre-tweet sentiment	-0.023	0.044	[-0.127, 0.046]
Ad stock	0.023	0.125	[-0.213, 0.279]
Post-Opening			
Tweet stock	0.055**	0.033	[0.015, 0.145]
Pre-tweet sentiment	0.071	0.115	[-0.128, 0.317]
Positive sentiment change	0.153**	0.062	[0.031, 0.272]
Negative sentiment change	-0.065	0.121	[-0.378, 0.057]
Ad stock	0.335**	0.143	[0.09, 0.65]

Franchise vs Non-Franchise

	Series Movies		Non-Series Movies	
	Estimate	95% CI	Estimate	95% CI
Opening Weekend				
Tweet stock	0.154**	[0.007, 0.345]	0.017	[-0.012, 0.197]
Pre-tweet sentiment	0.042	[-0.122, 0.224]	-0.067	[-0.209, 0.046]
Ad stock	0.189	[-0.332, 0.84]	0.000	[-0.247, 0.261]
Post-Opening				
Tweet stock	0.097	[-0.07, 0.252]	0.052**	[0.014, 0.346]
Pre-tweet sentiment	0.21	[-0.294, 0.485]	0.328**	[0.042, 0.585]
Positive sentiment change	-0.096*	[-0.333, 0.022]	0.188***	[0.083, 0.357]
Negative sentiment change	-0.06	[-0.272, 0.313]	-0.341**	[-0.652, -0.052]
Ad stock	0.263	[-0.138, 0.951]	0.384**	[0.092, 0.742]

3/ Effects of Influencer Advertising Disclosure Regulations

The Influencer Market: Stylized Facts

- **Large and growing**, approx 9.7 billion USD in 2020
  - 2016: 1.7 billion USD, 2025: (expected) > 20 billion USD
- **Approx. 3/4 companies use influencer marketing** to some extent, mainly Instagram
  - **Only 65%** of those who use it, **track ROI**
- Becoming ‘centralized’ through influencer marketing agencies
  - Interesting incentive problems here...
- Large **growth in the use of “micro” influencers**
- Growing concern: **compliance issues**
  - Only approx 20% of (US/UK) posts meet regulatory guidelines

Influencers as Advertisers

Strategic Question: What do advertising disclosure regulations do to positng behavoiur of influencers and consumer content engagement?

Why Relevant?

- Sponsored influencer posts is still the “wild west” of advertising
  - i.e. unregulated, left to consumers to discern
  - Unlike most advertising markets
- Increase regulation mandated by governments …
- … But actual uptake is still low

How are we going to answer the question?: Natural Experiment on Instagram

- Introduction of disclosure laws

The Experiment & Regression

The **Natural Experiment**: Introduction of strict Ad Disclosure Laws

- **Before/After**: New laws passed in Germany in late 2016 $\rightarrow$ `Treated Periods`
- **Treated Group**: Germany influencer market $\rightarrow$ `Germany`
- **Untreated Group**: Spanish influencer market

Data: 6,000 local influencers in Spain and Germany

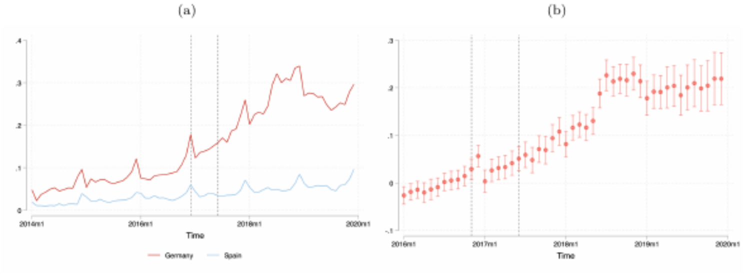
The **regression framework**:

$$y_{it} = \alpha Germany_i \times TreatedPeriods_t + \beta X_{it} + \delta_i + \delta_t + \varepsilon_{it}$$

We are interested in α

Disclosure Patterns

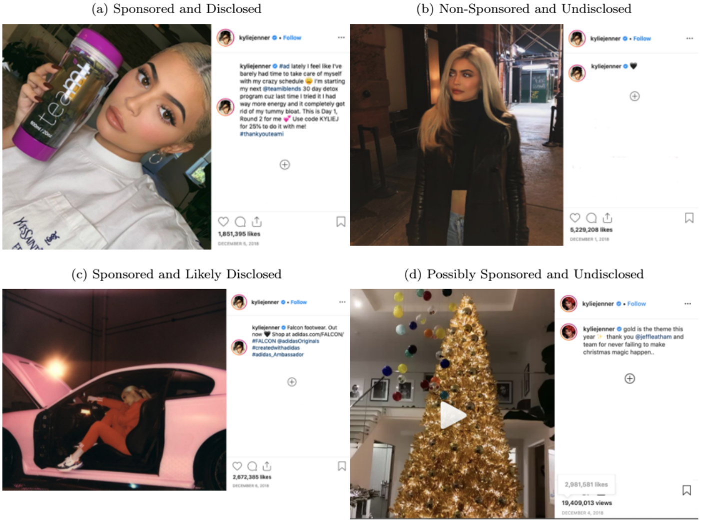
Figure 1: Disclosed Post Shares in Germany and Spain



Notes: Each line in panel (a) shows the total number of posts labelled as "disclosed" advertising over the total number of posts in month t in Germany or Spain ($\frac{NDisclosedPosts}{NPosts}$). A post is labelled as disclosed if it includes one of the disclosure words from Appendix A.3.1. CEM-matched sample of influencers is used. Panel (b) shows estimates from a difference-in-differences regression with heterogeneous monthly treatment effects and with placebo effects starting in January 2016 (Eqn. 3). The baseline periods for this regression are 2014 and 2015. 95% confidence intervals shown. The first dashed vertical line represents the initial changes to German disclosure regulations in November 2016 (see Section 2.1). The second dashed vertical line represents the first fines handed out to German influencers in mid 2017.

Detecting Disclosure?

Figure 2: Examples of Non-Sponsored, Sponsored and Disclosed Posts



Disclosure Before & After

Table 1: Influencer/Month DiD Estimates - Sponsored Share

	(1)	(2)	(3)
Outcome:	Predicted Sponsored Shares		
Classifier:	SGD L1	Manual	SDD L1 + Manual
Germany × Treated Period	0.046***	0.019**	0.045***
	(0.008)	(0.009)	(0.008)
Pre-Treatment Mean	0.382	0.452	0.216
Observations	67,235	67,235	67,235
R-squared	0.522	0.556	0.571
Outcome:	Predicted Sponsored Shares| Non-Disclosure		
Classifier:	SGD L1	Manual	SDD L1 + Manual
Germany × Treated Period	0.025***	0.004	0.019**
	(0.008)	(0.010)	(0.008)
Pre-Treatment Mean	0.373	0.444	0.207
Observations	65,984	65,984	65,984
R-squared	0.474	0.519	0.515
Country Controls	YES	YES	YES
Influencer FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Account Age FE	YES	YES	YES
Account Age × First-Account-Year FE	YES	YES	YES

- Increase in share of disclosed sponsored content, between 2 and 4.6 percentage points
- PLUS an unintended consequence: share of undiscloses also increases

Disclosure Before & After

Table 2: Influencer/Month DiD Estimates - Additional Post Content Outcomes

Outcome:	(1) Disclosed Share	(2) N Posts
Germany × Treated Period	0.091***	0.974
	(0.007)	(0.777)
Pre-Treatment Mean	0.0509	19.29
Country Controls	YES	YES
Influencer FE	YES	YES
Year-Month FE	YES	YES
Account Age FE	YES	YES
Account Age × First-Account-Year FE	YES	YES
Observations	67,235	67,235
R-squared	0.576	0.579

Engagement Metrics

Table 3: Influencer/Month DiD Estimates - Engagement

Outcome:	(1) Mean N Likes	(2) Mean N Comments	(3) Mean N Followers
Germany × Treated Period	-483.217***	-22.663***	-4,693
	(157.693)	(7.232)	(8,275)
Pre-Treatment Mean	769.1	17.10	76,790
Country Controls	YES	YES	YES
Influencer FE	YES	YES	YES
Year-Month FE	YES	YES	YES
Account Age FE	YES	YES	YES
Account Age × First-Account-Year FE	YES	YES	YES
Observations	67,235	67,235	14,165
R-squared	0.637	0.251	0.906

Takeaways

- Advertising disclosure regulation has real effects
    - Disclosure rates of sponsored posts increase after regulation introduced
    - Important given skepticism about its impact
- But, engagement per post falls (on average)
- What we still don’t know: Does the type of content an influencer posts change after the introduction of regulation?

4/ BONUS: Do influencers impact product demand?

Motivation

Strategic Question: What is the effectiveness of influencer marketing on generating consumer demand?

Why relevant?

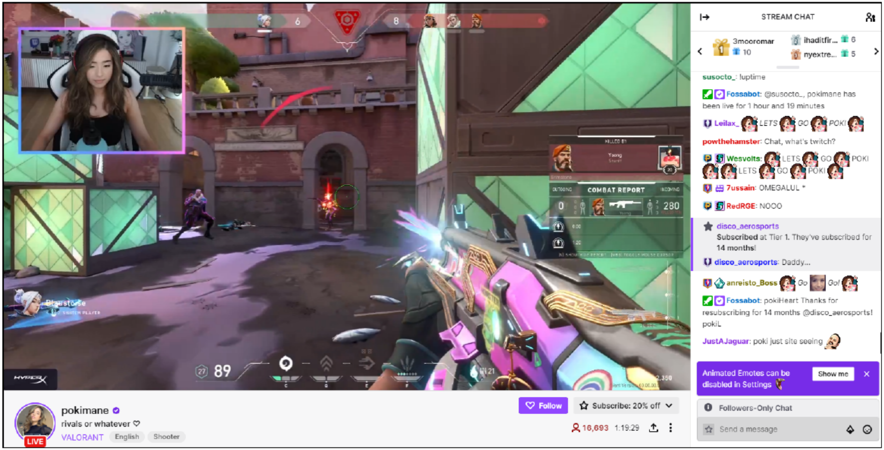
- (Over-) Excitement about influencer marketing
- Limited evidence beyond anecdotes
- It's **hard** to measure, but **worth** trying

Methods:

- No experiment but **clever empirical strategy**
- $\implies$ think of most estimates as **causal**

Application: Twitch & Video Games

Twitch


STREAM CHAT
susocto_: !uptime
Fossabot: @susocto_, pokimane has been live for 1 hour and 19 minutes
Leilax_: LETS GO POKI
powthehamster: Chat, what's twitch?
Wesvolts: LETS GO POKI LETS GO POKI
7ussain: OMEGALUL
RedRGE: NOOO
disco_aerosports
Subscribed at Tier 1. They've subscribed for 14 months!
disco_aerosports: Daddy...
anreisto_Boss: Go Go!
Fossabot: pokiHeart Thanks for resubscribing for 14 months @disco_aerosports! pokiL
JustAJaguar: poki just sits seeing
Animated Emotes can be disabled in Settings
Show me
Followers-Only Chat
Send a message
COMBAT REPORT
pokimane
rivals or whatever
VALORANT
English
Shooter
LIVE
Follow
Subscribe: 20% off
16,693
1:19:29

What we want to know

.center[How **Viewership of streams impacts demand**]

- **Question:** Why won’t linear regression suffice?

The Key Idea of the Paper

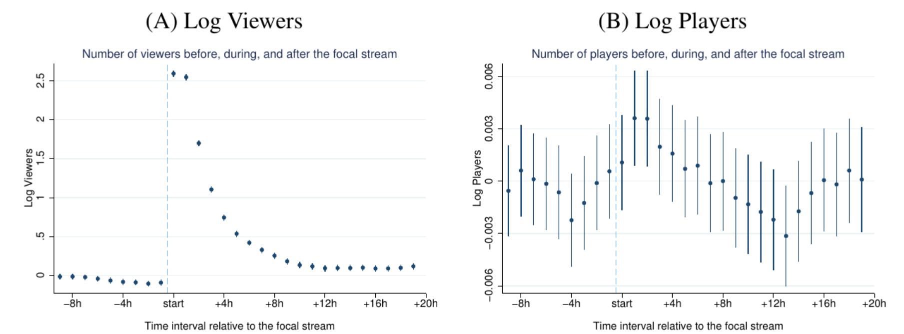
Ideal Experiment:

Activate streamers at **random times** to generate **exogenous variation** in **viewers**

Paper's "best" approximation:

- Variation in timing of when top streamers broadcast
- ... that should be independent of shocks to game popularity
- Essentially exploiting that streamers work "irregular" hours

Model Free Evidence



Observations:

- Players does not immediately peak
- Slowly returns to baseline

$\implies$ potential for short-lived effects on game usage

Empirical Strategy

Regression Equation:

$$\log(1+\text{Players}_{jt}) = \beta \log(1+V_{jt}(\delta)) + \lambda_{j,d(t)} + \mu_{j,h(t)} + \eta_t + \varepsilon_{jt},$$

where:

$$V_{jt}(\delta) = \sum_{\tau=0}^{T} \delta^{\tau} \text{Viewers}_{j,t-\tau},$$

i.e. is a weighted sum of recent viewers ...

- where the eights decay geometrically, over time
- Set T = 72 hours (!!)

$\delta \rightarrow$ persistant effect of viewership

- "carryover" relative to "immediate" effects

Empirical Strategy

How do they use streamer's timing of broadcast?

- Technically: **instrumental variables**
- To overcome omitted variable bias and simultaneity
- See diagram in class ...

What are the instruments?

- **Number of top streamers broadcast game** j **at time** t
  - Measure this for **each of the last 12 hours**

Main Results

Table 3: **The effect of Twitch viewership on video game usage**

		OLS	2SLS	2SLS
Streaming Elasticity	β	0.561	0.013	0.027
		(0.002)	(0.002)	(0.001)
Persistence Parameter	δ			0.712
				(0.060)
Game-Date FE		No	Yes	Yes
Game–Hour-of-Day FE		No	Yes	Yes
Time FE		No	Yes	Yes
First-Stage F-Statistic			654.7	118.6
Observations		3,257,904	3,257,904	3,257,904

Column 1 shows results from an OLS regression that fixes the persistence parameter to zero ($\delta = 0$) and does not control for any fixed effects. Columns 2-3 show results from our main specification in (1), without the persistence parameter (column 2) and with this parameter (column 3). The last two rows show the first-stage F-statistic for excluded instruments as well as the number of game-time period combinations used to estimate each model. We estimate all three models using the full sample of 599 games. Bootstrap standard errors are clustered at the game-date level.

Findings:

- **Small**, positive & statistically significant effects $\rightarrow \beta$
    - Interpret!
- Effect **persists over time** $\rightarrow \delta$

Heterogeneity in Effect Sizes

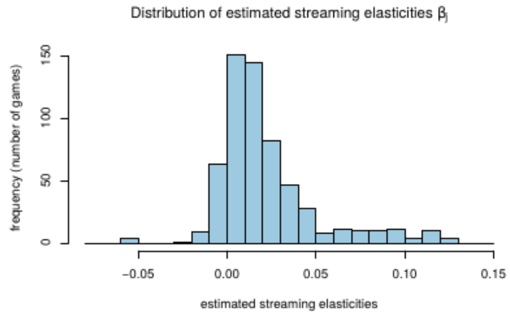


Figure 4: **Distribution of streaming elasticities.** This figure visualizes the distribution of estimated streaming elasticities $\hat{\beta}(x_j)$ from the generalized random forests (GRF). See Section 4 and Appendix D.2 for details of our GRF estimation procedure.

Heterogeneity Across Products

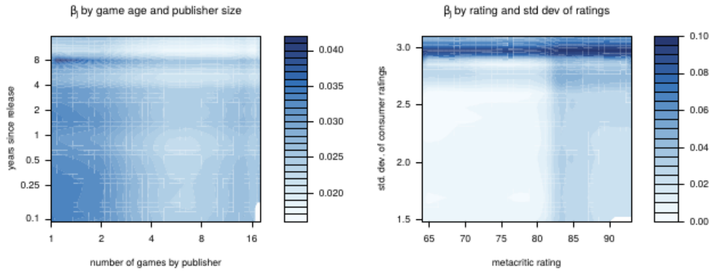


Figure 5: **Estimated streaming elasticities from Generalized Random Forests.** These graphs visualize the estimated function $\hat{\beta}(x_j)$ for two game attributes at a time, while holding all other attributes x_j fixed at their average levels.

Largest Effects:

- **Small** publishers
- **High variance** in ratings
- **Older** games

Sponsored Versus Organic

Appendix Table E.1: **Streaming elasticities of sponsored and partnered streams**

Parameter		Sponsored streams	Partner programs
		IV	IV
Organic stream elasticity	β^{org}	0.042	0.053
		(0.004)	(0.007)
Sponsored elasticity factor	ω	0.157	0.125
$(\beta^{spons} = \omega \cdot \beta^{org})$		(0.099)	(0.089)
Persistence parameter	δ	0.849	0.900
		(0.146)	(0.118)
Game-Date FE		Yes	Yes
Game–Hour-of-Day FE		Yes	Yes
Time FE		Yes	Yes
Observations		1,485,120	1,485,120

The table shows the estimates of β^{org}, ω, and δ from equation (6). To make the estimates of organic and sponsored elasticities comparable, we limit the estimation sample to 272 games that are sponsored on Twitch at least once. The first-stage F statistics are 62.9 and 67.1 for the log number of viewers in sponsored and non-sponsored streams, and 11.4 and 66.5 for the log number of viewers in partnered and non-partnered streams. Standard errors are clustered at the game-date level. All specifications include game-date, game–hour-of-day, and time fixed effects.

$\omega < 1 \rightarrow$ **sponsored content** effectiveness is a fraction of organic

- **15.7% as effective as organic streams!**

Cascades

Appendix Table E.2: **Estimates of the cascade effect parameters**

	Number of viewers	Number of organic streams by top streamers
	OLS	IV
Number of sponsored streams by top streamers (θ_1)	2,558.8 (373.8)	
Number of organic streams by top streamers (θ_2)	2,288.4 (59.0)	
Number of viewers in 1,000s (γ)		0.317 (0.024)
Fixed effects:	game-date, time, game-hour-of-day	game-week, time, game-day-of-week
First-Stage F-Statistic		25.1
Observations	1,487,304	3,257,904

This table shows the estimates of parameters θ_1 and θ_2 from equation (8) (column 1) and parameter γ from equation (7) (column 2). In column 1, the outcome variable is the absolute number of viewers of game j on Twitch. In column 2, the outcome variable is the number of organic (non-sponsored) live streams of game j by the top 5% streamers on Twitch in time t. The IV estimate is obtained using the IV strategy described in Section 5.1.2, which instruments viewership using the current price of the game as well as its one-day and two-day lagged prices. Standard errors are clustered at the game-date level.

Cascades: sponsored stream generates additional organic content

Cascades

Suppose a top streamer is sponsored to broadcast:

- **Directly increases viewership** by $\hat{\theta}_1$
- **Induces** an additional $\hat{\theta}_1\hat{\gamma}$ to **organic broadcasters**
- Which **generates** $\hat{\theta}_1\hat{\gamma} \times \hat{\theta}_2$ **more viewers**
- ...

$$\theta_{\text{total}} = \theta_1 + \theta_1 \cdot (\gamma\theta_2) + \theta_1 \cdot (\gamma\theta_2)^2 + \theta_1 \cdot (\gamma\theta_2)^2 + \ldots = \underbrace{2,559}_{\text{direct effect } \theta_1} + \underbrace{6,755}_{\text{cascade effect}}$$

End result (Viewers):

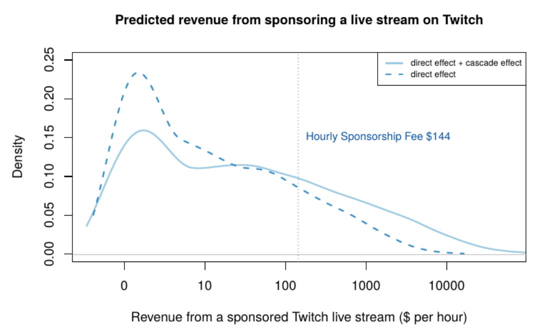
- Induces approx. 3 organic streams
- Which is 2.6 times as many viewers than sponsored stream

End result (Sales):

- 1.6 sales per 1000 viewers

Return on Investment (ROI)

$$\mathrm{ROI}_j = \frac{\Delta\mathrm{Purchase}_j \times \mathrm{Profit\ Margin}_j - \mathrm{Sponsorship\ Fee}}{\mathrm{Sponsorship\ Fee}},$$



- Median game has additional revenue of $19.50 due to sponsored streams
    - Median ROI = -95%
- 16% of games have positive ROI
- 90th percentile has ROI of 138%

Takeaways

- **Small, positive effects** of organic influencer content **on demand**
  - Similar in magnitude to OWoM volume and advertising
  - ... which helps these numbers feel credible
- **Sponsoring influencers** leads even **smaller effects** (approx 1/5th!)
  - But does lead to cascades of organic content
- **ROI** for sponsored content is **negative for 2/3rds of games**

License and Citation

Suggested Citation:

```
@misc{smwa2025_didapps,
      title={"Social Media and Web Analytics: Diff in Diff - Applicatio
      author={Lachlan Deer},
      year={2025},
      url = "https://tisem-digital-marketing.github.io/2025-smwa"
}
```

This course adheres to the principles of the Open Science Community of Tilburg University. This initiative advocates for transparency and accessibility in research and teaching to all levels of society and thus creating more accountability and impact.