DUKE FUQUA SCHOOL OF BUSINESS

Empirical Analysis for Strategy

Professor McDevitt

Winter 2021

Class 1

Course Overview

Our Objective

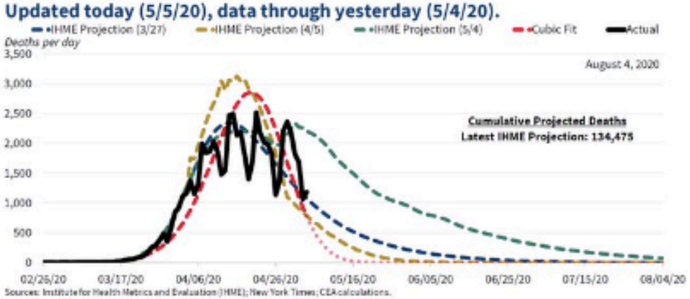
Introduce statistical methods for measuring the <u>true</u> effectiveness of strategies

Bad Stats Training → Bad Decision Making

Jason Furman @jasonfurman

This might be the lowest point in the 74 year history of the Council of Economic Advisers. The stakes on the epidemiological questions are so high that this utterly superficial and misleading "modeling" has no place whatsoever in any discussion of the government's response.

> **CEA45 Archived** @WhiteHouseCEA45
> Replying to @WhiteHouseCEA45
> To better visualize observed data, we also continually update a curve-fitting exercise to summarize COVID-19's observed trajectory. Particularly with irregular data, curve fitting can improve data visualization. As shown, IHME's mortality curves have matched the data fairly well.
>
> **United States Daily COVID-19 Deaths: Actual Data, IHME/UW Model Projections, & Cubic Fit.**
> **Updated today (5/5/20), data through yesterday (5/4/20).**

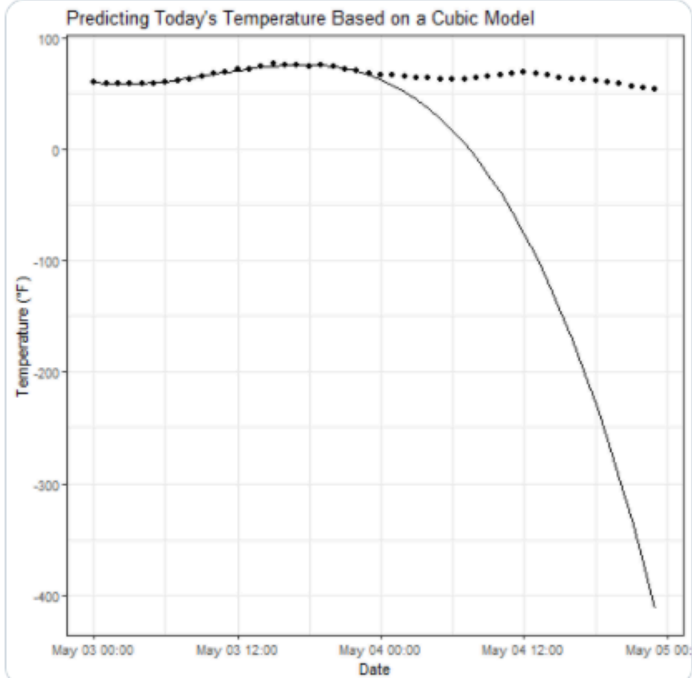


12:24 PM · May 5, 2020

♡ 3.9K   See the latest COVID-19 information on Twitter

John Voorheis @john_voorheis

According to predictions from my "cubic model" fitted on yesterday's data, today's temperature is approaching absolute zero and we are all dead, mercifully.



8:39 PM · May 4, 2020

♡ 8.3K   1.5K people are Tweeting about this

Our Approach

Answer specific causal questions, rather than provide a general model of economic variables like GDP or exchange rates

- Focus on policy questions, like “Does advertising cause an increase in sales?”
- Establish credible research designs to answer these questions
- Will not obsess over technical concerns, like functional forms or distributional assumptions for error terms

Our Topics

Experimental Design	RCTs are the gold standard for causal inference but may not always be feasible or ethical
Regression + Fixed Effects	Fixed effects regressions account for some unobserved confounds → make “all else equal”?
Matching Models	Matching models approximate an RCT using observational data by pairing treatment/control
Instrumental Variables	IV regressions establish causality by using exogenous variation to overcome confounding
Regression Discontinuity	RDD approximates RCT by comparing similar subjects who just did/did not receive treatment
Differences-in-Differences	DD compares outcomes after treatment to what would have happened otherwise

Our Structure

Recorded Videos

- Introduce concepts + provide examples
- Supplements Mastering Metrics (not substitute)
- Supplements some articles (may be substitute)

Live Lectures

- Review topics + answer questions
- No new material presented, reinforce videos
- Q+A at end, can submit questions ahead of time

Case Discussions

- Articles highlight key concepts
- Provides context for when to use each method
- Intuition more important than technical details

Our Materials

Materials for Each Class

- Videos
- Discussion Questions
- Discussion Articles
- Class Slides
- Lecture Summaries

Available on Canvas

Supplemental Articles

- Posted on Twitter @ryanmcdevitt
- Not required, solely for your "enjoyment"

Textbook

- Angrist & Pischke Mastering Metrics
- Will closely follow chapters 1-5

Your Grade

Class Participation
20%

- Participation in each class scored from 0 – 5
- Typically inverted-U shape with respect to frequency
- Final grade based on handful of best classes
- Discussion questions should guide preparation

Online Assignments
40%

- Write up 2-3 paragraphs (or less) for each
- Submit (individually) on Canvas
- Don't stress: low stakes, can work as group
- Best four of six count for your final grade

Final Exam
40%

- During final exam period
- Covers three academic articles
- Similar in style/difficulty to online assignments
- Not trying to kill anybody with this

Our Expectations

- Maintain Fuqua (virtual) classroom norms
- Participate in class discussions
- Respect classmates
- Abide by honor code at all times
- Provide feedback whenever possible (this is a brand new course!)

Announcements

OA2 due Feb 13 8:59am

OA1 graded next week

TA review session Feb 11

Agenda

Case Are Eggs Bad for You?

Lecture Experimental Design + Causal Inference

Case Does Advertising Work?

Roadmap

Next Regression Review + Fixed Effects

Are Eggs, Alcohol, & Nutrasweet Really That Bad for You?

The New York Times

How Much Alcohol Can You Drink Safely?

The Lancet published a study thought to be the most comprehensive global analysis of the risks of alcohol consumption. Its conclusion, which the media widely reported, sounded unequivocal: "The safest level of drinking is none."

But because epidemiologists can only observe — not control — the conditions in which their subjects behave, there are also a vast and unknown number of variables acting on those subjects, which means such studies can't say for certain that one variable causes another.

The New York Times

How Much Alcohol Can You Drink Safely?

Key Facts

- A recent study concluded that "The safest level of drinking is none"
- Epidemiological studies allow us to discover relationships between variables and how they change over time: they can include millions of people, far more than could be entered into an RCT

Conceptual Questions

- Moderate drinkers tend to have a lower risk of mortality than abstainers — does this mean that a certain amount of alcohol offers a "protective" effect?
- "This problem of the reference group in alcohol epidemiology affects everything" — what does this mean?
- How did the Lancet study construct its control group? Why is that important?
- Does removing "former drinkers" from the study adequately resolve the issue of selection bias?
- Is this study conclusive? What's missing?

The New York Times

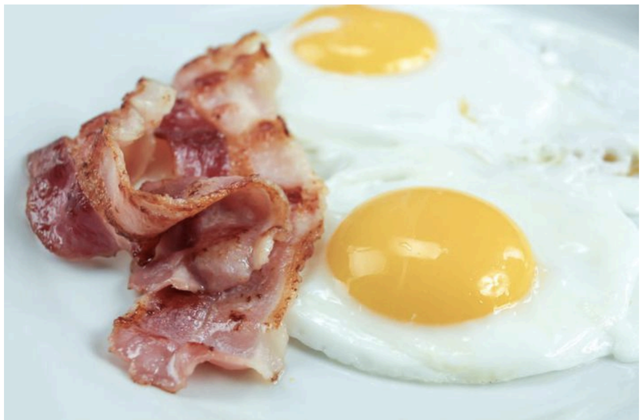
How Much Alcohol Can You Drink Safely?

Over the years, compared with abstinence, moderate drinking has been associated with conditions it couldn't logically protect against: a lower risk of deafness, hip fractures, the common cold and even alcoholic liver cirrhosis. All of which advances a conclusion that health determines drinking rather than the other way around.

Then again, it might be your innate biological resilience that kept you healthy enough to drink. The data still can't say.

Bottom line: are probably ok (in moderation)

Three or more eggs a week increase your risk of heart disease and early death, study says



People who eat an added three or four eggs a week or 300 mg of dietary cholesterol per day, have a higher risk of both heart disease and early death compared with those who eat fewer eggs, new research finds.

A potential reason for inconsistent results in the past was the fact that other studies did not take into account that egg consumption may be related to other unhealthy behaviors, such as low physical activity, smoking and an unhealthy diet.

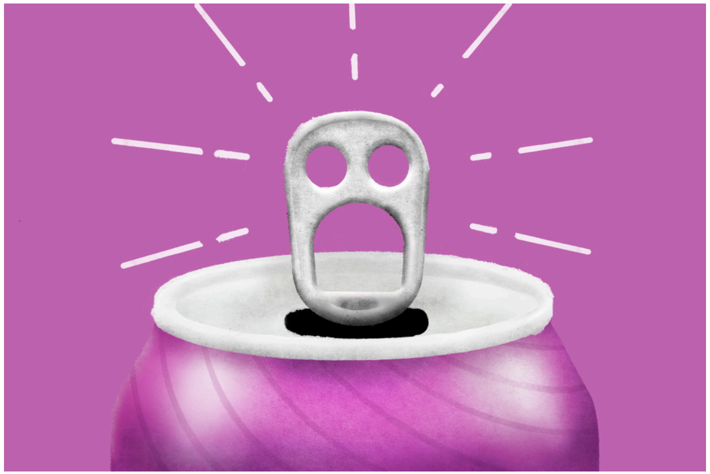
Eating Eggs Correlated with Doing Other “Bad” Things

eFigure 1. Association Between Each Additional 300 mg of Dietary Cholesterol Consumed per Day and Coronary Heart Disease

	Hazard ratio (95% CI)
Model 1	1.13 (1.04-1.23)
Model 2	1.10 (0.98-1.23)
Model 3	1.12 (1.00-1.26)
Model 2 + nutrients	
SFA	1.17 (1.03-1.33)
SFA, unsaturated fat, trans fat	1.19 (1.04-1.35)
Animal protein	1.04 (0.90-1.20)
Fiber	1.12 (0.99-1.26)
Sodium	1.08 (0.97-1.22)
All previous nutrients	1.13 (0.97-1.32)
Model 2 + cholesterol-containing foods	
Eggs	1.04 (0.84-1.29)
Unprocessed red meat	1.09 (0.97-1.23)
Processed meat	1.10 (0.98-1.24)
Poultry	1.09 (0.97-1.23)
Fish	1.10 (0.98-1.23)
Total dairy products	1.10 (0.98-1.23)
Eggs, total red meat*	0.99 (0.78-1.25)
Poultry, fish, total dairy products	1.09 (0.97-1.22)

The New York Times on July 27, 2015

The Evidence Supports Artificial Sweeteners Over Sugar



The initial studies showed that aspartame didn't cause cancer in animals, so it was deemed safer than saccharin. But in 1996, a study was published with the title "Increasing Brain Tumor Rates: Is There a Link to Aspartame?"

Most people ignored the question mark. Instead, they noted that the paper stated that (1) brain cancer had become more common from 1975 to 1992 and (2) that more people had started consuming aspartame recently.

The New York Times on Sept 6, 2019

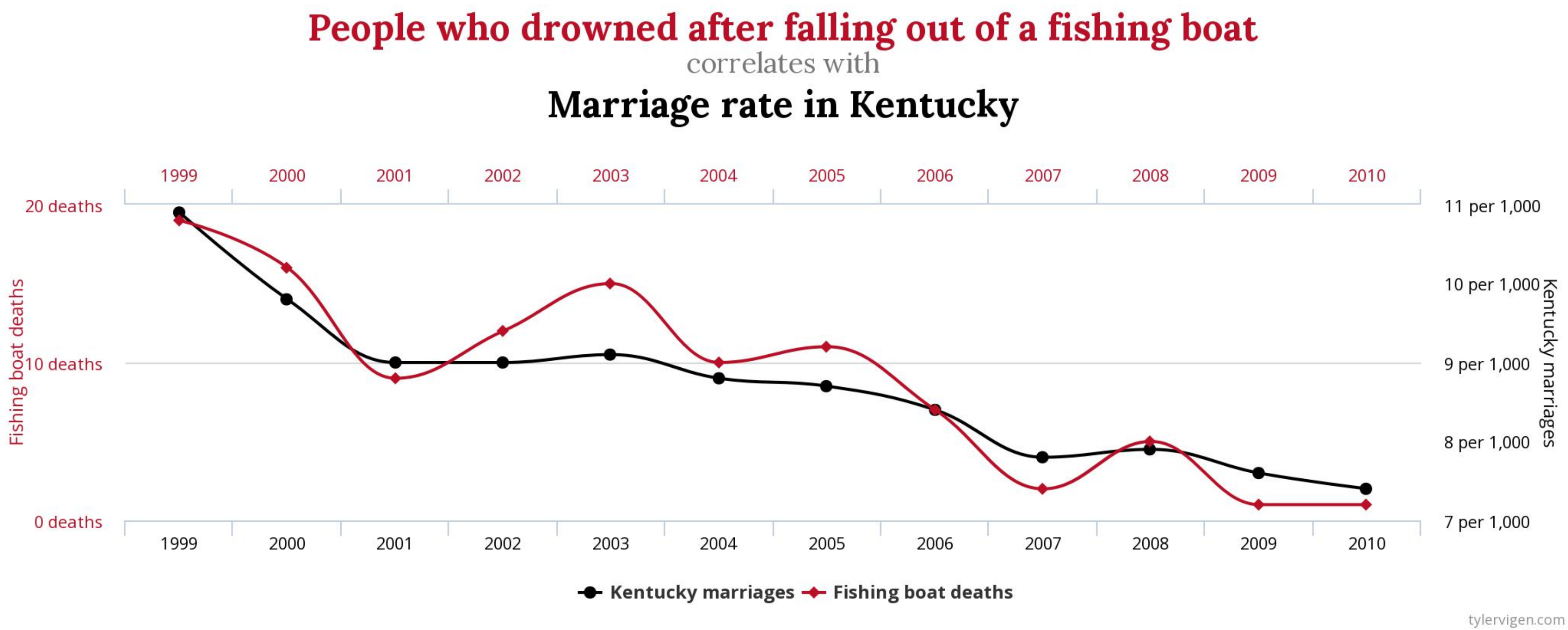
Death by Diet Soda?

A new study found prodigious consumers of artificially sweetened drinks were 26 percent more likely to die prematurely than those who rarely drank sugar-free beverages.

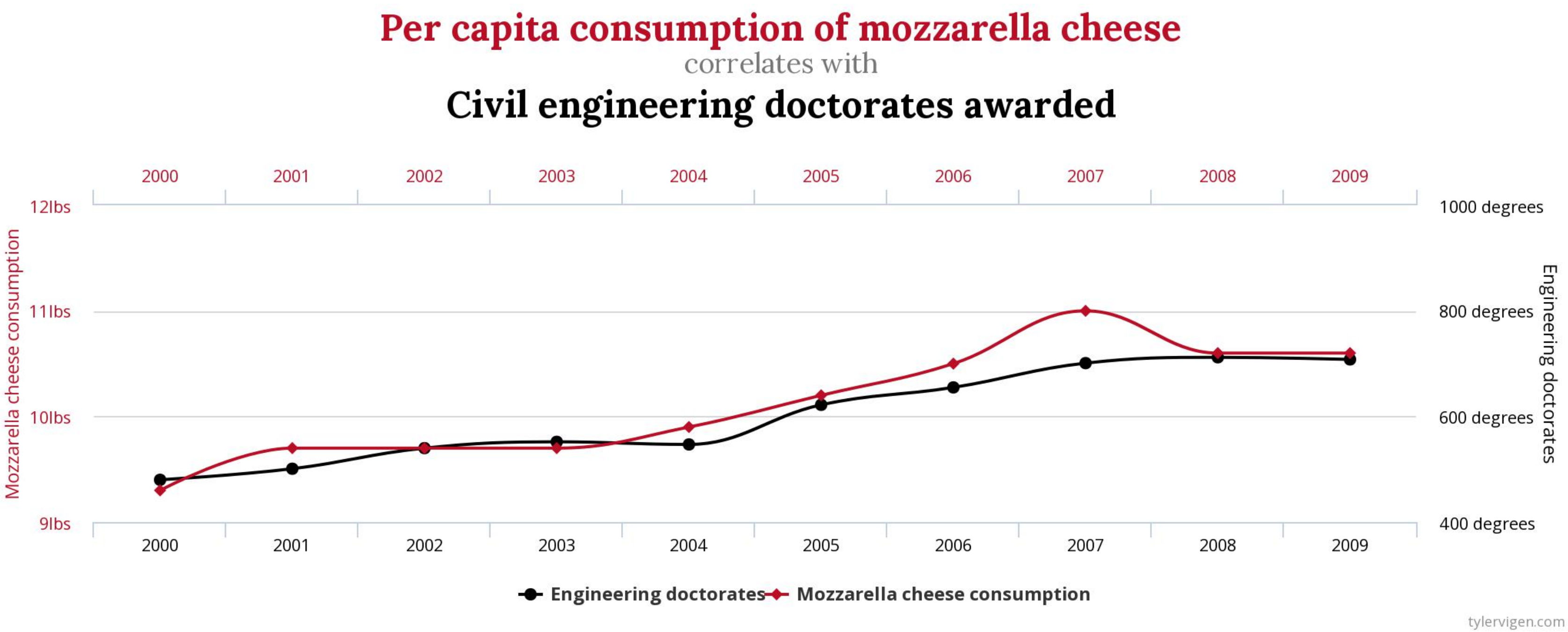
The problem, experts say, is that these studies have been unable to resolve a key question: Does consuming drinks sweetened with aspartame or saccharin harm your health? Or could it be that people who drink lots of Diet Snapple or Sprite Zero lead a more unhealthy lifestyle to begin with?

“Maybe it’s just that people with an increased risk of mortality, like those with overweight or obesity, are choosing to drink diet soda but, in the end, this doesn’t solve their weight problem and they die prematurely.”

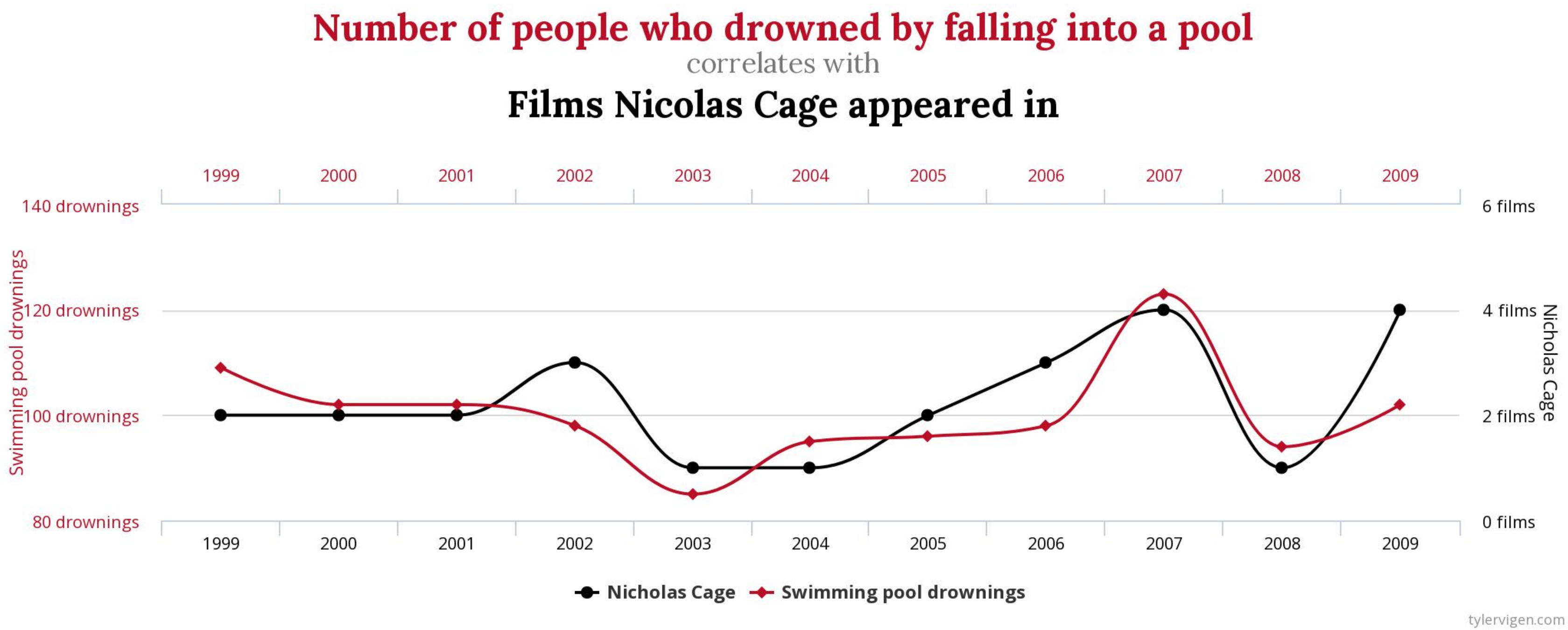
Beware of Spurious Correlations



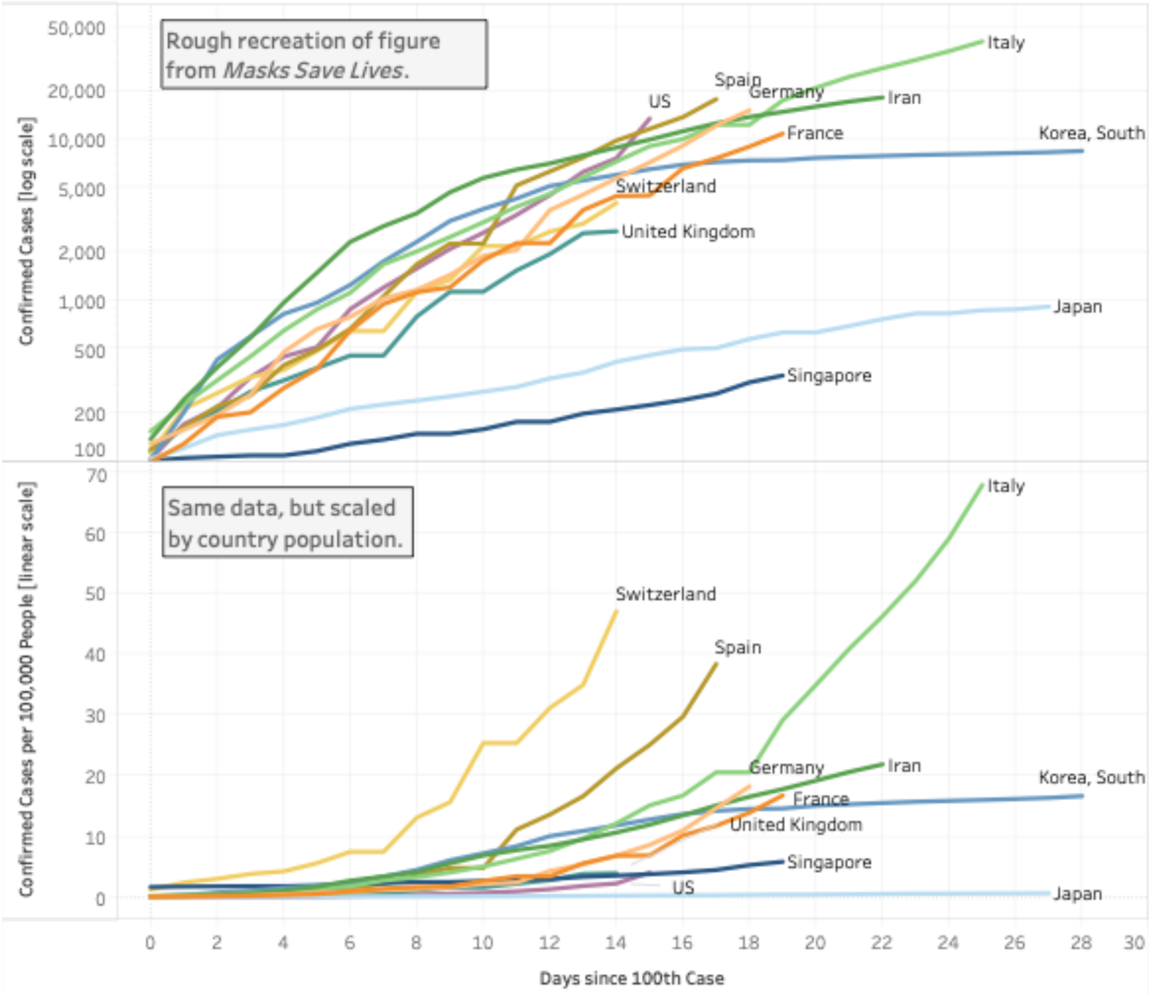
Beware of Spurious Correlations



Beware of Spurious Correlations



Beware of Spurious Correlations: Wearing Masks Stops Covid?



Not masking a lot

Growth of covid-19, by country or region

First 60 days after reaching 30 confirmed cases

Region/ country	Total cases	Avg. daily growth rate, %	Measures taken
Beijing	558	4.7	Lockdown, masks
Hong Kong	989	5.6	Masks
Japan	4,618	6.9	Masks
S. Korea	10,635	10.3	Masks
Germany	158,758	14.5	Lockdown
France	164,589	15.0	Lockdown
Britain	171,253	15.2	Lockdown
America	903,882	17.7	Lockdown

Sources: De Kai et al.; Johns Hopkins University CSSE; *The Economist*

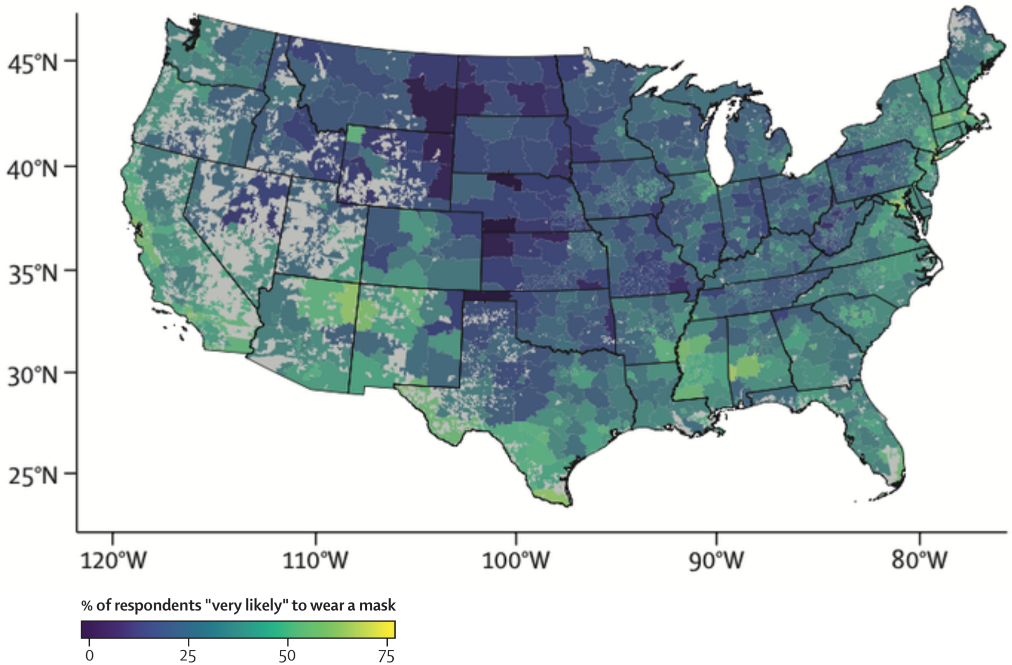
The Economist

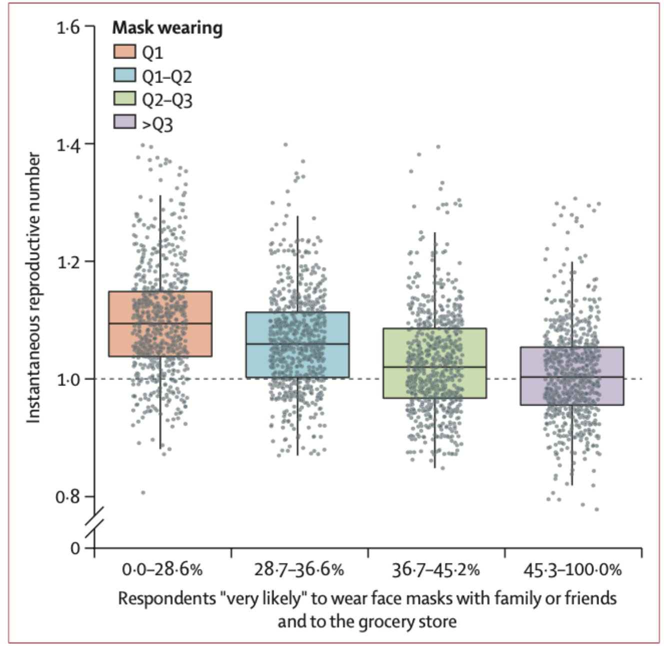
Beware of Spurious Correlations: Wearing Masks Stops Covid?

Mask-wearing and control of SARS-CoV-2 transmission in the USA: a cross-sectional study

Benjamin Rader, MPH • Laura F White, PhD • Michael R Burns • Jack Chen, PhD • Joseph Brilliant, MBA • Jon Cohen, MA • et al. Show all authors

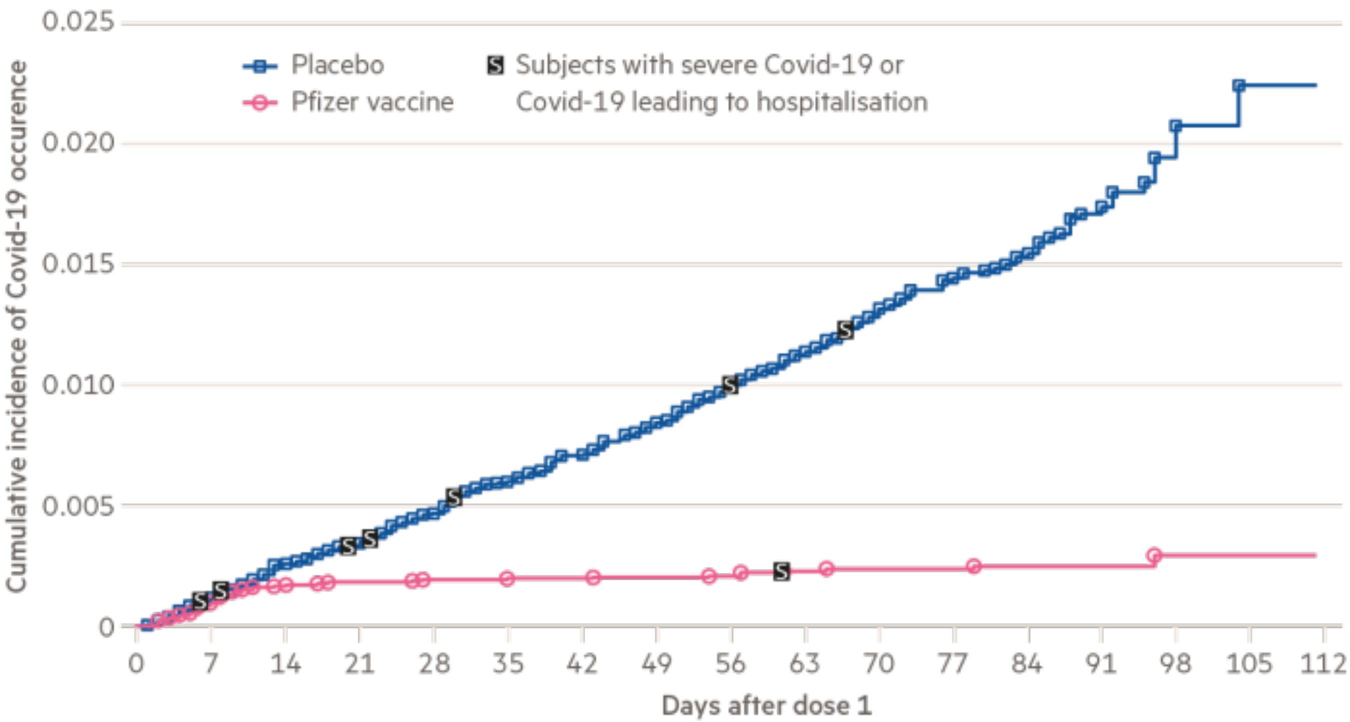
Open Access • Published: January 19, 2021 • DOI: https://doi.org/10.1016/S2589-7500(20)30293-4





RCTs Used to Distinguish Causation from Correlation

Covid-19 cases in the placebo group overtake the vaccine group soon after first dose



Source: Pfizer/BioNTech

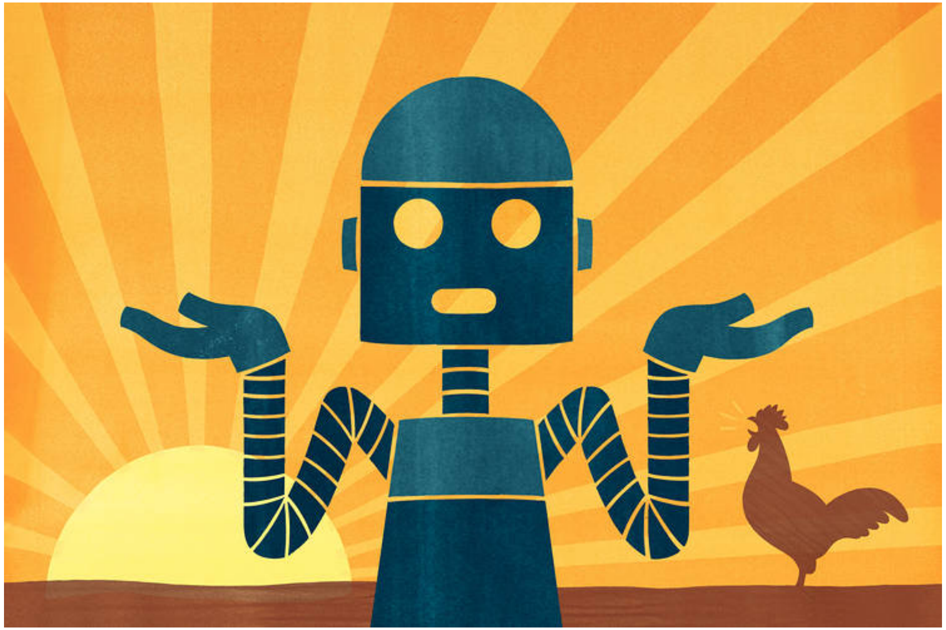
Experimental Design
+ Causal Inference

Do Roosters Make the Sun Rise, Or The Other Way Around?

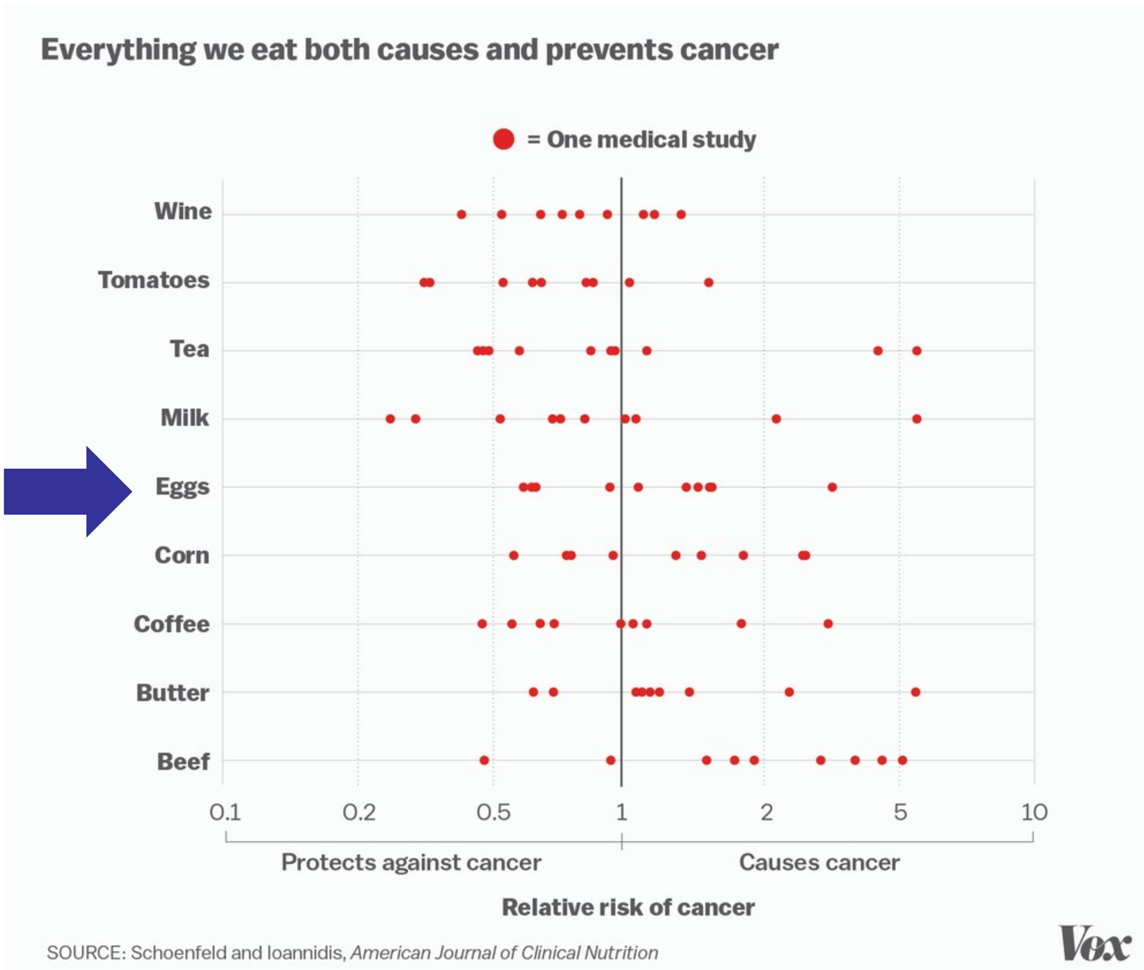
THE WALL STREET JOURNAL.

AI Can't Reason Why

The current data-crunching approach to machine learning misses an essential element of human intelligence



Many Health Studies Confuse Correlation with Causation



Goal: Distinguish Between Cause and Effect

How do we know if X causes Y?

- Most important questions we want to answer with data are causal
- Need to understand common research designs for uncovering causality when experiments aren't possible
- Important not to get hoodwinked by causal claims based on mere correlations
- We often encounter correlations that aren't causal (e.g., roosters → sun rise)

Causal inference, the main topic of this course, uses various statistical methods to deduce that one thing causes another

Working Definition of Causality

We say that X **causes** Y if an intervention that changes X (without changing anything else) results in a change in Y as well

- Does not mean that X is the only thing that causes Y
  - Smoking causes lung cancer, but not everyone with lung cancer gets it from smoking
- Does not mean that X invariably causes Y
  - Some smokers never get lung cancer, but smoking does elevate the risk

Using Data to Infer Causality Is Difficult

We often don’t observe the counterfactual outcome, so we may struggle to conclude that X truly caused Y (maybe W caused Y, or even Y caused X)

- Consider the example of health insurance and health outcomes: we want to know if having health insurance caused my health to improve
- But I’ve always had health insurance, so we don’t know if my health is good from having health insurance or some other factor, like genes, exercise, etc.
- We could compare my health to someone who doesn’t have insurance, but we might be different for other, hard to measure reasons (I’m one of a kind)

Main goal of causal inference: approximate as closely as possible what Y would have been if X had been different → the counterfactual

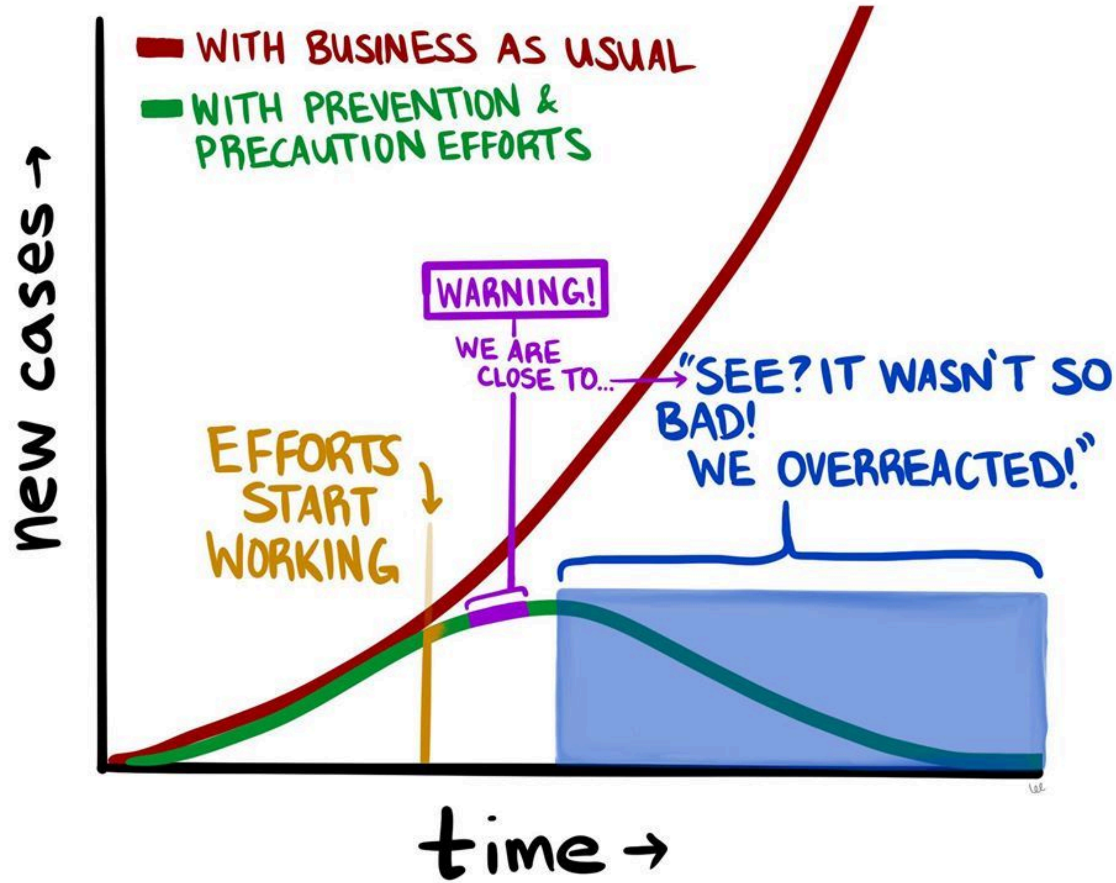
Covid Lockdown Counterfactual

Kurt Andersen @KBAndersen · Apr 15, 2020

[redacted] (Heritage, Club for Growth founder, ex WSJ editorial board) Stephen Moore: "You find the Zip codes that don't have the disease, and you get those things opened up. They should probably have never been shut down. This isn't rocket science."

Show this thread

Covid Lockdown Counterfactual



Covid Lockdown Counterfactual

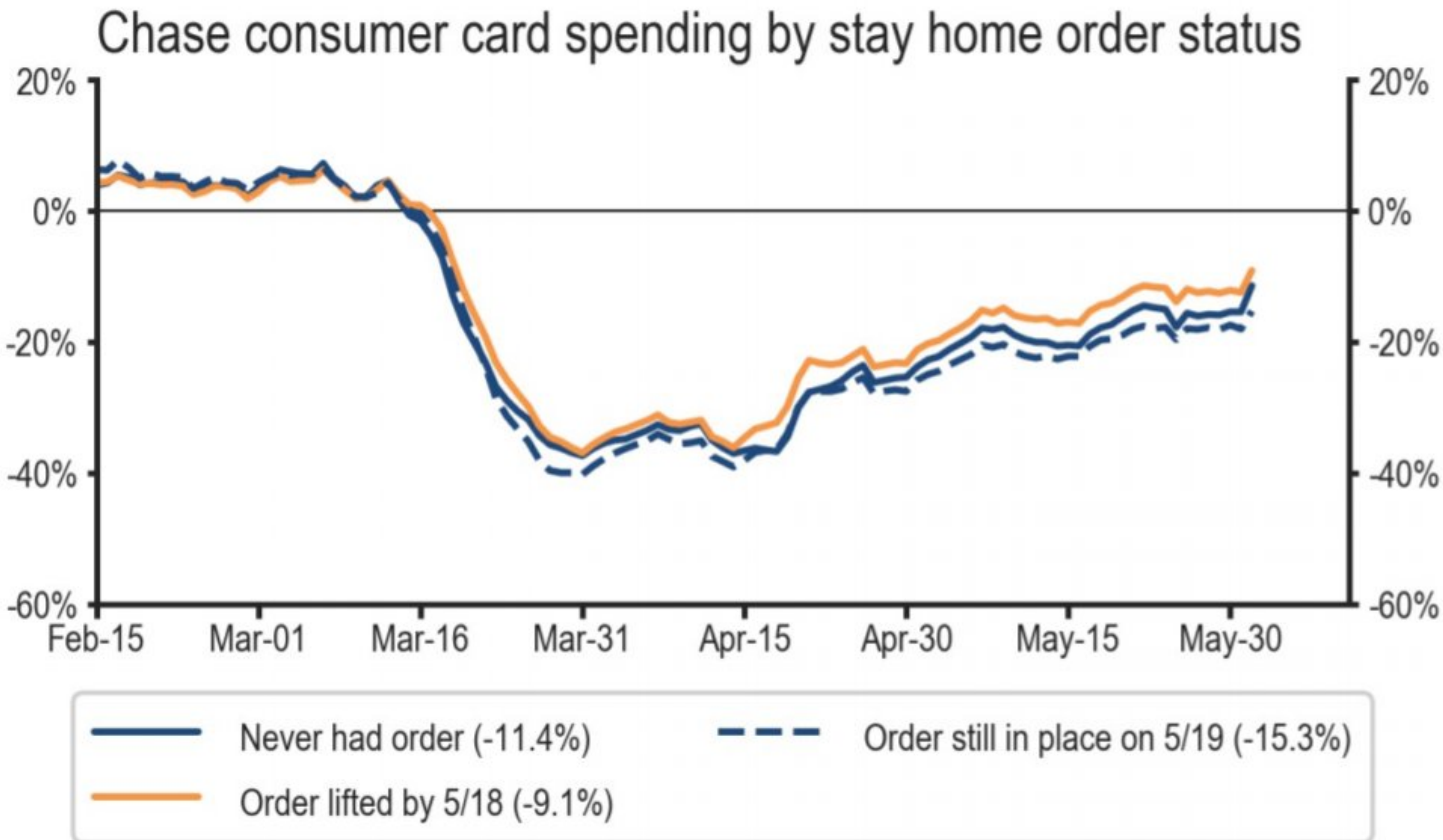
Daniel Bergstresser @dbergstresser · Apr 15, 2020

It's crazy that Deion Sanders made millions of dollars a year for playing cornerback on a side of the field that nobody ever threw to.

> **Kurt Andersen** @KBAndersen · Apr 15, 2020
>
> [redacted] (Heritage, Club for Growth founder, ex WSJ editorial board) Stephen Moore: "You find the Zip codes that don't have the disease, and you get those things opened up. They should probably have never been shut down. This isn't rocket science."
>
> Show this thread

10 113 507

Covid Lockdown Counterfactual



Potential Outcomes Example

TABLE 1.2
Outcomes and treatments for Khuzdar and Maria

	Khuzdar Khalat	Maria Moreño
Potential outcome without insurance: Y_{0i}	3	5
Potential outcome with insurance: Y_{1i}	4	5
Treatment (insurance status chosen): D_i	1	0
Actual health outcome: Y_i	4	5
Treatment effect: $Y_{1i} - Y_{0i}$	1	0

K has insurance & health = 4, M does not have insurance & health = 5
→ naive conclusion that insurance causes worse health!

Potential Outcomes and Selection Bias

Health insurance is not randomly assigned (both K & M choose whether or not to be covered) and we don't observe the counterfactual outcome (K not buying insurance, M buying insurance)

- In the data we only observe $Y_{1,K}$ and $Y_{0,M}$ so it looks like having health insurance results in worse health outcomes: $Y_{1,K} - Y_{0,M} = 4 – 5 = -1$

- This is what a naive OLS regression would tell us due to selection bias
  - Occurs because we haven't made "all else equal"
  - Also see this in Table 1.1: wealthier people more likely to have insurance

$$\underbrace{(Y_{1,K} - Y_{0,K})}_{\text{Causal Effect}} + \underbrace{(Y_{0,K} - Y_{0,M})}_{\text{Selection Bias}} = (4 - 3) + (3 - 5) = -1$$

Health Insurance Is Correlated with Health Outcomes

Table 1.1
Health and demographic characteristics of insured and uninsured couples in the NHIS

	Husbands			Wives		
	Some HI (1)	No HI (2)	Difference (3)	Some HI (4)	No HI (5)	Difference (6)
A. Health						
Health index	4.01 [.93]	3.70 [1.01]	.31 (.03)	4.02 [.92]	3.62 [1.01]	.39 (.04)
B. Characteristics						
Nonwhite	.16	.17	−.01 (.01)	.15	.17	−.02 (.01)
Age	43.98	41.26	2.71 (.29)	42.24	39.62	2.62 (.30)
Education	14.31	11.56	2.74 (.10)	14.44	11.80	2.64 (.11)
Family size	3.50	3.98	−.47 (.05)	3.49	3.93	−.43 (.05)
Employed	.92	.85	.07 (.01)	.77	.56	.21 (.02)
Family income	106,467	45,656	60,810 (1,355)	106,212	46,385	59,828 (1,406)
Sample size	8,114	1,281		8,264	1,131	

Health Insurance Is Correlated with Health Outcomes

TABLE 1.1
Health and demographic characteristics of insured and uninsured couples in the NHIS

	Husbands			Wives		
	Some HI (1)	No HI (2)	Difference (3)	Some HI (4)	No HI (5)	Difference (6)
A. Health						
Health index	4.01 [.93]	3.70 [1.01]	.31 (.03)	4.02 [.92]	3.62 [1.01]	.39 (.04)
B. Characteristics						
Nonwhite	.16	.17	−.01 (.01)	.15	.17	−.02 (.01)
Age	43.98	41.26	2.71 (.29)	42.24	39.62	2.62 (.30)
Education	14.31	11.56	2.74 (.10)	14.44	11.80	2.64 (.11)
Family size	3.50	3.98	−.47 (.05)	3.49	3.93	−.43 (.05)
Employed	.92	.85	.07 (.01)	.77	.56	.21 (.02)
Family income	106,467	45,656	60,810 (1,355)	106,212	46,385	59,828 (1,406)
Sample size	8,114	1,281		8,264	1,131	

But Other Factors Correlated with Health Outcomes Too...

Table 1.1
Health and demographic characteristics of insured and uninsured couples in the NHIS

	Husbands			Wives		
	Some HI (1)	No HI (2)	Difference (3)	Some HI (4)	No HI (5)	Difference (6)
			A. Health			
Health index	4.01 [.93]	3.70 [1.01]	.31 (.03)	4.02 [.92]	3.62 [1.01]	.39 (.04)
			B. Characteristics			
Nonwhite	.16	.17	−.01 (.01)	.15	.17	−.02 (.01)
Age	43.98	41.26	2.71 (.29)	42.24	39.62	2.62 (.30)
Education	14.31	11.56	2.74 (.10)	14.44	11.80	2.64 (.11)
Family size	3.50	3.98	−.47 (.05)	3.49	3.93	−.43 (.05)
Employed	.92	.85	.07 (.01)	.77	.56	.21 (.02)
Family income	106,467	45,656	60,810 (1,355)	106,212	46,385	59,828 (1,406)
Sample size	8,114	1,281		8,264	1,131	

Which way does the bias go?

Challenges with Using Observational Data for Causal Inference

Using observational data to determine causality comes with many challenges

- Hard to ensure "all else equal" between treatment and control groups
  - So may not have the right counterfactual
- Can try to correct for observable differences using controls in a regression
  - Like controlling for age, education, and income in previous table
- But probably won't have data for every important factor
  - Omitted variable bias (coming next class)

Throughout the course we will attempt to approximate RCTs with research designs that use various statistical methods to analyze observational data

RCTs: The Gold Standard for Causal Inference

Randomized Controlled Trials (RCTs) are the ideal way to estimate causal effects

- If we randomly assign X and measure the resulting Y, any differences in Y must have been caused by the differences in X
  - All else is equal across the treatment and control groups by design
  - We get to see the true counterfactual Y for each value of X
- But in many cases RCTs aren't possible
  - Could be costly, impractical, or unethical

RAND + Oregon Medicaid RCTs for Health Care Demand

RAND Health Insurance Experiment (HIE)

The RAND HIE randomly assigned 2,000 families from six US cities into four different insurance contracts that varied by level of coinsurance

- Coinsurance set at 0%, 25%, 50%, 95%
  - It's the % of total expenses paid by the patient
  - Reflects the price each family pays when using health care
- Run from 1974-1982

Why is an RCT necessary to measure demand for health care?

RAND Health Insurance Experiment Setup

Table 1.3
Demographic characteristics and baseline health in the RAND HIE

	Means	Differences between plan groups			
	Catastrophic plan (1)	Deductible – catastrophic (2)	Coinsurance – catastrophic (3)	Free – catastrophic (4)	Any insurance – catastrophic (5)
A. Demographic characteristics					
Female	.560	−.023	−.025	−.038	−.030
		(.016)	(.015)	(.015)	(.013)
Nonwhite	.172	−.019	−.027	−.028	−.025
		(.027)	(.025)	(.025)	(.022)
Age	32.4	.56	.97	.43	.64
	[12.9]	(.68)	(.65)	(.61)	(.54)
Education	12.1	−.16	−.06	−.26	−.17
	[2.9]	(.19)	(.19)	(.18)	(.16)
Family income	31,603	−2,104	970	−976	−654
	[18,148]	(1,384)	(1,389)	(1,345)	(1,181)
Hospitalized last year	.115	.004	−.002	.001	.001
		(.016)	(.015)	(.015)	(.013)
B. Baseline health variables					
General health index	70.9	−1.44	.21	−1.31	−.93
	[14.9]	(.95)	(.92)	(.87)	(.77)
Cholesterol (mg/dl)	207	−1.42	−1.93	−5.25	−3.19
	[40]	(2.99)	(2.76)	(2.70)	(2.29)
Systolic blood pressure (mm Hg)	122	2.32	.91	1.12	1.39
	[17]	(1.15)	(1.08)	(1.01)	(.90)
Mental health index	73.8	−.12	1.19	.89	.71
	[14.3]	(.82)	(.81)	(.77)	(.68)
Number enrolled	759	881	1,022	1,295	3,198

Randomization was effective
→ groups appear balanced
on observable characteristics

RAND Health Insurance Experiment Outcomes

TABLE 1.4
Health expenditure and health outcomes in the RAND HIE

	Means	Differences between plan groups			
	Catastrophic plan (1)	Deductible – catastrophic (2)	Coinsurance – catastrophic (3)	Free – catastrophic (4)	Any insurance – catastrophic (5)
A. Health-care use					
Face-to-face visits	2.78	.19	.48	1.66	.90
	[5.50]	(.25)	(.24)	(.25)	(.20)
Outpatient expenses	248	42	60	169	101
	[488]	(21)	(21)	(20)	(17)
Hospital admissions	.099	.016	.002	.029	.017
	[.379]	(.011)	(.011)	(.010)	(.009)
Inpatient expenses	388	72	93	116	97
	[2,308]	(69)	(73)	(60)	(53)
Total expenses	636	114	152	285	198
	[2,535]	(79)	(85)	(72)	(63)
B. Health outcomes					
General health index	68.5	−.87	.61	−.78	−.36
	[15.9]	(.96)	(.90)	(.87)	(.77)
Cholesterol (mg/dl)	203	.69	−2.31	−1.83	−1.32
	[42]	(2.57)	(2.47)	(2.39)	(2.08)
Systolic blood pressure (mm Hg)	122	1.17	−1.39	−.52	−.36
	[19]	(1.06)	(.99)	(.93)	(.85)
Mental health index	75.5	.45	1.07	.43	.64
	[14.8]	(.91)	(.87)	(.83)	(.75)
Number enrolled	759	881	1,022	1,295	3,198

Utilization went up in some cases, but health didn't improve as a result

Oregon Medicaid Experiment

A natural experiment in 2008 that gave some low-income adults who won a lottery the opportunity to apply for public health insurance through Medicaid

- Compare health care choices and outcomes for two populations
    1. Low-income family lottery winners
    2. Low-income family lottery losers

- Only 25% take-up of insurance in “treatment” group
    - This mutes treatment effect and suggests the possibility of selection bias

Oregon Experiment Outcomes

Table 1.5
OHP effects on insurance coverage and health-care use

Outcome	Oregon: Control mean (1)	Oregon: Treatment effect (2)	Portland area: Control mean (3)	Portland area: Treatment effect (4)
A. Administrative data				
Ever on Medicaid	.141	.256 (.004)	.151	.247 (.006)
Any hospital admissions	.067	.005 (.002)		
Any emergency department visit			.345	.017 (.006)
Number of emergency department visits			1.02	.101 (.029)
Sample size	74,922		24,646	
B. Survey data				
Outpatient visits (in the past 6 months)	1.91	.314 (.054)		
Any prescriptions?	.637	.025 (.008)		
Sample size	23,741			

Utilization went up, even for ER visits

Oregon Experiment Outcomes

OHP effects on health indicators and financial health

Outcome	Oregon: Control mean (1)	Oregon: Treatment effect (2)	Portland area: Control mean (3)	Portland area: Treatment effect (4)
A. Health indicators				
Health is good	.548	.039 (.008)		
Physical health index			45.5	.29 (.21)
Mental health index			44.4	.47 (.24)
Cholesterol			204	.53 (.69)
Systolic blood pressure (mm Hg)			119	−.13 (.30)
B. Financial health				
Medical expenditures >30% of income			.055	−.011 (.005)
Any medical debt?			.568	−.032 (.010)
Sample size	23,741		12,229	

Health status improved somewhat, mental health most prominent, and financial well-being improved

The New York Times

The I.R.S. Sent a Letter to 3.9 Million People. It Saved Some of Their Lives.

More in Class 4→

The study was an accident. The results show the positive effects of health insurance.

Three years ago, 3.9 million Americans received a plain-looking envelope from the Internal Revenue Service. Inside was a letter stating that they had recently paid a fine for not carrying health insurance and suggesting possible ways to enroll in coverage.

New research concludes that the bureaucratic mailing saved lives.

Three Treasury Department economists have published a working paper finding that these notices increased health insurance sign-ups. Obtaining insurance, they say, reduced premature deaths by an amount that exceeded any of their expectations. Americans between 45 and 64 benefited the most: For every 1,648 who received a letter, one fewer death occurred than among those who hadn't received a letter.

In all, the researchers estimated that the letters may have wound up saving 700 lives.

Department of the Treasury
Internal Revenue Service
IRS

January 12 , 2017

Why am I getting this letter?
The law requires people to have a minimum level of health coverage, qualify for an exemption, or pay a penalty when they file their taxes. Our records show you reported owing this penalty when you filed your 2015 taxes because you or someone in your family did not have health insurance during that year. If you don't have health insurance or an exemption next year, you'll likely owe a penalty for 2017 as well. We are writing to make sure you know how you can avoid this penalty by signing up for health insurance.

How do I avoid the penalty next year?
If you don't have health coverage, you can avoid owing a penalty for most or all of 2017 by signing up for health insurance soon. One way to get insurance is to sign up at HealthCare.gov **before January 31, 2017**. If you already have health coverage, you won't owe a penalty as long as you stay covered.

How much will my penalty be next year if I don't sign up?
The penalty for not having any health coverage in 2017 will be about if your income and family size have not changed since 2015.

How much does health insurance at HealthCare.gov cost?
Most people who enroll in a plan through HealthCare.gov can find plans for **$75 a month or less** after financial help. At HealthCare.gov, you can compare plans to find one that meets your needs and budget.

How do I sign up for health insurance or get help finding a plan?
You can apply online by computer or mobile device, or you can get help in-person or by phone.

- Visit HealthCare.gov, select your state, and follow the step-by-step directions.
- Find in-person help from someone in your community at LocalHelp.HealthCare.gov.
- For questions or help signing up, call

When is the deadline to sign up?
January 31, 2017, is the last day to enroll in a 2017 plan on HealthCare.gov.

Does Advertising Actually Work?

FREAKONOMICS

Does Advertising Actually Work?

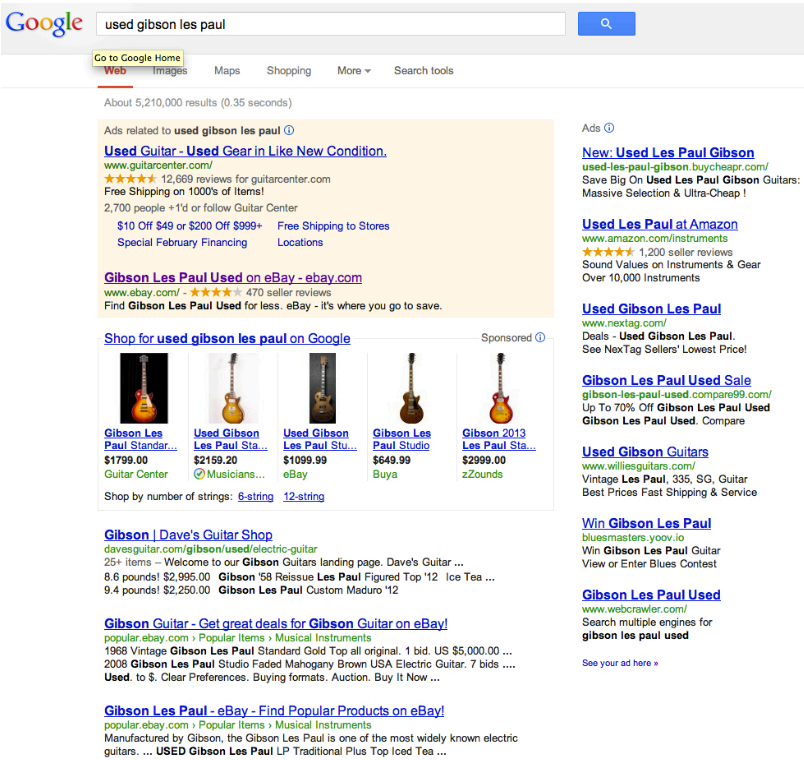
Key Facts

- Executives told Professor Levitt that there was one thing they knew to be true: that TV ads were much more effective, dollar-for-dollar, than their newspaper ads
- They also said that they'd been advertising in every big Sunday newspaper in the U.S., every week, for the past 15 years
- It's foolish to try to "out-jargon" an economics professor like Steve Tadelis

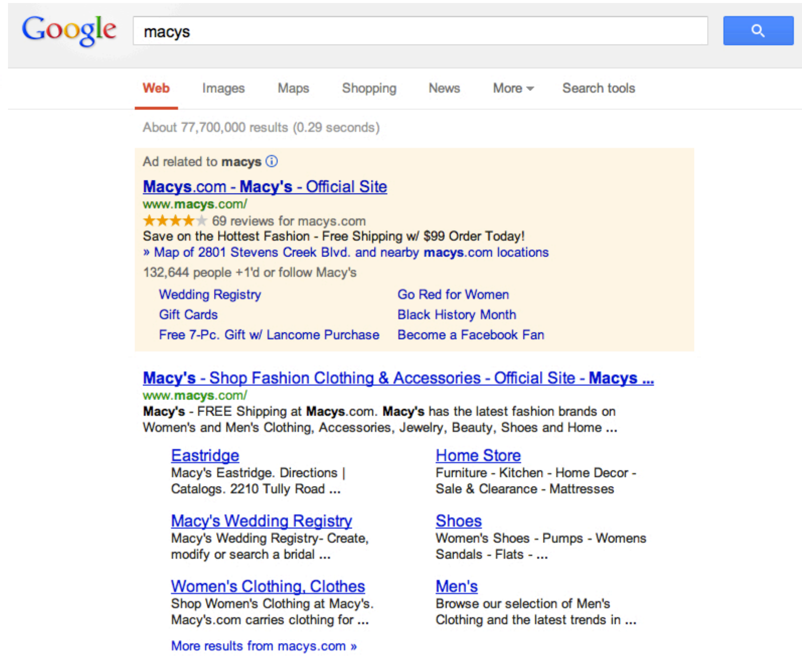
Conceptual Questions

- Why would it be challenging to determine which form of advertising is more effective given this company's advertising strategy — and even hard to tell if advertising matters at all?
- How did the research design of Professor Tuchman's study resemble an RCT? Does it allow her to uncover the causal effects of advertising?
- How did eBay's "natural experiment" allow the company to distinguish causation from correlation regarding the impact of keyword advertising on sales?

Different forms of Keyword Advertising

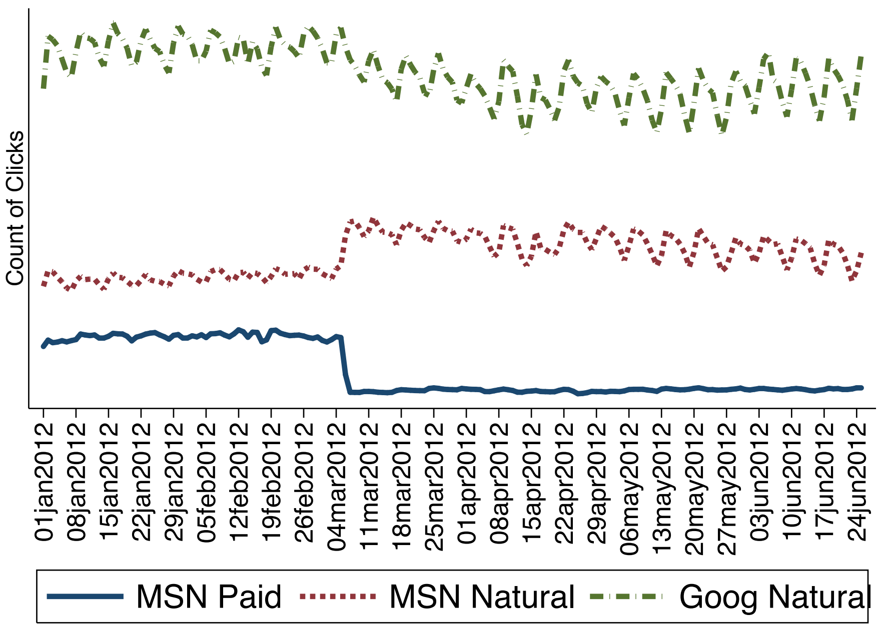


(a) Used Gibson Les Paul

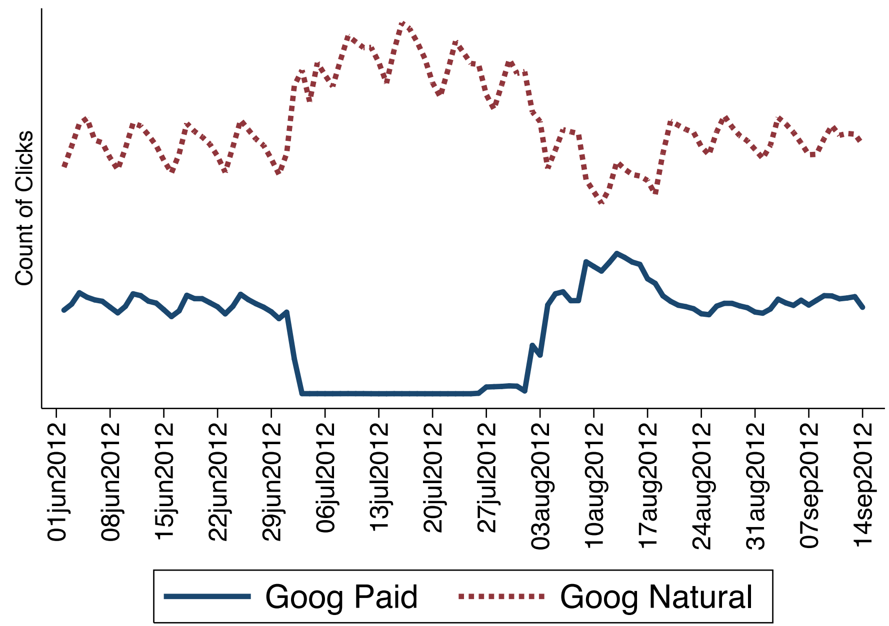


(b) Macys

Experimental Design

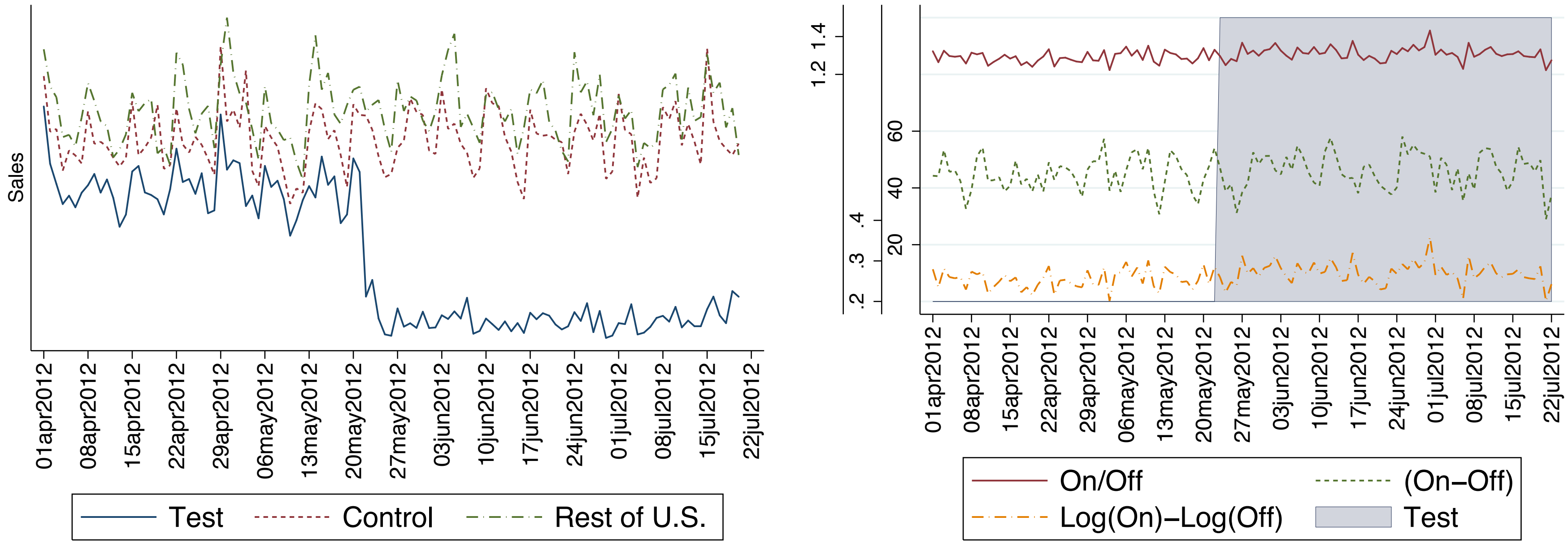


(a) MSN Test



(b) Google Test

Measuring Experimental Effects



(a) Attributed Sales by Region

(b) Differences in Total Sales

Regression 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.

Q + A