w241: Experiments and Causality

Unit 2

David Reiley, David Broockman, D. Alex Hughes
UC Berkeley, School of Information
Updated: 2021-04-29

Introduction

Introduction

- Comparing apples to apples.
- Potential Outcomes
- Measuring Effects of Online Ads

Comparing Apples to Apples

Comparing Apples to Apples

Main topics this week

- Potential outcomes (FE 2.1)
- Average treatment effects (FE 2.2)
- Random sampling and expectations (FE 2.3)
- Random assignment and unbiased inference (FE 2.4–5)
- Biased results in observational data (FE 2.6)
- Measuring the effects of online advertising (Lewis and Reiley, 2014)

Why experimentation?

- Experimentation delivers much more reliable causal inference than any observational method.
  - Allows us to compare two identical populations in which all that varies is treatment of interest
- Conducting experiments correctly isn't easy.
- Concept of "potential outcomes" shows us what can go right and wrong.

Defining Potential Outcomes

Potential Outcomes: Theoretical concepts useful for thinking about what an experiment could show

- Example: Table 2.1
  - Could never be derived from real data
  - Assumes an impossible amount of information
- *In Practice*: Only treatment group observed in treatment, and only control group observed in control
- *In theory*: Can imagine a group in two counterfactual states, but can actually observe only one

Potential Outcomes Notation

- $Y_i(1) =$ Outcome if you were to be in treatment
- $Y_i(0) =$ Outcome if you were to be in control
- $\tau_i = Y_i(1) - Y_i(0) =$ Treatment effect for individual *i*
- In village *i* only: How many more budget percentage points would be devoted to water sanitation if you were in treatment versus control?
- Not directly observable, but useful to think about hypothetically

No Causation Without Manipulation

- If you can't imagine a manipulation that answers your question, it may not have a causal answer.
- What would the same person do if in one treatment versus another?
- Intervention is required to generate needed data, but sometimes imagining an intervention is impossible.

Fundamentally Unanswerable Questions

Example

What is the effect on mortality rates of being born in Africa?

- What does this even mean for a particular person?
  - $Y_i(1) =$ outcome if person born in Africa?
  - $Y_i(0) =$ outcome if same person born in the United States?
- Born in African hospital?
- Lived entire life in Africa?
- Question not posed well -- *FUQ'd*

Of Ideals and Experiments

- What is the ideal experiment?
- What is the implied manipulation?

FE 2.2: Reading Guidelines

- $d_i =$ treatment *"dosage"*
- Box 2.1:
  - D_i versus d_i
  - Should remind you of statistics and random variables: a realization x_i of a random variable X_i
- Equation 2.2
  - Uses multiplication to express conditionals
  - $Y_i(1)$ if in treatment -- this is where $d_i = 1$ and $(1 - d_i) = 0$]
  - $Y_i(0)$ if in control -- this is where $d_i = 0$ and $(1 - d_i) = 1$]

Measuring the Effects of Advertising

Measuring the Effects of Advertising

Brand Image Advertising

- Difficult to measure the effects of brand image advertising
  - Advertisements that don't solicit direct response;
  - Rather, increase awareness of and positive association with a particular brand
- Consider observational methodology published in Harvard Business Review by founder and president of ComScore (Abraham, 2008)
  - Panel of one million people.
  - Compare buying behavior of people who did and did not see a given ad campaign.
  - Is treated population more likely to shop at the retailer than those not exposed to the ad?

Potential Problem

- The two samples don't come from same population.

Brand Image Advertising Case Study:

E*Trade

- Increases sales by over 200%, according to ComScore's analysis.
- Comparing people who did and did not see an E*Trade ad on Google search results.
    - People who see the ad have searched keywords such as "online brokerage."
    - Could there be other differences between those who did and did not execute such searches, aside from seeing the ad?

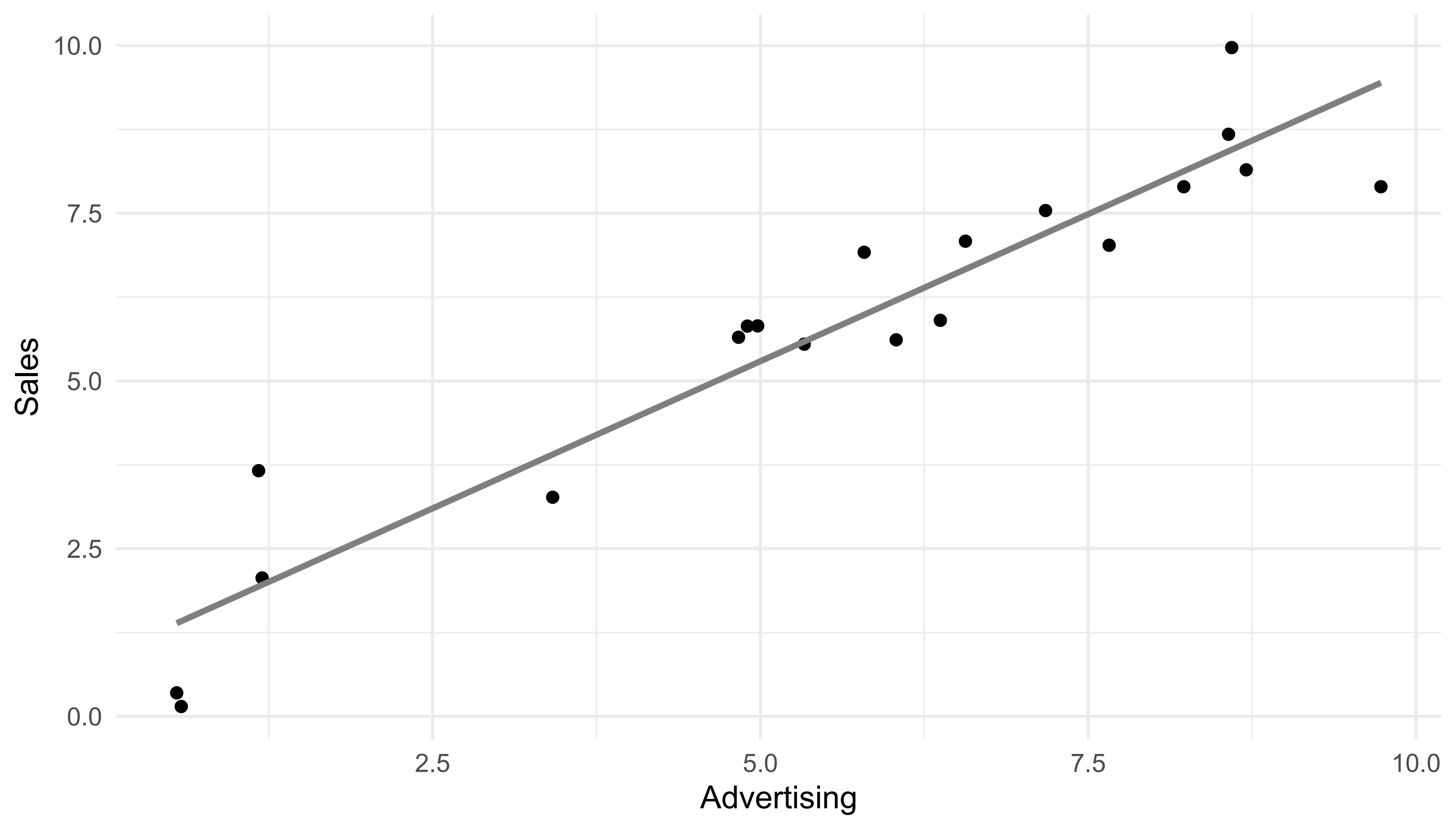
Potential Problem

- Group who sees the ad already interested in online brokerage.
- Correlation not the same as causality.

The Marketing Two-Step

- How is advertising effectiveness measured?
- Online ad firm shows ads only to people most likely to buy a company's product.
- Determining effect of the campaign:
  - Comparing behavior of those who saw ads with those who didn't is not apples-to-apples.
  - Choosing who gets the treatment often has a lot to do with the very outcome we're intending to measure.
- Beware of bias in measuring effects.

Measuring Effects of Advertising on

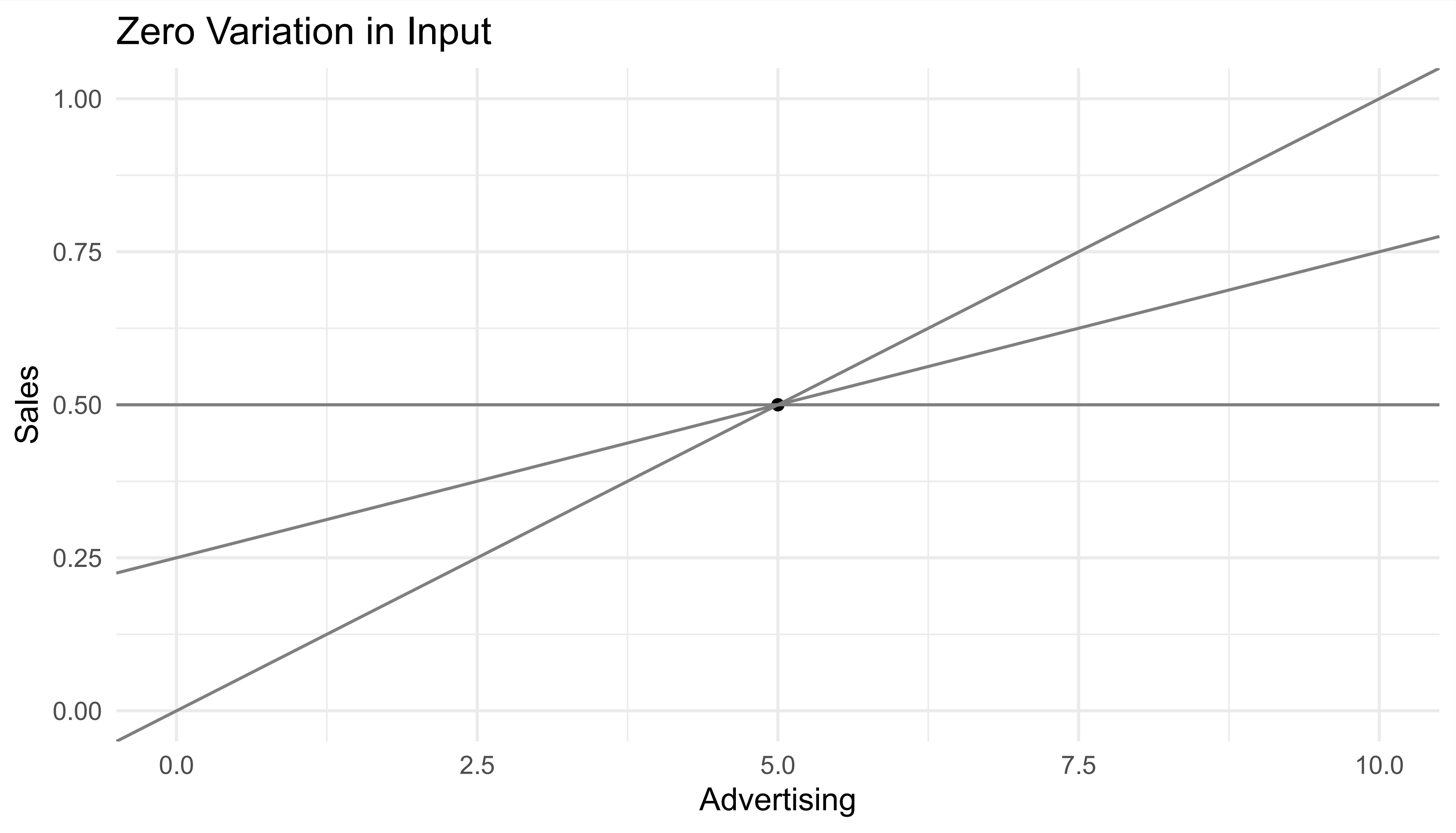


Measuring Advertising on Sales

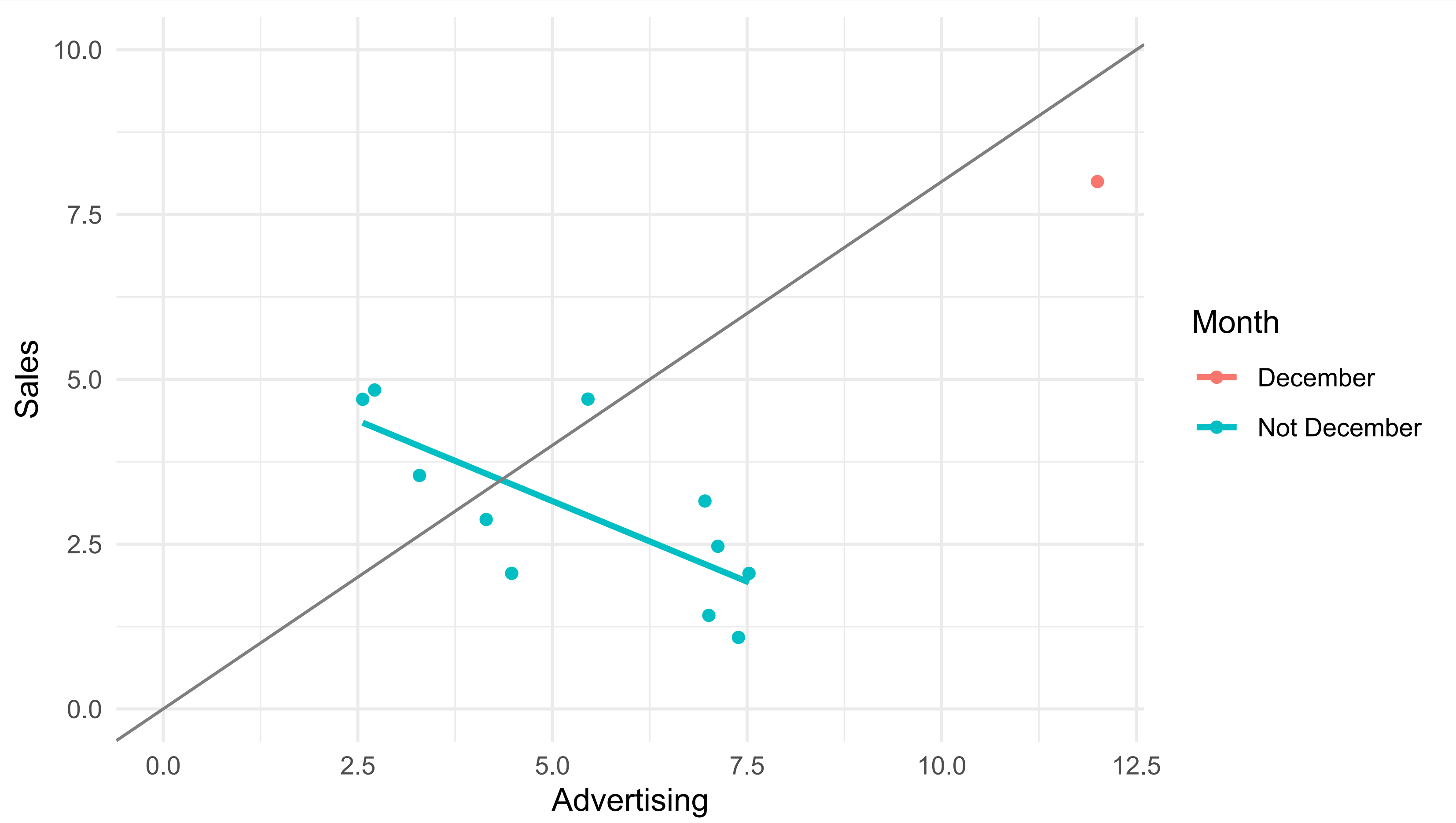
- Econometric regressions of aggregate sales versus advertising
- "*Endogeneity*" problem
    - Amount of advertising not randomly determined.
    - Sales and advertising both influence each other.
    - Potential for reverse causality.
- Need a situation where advertising varies independently of other factors that could cause sales

Needs an Experiment!

No Variation in X



Bad Variation In X



Experimentation vs. Observational Data

- Regressing sales on advertising:
  - If advertising doesn't vary, regression doesn't convey much useful information.
  - Experiments generate variation.
- Advertising must vary somehow or slope of regression wouldn't be measurable.
  - More advertising in December
  - More likely to overestimate or underestimate effects of increased advertising in December?

Conclusions: Christmas Advertising

- Key question in measuring causal effects of X on Y: How does X vary?
- Omitted variable—Christmas—causes increased advertising and increased sales.
- Blindly running regression on observational data implicitly assumes advertising to be only variable responsible for increased sales.
- Effects of advertising overestimated due to omitted-variable bias.
- Using observational data; not comparing apples to apples.

Review

Three Examples

- We've discussed three examples of observational data providing inaccurate results.
- Aggregate time-series data
    - Advertising doesn't vary systematically over time.
    - **Reverse causality** problem: we only have draws from the joint distribution, not directions.
    - **Omitted-variable bias** problem: there are other variables that might confound relationships that naive regressions *estimate* exist.
- Individual cross-sectional data
    - **Selection bias** problem: type of people who see ads not the same population as those who don't.

Even in absence of ads, shopping behavior might be different.

Online Ads and Offline Sales

Rudimentary Understanding

- Advertising today = physics in the 1500s

> "Do heavy bodies fall at faster rates than light ones?" - Galileo

- Manipulate mass while keeping shape and size constant.
- Used experimental method to prove objects fell at same rate despite different masses.
- Huge advance over observational data.

Online-Advertising Field Experiment

- Lewis and Reiley, "Online Ads and Offline Sales," Quantitative Marketing and Economics, 2014.
- One of largest field experiments ever conducted.
- Read through Section III.B.

Online-Advertising Field Experiment

Positive Increase in Sales Due to Ads

	During Campaign
Control	R$1.84
	(0.03)
Treatment	1.89
	(0.02)

Design

- 1.2 million in treatment group
- 400,000 in control
- Two weeks

Results

- Effect not statistically significant

Online-Advertising Field Experiment

Observational Comparison

- Treatment-group Members Exposed vs. Not Exposed to Ads

	During Campaign
Control	R$1.84
	(0.03)
Treatment	1.89
	(0.02)
Shown Retailer's Ads	1.81
(64% of Treatment Group)	(0.02)
Not Shown Retailer's Ads	2.04
(36% of Treatment Group)	(0.03)

- Could conclude that advertising reduced sales by R$0.23
- Not comparing apples to apples

Nonexperimental Sales Differences

Unrelated to Ad Exposure

	Before Campaign	**During Campaign**
Control	R\$1.95	R\$1.84
	(0.04)	(0.03)
Treatment	1.93	1.89
	(0.02)	(0.02)
Shown Retailer's Ads	1.81	1.81
(64% of Treatment Group)	(0.02)	(0.02)
Not Shown Retailer's Ads	2.15	2.04
(36% of Treatment Group)	(0.03)	(0.03)

- Those who browse enough to see ads also have lower baseline propensity to purchase from the retailer.
- Potential mistake solved with experiment.

Experiments Eliminate Selection Bias

- To measure effect of X on Y, we compare Y among units with different values of X.

Why do units have different values of X?

- With no experiment, inference difficult because units obtain different values of X for reasons related to Y.
    - Experiments generate variation in X independent of Y.
    - Populations should be identical in all ways other than the value of X.
- Random assignment generates apples-to-apples comparison.
- Always ask yourself how group divisions came to be.

Example

Does Playing Outside Improve Eyesight?

- Study conducted by Australian doctors
  - Kids who play outside are less likely to need glasses.
  - *Possible explanation*: More sunlight exposure causes better retinal development?
- **Better question**:
  - Why do kids choose to play outside or inside in the first place?
  - Maybe kids with worse eyesight don't like to play outside.
  - Need an experiment to establish causality.

Abstracting from the Example

Reading Assignment

- Read Sections 2.3–2.6.
- Bring any questions to this week's live session.

Key Points to Remember

- Observational data can easily compare apples to oranges.
- *Selection bias*: Without a clean experiment, other factors can seem like treatment effects.
  - Those who select treatment often differ in other ways.
- In Lewis-Reiley advertising study, naive observational measurement has wrong sign and is three times larger than estimate given by experiment.
- Experimentation more reliably estimates causal effects than observation.
- Random assignment is gold standard.
- Measuring effect of X on Y.
- What are the potential outcomes for a given person?
- What is the ideal experiment?
- What causes the variation in X?