Applied Causal Inferece

Lecture 2

Alex Stephenson

8-27-2021

What did we do last time?

1. Overview and Logistics
2. For those who have not taken the final exam or the pre-class survey do so by the end of today.

What are we doing today?

1. Thinking about different kinds of research projects
2. Introducing the basics of Research Design
3. Priming some concepts that we will cover in more depth starting on Monday

Three Types of Research Projects

We can bucket research into one of three types:

1. Descriptive research
    - e.g., how many students attend UC Berkeley
2. Predict the future
    - Much machine learning work falls in this bucket
3. Predict the result of an intervention
    - **Our class is devoted to these types of questions**

Framework for a Causal Question

In general, causal questions take the form "What is the effect of X on Y?" Here are some examples:

"What is the effect of co-ethnicity of judge appeal decisions in Kenya?" (Shen Bayh, Choi, and Harris 2021)

"What is the effect of welfare transfers on citizens' demands for services in India?" (Kumar 2021)

"What is the effect of resource extraction on local governance?" (Mitchell Elder 2021)

What is a Causal Research Design

A causal research design is a statement of how a study will estimate a relationship between variables that is causal in nature.

The design describes to the audience how:

- An intervention causes Some variation in an outcome of interest
- Either explicitly through a randomized experiment or a situation that well approximates one.

When we consider a design, we should always ask, "How would I do the study if it was possible to run an experiment?"

How do we do this?

We can fully describe a research design in five sentences (Lundberg et al., 2021). The first three are conceptual questions:

1. What is the estimand I care about?

2. What is the target population of interest?

3. Why is learning this quantity useful?

How do we do this?

The next two are implementation questions:

1. What is the empirical estimator that best approximates the estimand of interest?

2. What is the strategy to estimate the empirical estimator?

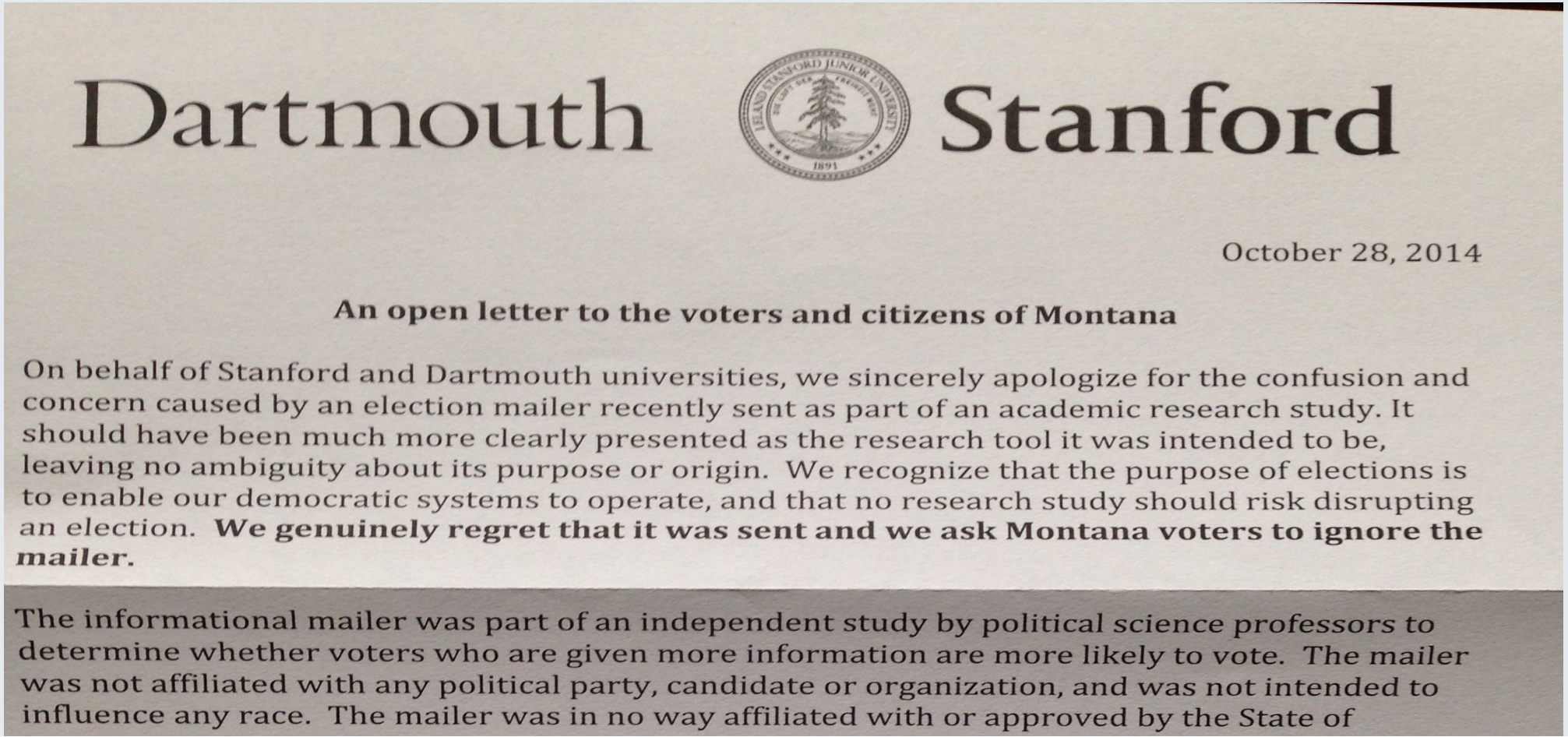
Research Ethics

In academic research, many issues are studied because they lend themselves well to specific designs.

Since you are not currently academics, it is worth asking if studying a question is socially useful.

Example 1: Bonica *et al.* (2014)

Example 1: Bonica *et al.* (2014)



Dartmouth Stanford

October 28, 2014

An open letter to the voters and citizens of Montana

On behalf of Stanford and Dartmouth universities, we sincerely apologize for the confusion and concern caused by an election mailer recently sent as part of an academic research study. It should have been much more clearly presented as the research tool it was intended to be, leaving no ambiguity about its purpose or origin. We recognize that the purpose of elections is to enable our democratic systems to operate, and that no research study should risk disrupting an election. **We genuinely regret that it was sent and we ask Montana voters to ignore the *mailer*.**

The informational mailer was part of an independent study by political science professors to determine whether voters who are given more information are more likely to vote. The mailer was not affiliated with any political party, candidate or organization, and was not intended to influence any race. The mailer was in no way affiliated with or approved by the State of

Example 1: Bonica, Rodden, Dropp (2014)

Bonica, Rodden, and Dropp were interested in assessing why most voters choose to abstain in certain races.

- Assuming that this is a worthwhile question, is this a helpful research design to get an answer?
- Could the same experiment have been conducted without using official state seals?

Example 2: Bursztyn *et al.* (2021)

"We study the causes of sustained participation in political movements. **To identify the persistent effect of protest participation, we randomly, indirectly incentivize Hong Kong university students in participation in an anti-authoritarian protest.** To identify the role of social networks, we randomize this treatment's intensity across major-cohort cells. We find that incentives to attend one protest within a political movement increase subsequent protest attendance, but only when a sufficient fraction of an individual's social network is also incentivized to attend the initial protest."

Example 2: Bursztyn *et al.* (2021)

Some ethical worries here:

1. The students themselves could be imprisoned for protesting, and families in mainland China targeted by the CCP

2. The study puts non-participants in the research at risk. It is almost certainly more complicated to work at these institutions than before.

Lots of additional conversation of these and other issues at https://twitter.com/SheenaGreitens/status/1148382146781949953

Example 2: Bursztyn *et al.* (2021)

What are the benefits of doing this research?

1. Are these points already known by the literature on social movements? **Absolutely, and in great detail.**

2. If the risks are minimal, does this research design get at the question of interest? **Probably not.**

Research Ethics

- Ethics is a fraught concept.
- Some of you may find either or both of these examples perfectly reasonable and ethical. *This is fine. You need a justification*
- At a bare minimum, every ethical study has to pass the basic bar of **"Is this useful in some way such that the costs of going about this design are worth it?"**

What are we doing on Monday?

- Defining Causal Inference
- Introducing Potential Outcomes
- Introducing our first estimand