Causal Inference

Justin Grimmer

Associate Professor
Department of Political Science
University of Chicago

March 26th, 2018

Who Take This Class and Why They Take It

Who:

Who Take This Class and Why They Take It

Who:

1) You have taken the prior two classes in the sequence (covering Math, Probability, and Linear Regression) *or their equivalent courses*

Who Take This Class and Why They Take It

Who:

1) You have taken the prior two classes in the sequence (covering Math, Probability, and Linear Regression) *or their equivalent courses*
2) Familiar with the `R` programming language

Who Take This Class and Why They Take It

Who:

1) You have taken the prior two classes in the sequence (covering Math, Probability, and Linear Regression) *or their equivalent courses*
2) Familiar with the `R` programming language

Why:

Who Take This Class and Why They Take It

Who:

1) You have taken the prior two classes in the sequence (covering Math, Probability, and Linear Regression) *or their equivalent courses*
2) Familiar with the `R` programming language

Why:

1) Learn about methods for causal inference: the effect of some intervention on some outcomes

Who Take This Class and Why They Take It

Who:

1) You have taken the prior two classes in the sequence (covering Math, Probability, and Linear Regression) *or their equivalent courses*
2) Familiar with the `R` programming language

Why:

1) Learn about methods for causal inference: the effect of some intervention on some outcomes
2) Topics: experiments, matching, regression, sensitivity analysis, difference-in-differences, panel methods, instrumental variable estimation, and regression discontinuity designs (and if time, machine learning + causal inference)

Who Take This Class and Why They Take It

Who:

1) You have taken the prior two classes in the sequence (covering Math, Probability, and Linear Regression) *or their equivalent courses*
2) Familiar with the `R` programming language

Why:

1) Learn about methods for causal inference: the effect of some intervention on some outcomes
2) Topics: experiments, matching, regression, sensitivity analysis, difference-in-differences, panel methods, instrumental variable estimation, and regression discontinuity designs (and if time, machine learning + causal inference)
3) Applications from across the social sciences, with a focus on public policy questions in the social sciences.

Who Take This Class and Why They Take It

Who:

1) You have taken the prior two classes in the sequence (covering Math, Probability, and Linear Regression) *or their equivalent courses*
2) Familiar with the R programming language

Why:

1) Learn about methods for causal inference: the effect of some intervention on some outcomes
2) Topics: experiments, matching, regression, sensitivity analysis, difference-in-differences, panel methods, instrumental variable estimation, and regression discontinuity designs (and if time, machine learning + causal inference)
3) Applications from across the social sciences, with a focus on public policy questions in the social sciences.

Goal:

Provide you with adequate methodological skills to conduct cutting-edge empirical research

Causal Inference

- Statistics can be used for many purposes:
  - Discovery
  - Measurement
  - Causal inference
    - Relatively new subfield within statistics
    - Highly interdisciplinary, rapidly expanding

Causal Inference

- Statistics can be used for many purposes:
  - Discovery
  - Measurement
  - Causal inference
    - Relatively new subfield within statistics
    - Highly interdisciplinary, rapidly expanding

Causal Inference

- Statistics can be used for many purposes:
  - Discovery
  - Measurement
  - Causal inference
    - Relatively new subfield within statistics
    - Highly interdisciplinary, rapidly expanding

$\boxed{\text{Causal Inference is a Hard Problem}}$

Traditional Approaches

John Stuart Mill's **Method of Difference**:

> *If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring only in the former; the circumstance in which alone the two instances differ is the effect, or the cause, or an indispensable part of the cause, of the phenomenon.*

It would be fantastic if two cases were exactly alike and only differ in one single characteristic, but we rarely have such comparisons in the social sciences

Traditional Approaches

John Stuart Mill's **Method of Difference**:

> *If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance in common save one, that one occurring only in the former; the circumstance in which alone the two instances differ is the effect, or the cause, or an indispensable part of the cause, of the phenomenon.*

It would be fantastic if two cases were exactly alike and only differ in one single characteristic, but we rarely have such comparisons in the social sciences

More from Mill

If so little can be done by the method to determine conditions of an effect of many combined causes in the case of the medical sciences, still less is the method applicable to a class of phenomena more complicated than even those of physiology, the phenomena of politics and history. There, Plurality of Causes exists in almost boundless excess, and effects are, for the most part, inextricably interwoven with one another. To add to the embarrassment, most of the inquiries in political science relate to the production of effects of a most comprehensive description, such as the public wealth, the public security, public morality, and the like: results liable to be affected directly or indirectly either in plus or minus by nearly every fact which exists, or event which occurs, in human society. The vulgar notion that, the safe methods on political subjects are those of Baconian induction [...] will one day be quoted as among the most unequivocal marks of a low state of the speculative faculties in any age in which it is accredited.

- Causes of Effects $\rightsquigarrow$ Why Did Trump Win? (Headache question (Dawid 2000): Headache is gone, is that because I took asprin?)

- Causes of Effects $\rightsquigarrow$ Why Did Trump Win? (Headache question (Dawid 2000): Headache is gone, is that because I took asprin?)
- Effects of Causes $\rightsquigarrow$ How did a "China Shock" effect support for Trump? (Dawid 2000: I have a headache, will it help if I take aspirin?)

A Case Study in Failed Causal Inference

I Just Ran Two Million Regressions

By XAVIER X. SALA-I-MARTIN*

Empirical Growth Literature

Levine and Ross (1992):

> *There are many econometric specifications in which measures of economic policy are significantly correlated with long-run per capita growth rates.*
>
> *The cross-country statistical relationships between long-run average growth rates and almost every particular policy indicator considered by the profession are fragile: small alterations in the "other" explanatory variables overturn past results.*

Empirical Growth Literature

Levine and Ross (1992):

> *There are many econometric specifications in which measures of economic policy are significantly correlated with long-run per capita growth rates.*
>
> *The cross-country statistical relationships between long-run average growth rates and almost every particular policy indicator considered by the profession are fragile: small alterations in the "other" explanatory variables overturn past results.*

Is Causal Inference Possible?

Why We Learn Nothing from Regressing Economic Growth on Policies*

Dani Rodrik*

Why is this a widespread sentiment?

- Important topic
- A lot of data
- A lot of theory
- Many smart people working on this problem

Is Causal Inference Possible?

Why We Learn Nothing from Regressing Economic Growth on Policies*

Dani Rodrik *

Why is this a widespread sentiment?

- Important topic
- A lot of data
- A lot of theory
- Many smart people working on this problem

Case Study: Post-Menopausal Estrogen

- Postmenopausal estrogen plus progestin showed large reduced risk in coronary heart disease.
  - High quality observational studies using regression and related techniques on large samples provided the basis for wide scale prescription of estrogen.

Grodstein and Stampfer (1998):

> *"Consistent evidence from over* **40** *epidemiologic studies demonstrates that postmenopausal women who use estrogen therapy after the menopause have significantly lower rates of heart disease than women who do not take estrogen. . . the evidence clearly supports a clinically important protection against heart disease for postmenopausal women who use estrogen."*

Case Study: Post-Menopausal Estrogen

- Postmenopausal estrogen plus progestin showed large reduced risk in coronary heart disease.
  - High quality observational studies using regression and related techniques on large samples provided the basis for wide scale prescription of estrogen.

Grodstein and Stampfer (1998):

> *"Consistent evidence from over* **40** *epidemiologic studies demonstrates that postmenopausal women who use estrogen therapy after the menopause have significantly lower rates of heart disease than women who do not take estrogen. . . the evidence clearly supports a clinically important protection against heart disease for postmenopausal women who use estrogen."*

Case Study: Post-Menopausal Estrogen

The New York Times

This copy is for your personal, noncommercial use only. You can order presentation-ready copies for distribution to your colleagues, clients or customers, please click here or use the "Reprints" tool that appears next to any article. Visit www.nytreprints.com for samples and additional information. Order a reprint of this article now. »

April 22, 2003

Hormone Studies: What Went Wrong?

By GINA KOLATA

- Women's Health Initiative (WHI) was a large randomized trial studying the question
- Experiment showed no benefits from the therapy and increased risk from Breast Cancer
- Experiment was halted early and prescription of therapy plummeted immediately thereafter

Case Study: Voter ID Laws

- Voter ID laws: require government issued identification (proof of citizenship) to vote

Case Study: Voter ID Laws

- Voter ID laws: require government issued identification (proof of citizenship) to vote
- Disproportionately Burdens: (1) poor, (2) persons of color, and (3) elderly (Ansolabahere and Hersh 2016)

Case Study: Voter ID Laws

- Voter ID laws: require government issued identification (proof of citizenship) to vote
- Disproportionately Burdens: (1) poor, (2) persons of color, and (3) elderly (Ansolabahere and Hersh 2016)
- Causal effect(?):

Case Study: Voter ID Laws

- Voter ID laws: require government issued identification (proof of citizenship) to vote
- Disproportionately Burdens: (1) poor, (2) persons of color, and (3) elderly (Ansolabahere and Hersh 2016)
- Causal effect(?):

Case Study: Voter ID Laws

- Voter ID laws: require government issued identification (proof of citizenship) to vote
- Disproportionately Burdens: (1) poor, (2) persons of color, and (3) elderly (Ansolabahere and Hersh 2016)
- Causal effect(?):

- And yet:

Case Study: Voter ID Laws

- Voter ID laws: require government issued identification (proof of citizenship) to vote
- Disproportionately Burdens: (1) poor, (2) persons of color, and (3) elderly (Ansolabahere and Hersh 2016)
- Causal effect(?):


Latino Decisions @LatinoDecisions · 4h
How about this for an effect. If you lack a valid ID, a necessary credential to vote, you have a 0% probability of turnout.

- And yet:
  "To date, there is little convincing evidence that voter identification laws influence turnout substantially" Highton 2017.

Case Study: Voter ID Laws

- Voter ID laws: require government issued identification (proof of citizenship) to vote
- Disproportionately Burdens: (1) poor, (2) persons of color, and (3) elderly (Ansolabahere and Hersh 2016)
- Causal effect(?):

  > **Latino Decisions** @LatinoDecisions · 4h
  > How about this for an effect. If you lack a valid ID, a necessary credential to vote, you have a 0% probability of turnout.

- And yet:
  "To date, there is little convincing evidence that voter identification laws influence turnout substantially" Highton 2017.
- Or:

Case Study: Voter ID Laws

- Voter ID laws: require government issued identification (proof of citizenship) to vote
- Disproportionately Burdens: (1) poor, (2) persons of color, and (3) elderly (Ansolabahere and Hersh 2016)
- Causal effect(?):

- And yet:
  "To date, there is little convincing evidence that voter identification laws influence turnout substantially" Highton 2017.
- Or:
  "When errors are corrected [in Hajnal, Lajevardi, and Nielson 2017], one can recover positive, negative, or null estimates of the effect of voter ID laws on turnout, precluding firm conclusions" Grimmer, Hersh, Meredith, Mummolo, and Nall (2018)

Case Study: Voter ID Laws

Why?

Case Study: Voter ID Laws

Why?

- Uneven enforcement ("For example, Ansolabehere (2009) reports that white voters are less likely to be asked for a photo ID than either Latinos or African Americans by about seven percentage points" (see also White et. al 2015))

Case Study: Voter ID Laws

Why?

- Uneven enforcement ("For example, Ansolabehere (2009) reports that white voters are less likely to be asked for a photo ID than either Latinos or African Americans by about seven percentage points" (see also White et. al 2015))
- Psychological motivation (Valentino and Neuner (2017))

Case Study: Voter ID Laws

Why?

- Uneven enforcement ("For example, Ansolabehere (2009) reports that white voters are less likely to be asked for a photo ID than either Latinos or African Americans by about seven percentage points" (see also White et. al 2015))
- Psychological motivation (Valentino and Neuner (2017))
- Compensatory campaign spending

Doesn't Regression Solve the Problem?

$$Y = \alpha + \tau D + \beta_1 X_1 + ... + \beta_k X_k + u$$

Key problems with regression:

Doesn't Regression Solve the Problem?

$$Y = \alpha + \tau D + \beta_1 X_1 + ... + \beta_k X_k + u$$

Key problems with regression:

- Endogeneity and omitted variable bias

Doesn't Regression Solve the Problem?

$$Y = \alpha + \tau D + \beta_1 X_1 + ... + \beta_k X_k + u$$

Key problems with regression:

- Endogeneity and omitted variable bias
- Misspecified functional form

Doesn't Regression Solve the Problem?

$$Y = \alpha + \tau D + \beta_1 X_1 + ... + \beta_k X_k + u$$

Key problems with regression:

- Endogeneity and omitted variable bias
- Misspecified functional form
- Heterogeneous treatment effects

Let's be adults

Let's be adults

Language games don't change that you're estimating a causal effect

Let's be adults

Language games don't change that you're estimating a causal effect

So talk about it like a causal effect

Causal Identification (Manski 1995)

Causal Identification (Manski 1995)

Studies of **identification** *seek to characterize the conclusions that could be drawn if one could use the sampling process to obtain an unlimited number of observations.*

Causal Identification (Manski 1995)

Studies of **identification** *seek to characterize the conclusions that could be drawn if one could use the sampling process to obtain an unlimited number of observations.*

Studies of **statistical inference** *seek to characterize the generally weaker conclusions that can be drawn from a finite number of observations.*

Causal Identification (Manski 1995)

Studies of **identification** *seek to characterize the conclusions that could be drawn if one could use the sampling process to obtain an unlimited number of observations.*

Studies of **statistical inference** *seek to characterize the generally weaker conclusions that can be drawn from a finite number of observations.*

In statistics, something is **identifiable** if you can get it from the joint distribution of observable random variables.

Causal Identification (Manski 1995)

> *Studies of* **identification** *seek to characterize the conclusions that could be drawn if one could use the sampling process to obtain an unlimited number of observations.*
>
> *Studies of* **statistical inference** *seek to characterize the generally weaker conclusions that can be drawn from a finite number of observations.*

In statistics, something is **identifiable** if you can get it from the joint distribution of observable random variables.

Informally, **causal identification** describes the combination of assumptions, estimation approach, and data necessary to reach a conclusion about a causal effect:

Causal Identification (Manski 1995)

> *Studies of* **identification** *seek to characterize the conclusions that could be drawn if one could use the sampling process to obtain an unlimited number of observations.*

> *Studies of* **statistical inference** *seek to characterize the generally weaker conclusions that can be drawn from a finite number of observations.*

In statistics, something is **identifiable** if you can get it from the joint distribution of observable random variables.

Informally, **causal identification** describes the combination of assumptions, estimation approach, and data necessary to reach a conclusion about a causal effect:

$$\text{assumptions} + \text{data} \rightarrow \text{causal conclusions}$$

Causal Identification (Manski 1995)

> *Studies of* **identification** *seek to characterize the conclusions that could be drawn if one could use the sampling process to obtain an unlimited number of observations.*

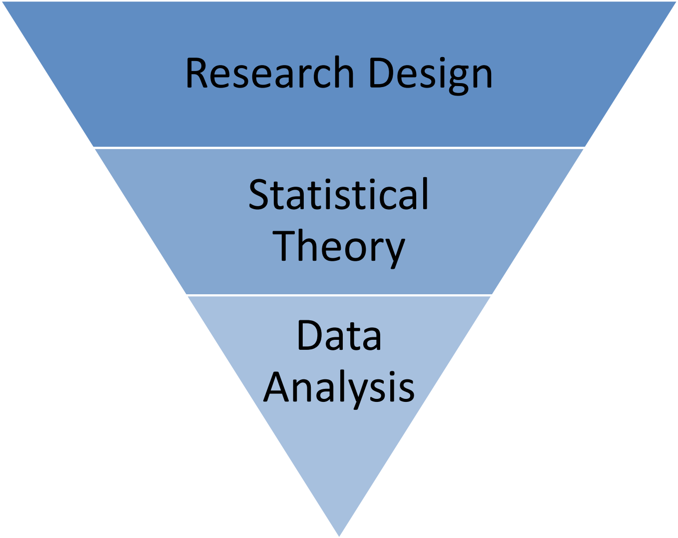
> *Studies of* **statistical inference** *seek to characterize the generally weaker conclusions that can be drawn from a finite number of observations.*

In statistics, something is **identifiable** if you can get it from the joint distribution of observable random variables.

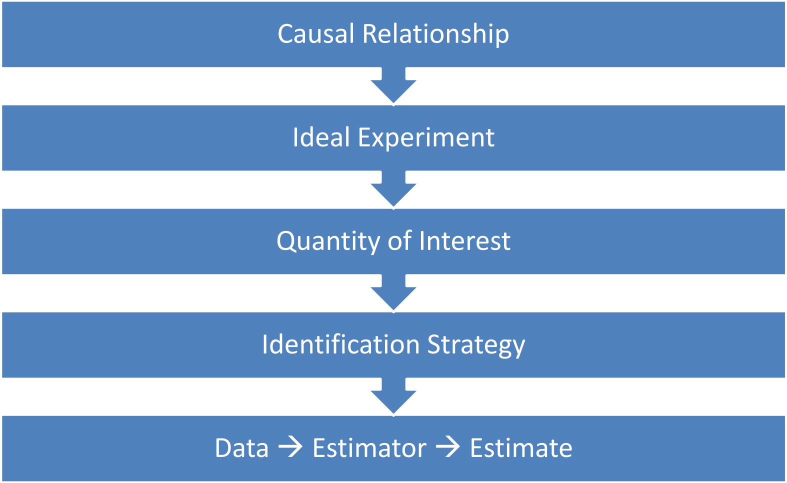
Informally, **causal identification** describes the combination of assumptions, estimation approach, and data necessary to reach a conclusion about a causal effect:

$$\text{assumptions} + \text{data} \rightarrow \text{causal conclusions}$$

How can we conduct causal inference?



Causal Inference Workflow



Roadmap for the Course

- Potential Outcomes Model
- Random Assignment
  - Design and Analysis of Experiments
- Selection on Observables
  - Matching, Regression
  - Sensitivity Analyses
- Selection on Unobservables
  - Longitudinal Research Designs: Difference-in-Differences, Panel Methods, and related methods
  - Cross-Sectional Designs: Instrumental Variables, Regression Discontinuity Design

If there is time...

Machine Learning + Causal Inference

- Heterogeneous Treatment Effects, Extrapolation of Effects
- Text + Causal Inference.

Prerequisites

- This course assumes a graduate level knowledge of linear regression, probability, and statistical computing in `R` .
- A willingness to work hard on possibly unfamiliar material
- A correlate of whether you have the background:
  - $\beta = (X'X)^{-1}X'Y$
  - $Var[X + Y] = Var[X] + Var[X] + Cov[X + Y]$
- Computing: `R`

Requirements

- Readings (best read before lecture)
  - Read slow, some material should be read multiple times, and do not skip equations.
  - Read at least one application of every technique.
- Homework assignments (35% of the final grade)
  - Posted on Wednesday, due following Wednesday before class.
  - Can work in groups, but solo attempt first.
  - Provide your own printed write-up and submit code files.
- Midterm exam (30% of the final grade). Date is April 30th.
- Final exam (30% of the final grade).
- Participation and presentation (5% of the final grade).

Housekeeping

- Piazza course website will have slides, homework, data sets, and some additional readings:
  `piazza.com/uchicago/spring2018/plsc30600/home`

- You can sign up on the Piazza course page directly from the above address. There are also free Piazza apps for mobile devices.

- The course github is at `https://github.com/justingrimmer/CausalInf` . There you'll find all the content that I post here, homeworks, code, etc.
- Use OHs and Piazza site to ask questions about the course and homework. Do not email your questions directly to the instructor or TAs.

- Office hours:
  - Justin: Open Door

Readings

- Books:
  - Angrist, Joshua D. and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
  - Morgan, Stephen L. and Christopher Winship. 2015. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*, **Second Edition**. Cambridge University Press.
- Assigned articles
  - Most linked from syllabus, a few will be posted on course website

Scheduling to Keep in Mind

1) No class on 5/28
2) I'll be absent on 4/2 and 4/4. I'll hold make up sessions on 3/30 and 4/6. Time and location TBA
3) Last day of Class is 5/30
4) Midterm is 4/30. you must be here
5) Final exam is 6/4 - 6/7. You can take it remotely.