Final Project
Stat186/Gov2002 Causal Inference

Kosuke Imai
Professor of Government and of Statistics
Harvard University

Fall 2018

The goal of the final project for this course is to write a report that can be *eventually* developed into a high-quality publishable paper. During this process, you should always have the following question in mind: *Whose mind are you trying to impact in what way?* Your contribution is judged based on the new knowledge you bring to the relevant scientific community. Why should one care about your paper? What makes your paper different from the existing research? When writing your paper, you should think about how best to convince your (skeptical) colleagues to change their way of thinking.

I encourage you to keep working on your project after this course, and I am happy to continue to provide guidance along the way. In the past, when I taught a similar course at Princeton, some papers based on final projects appeared in top refereed journals such as American Political Science Review and .

Submission

A pdf copy of your paper should be uploaded to Canvas by 5pm on December 19. No late submission is allowed.

Content Suggestions

Below are suggestions on content to include in the final report writeup:

- **Research Question and Motivation:** What is the primary research question you seek to answer? Why is it important?
- **Contributions:** What are the contributions of the paper to our understanding? What does your paper add to the existing literature?
- **Hypotheses:** What are your hypotheses? What are testable implications?
- **Data:** What data will you use to test your hypotheses? Do they reflect the population of interest?
- **Causal Identification Strategy:** What research design are you using to make causal claims? What assumptions do you require? Are the assumptions plausible in your context? Do you have empirical evidence to show the plausibility? Are there placebo tests you can conduct?

- **Results**: What are the main findings? What inferential procedures are you using? Are the results of meaningful substantive significance?
- **Robustness:** Within the context of your study, are there analyses you can conduct to see if the assumptions might hold? If they do not, what direction is the bias? How robust are your findings to the departure from the key identification assumptions?

Stylistic Requirements

Your final report should meet the following stylistic requirements. Deviating from these requirements may lower your grade.

- **Abstract:** There should be an abstract of no more than 150 words.
- **Length:** The report should not exceed 20 pages including title page, bibliography, footnotes, tables, figures, and appendices. The font size should be 12 point and the paper should be fully double spaced. Given this length limitation, I do not expect you to develop complete theoretical arguments and present comprehensive literature review. Instead, you should briefly summarize your substantive arguments and hypotheses and focus on the presentation of methodological and empirical results.
- **Section formatting:** Each section and subsection should be numbered appropriately.
- **Tables and figures:** Tables and figures are important tools for communicating your empirical results. Here are a few ground rules.
    1. Carefully choose which results are central to your study and only use tables and figures to present these results. Once you know what results to present, then choose most effective graphical methods and table formats.
    2. Tables and figures should be numbered.
    3. Tables and figures should appear where their discussions are given. They should not appear at the end of the report.
    4. Tables and figures should come with detailed captions so that they speak for themselves. Readers should be able to understand them without referring to the main text.
    5. Create figures and tables such that readers who are unfamiliar with your research can immediately understand the results. Too many numbers and digits in a table and too many lines in a figure, for example, can confuse readers. For figures, each axis should be labeled clearly and the legends should be avoided whenever possible. For tables, do not simply copy and paste outputs from statistical software. Think carefully what are the quantities of interest and present them in an intuitive manner. Often, figures are more effective means of conveying the empirical results than tables.

General Tips for Writing a Scientific Paper

Good writing skill is an essential requirement for becoming a successful researcher. Below are the guidelines I follow when writing an empirical research paper. I hope they will be helpful.

1. Do not start writing a paper until you finish all of your empirical analyses and finalize *tables and figures* first. Scientific writing is different from writing a novel. You need to know the exact contents of the paper before writing it. It is also most efficient way of writing a paper because you do not need to revise it.

2. Once figures and tables are done, determine the *title* of the paper. The title should be informative and you should not choose a title that is (or is meant to be) funny and yet does not tell readers what the paper is about.

3. Next, draft an *abstract* which should concisely describe the problem and your solution to it (or the question and your answer to it) and explain why this is a novel contribution. Carefully draft each sentence in the abstract to efficiently convey all the important information about your paper. With the 150 word limit, you do not have any sentences to waste. The abstract is typically the most difficult part of paper writing. Spend a lot of time before proceeding to the main text.

4. Now, the *introduction section* of the paper can be written by simply elaborating each sentence of the abstract. Use a couple of paragraphs to expand what you wrote with one or two sentences in the abstract. The introduction should start with a brief discussion of the motivation of your paper immediately followed by a concise summary of the main contributions. After that, you can further discuss the ways in which your methods or empirical findings contribute to the relevant literature. Do not reverse the order. The description of your findings and contributions should come before explaining the existing research.

5. Do not write a "*literature review*" by simply summarizing the existing work. Instead, describe the literature by explaining how your new methods or empirical results differ from what have been already done. The literature review should therefore be another way to demonstrate your contributions. Do not ignore prior work. Your contributions can only be understood if you tell readers how your research differs from those done by others.

6. Once the introduction is written, the rest of the paper should follow naturally. Each paragraph in the introduction becomes a section of the paper where you give further details about your methods and empirical results. Use subsections to effectively navigate readers so that they do not lose the big picture. Importantly, the paper should be written in a *top-down* manner: the beginning of a section (a subsection or even a paragraph) should give the main message and the rest should elaborate it. Try to be direct and explicit while avoiding unclear and vague sentences. Be as concise as possible.

7. For an empirical paper, be sure to precisely descrive all the details of your analysis. Readers of your paper should be able to replicate your analysis without talking to you. Transparency is one of the most important principles of scientific research.

8. During the entire process, put yourself in the position of a reader who is skeptical of your results and anticipate potential criticisms of your paper. A good scientific paper is convincing. You must justify all the major decisions you made (e.g., sample selection, choice of methods, variables and their measurements, interpretation of results) and explain why your approach is superior to other alternatives or at least reasonable. It is also important to acknowledge the limitations of your approach. Finally, the paper should give all the information necessary for replication and all data and code should be posted online upon publication.