

## Stat 156/256 Project Guidelines

The goal of this assignment is to apply methods learned in this course to analyze real-world datasets and critically appraise causal claims made in academic publications. This project is a group assignment with each group consisting of three students.

It is recommended that the project follow one of the two below forms:

1. **Replication and re-analysis** Students replicate and re-analyze the results of an academic paper whose original datasets or similar datasets are publicly available. Datasets provided by authors on the publication website are cleaned already and should match the authors published results exactly, so please do not use the cleaned datasets on the publication website unless the paper is an experimental study. As a part of the replication exercise, you and your group should download the original dataset and clean it to approximately match the sample selection used in the published paper.
2. **Applied data analysis** Students propose an applied research question for a publicly available dataset (e.g., one in an academic paper). Students select a method from class suitable for answering this question. For this type of project, you may use a cleaned dataset.

For other forms of final project assignments, such as a literature review with simulation studies to compare multiple methods, please email and schedule an office hour appointment with the GSIs.

## SUGGESTED WRITEUP FOR REPLICATION & RE-ANALYSIS

The final project writeup should be understood as a formal replication paper that is ready to be submitted to a journal. Our suggested structure of the writeup contains the following parts:

### **Paper summary and summary statistics table**

1. Summarize the paper's research question and its answer;
2. Describe the datasets used in answering the question;
3. Clean the dataset;
4. Replicate and interpret a summary statistics table that presents distributional characteristics (mean, median, IQR, etc) of key variables and covariates used in the empirical analysis.

Note: The summary table need not replicate exactly as the table produced in the paper. They say that 90% of the effort in an observational empirical paper is cleaning the data. It is fine if you are not able to clean the data exactly as described in the paper.

### **Replicate the main results**

1. Describe the empirical method in identifying the causal effect (for instance, whether the researchers conduct a randomized experiment or use policy changes to answer their research questions) and state their assumptions for identifying the causal effect clearly in both English and mathematics;
2. Replicate the main result of the paper and interpret it in English;
3. Critically appraise the stated assumptions for causal identification.

For instance, if the paper is carrying out an experiment, consider whether the experiment is balanced or if it achieves the stated goal of the author. If the paper is using a policy change or another form of "natural experiments," consider whether there would be confounding factors.

### **Replicate robustness checks/extensions**

In economic and other social sciences, after the main result, there will be a section titled "Robustness Checks" in papers about observational studies or "Extensions" in papers about experimental studies. The purpose of these sections is to convince readers that the main result holds up against various critiques on the identification assumption, or to illustrate subtleties in interpreting the main result. The writeup should complete the following:

1. Pick at least one of the robustness checks or extensions from the paper and replicate it;
2. Include a writeup explaining what the robustness check or extension achieves.

### **Re-analyze**

1. Re-analyze the main result in the paper using methods taught in this class (for example, IPW estimators, treatment effect bounds, Fisherian rerandomization tests);
2. Justify why these methods can be applied to the setting in the paper;
3. Compare and contrast your findings with the main result. If the results differ significantly, conjecture or analyze the source of discrepancies.

## SUGGESTED WRITEUP FOR APPLIED DATA ANALYSIS

### **Problem formulation**

1. Introduce and motivate the applied research problem.
2. Describe related literature and summarize their findings
3. Clearly state the research question.

### **Method**

1. Propose a method(s) from class to answer this research question. Describe the method conceptually and in math.
2. Justify why this is a suitable method for the question.
3. State in English and math the assumptions under which this method answers the question.
4. Discuss limitations with the method.

### **Results & Discussion**

1. Present results and interpret them in English.
2. Connect the results to the original research question
3. Compare the results to results of prior work
4. Revisit limitations of the method

### **Robustness Check**

In economic and other social sciences, after the main result, there will be a section titled “Robustness Checks” in papers about observational studies or “Extensions” in papers about experimental studies. The purpose of these sections is to convince readers that the main result holds up against various critiques on the identification assumption, or to illustrate subtleties in interpreting the main result.

1. Describe and conduct a robustness check against one of the key assumptions of your paper.
2. Include a discussion explaining what the robustness check achieves.

## HOW TO FIND ACADEMIC PAPERS FOR REPLICATION

### **Economics**

It is common for students to replicate papers from the American Economic Association as the Association has required authors to upload datasets and code (if it can be shared publicly) beginning in 2005. In particular, we recommend students look for empirical papers in the American Economics Review, American Economic Journal: Applied Economics, or American Economic Journal: Economic Policy. Three other major outlets for empirical papers in economics that have adopted data-sharing policies within the past five years are the Quarterly Journal of Economics, Journal of Political Economy, and Journal of Labor Economics.

### **Epidemiology and Biostatistics**

Top journals in epidemiology and biostatistics include: Biostatistics, Biometrics, American journal of Epidemiology, Statistics in Medicine, Statistical Methods in Medical Research, PLOS ONE Epidemiology.

### **Others**

Top journals in sociological and political science that have also adopted and enforced the data sharing policies are American Journal of Political Science (since 2014) and Sociological Methods & Research (since 2009).

## **Suggested (Observational Studies) Papers for Replication**

- Daron Acemoglu and Joshua D. Angrist, “Consequences of Employment Protection? The Case of the Americans with Disabilities Act”, Journal of Political Economy, 2001
- Daron Acemoglu, David H. Autor and David Lyle, “Women, War and Wages: The Effect of Female Labor Supply on the Wage Structure at Mid-Century”, Journal of Political Economy, 2004
- Elizabeth O. Ananat, “The Wrong Side(s) of the Tracks: The Causal Effects of Racial Segregation on Urban Poverty and Inequality”, American Economic Journal: Applied Economics, 2011
- Maximilian Auffhammer and Ryan Kellogg, “Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality”, American Economic Review, 2011
- Patricia Cortes and Jose Tessada, “Low-Skilled Immigration and the Labor Supply of Highly Skilled Women”, American Economic Journal: Applied Economics, 2011
- David Deming, “Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start”, American Economic Journal: Applied Economics, 2009
- Daniel K. Fetter, “How Do Mortgage Subsidies Affect Home Ownership? Evidence from the Mid-Century GI Bills”, American Economic Journal: Economic Policy
- Alexander M. Gelber, “How Do 401(k)s Affect Saving? Evidence from Changes in 401(k) Eligibility”, American Economic Journal: Economic Policy, 2011
- Jonathan Gruber and Samuel A. Kleiner, “Do Strikes Kill? Evidence from New York State”, American Economic Journal: Economic Policy, 2012
- Hilary W. Hoynes and Diane Schanzenbach, “Consumption Responses to In-Kind

Transfers: Evidence from the Introduction of the Food Stamp Program”, American Economic Journal: Applied Economics, 2009

- David S. Johnson, Jonathan A. Parker and Nicholas S. Souleles, “Household Expenditure and Income Tax Rebates of 2001”, American Economic Review, 2006
- David S. Johnson, Robert McClelland, Jonathan A. Parker and Nicholas S. Souleles, “Consumer Spending and the Economic Stimulus Payments of 2008”, American Economic Review, 2013
- Melissa S. Kearney and Phillip B. Levine, “Early Childhood Education by Television: Lessons from Sesame Street”, American Economic Journal: Applied Economics, 2019
- Jeanne LaFortune, “Making Yourself Attractive: Pre-marital Investments and the Returns to Education in the Marriage Market”, American Economic Journal: Applied Economics, 2013
- Phillip B. Levine, Robin McKnight and Samantha Heep, “How Effective Are Public Policies to Increase Health Insurance Coverage among Young Adults?”, American Economic Journal: Economic Policy, 2011
- Gianmarco Ottaviano, Giovanni Peri and Greg C. Wright, “Immigration, Offshoring and American Jobs”, American Economic Review, 2013
- Albert Saiz and Susan Wachter, “Immigration and the Neighborhood”, American Economic Journal: Economic Policy, 2011
- Betsey Stevenson and Justin Wolfers, “Bargaining in the Shadow of the Law: Divorce Laws and Family Distress”, Quarterly Journal of Economics, 2006
- Li L, Greene T. ”A weighting analogue to pair matching in propensity score analysis”, The international journal of biostatistics. 2013 Jul 31;9(2):215-34.
- Baiocchi M, Cheng J, Small DS. Tutorial in Biostatistics: Instrumental Variable Methods for Causal Inference.
- Kurth, Tobias, et al. ”Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect.” American journal of epidemiology 163.3 (2006): 262-270.
- Franklin, Jessica M., et al. ”Comparing the performance of propensity score methods in healthcare database studies with rare outcomes.” Statistics in medicine 36.12 (2017): 1946-1963.
- Austin, Peter C., and Elizabeth A. Stuart. ”Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies.” Statistics in medicine 34.28 (2015): 3661-3679.

#### HOW TO FIND DATASETS

- ICPSR – a user-supported data repository at the University of Michigan that contains thousands of datasets. Most major universities subscribe to ICPSR and datasets are downloadable from the website after registration.
- IPUMS – a data repository at the University of Minnesota that contains harmonized census and survey datasets from around the world.
- Panel Study of Income Dynamics – the longest running longitudinal household survey in the world, abundantly used by social scientists.
- National Longitudinal Survey of Youth – longitudinal data of young adults, also commonly used by social scientists.

- National Health and Nutrition Examination Surveys – a health study commonly used in epidemiology and public health.
- Harvard Dataverse – a code and data repository for a broad selection of academic papers. When using the Dataverse, be careful to find data that is not already cleaned.
- Home of US Government’s open data - Many US Government’s open data set. For example, there are many Financial data sets on data.gov.
- Applied Econometrics Data Archive.

## GRADING

### Presentation

- Relevance (10%): Is the subject of the project relevant to causal inference? This should be straightforward to get full credit for. If you have questions about relevance for your particular project, ask the instructor or GSI.
- Correctness (20%): Is the material presented technically sound? Points will be deducted from this category if the presentation makes major technical mistakes.
- Content (20%): The presentation should motivate the problem setting, briefly mention 1-3 related works, introduce notation, describe the methodology, and present preliminary results.
- Presentation (50%): Is the presentation well-communicated? The presentation should be accessible to someone who has sat through STAT 156/256. The speakers should make eye contact and appropriately pace their talk. This component includes presentation of the slides which should be informative and not overloaded with text.

### Report

- Relevance (10%): Is the subject of the project relevant to causal inference? This should be straightforward to get full credit for. If you have questions about relevance for your particular project, ask the instructor or GSI.
- Correctness (20%): Is the material presented technically sound? Points will be deducted from this category if the report makes major technical mistakes.
- Content (50%): The report should motivate the problem setting, provide relevant related work, introduce notation, describe the methodology undertaken in technical detail and explain why it is relevant for the problem at hand, report results and quantify uncertainty, discuss limitations of the method, and provide a conclusion that highlights the key contribution. Points will be deducted from this category if the paper fails to adequately engage in these areas. For more detailed suggestions, see the project guidelines.
- Clarity (20%): The report should be well-written. The report should use precise statistical terminology and mathematical notation. The report content should be accessible to someone who has sat through STAT 156/256. We recommend using LaTeX to provide well-formatted papers that make use of tables, figures, section headings, etc as appropriate.

*Modified from Peng Ding’s project assignment in fall 2022.*