Problem Set 3

STAT II (Spring 2025)

Disclaimer: *Please read the guidelines below carefully. Good luck!*

Guidelines

- Please upload your answers to Brightspace by **Thursday, April 15th, at 11:59 PM**.
- Direct all questions about the problem set to Brightspace under Contents > Discussions > Class Feed > Problem Set 3.
- For precise and expedited responses, we will address questions about this problem set on Brightspace until 6:00 PM on April 14th.
- Collaboration is allowed, but I encourage you to attempt the problems on your own before seeking help from others. Regardless of collaboration, you must individually write up and submit your answers.
- Late submissions will not be accepted unless prior approval is obtained from the instructor at least one day before the due date.
- **The total points for the problem set are 150**.
- **Grading:** Show every step of your derivations. We grade the steps of derivations as well as the final answers. If you are unable to solve the problem completely, partial credit can be given for your derivations. Conversely, a correct final answer without complete derivations may not receive full credit.
- **Stylistic Requirements:** Adhere to the guidelines for the format of your submitted answers. Ensure you follow these rules:

    1. Submit your answers as a PDF file compiled from LaTeX, R Markdown, or (ideally) a Quarto Markdown document.
    2. If using raw LaTeX(.tex) for your answers, submit an accompanying .R file for any computational tasks, with referenced line numbers corresponding to each specific task.
    3. If using R Markdown (.Rmd) or Quarto Markdown (.qmd), include your code as code chunks in the source file. Additionally, submit the source .Rmd or .qmd file along with the compiled PDF to allow us to run your code easily.
    4. To ensure reproducibility of your simulation results, use `set.seed()` at the beginning of your document.

Problem 1. Propensity Score and Weighting (25 points)

In this question, we learn more about the properties of propensity score and weighting estimators. We consider the binary treatment variable $T_i \in \{0, 1\}$, which takes 1 if unit i is in the treatment group and 0 if unit i is in the control group. We define two potential outcomes $\{Y_i(1), Y_i(0)\}$ for each unit and make the standard consistency assumption, i.e., $Y_i = T_i Y_i(1) + (1 - T_i)Y_i(0)$. We observe pre-treatment covariates $\mathbf{X}_i$ for each unit i. We assume throughout this question that we obtain n i.i.d. samples $\{Y_i, T_i, \mathbf{X}_i\}_{i=1}^n$ where n is the sample size.

(a) We learned in the class that we can identify the ATE by conditioning only on propensity scores. To prove this result, we first establish an important lemma. Suppose the following two identification assumptions hold:

1. **Conditional Ignorability:** $\{Y_i(1), Y_i(0)\} \perp\!\!\!\perp T_i \,|\, \mathbf{X}_i = \mathbf{x}$ for any $\mathbf{x} \in \mathcal{X}$ where $\mathcal{X}$ is the support of $\mathbf{X}_i$.
2. **Positivity:** $0 < \Pr(T_i = 1 \,|\, \mathbf{X}_i = \mathbf{x}) < 1$ for any $\mathbf{x} \in \mathcal{X}$.

Under these two assumptions, please prove that the following conditional independence holds:

$$\{Y_i(1), Y_i(0)\} \perp\!\!\!\perp T_i \;\mid\; \pi(\mathbf{X}_i),$$

where $\pi(\mathbf{X}_i)$ is the propensity score, i.e., $\pi(\mathbf{X}_i) \equiv \Pr(T_i = 1 \,|\, \mathbf{X}_i)$. This conditional independence means that the treatment variable is independent of the potential outcomes conditional only on the propensity score.

(b) Please explain in words the benefit of the approach based on propensity score $\pi(\mathbf{X}_i)$ as opposed to adjustment for $\mathbf{X}_i$.

Note: *There is no "one correct" answer. We will evaluate your answers primarily based on how well you apply concepts we learn in the class to investigate the benefit of propensity score.*

Problem 2. Instrumental Variables (25 points)

In this problem, we explore formal properties of the instrumental variable (IV) approach. We assume a binary treatment variable as in class, and adopt the following notation:

- **Units:** Indexed by $i \in \{1, \ldots, n\}$, where n is the total number of units.
- **Treatment:** T_i is a binary treatment variable for unit i, with $T_i \in \{0, 1\}$.
- **Instrument:** Z_i is a binary instrumental variable for unit i, with $Z_i \in \{0, 1\}$.
- **Potential Treatment:** $T_i(z)$ is the treatment received by unit i when $Z_i = z$.
- **Potential Outcome:** $Y_i(z, t)$ is the outcome for unit i when $Z_i = z$ and $T_i = t$.

For formalization, define the compliance type S_i as follows:

- **Complier ($S_i = C$):** $T_i(1) = 1$ and $T_i(0) = 0$.
- **Never-Taker ($S_i = N$):** $T_i(1) = 0$ and $T_i(0) = 0$.
- **Always-Taker ($S_i = A$):** $T_i(1) = 1$ and $T_i(0) = 1$.
- **Defier ($S_i = D$):** $T_i(1) = 0$ and $T_i(0) = 1$.

**Hint**: Slides 22-23 on Instrumental Variables should be useful for this problem.*

(a) Define the local average treatment effect (LATE) in terms of the potential outcomes. Prove (showing all steps, not just the final result) that the IV estimand

$$\frac{\mathbb{E}(Y_i \mid Z_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0)}{\mathbb{E}(T_i \mid Z_i = 1) - \mathbb{E}(T_i \mid Z_i = 0)}$$

is equal to the LATE under the four core assumptions discussed in class. Clearly state each assumption in terms of the potential outcomes and indicate at each step of your proof which assumption you are using.

(b) Now, relax some of the assumptions used in part (a). Specifically, assume that the exclusion restriction holds **only for compliers**. Prove (showing all steps, not just the final result) that the IV estimand does not equal the LATE. Derive the exact expression for the bias, defined as the IV estimand minus the LATE.

(c) Finally, consider the case where the **monotonicity** assumption is violated *while maintaining the exclusion restriction*. Explain *in words* what is the source of bias of IV estimand (compared to the LATE) in this case?

Problem 3. Instrumental Variables (Group Assignment; 50 points)

In this problem, you will need to split into groups of 2-3 people and then find and examine an observational study where researchers rely on the instrumental variable approach. It should satisfy the following criteria:

1. **Journals**: It should be from the following three journals: *American Political Science Review, American Journal of Political Science*, or *Journal of Politics*.
2. **Replication Data**: The paper should have a publicly available replication file. For the above three journals, they require authors to upload replication files to the following archives in recent years. So, if you pick a paper from recent years, it should be straightforward to find a replication file.
    - American Political Science Review: https://dataverse.harvard.edu/dataverse/the_review
    - American Journal of Political Science: https://dataverse.harvard.edu/dataverse/ajps
    - Journal of Politics: https://dataverse.harvard.edu/dataverse/jop
3. **Google Spreadsheet**: Each group should find a different paper. To make sure every group finds a different paper, as soon as your group finds a paper, please write it down in this Google Spreadsheet (https://docs.google.com/spreadsheets/d/1ElMZMC7zMIw73RIqm3PGVzRDmg3clA28rac_NPEosBU/edit?usp=sharing). We use the "first come first served" rule.
4. **New Application**: You cannot choose papers that are discussed or mentioned in lectures or sections. You cannot choose a paper from a list at the end of slides on Instrumental Variables.

Note: *Do not spend too much time looking for a "perfect" paper that perfectly fits your research interests, rather be less selective and treat this exercise as if you were asked to review the paper by a journal or for a conference. There is no "one correct" answer to most of the following questions. We will evaluate your answers primarily based on how well you apply concepts we learn in the class to investigate a study you choose.*

(a) Please first explain the overall research question of the paper, the data and contexts of the study, and the main findings.

Then, please explain what are the outcome variable, the treatment variable, and the instrumental variable in the study.

(b) Then, please examine the identification assumptions necessary for the IV analysis. How do the authors justify those IV assumptions? Please investigate each of the four assumptions we covered in the class. Do they discuss potential violations of assumptions? If so, what do they do to address the potential violation of those IV assumptions?

Note: *Please focus your answer to this question on how* the original authors *discuss and address the IV assumptions. You will evaluate them in the next question!*

(c) Please evaluate the identification assumptions the authors make. Are they plausible? If so, why? If not, please provide a concrete substantive argument that suggests potential violations of the assumptions. Again, please investigate all the four assumptions we covered in the class.

(d) What methodological limitations do *you* think the study has? You can discuss points related to the IV or any other methodological problems. What would you do to address the methodological limitations you identified?

Problem 4. Difference-in-Differences Design (50 points)

Bechtel and Hainmueller (2011) use the difference-in-differences design to estimate the causal impact of beneficial policies. The data for the paper is available online at the Harvard Dataverse (Link). Please check their replication file for definitions of each variable in the data.

(a) Please read it carefully before you work on this question. Define each of the following variables in the context of this study. Note that you might need to define some of the variables separately for the short- and long-term effects.

1. Treatment group
2. Treatment variable
3. Pre-treatment period and post-treatment period
4. Outcome variable

(b) Some scholars might argue that the 2002 Elbe flooding in Germany can be considered quasi-random. Suppose that they then propose to use a difference-in-means estimator, rather than the difference-in-differences design, to estimate the causal effect of beneficial policies.

Explain why the difference-in-differences design might be a more plausible approach to make causal inference in this study. After describing your reasoning, provide empirical evidence using the data.

(c) The difference-in-differences design requires the "parallel trends assumption." Discuss the substantive meaning of this assumption in the context of this study. How did the authors check the plausibility of this assumption? (Note: they could not directly test this assumption, but they checked its plausibility). Provide a plot that assesses the parallel trends assumption in this study. Finally, provide a formal statistical test that assesses the parallel trends assumption.

(d) What are the main quantities of interest in this study? Provide formal definitions and then interpret them in the context of this study. Note that the authors consider not only one but several estimands in Table 1 (page 857).

(e) Estimate these main quantities of interest defined in (d) using the difference-in-differences estimator without including any control variables. Please report point estimates, standard errors, and the 95% confidence intervals. When computing standard errors, please cluster standard errors at the district level. Finally, please briefly interpret the result.

(f) Estimate these main quantities of interest defined in (d) using the difference-in-differences estimator while including control variables that the original authors included (control variables used in Columns (3) and (6) of Table 1 in the paper). Due to the data limitation, you can focus on the effect on 2002 and 2005 (excluding 2009). Please report point estimates, standard errors, and the 95% confidence intervals. When computing standard errors, please cluster standard errors at the district level. Finally, please briefly interpret the result.

References

Bechtel, Michael M, and Jens Hainmueller. 2011. "How Lasting Is Voter Gratitude? An Analysis of the Short-and Long-Term Electoral Returns to Beneficial Policy." *American Journal of Political Science* 55 (4): 852–68.