PROBLEM SET 1

Jacob Jameson

Due on Tuesday, February 14, 2023

IDENTIFICATION

(1) Your information

Jacob Jameson

(2) Group Members (please list below the classmates you worked with on this problem set):

Bohan Li, Jenna Rogers

(3) Compliance with Harvard Kennedy School Academic Code

I certify that my work in this problem set complies with the Harvard Kennedy School Academic Code

Conceptual Questions (30 points + 8 extra points)

1. Read the paper. Clearly state the primary research question that the authors are trying to answer. What makes this an interesting question? (2 points)

2. Explain the main finding of the paper (including what the 'flypaper effect' is) using non-technical jargon, as if you were writing a brief policy memo. Provide two versions:

a. In 3-5 sentences (without copying the article abstract or the text): (2 points)

b. In a tweet, i.e. 280 characters or less: (1 point)

3. Summarize the specific details of the treatment that participating businesses in this country underwent. (2 points)

4. The authors note that their experimental design is very similar to an earlier study in Sri Lanka. What reasons do they give to conduct a separate study, rather than expecting the findings of that research to apply in this context? (2 points)

5. The authors used a randomized control trial because they believed an observational analysis of similar policies would be insufficient. Imagine that another country implemented the investment program without randomizing treatment, and that you were trying to understand the effect of this program on food security in that country. What are two possible confounders (omitted variables) that would bias the results from your observational analysis? Explain the mechanism of the omitted variable and use the omitted variable bias formula to argue whether it would lead to an understatement or overstatement of the true effect. (3 points)

6. Let Y_{0i} be the business profits in the absence of the grants program and let $D_i = 1$ denote participation in the program. If there were no RCT and individuals were allowed to opt-in to the program, critics might point out that: (3 points)

$$E(Y_{0i}|D_i = 1) < E(Y_{0i}|D_i = 0)$$

Explain this equation in words, explain why it is a problem, and give a plausible scenario in which that may be the case.

7. Defining treatment as being assigned to the grants program, what is the difference between the ITT and TOT in this context? Which do the authors report and why do they make this decision? Write an equation (using potential outcomes notation) that shows what the authors are trying to estimate. (3 points)

8. At what level do the authors clustered their standard errors of the main results of the paper (if at all)? Briefly note why the authors cluster the standard errors and why this is the appropriate level to cluster at. (2 extra points)

9. To assess whether treatment was actually randomly assigned, we can examine the results of a balance test, presented in Table 2. Do the results in this table make you more or less confident about the validity of the paper's results? Interpret one of the p-values from column (5). (3 points)

10. Attrition in experiments like this one is often a concern for internal validity. Does the particular sort of attrition mentioned in this article give you reason to be concerned about the validity in this study? Describe using particular aspects of the experiment or its implementation. (1 point)

11. What other threats to internal validity may have affected this this experiment? Choose one threat and explain how it might bias the coefficient of interest. (2 points)

12. Describe four specific problems involved with generalizing the results of this study as a result of using an RCT. Hint: review the Muralidharan and Niehaus (2017) or Banerjee et al. (2017) papers discussed in class. (4 points)

13. List at least two strategies the authors use to address some of the concerns you described above. (2 points)

14. Why do the authors include Table 4? (2 extra points)

15. Do you think the main results would be the same if this experiment were expanded to slightly larger businesses? Describe plausible scenarios in which providing the same transfer to these different businesses could both lead to (i) a larger increase in profits, and (ii) a smaller increase in profits. (2 extra points)

16. If you were a researcher at the World Bank interested in scaling up one or several of these treatments, what follow-up study would you propose to expand on these findings? Explain in 4-5 sentences as if you were trying to convince a policymaker of the need for additional research. (2 extra points)

Data Analysis Questions (22 points + 2 extra points)

17. Produce a well-organized descriptive statistics table that includes (i) the number of households, (ii) the number of geographic units, (iii) the number of units of randomization, alongside (iv) the sample mean and standard error of the income and revenue index in the control group, and (v) the same sample mean and standard error of the same index but in the treatment group. In other words, the table should have one row (countries), and six columns (including country).

a. Print your table below. (6 points)

b. Are the differences in baseline profits between the control and treatment groups significant at the 0.05 level? (3 points)

18. Reproduce the coefficient estimate and standard error estimates in columns (1-2) of Table 3: these correspond to the main pooled OLS specifications (Equation 5 in the paper) without and with sample trimming respectively. (3 points)