
EC 320: Introduction to Econometrics

Problem Set 2

Due: On canvas, January 27th at 5 pm

Learning Outcomes:

- Understanding individual and average treatment effects
- Understanding the fundamental problem of econometrics and causality
- Run regressions and interpret results

Checklist Before Handing In:

- Did you answer all questions?
- Did you answer all parts for each question?
- Were your answers too vague? If so, make them more precise to make sure they really answer the question being asked.

Instructions: You are encouraged to work with other students in the class, but you must provide original responses. To receive full credit, justify your answers and list your collaborators. For full credit on the computational exercises, include your code and output in addition to your answers. You will turn in digital copies of your responses on Canvas. Please note the list of acceptable file types on the submission page.

Name:

Collaborator 1:

Collaborator 2:

Collaborator 3:

Analytical Questions

1. Consider hypothetical data on the counterfactual outcomes of 6 individuals:

i	Treatment $_i$	$Y_{1,i}$	$Y_{0,i}$
1	0	4	8
2	0	9	9
3	0	3	7
4	1	10	14
5	1	11	12
6	1	14	19

- (a) Calculate the individual treatment effects τ_i . Are the treatment effects constant? In other words, is it the case that $\tau_i = \bar{\tau}$ for each individual?
 - (b) Calculate the average treatment effect.
 - (c) Why is it impossible to observe data on counterfactual outcomes in real life? How does the inability to observe those data relate to the fundamental problem of econometrics?
 - (d) Estimate the average treatment effect by comparing the mean of the treatment group to the mean of the control group.
 - (e) Do you think that the estimator in part (d) is unbiased? Why or why not?
2. Consider the research question, “*does binge drinking induce violent behavior among college students?*” Use it to address the following:
- (a) What is the outcome variable? What is the treatment?
 - (b) Define the counterfactual outcomes $Y_{1,i}$ and $Y_{0,i}$.
 - (c) What are some plausible causal mechanisms that run directly from the treatment to the outcome?
 - (d) What are some potential sources of selection bias in the raw comparison of outcomes by treatment status? Do you expect the bias to be positive or negative? Why?

-
3. Suppose that a school district conducts a randomized control trial of a reading intervention for low-performing 8th-grade students. To determine whether the intervention improved reading test scores, the district's data analysts estimated a regression using ordinary least squares (OLS). The results are summarized in the table below:

	Reading Test Score
Treatment	5.231 (2.015)
Intercept	50.812 (12.887)
Observations	956

Note: Standard errors in parentheses.

- (a) Write down the regression model the analysts estimated.
- (b) What is the average reading test score of students in the control group after the intervention?
- (c) What is the estimated treatment effect?
- (d) Using the rule of thumb discussed in class, determine whether the treatment effect is statistically distinguishable from zero. Explain.
- (e) What is the average reading test score of students in the treatment group after the intervention?
- (f) Under what conditions would the randomized control trial isolate the causal effect of the reading intervention on reading test scores?
- (g) Suppose that the randomly selected control group happened to have twice as many non-native English speakers as the randomly selected treatment group. Non-native speakers score lower on standardized (English) reading tests, on average. In light of this information, is it likely that the estimated treatment effect isolates the effect of the intervention? Why or why not? If not, what is the sign of the bias?

Computational Questions

For this portion of the problem set, you will use the file `bm.csv` in the `Problem_Set_2` folder on Canvas. The file contains experimental data from the influential study *Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination* by Bertrand and Mullainathan (2004). The authors study the effect of job applicant race on hiring managers' callback decisions by sending fictitious résumés to real job postings. To isolate the effect of race, the authors randomly determine whether each résumé has a “black-sounding” name or a “white-sounding” name.

Variable Name	Description
<code>black</code>	= 1 if résumé has a “black-sounding” name
<code>female</code>	= 1 if résumé has a “female-sounding” name
<code>call</code>	= 1 if résumé received a callback from a real employer
<code>educ</code>	Level of education reported on the résumé
<code>num_jobs</code>	Number of jobs listed on the résumé
<code>exper</code>	Years of work experience on the résumé
<code>comp_skill</code>	= 1 if computer skills are mentioned on the résumé

Use the data to complete the tasks and questions below.

1. What percentage of résumés received a callback?

Hint: The mean of a binary variable gives the fraction of observations with a value of 1.

2. Calculate the callback rate for each race (in this case, black and white). Do employers appear to consider applicants' race when making callback decisions? Explain.
3. What is the difference in average callback rates between black and white applicants? Can you conclude that employers discriminate against black applicants?
4. Without running a regression, conduct a t -test for the difference in callback rates by race. Is the difference in callback rates statistically significant?
5. Run a regression of `call` on `black`.
 - (a) Does the estimated coefficient on `black` match the difference-in-means you estimated in exercise 3?
 - (b) Using the rule of thumb discussed in class, determine whether the estimated coefficient on `black` is statistically significant.
6. Run a regression of `call` on `black` with controls for education and experience.
 - (a) Report the results of the new regressions and the regression you estimated in exercise 5 in a `stargazer` table.

-
- (b) Does the coefficient on **black** change with the additional controls?
7. Are the results of the field experiment consistent with the notion that employers discriminate based on race? Explain.