A-164/API-211

Rucinski

Program Evaluation (Spring 2025)

**Problem Set 1:**

**Interpreting Regression Coefficients and Omitted Variable Bias**

**Problem Set Policies:**

* This assignment is due by **5pm on Friday, February 7th**. Submit your solutions to the assignment on Canvas as a Word doc or PDF. No coding is required on this problem set.
* You may take an automatic extension of 24 hours (to 5pm on Saturday, February 8th) without contacting the teaching team, but we encourage you to submit by the deadline to give yourself a break over the weekend.
* You are welcome to discuss the problem set in small groups, and to submit questions to the teaching team and other students via email or Slack. Please write the names of any students you worked closely with at the top of your submission. **You must submit individually, and** **all work you submit should be your own**. To avoid misunderstandings, we ask that you **do not share any *written* answers to these questions** in shared documents, emails, chats, etc.
* Please familiarize yourself with [HGSE policy on the use of generative AI](https://registrar.gse.harvard.edu/AI-policy). You may use ChatGPT, Claude, etc. to refine or translate your ideas. You may **not** use AI to *generate* ideas for this or any other assignment. If you do use AI to assist you in completing this assignment, please credit the tool you used as you would a human collaborator, and explain how you used it.
* If you have questions about appropriate use of AI or find yourself turning to AI to write answers for this assignment, or if you are considering intentionally copying an answer written by another student, please be in touch with a member of the teaching team. We would far prefer to help you in office hours or grant you an extension than to have a conversation about academic misconduct.
* Once your assignment is graded, please review the feedback carefully to support your success on future assignments.

Suppose you observe test scores and class size for students in Tennessee. You run the following regression with this data:

(eq. 1)

In this regression equation,   represents the scaled standardized test score for student *i*, andisan indicator variable that equals 1 if a student was in a small class, and 0 otherwise.

1. Express and  using the notation of the potential outcomes framework. Do you believe that represents the causal effect of class size on student achievement? Explain why or why not in 1-2 sentences.
2. Suggest one omitted variable that would bias your estimate of the causal effect of class size in question 1. Explain if this omitted variable would lead you to overestimate or underestimate the effect of small class size.

The Tennessee STAR experiment randomly assigned students to small and large classes. Krueger (1999) used data from this experiment to measure the effect of class size on student achievement.

1. Table V in the paper (page 512) shows results from several versions of the regression in Question 1 (eq.1). Use regression (8) in Table V to report the experimental estimate of for **kindergarten** students.
2. Interpret Krueger’s estimate of in a sentence**.** Does this estimate of represent a correlational or causal estimate of the impact of class size on test scores? Briefly explain why.

1. Review Tables I and II in Krueger (1999). These tables report results from balance tests for student characteristics across small and large classrooms. Is there any evidence in this table that indicates a potential issue with random assignment? Explain.

1. In the Tennessee STAR experiment, both students and teachers were randomly assigned to small and large classrooms. Explain why it is important to randomly assign not just students but also teachers to estimate the causal effect of class size.

1. Krueger estimates a positive effect of class size on student achievement. If teachers were **not** randomly assigned to classrooms, and teachers with **more** experience tended to be assigned to **larger** classes, would this estimated effect be biased upward or downward relative to the true effect?  Explain.