
API-202B Empirical Methods II

Session #6: Omitted variables case study

miguel_santos@hks.harvard.edu
@miguelsantos12

Omitted variable bias (OVB): Case Study

- Today we'll discuss the IMPACT system for teacher evaluation that was adopted by the Washington, D.C. public schools
- We will focus on the regression-based component of the system
- The goal is to measure teacher quality
- For purposes of today's discussion, we'll adopt a narrow definition of quality, namely how much a teacher improves his or her students' test scores (we assume student's test scores are a good indicator of teachers' quality)

The IMPACT Tool

- Below is the student-level regression used by IMPACT:

$$\begin{aligned} \text{Math}_{ig} = & \beta_1 \text{Math}_{i(g-1)} + \beta_2 \text{Read}_{i(g-1)} + \beta_3 \text{Female}_{ig} + \beta_4 \text{FreeLunch}_{ig} + \beta_5 \text{SpecialEd}_{ig} + \beta_6 \text{LEP}_{ig} + \beta_7 \text{Abs}_{ig} + \\ & \alpha_1 \text{Teacher1}_{ig} + \alpha_2 \text{Teacher2}_{ig} + \alpha_3 \text{Teacher3}_{ig} + \dots + \alpha_N \text{TeacherN}_{ig} + \varepsilon_{ig} \end{aligned}$$

- Why might coefficients on teacher dummies measure teacher quality while keeping constant other relevant factors? Why is this called value-added modeling?

The IMPACT Tool

- Below is the student-level regression used by IMPACT:

$$\begin{aligned} \text{Math}_{ig} = & \beta_1 \text{Math}_{i(g-1)} + \beta_2 \text{Read}_{i(g-1)} + \beta_3 \text{Female}_{ig} + \beta_4 \text{FreeLunch}_{ig} + \beta_5 \text{SpecialEd}_{ig} + \beta_6 \text{LEP}_{ig} + \beta_7 \text{Abs}_{ig} + \\ & \alpha_1 \text{Teacher1}_{ig} + \alpha_2 \text{Teacher2}_{ig} + \alpha_3 \text{Teacher3}_{ig} + \dots + \alpha_N \text{TeacherN}_{ig} + \varepsilon_{ig} \end{aligned}$$

- What advantages was IMPACT supposed to have over prior system? What are its shortcomings?

The IMPACT Tool: Factors affecting student test scores

- We want to isolate an individual teacher's impact on test scores

$$\text{Math}_{ig} = \beta_1 \text{Math}_{i(g-1)} + \beta_2 \text{Read}_{i(g-1)} + \beta_3 \text{Female}_{ig} + \beta_4 \text{FreeLunch}_{ig} + \beta_5 \text{SpecialEd}_{ig} + \beta_6 \text{LEP}_{ig} + \beta_7 \text{Abs}_{ig} + \alpha_1 \text{Teacher1}_{ig} + \alpha_2 \text{Teacher2}_{ig} + \alpha_3 \text{Teacher3}_{ig} + \dots + \alpha_N \text{TeacherN}_{ig} + \varepsilon_{ig}$$

IMPACT does control for

IMPACT doesn't control fo

Is OVB a problem here?

“The teacher matters, absolutely. But they’re not the only thing that matters... is it only the teacher’s responsibility for student achievement, or ... does the school, does the district, does the community play a role?”

Rob Weil (union representative)

“Teachers are not afraid of accountability. What they’re afraid of is being treated unfairly. And when you can explain to them... how the value-added growth model actually controls for factors that are outside of their control... the vast majority that I talk to say, ‘Okay, that makes sense.’”

Michelle Rhee

- Let’s assume you are a teacher in DC and IMPACT has determined you did a poor job of improving your students’ scores.
- Which of these would be a valid objection to this finding?

Is OVB a problem here?

“The new system did not take into account that some classes might have more low-skilled students”

Is OVB a problem here?

“My class this year had an unusually large number of students whose parents have been unemployed for years.”

Is OVB a problem here?

“My class this year had an unusually large number of students whose parents recently lost their jobs.”

Policy implications

What would you do? What are the implications of this case for any of you aiming at having an impact in the quality of education?

Takeaways

- Prior test scores may serve to control for factors not explicitly controlled for: The extent of OVB depends on how well prior test scores account for other factors correlating with test scores and probability of being in a given classroom
- There's a difference between successfully identifying each teacher's quality and getting it right on average (bias vs. noise)
- IMPACT may be flawed, but still may be better than other alternatives
- New system might incentivize new and perverse behavior (teaching to the test, cheating, etc.)
- Great example of government using multivariate regression for policy ends
- Most economists think value-added measures contain information, though should be used in conjunction with other measures
- Some research at Harvard (John Friedman, Raj Chetty, Jonah Rockoff) shows these measures of teacher quality are good predictors of students' college admission and earnings later in life

Takeaways

Changing things is hard! Even if we are doing things right, and they will improve the targets (students) prospects... there are always winners and losers...