Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India

ECON 280 – Part 2

Michael Partridge

October 25, 2025

1 Summary of paper

1.1 Main question and contribution

This paper investigates the effectiveness of technology-based learning amplifiers in a developing country setting. In particular, it studies a personalized technology-assisted afterschool instruction program called Mindspark in Delhi, India. The main question of the paper is: can a personalized and digitalized program benefit students when there is large heterogeneity of learning levels across these students in a particular grade? They find that the answer to this question is yes and that there are actually large increases in test scores across all students. Students with the low learning levels prior see the biggest relative boost, while students at the top perform better with their current grade-level material, as a result. There appear to be large potential returns to scale with this program and could be a solution to help with this heterogeneity of learning level issue of students in India (and other developing countries). Other after school programs, such as tutoring, have not been shown to have the same benefit as this program.

1.2 Empirical framework

This study ran an experiment where they first recruited families that would like to be enrolled into this after-school personalized program, and they then randomly allowed half of the students to enroll for free while the other students were not allowed to enroll. The program lasted about 90 days, which is half the school year. The average attendance was 58%. Their first set of analysis looks at ITT estimates using this 58% baseline, which shows a positive impact on students' test scores. They then want to investigate the effect on attending an additional day on test scores. To do this, they instrument for attendance

with the randomized offer of a voucher. They then identify the average effect of an extra day of attending the Mindspark centers on test scores. They call this the average causal response of the treatment.

2 Dataset description

The first dataset used in the analysis code of the project contains info of students, such as test scores and other characteristics of them and their households. There are 46 variables and 1,158 observations in the dataset. The unit of observation is a student by round (either baseline or endline of the treatment), so it would contain test scores of both of the baseline and endline exams. Some students have missing values for certain variables, such as age, but these don't seem to be an issue for the analysis.

Below, I create a histogram to view test scores of students for their baseline vs endline math exams. The continuous variable I observe here is the percent score on the math exam.

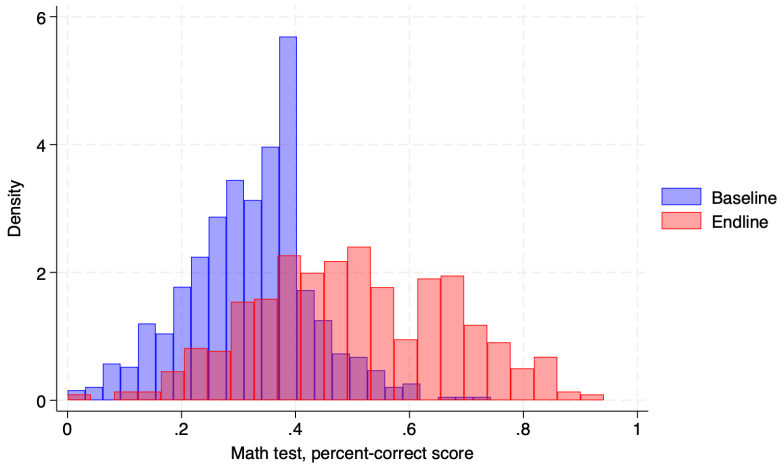

Figure 1: Distribution of math test scores for baseline vs endline exams