

This paper investigates the impact of a personalized technology-aided after-school instruction program called Mindspark on middle-school students in urban India using a randomized lottery design. The study finds that lottery winners scored  $0.37\sigma$  higher in math and  $0.23\sigma$  higher in Hindi over a 4.5-month period, with similar absolute gains across all student ability levels but particularly large relative gains for academically-weaker students. The key contribution is demonstrating that well-designed computer-aided learning (CAL) programs can dramatically improve educational productivity by using technology to personalize instruction and "teach at the right level" for each student, thereby addressing a critical binding constraint in developing country education systems: the mismatch between classroom instruction levels and the wide heterogeneity in actual student learning levels within the same grade.

The study recruited 619 middle-school students from public schools in Delhi, with approximately half randomly selected via lottery to receive vouchers providing free access to Mindspark centers for 4.5 months. Student learning was measured using independently-conducted paper-and-pencil tests in math and Hindi at baseline and endline, with tests designed using Item Response Theory (IRT) to be comparable on a common scale across grades and testing rounds. The authors estimate intent-to-treat (ITT) effects by comparing endline test scores between lottery winners and losers, controlling for baseline scores and stratification fixed effects, and also present instrumental variables estimates using the lottery as an instrument for actual program attendance to estimate dose-response relationships.

The dataset `ms_blel_jpal_long.dta` contains 1,158 observations in a long format with 36 variables. The unit of observation is student-round, where students can appear once or twice depending on their presence at endline testing. There are 619 unique students in the dataset: all 619 students completed baseline testing in September 2015, but only 539 students (87.08%) were present for endline testing in February 2016, resulting in 80 students appearing only once in the dataset and 539 students appearing twice (once at baseline and once at endline). The dataset includes student demographic characteristics (age, gender, grade level), treatment assignment status from the randomization (314 treatment, 305 control), test performance measures in math and Hindi (both raw scores and IRT-scaled scores), socioeconomic status indicators, and participation in private tutoring. This longitudinal structure allows researchers to analyze both cross-sectional differences between treatment and control groups at baseline and changes in student achievement over the 4.5-month intervention period, while also documenting the 13% attrition rate between baseline and endline that the authors address through robustness checks in the paper.