

README — Mindspark Replication Extension (Ordered Grade-Index Heterogeneity)

Author: Carl Romer

Date: 2025/11/30

1. Overview

This replication package contains my extension to the J-PAL replication materials for Muralidharan, Singh, and Ganimian (2019), “Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India.” The original paper evaluates the impact of lottery-based access to the Mindspark after-school learning software on student learning outcomes in Delhi. My extension examines whether the program’s effects systematically vary across grade levels by creating an ordered grade index and interacting it with treatment status.

2. Purpose of the Extension

Although the original paper documents substantial heterogeneity by baseline achievement, it does not examine whether impacts differ across students enrolled in different grades. Because the distribution of grades in the sample is heavily skewed—589 students in grades 6–9 and only 15 in grades 4–5—simple subgroup analysis is severely underpowered.

To address this, I construct an ordered grade index:

- Grade $< 7 \rightarrow -1$
- Grade 7 $\rightarrow 0$
- Grade 8 $\rightarrow 1$
- Grade 9 $\rightarrow 2$

I then interact this index with treatment assignment to test for a linear trend in treatment effects across grade levels.

3. Data Description

All data come from the J-PAL replication package. The primary file used is:

ms_blel_jpal_long.dta

Located in: "6. Replication package/data/"

Key variables:

- round: 1 = baseline, 2 = endline
- m_theta_mle: math IRT score
- h_theta_mle: Hindi IRT score
- treat: treatment assignment (voucher lottery)
- strata: randomization strata
- st_grade1: baseline enrolled grade
- in_r2: indicator for inclusion in the main analysis sample

4. Methodology and Estimating Equation

I create a student-level dataset by merging baseline (round 1) and endline (round 2) scores. I generate the ordered grade index and estimate the following model separately for math and Hindi:

$$Y_2 = \alpha + \gamma Y_1 + \beta T + \eta \cdot \text{GradeIndex} + \delta(T \times \text{GradeIndex}) + \text{StrataFE} + \varepsilon.$$

Interpretation:

- β : treatment effect for grade index = 0 (grade 7)
- $\beta - \delta$: effect for grade < 7
- $\beta + \delta$: effect for grade 8
- $\beta + 2\delta$: effect for grade 9
- δ : slope of treatment effects across the ordered grade index

Robust (HC1) standard errors are used.

5. Main Findings

The interaction between treatment and the grade index is small and statistically insignificant for both math and Hindi. This indicates that Mindspark's treatment effects do not meaningfully vary by enrolled grade. While higher-grade students score higher overall at endline, the estimated treatment effect is effectively flat across grade levels.

This result is consistent with the original paper's broader conclusion that Mindspark benefits students across the distribution rather than targeting specific grade groups.