Preanalysis plan

Helen Baker Henningham and Costas Meghir

December 2025

There are two rounds of data collection. We refer to the first one as cohort 1 and to the second one as cohort 2.

Cohort 1 includes children who at the start of the experimental study were 3-year-old. These will be referred to as cohort 1/3. It also includes children who at the start of the intervention in January 2024 were 4 years old. These will be referred to as Cohort 1/4.

Cohort 2 includes children who were 3 years old in January 2025. They are referred to as Cohort 2/3

| Timeline of Data Collection | | | |
|---|---|---|---|
| | January-February 2024 | October -December 2024 | September-December 2025 |
| Cohort 1/3 | Yes - Baseline | Yes - Midline | Yes - Endline |
| Cohort 1/4 | Yes - Baseline | Yes – Midline | Yes - Endline |
| Cohort 2/3 | No | Yes - Baseline | Yes - Endline |

**Process of Randomization.**

Out of 118 eligible schools, 60 were randomized to treatment and 58 to control.

**Cohort 1/3**

In a radius of 4 km from each school (treatment or control) we chose up to 15 children of the eligible age from families that have consented to participate in the study.

If the corresponding school was assigned to treatment, we selected randomly up to 5 children to attend the school and 5 kept back as controls

If the school was a control school, we selected randomly 5 children to be observed, but who would not receive any intervention activities.

**Cohort 1/4**

In principle all four-year-olds attend pre-school. Within each treatment and control school we randomly selected 5 four-year-olds to be assessed.

**Cohort 2/3**

For Cohort 2/3 we identified the five families with children who will be 3 years old in January 2025 who were closest to a study school (treatment or control) and who consented to participate in the study. Those in the catchment area of a treatment school were assigned to attend the catchment area school. Information about the treatment status of the school to be assigned was revealed only after consent was obtained.

**Analysis of Results**

For all cohorts, there are two primary outcomes: Child IQ using the Weschler Preschool and Primary Scale of Intelligence (W PPSI) full scale IQ scores and child school readiness using the International Development of Early Learning Assessment (IDELA) total score. These will be scored based on the algorithm provided by the test publishers. The scores will be the outcome we will analyze. In further analysis we will also use the Item Response Theory (IRT) method of scoring, which weighs questions according to the performance of this population.

Secondary outcomes:

- The WPPSI verbal comprehension index
- The five individual subscales of the IDELA (emergent numeracy, emergent literacy, social emotional. motor skills, and executive function)
- Child behavior during the test using the Preschool Self-Regulation Assessment (PSRA) rating scale.
- Quality of the classroom environment through independent observation using an adapted version of the MELE (Measuring Early Learning Environments)
- Teacher satisfaction by teacher report.

**Analysis for Cohort 1/3**

The analysis of impacts will take place using simple regression analysis. We will control for tester effects, age and sex of the child, baseline test scores for school readiness (measured using and adapted version of the Daberon Screening for School Readiness) and PSRA. Because of the subsequent analysis of heterogeneity, we will also present a version of results where we control for family characteristics, including mother's education, household wealth and presence of father.

We will use three comparisons:

**Comparison A:** Any child assigned to treatment will be compared to any child assigned to control.

**Comparison B**: Children assigned to a treatment school versus children not assigned to a treatment schools in the catchment area of a treatment school.

**Comparison C:** Children in the catchment area of a treatment school and assigned to attend it versus children living in the catchment area of control schools.

The above is an intention to treatment analysis. We will also present Local Average treatment effects (LATE) using instrumental variables. The regressor in this case will be the actual attendance, measured either as ever attended versus never attended or as a proportion of classes attended. The instrument will be the original randomized treatment assignment.

The same approach will be used for the midline and endline data.

**Inference**

Computation of standard errors and p-values will be performed using the bootstrap and respecting the clustering structure by school.

p-values for the primary outcomes will be corrected for multiple testing, where two hypotheses will be included (effects on IDELA total score and effects on WPPSI FSIQ). We will use the Romano-Wolf stepdown procedure for adjusting the p-values for multiple testing.

**Analysis for Cohort 1/4**

The test scores for both IDELA and WPPSI will be regressed on the treatment dummy, dummy variables for the tester, child age and sex, and family characteristics. We will not use baseline test scores because the assessments were made after the intervention started. Because of the subsequent analysis of heterogeneity we will also present a version of results where we control for family characteristics, including mother's education, household wealth and presence of father.

The comparison will be between 4-year-olds in treatment schools versus 4-year-olds in control schools.

Since there is no attendance/compliance issue for this group no LATE analysis will be used.

The same approach will be used for the midline and endline data.

**Analysis for Cohort 2/3**

For Cohort 2/3 we compare children assigned to attend a treatment school versus children assigned to attend a control school.

The analysis of impacts will take place using simple regression analysis. We will control for tester effects, age and sex of the child, baseline test scores for Daberon and PSRA. Because of the subsequent analysis of heterogeneity, we will also present a version of results where we control for family characteristics, including mother's education, household wealth and presence of father.

**p-values and inference:** In all cases p-values will be corrected for multiple testing, where two hypotheses will be included (effects on IDELA and effects on WPPSI). We will use the Romano-Wolf (RW) stepdown procedure for adjusting the p-values for multiple testing. The RW procedure will also be used when testing for heterogeneity of impacts

**Further analysis: Mechanisms and mediators and moderators**

**Heterogeneity analysis (moderators):** As exploratory analysis we will consider heterogeneity of effects based on gender and based on family background characteristics such as mothers' education and household wealth.

**Mediators:** Mediation analysis will be carried out by regressing the outcome on the treatment status and on potential mediators that themselves will have been affected by the intervention. In this way we will explore mechanisms explaining the impacts of the intervention and specifically, the role of classroom observations.