Details of Evaluation Approach

We evaluate the impact of our intervention on a subset of the population of particular interest to our implementing partner: high-vulnerability schools in the Chilean Metropolitan Region. The number of treated schools was set to five because of budgetary reasons. The program has been on-going since 2022 and solicits school participation through an open call disseminated through the foundation's social media and school visits. Conditional on being a high-vulnerability school in the Chilean Metropolitan Region, schools were identified to be in our study, satisfying the following additional criteria: (1) there are at least 2 classes of the third grade in the school, and (2) schools are locally clustered to optimize on the foundation's resources (e.g., the psychopedagogue's time). Five school pairs were identified such that schools in the same match-pair have approximately the same "schooling vulnerability index" (IVE), the same size (in terms of students enrolled), and located in the same (or nearby) communes. Our implementing partner randomly selected one member of each pair to participate in the intervention. Students and teachers from both treatment and control schools were surveyed before and after the intervention. We also obtain additional standardized test information for all students.

Given that our intervention sessions take place during normal school hours, all students enrolled in third grade of the treated schools participate in them. We can therefore simply compare the change in outcomes for students in treated schools and those in control schools, within the same matched-pair strata, without worrying about issues related to selection into treatment. We include individual- (e.g., gender, repeater status, initial risk of dyslexia) and school-level controls in a regression of the outcome on treatment status and matched-pair identifiers to improve on statistical efficiency. We may also include teacher-specific controls in a separate regression. We report standard errors clustered at both the matched-pair strata and school level, computed using a wild bootstrap procedure. Clustering at the strata-level is a more conservative approach to inference, but as we have a small number of strata, we may be underpowered. We also consider clustering at the school level as we believe it is plausible unobserved shocks to school are uncorrelated, even within strata.

As part of the data collection at baseline, we measured the student-level risk of dyslexia through an in-app screening tool. Moreover, we collected information on whether the child is a repeater or not. We will study heterogeneity in the effect of our intervention based on the initial risk of dyslexia, whether the child is a repeater, and sex of the child.

To explore outcomes of interest we will construct indices (families) following either Kling, Liebman and Katz (2007) or Anderson (2008). We detail the families in the Outcomes section. This approach is attractive because (i) it increases our power to detect significant effects and (ii) it helps in dealing with concerns of multiple hypothesis testing. We will also separately explore the various elements of the family individually. In that case we will correct for multiple hypothesis testing employing False Discovery Rate and Family Wise Error Rate corrections (e.g., Romano and Wolf, 2016).

References:

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. Journal of the American Statistical Association, 103(484):1481–1495.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. Econometrica, 75(1):83–119.

Romano, J. P. and Wolf, M. (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. Statistics & Probability Letters, 113:38–40.

Outcomes (Endpoints)
[Primary] Academic performance will be measured by the Language and Mathematics tests of the Diagnostico Integral de Aprendizajes (DIA), a standardized testing tool crafted by the Agencia de Calidad de la Educacion of the Ministry of Education of the Chilean government. We will use a continuous measure of both the overall score and the performance at the various competences within a subject.

[Primary] Non-cognitive skills and effort will be measured by: (1) academic aspirations, (2) self-perceived performance (absolute and relative to peers) in reading, mathematics and language and by the self-perceived potential to reach university, (3) taste for reading, mathematics, language and for attending school, (4) grit and locus-of-control, (5) time use (studying and doing chores), (6) individual wellbeing (feelings of happiness, of worry, of sadness, demotivation, and of easiness of getting angry) and (7) social wellbeing (feelings of isolation, feeling safe at school and being treated well by peers). These measures are collected through our specially-crafted questionnaire distributed at baseline and at endline. The Spanish version of the survey instruments used are reported in the "Data" section of this registry.

More specifically, we use the following variables to construct the families:

(1) academic aspirations: up to which academic level would you like to study? [single-element family]

(2) self-perceived performance: do you find math/language/reading hard?; perceived ability in math/language/reading relative to peers.

(3) taste for academic subjects and school: how much do you like math/language/reading?; do you like attending school?

(4) grit and locus-of-control: the first 4 subquestions in question 15 of the baseline questionnaire for grit. The remaining subquestions are for locus-of-control.

(5) time use: time devoted to studying and to doing chores in a normal week (weekdays and weekend).

(6) individual wellbeing:

(a) I feel happy; many things worry me; I feel sad; I get angry easily; oftentimes I do not feel like doing anything.
(b) oftentimes I feel I do things wrong.
(c) oftentimes I find problems focusing.

(7) social wellbeing: I feel lonely; my classmates treat me with respect; I feel safe at school.

All items above (when applicable) are elicited on a 5-point scale. We reverse the scale, when appropriate, to make the individual items within a family to indicate the same direction.

As an additional secondary outcome, to help us better understand possible mechanisms, we will look into measures of parental investment in the child. For this, we will exploit information on frequency in: helping the child with school work, worrying about his/her academic

performance, reading with him/her, playing, and eating with the child. The first two dimensions are available at both baseline and endline while the remaining items are only available at the endline.