Summer School: An Evaluation of a Government-Run Program to Remediate Foundational Learning

[Zero draft. Please do not cite or circulate]

Carolina Concha-Arriagada (Columbia University)
Guthrie Gray-Lobe (University of Chicago)
Michael Kremer (University of Chicago)

Abstract

Many education systems struggle to remediate foundational literacy and numeracy (FLN) deficits that emerge early in primary school. Grade retention is the most common policy response, but it is costly to governments, costly to students in foregone years of schooling, and may harm motivation. We study an alternative approach: a short, intensive remedial summer camp implemented at scale by the Dominican Republic's Ministry of Education. The program provided three to four weeks of full-day instruction focused on foundational literacy and numeracy to low-performing students in grades 3–5. We evaluate two pilot cohorts implemented in July 2024 and July 2025 using randomized invitations to attend camp. Despite substantial non-compliance, we find that fifteen days of attendance increased test scores by roughly 0.07–0.22 standard deviations, with effects generally larger for students who had not previously been retained. Back-of-the-envelope calculations suggest the program is highly cost-effective relative to business-as-usual instruction and may be particularly attractive where it substitutes for grade repetition.

Introduction

A large share of students in low- and middle-income countries reach the later primary grades without mastering foundational literacy and numeracy. Once these gaps emerge, they are difficult to close within the regular school day. The most common policy response is grade retention, yet retention is expensive, increases time spent in school per completed grade, and may negatively affect students' self-image and motivation. Moreover, many education systems restrict or prohibit retention in early grades, allowing students with weak foundational skills to advance without remediation.

A growing body of evidence shows that remedial education can substantially improve foundational skills, particularly when delivered through structured pedagogy with clear scope and sequence, teacher guides, and level-appropriate materials. Implementing remediation during the school year, however, raises practical challenges: it displaces regular instruction, requires tracking or differentiation, may increase stigma, and often demands smaller class sizes. These constraints motivate interest in remedial instruction delivered outside the regular school year.

Summer school programs are widely used in high-income countries but remain relatively uncommon in low- and middle-income settings. Summer programs face a fundamental design tension: learning gains require instructional time, but time outside the school year is costly for both students and teachers. This paper evaluates whether a short, intensive summer program can meaningfully improve foundational skills and whether such a program is better targeted at students who are already overage or those at risk of future retention.

Related Literature

A large literature—primarily from the United States—studies the impacts of grade retention using quasi-experimental variation induced by test-based promotion rules. A central challenge in interpreting this evidence is that grade retention is rarely implemented in isolation. Under many accountability regimes, students who fall below promotion cutoffs are simultaneously assigned to mandatory summer school and face a sharply increased probability of being retained if they fail to meet performance standards. As a result, widely cited estimates often identify the impact of a bundled policy rather than the causal effect of retention alone.

Seminal work by Jacob and Lefgren (2004, 2009) exploits regression discontinuities in Chicago Public Schools' promotion policies and documents short-run test score gains that fade over time, alongside longer-run increases in dropout risk for retained students. Importantly, the authors emphasize that falling below the cutoff induces both summer school participation and retention, and that their estimates therefore capture the net effect of exposure to this combined treatment. Subsequent work using similar designs reaches comparable conclusions. For example, Manacorda (2012) and Eren, Depew, and Barnes (2017) analyze test-based promotion regimes and show heterogeneous and often transitory achievement effects, with adverse longer-run outcomes for some students. In these settings, summer remediation is an integral part of the promotion system, making it difficult to interpret estimated effects as reflecting retention per se.

A smaller set of studies attempts to disentangle the effects of summer school and retention by exploiting institutional details or functional-form assumptions. For example, Matsudaira (2008) studies summer remediation programs and finds modest positive effects on achievement, while Schwerdt, West, and Winters (2017) and related work use structural or constant-effects assumptions to separate the contribution of remediation from repetition. Across these studies, estimated impacts of summer remediation typically range from near zero to roughly 0.1

standard deviations, though identification relies on strong assumptions about treatment homogeneity and counterfactual learning trajectories. Taken together, this literature suggests that remediation can improve short-run outcomes, but that isolating its effect from retention within promotion systems remains empirically challenging.

Nearly all evidence in this literature comes from high-income countries, where grade retention is relatively uncommon and progression rules differ substantially from those in low- and middle-income settings. In these contexts, retention is often used sparingly and late, whereas in many developing countries large shares of students become overage early in primary school, often due to automatic promotion rules that limit retention in early grades. This raises questions about the external validity of existing estimates for low- and middle-income countries.

A separate literature evaluates remedial education programs, often implemented through short-term learning camps, most prominently in the Teaching at the Right Level (TaRL) tradition (Banerjee et al. 2016). These programs group students by current learning level rather than grade and deliver structured, level-appropriate instruction in foundational literacy and numeracy. A number of evaluations—including large randomized trials in India and Africa—show that such interventions can generate substantial learning gains over short periods.

Several TaRL evaluations include camp-style models delivered outside regular instructional time, such as during school holidays or as short intensive bursts. However, these programs differ from ours along several important institutional dimensions. First, TaRL camps are typically implemented by NGOs using local volunteers or community instructors, rather than by public school teachers. Second, they are generally not government-run programs embedded within the public education system, but instead operate as parallel interventions. Third, while TaRL is motivated by the reality that many students are far below grade-level expectations—often in part because systems restrict grade retention in early grades—these programs are not explicitly designed as substitutes for grade retention, nor are they typically evaluated with retention avoidance as a primary outcome. In many cases, these programs are instituted as a response to students falling behind in grade in settings where grade retention is prohibited.

Our study bridges these two literatures by providing clean evidence on the impacts of government-run summer remediation delivered by public school teachers in a context where grade retention is a salient policy margin. Unlike the U.S. promotion literature, we evaluate summer remediation separately from retention, allowing us to isolate the causal effect of short-course remedial instruction without bundling it with repetition of the grade. Unlike the TaRL literature, we study a program implemented at scale by a ministry of education and explicitly motivated by concerns about retention and over-age enrollment. This combination allows us to assess whether a short, intensive summer program can serve as a cost-effective alternative to grade retention in settings where foundational learning gaps are widespread.

Program Description

Aprender es Divertido Summer Camps

The program we study, *Aprender es Divertido* ("Learning is Fun"), was developed and implemented by the Dominican Ministry of Education (MINERD). Two pilot cohorts were implemented in July 2024 and July 2025.

The camps lasted three to four weeks (15 instructional days in 2024 and 19 in 2025) and operated as full-day programs from 8:30 a.m. to 3:30 p.m. Camps targeted low-performing students in grades 3 to 5. Class sizes were approximately 25 students.

The curriculum focused on foundational literacy (phonics, fluency, and reading comprehension) and foundational numeracy (number sense, basic operations, and problem solving). Instruction combined structured lessons with games, small-group work, and individualized practice. Teachers were selected based on assessed performance using administrative evaluations and local recommendations.

Summer camps were organized around a hub-and-spoke model (See Figure **1**). Each camp was hosted at a designated "central" school within a cluster of nearby primary schools. Eligible students enrolled in the central school, as well as students attending surrounding schools in the cluster, were invited to attend the camp at the central location. As a result, the distance between students' homes (or enrolled schools) and the camp varied, and in some cases students were required to travel substantial distances to attend. In the first year, some students were enrolled in schools up 25 kilometers from the camp. In the second year of implementation, the average distance between students and the assigned camp declined substantially, with nearly all students enrolled in schools less than 5 kilometers from the camp.[1]

[1] While we are still awaiting administrative confirmation, discussions with program staff suggest that this reduction likely reflects a deliberate design change in which the Ministry reduced the geographic radius used to assign surrounding schools to each camp center, thereby locating camps closer to participating students on average.

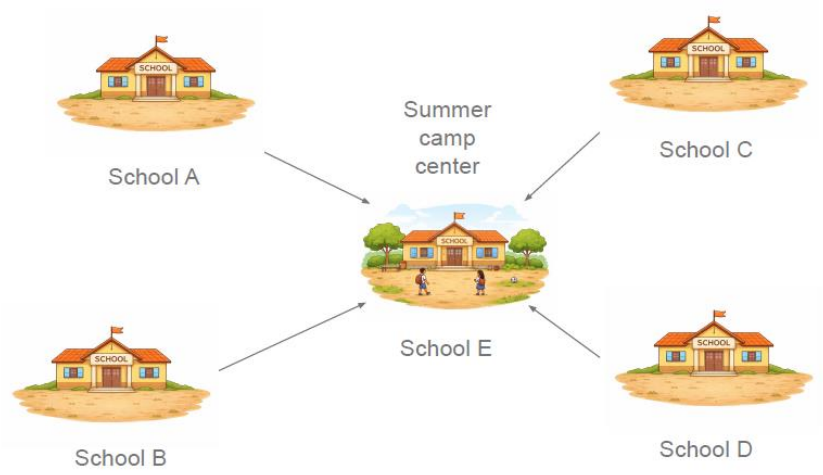


Figure 1 Hub-and-spoke summer camp model. Students were invited from schools surrounding the school where the summer camp was held.

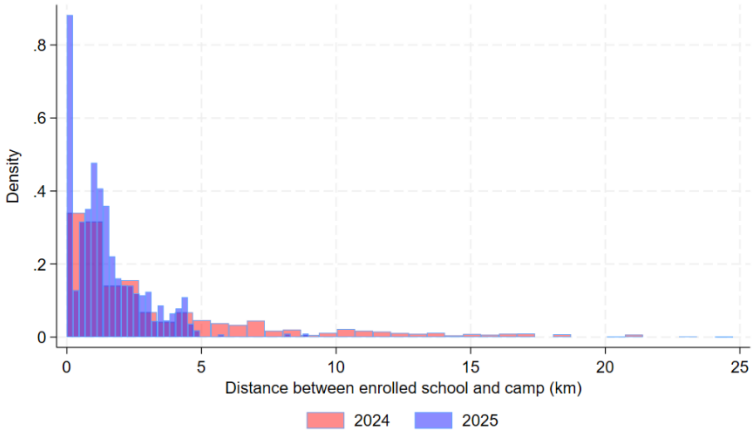


Figure 2 Histogram of the distance between students enrolled school and camp school (km) in 2024 and 2025.

Targeting and Scale

Targeting differed across cohorts. In 2024, teachers nominated students they viewed as struggling academically; approximately 70 percent of nominated students were overage for grade. In 2025, eligibility was based exclusively on administrative data, and all enrolled students were overage.

The program scaled from 30 summer camp centers in 2024 to 52 centers in 2025. Students attended camps located at designated schools, often different from their regular school. Distance to the assigned camp varied and played an important role in attendance decisions.

Research Design and Data

Randomized Invitations and Non-Compliance

In both years, eligible students were randomly assigned an invitation to attend the summer camp. Take-up was incomplete, generating two-sided non-compliance: some invited students attended few or no days, while a small number of non-invited students attended.

Attendance is measured as the fraction of scheduled days attended. In 2024, invited students attended on average fewer than half of program days, and only a small fraction attended all 15 days. Attendance improved substantially in 2025, when nearly half of invited students attended at least 15 days and 30 percent attended all 19 days. Reduced attendance among non-invited students contributed to a stronger first stage in 2025.

Outcomes

Primary outcomes are standardized endline test scores in Spanish language and mathematics, administered shortly after the program ended. Baseline assessments were conducted prior to the start of the camps. We also examine selected non-academic outcomes in 2025, including measures of emotional regulation, academic self-concept, growth mindset, empathy, and bullying victimization.

Empirical Strategy

We estimate the causal effect of camp attendance using an instrumental variables framework. Let $Z_i \in \{0,1\}$ denote random assignment to an invitation, D_i denote days attended normalized by the scheduled program length, and Y_i denote endline test scores. Our parameter of interest is the average causal response of attendance:

[Y_i = + D_i + _i.]

Because attendance is endogenous, we instrument D_i with the randomized invitation Z_i. Identification relies on random assignment of invitations and the assumption that the effect of the program scales linearly with attendance. All specifications control for baseline test scores and randomization strata defined by school and grade.

We also report intent-to-treat (ITT) estimates of the effect of being invited to attend camp and ordinary least squares (OLS) estimates that condition on baseline scores.

Results

First Stage

The invitation substantially increased attendance in both cohorts, though the first stage was considerably stronger in 2025. In 2024, the summer camp increased the number of days that students attended the summer camp by around 4 days (Figure 3). The weak first stage is partly due to non-compliance in the comparison group. Non-compliance in the comparison group was due to concerns about the program failing to meet enrollment targets. After observing low take-up in the invited group, the government invited applicants that had been assigned to the comparison group. Invited applicants were around 9 percentage points more likely to attend all 15 days of the camp, over a base of around 4 percent in the comparison group.

The first stage effect on camp attendance was substantially higher in 2025. Invited students attended more than 8 days more than coparison students. Comparison students attended less than half a day on average. The fraction of invited students who attended at least 15 days increased to nearly 45 percent, and 30 percent attended all 19 days.

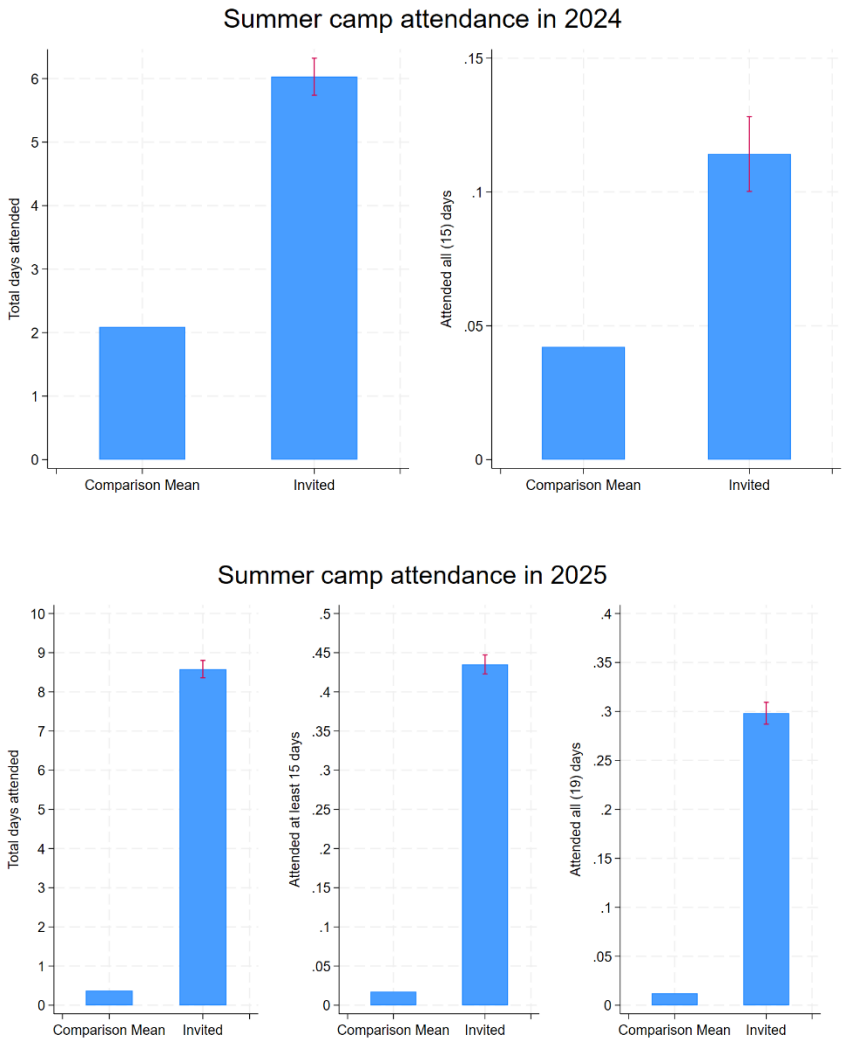


Figure 3 First stage effect of invitation on summer camp attendance in 2024 and 2025.

Impacts on Test Scores

In 2024, fifteen days of attendance increased Spanish language scores by approximately 0.22 standard deviations in 2SLS specifications, with smaller and statistically insignificant effects in mathematics (Table 1). Estimates are noisy due to a weak first stage and smaller sample size.

In 2025, effects are more precisely estimated. Fifteen days of attendance increased Spanish scores by about 0.07 standard deviations and mathematics scores by about 0.10 standard deviations. Although the ITT effects are slightly smaller than in 2024, this reflects a larger first stage and longer program duration. We cannot reject equality of the 2024 and 2025 effects.

Table 1 Effect of summer camp on test scores

	2024 Cohort			2025 Cohort		
	ITT	2SLS	OLS	ITT	2SLS	OLS
			Panel A: Spanish language			
Estimate	0.07	0.22	0.18	0.05	0.07	0.11
SE	(0.03)	(0.11)	(0.05)	(0.02)	(0.03)	(0.02)
Observations	2,445	2,445	1,569	7,941	7,941	6,547
			Panel B: Mathematics			
Estimate	0.00	0.00	0.12	0.06	0.10	0.11
SE	(0.03)	(0.10)	(0.06)	(0.02)	(0.03)	(0.02)
Observations	2,346	2,346	1,429	7,903	7,903	6,328

Heterogeneity by Overage Status

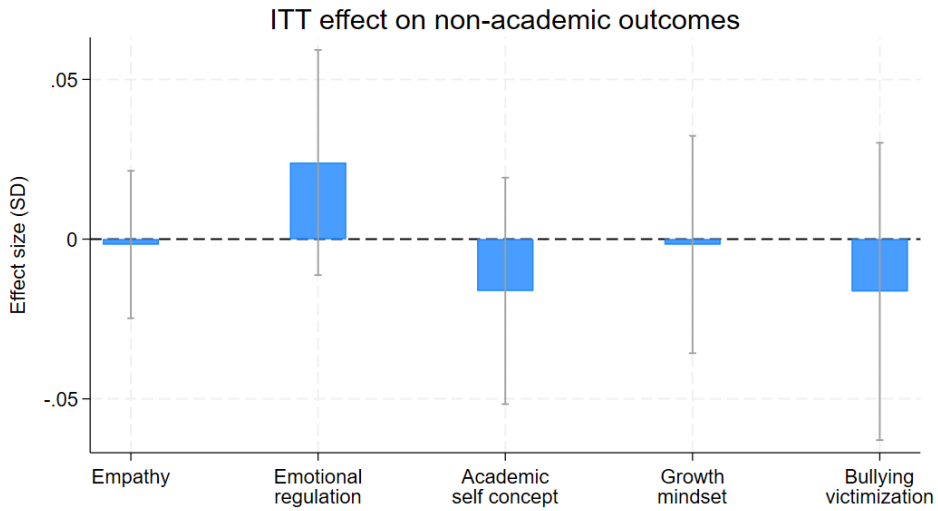
In 2024, effects are substantially larger for students who were age-appropriate for grade than for students who were already overage. For overage students, estimated impacts are small and statistically indistinguishable from zero. This pattern is consistent with the weaker results observed in 2025, when all participants were overage.

Table 2 Effect of summer camp on learning for overage students in 2024 cohort

	ITT	2SLS	OLS
		Panel A: Spanish language	
Estimate	0.04	0.14	0.14
SE	(0.04)	(0.14)	(0.06)
Observations	1,756	1,756	1,140
		Panel B: Mathematics	
Estimate	-0.02	-0.06	0.06
SE	(0.04)	(0.14)	(0.07)
Observations	1,694	1,694	1,042

Non-Academic Outcomes

We find limited evidence of impacts on non-academic outcomes in 2025. Point estimates suggest small improvements in emotional regulation and small declines in academic self-concept and bullying victimization, but confidence intervals include zero.



Cost-Effectiveness

A rough cost-effectiveness calculation suggests the program may be cost effective. The program generated approximately 0.1 standard deviations of learning gains in about fifteen days of instruction. Mapping test score gains to equivalent years of schooling implies roughly 0.30–0.35 standard deviations per 180-day school year.[2] Under conservative assumptions that per-pupil costs scale linearly with instructional days, the program delivers substantially more learning per dollar than business-as-usual instruction. Final analysis of cost effectiveness will rest on empirical evidence of the impact of grade retention on learning. However, existing evidence from other studies on the impact of grade retention (which typically combine the impact of grade retention with other remedial supports) find smaller impacts (E.g. Jacob & Lefgren 2004, Berne et al. 2025), and potentially higher rates of drop out over the long-run (Jacob & Lefgren 2004).

[2] The estimate of EYS is obtained from the comparison group in the sample, and therefore reflects the learning gradient for students who are overage.

Discussion

These results demonstrate that a short, government-run remedial summer program can meaningfully improve foundational skills, even with incomplete attendance. The findings also highlight the importance of targeting: students who have not yet been retained appear to benefit more than those who are already overage, suggesting diminishing returns to remediation once students have fallen far behind.

Attendance remains a key constraint. Distance to camp substantially reduces take-up, implying that future scale-up may require more decentralized centers or transportation support.

Conclusion

Short-course remedial summer camps offer a promising, scalable alternative to grade retention for addressing foundational learning gaps. Evidence from two pilot cohorts in the Dominican Republic shows robust test score gains over a short instructional period and suggests high cost-effectiveness. Future work should examine longer-run impacts on grade progression and retention and explore how best to target such programs.

References

Jacob, B. A., & Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. Review of Economics and Statistics.

Jacob, B. A., & Lefgren, L. (2009). The effect of grade retention on high school completion. American Economic Journal: Applied Economics.

Manacorda, M. (2012). The cost of grade retention. Review of Economics and Statistics.

Matsudaira, J. D. (2008). Mandatory summer school and student achievement. Journal of Econometrics.

Eren, O., Depew, B., & Barnes, S. (2017). Test-based promotion policies, retention, and student outcomes. Journal of Public Economics.

Banerjee et al. (2016). Mainstreaming an effective intervention: Evidence from randomized evaluations of “Teaching at the Right Level” in India. American Economic Journal: Applied Economics.

Schwerdt, G., West, M. R., & Winters, M. A. (2017). The effects of test-based retention on student outcomes over time: Regression discontinuity evidence from Florida. Journal of Public Economics, 152, 154–169.

This paper is a preliminary working draft. Comments welcome.

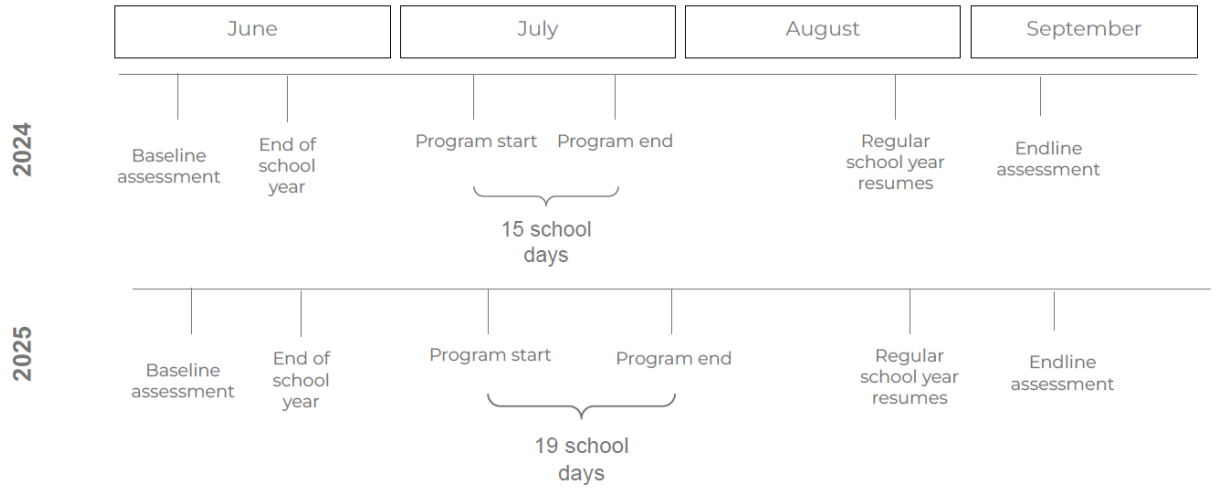


Figure 4 Study timeline