COVID-19 Learning loss and recovery: Panel data evidence from India*

Abhijeet Singh† Mauricio Romero‡ Karthik Muralidharan§

March 4, 2024

Abstract

We use a panel survey of ∼19,000 primary-school-aged children in rural Tamil Nadu to study 'learning loss' after COVID-19-induced school closures, and the pace of recovery after schools reopened. Students tested in December 2021 (18 months after school closures) displayed learning deficits of $\sim 0.73\sigma$ in math and 0.34σ in language compared to identically-aged students in the same villages in 2019. Two-thirds of this deficit was made up within 6 months after schools reopened. Further, while learning loss was regressive, recovery was progressive. A government-run after-school remediation program contributed ∼24% of the cohort-level recovery, likely aiding the progressive recovery.

Keywords: COVID-19, school closures, learning loss, remediation, increasing instructional time

JEL Codes: H52, I21, I25, O15

---

*We are grateful for comments from Alison Andrew, Noam Angrist, Jishnu Das, Jean Dreze, Peter Hull, Ajinkya Keskar, Ben Piper and Justin Sandefur. Ramamurthy Sripada, Dwithiya Raghavan, Akanksha Soni and Indira Tirumala provided excellent field management and research assistance. We received IRB approval from the Institute for Financial Management and Research (#113401). Financial support from the RISE Programme, — funded by the UK's Foreign, Commonwealth and Development Office, Australia's Department of Foreign Affairs and Trade, and the Bill and Melinda Gates Foundation — is gratefully acknowledged. Romero gratefully acknowledges financial support from the Asociación Mexicana de Cultura, A.C. All errors are our own. Replication data and code is available at Singh et al. (2024).

†Abhijeet Singh is a professor at the Stockholm School of Economics. He is also affiliated with the Jameel Poverty Action Lab (J-PAL), CESifo Economics of Education Research Network in Munich, the Center for Global Development, the Mistra Centre for Sustainable Markets (MISUM) at SSE, and the Young Lives study.

‡Mauricio Romero is a professor of at Instituto Tecnológico Autónomo de México (ITAM). He is also affiliated to the Jameel Poverty Action Lab (J-PAL), Innovations for Poverty Action (IPA), Experiments in Governance and Politics (EGAP), the Center for Global Development (CGD), The Bureau for Research and Economic Analysis of Development (BREAD), and The Institute for Labor Economics (IZA). E-mail: mtromero@itam.mx

§Karthik Muralidharan is the Tata Chancellor's Professor of economics at the University of California, San Diego. He is also a Research Associate of the National Bureau of Economic Research (NBER), a Fellow and Board Member of the Bureau for Research and Economic Analysis of Development (BREAD), a Board member and co-chair of the Education program at the Jameel Poverty Action Lab (J-PAL), an Affiliate at the Center for Effective Global Action (CEGA), and a Research Affiliate with Innovations for Poverty Action (IPA).

doi:10.3368/jhr.0723-13025R2

I Introduction

COVID-19 disrupted education systems worldwide. This shock was more severe in low- and middle-income countries (LMICs), which had longer school closures than OECD countries, and where schools and parents were less equipped to pivot to remote instruction (Agarwal, 2022; UNESCO, 2022). Poor households were particularly limited in their ability to compensate for school closures and more vulnerable to severe economic and health shocks (Patrinos et al., 2022). Thus, the COVID-19 crisis may have substantially exacerbated the 'learning crisis' in LMICs and increased educational inequality (World Bank, 2020).

India offers a leading example of such concerns. Compared to other establishments, schools were first to close and last to open, resulting in about 18 months of school closures (Andrew & Salisbury, 2023). Households faced significant economic hardship due to stringent lockdowns (Kesar et al., 2021). Health shocks were also severe: independent estimates indicate excess mortality of 3.2 million people between March 2020 and September 2021 (Jha et al., 2022). These shocks occurred in a context where, even before the pandemic, 50% of rural children in Grade 5 could not read a Grade 2 level text (Pratham, 2019). Evidence on past natural disasters and epidemics suggests that their negative effects on student learning, and potentially outcomes later in life, could be long-lasting (Andrabi et al., 2021; Bandiera et al., 2020). Yet, despite the scale of the disruption, and the large potential costs, relatively little is known about the magnitude and persistence of COVID-19-induced learning losses.

This paper presents new evidence on these questions using a large panel dataset from a sample of rural children in a large Indian state (Tamil Nadu). We use a household-based *census* of 25,126 children across 220 villages (conducted in 2019), which includes cognitive tests for all children aged 2-7 years, as a baseline. In 2021-22, we retested 19,289 of them using comparable assessments. These tests were administered over three survey waves between December 2021 (soon after schools reopened) and May 2022. Each student was revisited once in 2021-22 and the timing of these revisits was randomized within village. Thus, we observe population-level test score distributions four times (2019, December 2021, February 2022, and April-May 2022), and observe individual students twice (in 2019, and once in 2021-22).

We use these data to conduct three exercises. First, we quantify the magnitude of learning loss in December 2021, using comparable assessments linked via Item Response Theory (IRT) models, for students in early grades of primary schooling (a crucial stage for achieving foundational skills). We find large learning losses in December 2021, after 18 months of school closures. On average, students between 5–7 years were 0.73 and 0.34 standard deviations (σ) behind in mathematics and language, respectively, compared to students of the same age in the same villages in 2019. This is equivalent to 1–2 years of schooling in this context. These learning losses were regressive, consistent with a significant positive socioeconomic gradient in educational inputs received by children during school closures. The magnitude of this heterogeneity is, however, small relative to the size of the learning loss in the overall population.

Second, we estimate the pace of recovery and find a rapid catch-up in learning. Two-thirds of the learning loss documented in December 2021 was made up for by May 2022 (after 5–6 months of schools reopening). This recovery was modestly larger for children from more disadvantaged backgrounds, compensating fully for the socioeconomically unequal learning loss found after 18 months of school closures.

Finally, we evaluate the effectiveness of the state government's flagship COVID-recovery intervention in education. To address learning loss when schools reopened, the Government of Tamil Nadu introduced an after-school remedial program run by community volunteers for 60-90 minutes daily. This program, called *Illam Thedi Kalvi* ("Education at Doorstep", or ITK), was rolled out state-wide in January 2022 and employed approximately 200,000 volunteers. These volunteers were typically not trained or credentialed teachers, but had at least a high-school degree. It was the largest supplementary instruction program for COVID learning loss recovery in India (providing supplementary instruction to 3.3 million students) and among the largest COVID education response initiatives globally. This model of after-school remedial camps, led by locally-hired community volunteers, with content de-linked from school curricula, is similar to interventions studied in non-pandemic settings by Banerjee et al. (2017) and Duflo et al. (2020).

The program was salient: ~57% of households reported sending their children to these sessions; and of those doing so, ∼90% reported sending their children for 4 days or more per week. Within villages, children from less-advantaged households were more likely to attend ITK centers than students from better-off households. In particular, 74% of children enrolled in public schools attended ITK centers, compared to 19% of children in private schools. This contrasts with other mechanisms to mitigate learning loss *during* school closures, such as technology-based remote instruction or private tutoring, which display a positive socioeconomic gradient.

We estimate the effects of attending the ITK program using value-added models that incorporate rich measures of pre-pandemic achievement and household characteristics.[i] Attending ITK classes increased student test scores by 0.17σ and 0.09σ in mathematics and Tamil language over 3-4 months, with the treatment effects concentrated in students enrolled in public schools (who account for 90% of enrolled students in ITK centers). These results are robust to including extensive vectors of educational resources available to the child, compensatory inputs provided by schools and parents during school closures, or measures of child activities during school closures. Bounds computed as in Oster (2019) indicate robustness to relatively extreme selection from omitted variables.[ii]

These gains from a statewide program are noteworthy given the well-documented tendency for treatment effects to be smaller for programs implemented by governments at larger scales (Vivalt, 2020; Bold et al., 2018). Adjusting for the 57.3% attendance rate, the ITK program accounts for 28% of the population-level catch-up in Tamil and 20.7% of the catch-up in math. Thus, about half of the initial learning losses documented in December 2021 would have been recovered after 6 months of school reopening even without the ITK program, and we estimate that the ITK program increased this to two-thirds.

This paper contributes evidence to three key areas of the discourse on the impact of the COVID-19 pandemic on learning outcomes, which were forecast to cost up to $17 trillion in lost lifetime earnings (World Bank, UNESCO and UNICEF, 2021). First, despite substantial policy interest, the evidence to date on the extent of *actual* COVID-19 learning losses in LMICs remains limited (see reviews by Patrinos et al. (2022), Moscoviz & Evans (2022), and Betthäuser et al. (2023)). For example, Betthäuser et al. (2023) conclude that "there is a strong over-representation of [learning loss] studies from high-income countries, a dearth of studies from middle-income countries and no studies from low-income countries”. In particular, given the difficulties of in-person testing during the pandemic, most estimates of the impact on learning levels have relied on simulations or phone-based testing in non-representative samples.[iii] We join the few papers documenting learning losses using in-person testing in broad samples (e.g., Hevia et al. (2022); Alasino et al. (2024) who study Guanajuato and the Yucatan peninsula in Mexico, Guariso & Björkman Nyqvist (2023) who study public schools in Assam, India, Ardington et al. (2021) who study under-resourced schools in South Africa, and Lichand et al. (2022) who study secondary school students in São Paulo State, Brazil). Unlike previous studies, given the structure of our data and the design of the tests, we can show learning losses for a wide range of ages, as well as the distributional effects of the pandemic on learning losses.

Second, we contribute to the evidence of recovery (or lack thereof) after the pandemic. Few studies have documented the recovery in middle-income countries — notable exceptions are Lichand & Doria (2022) who document that secondary students recovered 25% of the losses after a year of returning to in-person classes in São Paulo State in Brazil and Alasino et al. (2024) who document that students in fifth grade in Guanajuato, Mexico recovered $\sim$60% of the losses 21 months after schools re-opened — and there is no evidence of the pace of recovery from a low-income country. We provide some of the first measurements of recovery in representative samples and with IRT-linked measurement of primary-school learning outcomes in a low-income setting. Given the potential long-term costs of even limited durations of school closure (Andrabi et al., 2021), this is a major gap in our understanding of learning trajectories in the aftermath of the pandemic. One reason for our contrasting results with those reported in Andrabi et al. (2021), may be that policymakers in Tamil Nadu recognized the extent of disruption to schooling, and prioritized remediation in both regular school instruction and in creating the ITK program.[iv]

Third, while there is evidence on the impacts of specific remote tutoring and technology interventions on mitigating learning losses *during* school closures (Carlana & La Ferrara, 2021; Hassan et al., 2021; N. Angrist et al., 2022, 2023; Crawfurd et al., 2023), there is much less evidence on the effectiveness of attempts to remediate learning losses upon school opening at scale. Since schools have now reopened, interventions based on in-person instruction may be more relevant for remedying learning loss at scale. Our results suggest that the ITK program, designed and implemented by a government in a short time and implemented state-wide, may provide a useful template for LMIC governments seeking to mitigate COVID-19 learning loss.

Beyond the literature on the impact of COVID-19 on education, we also contribute to the literature on after-school programs in LMICs, where programs run by non-profit organizations have been shown to be effective at improving learning outcomes (e.g., Banerjee et al. (2017), Muralidharan et al. (2019)). Our results suggest the ITK program may provide a template that can be followed by *governments* to run a highly-effective and cost-effective after-school program at scale. Such programs can contribute to ensuring universal foundational numeracy and literacy, and reducing socioeconomic gaps in learning even in non-pandemic recovery settings.[v]

II Data

II.A Sampling

Our study is based in 220 villages in 4 districts of Tamil Nadu (see map in Figure A.1 in the Online Appendix). These districts were chosen based on probability proportional to size sampling and are representative of rural Tamil Nadu. In these villages, we conducted a census of households and tested *all* students between the ages of 24-95 months in August 2019 (i.e., between 2–7 completed years).[vi] Although the villages sampled within the district were not randomly selected (the study universe is restricted to blocks with at least two government preschool centers (*anganwadis*) co-located with middle schools), our baseline sample is similar on observable characteristics to the rural population of the state, albeit slightly poorer in terms of asset ownership (see Table A.1 in the Online Appendix).

We revisited these communities and households between December 2021 and May 2022, administering a comparable test of student achievement to all children between ages 36–131 months (i.e., between 3–10 completed years) and collecting detailed information about household experiences and educational inputs during the COVID-19 pandemic. Of 25,126 children (18,457 households) with completed baseline tests in 2019, we were able to retest over 77% of the original sample (19,467 children, 14,648 households). This attrition does not vary by gender or SES, but does vary by maternal education and age (see Table A.2 in the Online Appendix). The principal reason for not being able to interview the rest was that the household had either moved or could not be found in the revisit.[vii] We restrict our sample to the 19,299 students aged between 48–131 months (i.e., between 4–10 completed years) at the time of the 2021-22 survey rounds for whom we also have baseline scores. This window covers the period leading up to school entry — which is mandated from 6 years of age — until the end of primary schooling in Grade 5.[viii]

II.B Waves of measurement

Our surveys in 2021-22 were designed to (a) measure 'learning loss', which we define as the deficit between what students know and what they might have been expected to know in the absence of the pandemic and (b) the pace at which they recover (or not) to pre-pandemic learning trajectories after schools reopened.

We randomized the initial sample within each village into an "early" and "late" follow-up group. The two groups are balanced on observables, as expected (see Table 1, Columns 1–3). The fieldwork for the "early" follow-up group was divided into two phases: 5,554 children were tested between December 20 and January 7 (Wave 1), following which fieldwork was paused due to the spread of the Omicron variant. Fieldwork was resumed after two months and 3,993 children were tested between February 25 and March 23 (Wave 2). Fieldwork for the "late" follow-up group was started immediately after completing Wave 2 in each district. 9,752 students were tested in a single contiguous round from March 11 to May 7 (Wave 3) — see Figure 1 for a timeline of the fieldwork alongside key dates of school closures and reopening. Although splitting the "early" follow-up group into two phases was not by design, respondents are balanced on observable characteristics across these three survey waves (see Table 1, Columns 4–7). Therefore, our analyses treats the waves as exogenously assigned and focus primarily on comparing Wave 1 (Dec 2021) to Wave 3 (April 2022).

II.C Learning Assessments

This paper focuses on student learning, which we assess through tests of cognitive skills that we designed tailored for the study objectives. Surveyors administered these to children individually and in person at the time of household visits.

In 2019, reflecting our principal focus on students of preschool and school-entry age, we administered assessments of basic numeracy and language skills to all children between 2–7 completed years of age. These were based on assessments used in a complementary project in the same state by Ganimian et al. (in press). All students were tested using the same survey tool. In 2021-22, we redesigned our assessments to accommodate the full range of student achievement by developing age-specific test booklets with an overlap of items between successive ages. At younger ages, our assessment items are mostly taken from the baseline test; at older ages ($\geq$5 years), we introduce additional items in math and Tamil to ensure better coverage of school-level competencies (and to address issues of ceiling and floor effects). Identical tests were used across the three survey waves in 2021-22.

The common items across rounds and ages, allow us to link achievement on a common metric using Item Response Theory (IRT) models (Das & Zajonc, 2010). We estimate these pooling all test observations across rounds, separately for math and language. We standardize test scores to have mean zero and standard deviation of one in the sample of children aged 60–72 months at baseline. See Online Appendix A.2 for details on test construction, psychometric properties of individual test questions, and distribution of student scores (to examine floor and ceiling effects).

II.D Household characteristics and educational inputs

In both years, we collected extensive data from households about their socioeconomic status and children's education. From 2019, we mainly use household socioeconomic status, measured using information about household ownership of various assets, and maternal education. In 2022, we also collected information about the educational inputs students received during school closures (e.g., video lectures, audio lectures, homework assignments, parental support for instruction, private tutoring, and the use of other online resources).

In Wave 3, surveyed in April-May 2022, we collected extensive information about the *Illam Thedi Kalvi* (ITK) program. This includes parental reports of awareness about the program and availability in the village, whether children from the household attend the ITK centers (and how frequently), when children started attending the ITK center, and what parents believe the ITK volunteers do in the remedial sessions.

III Measuring learning loss and post-pandemic recovery

III.A Learning loss in December 2021

Figure 2a presents non-parametric learning profiles of test scores with respect to age (at the time of testing) separately for the July-August 2019 and December 2021 rounds. Test scores increase monotonically with age in both rounds, but the gradient is markedly less steep in 2021.

With test scores on the same IRT-equated scale across ages and rounds, we can compute two measures of learning loss in each subject. The vertical distance between the 2019 and 2021 learning profiles provides an absolute measure of learning loss, expressed in standard deviations, at every age. The horizontal distance between the two learning profiles provides an alternative measure, namely how much older a student in 2021 was relative to a student who achieved the same score in 2019 (i.e. a *development lag*).[ix]

Both measures indicate learning losses of substantial magnitude, which we present at key ages in Table 2, Panel A. In mathematics, we estimate an absolute learning loss of $\sim 0.43\sigma$ at 60 months, equaling a development lag of about 10 months; by 84 months, this loss expands to ~ 0.74 SD, a development lag of 14.5 months. In Tamil, absolute learning losses are smaller in the standard deviations metric ($\sim 0.15\sigma$ at 60 months and $\sim 0.4\sigma$ at 84 months), but similar in terms of developmental lag for 5–8-year-olds.

Larger learning losses for older students may result from several sources. In this setting, it likely reflects that older students lost more months of in-person formal schooling compared to younger students (who would have been out-of-school or in preschool centers even in the absence of pandemic-induced closures).[x] Public preschools (ICDS centers) provide very little educational instruction and function largely as daycare and supplementary nutrition centers in this setting (Ganimian et al., in press). Thus, even though preschools were also affected by pandemic-induced closures, the direct effects on lost instructional time are likely insignificant. Further, parents may be able to compensate more at home for inputs provided to preschool children compared to children at later stages of schooling.

Table 2, Panel B further investigates absolute learning loss using the following specification:

$$Y_{it} = \alpha_v + \beta_1 Dec2021_t + \beta_2 \mathbf{X}_{it} + \varepsilon_{it} \tag{1}$$

where α_v is a vector of village-specific intercepts, $Dec2021$ is an indicator variable for being in the December 2021 survey round (with the 2019 round as the base category), and X is a vector of characteristics that includes the age of the child at the time of the test, their gender, maternal education (in categories) and their socioeconomic status (measured in percentiles of the 2019 distribution).[xi] We then examine how learning loss differs by observed student/household characteristics using linear interactions. Standard errors in all regressions are clustered at the village level. The sample is restricted to students between 55–95 months of age at the time of the test to ensure common support across the two years in the age of children.

Children score 0.73σ lower in math, and 0.35σ lower in Tamil language in December 2021 compared to similarly-aged children in the same villages in August 2019 (Columns 1 and 5).

Learning loss appears to have been severe for students of all backgrounds, and we do not find heterogeneity by gender. We find greater learning losses among children whose mothers had not completed high school (12th grade). Mothers' education is both a direct input into child learning, and a key determinant of the intergenerational transmission of human capital. It is also a marker of socio-economic status that correlates with other education inputs. Indeed, mothers' education is significantly correlated with student access to many educational inputs during school closures (see Table A.7 in the Online Appendix), with most of these inputs being significantly predictive of learning changes during the 18 months of school closures. While we do not find significant differences in learning loss by SES, as measured by ownership of consumer durables, the point estimates suggest greater learning loss among lower SES children. We find similar differences in access to most education inputs when children are ordered in terciles of socio-economic status, as measured by consumer durables (Table A.8 in the Online Appendix). Together, these results support the widely-held conjecture that learning losses during the pandemic would be regressive.

III.B Partial recovery after December 2021

The severity of estimated learning losses corroborates concerns about the worsening of the learning crisis as a result of the COVID-19 pandemic (Pratham, 2021a,b). Yet, an unanswered question is whether, after schools reopened, students "caught up" and recovered to pre-pandemic learning trajectories or whether the initial learning losses persisted or even expanded due to the potential worsening of the mismatch between student preparation and overambitious curricula (Banerjee et al., 2017; Pritchett & Beatty, 2015; Muralidharan et al., 2019; Bau, 2022).

Figure 2b generates learning profiles, as previously, for all four survey waves over the full age range tested in 2021. There are three main results, also shown numerically at key ages in Table 3, Panel A. First, the absolute learning loss documented in December 2021 is substantially reduced in the February 2022 wave and further still in the April 2022 wave. By this point, about two-thirds of the learning loss appears to be compensated in math and Tamil. Second, the shift across the three survey waves in 2021/22 is a shift in intercepts rather than of gradient — i.e., recovery was largely uniform regardless of age. Third, this shift in learning profiles in the post-pandemic period happens over the entire span of primary school ages.

We investigate recovery in greater detail in Table 3. Students score 0.24σ higher in February 2022 and 0.47σ higher in April 2022 in mathematics (Columns 1), and $0.12/0.19\sigma$ higher in Tamil in February/April (Columns 5), than those tested in December 2021 (the omitted category). This recovery by April-May 2022 compensates for $\sim$67% of the estimated learning loss of 0.73σ in December 2021 in mathematics and $\sim$56% of the initial loss of 0.34σ in Tamil.[xii] All regressions include background covariates for precision; however, since these are balanced between survey waves, the results are similar to those obtained from only controlling for age. Investigating heterogeneity by background covariates, recovery was *faster* for children with less-educated mothers and from poorer households (Columns 2-4 and Columns 6-8). We find no consistent evidence of heterogeneity by gender.

This pattern of rapid recovery — which is potentially surprising in light of evidence of persistence of losses from school closures in localized disasters (Andrabi et al., 2021) — is unlikely to be an artifact of test content, administration, or aggregation. Unlike typical tests in school settings, our tests are administered at home in a one-to-one setting by enumerators. They also include a substantial fraction of questions that rely on visual stimuli without requiring the ability to write. Thus, they are unlikely to reflect test-taking practice or familiarity. Studying learning loss and recovery in different competencies (Table A.6 in the Online Appendix) reveals both the losses and the recovery were generalized across age groups and domains. However, the losses were less pronounced (and the recovery was faster) for more basic skills and for younger children.

Recovery is also unlikely to be explained by issues of aggregating test scores — it is evident not only using IRT scaled scores but also on individual test items and in the percentage correct on common items (see Figure A.2 and Table A.6 in the Online Appendix). More generally, the tests have good psychometric properties: they have good internal consistency (as measured by Cronbach (1951)'s alpha), the questions are able to discriminate students with different abilities, the IRT model has a good empirical fit, and ceiling and floor effects are not an issue in the 2021-22 rounds (see Online Appendix A.2).

IV Evaluating the ITK policy to remedy learning losses

The rapid recovery we document likely reflects both "natural" catch-up after schools reopened and the effect of interventions designed to combat learning loss. In particular, the Government of Tamil Nadu implemented an ambitious statewide remediation program to help mitigate learning losses due to COVID-induced school closures called *Illam Thedi Kalvi* ("Education at Doorstep", or ITK). In this section, we first estimate the effects of attending ITK centers and then use them, together with program participation rates, to estimate the portion of the catch-up that may have occurred even in the absence of the ITK program.

IV.A The *Illam Thedi Kalvi* Program

The Government of Tamil Nadu introduced the ITK program as a pilot in selected geographies in November 2021 and then universalized it state-wide in January 2022. The program uses community volunteers to provide remedial instruction for 60–90 minutes in the evening. Instruction is delivered in small groups of 15–20 students and organized in school premises, preschool centers, or volunteers' homes. Volunteers were required to be local residents who had at least completed high school (Grade 12) to teach primary school children, and a Bachelor's degree to teach middle school children. They are paid a stipend of INR 1,000 ($\sim$ 12 USD) per month for incidental expenses — compared to an average primary teacher salary of INR 28,660 in 2014 (Ramachandran et al., 2015). In practice, nearly all the volunteers were women (who were given explicit preference in recruitment). See Online Appendix A.3 for further details on the design and implementation of the program.

Although initially conceived to last until June 2022, the ITK program has been extended to March 2023. It was estimated to have covered 3.3 million children, and employed over 200,000 volunteers by June 2022.

IV.B Take-up and selection into the program

The program was very salient: 91.3% of respondents reported having heard of it, and 57% of parents reported that their children attend the sessions. Approximately 87% of the households reported the program as having started in January or February 2022, with about 10.5% reporting the program having started in December. 92% of the children who attended the center were reported to attend for at least 4 days per week.

Children attending ITK centers differ from those who do not on observed characteristics (Table 4). They are slightly more likely to be female and older by 7–8 months on average (higher participation among older children could reflect the need to travel to the ITK centers after school hours). Importantly, they are from less-advantaged backgrounds: their mothers are 13 percentage points less likely to have completed 12 or more years of education, and their households were significantly poorer. Adjusting for age differences, ITK participants scored significantly lower in math and Tamil in 2019.[xiii] Overall, ITK participation was highly progressive.

The progressive participation in ITK largely tracks differences in ITK enrollment rates across public and private schools: In 2021-22, ITK attendees were much less likely to be enrolled in private schools (by 35 percentage points) than students who did not attend ITK centers.[xiv] This greater propensity of public school students to enroll in ITK classes is likely to be explained by the active role of public functionaries (including teachers) in promoting the program, as well as the higher take-up of public services by poorer households. While take-up is progressive within students enrolled in private schools — ITK participants have less educated mothers than non-participants — students in private schools are from substantially better-off households than students in public schools. However, ITK participants and non-participants within public/private schools are more similar on fixed background characteristics and lagged achievement than in the population overall.

The principal challenge for evaluating the causal effects of attending ITK centers is addressing these non-random patterns of enrollment, which we turn to next.

IV.C Evaluating the causal effect of attending ITK

We estimate the effect of ITK using value-added models that control for lagged achievement and child/household characteristics. Specifically, we estimate the following regression(s):

$$Y_{it} = \alpha_v + \beta.AttendITK_{it} + \gamma.\mathbf{X_i} + \sum_a \phi_a.I_a.\mathbf{Y_{i,t-1}} + \varepsilon_{it} \tag{2}$$

Here, Y_{it} is achievement in 2022; α_v is a vector of village-level dummy variables; $AttendITK_{it}$ is an indicator variable for whether child i attends an ITK center; $\mathbf{X_i}$ is a vector of child and household background characteristics including SES, maternal education, age at the time of the test, and enrollment in government or private school; $\mathbf{Y_{i,t-1}}$ is a vector of lagged achievement measures in math and Tamil in 2019, which we interact with age dummies (I_a) to allow the effect of lagged scores to vary flexibly with age as in Chetty et al. (2014); and ε_{it} is an error term.[xv] We enter the control variables sequentially to assess the direction of likely bias.

Specification (2) is a dynamic OLS lagged value-added model (VAM) which relies on an assumption of conditional exogeneity for identification of the causal effect of attending ITK centers (see e.g. Todd & Wolpin (2003, 2007)). Whether this assumption is satisfied in practice depends on the nature of selection in the specific context and the extent to which lagged achievement measures baseline ability accurately. In our setting, the major source of non-random selection is whether children were enrolled in a private or a government school. We control for this pre-program choice. Further, value-added models have been shown to be reliable in addressing selection biases correlated with non-random school choice in South Asia (Andrabi et al., 2011, 2022; Singh, 2015), as well as more generally.[xvi] Thus, even though we do not have exogenous sources of variation for ITK participation, our prior is that these estimates are likely to approximate the causal effect of interest. In the population overall, given the substantial *negative* selection into attending ITK centers, we expect any residual confounding factor to bias our estimates downwards and to be conservative approximations of the true causal effect of attending ITK centers.

The value-added model estimates may still be biased if inputs that are not proxied by lagged achievement and included controls are correlated with program enrollment. For example if, even conditional on covariates, ITK participation was correlated with compensatory inputs provided by schools or households, effort and time invested by students, or other inputs determining achievement, our estimates would be biased. We will investigate the potential magnitude of these biases using two complementary strategies.

Our first strategy resembles validation exercises in Chetty et al. (2014). Specifically, we collected detailed information on inputs provided to children during school closures. We include extensive data on compensatory inputs and other resources, which are not included in our benchmark specifications, in our value-added models to examine the sensitivity of our program estimates to their exclusion. To the extent that confounding household- or child-specific factors that raise the probability of enrollment in ITK centers also led to increased compensatory investments in the period of school closures, we would expect the inclusion of these investments to reduce bias.

Our second strategy takes this intuition of coefficient stability further and computes Oster (2019) bounds for treatment effects assuming the selection on unobservables is in the same direction as the selection on observables. We do this separately in the subsamples of students in government and private schools (as selection is much less stark within each sector and we control for attending private or public school when estimating the effect of ITK, see Table 4).

Our benchmark results are presented in Table 5. Column 1 presents a "naive" regression controlling for village fixed effects, gender and age, and shows that attending an ITK center is associated with an increase of 0.083σ in math and 0.073σ in Tamil. Column 2 presents a conventional value-added specification which includes the lagged test score and basic background characteristics (maternal education, SES and whether the child was enrolled in a government/private school). We estimate the effect of attending an ITK center to be 0.17σ in Math and 0.093σ in Tamil. The increase relative to the naive estimates in Column 1 is consistent with the negative selection into ITK observed in Table 4. This is our preferred estimate and is similar to lagged score value-added models in Chetty et al. (2014) and J. D. Angrist et al. (2017); J. Angrist et al. (2021). These results are robust to controlling for age fixed-effects, instead of age linearly (See Table A.10 in the Online Appendix). These results are also similar if we use a panel data approach, like the one used to estimate the pace of recovery in Table 3, and estimate the pace of recovery by whether students are attending ITK or not in wave 3 (see Table A.13 in the Online Appendix).

In Columns 3–5, we supplement our preferred value-added estimates above with three vectors of inputs during school closures, entered sequentially, and examine the stability of treatment effect estimates. First, a vector of *resources for remote learning* available to children including TV, smartphone, internet, computers and WiFi. Second, *compensatory actions from schools and households*, including video lessons, audio lessons, in-person classes, school-assigned homework, home-based help by household members, and private tutoring. Third, *compensatory activities by the child* including accessing YouTube for educational content, educational programs on TV, using books from school, using books from home, and using other internet resources. Table A.9 in the Online Appendix provides summary statistics of these three vectors, separately by individual's participation status in ITK.[xvii] Including vectors for resources for remote instruction and inputs provided by schools and parents, or compensatory activities undertaken by children does not affect our estimates (see Columns 3-5 in Table 5 and Table A.12 in the Online Appendix).

Next, we estimate the sensitivity of our results to further omitted variables bias, following the procedure of Oster (2019) (see Figure 3). We assume that selection-on-unobservables equals selection on observed variables in Table 5 (other than village fixed effects and age which are treated as orthogonal).[xviii] In the sample of private school students, given the negative selection of participants on observed characteristics, this procedure is informative of the extent to which our value-added estimates may understate the program effects of attending ITK. In the public school sample, where selection is mildly positive (see Table 4), this procedure provides bounds of likely upwards bias in our estimates.

We provide estimates for a wide range of parameter values going from 10% to 130% additional variation in Tables A.14-A.15 in the Online Appendix. In practice, we expect much lower incremental variation given the rich set of covariates. For instance, even the extensive vector of inputs added in our validation exercise above, most of which are statistically significant, only raises R^2 by 0.01 (which is 15% additional variation). Thus, the exercise provides an extreme scenario for bias. In the public school sample, where treatment effects are concentrated, assuming that the unobserved variables further increase R^2 by 50% as much as all controls did over the "naive" specification with only village fixed effects and age, reduces the effect size to 0.10σ in language and 0.17σ in math (from 0.11σ and 0.21σ respectively).

IV.C.1 Heterogeneity in ITK program effects

We first investigate whether the effects of ITK differ across public and private school students. We anticipate heterogeneity across this dimension for several reasons. First, students in private schools differ substantially in their characteristics: they have more educated mothers, are from much richer families and have much higher test scores at baseline. As such, they may be more likely to have home support for education. Second, private schools provided substantially more inputs for remote instruction during school closures, including video and audio lectures (see Table A.9 in the Online

Appendix). Thus, even initial learning loss may have been much less severe. Third, it is also likely that school responses to learning loss after reopening also differed between private and public schools. This affects whether ITK centers supplement school-based efforts or provide them in the absence (or ineffectiveness) of such efforts in public school settings.

In Table 6, we estimate value-added models separately for the subsamples of students in public and private schools. Estimates are stable across specifications (including from the naive to the basic and the augmented value-added models), suggesting differences in school type drive most of the selection into ITK enrollment. The effect of the program, however, appears to be concentrated entirely in children enrolled in public schools (with an average effect of 0.2σ in math and 0.11σ in Tamil), with private school effects being close to zero and statistically insignificant in all specifications.

Next, we investigate heterogeneity in program effects across four student characteristics: maternal education, gender, socioeconomic status and age. We estimate this for the sample overall and for private and public school students separately (see Table 7). Overall, we find no significant evidence of heterogeneity. Thus, ITK appears to contribute to the progressivity of cohort-level learning recovery (seen in Table 3) more through the greater participation of disadvantaged students in the ITK program (seen in Table 4) than through differential effects for more-disadvantaged participants.

Finally, we investigate heterogeneity in ITK effects across competencies.[xix] We find that ITK improved student performance across various competencies (Table A.16 in the Online Appendix). This suggests that the program was able to support students across the (baseline) distribution of ability.

IV.D Estimating the contribution of ITK to recovery from learning losses

The sensitivity checks above suggest our estimates approximate the causal effect of ITK. The ITK effects in Table 5 equal ∼36.1% of the estimated recovery of 0.47σ in mathematics between January-May 2022 and ∼48.9% of the estimated recovery of 0.19σ in Tamil (see Table 3). However, our overall estimates of recovery are population-wide, whereas our ITK effects are estimated based on *attending* the after-school classes. Accounting for the 57.3% attendance rate, the ITK program accounts for about 20.7% of the *population-level* catch-up in mathematics and

28% of the catch-up in Tamil between January and May 2022.

Since two-thirds of the learning loss had been bridged, and ∼24.4% of this can be attributed to ITK (averaged across math and Tamil), this implies that around half the learning loss would have been made up even without ITK. This calculation assumes no spillovers from ITK to non-participants. In theory, spillovers could be positive (if ITK made classroom instruction more productive for all students by helping with remediation) or negative (if regular teachers reduced their classroom effort due to ITK). In practice, these spillovers are likely to be second-order since ITK was implemented outside school hours.

Note, however, that the COVID-19 pandemic and its immediate aftermath were a period of substantial flux in the education system. When schools reopened, there was a focus on remediation in schools and a rationalization of the syllabus (recognizing that only a portion of the school year remained and that students had effectively been out of school for a long time). Thus, it is likely that the "non-ITK" recovery reflects, at least in part, the effect of these other measures rather than a "natural" recovery that might have been expected if the education system reopened with business-as-usual.

Finally, the estimated effect of ITK is a *composite* effect that combines additional instructional time, specific remedial content, and a focus on supporting students in (re-)adjusting to in-person instruction. Although we cannot causally decompose the effect of the program into these individual components, both the absolute and relative treatment effects of ITK appear to be consistent with additional instructional time being the primary channel of program impacts. Scheduled time for ITK, of ∼90 minutes per day, is roughly 30% of scheduled instructional time in government primary schools in Tamil Nadu, a proportion also commensurate with the relative contribution of ITK to the overall recovery. Further, Lavy (2015) reports, using data from PISA, that an additional hour of instruction in developing countries increases test scores in a given subject by 0.025 standard deviations.[xx] Attending ITK adds ∼7 hours of instructional time per week. Extrapolating from Lavy (2015)'s estimates should increase test scores by 0.175σ, which is also similar to the treatment effect we estimate for ITK.

V Discussion

We present direct evidence of the severity of learning losses during the COVID-19 pandemic using in-person testing in a near-representative sample of students and IRT-linked test items. While estimated learning losses suggest a developmental lag of one to two years, our results also provide grounds for cautious optimism. Much of the learning loss was recovered within 5–6 months after schools reopened. This recovery was accelerated by a supplemental remedial instruction program implemented by the government on a state-wide scale. We draw three broader lessons from these results.

First, even though the pandemic has affected student achievement adversely (from an already low base), compensating for these losses is possible, even at scale. The most important policy action was simply to reopen schools (which accounted for the majority of the recovery). In addition, programs that provide supplemental remedial instruction can meaningfully accelerate recovery and compensate for regressive learning losses during the pandemic. With sufficient prioritization within the education system, similar programs could be successfully implemented more broadly. Given the breadth of COVID learning losses, this is urgent for the global education community.

Second, continuing such remediation programs may be a cost-effective tool for remedying the 'learning crisis' in developing countries, even beyond the period of post-pandemic recovery (World Bank, 2017). The program has a yearly budget allocation of $\sim$25 million USD (INR 2 billion) and is estimated to have benefited 3.3 million children, yielding an *annual* per-child cost of USD 7.6, and a half-yearly cost of USD 3.8. We estimate substantial gains ($\sim 0.13\sigma$, averaged across subjects) even in 4-5 months of exposure (which is around half a school year), implying a gain of $\sim$3.4 standard deviations per 100 USD, which would be very cost-effective relative to other interventions around the world (Kremer et al., 2013).[xxi] Beyond cross-country and cross-study comparisons, we can also assess the cost-effectiveness of the ITK program relative to the default patterns of education spending in the *same* context. We estimate that the marginal learning gains per unit of expenditure on the ITK program were over 10 times greater than the average returns to status quo spending.[xxii]

This cost-effectiveness is driven by volunteers being paid only modest stipends, that are much lower than regular teacher salaries. However, there were nearly four applicants for every opening, suggesting that the supply of volunteers is unlikely to be a constraint for continuing the program at scale in this setting.[xxiii] Field reports by officials suggest that a key attraction of the program for volunteers was the recognition and respect it provided them in the community, and the empowerment from having a reason to leave the home, and having an independent source of income. Thus, beyond the benefits to students, the ITK program may have also benefited the tutors, and contributed to boosting female labor force participation, which is very low in India (Mehrotra & Parida, 2017).

The combination of effectiveness, cost-effectiveness, positive effects on reducing learning inequality, and likely positive effects on female empowerment and labor force participation, makes the ITK program an attractive candidate for policymakers to continue as a long-term after-school program beyond its short-term use for mitigating COVID-19 learning loss. Further, since this program has *already* been deployed at scale across the state by the government of Tamil Nadu, it suggests that it may be possible for other governments to implement similar programs at scale (Al-Ubaydli et al., 2017).

Finally, understanding the effects of the pandemic and school closures on student human capital will require *repeated* follow-ups in representative samples. The effects of the COVID-19 pandemic on education are expected to be long-lasting, and understanding whether they persist, and how they affect outcomes later in life, are questions of substantial importance. More generally, learning trajectories and persistence in LMICs remain poorly understood (Bau et al., 2021). Yet, the data to generate such evidence, whether through long-run panels such as the NAEP and ECLS in the US or reliable administrative registers as in Scandinavia, do not exist in most LMIC outside of Latin America (Das et al., 2022; Singh, 2020b). Remedying this data deficit should be a priority for public research investment.

References

Agarwal, R. (2022). Pandemic scars may be twice as deep for students in developing countries. *IMF Blog: Insights into Economics and Finance, Feb 3*. Retrieved from https://blogs.imf.org/2022/02/03/pandemic-scars-may-be-twice-as-deep-for-students-in-developing-countries/

Alasino, E., Ramírez, M. J., Romero, M., Schady, N., & Uribe, D. (2024). Learning losses during the COVID-19 pandemic: Evidence from Mexico. *Economics of Education Review*, *98*, 102492.

Al-Ubaydli, O., List, J. A., LoRe, D., & Suskind, D. (2017). Scaling for economists: Lessons from the non-adherence problem in the medical literature. *Journal of Economic Perspectives*, *31*(4), 125–144.

Andrabi, T., Bau, N., Das, J., & Khwaja, A. I. (2022, November). *Heterogeneity in school value-added and the private premium* (Working Paper No. 30627). National Bureau of Economic Research. doi: 10.3386/w30627

Andrabi, T., Daniels, B., & Das, J. (2021). Human capital accumulation and disasters: Evidence from the Pakistan earthquake of 2005. *Journal of Human Resources*, 0520–10887R1.

Andrabi, T., Das, J., & Khwaja, A. I. (2013). Students today, teachers tomorrow: Identifying constraints on the provision of education. *Journal of public Economics*, *100*, 1–14.

Andrabi, T., Das, J., Khwaja, A. I., & Zajonc, T. (2011). Do value-added estimates add value? accounting for learning dynamics. *American Economic Journal: Applied Economics*, *3*(3), 29–54.

Andrew, A., & Salisbury, A. (2023). The educational experiences of Indian children during COVID-19. *Economics of Education Review*, *97*, 102478. Retrieved from https://www.sciencedirect.com/science/article/pii/S0272775723001255 doi: https://doi.org/10.1016/j.econedurev.2023.102478

Angrist, J., Hull, P., Pathak, P. A., & Walters, C. (2021). Credible school value-added with undersubscribed school lotteries. *The Review of Economics and Statistics*, 1–46.

Angrist, J. D., Hull, P. D., Pathak, P. A., & Walters, C. R. (2017). Leveraging lotteries for school value-added: Testing and estimation. *The Quarterly Journal of Economics*, *132*(2), 871–919.

Angrist, N., Ainomugisha, M., Bathena, S. P., Bergman, P., Crossley, C., Cullen, C., ... Sullivan, T. (2023, May). *Building resilient education systems: Evidence from large-scale randomized trials in five countries* (Working Paper No. 31208). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w31208 doi: 10.3386/w31208

Angrist, N., Bergman, P., & Matsheng, M. (2022). Experimental evidence on learning using low-tech when school is out. *Nature Human Behaviour*, 1–10.

Angrist, N., de Barros, A., Bhula, R., Chakera, S., Cummiskey, C., DeStefano, J., ... Stern, J. (2021). Building back better to avert a learning catastrophe: Estimating learning loss from COVID-19 school shutdowns in Africa and facilitating short-term and long-term learning recovery. *International Journal of Educational Development*, *84*, 102397.

Angrist, N., Djankov, S., Goldberg, P. K., & Patrinos, H. A. (2021). Measuring human capital using global learning data. *Nature*, *592*(7854), 403–408.

Ardington, C., Wills, G., & Kotze, J. (2021). COVID-19 learning losses: Early grade reading in South Africa. *International Journal of Educational Development*, *86*, 102480. Retrieved from https://www.sciencedirect.com/science/article/pii/S0738059321001334 doi: https://doi.org/10.1016/j.ijedudev.2021.102480

Bandiera, O., Buehren, N., Goldstein, M., & Rasul, I. (2020). Do school closures during an epidemic have persistent effects? evidence from Sierra Leone in the time of Ebola. *J-PAL working paper*.

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., ... Walton, M. (2017). From proof of concept to scalable policies: Challenges and solutions, with an application. *Journal of Economic Perspectives*, *31*(4), 73–102.

Bau, N. (2022). Estimating an equilibrium model of horizontal competition in education. *Journal of Political Economy*, *130*(7), 000–000.

Bau, N., & Das, J. (2020). Teacher value added in a low-income country. *American Economic Journal: Economic Policy*, *12*(1), 62–96.

Bau, N., Das, J., & Chang, A. Y. (2021). New evidence on learning trajectories in a low-income setting. *International Journal of Educational Development*, *84*, 102430.

Betthäuser, B. A., Bach-Mortensen, A. M., & Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. *Nature Human Behaviour*, *7*(3), 375–385.

Bold, T., Kimenyi, M., Mwabu, G., Ng'ang'a, A., & Sandefur, J. (2018). Experimental evidence on scaling up education reforms in Kenya. *Journal of Public Economics*, *168*, 1-20. Retrieved from https://www.sciencedirect.com/science/article/pii/S0047272718301518 doi: https://doi.org/10.1016/j.jpubeco.2018.08.007

Carlana, M., & La Ferrara, E. (2021). *Apart but connected: Online tutoring and student outcomes during the COVID-19 pandemic* (Tech. Rep. No. IZA DP No. 14094). Institute of Labor Economics (IZA). Retrieved from https://www.hks.harvard.edu/publications/apart-connected-online-tutoring-and-student-outcomes-during-covid-19-pandemic

CBGA. (2018). *Budgeting for school education in Tamil Nadu: What has changed and what has not?* (Tech. Rep.). Centre for Budget and Governance Accountability (CBGA) and Child Rights and You (CRY). Retrieved from https://www.cry.org/wp-content/uploads/2019/03/Budgeting-for-School-Education-inTamil-Nadu.pdf

Chalmers, R. P. (2012). mirt: A multidimensional Item Response Theory package for the R environment. *Journal of Statistical Software*, *48*(6), 1–29. doi: 10.18637/jss.v048.i06

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104*(9), 2593–2632.

Crawfurd, L., Evans, D. K., Hares, S., & Sandefur, J. (2023). Live tutoring calls did not improve learning during the COVID-19 pandemic in Sierra Leone. *Journal of Development Economics*, *164*, 103114.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.

Das, J., Singh, A., & Chang, A. Y. (2022). Test scores and educational opportunities: Panel evidence from five low-and middle-income countries. *Journal of Public Economics*, *206*, 104570.

Das, J., & Zajonc, T. (2010). India shining and bharat drowning: Comparing two Indian states to the worldwide distribution in mathematics achievement. *Journal of Development Economics*, *92*(2), 175–187.

Deming, D. J. (2014). Using school choice lotteries to test measures of school effectiveness. *American Economic Review*, *104*(5), 406–11.

Deming, D. J., Hastings, J. S., Kane, T. J., & Staiger, D. O. (2014). School choice, school quality, and postsecondary attainment. *American Economic Review*, *104*(3), 991–1013.

Duflo, A., Kiessel, J., & Lucas, A. (2020, June). *Experimental evidence on alternative policies to increase learning at scale* (Working Paper No. 27298). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w27298 doi: 10.3386/w27298

Filmer, D., Rogers, H., Angrist, N., & Sabarwal, S. (2020). Learning-adjusted years of schooling (LAYS): Defining a new macro measure of education. *Economics of Education Review*, *77*, 101971. Retrieved from https://www.sciencedirect.com/science/article/pii/S0272775719300263 doi: https://doi.org/10.1016/j.econedurev.2020.101971

Ganimian, A. J., Muralidharan, K., & Walters, C. R. (in press). Augmenting state capacity for child development: Experimental evidence from India. *Journal of Political Economy*. Retrieved from https://doi.org/10.1086/728109 doi: 10.1086/728109

Guariso, A., & Björkman Nyqvist, M. (2023). The impact of the COVID-19 pandemic on children's learning and wellbeing: Evidence from India. *Journal of Development Economics*, *164*, 103133. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304387823000883 doi: https://doi.org/10.1016/j.jdeveco.2023.103133

Hassan, H., Islam, A., Siddique, A., & Wang, L. C. (2021). *Telementoring and homeschooling during school closures: A randomized experiment in rural Bangladesh* (Tech. Rep.). TUM School of Governance at the Technical University of Munich. Retrieved from https://ideas.repec.org/p/aiw/wpaper/13.html

Hevia, F. J., Vergara-Lope, S., Velásquez-Durán, A., & Calderón, D. (2022). Estimation of the fundamental learning loss and learning poverty related to COVID-19 pandemic in Mexico. *International Journal of Educational Development*, *88*, 102515.

Jha, P., Deshmukh, Y., Tumbe, C., Suraweera, W., Bhowmick, A., Sharma, S., ... Brown, P. (2022). COVID mortality in india: National survey data and health facility deaths. *Science*, *375*(6581), 667-671. Retrieved from https://www.science.org/doi/abs/10.1126/science.abm5154 doi: 10.1126/science.abm5154

Kesar, S., Abraham, R., Lahoti, R., Nath, P., & Basole, A. (2021). Pandemic, informality, and vulnerability: Impact of COVID-19 on livelihoods in India. *Canadian Journal of Development Studies/Revue canadienne d'études du développement*, *42*(1-2), 145–164.

Kim, J. S., & Quinn, D. M. (2013). The effects of summer reading on low-income children's literacy achievement from kindergarten to grade 8: A meta-analysis of classroom and home interventions. *Review of Educational Research*, *83*(3), 386–431.

Kremer, M., Brannen, C., & Glennerster, R. (2013). The challenge of education and learning in the developing world. *Science*, *340*(6130), 297–300.

Lavy, V. (2015). Do differences in schools' instruction time explain international achievement gaps? evidence from developed and developing countries. *The Economic Journal*, *125*(588), F397–F424.

Lichand, G., & Doria, C. A. (2022). *The lasting impacts of remote learning in the absence of remedial policies: Evidence from Brazil* (Tech. Rep.). SSRN. Retrieved from http://dx.doi.org/10.18235/0003344

Lichand, G., Doria, C. A., Leal-Neto, O., & Fernandes, J. P. C. (2022). The impacts of remote learning in secondary education during the pandemic in Brazil. *Nature Human Behaviour*, 1–8.

Lynch, K., An, L., & Mancenido, Z. (2022). The impact of summer programs on student mathematics achievement: A meta-analysis. *Review of Educational Research*, 00346543221105543.

Mehrotra, S., & Parida, J. K. (2017). Why is the labour force participation of women declining in India? *World Development*, *98*, 360–380.

Moscoviz, L., & Evans, D. K. (2022). Learning loss and student dropouts during the COVID-19 pandemic: A review of the evidence two years after schools shut down. *Center for Global Development, Working Paper*, *609*.

Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting education? experimental evidence on technology-aided instruction in India. *American Economic Review*, *109*(4), 1426–60.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187-204. Retrieved from https://doi.org/10.1080/07350015.2016.1227711 doi: 10.1080/07350015.2016.1227711

Patrinos, H. A., Vegas, E., & Carter-Rau, R. (2022). *An analysis of COVID-19 student learning loss* (Policy Research Working Paper Series No. 10033). The World Bank.

Pratham. (2019). *Annual status of education report* (Tech. Rep.). Author. Retrieved from http://img.asercentre.org/docs/ASER%202019/ASER2019%20report%20/aserreport2019earlyyearsfinal.pdf

Pratham. (2021a). *Annual status of education report Chhattisgarh (rural) 2021* (Tech. Rep.). Author. Retrieved from http://img.asercentre.org/docs/asercg2021_fullreport_11.01.2021.pdf

Pratham. (2021b). *Annual status of education report West Bengal (rural) 2021* (Tech. Rep.). Author. Retrieved from http://img.asercentre.org/docs/aserwb20215-pager08.02.202211.36amfinal1.pdf

Pritchett, L., & Beatty, A. (2015). Slow down, you're going too fast: Matching curricula to student skill levels. *International Journal of Educational Development*, *40*, 276–288.

Ramachandran, V., Chatterjee, P., Ravi, A., Linden, T., Beteille, T., Dey, S., … Mehta, C. (2015). Teachers in the Indian education system: Synthesis of a nine-state study. *National University of Educational Planning and Administration, NUEPA, New Delhi*.

Schaub, S. (2020). Robomit: robustness checks for omitted variable bias. *R package version*, *1*(3).

Singh, A. (2015). Private school effects in urban and rural India: Panel estimates at primary and secondary school ages. *Journal of Development Economics*, *113*, 16–32.

Singh, A. (2020a). Learning more with every year: School year productivity and international learning divergence. *Journal of the European Economic Association*, *18*(4), 1770–1813.

Singh, A. (2020b). Myths of official measurement: Auditing and improving administrative data in developing countries. *Research on Improving Systems of Education (RISE) Working Paper*, *42*.

Singh, A., Romero, M., & Muralidharan, K. (2024). *Replication Data for: "COVID-19 Learning loss and recovery: Panel data evidence from India"*. Harvard Dataverse. Retrieved from https://doi.org/10.7910/DVN/XGY7CV doi: 10.7910/DVN/XGY7CV

Todd, P. E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, *113*(485), F3–F33.

Todd, P. E., & Wolpin, K. I. (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital*, *1*(1), 91–136.

UNESCO. (2022). *UNESCO global dataset on the duration of school closures.* Retrieved from https://covid19.uis.unesco.org/global-monitoring-school-closures-covid19/

Vivalt, E. (2020). How much can we generalize from impact evaluations? *Journal of the European Economic Association*, *18*(6), 3045–3089.

World Bank. (2017). *World development report 2018: Learning to realize education's promise*. The World Bank.

World Bank. (2020). *The COVID-19 pandemic: Shocks to education and policy responses*. Author.

World Bank, UNESCO and UNICEF. (2021). *The state of the global education crisis: Pathways to recovery*. Author.

Table 1: Balance on respondent characteristics across survey rounds

	As randomized			Actual timing			
	(1) Early follow-up	(2) Late follow-up	(3) p-value H_0: Equality	(4) Dec/21- Jan/22	(5) Feb/22- Mar/22	(6) Mar/22- May/22	(7) p-value H_0: Equality
Male	0.51	0.51	0.28	0.51	0.52	0.51	0.553
	(0.50)	(0.50)		(0.50)	(0.50)	(0.50)	
	[9,547]	[9,752]		[5,554]	[3,993]	[9,752]	
Mother Edu: $<$ Gr. 9	0.35	0.34	0.05^{**}	0.35	0.36	0.34	0.121
	(0.48)	(0.47)		(0.48)	(0.48)	(0.47)	
	[9,480]	[9,672]		[5,517]	[3,963]	[9,672]	
Mother Edu: Gr. 9-11	0.31	0.33	0.04^{**}	0.32	0.31	0.33	0.097^{*}
	(0.46)	(0.47)		(0.47)	(0.46)	(0.47)	
	[9,480]	[9,672]		[5,517]	[3,963]	[9,672]	
Mother Edu: Gr. 12+	0.33	0.33	0.76	0.33	0.34	0.33	0.486
	(0.47)	(0.47)		(0.47)	(0.47)	(0.47)	
	[9,480]	[9,672]		[5,517]	[3,963]	[9,672]	
SES Decile	4.93	4.97	0.38	4.99	4.85	4.97	0.563
	(2.84)	(2.84)		(2.79)	(2.92)	(2.84)	
	[9,547]	[9,752]		[5,554]	[3,993]	[9,752]	
Math (2019)	-0.00	0.00	0.43	-0.01	0.01	0.00	0.725
	(1.10)	(1.08)		(1.10)	(1.11)	(1.08)	
	[9,547]	[9,752]		[5,554]	[3,993]	[9,752]	
Tamil (2019)	0.00	0.00	0.71	-0.00	0.01	0.00	0.908
	(0.65)	(0.64)		(0.64)	(0.65)	(0.64)	
	[9,547]	[9,752]		[5,554]	[3,993]	[9,752]	
Government school (2020-21)	0.51	0.50	0.25	0.51	0.51	0.50	0.493
	(0.50)	(0.50)		(0.50)	(0.50)	(0.50)	
	[9,301]	[9,751]		[5,312]	[3,989]	[9,751]	
Private school (2020-21)	0.29	0.27	0.11	0.29	0.29	0.27	0.281
	(0.45)	(0.45)		(0.45)	(0.45)	(0.45)	
	[9,301]	[9,751]		[5,312]	[3,989]	[9,751]	
Age at baseline (months)	55.98	55.76	0.47	55.87	56.13	55.76	0.293
	(19.39)	(19.54)		(19.35)	(19.45)	(19.54)	
	[9,547]	[9,752]		[5,554]	[3,993]	[9,752]	

Notes: The first three columns of this table present the mean and the standard deviation (in parenthesis) for children randomly assigned to be surveyed early (Column 1) and those randomly assigned to be surveyed late (Column 2). The number of observations appears in square brackets. The p-value in Column 3 is for a statistical test where the null is that the means within village (i.e., taking into account village fixed effects) are equal, clustering standard errors at the village level. The p-value of the joint F-statistic across rounds is 0.28. The last four columns of this table present the mean and the standard deviation (in parenthesis) for each of the actual survey waves (Columns 4-6). The p-value in Column 7 is for a statistical test where the null is that all three means within village (i.e., taking into account village fixed effects) are equal, clustering standard errors at the village level. The p-value of the joint F-statistic across wave 1 and wave 2 is 0.45. The p-value of the joint F-statistic across wave 1 and wave 3 is 0.68. The p-value of the joint F-statistic across wave 2 and wave 3 is 0.34. Math and Tamil (2019) baseline scores correspond to the residuals after regressing the original scores on age brackets (in discrete years) and the age in months. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 2: Learning loss between August 2019 and December 2021

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Learning loss at different ages								
	Math				Tamil			
Age (in months)	60	72	84	96	60	72	84	96
IRT score (Aug 2019)	-0.61	0.23	0.80	1.02	-0.14	0.34	0.68	0.84
IRT score (Dec 2021)	-1.04	-0.46	0.06	0.28	-0.29	0.02	0.28	0.42
Absolute loss (in SD)	0.43	0.69	0.74	0.75	0.15	0.32	0.40	0.41
Developmental lag (in months)	10.0	10.0	14.5	23.5	5.5	8.0	13.5	21.5
Panel B: Learning loss in regression form								
	Math score (in SD)				Tamil score (in SD)			
Wave 1 (Dec 2021)	-.73***	-.74***	-.76***	-.75***	-.35***	-.35***	-.37***	-.38***
	(.031)	(.038)	(.042)	(.049)	(.02)	(.023)	(.027)	(.029)
Male $\times$ Dec 21		.023				-.0074		
		(.041)				(.022)		
Mother Edu: Gr. 9-11 $\times$ Dec 21			.019				.0015	
			(.053)				(.03)	
Mother Edu: Gr. 12+ $\times$ Dec 21			.09*				.06**	
			(.049)				(.025)	
SES Decile $\times$ Dec 21				.0046				.0061
				(.0075)				(.0039)
N. of obs.	13,083	13,083	13,083	13,083	13,083	13,083	13,083	13,083
R-squared	.33	.33	.33	.33	.31	.31	.31	.31

Notes: Panel A presents, for children of different ages, the raw IRT score in wave 0 (Aug 2019) and wave 1 (Dec 2021), as well as the difference between the two (the absolute learning loss in standard deviations), and the developmental lag (i.e., how much longer, in months, it took a student in 2021 to achieve the same score as a student in 2019). Panel B estimates the learning loss following Equation 1. The estimation sample is restricted to individuals tested in Aug 2019 (Wave 0) or December 2021 (Wave 1) who were aged between 55–95 months at the time of the test. All regressions in Panel B include village fixed effects and control for age, gender, maternal education, and SES percentile. Test scores are normalized for age 60–72 months in 2019. Standard errors are clustered at the village level. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 3: Recovery from learning loss

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Recovery at different ages								
	Math				Tamil			
Age (in months)	60	72	84	96	60	72	84	96
IRT score (Aug 2019)	-0.61	0.23	0.80	1.02	-0.14	0.34	0.68	0.84
IRT score (Dec 2021)	-1.04	-0.46	0.06	0.28	-0.29	0.02	0.28	0.42
IRT score (Feb 2022)	-0.72	-0.18	0.31	0.66	-0.13	0.17	0.42	0.69
IRT score (Apr 2022)	-0.62	-0.02	0.55	0.88	-0.10	0.20	0.48	0.75
Absolute loss (in SD)	0.43	0.69	0.74	0.75	0.15	0.32	0.40	0.41
Absolute recovery (in SD) by Feb 22	0.32	0.28	0.26	0.38	0.16	0.15	0.14	0.26
Absolute recovery (in SD) by Apr 22	0.42	0.44	0.49	0.60	0.19	0.17	0.20	0.32
Panel B: Recovery in regression form								
	Math score (in SD)				Tamil score (in SD)			
Wave 2 (Feb 2022)	.24***	.27***	.23***	.26***	.12***	.11***	.12***	.14***
	(.043)	(.047)	(.056)	(.061)	(.024)	(.026)	(.031)	(.031)
Wave 3 (April 2022)	.46***	.48***	.48***	.55***	.19***	.19***	.2***	.23***
	(.025)	(.03)	(.037)	(.043)	(.013)	(.016)	(.02)	(.021)
Interactions:								
Male × Feb 22		-.068				.023		
		(.045)				(.023)		
Male × Apr 22		-.044				-.0024		
		(.033)				(.017)		
Mother Edu: Gr. 9-11 × Feb 22			.027				.0047	
			(.057)				(.029)	
Mother Edu: Gr. 9-11 × Apr 22			.068				.023	
			(.047)				(.025)	
Mother Edu: Gr. 12+ × Feb 22			-.014				-.023	
			(.061)				(.031)	
Mother Edu: Gr. 12+ × Apr 22			-.13***				-.061**	
			(.043)				(.024)	
SES Decile × Feb 22				-.0055				-.0045
				(.0089)				(.0042)
SES Decile × Apr 22				-.017**				-.008**
				(.0069)				(.0034)
N. of obs.	18,978	18,978	18,978	18,978	18,978	18,978	18,978	18,978
R-squared	.4	.4	.4	.4	.46	.46	.46	.46

Notes: Panel A presents, for children of different ages, the raw IRT score in wave 1 (Dec 2021), wave 2 (Feb 2022), and wave 3 (Apr 2022), as well as the difference between the wave 2 and 3 with wave 1 (the absolute recovery in standard deviations). Panel B estimates the rate of recovery via regressions by comparing test scores in wave 1, 2 and 3. The estimation sample is restricted to individuals who were aged between 55–131 months at the time of the survey and tested in December 2021 (Wave 1), February 2022 (Wave 2), or April 2022 (Wave 3). Standard errors are clustered at the village level. All regressions include village fixed effects and control for test scores in 2019, age, gender, maternal education, and SES percentile. Test scores are normalized for age 60–72 months in 2019. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 4: Difference in characteristics across *Illam Thedi Kalvi (ITK)* participants and non-participants

	Overall			Public			Private		
	(1) Does not attend ITK	(2) Attend ITK	(3) Difference (village FE)	(4) Does not attend ITK	(5) Attend ITK	(6) Difference (village FE)	(7) Does not attend ITK	(8) Attend ITK	(9) Difference (village FE)
Male	0.52	0.49	-0.03**	0.49	0.49	-0.00	0.55	0.55	-0.00
	(0.50)	(0.50)	(0.01)	(0.50)	(0.50)	(0.02)	(0.50)	(0.50)	(0.03)
	[3,830]	[5,136]	[8,966]	[1,596]	[4,616]	[6,212]	[1,789]	[421]	[2,198]
Age in months	86.66	93.79	8.05***	89.04	94.48	6.17***	90.04	93.11	3.37***
	(19.11)	(17.45)	(0.49)	(18.11)	(16.98)	(0.56)	(18.45)	(18.46)	(1.25)
	[3,830]	[5,136]	[8,966]	[1,596]	[4,616]	[6,212]	[1,789]	[421]	[2,198]
Mother Edu: $<$ Gr. 9	0.29	0.39	0.09***	0.44	0.41	-0.00	0.14	0.22	0.07**
	(0.45)	(0.49)	(0.01)	(0.50)	(0.49)	(0.02)	(0.35)	(0.41)	(0.03)
	[3,806]	[5,096]	[8,902]	[1,584]	[4,578]	[6,162]	[1,779]	[419]	[2,185]
Mother Edu: Gr. 9-11	0.31	0.35	0.03**	0.33	0.35	0.01	0.31	0.32	0.01
	(0.46)	(0.48)	(0.01)	(0.47)	(0.48)	(0.02)	(0.46)	(0.47)	(0.04)
	[3,806]	[5,096]	[8,902]	[1,584]	[4,578]	[6,162]	[1,779]	[419]	[2,185]
Mother Edu: Gr. 12+	0.39	0.26	-0.13***	0.24	0.24	-0.01	0.55	0.46	-0.08**
	(0.49)	(0.44)	(0.01)	(0.43)	(0.43)	(0.02)	(0.50)	(0.50)	(0.04)
	[3,806]	[5,096]	[8,902]	[1,584]	[4,578]	[6,162]	[1,779]	[419]	[2,185]
SES Decile	5.42	4.59	-0.77***	4.40	4.43	-0.01	6.53	6.31	-0.30
	(2.91)	(2.73)	(0.09)	(2.80)	(2.66)	(0.11)	(2.62)	(2.83)	(0.18)
	[3,830]	[5,136]	[8,966]	[1,596]	[4,616]	[6,212]	[1,789]	[421]	[2,198]
Math (2019)	0.08	-0.05	-0.11***	-0.10	-0.07	0.04	0.26	0.21	-0.03
	(1.12)	(1.09)	(0.03)	(1.09)	(1.08)	(0.04)	(1.19)	(1.16)	(0.07)
	[3,830]	[5,136]	[8,966]	[1,596]	[4,616]	[6,212]	[1,789]	[421]	[2,198]
Tamil (2019)	0.04	-0.02	-0.06***	-0.04	-0.03	0.01	0.11	0.06	-0.05
	(0.65)	(0.65)	(0.02)	(0.66)	(0.65)	(0.03)	(0.67)	(0.70)	(0.04)
	[3,830]	[5,136]	[8,966]	[1,596]	[4,616]	[6,212]	[1,789]	[421]	[2,198]
Gov. school (2021-22)	0.42	0.90	0.47***						
	(0.49)	(0.30)	(0.02)						
	[3,830]	[5,136]	[8,966]						
Private school (2021-22)	0.47	0.08	-0.35***						
	(0.50)	(0.27)	(0.02)						
	[3,830]	[5,136]	[8,966]						
AWC (2021-22)	0.10	0.02	-0.10***						
	(0.30)	(0.13)	(0.01)						
	[3,830]	[5,136]	[8,966]						

Notes: This table presents the mean and the standard deviation (in parenthesis) for children who do not attend ITK (Columns 1, 4, and 7) and those who attend (Columns 2, 5, and 8). The number of observations appears in square brackets. Columns 3, 6, and 9 have the difference in means within village (i.e., after taking into account village fixed effects), as well as the standard error, clustered at the village level, of the difference (in parenthesis). Columns 1-3 use the full sample, while Columns 4-6 restrict the sample to children enrolled in public schools and Columns 7-9 restrict the sample to children enrolled in private schools. The estimation sample is restricted to individuals who were aged between 55–131 months at the time of the survey. Math and Tamil (2019) baseline scores correspond to the residuals after regressing the original scores on age brackets (in discrete years) and the age in months. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 5: Assessing effect of *Illam Thedi Kalvi (ITK)*

	Naive	VAM	Augmented		
	(1)	(2)	(3)	(4)	(5)
Panel A: Math					
ITK effect	.08***	.17***	.16***	.17***	.16***
	(.027)	(.026)	(.026)	(.025)	(.025)
N. of obs.	8,966	8,902	8,901	8,901	8,901
R-squared	.32	.38	.39	.39	.39
Panel B: Tamil					
ITK effect	.073***	.093***	.09***	.092***	.083***
	(.015)	(.015)	(.015)	(.015)	(.014)
N. of obs.	8,966	8,902	8,901	8,901	8,901
R-squared	.4	.45	.45	.46	.46
Child demographic characteristics	Yes	Yes	Yes	Yes	Yes
Household characteristics	No	Yes	Yes	Yes	Yes
Lagged achievement	No	Yes	Yes	Yes	Yes
Enrollment type	No	Yes	Yes	Yes	Yes
Resources for remote instruction	No	No	Yes	Yes	Yes
Compensatory inputs from parents and schools	No	No	No	Yes	Yes
Child educational activities	No	No	No	No	Yes

Notes: The estimation sample is restricted to individuals tested during wave 3 (March-May of 2022) who were aged between 55–131 months at the time of the test. Column 1 has a naive specification that only controls for children's demographic characteristics (age and gender). Column 2 has the standard value-added model (VAM) specification, which controls for children's demographic characteristics, for household characteristics (maternal education and SES percentile), for lagged tests scores (in math and Tamil) allowing the effect of the lagged score to vary by age, and for enrollment type (private, public or out of school). Columns 3-5 have augmented specifications that also control for resources during remote instruction, compensatory inputs from parents and schools, and child educational activities. Table A.9 in the Online Appendix presents mean values for these inputs and Table A.12 in the Online Appendix presents the full list of estimated coefficients. Standard errors are clustered at the village level. All regressions include village fixed effects. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 6: Assessing effect of *Illam Thedi Kalvi* (ITK) for children in public and private schools

	Naive	VAM	Augmented		
	(1)	(2)	(3)	(4)	(5)
Panel A: Public					
ITK effect on math test scores	.21***	.2***	.19***	.19***	.18***
	(.031)	(.031)	(.031)	(.03)	(.03)
N. of obs.	6,212	6,162	6,161	6,161	6,161
R-squared	.28	.32	.33	.33	.33
ITK effect on Tamil test scores	.11***	.11***	.11***	.11***	.099***
	(.018)	(.017)	(.017)	(.017)	(.017)
N. of obs.	6,212	6,162	6,161	6,161	6,161
R-squared	.39	.43	.43	.43	.44
Panel B: Private					
ITK effect on math test scores	-.044	-.015	-.014	.014	.014
	(.065)	(.067)	(.067)	(.069)	(.067)
N. of obs.	2,198	2,185	2,185	2,185	2,185
R-squared	.33	.36	.36	.37	.38
ITK effect on Tamil test scores	-.021	-.01	-.01	-.0008	-.006
	(.043)	(.043)	(.044)	(.044)	(.043)
N. of obs.	2,198	2,185	2,185	2,185	2,185
R-squared	.39	.41	.41	.42	.43
Child demographic characteristics	Yes	Yes	Yes	Yes	Yes
Household characteristics	No	Yes	Yes	Yes	Yes
Lagged achievement	No	Yes	Yes	Yes	Yes
Enrollment type	No	Yes	Yes	Yes	Yes
Resources for remote instruction	No	No	Yes	Yes	Yes
Compensatory inputs from parents and schools	No	No	No	Yes	Yes
Child educational activities	No	No	No	No	Yes

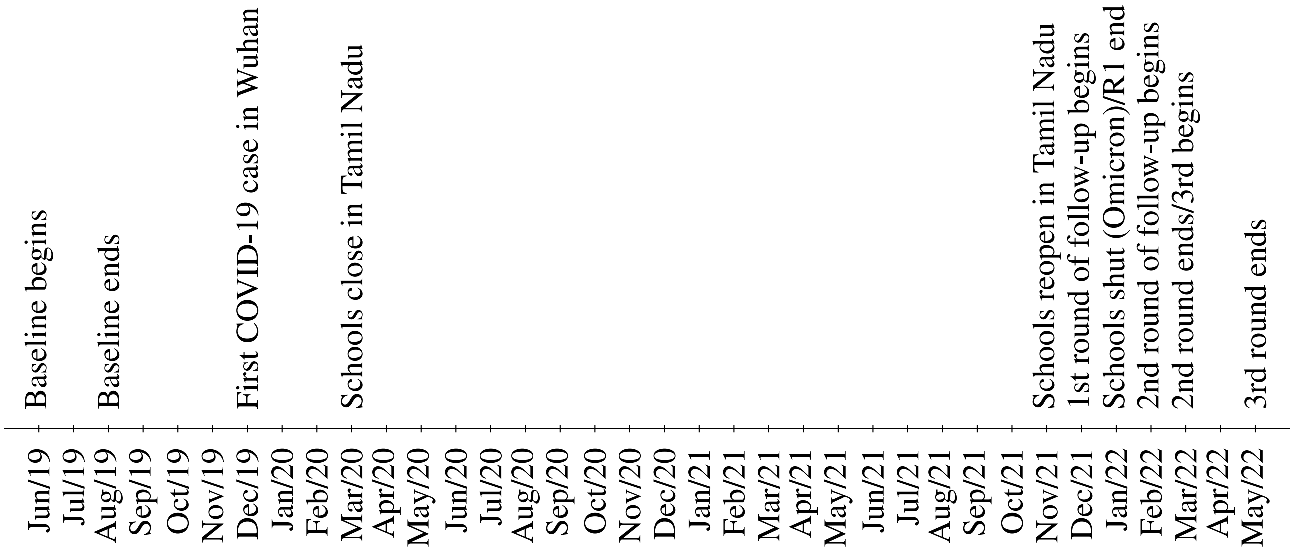
Notes: The estimation sample is restricted to individuals tested during wave 3 (March-May of 2022) who were aged between 55–131 months at the time of the test. Column 1 has a naive specification that only controls for children's demographic characteristics (age and gender). Column 2 has the standard value-added model (VAM) specification, which controls for children's demographic characteristics, for household characteristics (maternal education and SES percentile), for lagged tests scores (in math and Tamil) allowing the effect of the lagged score to vary by age, and for enrollment type (private, public or out of school). Columns 3-5 have augmented specifications that also control for resources during remote instruction, compensatory inputs from parents and schools, and child educational activities. Panel A presents results for children enrolled in public schools, while Panel B for children enrolled in private schools. Standard errors are clustered at the village level. All regressions include village fixed effects. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

Table 7: Heterogeneity in effect of *Illam Thedi Kalvi* (ITK)

	Math					Tamil				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Overall										
If child attends ITK	.21***	.19***	.19***	.31***	.17***	.11***	.093***	.11***	.11*	.17***
	(.041)	(.049)	(.032)	(.11)	(.026)	(.022)	(.028)	(.018)	(.06)	(.026)
Interactions:										
ITK × Mother Edu: Gr. 9-11	-.021					-.015				
	(.057)					(.029)				
ITK × Mother Edu: Gr. 12+	-.1*					-.048				
	(.054)					(.032)				
ITK × SES Decile		-.0033					-.000019			
		(.008)					(.0043)			
ITK × Male			-.04					-.036		
			(.038)					(.022)		
ITK × Age				-.0015					-.0002	
				(.0012)					(.00066)	
ITK × Baseline score					.0023					.032
					(.02)					(.038)
N. of obs.	8,902	8,902	8,902	8,902	8,902	8,902	8,902	8,902	8,902	8,902
R-squared	.38	.38	.38	.38	.38	.45	.45	.45	.45	.38
Panel B: Public										
If child attends ITK	.22***	.23***	.23***	.39***	.2***	.11***	.1***	.14***	.2**	.2***
	(.047)	(.06)	(.037)	(.14)	(.031)	(.026)	(.032)	(.022)	(.079)	(.031)
Interactions:										
ITK × Mother Edu: Gr. 9-11	-.011					.0097				
	(.069)					(.038)				
ITK × Mother Edu: Gr. 12+	-.09					-.013				
	(.077)					(.043)				
ITK × SES Decile		-.0076					.0014			
		(.011)					(.0057)			
ITK × Male			-.066					-.053*		
			(.048)					(.029)		
ITK × Age				-.0022					-.00094	
				(.0016)					(.00084)	
ITK × Baseline score					-.014					.0025
					(.03)					(.054)
N. of obs.	6,162	6,162	6,162	6,162	6,162	6,162	6,162	6,162	6,162	6,162
R-squared	.32	.32	.32	.32	.32	.43	.43	.43	.43	.32
Panel C: Private										
If child attends ITK	-.17	.00087	-.034	-.042	-.0068	.022	.05	-.021	.078	-.012
	(.12)	(.13)	(.088)	(.3)	(.067)	(.068)	(.086)	(.048)	(.17)	(.066)
Interactions:										
ITK × Mother Edu: Gr. 9-11	.19					-.083				
	(.16)					(.087)				
ITK × Mother Edu: Gr. 12+	.19					-.01				
	(.14)					(.084)				
ITK × SES Decile		-.0024					-.0095			
		(.018)					(.011)			
ITK × Male			.035					.02		
			(.1)					(.058)		
ITK × Age				.0003					-.00096	
				(.0034)					(.0019)	
ITK × Baseline score					-.033					-.043
					(.04)					(.085)
N. of obs.	2,185	2,185	2,185	2,185	2,185	2,185	2,185	2,185	2,185	2,185
R-squared	.36	.36	.36	.36	.36	.41	.41	.41	.41	.36

Notes: Standard errors are clustered at the village level. All regressions include village fixed effects and control for lagged tests scores (in math and Tamil) allowing the effect of the lagged score to vary by age. Statistical significance at the 1, 5, 10% levels is indicated by ***, **, and *.

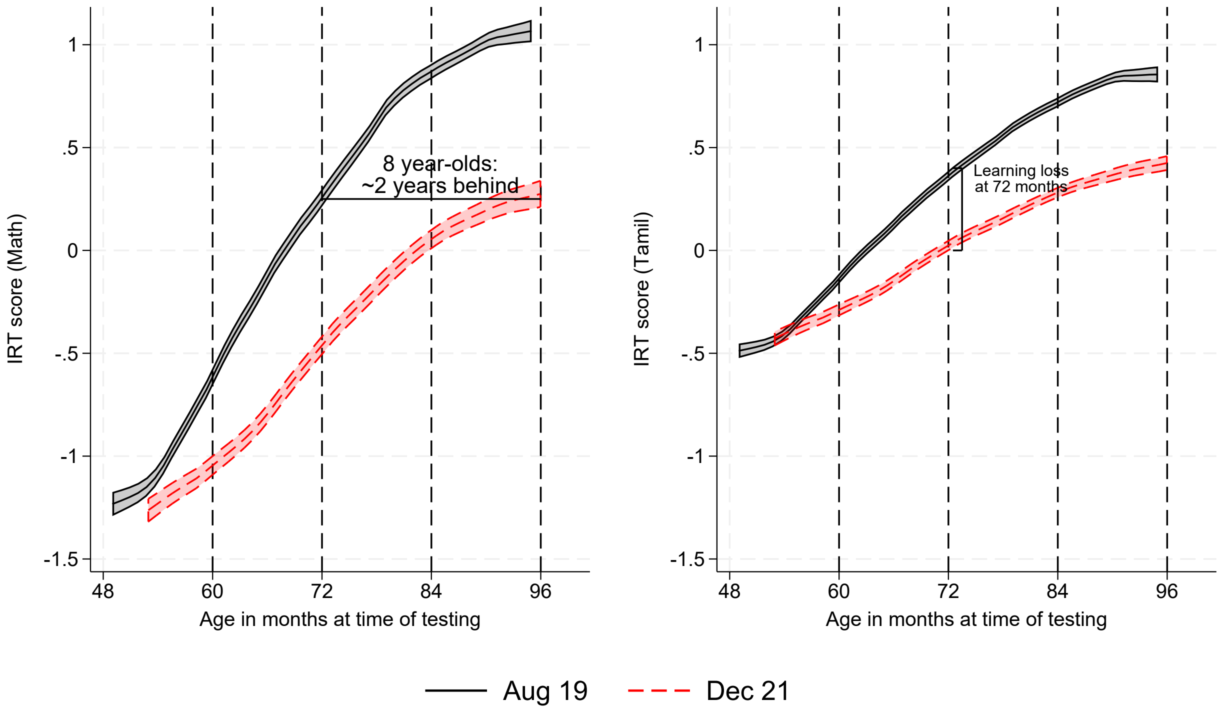
Figure 1: Timeline



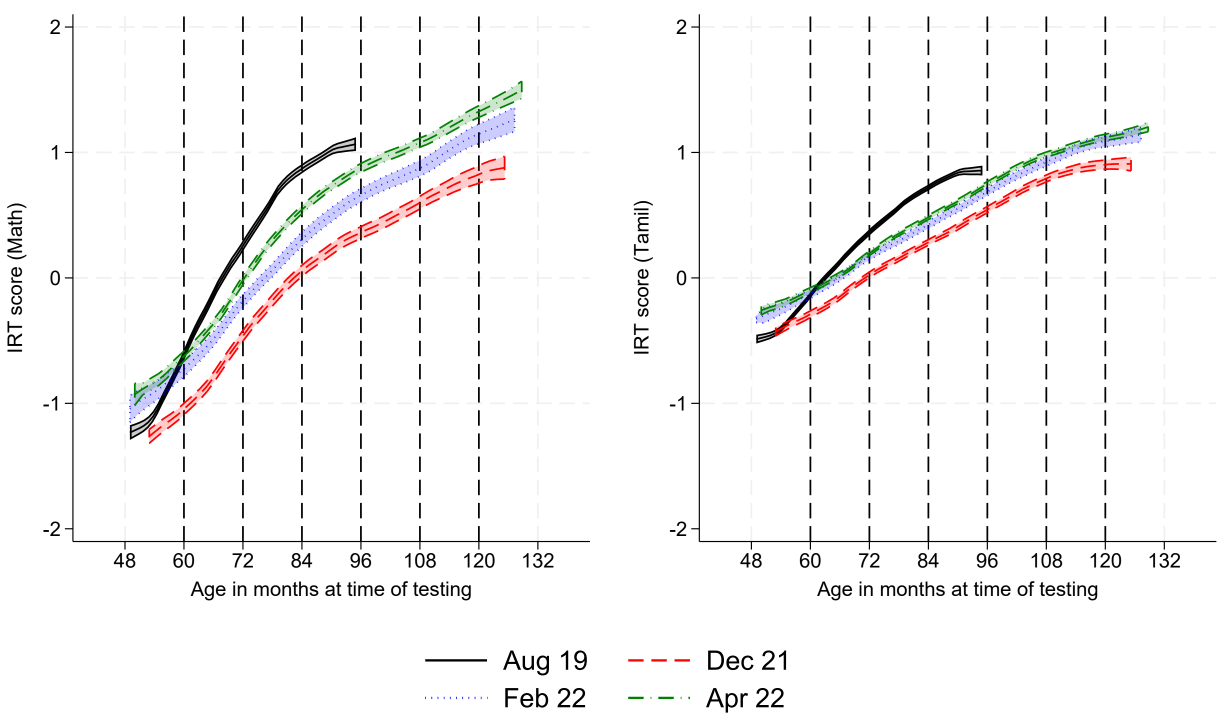
Note: This figure shows the timeline of data collection and of key events during the COVID-19 pandemic and school closures.

Figure 2: Learning loss and recovery in test scores across survey waves

(a) Learning loss in December 2021



(b) Recovery between December 2021 and May 2022

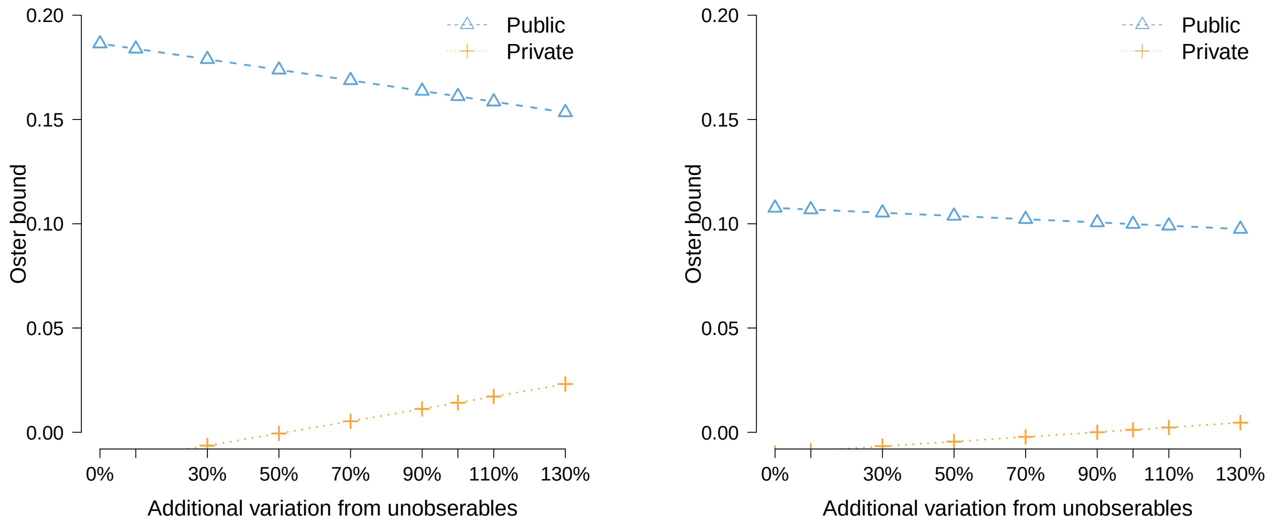


Note: These figures present local polynomial regressions with respect to age at the time of test-taking across the four survey waves in the data. At any age, the decline in scores from Aug 2019 to Dec 2021 measures learning loss. The shift from December 2021 to the two subsequent survey waves measures the degree of recovery for children of a particular age at the time of testing (horizontal axis).

Figure 3: Oster (2019) bounds of *Illam Thedi Kalvi (ITK)*

(a) Math

(b) Tamil



Note: These figures present bias-adjusted treatment effects (i.e., Oster (2019) bounds). The x-axis represents the additional variation (i.e., increase in R^2) from controlling by unobservables (as a function of the increase from controlling by observables). As long as the selection on unobservables is at most as large as the selection on observables (i.e., $\delta = 1$ in Oster (2019)), the treatment effect is between the original estimate from controlling by observables (represented in the figure by the point when the additional variation from unobservables is 0%) and the Oster (2019) bound. Tables A.14 and A.15 in the Online Appendix present additional details from the bounds estimation.

Notes

[i]Value-added models have been shown to recover similar effects as estimates based on experiments, lottery-induced variation, regression discontinuity designs, and dynamic panel models, both in the US (Chetty et al., 2014; Deming et al., 2014; J. D. Angrist et al., 2017; J. Angrist et al., 2021) and in developing countries (Andrabi et al., 2011; Bau & Das, 2020; Singh, 2015, 2020a).

[ii]Since we control for attending private or public school, this exercise considers robustness to selection on omitted variables in ITK participation *within* public and private schools.

[iii]Of the 36 studies reviewed in Patrinos et al. (2022), only one features representative samples of primary school students with in-person testing in an LMIC (Hevia et al. (2022) in Mexico).

[iv]The Andrabi et al. (2021) results may reflect a setting where pedagogy had not adjusted for school disruptions (because only a small fraction of students were affected). This may have contributed to their finding that a 4-month

school closure led to even larger learning gaps over time for children of less educated mothers. In contrast, the COVID-19 education shock affected all children, and the Government of Tamil Nadu prioritized remediation when schools reopened.

[v]There is a large literature on after-school programs in the US, where reviews of the evidence suggest that summer and after-school programs can help improve learning outcomes, mitigate summer learning loss, and reduce socioeconomic gaps in learning (Lynch et al., 2022; Kim & Quinn, 2013).

[vi]This round of fieldwork was done as a baseline for an experimental evaluation of a government program to improve preschool education. Given the onset of the pandemic, and subsequent preschool and school closures from March 2020, the intervention and the evaluation were canceled. See https://doi.org/10.1257/rct.5599 for more details.

[vii]The main reasons for not being able to interview the household was that it had moved to a different location (∼48% of the incomplete surveys) or that the household was not in their residence when we visited multiple times (∼25% of the incomplete surveys). Roughly ∼18% of the incomplete surveys are because households refused consent. The rest are mostly related to children not being available for different reasons.

[viii]While we focus on estimates by age of the child, is important to note that grade progression is automatic in this context. That is, students progress from one grade to the next regardless of academic performance, until the completion of Grade 8. This policy on grade progression did not change during or after the pandemic. This potentially makes learning losses more consequential, as students lagging behind are not held back. Instead, students may need remedial education in all future grades.

[ix]Our measures of learning loss potentially combine an accelerated deterioration of previously acquired skills and an "opportunity cost" portion — i.e., skills which students would have learned ordinarily but did not due to the pandemic. This distinction between forgotten and foregone learning is prominent in simulations of COVID-19 learning losses (see, e.g., N. Angrist, de Barros, et al. (2021)) but is not crucial for understanding the *aggregate* effect of the pandemic on test scores, our principal object of interest.

[x]School enrollment in our sample is near-universal after 72 months of age — schooling is compulsory in India from 6–14 years — and rates of formal enrollment are unchanged between 2019 and 2021.

[xi]By imposing a common coefficient on age for all children, this regression specification is less general than Figure 2a or the estimates presented in Panel A of Table 2. In particular, it does not allow learning loss to differ by age or account for a change in the age gradient across rounds. Our learning loss and recovery results are robust to controlling for age fixed-effects, instead of age linearly (see Table A.3 in the Online Appendix).

[xii]To maximize coverage, we include all children over 55 months old at the time of the survey in Table 3. However, the results are similar if we focus only on children 55–95 months old at the time of the survey, which is the common support across all rounds (see Table A.5 in the Online Appendix).

[xiii] We adjust for age differences because test scores increase with age, and older children are more likely to attend ITK.

[xiv] Decisions on enrollment for the 2021-22 school year would have been taken in June-July 2021, substantially before the introduction of the program. The proportion of students enrolled in government or private schools does not differ across our different survey waves (Table 1).

[xv] The treatment effects of ITK are robust to controlling for age fixed-effects, instead of age linearly. See Table A.10 in the Online Appendix. They are also robust to restricting the sample to children 55–95 months old at the time of the survey, which is the common support across all rounds (see Table A.11 in the Online Appendix).

[xvi] In developing countries, Andrabi et al. (2011), Singh (2015) and Singh (2020a) studying school effects, Bau & Das (2020) studying teacher effects, and Muralidharan et al. (2019) studying the dose-response of (endogenously-chosen) usage of an after-school intervention, all find that value-added specifications yield similar estimates as those based on experimental variation, regression discontinuity, or dynamic student-level panel estimates. In the United States, Chetty et al. (2014) show similar reliability for teacher effects, as do J. D. Angrist et al. (2017); J. Angrist et al. (2021) and Deming (2014) for school effects.

[xvii] On nearly all measures of school and parental inputs and resources for remote learning, participants in ITK have access to fewer inputs. In contrast, on child activities, we see higher reported usage of educational TV programs and school books during school closures for ITK participants — this could represent a mechanism for impacts if children were encouraged by ITK volunteers to access these materials when schools closed due to the Omicron variant (after ITK introduction in many villages). We take a conservative view and attribute these differences to unobserved individual-specific propensity for education and examine if the treatment effects are substantially moderated by their inclusion.

[xviii] These controls include gender, maternal education, socioeconomic status, and lagged achievement in math and language.

[xix] We restrict ourselves to common items across multiple ages. These are, de facto, basic skills. More complex items (e.g., word problems requiring multiplication) were only administered to older students.

[xx] We are not aware of well-identified India-specific estimates of the causal effect of one hour of scheduled instructional time on student achievement, thus making Lavy (2015) the most appropriate benchmark for extrapolation.

[xxi] As an alternative benchmark, we estimate the impact of ITK in terms of learning-adjusted years of schooling (LAYS) (Filmer et al., 2020). Specifically, in our context, each additional year of age (and thus, each grade) increases test scores by 0.65σ in math and 0.36σ in Tamil. Therefore, the effect of ITK is equivalent to 0.25 additional years of schooling in math and Tamil. Using the Harmonized Learning Outcomes (HLO) Database, we estimate one year of schooling learning in India is equivalent to 0.6 years in a high-performing education system. Thus, our treatment effects are equivalent to

0.3 LAYS (assuming a treatment effect that persists throughout the school year) or a gain of almost 4 LAYS per 100 USD spent, which compare favorably to other interventions (N. Angrist, Djankov, et al., 2021).

[xxii]Annual program costs were $\sim$2% of the per-student spending in the public school system in Tamil Nadu (which is estimated at $\sim$USD 350 per-child (CBGA, 2018)), but delivered learning gains of over 30% relative to the recovery in the absence of the program.

[xxiii]Many other low- and middle-income countries also feature low economic opportunities for educated women in rural areas (see, e.g., Andrabi et al. (2013)).