School Closures and Learning Loss during the Pandemic: Evidence from Rural India*

Muchin Bazan†, Shatakshee Dhongde‡, Abu S. Shonchoy§

Abstract

Large-scale interruptions, such as the pandemic, may appear as "equalizing agents" not because they provide an impetus to the laggards but often because they pull back the leaders. We provide an illustration by analyzing the effect of school closures during the COVID-19 pandemic on primary-school learning in rural India. Using multiple waves of the Annual Status of Education Report (ASER) for children aged 5–12 and a difference-in-differences design comparing historically high- versus low-performing states, we show that the pre-COVID cross-state learning gap of 0.43 standard deviations (SD) narrowed by 0.12 SD in math and 0.13 SD in reading after school closures, driven by larger losses in historically high-performing states. Exploiting cross-state variation in the duration of school closures, we find that extended closures (12 months or greater) disproportionately reduced scores in high-performing states. We find little systematic heterogeneity by village- or household-level digital access, parental education, or gender. Thus, crisis-era "equalization" reflects declines at the top rather than catch-up at the bottom, underscoring the need for targeted remediation that protects prior gains while accelerating recovery in lower-performing settings.

Keywords: COVID-19, school closure, learning loss, rural India.

JEL Classification codes: I24, I25, I28.

*We would like to thank conference participants at the Southeast Exchange of Development Studies (SEEDS), and the Southern Economic Association (SEA) for their thoughtful comments.

†Department of Industrial Engineering, Pontificia Universidad Católica de Chile and Universidad de Piura, GLO Affiliate. Av. Vicuña Mackenna 4860, Macul, Región Metropolitana. Email: muchin.bazan@uc.cl.

‡Corresponding author. School of Economics. Georgia Institute of Technology. Old C.E. Building, 221 Bobby Dodd Way, Atlanta, GA 30332. Email: shatakshee.dhongde@econ.gatech.edu.

§The Department of Economics, Florida International University. 11101 S.W. 13 ST., Miami, FL 33199. Email: shonchoy@fiu.edu.

1 Introduction

It is well-documented that foundational skill deficits have persistent consequences for productivity, growth, and inter-generational mobility, as well as for children's subsequent schooling and labor market opportunities (ASER, 2024). Despite sustained gains in enrollment, foundational learning remains low in many low- and middle-income settings, including India (Chatterjee and Poddar, 2021; Das and Zajonc, 2010; Muralidharan et al., 2019). Although there are well-established components contributing to the foundation skill deficit, one less researched contributing factor is system-wide interruptions during crisis times, such as pandemics. During crisis moments, such an interruption can compress pre-existing performance gaps — not because lagging systems catch up, but because leading systems lose ground. We study this leveling-down dynamics in primary-school learning in rural India. When COVID-19 precipitated prolonged school closures and a rapid shift from in-person to remote modalities, many education experts expressed concern that learning would decline and that inequalities might widen, particularly where access to digital technologies and home learning support was inadequate. In this paper, we examine how the pandemic and subsequent school closures affected math and reading outcomes in rural India and whether the apparent post-pandemic "equalization" reflects meaningful catch-up at the bottom or disproportionate losses at the top.

India provides an important context to study this question. In March 2020, approximately 1.5 million schools were closed nationwide, and many remained closed—fully or partially—through mid-2022.[1] At the same time, nationally representative assessments have repeatedly documented high enrollment among children aged 6–14 (exceeding 95%) but low proficiency in basic competencies; for example, more than half of children struggle with simple division typically expected by grades 3–4 (ASER, 2024). Against this backdrop, we ask whether the pandemic narrowed cross-state achievement gaps by eroding prior advantages

[1] UNICEF press release: https://www.unicef.org/india/press-releases/covid-19-schools-more-168-million-children-globally-have-been-completely-closed

in historically high-performing (HP) states rather than by elevating outcomes in historically low-performing (LP) ones.

India was also one of the few countries where data on learning outcomes was collected as early as 2022 which would help measure the impact of the pandemic on children's education. We compile data from nationally representative household-based assessments from the Annual Status of Education Report (ASER) for 2018 (pre-pandemic) and 2022 (the first public round after reopening). Our data cover 217,070 households and 261,778 children aged 5–12 enrolled in rural public primary schools, with rich information on household, village, and district characteristics. ASER tests are administered at home using comparable instruments across years, and across villages and follow a rotating panel design.[2][3] For comparability, we focus on districts surveyed in 2016, 2018, and 2022, which together represent roughly 95% of India's population (see Table A1). We classify states as historically high- or low-performing using pre-pandemic ASER math and reading scores (see Figure 1). Prior to COVID-19, the average standardized test score gap between these groups was about 0.43 standard deviation (SD) points.

Our empirical strategy uses a difference-in-differences (DiD) design that compares changes in outcomes from 2018 to 2022 in historically HP and LP states, controlling for individual and household characteristics and including district fixed effects. The identifying assumption is that, absent the pandemic, trends would have been parallel across these two sets of states—a condition we examine using pre-pandemic waves (2016–2018). To probe mechanisms, we exploit cross-state variation in the duration of school closures using the Oxford COVID-19 Government Response Tracker (OxCGRT) Stringency Index and its school-closing policy sub-component to compute effective months of closure from March 2020

[2] Annual Status of Education Report (ASER) Project: https://asercentre.org/.

[3] ASER uses a rotating panel of villages. Each year, one-third of sampled villages within a district are newly introduced.

to August 2022, aligned with each state's academic calendar.[4,5]

We document three main findings. *First*, the cross-state achievement gap narrowed after the pandemic, mostly driven by historically HP states experiencing larger losses: relative to LP states, children in high-performing states saw declines of 0.126 SD in math and 0.134 SD in reading. In effect, a sizable share of the pre-COVID advantage was eroded, consistent with an "equalization by leveling down" rather than bottom-end catch-up. *Second*, a longer duration of closures was strongly associated with larger losses in high-performing states. HP states averaged roughly 13 months of closure compared to about 12 months in LP states (see Table 3); cross-sectional estimates from 2022 indicate that each additional month of closure was associated with declines of approximately 0.06 SD in math and 0.03 SD in reading in HP states (both significant at the 1% level), and DiD estimates show 0.15–0.20 SD larger losses in states above the median closure duration (12 months). *Third*, we find little systematic heterogeneity by village- or household-level digital access, parental education, or gender, suggesting that closure duration—interacting with states' pre-pandemic performance levels—was the dominant driver of the observed compression.

Our paper adds to the growing literature studying the negative impacts of the COVID-19 pandemic, including economic (Rodríguez-Planas, 2022, Deb et al., 2022, Mottaleb et al., 2020), social (Abrams, 2021, Leslie and Wilson, 2020), and physiological impacts (Engzell et al., 2021), along with the impact of school closure on education and health (Lipkin and Crepeau-Hobson, 2022; Guariso and Nyqvist, 2023, Liao et al., 2022, Grewenig et al., 2021, Auger et al., 2020, Hawrilenko et al., 2021).

To our knowledge, this paper is the first to use household survey data and a representative sample of rural Indian states to examine the effect of school closures due to COVID-19 on learning losses in an LMIC setting while focusing on state inequalities and performance

[4] The Oxford Stringency Index summarizes government responses to COVID-19, including school closures, travel restrictions, and stay-at-home orders, on a 0–100 scale, with higher values indicating stricter measures.

[5] Oxford COVID-19 Government Response Tracker (OxCGRT) Stringency Index: https://ourworldindata.org/covid-stringency-index. We use state-level information on school-closure policies to calculate the number of months schools were effectively closed; see Figure A1.

measures. In the context of rural India, some states might be consistently high-performing in math and reading, even in the absence of a COVID-19 pandemic; therefore, exploring how HP and LP states responded to school closures and the consequences for students' learning is key. In addition, whereas most LMICs did not collect survey data and test scores during the pandemic, our analysis was made possible because ASER India administered a household survey right after schools reopened in September 2022, which allows us to compare test scores trends before and after 2020. Finally, our paper not only evaluates the effect of state performance and the COVID-19 pandemic on education outcomes but also speaks to the duration of school shutdowns. Using rich data on village and household characteristics, we also explore heterogeneous effects by parental education, gender, and access to digital media.

The paper proceeds as follows. Section 2 describes the data we used in our analysis. Section 3 explains our empirical framework to examine the impact of state performing and school closures on main outcomes. We interpret our main findings in Section 4, and concluding remarks are provided in Section 5.

2 Data

2.1 ASER Data

We use data from the Annual Status of Education Report (ASER), a nationwide household survey that captures the status of children's enrollment and learning outcomes in rural India. From 2016, nationwide ASER is conducted every other year. The report was not published in 2020, due to the COVID-19 pandemic and resulting school closures. In 2022, ASER was conducted across the country four years after 2018. The data include information on individuals between 3-16 years old. For our analysis we form a panel of Indian districts that are present in all years between 2016 and 2022. Our sample contains information for ASER on 28 states and on 533 districts reflecting a nationwide sample of rural districts. We restrict our sample to age appropriate children in public, primary schools.[6]

Children in the age group 5-16 are tested in basic reading and basic math. The same test is administered to all children, regardless of enrollment status, age or grade. The highest level of the reading test is a Grade 2-level text. In math, the highest level is a 3-digit by 1-digit division problem, usually taught in Grade 3 or Grade 4. Raw math and reading scores are measured using a scale from 1 (the lowest) to 5 (the highest). For our analysis we also construct z math (reading) scores for 2022 using 2018 as a baseline year by subtracting from the raw math (reading) scores the average 2018 math (reading) score by age and divided it by the 2018 standard deviation. Our main outcome variables of interest are raw scores and z-scores for math and reading. In addition, ASER household survey contains information on household size, child demographic information (gender, age), parental education, some information on household assets, and village characteristics. We use these variables as control variables in our estimations.[7]

Figures 2 shows z math and reading test scores in 2022 using 2018 as benchmark. Test scores declined in 2022 compared to 2018 levels. We believe that this decline in test scores can be attributed to the COVID-19 pandemic and the resulting lockdown restrictions. Table 1 and 2 show descriptive statistics for some key variables. In Table 1 average values are reported by year while Table 2 shows descriptive statistics by treatment status before and after the COVID-19 pandemic. The average age for our sample of children is 8.41 in 2022. 50% of the sample is female, 9% of households have a person who can use a computer, 76% of children reported having a father who attended school, 34% of children have an internet cafe in their village, and 93% of households have electricity connection. In 2022, there is a higher proportion of villages with internet cafe, and households with electricity connectivity compared to pre-COVID years. There are some imbalances across years, mostly occurring

[6] For instance, for class grade 1, we keep in our sample students older than 4; for class grade 2, we keep in our sample students older than 5, etc.

[7] Control variables include father education, mother education, public school indicator, indicator for test in Hindi, internet cafe in the village, electricity in the village, computer use at home, tuition cost, electricity connectivity at home, child's gender, child's class and child's age.

due to demographic variations across states. There are also some differences in household and village characteristics between treatment and control states. Table 2 shows difference in means between control (low performing) and treatment (high performing) states in 2018 and 2022. We control for demographic, household characteristics, and location-related variables in all our regression models to address imbalances between treatment and control states.

2.2 State Performance

We investigate how the COVID-19 pandemic affected high performing states and low performing states differently, specially looking at education outcomes (i.e. test scores). To construct a measure of state performance, we use the 2014, 2016, and 2018 waves of ASER Household survey. We compute average combined scores (raw math and reading together) between 2014, 2016 and 2018 for each state in our sample. Then, we rank our states and classify them in two groups based on sample median. High performing states are those with average combined score above the sample median of 5.94 out of 10. Low performing states are those with average combined score below or equal to the sample median of 5.94 out of 10 (see Figure 1). We exploit this variation in state performance and evaluate the impact of COVID-19 restrictions and the duration of school closures on students' learning.

2.3 Stringency Index and School Closure

The stringency index is calculated by the Oxford Coronavirus Government Response Tracker (OxCGRT). It is a composite measure of nine of the response metrics: school closures; workplace closures; cancellation of public events; restrictions on public gatherings; closures of public transport; stay-at-home requirements; public information campaigns; restrictions on internal movements; and international travel controls. The index on any given day is calculated as the mean score of the nine metrics, each taking a value between 0 and 100. A higher score indicates a stricter response.

To construct a measure of school closure by Indian state we use the school closing indicator, a component of the stringency index. The school closing indicator has daily values from 0 to 3. 0 indicates no measures, 1 indicates a recommendation to close schools or open schools with modified operations, 2 indicates a requirement to close some schools, and 3 requires closing schools at all levels. We construct a measure of school closures between March 2020 and August 2022 by counting the number of effective months schools were closed (had an indicator 3) in a given state during this period of time. We only counted months if schools were closed and if those months correspond to school calendar months for states.[8] Table 4 shows academic calendar months for all the states in our sample, the months schools were closed between March 2020 and August 2022 due to the COVID-19 pandemic, and the number of effective months of school closures by Indian state. There is variation in the number of months schools closed in both high performing and low performing states. With high performing states closing on average 13 months and low performing states closing on average 12 months.

3 Empirical Framework

3.1 Linear Regression, cross section analysis: 2022

To provide evidence that prolonged school closures led to a reduction in test scores among children, we estimate a linear regression model where our main regressor is the number of months of school closures in each state. The equation is as follows:

$$Y_{isd} = \alpha + \beta_1 number\_months\_sc_s + \beta_2 X_{isd} + \phi_d + \varepsilon_{isd} \tag{1}$$

Where Y_{isd} is the outcome of interest for individual i residing in state s and district d in 2022 , $number\_months\_sc_s$ is our continuous treatment and represents the number

[8]If all schools were closed in a school calendar month for that state, we counted it. We did not count months that were outside the school calendar period even if schools were closed in those months.

of months schools were closed between March 2020 and August 2022. ϕ_d are district fixed effects to account for time invariant district characteristics that can affect the quality of education provided to students and their performance, and X_{isd} is a vector of individual-level characteristics. Finally, in all our estimations standard errors are clustered at the district level.

We also show that our results are robust to other clustering methods such as cluster at the state level or bootstrapping. The coefficient of interest in Equation (1) is β_1, which represents the impact of one additional month of school closure on test scores among children between 5 to 12 years old.

Alternatively, we evaluate the effect of residing in states where schools closed for a longer period of time on test scores with a binary treatment. We estimate the following equation:

$$Y_{isd} = \gamma_0 + \gamma_1 above\_12\_months\_sch_s + \gamma_2 X_{isd} + \phi_d + \varepsilon_{isd} \tag{2}$$

Where all the terms are defined previously and $above\_12\_months\_sch_s$ is an indicator equal to one if the individual lives in a state where schools closed more than 12 months (sample median), and zero otherwise. The coefficient of interest in Equation (2) is γ_1, which represents the effect of residing in states where schools closed for longer period of time on test scores.

In addition, we are interested in evaluating the learning gap between high and low performing states and how the gap evolved due to the COVID-19 pandemic. We estimate the following equation:

$$Y_{isd} = \delta_0 + \delta_1 high\_state\_performance_s + \delta_2 X_{isd} + \delta_d + \varepsilon_{isd} \tag{3}$$

Where all the terms are defined previously and $high\_state\_performance_s$ is an indicator equal to one for high performing states and zero, otherwise. The coefficient of interest in Equation (3) is δ_1, which represents the effect of residing in high performing states on test scores.

Longer school closures can affect test scores more than shorter school closures and have different impacts for high versus low performing states. To investigate this, we estimate heterogeneous effects of school closures for high versus low performing states:

$$\begin{gathered} Y_{isd} = \alpha_0 + \alpha_1 high\_state\_performance_s + \alpha_2 above\_12\_months\_sch_s \\ +\alpha_3 above\_12\_months\_sch_s * high\_state\_performance_s + \alpha_4 X_{isd} + \delta_d + \varepsilon_{isd} \end{gathered} \tag{4}$$

α_3 is the coefficient of interest and represents the heterogeneous impact of the duration of school closures for high compared to low performing states in 2022.

3.2 Differences-in-Differences Method

We examine how high and low performing states faced COVID-19 restrictions and their implications on test scores. Our analysis focuses on differences in children' test scores trends across years before and after the COVID-19 pandemic and across high versus low performing states to identify the causal effect of COVID-19 restrictions on educational outcomes. We estimate the following differences-in-differences equation:

$$\begin{gathered} Y_{isdt} = \alpha_1 + \alpha_2 high\_state\_performance_s * year2022 \\ +\alpha_3 high\_state\_performance_s + \alpha_4 year2022 + \beta X_{isdt} + \phi_d + \varepsilon_{isdt} \end{gathered} \tag{5}$$

Where Y_{isdt} is an outcome of interest for individual i residing in state s, district d and year t. $high\_state\_performance_s$ is an indicator equal to one if the individual lives in a high performing state and zero if the individual lives in a low performing state. $year2022$ is an indicator equal to one if the children took the test in 2022 and zero if they took the test in 2018. ϕ_d are district fixed effects. X_{isdt} is a vector of individual-level characteristics, including child's age, child's gender, parental education, economic condition, household and village characteristics. ε_{isdt} is a random disturbance term clustered at the district level.

The district fixed effect ϕ_d allows us to control for time-invariant characteristics specific to each district, such as socio-economic conditions, access to resources, education infrastructure, that may affect students' learning.

The coefficient of interest in Equation (5) is α_2, which captures the effect of the COVID-19 pandemic on test scores for high performing states relatively to low performing states. Our outcomes of interest include raw and z test scores.

Furthermore, to assess the duration of school closures on test scores and its impact for high versus low performing states in a DID design, we re-estimate Equation (5) with slight adjustments:

$$\begin{aligned} Y_{isdt} = \alpha_0 + \alpha_1 high\_state\_performance_s + \alpha_2 above\_12\_months\_sch_s \\ + \alpha_3 above\_12\_months\_sch_s * high\_state\_performance_s + \alpha_4 X_{isdt} + \delta_d + \varepsilon_{isdt} \end{aligned} \tag{6}$$

Equation (6) estimates the effect of school closures for high performing states compared to low performing states in 2018 and 2022. The DID estimate is α_3 which captures the heterogeneous impact of the duration of school closures on students' learning.

4 Results

4.1 Effect of State Performance on Education Outcomes

4.1.1 Linear Regression Model: Only 2022

We estimate a linear regression model to illustrate the impact of state performance on students' learning. Table 5 shows the OLS estimates. Both math and reading test scores declined and OLS estimates are statistically significant at 1% level. On average students who lived in high performing states scored 0.358 standard deviations lower in math and 0.449 standard deviations lower in reading compared to students who lived in low performing states. Similarly, students living in high performing states scored 0.328 points lower in math and 0.521 points lower in reading compared to their peers in low performing states. This is equivalent to a 13% decline in math and 20% decline in reading ASER raw test scores. We attribute the decline in test scores of students in high performing states to longer school closures. As mentioned before, schools remained closed for longer periods of time between March 2020 and August 2022 in high performing states more than in low performing states (see Table 3 and Table 4). This could potentially explain why students in high performing states were more affected than their counterparts. Table 6 presents the OLS estimates using as regressor the number of months of school closures which varies by state between March 2020 and August 2022. As expected, one additional month of school closure leads to a 0.07 point decline in raw math and a 0.04 point decline in raw reading test scores (significant at 1% level). Changing the measure of school closures from continuous to a binary indicator produces similar findings. For instance, Table 7 shows that students living in states where schools closed for more than a year (12 months) scored 0.70 points less in math and 0.41 points less in reading than students living in less impacted states. The OLS coefficients are large in size and statistically significant at 1%. Table 8 evaluates how both school closures and state performance affected test scores. The interaction term is negative and statistically significant at 1%. While on average the impact of COVID-19 school closures is positive and significant, high performing states exposed to longer school closures performed worse than high performing state exposed to shorter school closures. This result goes in line with our previous finding that the longer duration of school closures in some high performing states affected students' performance in math and reading.

4.1.2 Differences-in-differences Estimates

We now quantify the effect of COVID-19 pandemic using a differences-in-differences estimation, as outlined in Equation (5). Table 10 presents the effect of COVID-19 restrictions on our education outcomes defined before controlling for covariates.

We focus on the impact of COVID-19 pandemic on z test scores reported in columns 1 and 2, and on raw test scores reported in columns 3 and 4. We find that in high performing states, learning outcomes declined due to COVID-19 pandemic in favor of low performing states. The interaction term is negative and statistically significant. The decline on z test scores is about 0.12-0.13 standard deviations and equivalent to a 80% decline from baseline levels. Figure 4 shows the DID coefficients for our main outcome variables. All of them are negative and statistically significant at 1% level. Notice that the effect is stronger for the reading test component of the ASER survey.

The identification strategy relies on the assumption that test scores evolve similarly in high performing and low performing states before the pandemic and any change in 2022 test scores should be attributed uniquely to the COVID-19 state restrictions. Table 9 evaluates whether the parallel trend condition holds by re-estimating equation (5) using the alternative period of 2016-2018. Since this period corresponds to pre-COVID years, we expect no to find any significant effects on test scores. The interaction term is no statistically significant, which confirms that test scores in high-performing and low-performing states were very similar in 2016 and 2018 (see Figure 3). Figure 5 depicts a convergence graph and plots z math and z reading test scores across years and for low and high performing states. Before 2020, test scores in high and low performing states followed a similar path reflecting an upward trend during 2014-2018. However, test scores drastically declined after 2020 in high performing states. In contrast, math test scores slightly increased and reading test scores slightly declined in low performing states. Table A4 also shows a simple difference-in-difference calculation of the main outcomes for 2018 and 2016 cohorts and for high versus low performing states. The simple difference-in-difference calculation is not statistically significant at 5% level and test scores trends had similar paths before the COVID-19 pandemic.

Figures 6 and 7 show kernel density functions of z test scores for all the states and separately for low and high performing states. From a visual inspection, there are more children with negative z scores in 2022 than in 2018. The effect is more pronounced in reading than in math. Separating the sample by state performance gives similar conclusions. While students scored lower in 2022 compared to 2018, children' performance in reading declined in 2022 compared to 2018 in high performing states more than in low performing states.

Then, we evaluate the effect of longer school closures and its impact for high and low performing states. The effects are very similar to the cross section results. Longer school closures reduced test scores in high performing states in favor of low performing states. The DID estimate is roughly 0.20 points out of 5 and represents a 7-8% change from baseline (see columns 3 and 4 of Table 11).

4.2 Robustness Checks

We conduct several robustness tests to assess the internal validity of our results. First, we explore alternative ways to construct our binary treatment. The main analysis uses a binary indicator equal to one for high performing states and 0, otherwise. We use the continuous treatment, the average state combined score between 2014, 2016, and 2018 and the estimates remain consistent. States with higher historically combined scores were more affected in the sense that students living in those states scored less in math and reading in 2022 compared to students in low performing states. In addition, we changed the sample of pre-COVID years and excluded the year 2018 to construct our binary treatment based on sample means, and the results remain robust to different time periods choices. Secondly, instead of number of months of school closures we calculate the number of days of school closures between March 2020 and August 2022 by Indian state. The results remain robust to daily measures of school closures. Finally, we rerun the regressions with different fixed effect structures for example state fixed effects instead of district fixed effects. In addition, we cluster standard errors at the state rather than at the district level. The estimates remain consistent across different specifications.

4.3 Subgroup Analysis and Heterogeneous Effects

To what extent the impact of the COVID-19 state restrictions vary across states and individuals? To answer this, we estimated separated treatment effects based on child's gender, parental education, assets in the household, and access to digital technology at home and at the village.[9] We separately estimated our linear regression and DID model for each group. Table 12 shows the effect of COVID-19 restrictions for high performing states relatively to low performing states and by group category. The COVID-19 pandemic affected female students' learning in high performing states more than male students' learning. Female students' z test scores declined by 0.14 standard deviations (significant at 1% level) in high performing states compared to low performing states. While male students' test scores also declined, this decline was lower in size. Household assets can also mitigate the impact of COVID-19 pandemic on learning outcomes. Panel C and D of Table 12 show the DID estimates for students in households with TV access and without TV access, respectively. It seems that having TV access at home mitigated the impact of the COVID-19 pandemic for math test scores. In contrast, reading test scores declined more in households with TV access. We believe that watching TV can be a substitute for reading textbooks. The presence of TV at home can increase the time children spend in front of the TV and reduce the time allocated to reading. In addition, while there is no a clear explanation of the positive relationship between TV access at home and math test scores, some educational television programs, like those focused on math concepts, may have a positive impact on learning. The Indian government developed educational television programming during school closures, therefore, having a TV at home could have mitigated the decline in math test scores. Digital access at the village can also benefit students' learning in times of school closures. In our study, digital access is a combination of internet cafe and electricity availability at the village. Students living in villages with digital access were less affected by the COVID-19 restrictions

[9] Digital access at the village is an indicator equal to one if there is internet cafe in the village and electricity connectivity, and zero otherwise. Digital access at home is an indicator equal to one if there is a smartphone or computer , electricity and internet connectivity at home, and zero otherwise.

(see Panel E and F of Table 12). From our subgroup analysis, we also find positive effects of parental education. Students with at least one parent educated (who attended school) presented a lower decline in test scores compared to students without educated parents. For instance, the DID estimate on z math test scores is -0.13 sd versus -0.15 sd. More educated parents might prioritize the education of their children, and help them in their education during school closures.

We also investigated the heterogeneous effects of COVID-19 state restrictions in Tables 13 to 15. While most of the interaction terms are no statistically significant and close to zero, female students' math performance was more sensitive to longer school closures. The coefficient of the interaction term is -0.04 and statistically significant at 5%. The interaction term suggests a heterogeneous significant effect based on gender.

5 Conclusions

India was one of the countries that closed schools for a longer period of time and implemented a strict school closure policy due to the COVID-19 pandemic. Strict COVID-19 restrictions and school closures helped prevent the rapid spread of the virus and significantly slowed down the rate of new infections. However, the COVID-19 pandemic led to a significant learning loss in math and reading subjects in rural states. This decline was not in the same magnitude for all rural states. States that were on average doing better before the pandemic (high performing states) were affected to a greater extent by school closures and statewide restrictions.

In this paper, we studied the impact of the COVID-19 pandemic, state performance, and the duration of school closures on education outcomes. Our linear regression model and differences-in-differences estimates documented a significant decrease in math and reading test scores in high performing states compared to low performing states. A longer duration of school closures was associated with greater learning loss in high performing states. We tested our models against multiple robustness checks and showed our findings to be consistent.

The paper contributes to the critical debate on the implementation of public policies such as school closures during health emergencies and their effect on education outcomes in developing countries, as well as the need to balance stringent restrictions while maintaining access to educational services. We believe our findings will encourage public institutions, school administrators, and policymakers to design alternative ways to maintain public health than closing schools for a long durations.

References

Abrams, David S. (2021) "COVID and crime: An early empirical look," *Journal of Public Economics*, 194.

ASER, Center (2024) "Annual Status of Education Report (Rural) 2023: beyond basics."

Auger, Katherine A., Samir S. Shah, Troy Richardson et al. (2020) "Association Between Statewide School Closure and COVID-19 Incidence and Mortality in the US," *JAMA*, 324 (9), 859–870.

Chatterjee, S. and P. Poddar (2021) "From safe motherhood to cognitive ability: exploring intrahousehold and intergenerational spillovers," *Economica*, 88 (352), 1075–1106.

Das, J. and T. Zajonc (2010) "India shining and Bharat drowning: comparing two Indian states to the worldwide distribution in mathematics achievement," *J Dev Econ*, 92 (2), 175–187.

Deb, Pragyan, Davide Furceri, Jonathan D. Ostry, and Nour Tawk (2022) "The Economic Effects of COVID-19 Containment Measures," *Open Economies Review*, 33, 1–32.

Engzell, Per, Arun Frey, and Mark D. Verhagen (2021) "Learning loss due to school closures during the COVID-19 pandemic," *Proceedings of the National Academy of Sciences*, 118 (17).

Grewenig, Elisabeth, Philipp Lergetporer, Katharina Werner, Ludger Woessmann, and Larissa Zierow (2021) "COVID-19 and educational inequality: How school closures affect low- and high-achieving students," *European Economic Review*, 140.

Guariso, A. and M. Björkman Nyqvist (2023) "The impact of the COVID-19 pandemic on children's learning and wellbeing: Evidence from India," *Journal of Development Economics*, 164, 1–10.

Hawrilenko, Matt, Emily Kroshus, Pooja Tandon, and Dimitri Christakis (2021) "The Association Between School Closures and Child Mental Health During COVID-19," *JAMA*, 4 (9).

Leslie, E. and R. Wilson (2020) "Sheltering in Place and Domestic Violence: Evidence from Calls for Service during COVID-19," *Journal of public economics*, 189.

Liao, Haoye, Sen Ma, and Hao Xue (2022) "Does school shutdown increase inequality in academic performance? Evidence from COVID-19 pandemic in China," *China Economic Review*, 75.

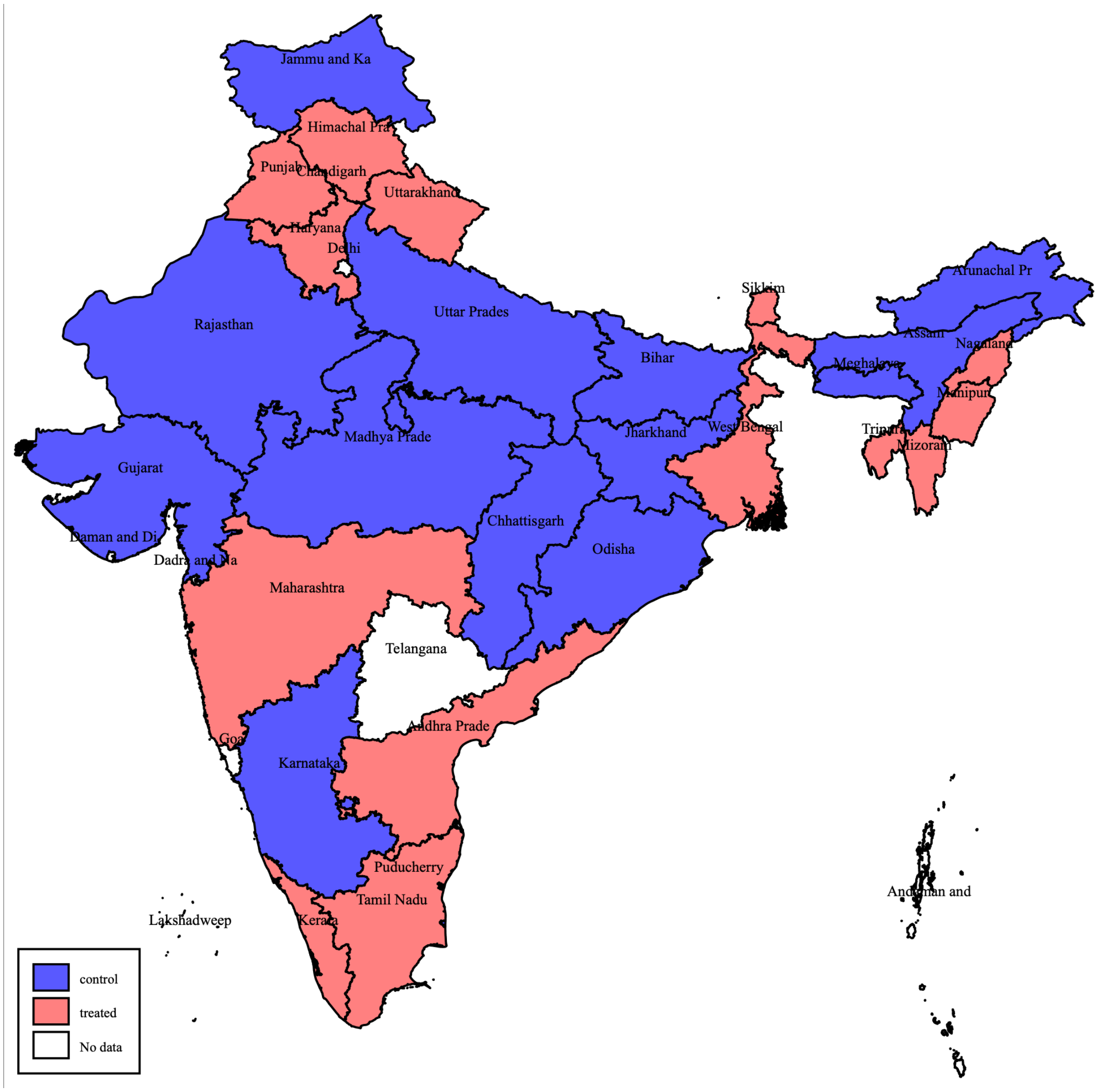
Lipkin, M. and F. Crepeau-Hobson (2022) "The impact of the COVID-19 school closures on families with children with disabilities: A qualitative analysis," *Psychol Sch*, 1–16.

Mottaleb, Khondoker Abdul, Mohammed Mainuddin, and Tetsushi Sonobe (2020) "COVID-19 induced economic loss and ensuring food security for vulnerable groups: Policy implications from Bangladesh," *PLoS ONE*, 15 (10).

Muralidharan, K., A. Singh, and AJ. Ganimian (2019) "Disrupting education? Experimental evidence on technology-aided instruction in India," *American Economic Review*, 109 (4), 1426–1460.

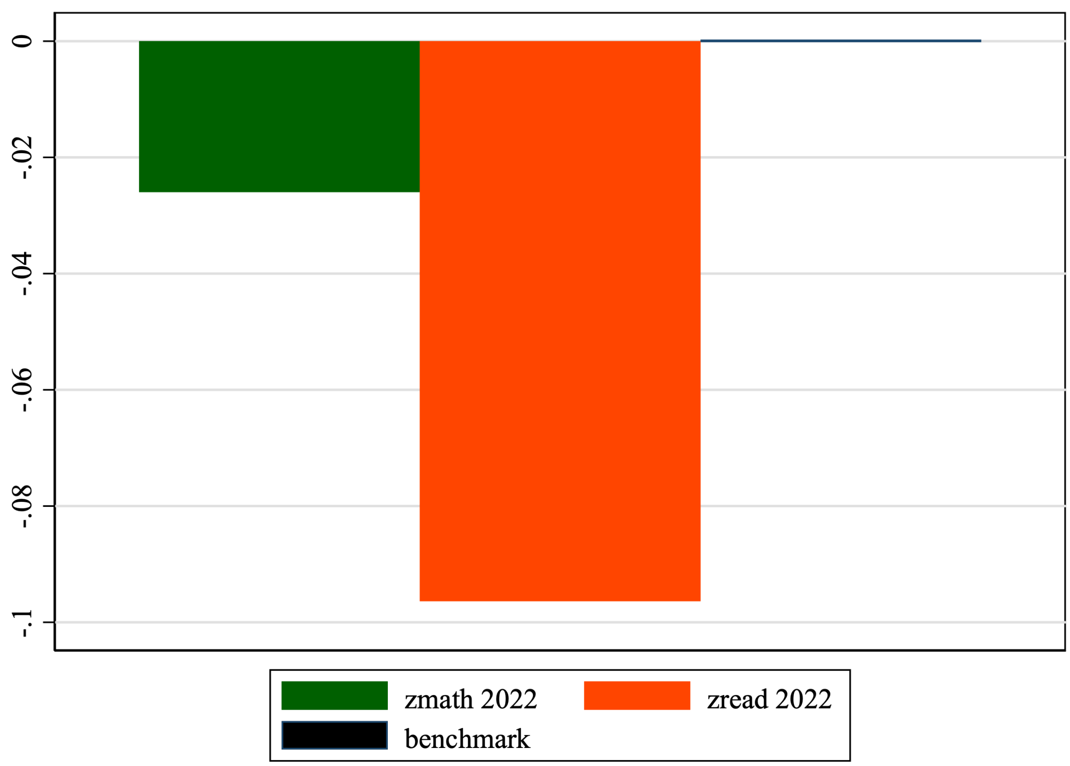
Rodríguez-Planas, Nuria (2022) "COVID-19, college academic performance, and the flexible grading policy: A longitudinal analysis," *Journal of Public Economics*, 207.

Figure 1: Low and High Performing States



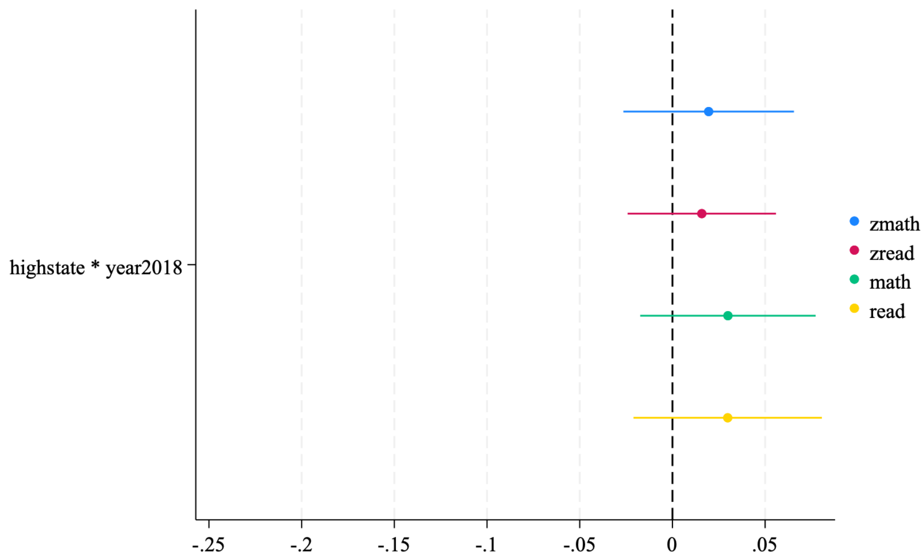
This figure shows low and high performing states constructed based on ASER average state combined raw score (math and reading) between 2014-2018 for individuals (5-12 years old) in primary public schools in rural India. A high performing state is defined as states with average combined raw (math and reading) test scores above the sample median of 5.94 between 2014, 2016, and 2018. We include states from which we have both ASER test scores and COVID-19 school closures (in number of months). Data on Telangana state are not included since we included a sample of panel of districts. Telangana state was formed in 2014 following the Act of Indian Parliament that split the state of Andhra Pradesh into Telangana and the residuary Andhra Pradesh state.

Figure 2: z (baseline 2018) Math and Reading Test Scores



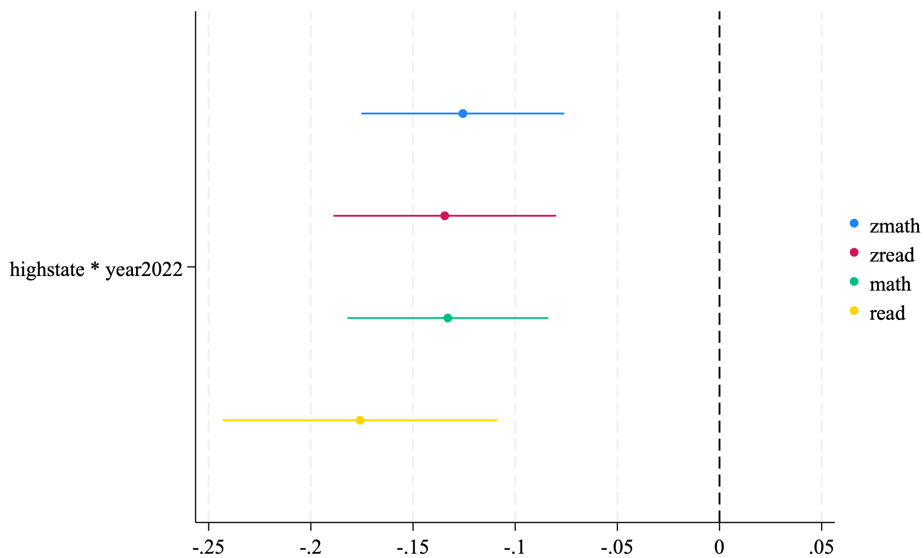
This figure shows the average math and reading z scores (baseline 2018) for individuals (5-12 years old) in primary public schools (grade 1 to 5). The data come from Indian ASER Project. Total number of observations is 157,348 (for reading scores) and 157,218 (for math scores) in 2022.

Figure 3: Primary Public Schools: Parallel trends 2016-2018



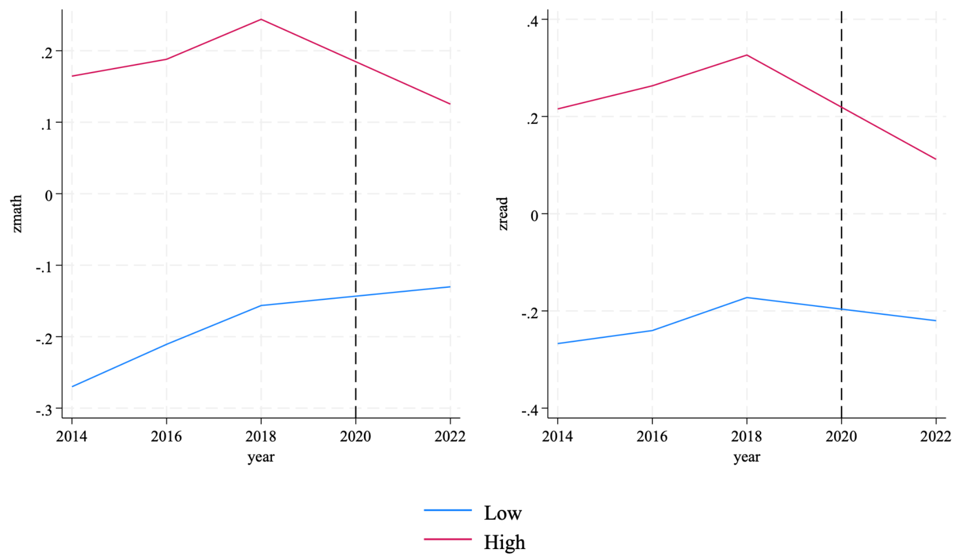
This figure shows the coefficient of the DID interaction term of a placebo test using DID method for 2016 and 2018. DID coefficients for high versus low performance states on test scores for the sample of primary public school going children (Grade 1 to 5) between 5 to 12 years old. The treatment group is a binary indicator equals to one for high performing states and zero for low performing states. A high performing state is defined as states with average combined raw (math and reading) test scores above the sample median of 5.94 between 2014, 2016, and 2018; and zero otherwise. The solid dots show the means and 95% confidence intervals are denoted by vertical capped bars.

Figure 4: Primary Public Schools: DID 2018-2022



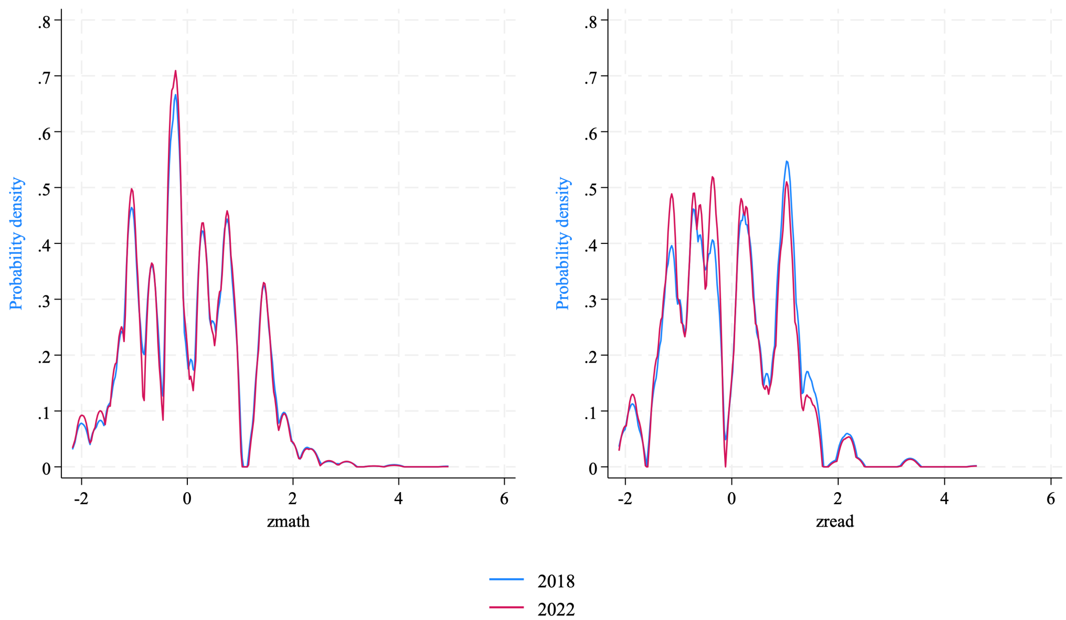
This figure shows the DID coefficients of the COVID-19 pandemic for high versus low performance states on test scores for the sample of primary public school going children (Grade 1 to 5) between 5 to 12 years old. The treatment group is a binary indicator equals to one for high performing states and zero for low performing states. A high performing state is defined as states with average combined raw (math and reading) test scores above the sample median of 5.94 between 2014, 2016, and 2018; and zero otherwise. The solid dots show the means and 95% confidence intervals are denoted by vertical capped bars.

Figure 5: DID Convergence Graph: Test Scores by State Performance Treatment



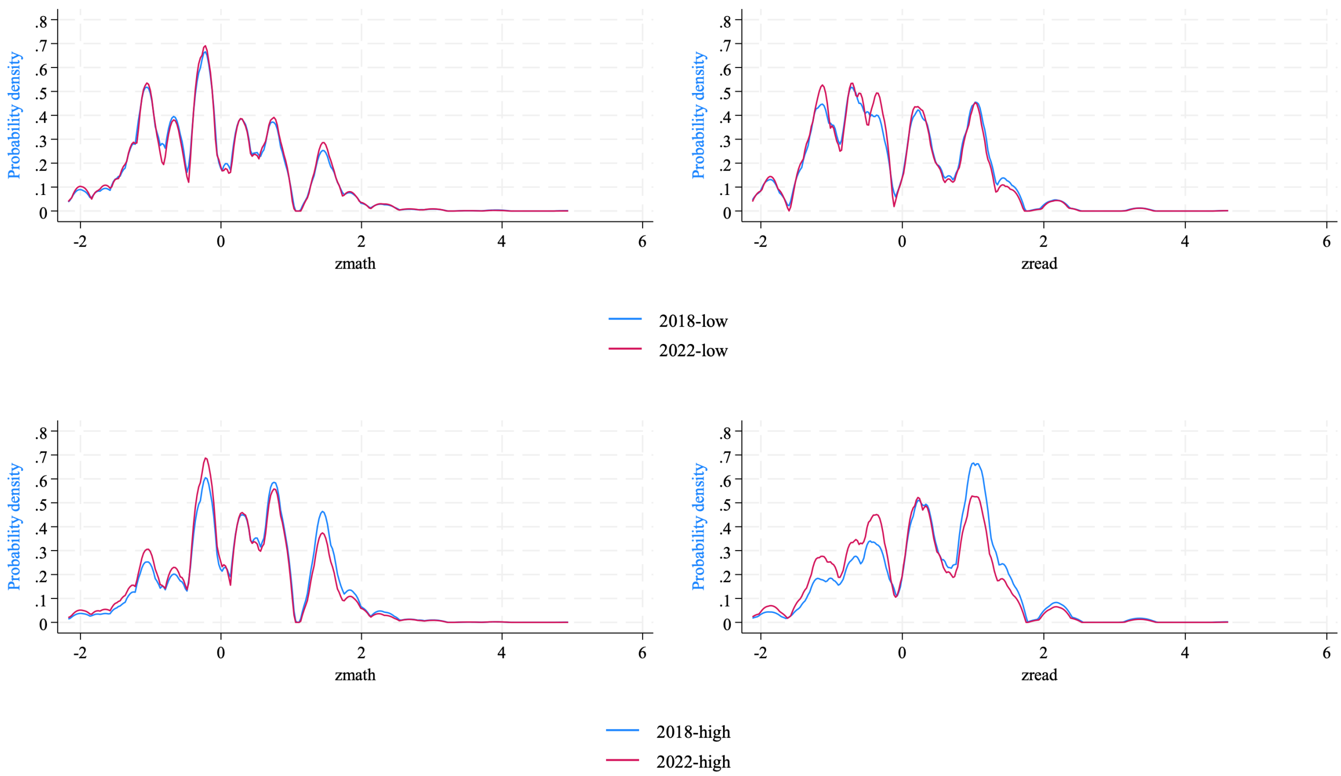
This figure shows average test scores by year and Treatment status (high versus low performing states). The sample consists of primary public school going children (Grade 1 to 5) between 5 to 12 years old. The treatment group is a binary indicator equals to one for high performing states and zero for low performing states. A high performing state is defined as states with average combined raw (math and reading) test scores above the sample median of 5.94 between 2014, 2016, and 2018; and zero otherwise. Data is not available for 2020 year.

Figure 6: Kernel Graph: Test Scores by year



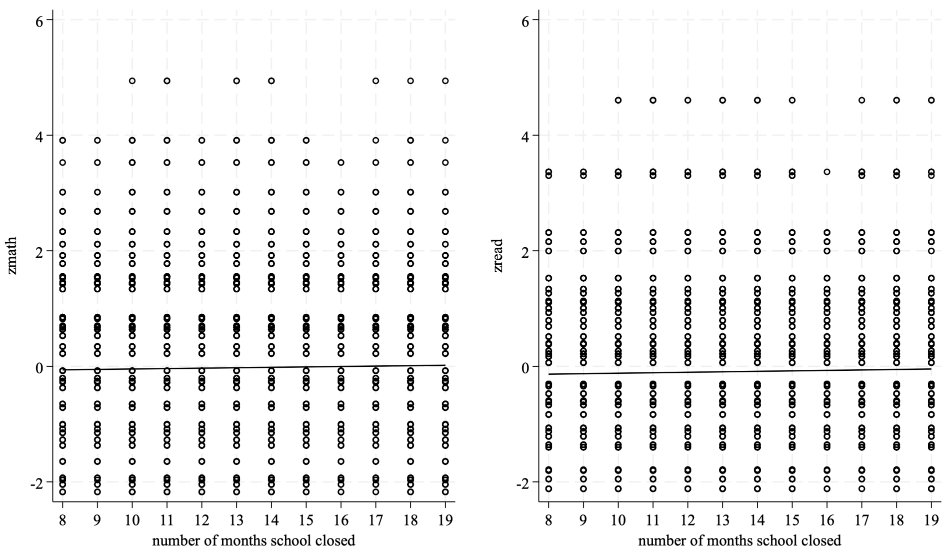
This figure shows epanechnikov kernel density function of our z test scores. The sample consists of primary public school going children (Grade 1 to 5) between 5 to 12 years old.

Figure 7: Kernel Graph: Test Scores by year and Treatment Status



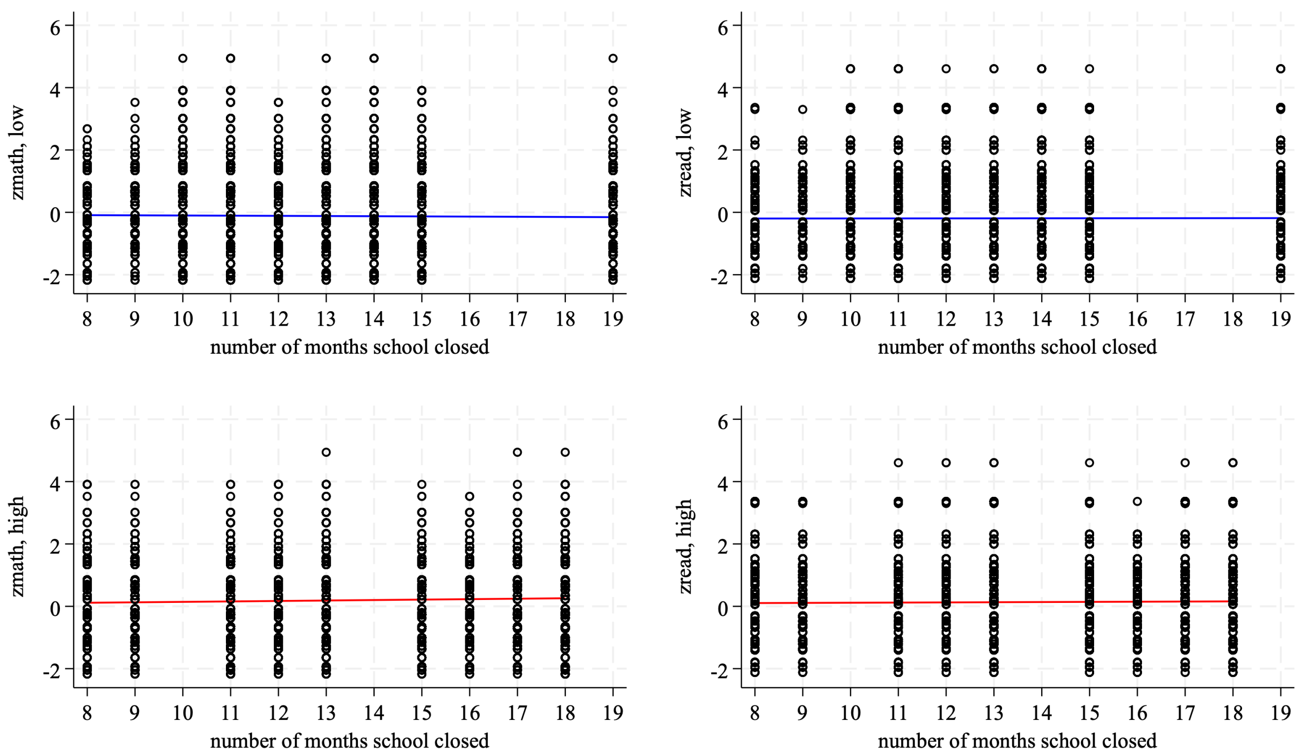
This figure shows epanechnikov kernel density function of our z test scores by year and for high versus low performing states. The sample consists of primary public school going children (Grade 1 to 5) between 5 to 12 years old.

Figure 8: Test Scores and School Closures (in number of months)



This figure shows scatter-plots of test scores and school closures between March 2020 and August 2022 (in number of months). The grey dots correspond to students average test scores for each value of school closure (in number of months) and the grey line is a refitted regression line illustrating the relationship between the two variables. The sample consists of primary public school going children (Grade 1 to 5) between 5 to 12 years old who answered ASER 2022 Household Surveys. A test for significant linear trend is performed by linear regression of z math (reading) score on number of months of school closure and evaluating the significance of the slope. The slope is positive and no statistically significant for z math (p-value=0.189), and positive and no statistically significant for z reading scores (p-value=0.134).

Figure 9: Test Scores and School Closures (in number of months): Low and High Performing States



This figure shows scatter-plots of test scores and school closures between March 2020 and August 2022 (in number of months). The grey dots correspond to district average test scores for each value of school closure (in number of months) and the blue and red line is a refitted regression line illustrating the relationship between the two variables in low and high performing states, respectively. The sample consists of primary public school going children (Grade 1 to 5) between 5 to 12 years old who answered ASER 2022 Household Surveys and live in low or high performing states. A test for significant linear trend is performed by linear regression of z math (reading) score on number of months of school closure and evaluating the significance of the slope. For low performing states, the slope is no statistically significant (for zmath the slope is -0.006 (p-value=0.367) and for z reading the slope is 0.0012 (p-value= 0.865)). For high performing states the slope is positive and no statistically significant for zmath (p-value=0.075) and positive and no statistically significant for z reading (p-value=0.426).

Table 1: Descriptive Statistics (2018-2022)

	2018	2022	Diff
	(1)	(2)	(3)
child_age	8.33	8.41	0.0803***
(in years)	(1.80)	(1.79)	
child_female	0.51	0.50	0.00659***
(child is female)	(0.50)	(0.50)	
father_attended_school	0.70	0.76	0.0627***
	(0.46)	(0.43)	
mother_attended_school	0.53	0.64	0.117***
	(0.50)	(0.48)	
test_in_Hindi	0.55	0.55	-0.00555**
(language child was tested in Hindi)	(0.50)	(0.50)	
d_vlg_internet_cafe	0.22	0.34	0.123***
(sampled village had an internet cafe)	(0.41)	(0.47)	
d_vlg_electricity	0.98	0.99	0.00980**
(sampled village has electricity)	(0.12)	(0.08)	
d_hh_computer_use	0.09	0.09	-0.00231
(household has a person who can use the computer)	(0.29)	(0.29)	
d_tuition	0.25	0.31	0.0562***
(child takes paid tuition)	(0.43)	(0.46)	
d_hh_electricity_conn	0.86	0.93	0.0757***
(household has electricity connection)	(0.35)	(0.26)	
school_class	2.96	3.00	0.0465***
(school class)	(1.43)	(1.42)	
math	2.65	2.65	0.00818
	(1.17)	(1.17)	
read	2.88	2.77	-0.0895***
	(1.48)	(1.42)	
zmath	-0.04	-0.06	-0.0142
	(1.00)	(1.01)	
zread	-0.03	-0.13	-0.0886***
	(1.01)	(0.98)	

This table shows descriptive statistics of key variables in our sample. Our sample comprises individuals between 5 to 12 years old going to public primary school in rural India. Reported values are annual averages and standard deviations are shown in parentheses. The coefficient of a regression of each covariable on 2022 year indicator is reported in column (3). *** p<0.01,** p<0.05, * p<0.1.

Table 2: Descriptive Statistics by intensity of Treatment

	Control (2018) (1)	Treatment (2018) (2)	Diff (3)	Control (2022) (4)	Treatment (2022) (5)	Diff (6)
child_age	8.31	8.40	0.0890**	8.38	8.48	0.0908***
(in years)	(1.83)	(1.69)		(1.83)	(1.70)	
child_male	0.51	0.51	-0.00542	0.50	0.50	-0.00466
(child is male)	(0.50)	(0.50)		(0.50)	(0.50)	
father_attended_school	0.67	0.81	0.141***	0.73	0.84	0.112***
	(0.47)	(0.40)		(0.44)	(0.36)	
mother_attended_school	0.44	0.77	0.324***	0.56	0.84	0.278***
	(0.50)	(0.42)		(0.50)	(0.36)	
test_in_Hindi	0.72	0.08	-0.641***	0.73	0.07	-0.655***
(language child was tested in Hindi)	(0.45)	(0.28)		(0.44)	(0.26)	
d_vlg_internet_cafe	0.20	0.26	0.0640***	0.32	0.39	0.0714***
(sampled village had an internet cafe)	(0.40)	(0.44)		(0.47)	(0.49)	
d_vlg_electricity	0.98	1.00	0.0160***	0.99	1.00	0.00325*
(sampled village has electricity)	(0.14)	(0.06)		(0.08)	(0.06)	
d_hh_computer_use	0.07	0.14	0.0662***	0.07	0.13	0.0600***
(household has a person who can use the computer)	(0.26)	(0.35)		(0.26)	(0.34)	
d_tuition	0.25	0.27	0.0232	0.31	0.31	-0.00261
(child takes paid tuition)	(0.43)	(0.44)		(0.46)	(0.46)	
d_hh_electricity_conn	0.83	0.95	0.124***	0.91	0.97	0.0649***
(household has electricity connection)	(0.38)	(0.22)		(0.29	(0.16)	
school_class	2.94	3.00	0.0605***	2.98	3.04	0.0552**
(school class)	(1.43)	(1.41)		(1.42)	(1.41)	
math	2.52	2.97	0.455***	2.57	2.87	0.300***
	(1.16)	(1.13)		(1.18)	(1.12)	
read	2.68	3.38	0.703***	2.64	3.12	0.483***
	(1.46)	(1.40)		(1.40)	(1.38)	
zmath	-0.16	0.24	0.400***	-0.13	0.13	0.256***
	(1.00)	(0.96)		(1.02)	(0.96)	
zread	-0.17	0.33	0.499***	-0.22	0.11	0.332***
	(1.00)	(0.94)		(0.97)	(0.95)	
Observations	89,135	32,316		124,168	50,784	

Column (1) and (2) of this Table show descriptive statistics of key variables in our sample by control and treatment states in 2018. Column (4) and (5) show descriptive statistics of key variables by control and treatment states in 2022. The treatment group is composed by individuals living in high performing states and the control group is composed by individuals living in low performing states. Our sample comprises school going individuals between 5 to 12 years old registered in primary public schools in rural households in India. Reported values are averages and standard deviations are shown in parentheses. The coefficient of a regression of each covariable on Treatment status for each corresponding year is reported in columns (3) and (6). *** p<0.01,** p<0.05, * p<0.1.

Table 3: Covid-19 school closures and state performing

	Average number of months schools closed	Correlation
Low Performing States	12	
High Performing States	13	0.1595

This table shows the relationship between duration of COVID-19 school closures and state performing status. The sample consists of 28 states in rural India.

Table 4: Number of Months of School Closures: March 2020-August 2022

State	Academic Calendar	Months of school closures	nmsc
Panel A: Low Performing States			
Arunachal Pradesh	June to April	April 2020-November 2020, May 2021-August 2021, January 2022	11
Assam	January to December	March 2020-September 2020, May 2021-August 2021	11
Bihar	April to March	March 2020-September 2020, April 2021-June 2021	10
Chhattisgarh	April to March	March 2020-February 2021, April 2021-July 2021, January 2022-March 2022	19
Gujarat	June to April	March 2020-December 2020, May 2021-July 2021	11
Jammu and Kashmir	November to October	April 2020-July 2020, January 2021, May 2021-August 2021	9
Jharkhand	April to March	March 2020-December 2020, May 2021-July 2021, January 2022-February 2022	15
Karnataka	June to April	March 2020-November 2020, May 2021-August 2021	11
Madhya Pradesh	June to April	March 2020-December 2020, April 2021-July 2021, January 2022	13
Meghalaya	February to December	April 2020-August 2020, May 2021-July 2021, January 2022	8
Odisha	June to May	March 2020-July 2020, September 2020, May 2021-July 2021, January 2022-February 2022	11
Puducherry	June to May	March 2020-December 2020, June 2021-August 2021	13
Rajasthan	May to April	March 2020- October 2020, December 2020-January 2021, Abril 2021, January 2022	12
Uttar Pradesh	April to March	March 2020-October 2020, April 2021-August 2021, January 2022	14
Panel B: High Performing States			
Andhra Pradesh	June to April	April 2020-September 2020, June 2021-August 2021	8
Haryana	April to March	March 2020-September 2020 December 2020, April 2021-July 2021, January 2022	13
Himachal Pradesh	January to December	April 2020-October 2020, December 2020, January 2021-February 2021 April 2021-September 2021, January 2022	17
Kerala	June to March	April 2020-October 2020, May 2021-August 2021	8
Maharashtra	June to April	April 2020-November 2020, February 2021-June 2021	11
Manipur	January to December	March 2020-January 2021, May 2021-October 2021, July 2022	18
Mizoram	April to March	March 2020-July 2020, December 2020-January 2021, May 2021-January 2022	16
Nagaland	January to December	March 2020-August 2020, October 2020, January 2021, May 2021	9
Punjab	April to March	March 2020- December 2020, March 2021-May 2021	13
Sikkim	February to December	March 2020-January 2021, April 2021-August 2021, January 2022	15
Tamil Nadu	June to May	April 2020-January 2021, May 2021-August 2021, January 2022	15
Tripura	January to December	July 2020-September 2020, April 2021-August 2021, January 2022	9
Uttarakhand	April to March	March 2020-October 2020, April 2021-July 2021	12
West Bengal	January to December	March 2020-January 2021, May 2021-November 2021	18

This table presents the states covered under ASER and included in our sample, along with their respective academic calendar months, months of COVID-19 school closures, and the number of months schools remained closed between March 2020 and August 2022. Panel A and B show the states classified as low and high performing, respectively. 'nmsc' denotes number of months of school closures.

Table 5: The Effect of State Performance on Test Scores: Sample 2022

	zmath (1)	zread (2)	math (3)	read (4)
high_state_performance	-0.358	-0.449	-0.328	-0.521
	(0.0413)***	(0.0248)***	(0.0445)***	(0.0322)***
	[0.0563]***	[0.0475]***	[0.0575]***	[0.0608]***
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	139,932	140,038	139,932	140,038
Adj R-squared	0.166	0.176	0.376	0.379
Mean Control for high_state_performance==0	-0.130	-0.220	2.566	2.641

This table shows the estimates of the effect of state performance on math and reading raw and z scores (baseline 2018). The sample includes children aged 5 to 12 in grade 1 to grade 5 who go to public school in 2022. *high_state_performance* is an indicator equals to 1 if the state average combined (math and reading) test score is above the sample median of 5.94, and zero otherwise. Control variables include father education, mother education, public school indicator, indicator for test in Hindi, internet cafe in the village, electricity in the village, computer use at home, tuition cost, electricity connectivity at home, school class, child's gender and child's age. Standard errors clustered at the state level in parentheses, standard errors clustered at the district level in squared brackets. *** p<0.01,** p<0.05, * p<0.1.

Table 6: The Effect of School Closures on Test Scores: Sample 2022

	zmath	zread	math	read
	(1)	(2)	(3)	(4)
number_months_sc	-0.0640	-0.0276	-0.0700	-0.0411
	(0.00145)***	(0.00148)***	(0.00145)***	(0.00180)***
	[0.000788]***	[0.000779]***	[0.000792]***	[0.000964]***
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	139,932	140,038	139,932	140,038
Adj R-squared	0.166	0.176	0.376	0.379

This table shows the estimates of the effect of COVID-19 school closures on math and reading raw and z scores (baseline 2018). The sample includes children aged 5 to 12 in grade 1 to grade 5 who go to public school in 2022. *number_months_sc* is the number of months schools closed between March 2020 and August 2022 in each rural Indian state of our sample. Control variables include father education, mother education, public school indicator, indicator for test in Hindi, internet cafe in the village, electricity in the village, computer use at home, tuition cost, electricity connectivity at home, school class, child's gender and child's age. Standard errors clustered at the state level in parentheses, standard errors clustered at the district level in squared brackets. *** p<0.01,** p<0.05, * p<0.1.

Table 7: The Effect of School Closures on Test Scores: Sample 2022 and binary treatment

	zmath	zread	math	read
	(1)	(2)	(3)	(4)
above_12_months_sch	-0.640***	-0.276***	-0.700***	-0.411***
	(0.00788)	(0.00779)	(0.00792)	(0.00964)
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	139,932	140,038	139,932	140,038
Adj R-squared	0.166	0.176	0.376	0.379

This table shows the estimates of the effect of COVID-19 school closures on math and reading raw and z scores (baseline 2018). The sample includes children aged 5 to 12 in grade 1 to grade 5 who go to public school in 2022. *above_12_months_sch* is an indicator equals to one for states with number of months of school closures above the sample median of 12 and zero, otherwise. Control variables include father education, mother education, public school indicator, indicator for test in Hindi, internet cafe in the village, electricity in the village, computer use at home, tuition cost, electricity connectivity at home, school class, child's gender and child's age. Standard errors clustered at district level in parentheses. *** p<0.01,** p<0.05, * p<0.1.

Table 8: The Effect of State Performance and School Closures on Test Scores: Sample 2022

	zmath	zread	math	read
	(1)	(2)	(3)	(4)
high_state_performance*above_12_months_sch	-1.362***	-0.923***	-1.424***	-1.224***
	(0.00782)	(0.00749)	(0.00792)	(0.00954)
high_state_performance	0.365***	0.198***	0.396***	0.291***
	(0.0565)	(0.0475)	(0.0577)	(0.0607)
above_12_months_sch	0.722***	0.646***	0.724***	0.812***
	(0.00564)	(0.00540)	(0.00565)	(0.00679)
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	139,932	140,038	139,932	140,038
Adj R-squared	0.166	0.176	0.376	0.379
Mean Control for high_state_performance==0	-0.112	-0.193	2.596	2.692

This table shows the estimates of the effect of state performance and COVID-19 school closures on math and reading raw and z scores (baseline 2018). The sample includes children aged 5 to 12 in grade 1 to grade 5 who go to public school in 2022. *high_state_performance* is an indicator equals to 1 if the state average combined (math and reading) test score is above the sample median of 5.94, and zero otherwise. *above_12_months_sch* is an indicator equals to one for states with number of months of school closures above the sample median of 12 and zero, otherwise. Control variables include father education, mother education, public school indicator, indicator for test in Hindi, internet cafe in the village, electricity in the village, computer use at home, tuition cost, electricity connectivity at home, school class, child's gender and child's age. Standard errors clustered at the district level in parentheses. *** p<0.01,** p<0.05, * p<0.1.

Table 9: Parallel Trends, State Performance on Test Scores: 2016-2018

	zmath	zread	math	read
	(1)	(2)	(3)	(4)
interaction_2018=	0.0195	0.0158	0.0299	0.0298
high_state_performance*year2018	(0.0465)	(0.0347)	(0.0475)	(0.0445)
	[0.0234]	[0.0204]	[0.0241]	[0.0259]
high_state_performance	0.636	0.305	0.655	0.398
	(0.0517)***	(0.0957)***	(0.0503)***	(0.125)***
	[0.0516]***	[0.0538]***	[0.0536]***	[0.0706]***
year2018	0.0109	0.0214	0.00863	0.0259
	(0.0412)	(0.0332)	(0.0428)	(0.0440)
	[0.0146]	[0.0130]	[0.0149]	[0.0167]
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	180,160	180,554	180,160	180,554
Adj R-squared	0.199	0.218	0.403	0.413
Mean Control 2016 & high_state_performance==0	-0.211	-0.240	2.464	2.593

This table shows the estimates of the effect of state performance on math and reading raw and z scores (baseline 2018) for Pre-covid years (2016 and 2018). The sample includes children aged 5 to 12 in grade 1 to grade 5 who go to public school. $high\_state\_performance$ is an indicator equals to 1 if the state average combined (math and reading) test score is above the sample median of 5.94, and zero otherwise. $year2018$ is an indicator equals to one for survey year 2018 and zero, otherwise. Control variables include father education, mother education, public school indicator, indicator for test in Hindi, internet cafe in the village, electricity in the village, computer use at home, tuition cost, electricity connectivity at home, school class, child's gender and child's age. Standard errors clustered at the state level in parentheses, standard errors clustered at the district level in squared brackets. *** p<0.01,** p<0.05, * p<0.1.

Table 10: DID State Performance on Test Scores: 2018-2022

	zmath	zread	math	read
	(1)	(2)	(3)	(4)
interaction_2022=	-0.126	-0.134	-0.133	-0.176
high_state_performance*year2022	(0.0576)**	(0.0688)*	(0.0576)**	(0.0855)**
	[0.0253]***	[0.0277]***	[0.0250]***	[0.0342]***
high_state_performance	0.00230	-0.111	0.0150	-0.129
	(0.0527)	(0.0646)*	(0.0557)	(0.0799)
	[0.0563]	[0.0461]**	[0.0573]	[0.0592]**
year2022	-0.0393	-0.109	-0.0409	-0.146
	(0.0453)	(0.0437)**	(0.0460)	(0.0555)**
	[0.0150]***	[0.0148]***	[0.0151]***	[0.0187]***
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	229,233	229,644	229,233	229,644
Adj R-squared	0.169	0.191	0.383	0.395
Mean Control 2018 & high_state_performance==0	-0.156	-0.172	2.518	2.680

This table shows the estimates of the effect of state performance on math and reading raw and z scores (baseline 2018). The sample includes children aged 5 to 12 in grade 1 to grade 5 who go to public school in 2018 and 2022. $high\_state\_performance$ is an indicator equals to 1 if the state average combined (math and reading) test score is above the sample median of 5.94, and zero otherwise. $year2022$ is an indicator equals to one for survey year 2022 and zero, otherwise. Control variables include father education, mother education, public school indicator, indicator for test in Hindi, internet cafe in the village, electricity in the village, computer use at home, tuition cost, electricity connectivity at home, school class, child's gender and child's age. Standard errors clustered at the state level in parentheses, standard errors clustered at the district level in squared brackets. *** p<0.01,** p<0.05, * p<0.1.

Table 11: DID State Performance and School Closures: 2018-2022

	zmath (1)	zread (2)	math (3)	read (4)
interaction_2022= high_state_performance*above_12_months_sch	-0.198*** (0.0372)	-0.152*** (0.0395)	-0.204*** (0.0367)	-0.195*** (0.0484)
high_state_performance	-0.0230 (0.0560)	-0.188*** (0.0464)	-0.0167 (0.0569)	-0.236*** (0.0593)
above_12_months_sch = 1	0.0464*** (0.0167)	-0.0363** (0.0170)	0.0435** (0.0171)	-0.0569*** (0.0214)
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	229,233	229,644	229,233	229,644
Adj R-squared	0.168	0.186	0.382	0.391
Mean Control high_state_performance==0 above_12_months_sch==0	-0.140	-0.180	2.547	2.685

This table shows the estimates of the effect of state performance on math and reading raw and z scores (baseline 2018). The sample includes children aged 5 to 12 in grade 1 to grade 5 who go to public school in 2018 and 2022. *high_state_performance* is an indicator equals to 1 if the state average combined (math and reading) test score is above the sample median of 5.94, and zero otherwise. *above_12_months_sch* is an indicator equals to one for states with number of months of school closures above the sample median of 12 and zero, otherwise. Control variables include father education, mother education, public school indicator, indicator for test in Hindi, internet cafe in the village, electricity in the village, computer use at home, tuition cost, electricity connectivity at home, school class, child's gender and child's age. Standard errors clustered at the district level in parentheses.*** p<0.01,** p<0.05, * p<0.1.

Table 12: The Effect of State Performance on Test Scores: DID By Groups

	zmath	zread	math	read
	(1)	(2)	(3)	(4)
Panel A: Only women				
interaction_2022=	-0.141	-0.138	-0.154	-0.186
high_state_performance*year2022	(0.0582)**	(0.0779)*	(0.0574)**	(0.0957)*
	[0.0281]***	[0.0339]***	[0.0276]***	[0.0418]***
high_state_performance	0.0815	-0.106	0.0712	-0.131
	(0.0789)	(0.0650)	(0.0806)	(0.0736)*
	[0.0636]	[0.0573]*	[0.0639]	[0.0685]*
year2022	-0.0257	-0.105	-0.0261	-0.139
	(0.0455)	(0.0470)**	(0.0465)	(0.0596)**
	[0.0155]*	[0.0156]***	[0.0159]	[0.0200]***
Observations	116,335	116,540	116,335	116,540
Adj R-squared	0.187	0.209	0.392	0.409
Mean Control 2018	-0.188	-0.161	2.492	2.704
high_state_performance==0				
Panel B: Only men				
interaction_2022=	-0.107	-0.129	-0.109	-0.163
high_state_performance*year2022	(0.0570)*	(0.0596)**	(0.0579)*	(0.0754)**
	[0.0275]***	[0.0272]***	[0.0273]***	[0.0336]***
high_state_performance	-0.0750	-0.112	-0.0413	-0.126
	(0.0523)	(0.0769)	(0.0538)	(0.101)
	[0.0701]	[0.0536]**	[0.0695]	[0.0715]*
year2022	-0.0534	-0.113	-0.0561	-0.152
	(0.0449)	(0.0400)***	(0.0452)	(0.0508)***
	[0.0169]***	[0.0167]***	[0.0166]***	[0.0206]***
Observations	112,898	113,104	112,898	113,104
Adj R-squared	0.156	0.175	0.378	0.381
Mean Control 2018	-0.124	-0.185	2.543	2.654
high_state_performance==0				
Panel C: TV access at home				
interaction_2022=	-0.107	-0.127	-0.116	-0.169
high_state_performance*year2022	(0.0630)*	(0.0651)*	(0.0634)*	(0.0812)**
	[0.0260]***	[0.0275]***	[0.0258]***	[0.0345]***
high_state_performance	0.124	-0.0484	0.142	-0.0534
	(0.0615)*	(0.0737)	(0.0597)**	(0.0842)
	[0.0587]**	[0.0514]	[0.0588]**	[0.0655]
year2022	-0.0624	-0.135	-0.0630	-0.176
	(0.0510)	(0.0448)***	(0.0523)	(0.0574)***
	[0.0184]***	[0.0186]***	[0.0185]***	[0.0237]***
Observations	124,289	124,460	124,289	124,460
Adj R-squared	0.147	0.165	0.385	0.397
Mean Control 2018	0.059	0.072	2.755	3.019
high_state_performance==0				
Panel D: No TV access at home				
interaction_2022=	-0.136	-0.104	-0.148	-0.146
high_state_performance*year2022	(0.0562)**	(0.0845)	(0.0568)**	(0.107)
	[0.0431]***	[0.0503]**	[0.0422]***	[0.0623]**
high_state_performance	-0.381	-0.402	-0.359	-0.432
	(0.0714)***	(0.0627)***	(0.0746)***	(0.0828)***
	[0.0885]***	[0.0765]***	[0.0869]***	[0.0944]***
year2022	-0.0169	-0.0866	-0.0180	-0.118
	(0.0413)	(0.0406)**	(0.0418)	(0.0517)**
	[0.0163]	[0.0162]***	[0.0164]	[0.0204]***
Observations	102,572	102,803	102,572	102,803
Adj R-squared	0.143	0.152	0.353	0.351
Mean Control 2018	-0.293	-0.326	2.368	2.468
high_state_performance==0				
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

The Effect of State Performance on Test Scores: DID By Groups (continue)

	zmath (1)	zread (2)	math (3)	read (4)
Panel E: Digital access at the village				
interaction_2022=	-0.0955	-0.0938	-0.106	-0.128
high_state_performance*year2022	(0.0620)	(0.0795)	(0.0602)*	(0.0944)
	[0.0385]**	[0.0428]**	[0.0386]***	[0.0541]**
high_state_performance	-0.175	-0.384	-0.142	-0.463
	(0.0658)**	(0.0839)***	(0.0635)**	(0.0995)***
	[0.0905]*	[0.0766]***	[0.0970]	[0.102]***
year2022	-0.0714	-0.132	-0.0679	-0.166
	(0.0419)*	(0.0330)***	(0.0422)	(0.0413)***
	[0.0225]***	[0.0234]***	[0.0225]***	[0.0298]***
Observations	59,567	59,663	59,567	59,663
Adj R-squared	0.182	0.203	0.390	0.401
Mean Control 2018	-0.140	-0.179	2.544	2.682
high_state_performance==0				
Panel F: No Digital access at the village				
interaction_2022=	-0.144	-0.154	-0.151	-0.202
high_state_performance*year2022	(0.0542)**	(0.0640)**	(0.0551)**	(0.0808)**
	[0.0270]***	[0.0294]***	[0.0266]***	[0.0358]***
high_state_performance	0.0816	-0.00415	0.0863	-0.000243
	(0.0561)	(0.0596)	(0.0619)	(0.0742)
	[0.0687]	[0.0541]	[0.0698]	[0.0683]
year2022	-0.0230	-0.0955	-0.0254	-0.130
	(0.0442)	(0.0454)**	(0.0449)	(0.0576)**
	[0.0159]	[0.0157]***	[0.0160]	[0.0197]***
Observations	169,666	169,981	169,666	169,981
Adj R-squared	0.169	0.191	0.384	0.396
Mean Control 2018	-0.161	-0.172	2.511	2.677
high_state_performance==0				
Panel G: At least one parent attended school				
interaction_2022=	-0.130	-0.136	-0.138	-0.177
high_state_performance*year2022	(0.0570)**	(0.0667)*	(0.0573)**	(0.0828)**
	[0.0249]***	[0.0268]***	[0.0246]***	[0.0332]***
high_state_performance	-0.0229	-0.156	-0.0126	-0.190
	(0.0523)	(0.0659)**	(0.0537)	(0.0816)**
	[0.0524]	[0.0484]***	[0.0531]	[0.0625]***
year2022	-0.0387	-0.116	-0.0407	-0.154
	(0.0462)	(0.0430)**	(0.0470)	(0.0548)***
	[0.0152]**	[0.0147]***	[0.0152]***	[0.0185]***
Observations	189,617	189,933	189,617	189,933
Adj R-squared	0.148	0.170	0.385	0.397
Mean Control 2018	-0.042	-0.046	2.621	2.826
high_state_performance==0				
Panel H: No parent attended school				
interaction_2022=	-0.151	-0.132	-0.160	-0.181
high_state_performance*year2022	(0.0524)***	(0.0700)*	(0.0502)***	(0.0862)**
	[0.0477]***	[0.0526]**	[0.0463]***	[0.0646]***
high_state_performance	0.240	0.141	0.336	0.312
	(0.119)*	(0.116)	(0.126)**	(0.162)*
	[0.193]	[0.143]	[0.192]*	[0.186]*
year2022	-0.0343	-0.0792	-0.0348	-0.107
	(0.0437)	(0.0462)*	(0.0443)	(0.0596)*
	[0.0237]	[0.0232]***	[0.0240]	[0.0295]***
Observations	39,616	39,711	39,616	39,711
Adj R-squared	0.121	0.119	0.318	0.306
Mean Control 2018	-0.456	-0.491	2.250	2.312
high_state_performance==0				
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

This table shows the estimates of the effect of state performance on math and reading raw and z scores (baseline 2018). The sample includes children aged 5 to 12 in grade 1 to grade 5 who go to public school in 2018 and 2022. $high\_state\_performance$ is an indicator equals to 1 if the state average combined (math and reading) test score is above the sample median of 5.94, and zero otherwise. $year2022$ is an indicator equals to one for survey year 2022 and zero, otherwise. Control variables include father education, mother education, public school indicator, indicator for test in Hindi, internet cafe in the village, electricity in the village, computer use at home, tuition cost, electricity connectivity at home, school class, child's gender and child's age. Standard errors clustered at the state level in parentheses, standard errors clustered at the district level in squared brackets.
*** p<0.01,** p<0.05, * p<0.1.

Table 13: The Effect of School Closures on Test Scores: Sample 2022, Heterogeneous Effects

	zmath (1)	zread (2)	math (3)	read (4)
Panel A: only Women				
number_months_sc*child_female	-0.00127	0.00267	-0.00191	0.00250
	(0.00310)	(0.00327)	(0.00301)	(0.00438)
number_months_sc	-0.0634***	-0.0290***	-0.0689***	-0.0425***
	(0.00198)	(0.00187)	(0.00193)	(0.00248)
child_female	-0.0161	0.0149	-0.00697	0.0363
	(0.0405)	(0.0407)	(0.0397)	(0.0543)
Observations	139,932	140,038	139,932	140,038
Adj R-squared	0.166	0.176	0.376	0.379
Panel B: TV access at home				
number_months_sc* tv_access_at_home	0.00705*	0.00366	0.00721*	0.00434
	(0.00418)	(0.00384)	(0.00428)	(0.00525)
number_months_sc	-0.0611***	-0.0229***	-0.0672***	-0.0350***
	(0.00271)	(0.00251)	(0.00278)	(0.00344)
tv_access_at_home	0.0652	0.102**	0.0650	0.135**
	(0.0547)	(0.0482)	(0.0559)	(0.0655)
Observations	138,241	138,344	138,241	138,344
Adj R-squared	0.170	0.180	0.380	0.382
Panel C: Digital Access at the Village				
number_months_sc* digital_access_village	0.00159	0.00202	0.000783	0.00143
	(0.00466)	(0.00437)	(0.00486)	(0.00541)
number_months_sc	-0.0647***	-0.0284***	-0.0703***	-0.0417***
	(0.00216)	(0.00204)	(0.00228)	(0.00253)
digital_access_village	-0.222*	-0.226	-0.227*	-0.266
	(0.131)	(0.153)	(0.131)	(0.185)
Observations	139,932	140,038	139,932	140,038
Adj R-squared	0.166	0.176	0.376	0.379
Panel D: Digital Access at home				
number_months_sc* digital_access_home	0.00352	0.00604	0.00290	0.00528
	(0.00670)	(0.00526)	(0.00659)	(0.00638)
number_months_sc	-0.0629***	-0.0275***	-0.0681***	-0.0406***
	(0.00119)	(0.00106)	(0.00119)	(0.00133)
digital_access_home	0.0509	-0.0126	0.0686	0.0216
	(0.0893)	(0.0793)	(0.0903)	(0.0974)
Observations	92,353	92,427	92,353	92,427
Adj R-squared	0.161	0.170	0.385	0.389
Panel E: At least one parent attended school				
number_months_sc*at_least_one_parent_attended_sch	0.00525	0.00640	0.00788	0.0118
	(0.00823)	(0.00716)	(0.00829)	(0.00904)
number_months_sc	-0.0683***	-0.0328***	-0.0763***	-0.0506***
	(0.00614)	(0.00531)	(0.00621)	(0.00674)
at_least_one_parent_attended_sch	-0.199*	-0.209**	-0.225**	-0.302***
	(0.105)	(0.0892)	(0.106)	(0.113)
Observations	139,932	140,038	139,932	140,038
Adj R-squared	0.166	0.177	0.377	0.380
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes

This table shows the estimates of the effect of COVID-19 school closures on math and reading raw and z scores (baseline 2018). The sample includes children aged 5 to 12 in grade 1 to grade 5 who go to public school in 2022. *number_months_sc* is the number of months schools closed between March 2020 and August 2022 in each rural Indian state of our sample. Control variables include father education, mother education, public school indicator, indicator for test in Hindi, internet cafe in the village, electricity in the village, computer use at home, tuition cost, electricity connectivity at home, school class, child's gender and child's age. Standard errors clustered at the district level in parentheses.*** p<0.01,** p<0.05, * p<0.1.

Table 14: The Effect of School Closures on Test Scores: Sample 2022, Heterogeneous Effects

	zmath (1)	zread (2)	math (3)	read (4)
Panel A: only Women				
above_12_months_sch*child_female	-0.0353**	-0.0150	-0.0366**	-0.0206
	(0.0152)	(0.0146)	(0.0150)	(0.0187)
above_12_months_sch	-0.621***	-0.268***	-0.680***	-0.400***
	(0.0121)	(0.0106)	(0.0120)	(0.0134)
child_female	-0.0158	0.0553***	-0.0141	0.0771***
	(0.0108)	(0.00899)	(0.0109)	(0.0115)
Observations	139,932	140,038	139,932	140,038
Adj R-squared	0.166	0.176	0.377	0.379
Mean Control	-0.062	-0.149	2.654	2.756
above_12_months_sch==0 & child_female==0				
Panel B: TV access at home				
above_12_months_sch* tv_access_at_home	0.0117	-0.000712	0.0136	0.000233
	(0.0203)	(0.0190)	(0.0207)	(0.0251)
above_12_months_sch	-0.575***	-0.206***	-0.636***	-0.324***
	(0.0155)	(0.0144)	(0.0159)	(0.0190)
tv_access_at_home	0.149***	0.149***	0.150***	0.190***
	(0.0141)	(0.0120)	(0.0143)	(0.0158)
Observations	138,241	138,344	138,241	138,344
Adj R-squared	0.170	0.180	0.379	0.382
Mean Control	-0.251	-0.349	2.459	2.494
above_12_months_sch==0 & tv_access_at_home==0				
Panel C: Digital Access at the Village				
above_12_months_sch* digital_access_village	-0.00138	0.00994	-0.00481	0.00589
	(0.0242)	(0.0237)	(0.0247)	(0.0300)
above_12_months_sch	-0.640***	-0.280***	-0.698***	-0.414***
	(0.0128)	(0.0123)	(0.0133)	(0.0156)
digital_access_village	-0.202*	-0.205	-0.215*	-0.251
	(0.118)	(0.144)	(0.117)	(0.173)
Observations	139,932	140,038	139,932	140,038
Adj R-squared	0.166	0.176	0.376	0.379
Mean Control	-0.072	-0.120	2.648	2.800
above_12_months_sch==0 & digital_access_village==0				
Panel D: Digital Access at home				
above_12_months_sch* digital_access_home	0.0121	0.0251	0.00812	0.0239
	(0.0370)	(0.0315)	(0.0367)	(0.0383)
above_12_months_sch	-0.627***	-0.272***	-0.679***	-0.403***
	(0.0102)	(0.00987)	(0.0103)	(0.0123)
digital_access_home	0.0900	0.0528	0.101*	0.0781
	(0.0553)	(0.0549)	(0.0572)	(0.0676)
Observations	92,353	92,427	92,353	92,427
Adj R-squared	0.161	0.170	0.385	0.389
Mean Control	-0.018	-0.068	2.708	2.874
above_12_months_sch==0 & digital_access_home==0				
Panel E: At least one parent attended school				
above_12_months_sch*at_least_one_parent_attended_sch	-0.0126	-0.00930	-0.00469	0.00175
	(0.0319)	(0.0304)	(0.0318)	(0.0384)
above_12_months_sch	-0.633***	-0.272***	-0.698***	-0.416***
	(0.0238)	(0.0227)	(0.0239)	(0.0288
at_least_one_parent_attended_sch	-0.128***	-0.127***	-0.126***	-0.158***
	(0.0271)	(0.0236)	(0.0272)	(0.0301)
Observations	139,932	140,038	139,932	140,038
Adj R-squared	0.166	0.177	0.377	0.380
Mean Control	-0.451	-0.551	2.314	2.307
above_12_months_sch==0 & at_least_one_parent_attended_sch==0				
Controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes

This table shows the estimates of the effect of COVID-19 school closures on math and reading raw and z scores (baseline 2018). The sample includes children aged 5 to 12 in grade 1 to grade 5 who go to public school in 2022. *above_12_months_sch* is an indicator equals to one for states with number of months of school closures above the sample median of 12 and zero, otherwise. Control variables include father education, mother education, public school indicator, indicator for test in Hindi, internet cafe in the village, electricity in the village, computer use at home, tuition cost, electricity connectivity at home, school class, child's gender and child's age. Standard errors clustered at the district level in parentheses.*** p<0.01,** p<0.05, * p<0.1.

Table 15: The Effect of State Performance on Test Scores: DID Heterogeneous Effects

	zmath (1)	zread (2)	math (3)	read (4)
Panel A: Only women				
high_state_performance*year2022*child_female	-0.0291	-0.00264	-0.0394*	-0.0151
	(0.0220)	(0.0254)	(0.0217)	(0.0319)
child_female * high_state_performance	0.120***	0.104***	0.127***	0.142***
	(0.0198)	(0.0206)	(0.0191)	(0.0249)
child_female * year2022	0.0279**	0.0127	0.0308***	0.0192
	(0.0117)	(0.0124)	(0.0112)	(0.0151)
high_state_performance*year2022	-0.110***	-0.132***	-0.112***	-0.167***
	(0.0277)	(0.0273)	(0.0274)	(0.0335)
child_female = 1	-0.0848***	0.00692	-0.0855***	0.0125
	(0.0105)	(0.0109)	(0.0105)	(0.0136)
high_state_performance = 1	-0.0600	-0.165***	-0.0506	-0.203***
	(0.0571)	(0.0470)	(0.0582)	(0.0603)
year2022 = 1	-0.0537***	-0.116***	-0.0568***	-0.156***
	(0.0167)	(0.0165)	(0.0164)	(0.0203)
Observations	229,233	229,644	229,233	229,644
Adj R-squared	0.170	0.191	0.384	0.395
Mean Control 2018	-0.124	-0.185	2.543	2.654
high_state_performance==0 &child_female == 0				
Panel B: TV access at home				
high_state_performance*year2022*tv_access_at_home	0.0259	-0.0310	0.0274	-0.0369
	(0.0362)	(0.0445)	(0.0354)	(0.0573)
tv_access_at_home * high_state_performance	0.0341	0.0643**	0.0317	0.0745**
	(0.0307)	(0.0289)	(0.0292)	(0.0343)
tv_access_at_home * year2022	-0.0392**	-0.0425**	-0.0362*	-0.0477*
	(0.0184)	(0.0189)	(0.0187)	(0.0245)
high_state_performance*year2022	-0.135***	-0.0985**	-0.144***	-0.134**
	(0.0396)	(0.0483)	(0.0390)	(0.0607)
tv_access_at_home = 1	0.157***	0.149***	0.155***	0.184***
	(0.0143)	(0.0136)	(0.0148)	(0.0177)
high_state_performance = 1	-0.0667	-0.202***	-0.0516	-0.238***
	(0.0643)	(0.0535)	(0.0645)	(0.0673)
year2022 = 1	-0.0213	-0.0909***	-0.0241	-0.125***
	(0.0165)	(0.0162)	(0.0167)	(0.0207)
Observations	226,861	227,263	226,861	227,263
Adj R-squared	0.173	0.194	0.386	0.397
Mean Control 2018	-0.293	-0.326	2.368	2.468
high_state_performance==0&tv_access_at_home == 0				
Panel C: Digital Access at the Village				
high_state_performance*year2022*digital_access_village	0.0527	0.0614	0.0516	0.0799*
	(0.0357)	(0.0390)	(0.0352)	(0.0479)
digital_access_village * high_state_performance	-0.0329	-0.0407	-0.0306	-0.0501
	(0.0313)	(0.0337)	(0.0318)	(0.0426)
digital_access_village * year2022	-0.0446**	-0.0405*	-0.0399*	-0.0448
	(0.0219)	(0.0232)	(0.0217)	(0.0291)
high_state_performance*year2022	-0.139***	-0.151***	-0.147***	-0.198***
	(0.0267)	(0.0289)	(0.0263)	(0.0352)
digital_access_village = 1	0.0584	0.0809	0.0630	0.120
	(0.0836)	(0.0950)	(0.0842)	(0.117)
high_state_performance = 1	0.00792	-0.103**	0.0203	-0.119**
	(0.0565)	(0.0466)	(0.0575)	(0.0596)
year2022 = 1	-0.0284*	-0.0995***	-0.0310*	-0.135***
	(0.0159)	(0.0157)	(0.0160)	(0.0197)
Observations	229,233	229,644	229,233	229,644
Adj R-squared	0.169	0.191	0.383	0.395
Mean Control 2018	-0.161	-0.172	2.511	2.677
high_state_performance==0&digital_access_village == 0				
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

The Effect of State Performance on Test Scores: DID Heterogeneous Effects (continue)

	zmath	zread	math	read
	(1)	(2)	(3)	(4)
Panel D: At least one parent attended school				
high_state_performance*year2022*at_least_one_parent_attended_sch	0.00879	-0.00664	0.0123	0.00342
	(0.0471)	(0.0442)	(0.0461)	(0.0544)
at_least_one_parent_attended_sch * high_state_performance	0.0960**	0.113**	0.0974**	0.139**
	(0.0441)	(0.0470)	(0.0449)	(0.0598)
at_least_one_parent_attended_sch * year2022	0.00620	-0.0293	0.00917	-0.0337
	(0.0220)	(0.0205)	(0.0221)	(0.0257)
high_state_performance*year2022	-0.140***	-0.130***	-0.151***	-0.181***
	(0.0486)	(0.0471)	(0.0473)	(0.0569)
at_least_one_parent_attended_sch= 1	-0.137***	-0.113***	-0.130***	-0.132***
	(0.0206)	(0.0208)	(0.0210)	(0.0262)
high_state_performance = 1	-0.0748	-0.209***	-0.0630	-0.248***
	(0.0709)	(0.0663)	(0.0717)	(0.0842)
year2022 = 1	-0.0435*	-0.0860***	-0.0472*	-0.119***
	(0.0241)	(0.0233)	(0.0244)	(0.0294)
Observations	229,233	229,644	229,233	229,644
Adj R-squared	0.170	0.192	0.384	0.395
Mean Control 2018	-0.456	-0.491	2.250	2.312
high_state_performance==0 &at_least_one_parent_attended_sch == 0				
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

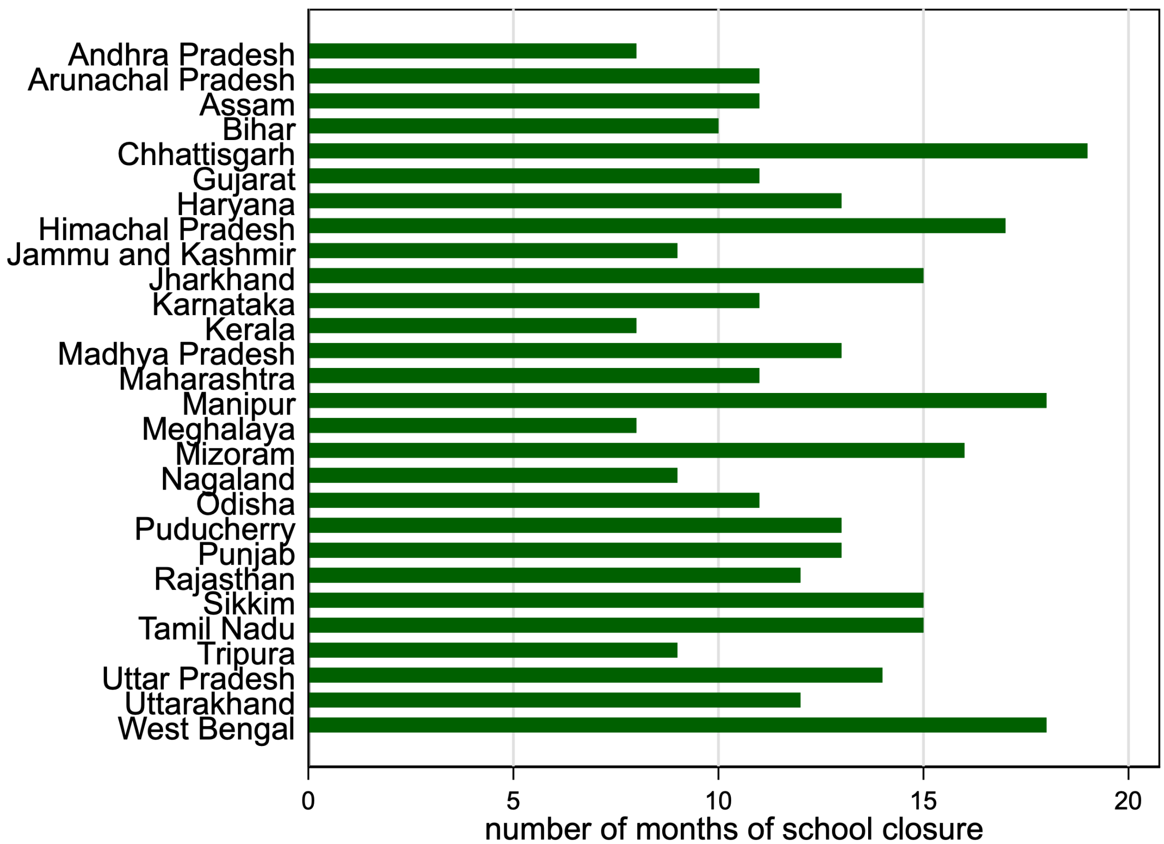
This table shows the heterogeneous effects of state performance on math and reading raw and z scores (baseline 2018). The sample includes children aged 5 to 12 in grade 1 to grade 5 who go to public school in 2018 and 2022. $high\_state\_performance$ is an indicator equals to 1 if the state average combined (math and reading) test score is above the sample median of 5.94, and zero otherwise. $year2022$ is an indicator equals to one for survey year 2022 and zero, otherwise. Control variables include father education, mother education, public school indicator, indicator for test in Hindi, internet cafe in the village, electricity in the village, computer use at home, tuition cost, electricity connectivity at home, school class, child's gender and child's age. Standard errors clustered at the district level in parentheses.
*** p<0.01,** p<0.05, * p<0.1.

Online Appendix

Figures in Appendix

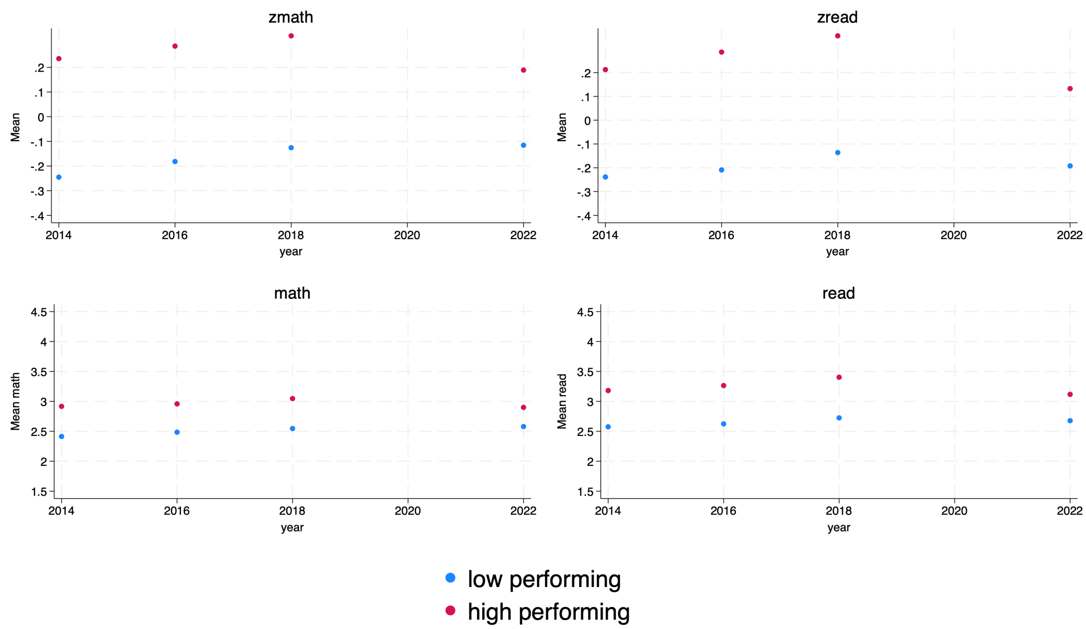
Tables in Appendix

Figure A1: Number of Months of School Closure in India (March 2020-August 2022), by state



This figure shows the number of effective months schools remained close in each state between March 2020 and August 2022 due to COVID-19 state restrictions. The data come from the Oxford COVID-19 Government Response Tracker (OxCGRT) Project. We account for effective months of school closures by taking into consideration only months pertaining to the state school academic calendar.

Figure A2: Test scores by year and treatment status



This figure shows the means of raw and z test scores for the sample of school going children. We restrict the sample to students aged 5 to 12 years old in public primary schools (grade 1 to 5). The treatment group corresponds to children living in high performing states with state average math and reading scores (2014-2018) above the sample median of 5.94. The solid dots show the means and 95% confidence intervals are denoted by vertical capped bars.

Table A1: List of states in the sample

State	Population	Population share (%)	math	read	math and read
Panel A: High Performing States					
Andhra Pradesh	52,787,000	3.87%	3.10	3.17	6.28
Haryana	29,483,000	2.16%	2.89	3.16	6.05
Himachal Pradesh	7,394,000	0.54%	3.29	3.67	6.96
Kerala	35,489,000	2.60%	3.22	3.73	6.96
Maharashtra	124,437,000	9.13%	2.71	3.32	6.03
Manipur	3,165,000	0.23%	3.40	3.15	6.55
Mizoram	1,216,000	0.09%	3.32	3.29	6.61
Nagaland	2,192,000	0.16%	3.15	3.13	6.29
Punjab	30,339,000	2.23%	3.12	3.40	6.52
Sikkim	677,000	0.05%	3.31	3.40	6.71
Tamil Nadu	76,402,000	5.61%	2.98	3.23	6.21
Tripura	4,071,000	0.30%	2.98	3.30	6.27
Uttarakhand	11,399,000	0.84%	2.74	3.21	5.95
West Bengal	98,125,000	7.20%	2.83	3.15	5.98
Panel B: Low Performing States					
Arunachal Pradesh	1,533,000	0.11%	3.08	2.70	5.78
Assam	35,043,000	2.57%	2.55	2.75	5.29
Bihar	123,083,000	9.03%	2.55	2.51	5.06
Chhattisgarh	29,493,000	2.16%	2.52	2.88	5.40
Gujarat	69,788,000	5.12%	2.58	3.09	5.67
Jammu and Kashmir	13,408,000	0.98%	2.82	2.86	5.68
Jharkhand	38,471,000	2.82%	2.40	2.48	4.88
Karnataka	66,845,000	4.90%	2.87	3.04	5.91
Madhya Pradesh	84,516,000	6.20%	2.26	2.43	4.69
Meghalaya	3,288,000	0.24%	2.87	3.07	5.93
Odisha	45,696,000	3.35%	2.75	3.18	5.92
Puducherry	1,572,000	0.12%	2.90	2.80	5.70
Rajasthan	79,281,000	5.82%	2.28	2.49	4.77
Uttar Pradesh	230,907,000	16.94%	2.15	2.22	4.38
Total	1,300,100,000	95.38%			

This table presents the states covered under ASER and included in our sample, along with their respective population, population share in India, and state average ASER test scores (2014-2018) in math and reading. Population projections are taken from the Government of india: https://mohfw.gov.in/sites/default/files/Population%20Projection%20Report%202011-2036%20-%20upload_compressed_0.pdf. Panel A and B show the states classified as high and low performing, respectively.

Table A2: Number of Months of School Closures: March 2020-August 2022

State	Academic Calendar	Months of school closures	nmsc
Panel A: High Performing States			
Andhra Pradesh	June to April	April 2020-September 2020, June 2021-August 2021	8
Haryana	April to March	March 2020-September 2020 December 2020, April 2021-July 2021, January 2022	13
Himachal Pradesh	January to December	April 2020-October 2020, December 2020, January 2021-February 2021 April 2021-September 2021, January 2022	17
Kerala	June to March	April 2020-October 2020, May 2021-August 2021	8
Maharashtra	June to April	April 2020-November 2020, February 2021-June 2021	11
Manipur	January to December	March 2020-January 2021, May 2021-October 2021, July 2022	18
Mizoram	April to March	March 2020-July 2020, December 2020-January 2021, May 2021-January 2022	16
Nagaland	January to December	March 2020-August 2020, October 2020, January 2021, May 2021	9
Punjab	April to March	March 2020- December 2020, March 2021-May 2021	13
Sikkim	February to December	March 2020-January 2021, April 2021-August 2021, January 2022	15
Tamil Nadu	June to May	April 2020-January 2021, May 2021-August 2021, January 2022	15
Tripura	January to December	July 2020-September 2020, April 2021-August 2021, January 2022	9
Uttarakhand	April to March	March 2020-October 2020, April 2021-July 2021	12
West Bengal	January to December	March 2020-January 2021, May 2021-November 2021	18
Panel B: Low Performing States			
Arunachal Pradesh	June to April	April 2020-November 2020, May 2021-August 2021, January 2022	11
Assam	January to December	March 2020-September 2020, May 2021-August 2021	11
Bihar	April to March	March 2020-September 2020, April 2021-June 2021	10
Chhattisgarh	April to March	March 2020-February 2021, April 2021-July 2021, January 2022-March 2022	19
Gujarat	June to April	March 2020-December 2020, May 2021-July 2021	11
Jammu and Kashmir	November to October	April 2020-July 2020, January 2021, May 2021-August 2021	9
Jharkhand	April to March	March 2020-December 2020, May 2021-July 2021, January 2022-February 2022	15
Karnataka	June to April	March 2020-November 2020, May 2021-August 2021	11
Madhya Pradesh	June to April	March 2020-December 2020, April 2021-July 2021, January 2022	13
Meghalaya	February to December	April 2020-August 2020, May 2021-July 2021, January 2022	8
Odisha	June to May	March 2020-July 2020, September 2020, May 2021-July 2021, January 2022-February 2022	11
Puducherry	June to May	March 2020-December 2020, June 2021-August 2021	13
Rajasthan	May to April	March 2020- October 2020, December 2020-January 2021, Abril 2021, January 2022	12
Uttar Pradesh	April to March	March 2020-October 2020, April 2021-August 2021, January 2022	14

This table presents the states covered under ASER and included in our sample, along with their respective academic calendar months, months of COVID-19 school closures, and the number of months schools remained closed between March 2020 and August 2022. Panel A and B show the states classified as high and low performing, respectively. 'nmsc' denotes number of months of school closures.

Table A3: The Effect of State Performance on Test Scores: Parallel Trends for different subgroups

	zmath (1)	zread (2)	math (3)	read (4)
Panel A: Only women				
interaction_2018= high_state_performance*year2018	0.0189 (0.0428) [0.0268]	0.00989 (0.0351) [0.0234]	0.0304 (0.0436) [0.0271]	0.0240 (0.0440) [0.0292]
high_state_performance	0.742 (0.0604)*** [0.0644]***	0.306 (0.0944)*** [0.0712]***	0.754 (0.0569)*** [0.0663]***	0.379 (0.121)*** [0.0920]***
year2018	0.0139 (0.0392) [0.0150]	0.0233 (0.0307) [0.0138]*	0.0136 (0.0406) [0.0153]	0.0311 (0.0401) [0.0176]*
Observations	91,711	91,931	91,711	91,931
Adj R-squared	0.221	0.245	0.412	0.431
Mean Control 2016 high_state_performance==0	-0.256	-0.242	2.420	2.595
Panel B: Only men				
interaction_2018= high_state_performance*year2018	0.0165 (0.0504) [0.0250]	0.0200 (0.0362) [0.0241]	0.0259 (0.0517) [0.0257]	0.0336 (0.0482) [0.0306]
high_state_performance	0.516 (0.0626)*** [0.0700]***	0.299 (0.106)*** [0.0699]***	0.540 (0.0640)*** [0.0715]***	0.409 (0.141)*** [0.0893]***
year2018	0.00845 (0.0429) [0.0166]	0.0190 (0.0362) [0.0158]	0.00419 (0.0446) [0.0168]	0.0201 (0.0485) [0.0199]
Observations	88,449	88,623	88,449	88,623
Adj R-squared	0.181	0.194	0.399	0.398
Mean Control 2016 high_state_performance==0	-0.164	-0.238	2.509	2.591
Panel C: TV access at home				
interaction_2018= high_state_performance*year2018	0.0306 (0.0428) [0.0266]	0.0332 (0.0299) [0.0249]	0.0351 (0.0442) [0.0267]	0.0440 (0.0365) [0.0308]
high_state_performance	0.449 (0.0701)*** [0.0707]***	0.0472 (0.116) [0.0756]	0.516 (0.0651)*** [0.0703]***	0.119 (0.151) [0.0933]
year2018	-0.00168 (0.0328) [0.0189]	0.00920 (0.0261) [0.0182]	-0.00241 (0.0337) [0.0190]	0.0119 (0.0330) [0.0228]
Observations	86,140	86,306	86,140	86,306
Adj R-squared	0.162	0.169	0.399	0.408
Mean Control 2016 high_state_performance==0	0.063	0.071	2.750	3.004
Panel D: No TV access at home				
interaction_2018= high_state_performance*year2018	0.00919 (0.0494) [0.0396]	-0.0118 (0.0404) [0.0344]	0.0299 (0.0504) [0.0396]	0.0112 (0.0559) [0.0424]
high_state_performance	0.728 (0.0491)*** [0.0774]***	0.367 (0.0724)*** [0.0617]***	0.724 (0.0523)*** [0.0819]***	0.484 (0.0969)*** [0.0831]***
year2018	0.0149 (0.0487) [0.0161]	0.0260 (0.0409) [0.0148]*	0.0114 (0.0511) [0.0167]	0.0301 (0.0548) [0.0193]
Observations	92,603	92,824	92,603	92,824
Adj R-squared	0.167	0.175	0.370	0.367
Mean Control 2016 high_state_performance==0	-0.335	-0.382	2.334	2.406
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

The Effect of State Performance on Test Scores: Parallel Trends for different subgroups (continue)

	zmath	zread	math	read
	(1)	(2)	(3)	(4)
Panel E: Digital access at the village				
interaction_2018=	0.0321	0.0266	0.0401	0.0464
high_state_performance*year2018	(0.0526)	(0.0436)	(0.0529)	(0.0513)
	[0.0428]	[0.0408]	[0.0430]	[0.0500]
high_state_performance	0.0749	0.0772	0.0584	0.0599
	(0.124)	(0.143)	(0.125)	(0.183)
	[0.116]	[0.124]	[0.118]	[0.158]
year2018	-0.0216	-0.00643	-0.0278	-0.0159
	(0.0382)	(0.0331)	(0.0399)	(0.0432)
	[0.0285]	[0.0272]	[0.0282]	[0.0335]
Observations	31,001	31,079	31,001	31,079
Adj R-squared	0.218	0.236	0.418	0.426
Mean Control 2016	-0.148	-0.197	2.543	2.671
high_state_performance==0				
Panel F: No Digital access at the village				
interaction_2018=	0.0228	0.0164	0.0337	0.0303
high_state_performance*year2018	(0.0497)	(0.0371)	(0.0512)	(0.0489)
	[0.0245]	[0.0213]	[0.0254]	[0.0276]
high_state_performance	0.743	0.347	0.771	0.462
	(0.0466)***	(0.0905)***	(0.0501)***	(0.120)***
	[0.0568]***	[0.0576]***	[0.0602]***	[0.0752]***
year2018	0.0190	0.0276	0.0178	0.0352
	(0.0437)	(0.0353)	(0.0450)	(0.0464)
	[0.0146]	[0.0133]**	[0.0148]	[0.0170]**
Observations	149,159	149,475	149,159	149,475
Adj R-squared	0.197	0.216	0.401	0.412
Mean Control 2016	-0.222	-0.249	2.450	2.579
high_state_performance==0				
Panel G: At least one parent attended school				
interaction_2018=	0.0265	0.0218	0.0364	0.0366
high_state_performance*year2018	(0.0479)	(0.0339)	(0.0486)	(0.0433)
	[0.0248]	[0.0210]	[0.0252]	[0.0266]
high_state_performance	0.583	0.230	0.595	0.290
	(0.0682)***	(0.108)**	(0.0682)***	(0.142)*
	[0.0533]***	[0.0611]***	[0.0544]***	[0.0803]***
year2018	0.00577	0.0165	0.00338	0.0199
	(0.0418)	(0.0332)	(0.0431)	(0.0435)
	[0.0152]	[0.0135]	[0.0155]	[0.0173]
Observations	139,998	140,292	139,998	140,292
Adj R-squared	0.172	0.189	0.402	0.415
Mean Control 2016	-0.076	-0.101	2.586	2.755
high_state_performance==0				
Panel H: No parent attended school				
interaction_2018=	-0.0368	-0.0316	-0.0276	-0.0303
high_state_performance*year2018	(0.0789)	(0.0724)	(0.0801)	(0.0913)
	[0.0462]	[0.0523]	[0.0463]	[0.0671]
high_state_performance	0.798	0.540	0.847	0.743
	(0.0467)***	(0.0578)***	(0.0547)***	(0.0699)***
	[0.0985]***	[0.0820]***	[0.0996]***	[0.102]***
year2018	0.0344	0.0412	0.0336	0.0516
	(0.0438)	(0.0381)	(0.0461)	(0.0514)
	[0.0187]*	[0.0190]**	[0.0196]*	[0.0246]**
Observations	40,162	40,262	40,162	40,262
Adj R-squared	0.149	0.145	0.343	0.326
Mean Control 2016	-0.516	-0.553	2.182	2.224
high_state_performance==0				
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

This table shows the estimates of the effect of state performance on math and reading raw and z scores (baseline 2018) during pre-COVID years. The sample includes children aged 5 to 12 in grade 1 to grade 5 who go to public school in 2016 and 2018. *high_state_performance* is an indicator equals to 1 if the state average combined (math and reading) test score is above the sample median of 5.94, and zero otherwise. *year*2018 is an indicator equals to one for survey year 2018 and zero, otherwise. Control variables include father education, mother education, public school indicator, indicator for test in Hindi, internet cafe in the village, electricity in the village, computer use at home, tuition cost, electricity connectivity at home, school class, child's gender and child's age. Standard errors clustered at the state level in parentheses, standard errors clustered at the district level in squared brackets.
*** p<0.01,** p<0.05, * p<0.1.

Table A4: Simple Difference-in-Differences: 2016 and 2018

			(1)
			z math
	High(A)	**Low(B)**	**Δ(A-B)**
Cohort 2018	0.409	-0.022	0.431
Cohort 2016	0.376	-0.053	0.429
Δ(Cohort2018-Cohort2016)	0.033	0.031	
		Diff-in-Diff	0.002
			(2)
			z read
	High(A)	**Low(B)**	**Δ(A-B)**
Cohort 2018	0.360	-0.084	0.444
Cohort 2016	0.302	-0.123	0.425
Δ(Cohort2018-Cohort2016)	0.058	0.039	
		Diff-in-Diff	0.019

This table presents a simple difference-in-differences calculation of the main outcomes. The sample includes children aged 5 to 12 in grade 1 to grade 5 who go to public school and with ASER test scores corresponding to year 2018 (Cohort 2018) and to year 2016 (Cohort 2016). High (A) represents high performing states, while Low (B) represents low performing states. A test for significant difference (2018-2016) is performed. The difference in test scores between high and low performing states is not statistically significant at 5% level. *** p<0.01,** p<0.05, * p<0.1.