THE
ECONOMIC
JOURNAL

The Economic Journal, **136** (*May*), 1254–1279 https://doi.org/10.1093/ej/ueaf089
Advance Access Publication Date: 27 September 2025

THE PRODUCTIVITY OF PUBLIC AND PRIVATE PRESCHOOLS (AND SCHOOLS): EVIDENCE FROM INDIA*

Petter Berg, Mauricio Romero and Abhijeet Singh

We study the relative productivity of private and public institutions at the preschool and primary school levels using panel data from 215 villages in Tamil Nadu. Private preschools show higher test score value-added in math and language (~0.59–0.74σ) and outperform government providers in nearly all villages. This productivity difference explains 60% of the socioeconomic test score gap before school entry. These results contrast starkly with primary schooling, where we find no evidence of a private-sector premium in math and negative effects in local language. Test score value-added is positively correlated between private and government options in a village, both at the preschool and primary school levels. Quality is also correlated across levels; villages with more productive primary schools also tend to have more productive preschools. Our findings inform debates on achieving universal foundational skills and highlight the need to improve the quality of preschools available to lower-income families.

JEL codes: H44, H52, I21, I25, I28, L10, L33

Private providers account for a substantial share of education services in low- and middle-income countries (LMICs). They are particularly important in preschool services where, globally, they account for ~37% of children enrolled in pre-primary institutions, compared to 19% of enrolment at the primary level in LMICs (UNESCO, 2021; 2022). Yet, evidence comparing the productivity of private to public preschools using broad samples remains scarce. This omission is in stark contrast to numerous studies on private primary school effects (Crawfurd *et al.*, 2024), and is especially surprising given broad recognition of the importance of early childhood education (Elango *et al.*, 2015; Holla *et al.*, 2021), the relative importance of private providers in the sector and international policy targets to universalise quality preschool services (see Sustainable Development Goals Target 4.2).

* Corresponding author: Petter Berg, Department of Economics, Stockholm School of Economics, Sveavägen 65, Sweden. Email: petter.berg@hhs.se

This paper was received on 10 September 2024 and accepted on 15 September 2025. The Editor was Basit Zafar.

The data and codes for this paper are available on the Journal repository. They were checked for their ability to reproduce the results presented in the paper. The authors were granted an exemption to publish parts of their data because access to these data is restricted. However, the authors provided the Journal with temporary access to the data, which enabled the Journal to run their codes. The codes for the parts subject to exemption are also available on the Journal repository. The restricted access data and these codes were also checked for their ability to reproduce the results presented in the paper. The replication package for this paper is available at the following address: https://doi.org/10.5281/zenodo.17119441.

We are grateful for comments from Jishnu Das and Ajinkya Keskar. Ramamurthy Sripada, Dwithiya Raghavan, Akanksha Soni and Indira Tirumala provided excellent field management and research assistance. We received IRB approval from the Institute for Financial Management and Research (#113401). Financial support for data collection from the Asian Development Bank Institute and the RISE Programme—funded by the UK's Foreign, Commonwealth and Development Office, Australia's Department of Foreign Affairs and Trade, and the Bill and Melinda Gates Foundation—is gratefully acknowledged. Abhijeet Singh acknowledges funding from the European Research Council Starting Grant (Grant ID 101076831, EDINCLUSION). M.R. gratefully acknowledges financial support from the Asociación Mexicana de Cultura, A.C. and the Jacobs Foundation. All errors are our own.

The authors declare that they have no competing interests to disclose.

In this paper, we focus on understanding the relative productivity of private and public preschools using a new child-level panel dataset of ∼19,000 children aged 3–10 years in 215 villages in the Indian state of Tamil Nadu. Importantly, the data include tests of student achievement in math and the local language (Tamil), using age-appropriate tests administered at home, as well as multiple measures of socioeconomic status and educational inputs; the test scores are vertically equated using item response theory models to be comparable over time and across ages. Although preschool is not compulsory in the setting we study, nearly all children are enrolled at age four. Primary school enrolment is compulsory and near-universal by age six. In our setting, private providers account for roughly one-third of enrolment at preschool age and one-quarter in primary schooling.

We use these data to conduct three exercises. First, following early studies of private schooling in South Asia (Andrabi *et al.*, 2011; Singh, 2015), we estimate the average private institution effect at the preschool and primary school stages using value-added models of student achievement. In addition to providing novel estimates at the preschool level, we can benchmark these estimates to primary school effects in our sample and to previous work. Second, we assess the extent to which the differential productivity of private institutions, combined with differential enrolment in these institutions by socioeconomic status (SES), can explain SES gaps in cognitive achievement. This reflects longstanding concerns about unequal access to quality early childhood education and preschooling. Finally, following recent evidence of substantial heterogeneity within and across markets in primary schooling (Andrabi *et al.*, 2025), we then allow the estimated private institution value-added to vary across villages and levels. This allows us to understand variation in the quality of public and private educational provision across villages.

We document three sets of results. The first of these focuses on the premium of the average private institution. Compared to public options, the private preschool premium in test score value-added is substantial, at 0.74 and 0.59 SDs (σ) of the test score distributions in math and Tamil language. This premium is roughly double the (cross-sectional) achievement difference in language between three- and four-year-olds enrolled in public preschools, and roughly four times the equivalent difference in math. These effects exceed those typically achieved by early childhood development interventions documented in the literature (e.g., Attanasio *et al.*, 2014; 2022; Andrew *et al.*, 2024). In contrast, we find no evidence of a positive effect for private primary schools: value-added in private schools is indistinguishable from that of government options in math and significantly lower in the local language (by $\sim 0.17\sigma$), results that are similar to those reported by Muralidharan and Sundararaman (2015) and Singh (2015) in neighbouring states.[1]

Second, we quantify the extent to which greater private-sector enrolment for students from wealthier households explains the test score gap between them and students from poorer households. Students from households in the top socioeconomic quartile are 35 percentage points more likely to attend private preschools and 39 percentage points more likely to attend private primary schools than students in the bottom quartile. This differential private-sector enrolment accounts for approximately 60% of the SES test score gap at preschool ages, but, reflecting the absence of average private school effects in primary schools, does not explain SES gaps at later ages.

Third, focusing on *village-specific* estimates, the value-added in government preschools is dominated by that of private options in virtually *all* villages. In both subjects, there is very limited common support in the distributions of value-added between government and private preschools. This is consistent with prior evidence that government care centres typically provide

[1] The negative effect in the local language likely reflects a greater focus on English over the local language in private primary schools (Muralidharan and Sundararaman, 2015; Singh, 2015).

little structured cognitive stimulation (Ganimian *et al.*, 2024), whereas private preschools focus much more on early childhood education (Singh, 2014; Dean and Jayachandran, 2019). In contrast, there is near-complete overlap in value-added in math and substantial overlap in Tamil at the primary school level. Public- and private-sector value-added estimates are positively correlated within villages: an increase of 1σ of student achievement in the value-added of public preschools in a village predicts 0.61–0.80σ higher private-sector productivity; in primary schooling, this correlation is between 0.88 and 0.99. Value-added estimates are also positively correlated across levels in the same sector in a village—i.e., higher value-added in government (private) preschools also predicts higher value-added in government (private) primary schooling.

Taken together, our analyses provide a unified treatment of the productivity of private and public options, and their link to socioeconomic inequality, over an extended span of education from preschool to the completion of primary schooling. These empirical results contribute to multiple distinct strands of economics research.

First, they contribute to an active literature that focuses on the development of cognitive skills in early childhood. Estimates suggest that over 250 million children under five, mostly in LMICs, do not fulfil their cognitive potential (Engle *et al.*, 2007; Grantham-McGregor *et al.*, 2007; Behrman *et al.*, 2014; Black *et al.*, 2017). In response, much of the economics literature has, appropriately, focused on evaluating novel interventions to address these deficits.[2] In contrast, we aim to identify differences in productivity between the options *currently used* by millions of children. This provides useful complementary information about the institutions where novel interventions may be most needed (in our context, public preschools), as well as the potential for achieving cognitive gains through reallocation of children across providers in the same village (e.g., through vouchers).

Second, these results contribute to a substantial literature examining the private sector in education.[3] This literature has mostly focused on primary schools; we contribute by examining preschools, a shift in focus that is particularly important in the Indian context. Existing estimates, which we confirm in our data, suggest that average private school effects in foundational math and local language are small at the primary level (Muralidharan and Sundararaman, 2015; Singh, 2015). In contrast, we show that the private-public margin is more consequential at the preschool level. These results are directly informative for current policy initiatives aimed at ensuring universal foundational literacy and numeracy in childhood, both in India and globally (see, e.g., World Bank, 2017; Muralidharan and Singh, 2021).

Relatedly, we contribute to the literature on socioeconomic inequality in early childhood cognitive skills. This primarily descriptive stream of research has documented the existence and evolution of disparities in the cognitive skills of young children from more- and less-advantaged backgrounds.[4] While differential access to effective preschools is a plausible contributor to these

[2] See, for example, evaluations of programs to support parents (e.g., Attanasio *et al.*, 2014; 2022; Andrew *et al.*, 2024), to improve public preschools (e.g., Evans *et al.*, 2024; Ganimian *et al.*, 2024) or to send children to private preschools (e.g., Dean and Jayachandran, 2019; Bjorvatn *et al.*, 2025).

[3] See Crawfurd *et al.* (2024) for a recent review in LMICs. Influential studies from LMICs include multiple studies using the LEAPS dataset in Pakistan (Andrabi *et al.*, 2011; 2024; 2025; Carneiro *et al.*, 2024), as well as studies by Muralidharan and Sundararaman (2015) and Singh (2015) in the state of Andhra Pradesh, India, and evaluations of private-public partnership initiatives in Pakistan (Barrera-Osorio *et al.*, 2022) and Liberia (Romero *et al.*, 2020; Romero and Sandefur, 2021).

[4] See, e.g., Engle *et al.* (2011); Fernald *et al.* (2012); Elango *et al.* (2015); Rubio-Codina *et al.* (2015); Schady *et al.* (2015) and Reynolds *et al.* (2017). The analogous literature in the United States studies racial disparities (Fryer and Levitt, 2004; 2006; 2013), as well as the income-achievement gap (Reardon, 2011; 2021; Nielsen, 2023). Our findings are consistent with evidence from high–income settings showing that SES gaps emerge before formal schooling (e.g.,

disparities, drawing a conclusive link has been difficult. Our main contribution here is to show that differential access to private preschools, which offer higher value-added, is a substantial driver of the socioeconomic disparities in achievement observed before school entry age.

Finally, our results also contribute to the literature on education markets. Specifically, our results on the near-universal productivity advantage of private preschools over public alternatives across villages (unlike in primary schools), as well as market-level correlations in the productivity of providers across sectors and educational stages, are both novel in the context of LMIC education systems. Although data limitations constrain us from attempting a comprehensive analysis of preschool markets—we do not observe provider-specific quality, prices, attributes and actions in these villages—these results suggest that understanding both the demand and supply sides of educational markets is as important for preschool services as it has been for our understanding of primary schooling in LMICs.[5]

1. Context and Data

1.1. *Context*

Our study is based in rural areas of Tamil Nadu, a large south Indian state with an estimated population of 74 million and an education system that serves 13 million children annually (Government of India, 2019). While Tamil Nadu's public early childhood education system is considered high performing within India, substantial quality gaps remain: fewer than 10% of children can read individual words when they enter primary school, and 60% cannot recognise individual letters (Pratham, 2022).

Primary education in India is mandatory, and enrolment is nearly universal (Pratham, 2022). The official school starting age is six, though many children (including in Tamil Nadu) begin Grade 1 at age five. Government primary schools are free and provide in-kind benefits, including midday meals, textbooks and uniforms. Instruction is usually in the state's official language (Tamil in Tamil Nadu). In contrast, private primary schools charge fees and often use English as the medium of instruction, and there is substantial heterogeneity in quality and costs (Singh, 2014; Kingdon, 2020). In 2022, private providers accounted for roughly one-quarter of primary school enrolment in rural areas nationally; this figure was similar in rural Tamil Nadu, where they accounted for ∼24% of enrolment (Pratham, 2022).

Preschool education (ages 3–5) is not compulsory. The main public option is the network of *anganwadi* centres under the Integrated Child Development Services (ICDS) program. ICDS anganwadis constitute the world's largest early childhood program, offering free pre-primary education alongside nutrition and health services to roughly 36 million children aged 3–6 at 1.35 million centres nationwide (Ganimian *et al.*, 2024).[6] Enrolment is non-selective and free of charge. Anganwadis are typically staffed by a single anganwadi worker (and a helper) per centre, who is responsible for delivering a mix of preschool instruction, child nutrition and health

Hart and Risley, 1995; Lee and Burkam, 2002; Noble *et al.,* 2005; Fernald *et al.*, 2013), although a large share in our setting is accounted for by the type of preschool attended.

[5] For influential examples of such work in LMICs in primary schooling, see, e.g., Andrabi *et al.* (2017), Allende (2019), Neilson (2021) and Bau (2022).

[6] Throughout this paper, we refer to all forms of structured pre-primary enrolment as preschooling. We are only investigating differences across sectors (public/private), rather than across individual providers or facilities. Hence, we are unable to account for the potential differences *within* sectors, such as between *anganwadi* centres and formal, public preschools (which account for 2% of enrolment of four-year-olds; Pratham, 2022). To avoid introducing additional terminology, we therefore refer to these simply as 'preschools'.

monitoring services. Anganwadi centres in Tamil Nadu typically operate for about 4 hours each morning, of which 2 hours are nominally allotted to pre-school education. In practice, however, only 38 minutes are devoted to instruction per day due to staff constraints (Ganimian *et al.*, 2024).

Private preschools, on the other hand, charge fees and focus on pre-primary instruction in nursery and kindergarten classes, often as part of or attached to private primary schools. These private pre-schools tend to emphasise early academic skills, have more structured instructional time and frequently introduce English (Dean and Jayachandran, 2019). Overall, in India, preschool enrolment has been rising (and reached 80% for four-year-olds in 2022; Pratham, 2022), but varies widely by state, with Tamil Nadu having nearly every four-year-old enrolled (over 99%) in some educational facility.[7] Private providers account for roughly 20% of preschool enrolment nationally (and 36% in Tamil Nadu).

The annual per-child expenditure differs both across sectors and levels. In 2021, private schools in our study districts charged an average annual fee of INR 7,100 (∼96 USD) at the preschool level and INR 8,460 (∼114 USD) at the primary school level.[8] Per-pupil expenditure in public elementary schools is much higher (∼INR 29,000 (∼391.5 USD) in 2019–20; Bordoloi *et al.,* 2020). The cost in public preschools per child per year was ∼76 USD (∼5,630 INR) across all of India in 2020 (UNESCO, 2020a,b,c,d,e,f). These differences in per-child public expenditure across primary school and preschool reflect substantially lower salaries paid to anganwadi helpers compared to public school teachers and constrained resources in the public preschool sector.

1.2. *Data*

1.2.1. *Sample*

Our data cover 215 villages in four districts of Tamil Nadu (see the map in Online Appendix Figure A.1). These districts were selected using probability-proportional-to-size sampling to ensure representativeness of rural Tamil Nadu. In these communities, we administered comparable achievement tests to students aged 3–10 in early 2022 (baseline) and 2023 (endline). Our core analysis sample includes the set of children aged 4–10 in 2023 for which we have access to baseline assessments (N = 19,021).[9] Although these villages were not randomly selected (the sample only includes blocks with two or more government preschool centres (*anganwadis*) co-located with middle schools), our sample is similar in observable characteristics to the state's rural population, though slightly less wealthy in terms of asset ownership (see Online Appendix Table A.1).[10]

Our sampling strategy involved enumerating every household located within approximately 2 km of the village *anganwadi*, designated as the reference point for data collection. Villages

[7] Near universal enrolment is not unique to Tamil Nadu: rates of non-enrolment among four-year-olds are under 2% in several other states, including Andhra Pradesh, Telangana and Karnataka in South India, and Maharashtra, Jharkhand and Odisha in other parts of the country (Pratham, 2022).

[8] These estimates average fees across private preschools in the four study districts using from data from the Tamil Nadu Private Schools Fee Determination Committee (https://tnfeecommittee.com/). We use 0.0135 USD/INR as the exchange rate, which was the average rate in 2021.

[9] Attrition between survey rounds was ∼25%, but does not vary by SES or test scores (see Online Appendix B). The survey waves were administered slightly more than a year apart. Since we only have data on age in completed years (and not months), we measure attrition for children aged 3–9 in 2021/2022. All our results are robust under specifications using inverse probability weighting (see Online Appendix B.3).

[10] These data were originally collected for an experimental evaluation of a government program to improve preschool education. The intervention and the evaluation were cancelled due to the COVID-19 pandemic and subsequent preschool and school closures. See https://doi.org/10.1257/rct.5599 for more details. We continued to collect data to study the learning loss during the pandemic and the pace of recovery afterward (Singh *et al.*, 2024).

typically consist of multiple distinct clusters of households and, in practice, either an entire cluster was included in the sample or not. Online Appendix Figure A.2 illustrates the sample in four villages to provide a sense of the coverage of this sampling strategy in practice. In most villages, we enumerated all households. As a result, while we claim that our villages provide examples of disjoint markets, with complete enumeration in sampled geographical areas, we do not claim to have a full enumeration of all students and providers in these markets.[11]

1.2.2. *Assessments*

The learning assessments were designed to capture student achievement across the preschool and primary populations. Children were tested in math and Tamil (the local language) using age-appropriate booklets and overlapping items. For preschoolers (ages 3–4), the tests captured oral comprehension, letter recognition, quantitative comparisons, number recognition and counting; at ages five and six, they also included word recognition, more complex counting and basic addition; for children aged 6–10, the tests additionally included more complex arithmetic computation and word problems in math and passage comprehension and reading exercises in Tamil.

Our focus on foundational math and local language skills closely reflects the targets of the National Initiative for Proficiency in Reading with Understanding and Numeracy (NIPUN), the national educational policy on early childhood education (Ministry of Education, 2020). Online Appendix Figure A.3 provides an overview of skills that are targeted by the policy at different ages from preschool through to early primary grades. Our assessments, thus, map directly to stated policy goals while also including broader competences that were either included in Ganimian *et al.* (2024) or are standard in evaluations of early primary education in this context.

Test booklets included common items across waves and ages, which allows us to link achievement on a common metric using item response theory (IRT) models (Das and Zajonc, 2010). We estimate these scores by pooling all test observations across rounds, separately for math and Tamil. We standardise test scores to have mean zero and an SD of one in the sample of children aged five in the 2022 survey wave. The test scores display, as expected, a shifting in the distribution of achievement with age (indicating skill acquisition over time; see Online Appendix Figure G.7) and no evidence of differential item functioning by age or by round (see Online Appendix Figures G.8–G.29). See Online Appendix G for more details on the test and the validation exercises we conduct.[12]

1.2.3. *Household survey*

We collected extensive household data about their socioeconomic status and children's education in both survey waves. We use detailed information on household ownership of various assets in 2022 to construct a socioeconomic status index using principal component analysis (PCA). We use a household's percentile rank in this index as the primary measure of SES (see Online

[11] In very dispersed villages, our sampling strategy omits clusters that are distant from the reference school. More generally, in this setting, an education market can span several villages. In data from 2019, ∼20%–25% of children, even at preschool level, go to school using school vans and, in 2023, nearly a quarter report attending school further than 2 km away. Our estimates are, therefore, best interpreted as the productivity of educational institutions accessed by children in our sample geographies, rather than only those providers located within the boundaries of our enumeration area.

[12] In 2022, we divided the sample into two randomly assigned groups within villages that administered the tests in a staggered manner between December 2021 and April 2022 (Singh *et al.*, 2024). We administered the tests with a similar staggering and the same assigned groups in 2023 to maintain a similar gap between assessments. In 2023, we observed signs of ceiling effects for school-age children in the first round of testing. We remedied this by adjusting the test booklets for the second (randomly assigned) testing round, retaining common items for linking (see Online Appendix Figures F.1–F.4). No results are sensitive to only using the second round (see Online Appendix Table F.1).

Appendix C for details). We also use maternal and paternal education information as additional measures of SES and control for them in our regressions.

1.2.4. *Strengths and limitations of the data*

Our dataset has several important strengths. The most important of these is the availability of panel data on achievement for children over the full span of preschool and primary school ages across a large number of spatially disjoint education markets. A related strength is the comparability of measurement over time and across ages: datasets with vertically linked IRT scores are uncommon in low- and middle-income countries, but are crucial to our goals of expressing preschool and primary school productivity on a common scale. The final strength is the complete enumeration of households in sampled geographical areas: this prevents the attrition typical in many school-based surveys due to student absence on the day of testing.[13]

However, the expansiveness of the dataset also imposes some trade-offs that limit our analysis. Most importantly, we can only provide sector-specific estimates for a village as a whole, rather than estimates for each facility separately. This issue arises for a combination of reasons. First, matching children to centres generates substantial measurement error since we only collected facility names (which are hard to map to individual facilities, especially government facilities that do not have distinctive names). Second, because the total number of children in each facility is often very small, any individual facility estimates would be very noisy even with complete matching. Finally, in large villages, we restricted our censuses to a radius of ∼2 km from a reference point (the anganwadi); this does not affect our interpretation of each village as a distinct market, but does affect our ability to interpret our survey as a complete enumeration of the full market.

Second, since we did not collect detailed facility surveys, we only have information about whether a child attends a private or public preschool, but no other characteristics, such as staffing, fees or instructional practices. This prevents us from decomposing the sources of any private/public advantage (as in, e.g., Angrist *et al.,* 2013 for Charter schools)—note, however, that even with extensive data on school inputs and characteristics, Andrabi *et al.* (2025) could only explain 5% of the heterogeneity in value-added across private schools in Pakistan. Finally, since our data collection focused on foundational math and local language skills, the explicit targets of government policy, we did not administer tests for English language skills (an important differentiator for private schools).

2. Pre- and Primary School Choices and Value-Added

The first part of our analysis focuses on selection into and average productivity differences across different schooling options.

2.1. *Selection and Educational Trajectories*

We investigate enrolment patterns and child characteristics by age in Table 1. Virtually all children (∼96%) are enrolled in private preschools or public care centres (*anganwadis*) at age four. At

[13] For instance, ASER reports indicate that student absence ranges from 10% to 45% in different states of India (Pratham, 2022). This absence-induced attrition is non-random: it is typically higher in the public sector and for children with lower test scores and from poorer households.

Table 1. *Child and Household Characteristics by Age and Enrolment Status.*

		Preschool			Primary school		
	No school	Public	Private	Diff.	Public	Private	Diff.
Panel A: age 4							
Female	0.55	0.48	0.45	−0.03	0.43	0.65	0.21
Mother educ.: <Gr. 9	0.26	0.24	0.16	−0.08***	0.27	0.12	−0.15
Mother educ.: ≥Gr. 12	0.29	0.35	0.60	0.25***	0.46	0.53	0.07
SES percentile	49.11	47.20	65.60	18.40***	55.78	66.18	10.39
Math IRT score in 2022	−1.34	−1.27	−1.10	0.17***	−0.73	−0.65	0.07
Tamil IRT score in 2022	−1.55	−1.36	−1.18	0.17***	−0.50	−0.67	−0.17
Share of students	0.04	0.61	0.33		0.02	0.01	
Observations	87	1,349	726		37	17	
Panel B: age 5							
Female	0.46	0.51	0.46	−0.05	0.51	0.50	−0.01
Mother educ.: <Gr. 9	0.28	0.23	0.16	−0.07***	0.25	0.15	−0.10***
Mother educ.: ≥Gr. 12	0.40	0.34	0.57	0.24***	0.31	0.56	0.25***
SES percentile	43.98	45.25	67.74	22.49***	43.87	66.58	22.71***
Math IRT score in 2022	−0.99	−1.04	−0.66	0.38***	−0.81	−0.36	0.45***
Tamil IRT score in 2022	−1.07	−1.08	−0.68	0.41***	−0.85	−0.36	0.49***
Share of students	0.02	0.24	0.20		0.39	0.15	
Observations	50	683	592		1,132	434	
Panel C: ages 6–10							
Female	0.44	0.40	0.34	−0.06	0.51	0.44	−0.07***
Mother educ.: <Gr. 9	0.26	0.00	0.12	0.12	0.31	0.15	−0.16***
Mother educ.: ≥Gr. 12	0.21	0.20	0.52	0.32**	0.24	0.51	0.27***
SES percentile	46.87	44.47	68.45	23.98***	42.66	63.25	20.60***
Math IRT score in 2022	0.29	−0.30	−0.44	−0.14	0.62	0.92	0.30***
Tamil IRT score in 2022	0.19	−0.17	−0.48	−0.32	0.71	0.83	0.12***
Share of students	0.00	0.00	0.00		0.73	0.26	
Observations	39	15	67		10,465	3,797	

Notes: This table reports average differences in child and household characteristics by type of enrolment, separately by age, in three panels. The types of enrolment are no school and private/public pre-/primary school. Columns 4 and 7 show the difference between children in the private and public sectors, respectively. Virtually all children attend preschool at age four. At age five, children start transitioning into primary school. Between the ages of six and ten, virtually all are enrolled in primary school. *** $p < .01$, ** $p < .05$.

age five, roughly half of all children begin enrolling in primary school.[14] Between the ages of six and ten, primary school enrolment becomes nearly universal.

Private operators serve a significant portion of the market in both pre- and primary schools. At age four, a third of all children are enrolled in a private preschool. The market share of private providers reduces to one-quarter in primary school.

There is a significant SES gap in private enrolment. Private school children are about 25% more likely to have mothers with completed secondary education. The average child in private preschool ranks 18 percentiles higher in the socioeconomic distribution compared to children in public preschools. In primary school, this gap increases to around 21 percentiles. Additionally,

[14] This pattern, of a substantial fraction of students already choosing to enrol in Grade 1 at age five (although the official age to start formal schooling is supposed to be six years), is common in most Indian states (Pratham, 2022). The ubiquity of this empirical pattern motivates our attempt to understand the relative productivity of the four types of institutions—public and private preschools and primary schools—on a comparable scale, since each of these is a viable choice for students between 4–6 years old, and especially so for five-year-olds.

there is a clear gender gap in private enrolment at the primary school level, amounting to 7 percentage points in favour of boys.

We also document a substantial gap in test scores, as measured in our baseline assessments, between children from private and public schools. For students aged four in 2023, this gap amounts to 0.17σ in the previous year's test score distribution in both math and Tamil. At age five, the gap increases to around 0.45σ in both subjects. During the main primary school ages (6–10), the gap in Tamil reduces to 0.12σ, but remains large in math (0.3σ).

2.2. *Estimating Value-Added by Sector*

We rely on conventional value-added models to measure test score improvements from attending a private pre- and primary school. Specifically, test scores are regressed on school characteristics (e.g., private/public indicators) while conditioning on lagged scores and student sociodemographic characteristics to account for student selection (see, e.g., Todd and Wolpin, 2003; 2007).

For each subject (math and Tamil), we estimate the equation

$$y_{iv}^{2023} = \lambda y_{iv}^{2022} + \beta Private_i + \Gamma \mathbf{X}_{iv} + \epsilon_{iv}, \quad (1)$$

where i denotes a child and v a village. The variable y_{iv}^{t} denotes student i's test score in a particular subject in year t, λ is the coefficient on lagged test scores (i.e., a persistence parameter) and $Private_i$ is an indicator for whether student i attended a private preschool or primary school between the assessment waves. Here $\mathbf{X}_{iv}$ is a vector of additional controls, including village fixed effects, deciles of the SES wealth index, paternal and maternal education levels and the child's gender. The coefficient of interest is β, which captures the effect of attending a private preschool/school on test scores. This benchmark specification is similar to the dynamic OLS specifications used by Andrabi *et al.* (2011) and Singh (2015), the two most closely related papers on private formal schooling in South Asia.

We estimate this equation separately for children aged 4, 5 and 6–10 years. This choice reflects multiple distinct concerns. Most importantly, given the enrolment patterns in Table 1, estimates of β at ages 4 and 6–10 can be interpreted as the private school premia at the preschool and primary school levels; at age five, in contrast, β represents a weighted average of the private premia at the two levels.[15] Second, this allows the coefficient on lagged achievement to differ across the preschool and schooling stages; this is potentially important because both true persistence in achievement and measurement error in test scores may plausibly differ for children at very young ages. In Section 2.4 below, we further relax the assumption of a common persistence parameter (λ), allowing λ to differ for every age, and find similar results. Finally, by allowing the coefficients on all variables in $\mathbf{X}_{iv}$ to differ across ages, this also provides flexibility in case the nature of selection across sectors, influenced by factors such as parental education or wealth, differs across levels of education.

The causal interpretation of β relies on a conditional exogeneity assumption: our estimates will be unbiased only if our controls are rich enough to account for the selection of children into private versus public operators. In other settings, both in the United States and in LMICs, similarly estimated value-added measures appear to agree closely compared to estimates using identification from design-based experimental or quasi-experimental variation (Andrabi *et al.*,

[15] A few children ($N = 54$) at age four are recorded as attending primary school (Table 1). This likely reflects response errors, and we do not exclude these children in our analyses. Excluding them has very little impact on any of the main results. The same holds for children aged six and above recorded as attending preschool ($N = 82$).

Table 2. *Private School Value-Added in Preschool and Primary School.*

	Age 4		Age 5		Ages 6–10	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: math						
Private school	0.807***	0.737***	0.322***	0.180***	0.125***	−0.005
	(0.047)	(0.051)	(0.044)	(0.048)	(0.022)	(0.019)
Math IRT score in 2022		0.177***		0.242***		0.306***
		(0.028)		(0.026)		(0.010)
Panel B: Tamil						
Private school	0.640***	0.588***	0.160***	0.053	−0.098***	−0.172***
	(0.052)	(0.054)	(0.045)	(0.047)	(0.023)	(0.022)
Tamil IRT score in 2022		0.199***		0.207***		0.338***
		(0.027)		(0.024)		(0.011)
Controls	Village FEs	All	Village FEs	All	Village FEs	All
Observations	1,839	1,839	2,841	2,841	14,344	14,344

Notes: Robust SEs, clustered at the village level, are reported in parentheses. Columns (1), (3) and (5) show raw means by private school attendance within villages. Columns (2), (4) and (6) include lagged scores, village fixed effects, and controls for deciles of the SES wealth index, paternal and maternal education and child gender. Test scores refer to equated IRT scores, standardised with respect to children aged five in the 2022 assessments. *** $p < .01$.

2011; 2025; Singh, 2015; 2020; Angrist *et al.*, 2017; 2023).[16] Thus, although the nature of selection may differ across settings, the value-added estimates in our data likely reflect true productivity differences rather than selection effects. Furthermore, we examine the potential for bias through two (related) strategies in Section 2.4 below. First, analogously to the investigation of bias in teacher value-added estimates in Chetty *et al.* (2014), we show that our estimates are invariant to the inclusion of multiple additional controls. Second, we report sensitivity analyses following Cinelli and Hazlett (2019) to examine the strength of the unobserved confounding needed to overturn our results.

2.3. *The Private Premium in Preschool and Primary School*

Table 2 reports the estimated private premium (β) from (1), separately for children aged 4 (preschool), 5 (transition) and 6–10 (primary school). Columns (1), (3) and (5) report differences in test scores by private school attendance conditional only on village fixed effects; columns (2), (4) and (6) further condition on lagged scores and covariates as specified in (1).

Raw differences in academic achievement by private school attendance are substantial at the preschool level, amounting to 0.81 SDs of the baseline test score distribution (σ) in math and 0.64σ in Tamil (column (1)). Most of these gaps reflect stark productivity differences between sectors: conditional on lagged test scores and socioeconomic characteristics, the average private

[16] Since our identifying assumption does not require assignment to *individual school units* to be random, it is weaker than the assumption underpinning school value-added models validated across multiple contexts (Angrist *et al.*, 2024; Andrabi *et al.*, 2025). Imagine a village with two public schools A and B and two private schools X and Y. Typical school value-added models require enrolment into *each* of these options to be conditionally ignorable; our specifications only require sector-level enrolment to be conditionally ignorable (i.e., enrolment in either of school A or B versus school X or Y, but not the choice of school within sector).

premium in the preschool market is 0.74σ in math and 0.59σ in Tamil (column (2)).[17] This is equivalent to almost *twice* the raw difference in Tamil achievement between children aged three and four in public preschools in our endline assessments, and *four times* the difference in math.[18]

These patterns differ substantially at the primary school level (columns (5) and (6)). The private premium in math is virtually zero (column (6)). In Tamil, the 'premium' at the primary level is negative, which likely reflects a greater focus on English teaching in private schools. Patterns in both subjects are similar to previous estimates of private primary school effects in India (Muralidharan and Sundararaman, 2015; Singh, 2015).[19]

We examine these value-added estimates semi-parametrically in Figure 1, following Cattaneo *et al.* (2024). Specifically, we adjust for the full set of covariates in (1) and allow the relationship between lagged achievement and subsequent test scores to vary non-linearly.[20] We plot these semi-parametric estimates separately for students in private and public sectors in each of the three age categories in Table 2. In each sub-figure, the distance between the estimated fits for the private and government sectors provides an analogue to the estimated private school premium for students in that segment of the achievement distribution. In both subjects, the private school premium at age four is present across most of the achievement distribution and not substantively different in magnitude for students with differing achievement levels. In contrast, and reflecting the results in Table 2, for ages 6–10, there is no premium across the baseline achievement distribution in math, and a negative one in Tamil. Figure 1 shows that the estimated premia in Table 2 are not sensitive to controlling for lag scores more flexibly or a potential lack of common support in achievement between the public and private sectors.[21]

Our core analyses identify the difference in productivity between public and private options; determining their absolute levels would require comparing them to no enrolment. The latter margin is less relevant in Tamil Nadu, as preschool enrolment is nearly universal, but remains

[17] Since children aged five are a mix of pre- and primary school students, their private premium is positive, but muted, compared to their younger peers. If we focus on children aged five—of which around half will have started primary school—and allow private school effects to differ at the pre- and primary levels, we obtain very similar estimates as for children aged 4 and 6–10 (see Online Appendix Table A.2).

[18] In Online Appendix Tables A.3 and A.4, we divide items by competencies being assessed. In math, private preschools increase the proportion correct on test items by 19–33 percentage points, relative to public school averages of 19%–58%. In Tamil, this figure is 10–22 percentage points, relative to public school averages of 36%–50%. These effects are largest, both in absolute and relative terms, in competencies where public preschool children are particularly weak.

[19] We also estimated the private premium separately for each age within primary grades (see Online Appendix Figure A.4). The premium is zero across all ages for math, except for nine-year-olds, for whom it is positive and statistically significant. For Tamil, the point estimate for the premium is negative for all ages, but closer to zero and statistically insignificant for nine- and ten-year-olds. The similarity between the results for six-year-olds and the overall primary sample (ages 6–10) is reassuring. Older cohorts faced substantial educational disruptions and learning losses due to COVID-19 (Singh *et al.*, 2024), which could potentially affect relative sectoral productivity if private schools coped with these disruptions more (or less) effectively than public schools. However, disruptions were less severe for six-year-olds; yet, their estimated private premium closely matches that of the full primary sample, suggesting limited bias from pandemic-related disruptions.

[20] We use within-age percentiles of lagged achievement in 2022 to ensure an even distribution of the sample.

[21] Even with significant mean differences, there is full common support between the distribution of prior achievement in the two sectors (Online Appendix Figure D.1). This is important because, if segments of the achievement distribution were unique to each group then controlling for baseline achievement would have involved extrapolation rather than comparison of observationally similar students. To address further concerns around the common support of covariates other than baseline test scores that predict private enrolment (e.g., socioeconomic status), we also estimate the private premium restricted to a sample of observationally similar children. In particular, we first estimate the probability of attending a private institution using a probit regression, separately by age groups, on the full set of value-added controls used in the main analysis. We then estimate the private premium when restricting the sample only to children within 5, 10 and 15 percentage points of the age-specific median probability of private enrolment (Online Appendix Tables D.1–D.3). These estimates yield results that are similar to those of the main analysis, and we cannot reject the equality at conventional levels of statistical significance.

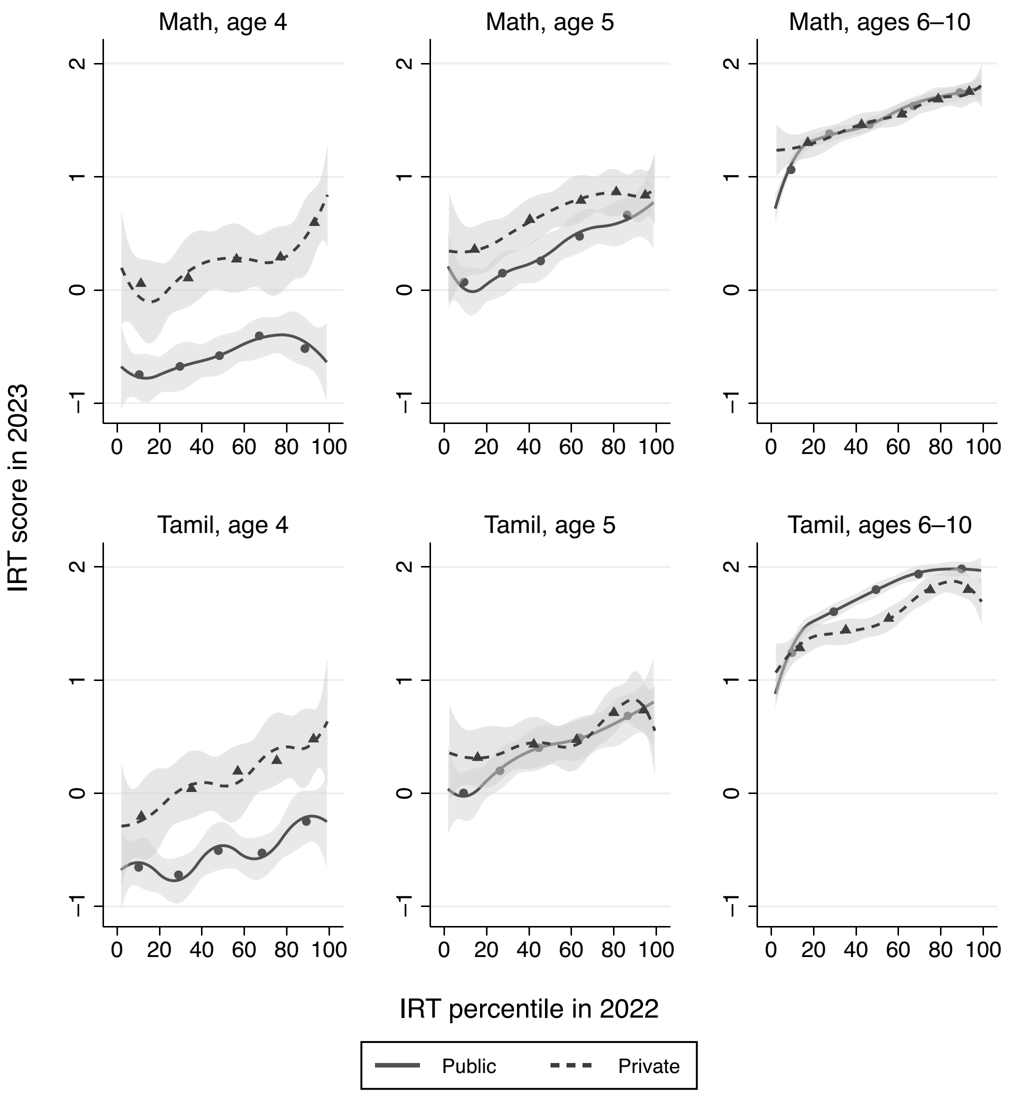

Fig. 1. *IRT Scores in 2023 by Percentile Scores in 2022 and Sector.*
Notes: These plots relate the IRT score of a student in the 2023 assessment to that of his or her within-age percentile rank in the 2022 assessment, separately by private/public enrolment. These semi-parametric estimations condition on village FEs, deciles of the SES index, child gender and parental education, using the approach of Cattaneo *et al.* (2024). The data are shown in five equally sized bins, and fitted lines are piecewise, quadratic polynomials with a smoothness parameter of 2, generated using the `binsreg` package in Stata.

relevant in many Indian states, such as Uttar Pradesh, where only 33% of children enrol in preschool (Pratham, 2022). In our sample, only ∼4% of four-year-olds ($N = 67$) are not enrolled in any level. Estimating the value-added of both public and private preschool facilities at age four, relative to the baseline of no enrolment, suggests that attending public care centres leads to learning gains of 0.24σ and 0.29σ in math and Tamil, respectively—roughly a quarter to a third of private preschool value-added (Online Appendix Table A.5).

Reflecting domestic and international policy targets, our data collection and analyses focus only on foundational numeracy and literacy in the local language. A resulting limitation is the absence of measures for English language proficiency, where prior research indicates that private primary schools significantly outperform public schools (Muralidharan and Sundararaman, 2015; Singh, 2015). Consequently, we likely underestimate the full private-sector premium in preschools and schools. Similarly, public or private preschools may place greater emphasis on socioemotional skills that are unmeasured in our data; our analysis is unable to estimate their differential productivity in these domains, which remain topics of independent interest.

2.4. *Robustness of the Private Premium*

The principal threat to our results is that our parsimonious specification of the value-added model (see (1)), with a limited set of covariates and a linear control for the subject-specific lagged score, does not fully account for selection into the two sectors. We examine the robustness of our estimates to these concerns in several ways.

First, we report results from richer specifications that include a battery of additional covariates, control for lagged achievement non-linearly and examine the stability of our estimated coefficient of the private school premium; this procedure is similar to validation exercises reported by Chetty *et al.* (2014). The additional covariates are quadratic polynomials in lagged scores in *both* math and Tamil; fixed effects at the level of survey month in 2022, 2023 and their interactions; controls for caste groups and educational inputs measured at baseline from our survey, including whether the child had recently received educational content via internet, TV or books. At the preschool level, the private premium is similar across specifications (0.72–0.74σ in math and 0.57–0.59σ in Tamil; see panel A of Table 3). Likewise, at the primary level, coefficients are stable across specifications, remaining close to zero in math and consistently negative in Tamil (between -0.17 and -0.19σ; see panel B of Table 3).

To explicitly quantify the sensitivity of our estimates to potential omitted variable bias, we further report robustness values (RVs) and bounds following Cinelli and Hazlett (2019) (see Table 3).[22] This procedure considers a potential omitted variable Z that predicts both private enrolment and test scores. Robustness values RV and $RV_{\alpha=0.05}$ measure how much of the residual variation in both test scores (Y) and private enrolment (D), after controlling for all other included covariates, needs to be explained by Z to (i) reduce the private premium to zero (RV) or (ii) make it statistically insignificant ($RV_{\alpha=0.05}$). These values range from 20% to 24% for Tamil and from 28% to 31% in math at the preschool level. To evaluate whether such a confounder Z is plausible, we consider the extreme scenario where Z is as powerful a predictor of Y and D, respectively, as our SES index, baseline test scores, parental education and child gender taken together. Even a confounder Z as strong as these covariates combined would explain at most $\sim$7% of remaining test-score variation and $\sim$16% of private enrolment variation, well below the threshold required to eliminate or substantially weaken our results. Hence, we view such an omitted variable as highly unlikely. For further details on the implementation of this exercise, see Online Appendix D.

[22] This procedure serves a similar purpose as Oster (2019) bounds, but offers a more intuitive approach to assessing the role of unobserved confounders.

Table 3. *Robustness of the Private School Premium.*

	Math				Tamil			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: age 4								
Private school	0.737***	0.736***	0.724***	0.724***	0.588***	0.585***	0.565***	0.573***
	(0.051)	(0.051)	(0.051)	(0.051)	(0.054)	(0.054)	(0.054)	(0.056)
RV	0.31	0.31	0.31	0.30	0.24	0.24	0.23	0.24
$RV_{\alpha=0.05}$	0.28	0.28	0.27	0.27	0.20	0.20	0.20	0.20
$R^2_{Y\sim D\mid X}$	0.12	0.12	0.12	0.12	0.07	0.07	0.07	0.07
$R^2_{Y\sim Z\mid D,X}$	0.06	0.02	0.02	0.02	0.07	0.02	0.02	0.02
$R^2_{D\sim Z\mid X}$	0.16	0.15	0.15	0.13	0.16	0.15	0.15	0.13
Observations	1,839	1,839	1,839	1,802	1,839	1,839	1,839	1,802
Panel B: ages 6–10								
Private school	−0.005	0.012	0.012	0.014	−0.172***	−0.192***	−0.192***	−0.178***
	(0.019)	(0.019)	(0.019)	(0.020)	(0.022)	(0.022)	(0.022)	(0.023)
RV	0.00	0.01	0.01	0.01	0.08	0.09	0.09	0.08
$RV_{\alpha=0.05}$	–	–	–	–	0.06	0.07	0.07	0.06
$R^2_{Y\sim D\mid X}$	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01
$R^2_{Y\sim Z\mid D,X}$	0.25	0.03	0.03	0.02	0.32	0.07	0.07	0.06
$R^2_{D\sim Z\mid X}$	0.18	0.17	0.18	0.15	0.17	0.17	0.17	0.14
Observations	14,344	14,344	14,344	13,932	14,344	14,344	14,344	13,932
Core controls	✓	✓	✓	✓	✓	✓	✓	✓
Lagged, squared Tamil & math scores		✓	✓	✓		✓	✓	✓
Survey month 2022 × 2023 FEs			✓	✓			✓	✓
Caste & home inputs in 2022				✓				✓

Notes: Robust SEs, clustered at the village level, are reported in parentheses. This table shows regression of IRT scores in 2023 on having attended a private rather than public preschool or primary school during the previous school year. Column (1) reports estimates of this private premium in our core specification (see (1)). Column (2) adds controls for baseline (2022) scores quadratically in both math and Tamil. Column (3) further includes fixed effects for the month in which the baseline and endline surveys were conducted, as well as their interactions. Finally, column (4) adds controls for caste and several educational inputs measured at baseline: whether the child had recently received educational content via (1) internet, (2) TV or (3) books at home. These measures are missing 37 (412) children at age 4 (6–10). Robustness values and partial R^2 under extreme scenarios of potential confounders are computed as described in Online Appendix D.1. *** $p < .01$.

As a final robustness check, we iteratively allow the coefficients on each of the covariates included in the main specification to vary one at a time by village. This addresses concerns about selection patterns being very different across villages. The estimated private premium changes very little when allowing for such interactions, both at the pre- and primary school levels (see Online Appendix Figure D.2).

2.5. *Socioeconomic Learning Gaps*

Concerns that private education might exacerbate inequality are central to public debates surrounding education policy (UNESCO, 2021). These concerns may be empirically grounded, as private-sector institutions disproportionately enrol students from higher-SES families (Table 1) and, at the preschool level, also have significantly higher productivity (Table 2). Therefore, we directly examine the contribution of differential productivity across sectors to SES gaps in achievement at early ages.

We focus on the gap in achievement between students from households in the top and bottom quartiles of socioeconomic status (omitting the middle half of the SES distribution). Specifically, we regress student test scores in 2023 on a dummy indicating whether the child belongs to the top SES quartile, and then sequentially condition on lagged achievement and a dummy variable for attending a private institution. As in our previous analyses, we estimate this separately for 4-, 5- and 6–10-year-olds (see Table 4).

Raw test score gaps between high- and low-SES students, at age four, are large (0.43σ and 0.36σ in math and Tamil, respectively). Differences in baseline achievement explain only around 5%–10% of this gap (columns (1) and (2)), but approximately 60% of it can be attributed to private preschool attendance (columns (2) and (3)). These patterns differ for primary school children. At ages 6–10, differences in baseline achievement account for around 30% of the test score gap in both subjects (columns (7) and (8)); private enrolment explains little of the SES gap in math and widens it in Tamil (columns (8) and (9)), which aligns with our previous findings on the primary private premium. Conditional on private preschool enrolment and lagged test scores, the remaining SES gap in test scores is relatively stable across age groups.

Overall, high-SES children enter primary school with substantially stronger academic achievement compared to their low-SES peers, a disparity that persists at later stages. These early gaps are primarily driven by differences in private preschool enrolment.

3. Spatial Variation in the Private Premium

The previous section estimated the *average* productivity differential across sectors. Yet, these averages likely conceal substantial heterogeneity across villages. Furthermore, within educational markets, the productivity of private and government preschools/schools is likely to interact; reducing productivity differentials to a sample-wide average restricts us from investigating such relationships. In this section, we investigate these market-level associations.

3.1. *Estimating Village-Level Value-Added*

Our empirical approach to estimating village- and sector-specific school productivity extends the value-added framework described in Section 2.2. We estimate test score gains for each level-sector-village cell, where level refers to pre-/primary schooling and sector to public/private operators. We define θ_{slv} as a set of dummy variables that indicate attendance at a private or public option s at the pre- or primary level l in village v. To improve precision, we pool all children aged 4–10 (adding subscript a for age) in the same regression and estimate the equation

$$y_{iaslv}^{2023} = \lambda_a y_{iaslv}^{2022} + \theta_{slv} + \Gamma \mathbf{X}_{iaslv} + \epsilon_{iaslv}. \quad (2)$$

The coefficient on the lagged score λ_a is allowed to differ by child age a. The vector $\mathbf{X}_{iaslv}$ contains controls for deciles of the SES wealth index, as well as paternal and maternal education, the child's age and gender. The public preschool sector of one of the villages is left out as the omitted category.

In this model, θ_{slv} captures the improvement in test scores between 2022 and 2023 for a child in a particular sector, level of schooling and village after controlling for baseline test scores and background characteristics. The difference in θ_{slv} across sectors allows us to identify the private

Table 4. *Decomposition of the SES Gap (Top/Bottom 25%) in 2023 Test Scores: Preschool and Primary Levels.*

	Age 4			Age 5			Ages 6–10		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: math									
Top 25% SES	0.434***	0.412***	0.169*	0.307***	0.250***	0.143*	0.238***	0.164***	0.155***
	(0.072)	(0.072)	(0.081)	(0.056)	(0.055)	(0.065)	(0.029)	(0.026)	(0.026)
Private school			0.680***			0.249**			0.024
			(0.080)			(0.075)			(0.026)
Panel B: Tamil									
Top 25% SES	0.356***	0.329***	0.135	0.147*	0.086	0.054	0.099**	0.053	0.105***
	(0.068)	(0.067)	(0.069)	(0.061)	(0.063)	(0.067)	(0.030)	(0.027)	(0.030)
Private school			0.541***			0.074			−0.142***
			(0.081)			(0.073)			(0.031)
Lagged score control		✓	✓		✓	✓		✓	✓
Observations	914	914	914	1,360	1,360	1,360	6,873	6,873	6,873

Notes: Robust SEs, clustered at the village level, are reported in parentheses. Village fixed effects and controls for child gender are included in all regressions. Test scores refer to the IRT EAP scores, standardised with respect to children aged five in the 2022 assessments. The SES index is based on questions regarding the availability of household amenities and computed with PCA. The omitted category contains students in households with an SES index below the 25th percentile. Households with an SES index between the 25th and 75th percentiles are excluded from the regressions. *** $p < .01$, ** $p < .05$, * $p < .1$.

premium at the level of schooling l and village v:

$$\beta_{lv} = \theta_{1lv} - \theta_{0lv}.$$

To interpret β_{lv} causally, we require selection into the private relative to the public sector in *each* village to be accounted for by our set of included controls. This is stronger than the assumption in Section 2.2 only in that this conditional ignorability assumption is imposed for each village individually (rather than the sample as a whole); it is weaker than the standard assumption underlying school value-added models wherein the choice of *each* school in the sample is assumed to be conditionally ignorable.

Our principal aim in this section is to compare the productivity of available options in the same market and estimate the distribution of the private premium across markets and levels. This goal is similar to Andrabi *et al.* (2025), who established the validity of school-level value-added using alternative sources of (within-village) identification. Like other applications of value-added models across many dispersed markets (see, e.g., Andrabi *et al.,* 2025; Einav *et al.,* 2025), we focus on reallocations only within the same market.[23] When we correlate estimated productivity across sectors in the same villages, we only interpret these as associations: positive correlations may be informative about spatial inequality, but could arise from institutional effects (e.g., through school competition) or through the effect of village-level unobservables not proxied for by our covariates.

The value-added estimates in θ_{slv} will be measured with uncertainty, introducing measurement error in both the individual parameters and, potentially, the private premium β_{lv}. This presents a challenge for several parts of our analysis. First, individual value-added estimates may be severely biased even under the assumption that our approach yields unbiased estimates *on average*. Second, the estimated slope in a regression that has value-added estimates in the right-hand side (e.g., private on public value-added, a core object of interest in our analysis) will be subject to attenuation bias due to measurement error.

We address this in two ways. We rely on empirical Bayes estimates of value-added when investigating individual productivity parameters. Our empirical Bayes procedure shrinks value-added estimates toward their level and sector averages proportionally to the uncertainty with which they are estimated (for details and the impact of shrinkage, see Online Appendix E.1). This shrinkage procedure is suitable when the goal is to improve unit-specific forecasts. However, it does not allow us to appropriately investigate moments of the value-added distribution—such as the covariance between private and public value-added (Walters, 2024).[24] Denoting the within-village covariance matrix of θ_{sljv} across sectors s, levels l and subjects j as Σ_θ, we estimate Σ_θ directly following Angrist *et al.* (2025), which corrects for bias due to estimation noise (for details, see Online Appendix E.2). In the results, we visually illustrate the relationships between our raw value-added estimates, and report both the unadjusted and bias-corrected slopes throughout.

[23] Comparing value-added estimates *across* villages is, however, much more complicated due to the possibility of spatial unobservables (e.g., neighbourhood effects, as in Chetty and Hendren, 2018). The relevant thought experiment here would be to move a child from, say, a private school in village A to a public school in village B; this would only be valid if village A or B themselves did not have an independent effect on student outcomes, beyond that proxied by covariates. We are not aware of results in any setting that validate value-added models against experimental or regression discontinuity estimates for this purpose.

[24] While the variance of the raw value-added estimates will be inflated due to estimation error, the shrunken empirical Bayes estimates will, in general, understate the variance of the true value-added parameters (Walters, 2024).

Finally, by pooling children aged 4–10 in the same specification, we also assume that (age-varying) lagged scores and background characteristics can address selection into early primary schooling (i.e., the *timing* of the transition, not just the sector). This is consequential for five-year-olds, around half of whom are enrolled in primary school (Table 1). The alternative strategy is to omit five-year-olds (Online Appendix Figure E.3), to which our results are robust.

3.2. *Value-Added across Markets, Sectors and Levels*

We now present the results of estimating the specification from (2), which yields four measures of test score gains for each village: the average value-added in private and public options at the preschool and primary levels.[25] The variation in test-score gains is slightly larger in the private sector compared to the public one, both at the pre- and primary school levels. We focus, first, on the difference in value-added between the public and private sectors (e.g., the private premium) across different markets, and then turn to within-market correlations of these estimates.

Figure 2 orders villages by their average value-added in the public sector, as estimated by the empirical Bayes method, along the horizontal axis. Circles (triangles) denote the average value-added in the public (private) sector for each village. The top panels show preschool results, separately by subject, and the bottom panels show primary school results.

No village has a public preschool sector that, on average, performs better than its private sector in math—with few exceptions, the same is true for Tamil.[26] The average private premium is 0.72σ in math and 0.57σ in Tamil, which is very similar to the results in Section 2.3. While productivity differences vary across markets, they are substantial almost everywhere. This pattern is very different at the primary school level. While the average private productivity premium is essentially zero in math and negative in Tamil, as noted previously, the magnitude of these differences remains modest across all villages.

Next, we investigate the association between value-added estimates across sectors and levels within villages. We show binned scatter plots of raw value-added estimates in Figure 3. As mentioned above, these relationships are likely subject to attenuation bias due to estimation noise in the value-added estimates. Hence, we report both raw and bias-corrected slopes throughout (see Online Appendix E.2 for details on the bias correction approach).

We identify a positive correlation between the value-added of the private and public sectors. Regressing raw private on public preschool value-added yields a coefficient of 0.45 in math and 0.36 in Tamil (the top panel of Figure 3). These slope coefficients are biased towards zero: correcting for estimation error increases these slopes to 0.80 in math and 0.61 in Tamil. These sectoral correlations are higher in primary school, with slopes of 0.88 in math and 0.99 in Tamil (after correcting for bias). However, the difference in slopes at the pre- and primary levels is only statistically significant in Tamil ($p = .48$ in math, $p = .02$ in Tamil). These associations in productivity across sectors are consistent with potential 'multiplier effects' in which, due to market-level incentives and competition, an increase in public-sector quality also leads to improvements in the private sector—Andrabi *et al.* (2024) presented experimental evidence of

[25] Online Appendix Figure A.5 presents density plots of these estimates, together with bias-corrected variances of the value-added distributions. For fifteen villages, we are unable to estimate village-level value-added in all combinations of subjects, sectors and levels. In most cases, this is because too few children are enrolled in private preschools. This leaves us with 200 villages for which we have complete value-added estimates.

[26] This is not primarily a consequence of empirical Bayes shrinkage. Focusing on raw (unshrunk) value-added estimates, only 7% (14%) of villages have a public preschool sector that outperforms its private sector on value-added in math (Tamil), despite raw value-added being estimated with a significant degree of noise (Online Appendix Figure E.4).

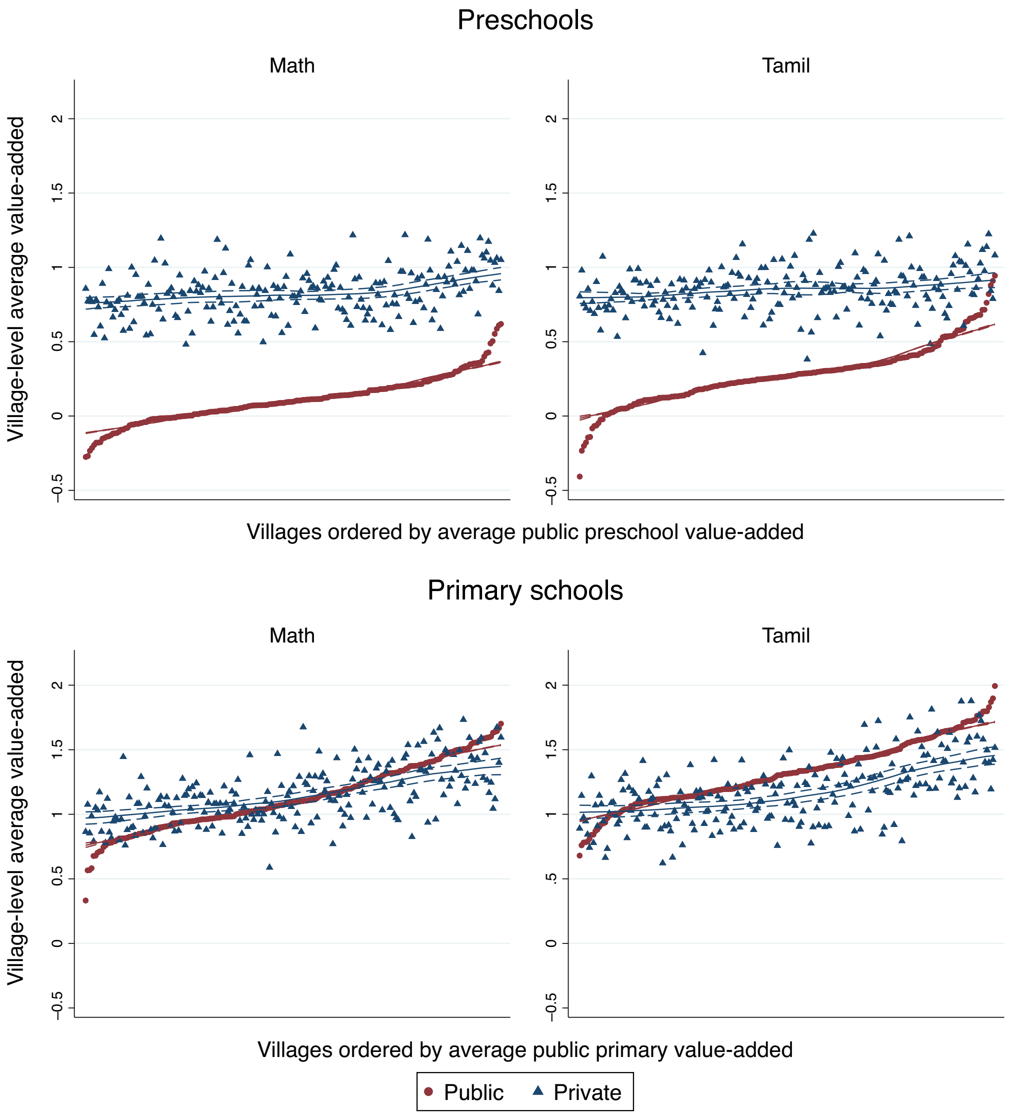

Fig. 2. *Village-Level Average Value-Added of Private and Public Options.*
Notes: These plots show village-level average school value-added by sector (public/private) and level (preschool/primary) using empirical Bayes measures, as described in Online Appendix E.1. Villages are ordered along the x axis by their average value-added in government schools. The regression specification generating these estimates is given by (2).

such a mechanism at work in primary schooling markets in Pakistan, although we are not aware of similar evidence at the preschool level. In short, villages with particularly low-quality public schools also tend to have weaker-performing *private* schools than other villages, and this pattern is stronger in primary relative to preschool markets, at least in Tamil.

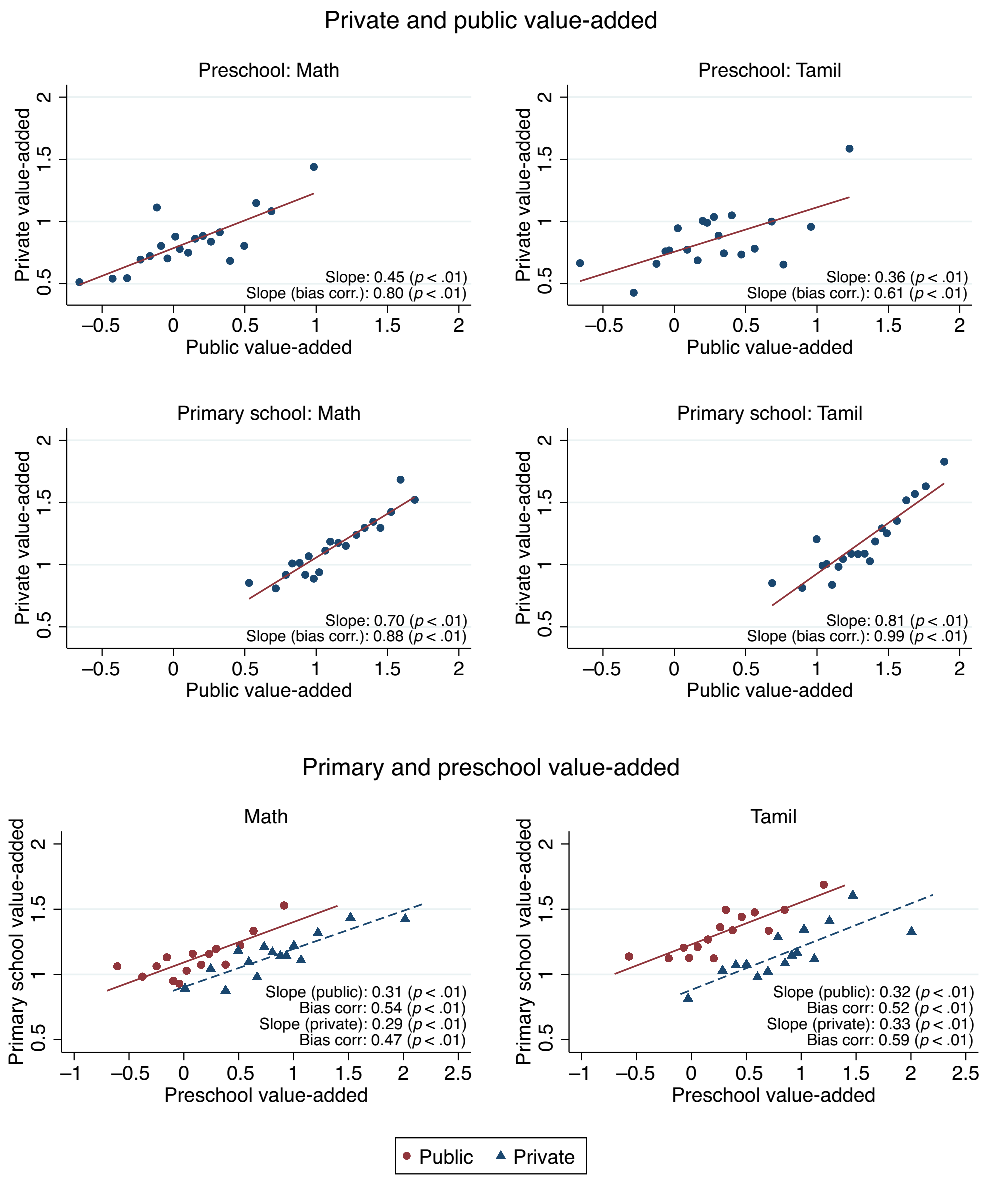

Fig. 3. *Correlations of Village Value-Added across Sectors and Levels.*

Notes: These plots show binned scatter plots of village-level empirical Bayes value-added estimates. In the top panel, we show the correlation between private- and public-sector value-added, separately by subject and level (preschool/primary school). In the bottom panel, we show correlations between preschool and primary school value-added, separately by subject and sector. Raw and bias-corrected slopes and p-values associated with a test of zero slopes are shown in the plots; robust SEs are used for inference (see Online Appendix E.2 for details on the bias-correction procedure).

We also investigate the within-sector correlation of value-added between pre- and primary schools within the same markets (bottom panel of Figure 3). In both the public and private sectors, productivity is clearly correlated across levels of schooling. On average, an increase of one SD of *preschool* value-added predicts an increase of roughly half a SD of *primary school* value-added. This correlation is not surprising in the private sector since private preschools are often vertically integrated with private schools. Government preschools, on the other hand, are managed by a parallel administrative set-up separate from the School Education Department (which oversees primary schools), but appear to display similar correlations.[27] As such, productivity differences across markets appear to be relatively persistent throughout early childhood and adolescence.

Turning to market shares at the preschool level, we find that private preschool enrolment does *not* increase with the size of the private premium (see Online Appendix Figure A.6). At the primary school level, in contrast, we *do* find that market shares reflect differences in the private premium. In math, a one-SD increase in the village-level private premium is associated with a 15-percentage-point increase in private enrolment. In Tamil, the correlation is essentially flat, which is not unsurprising, given that many households opt for private primary schools precisely because of their focus on English rather than the local language.[28]

Importantly, market shares are equilibrium outcomes. Thus, the lack of association between private preschool premia and market shares is consistent with distinct explanations. For example, it is possible that (i) households do not value cognitive skill production for very young children and/or (ii) higher-quality preschools also charge higher prices (which we do not measure), or (iii) higher-quality preschools want to keep enrolment low to maintain quality. Since market shares appear to respond at the primary level to value-added in math (which is emphasised in both public and government schools), explanations for this pattern are likely to be specific to preschools.

We provide further analyses of the correlates of village-level value-added in the Online Appendix, which reveal three additional findings. First, value-added is highly correlated across subjects (Online Appendix Figure A.7): within a village, private and public preschools that provide high value-added in math also tend to do so in Tamil. Second, private premia across markets are largely uncorrelated with market size, measured as the number of children aged 4–6 in each village, and village-level SES at any level of schooling (Online Appendix Table A.6). If anything, villages with weaker socioeconomic composition tend to have larger private premia compared to those with stronger composition.

4. Conclusion

This paper provides new insights into the study of public-private differences in education systems in LMICs. In nearly all villages in our sample, private preschools exhibit a substantial advantage in test score value-added over public options, which accounts for nearly two-thirds of socioeconomic inequality at the school entry age. This is in stark contrast to the (small) differences in productivity

[27] Public preschools are run by the Ministry of Women and Child Welfare at the national level, not the Ministry of Education (which runs primary schools). Staffing, pay, management and overall capacity all differ between these two structures.

[28] As shown in Chen (2024) and discussed in Walters (2024), using shrunken empirical Bayes estimates when correlating value-added with market shares may be inappropriate if the precision and magnitude of the value-added estimates are correlated. We address this issue by adopting an alternative empirical Bayes procedure that is robust to such precision dependence (Chen, 2024) and find similar results (see Online Appendix E.5).

between public and private schools at the primary school level. Within villages, we document a positive correlation between the value-added of educational options, suggesting spatial inequality in access to educational opportunity (akin to neighbourhood effects in other settings; see, e.g., Chetty and Hendren, 2018).

Why does the public sector underperform private alternatives so starkly at the preschool level, even while producing similar value-added at the primary school level? Our household-based dataset lacks the information on time use, pedagogical practices or other facility-based inputs needed for a structured analysis of this question. A likely explanation is that private preschools dedicate more instructional time to cognitive stimulation. In their control group, in the same districts, Ganimian *et al.* (2024) documented only 38 minutes of preschool instruction per day in anganwadis. Doubling this, using a part-time worker, led to gains of 0.28σ for children who attended the treatment centres. It is likely that, at the preschool level, private institutions effectively provide more instructional time than even the treatment group of Ganimian *et al.* (2024), which rationalises the large private preschool premia we find.[29] This limited instructional time in *anganwadis*, potentially combined with low productivity of such instruction, is a problem across Indian states.[30]

A potential explanation for our results is that, despite the *stated* importance given to early childhood cognition in policy documents, preschool instruction has been less central to anganwadi centres than their role as daycare centres or supplementary nutrition centres. Our results acknowledge this possibility, but also highlight that the ICDS system is (i) the principal government intervention at the preschool stage, (ii) the largest such system globally and (iii) explicitly intended to be part of the system-wide pivot towards achieving universal literacy and numeracy by Grade 3. Providing *quality* preschool to all children, as targeted by the Sustainable Development Goals and the National Education Policy, is unlikely to be feasible without improving the value-added of *anganwadis* in the skills related to functional literacy and numeracy.

Our results suggest that scaling effective interventions targeting learning in public preschools, beyond effects on achievement, could also improve socioeconomic equality. Promising anganwadi-based interventions include, for example, additional staffing in Tamil Nadu (Ganimian *et al.*, 2024) and WhatsApp-based instructional materials for anganwadi workers and parents (Keskar *et al.*, 2025). That we find positive correlations between private and public value-added also potentially suggests that these improvements might even have multiplier effects through market-level interactions (such as demonstrated by Andrabi *et al.*, 2024 at the primary school level). Developing, validating and scaling such interventions should, given our results, be an important priority for research and policy.

Our results also suggest that substantial gains may be possible from vouchers that enable children to attend private preschools. Dean and Jayachandran (2019) showed the effects of such a policy in India, finding gains of as much as 0.8σ for students induced to attend a particular

[29] Put differently, it is possible that public preschools could achieve similar gains as private preschools if they provided instruction of equivalent duration. That they are unable to do so reflects, at least in part, inadequate resources and staffing (Ganimian *et al.*, 2024). See Singh *et al.* (2024) for an example of sharp learning gains in response to government investments after COVID-19 in this setting.

[30] See, for instance, descriptive findings reported in the India Early Childhood Education Impact Study (Kaul and Bhattacharjea, 2019) and the FOCUS report (CIRCUS, 2006). Indeed, in the FOCUS report, anganwadis in Tamil Nadu were highlighted as being better functioning than in other states, suggesting that our results might understate the differential with private-sector preschools. De facto, preschool education has not been prioritised by the government: this is evident, for example, also in budgetary allocations, staffing and the social status of preschool workers compared to primary school teachers.

private preschool provider using a randomised voucher. Our value-added estimates are remarkably similar on average; however, we also observe significant differences in the private premium across villages. Voucher policies to move students to private preschools are also, de facto, already at scale: Romero and Singh (2022) showed that the principal effect of private-sector quotas in India's Right to Education Act 2009 on enrolment is to move some students from public preschools or no enrolment into private preschools. Understanding the effects of these policy-induced moves on skill acquisition, which has not been done outside of the COVID-19 pandemic, as well as ways to target the policies to make them more effective, are clearly avenues where further research would be productive.[31]

More generally, both in India and elsewhere among LMICs, there is remarkably little evidence on the organisation of preschool markets. Understanding the distribution of productivity among individual providers, the preferences and information sets of parents, constraints on the provision of quality instruction, and the nature of competition and market interactions between providers are all topics of substantial importance for academic research. We hope that the observational results presented in this paper will spur further work in this area, including interventions to improve the functioning of these markets.

Stockholm School of Economics, Sweden
Instituto Tecnologico Autonomo de Mexico, Mexico J-PAL, U.S.A.
Stockholm Schools of Economics, Sweden J-PAL, U.S.A.

Additional Supporting Information may be found in the online version of this article:

Online Appendix
Replication Package

References

Allende, Claudia. (2019). 'Competition under social interactions and the design of education policies', Working Paper No. 3979, Stanford Graduate School of Business. https://www.gsb.stanford.edu/faculty-research/working-papers/competition-under-social-interactions-design-education-policies (last accessed: 1 Sept. 2025).

Andrabi, Tahir, Natalie Bau, Jishnu Das, Naureen Karachiwalla, and Asim Ijaz Khwaja. (2024). 'Crowding in private quality: The equilibrium effects of public spending in education', *The Quarterly Journal of Economics*, vol. 139, pp. 2525–77.

Andrabi, Tahir, Natalie Bau, Jishnu Das, and Asim I. Khwaja. (2025). 'Heterogeneity in school value added and the private premium', *American Economic Review*, vol. 115(1), pp. 147–82.

Andrabi, Tahir, Jishnu Das, Asim I. Khwaja, and Tristan Zajonc. (2011). 'Do value-added estimates add value? Accounting for learning dynamics', *American Economic Journal: Applied Economics*, vol. 3(3), pp. 29–54.

Andrabi, Tahir, Jishnu Das, and Asim I Khwaja. (2017). 'Report cards: The impact of providing school and child test scores on educational markets', *American Economic Review*, vol. 107(6), pp. 1535–63.

Andrew, Alison, Orazio P. Attanasio, Raquel Bernal, Lina C. Sosa, Sonya Krutikova, and Marta Rubio-Codina. (2024). 'Preschool quality and child development', *Journal of Political Economy*, vol. 132(7), pp. 2304–45.

[31] These voucher-led policies could also be substantially cost effective. Ganimian *et al.* (2024) documented gains of 0.11σ incurred with a per-child cost of ~3,500 INR per year. Private preschool fees in our study districts are ~7,000 INR (~95 USD) per year on average. Taking our value-added estimates of private preschool premia of 0.55–0.7σ at face value, even if only 20% of voucher recipients could be induced to shift from public preschools (with the rest of voucher spending being inframarginal), this would be as or more cost effective. These calculations do not account for (unmeasured) potential treatment effects of private institutions on English, which are likely to be large (Muralidharan and Sundararaman, 2015; Singh, 2015). Incorporating these would further improve the measured cost effectiveness.

Angrist, Joshua, Parag A. Pathak, and Christopher R. Walters. (2013). 'Explaining charter school effectiveness', *American Economic Journal: Applied Economics*, vol. 5(4), pp. 1–27.
Angrist, Joshua, Peter D. Hull, and Christopher Walters. (2023). 'Methods for measuring school effectiveness', in (Eric A. Hanushek, Stephen Machin and Ludger Woessmann, eds.), *Handbook of the Economics of Education*, vol. 7, pp. 1–60, Amsterdam: Elsevier.
Angrist, Joshua, Peter D. Hull, Parag A. Pathak, and Christopher R. Walters. (2017). 'Leveraging lotteries for school value-added: Testing and estimation', *The Quarterly Journal of Economics*, vol. 132(2), pp. 871–919.
Angrist, Joshua, Peter D. Hull, Parag A. Pathak, and Christopher Walters. (2024). 'Credible school value-added with undersubscribed school lotteries', *The Review of Economics and Statistics*, vol. 106, pp. 1–19.
Angrist, Joshua, Peter D. Hull, Russell Legate-Yang, Parag A. Pathak, and Christopher R. Walters. (2025). 'Putting school surveys to the test', Working paper, National Bureau of Economic Research.
Attanasio, Orazio, Ricardo Paes de Barros, Pedro Carneiro, David K. Evans, Lycia Lima, Pedro Olinto, and Norbert Schady. (2022). 'Public childcare, labor market outcomes of caregivers, and child development: Experimental evidence from Brazil', Working paper, National Bureau of Economic Research.
Attanasio, Orazio, Camila Fernández, Emla O. A. Fitzsimons, Sally M. Grantham-McGregor, Costas Meghir, and Marta Rubio-Codina. (2014). 'Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in Colombia: cluster randomized controlled trial', *BMJ*, vol. 349, doi:10.1136/bmj.g5785.
Barrera-Osorio, Felipe, David S. Blakeslee, Matthew Hoover, Leigh Linden, Dhushyanth Raju, and Stephen P. Ryan. (2022). 'Delivering education to the underserved through a public-private partnership program in Pakistan', *Review of Economics and Statistics*, vol. 104(3), pp. 399–416.
Bau, Natalie. (2022). 'Estimating an equilibrium model of horizontal competition in education', *Journal of Political Economy*, vol. 130(7), pp. 1717–64.
Behrman, Jere R., Lia Fernald, and Patrice Engle. (2014). 'Preschool programs in developing countries', in (Paul Glewwe, ed.), *Education Policy in Developing Countries*, pp. 65–105, Chicago: University of Chicago Press.
Bjorvatn, Kjetil, Denise Ferris, Selim Gulesci, Arne Nasgowitz, Vincent Somville, and Lore Vandewalle. (2025). 'Childcare, labor supply, and business development: Experimental evidence from Uganda', *American Economic Journal: Applied Economics*, vol. 17, pp. 75–101.
Black, Maureen M., Susan P. Walker, Lia C. H. Fernald, Christopher T. Andersen, Ann M. DiGirolamo, Chunling Lu, Dana C. McCoy, Günther Fink, Yusra R. Shawar, Jeremy Shiffman, Amanda E. Devercelli, Quentin T. Wodon, Emily Vargas-Barón, and Sally Grantham-McGregor. (2017). 'Early childhood development coming of age: science through the life course', *The Lancet*, vol. 389(10064), pp. 77–90.
Bordoloi, M., S. Pandey, V. Irava, and R. Junnarkar. (2020). 'State education finances', Technical Report, Accountability Initiative, Centre for Policy Research.
Carneiro, Pedro, Jishnu Das, and Hugo Reis. (2024). 'The value of private schools: Evidence from Pakistan', *Review of Economics and Statistics*, vol. 106(5), pp. 1301–18.
Cattaneo, Matias D., Richard K. Crump, Max H. Farrell, and Yingjie Feng. (2024). 'On binscatter', *American Economic Review*, vol. 114(5), pp. 1488–514.
Chen, Jiafeng. (2024). 'Empirical Bayes when estimation precision predicts parameters', Preprint, https://arxiv.org/abs/2212.14444 (last accessed: 1 Sept. 2025).
Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. (2014). 'Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates', *American Economic Review*, vol. 104(9), pp. 2593–632.
Chetty, Raj, and Nathaniel Hendren. (2018). 'The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects', *The Quarterly Journal of Economics*, vol. 133(3), pp. 1107–62.
Cinelli, Carlos, and Chad Hazlett. (2019). 'Making sense of sensitivity: Extending omitted variable bias', *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 82(1), pp. 39–67.
CIRCUS. (2006). *Focus on Children Under Six: Abridged Report*. New Delhi, India. Right to Food Campaign.
Crawfurd, Lee, Susannah Hares, and Rory Todd. (2024). 'The impact of private schools, school chains and PPPs in developing countries', *The World Bank Research Observer*, vol. 39(1), pp. 97–123.
Das, Jishnu, and Tristan Zajonc. (2010). 'India shining and Bharat drowning: Comparing two Indian states to the worldwide distribution in mathematics achievement', *Journal of Development Economics*, vol. 92(2), pp. 175–87.
Dean, Joshua T., and Seema Jayachandran. (2019). 'Attending kindergarten improves cognitive but not socioemotional development in India', Working paper. https://seemajayachandran.com/kindergarten.pdf (last accessed: 1 Sept. 2025).
Einav, Liran, Amy Finkelstein, and Neale Mahoney. (2025). 'Producing health: Measuring value added of nursing homes', *Econometrica*, vol. 93(4), pp. 1225–64.
Elango, Sneha, Jorge Luis Garcia, James Heckman, and Andrés Hojman. (2015). 'Early childhood education', in (Robert A. Moffitt, ed.), *Economics of Means-Tested Transfer Programs in the United States*, vol. 2, pp. 235–97, Chicago: University of Chicago Press.
Engle, Patrice L., Maureen M. Black, Jere R. Behrman, Meena Cabral De Mello, Paul J. Gertler, Lydia Kapiriri, Reynaldo Martorell, and Mary E Young. (2007). 'Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world', *The Lancet*, vol. 369(9557), pp. 229–42.

Engle, Patrice L., Lia C.H. Fernald, Harold Alderman, Jere Behrman, Chloe O'Gara, Aisha Yousafzai, Meena Cabral De Mello, Melissa Hidrobo, Nurper Ulkuer, Ilgi Ertem, and Selim Iltus and Global Child Development Steering Group. (2011). 'Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries', *The Lancet*, vol. 378(9799), pp. 1339–53.

Evans, David K., Pamela Jakiela, and Amina Mendez Acosta. (2024). 'The Impacts of childcare interventions on children's outcomes in low- and middle-income countries: A systematic review', *AEA Papers and Proceedings*, vol. 114, pp. 463–6.

Fernald, Anne, Virginia A. Marchman, and Adriana Weisleder. (2013). 'SES differences in language processing skill and vocabulary are evident at 18 months', *Developmental Science*, vol. 16(2), pp. 234–48.

Fernald, Lia C.H., Patricia Kariger, Melissa Hidrobo, and Paul J. Gertler. (2012). 'Socioeconomic gradients in child development in very young children: Evidence from India, Indonesia, Peru, and Senegal', *Proceedings of the National Academy of Sciences*, vol. 109(supplement_2), pp. 17273–80.

Fryer, Roland G., Jr., and Steven D. Levitt. (2004). 'Understanding the black-white test score gap in the first two years of school', *Review of Economics and Statistics*, vol. 86(2), pp. 447–64.

Fryer, Roland G., Jr., and Steven D. Levitt. (2006). 'The black-white test score gap through third grade', *American Law and Economics Review*, vol. 8(2), pp. 249–81.

Fryer, Roland G., Jr., and Steven D. Levitt. (2013). 'Testing for racial differences in the mental ability of young children', *American Economic Review*, vol. 103(2), pp. 981–1005.

Ganimian, Alejandro J., Karthik Muralidharan, and Christopher R. Walters. (2024). 'Augmenting state capacity for child development: Experimental evidence from India', *Journal of Political Economy*, vol. 132(5), pp. 1565–602.

Government of India. (2019). *Unified District Information System for Education Plus (UDISE+)*, New Delhi: Ministry of Education, Government of India.

Grantham-McGregor, Sally, Yin Bun Cheung, Santiago Cueto, Paul Glewwe, Linda Richter, and Barbara Strupp. (2007). 'Developmental potential in the first 5 years for children in developing countries', *The Lancet*, vol. 369(9555), pp. 60–70.

Hart, Betty, and Todd R. Risley. (1995). *Meaningful Differences in the Everyday Experience of Young American Children*, Baltimore, MD: Paul H. Brookes Publishing Co.

Holla, Alaka, Magdalena Bendini, Lelys Dinarte, and Iva Trako. (2021). 'Is investment in preprimary education too low? Lessons from (quasi) experimental evidence across countries', Working Paper 9723, World Bank.

Kaul, Venita, and Suman Bhattacharjea. (2019). *Early Childhood Education and School Readiness in India*, Singapore: Springer.

Keskar, Ajinkya, Mauricio Romero, Abhijeet Singh, and Karthik Muralidharan. (2025). 'Using technology to deliver preschool services at scale: Experimental evidence from India', Working Paper. https://www.ajinkyakeskar.com/publication/rocket/ (last accessed: 1 Sept. 2025).

Kingdon, Geeta Gandhi. (2020). 'The private schooling phenomenon in India: A review', *The Journal of Development Studies*, vol. 56(10), pp. 1795–817.

Lee, Valerie E., and David T. Burkam. (2002). *Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School*, Washington, DC: Economic Policy Institute.

Ministry of Education. (2020). 'NIPUN Bharat mission', https://nipunbharat.education.gov.in/ (last accessed: 23 June 2025).

Muralidharan, Karthik, and Abhijeet Singh. (2021). 'India's new national education policy: Evidence and challenges', *Science*, vol. 372(6537), pp. 36–8.

Muralidharan, Karthik, and Venkatesh Sundararaman. (2015). 'The aggregate effect of school choice: Evidence from a two-stage experiment in India', *The Quarterly Journal of Economics*, vol. 130(3), pp. 1011–66.

Neilson, Christopher A. (2021). 'Targeted vouchers, competition among schools, and the academic achievement of poor students', Working Paper 2021-48.Princeton University. https://ideas.repec.org/p/pri/econom/2021-48.html (last accessed: 1 Sept. 2025).

Nielsen, Eric. (2023). 'The income-achievement gap and adult outcome inequality', *Journal of Human Resources*, doi: 10.3368/jhr.0519-10220R3.

Noble, Kimberly G., M. Frank Norman, and Martha J. Farah. (2005). 'Neurocognitive correlates of socioeconomic status in kindergarten children', *Developmental Science*, vol. 8(1), pp. 74–87.

Oster, Emily. (2019). 'Unobservable selection and coefficient stability: Theory and evidence', *Journal of Business and Economic Statistics*, vol. 37(2), pp. 187–204.

Pratham. (2022). *Annual Status of Education Report 2022*, New Delhi: ASER Centre.

Reardon, Sean F. (2011). 'The widening academic achievement gap between the rich and the poor: New evidence and possible explanations', *Whither Opportunity*, vol. 1(1), pp. 91–116.

Reardon, Sean F. (2021). 'The economic achievement gap in the US, 1960-2020: Reconciling recent empirical findings', Working Paper 21-09, Center for Education Policy Analysis.

Reynolds, Arthur J., Momoko Hayakawa, Suh-Ruu Ou, Christina F. Mondi, Michelle M. Englund, Allyson J. Candee, and Nicole E. Smerillo. (2017). 'Scaling and sustaining effective early childhood programs through school–family–university collaboration', *Child Development*, vol. 88(5), pp. 1453–65.

Romero, Mauricio, and Justin Sandefur. (2021). 'Beyond short-term learning gains: The impact of outsourcing schools in Liberia after three years', Economic Journal, vol. 132(644), pp. 1600–19.

Romero, Mauricio, Justin Sandefur, and Wayne A Sandholtz. (2020). 'Outsourcing education: Experimental evidence from Liberia', *American Economic Review*, vol. 110(2), pp. 364–400.
Romero, Mauricio, and Abhijeet Singh. (2022). 'The incidence of affirmative action: Evidence from quotas in private schools in India', Working paper 22/088, RISE, https://doi.org/10.35489/BSG-RISE-WP_2022/088 (last accessed: 1 Sept. 2025).
Rubio-Codina, Marta, Orazio Attanasio, Costas Meghir, Natalia Varela, and Sally Grantham-McGregor. (2015). 'The socioeconomic gradient of child development: Cross-sectional evidence from children 6–42 months in Bogota', *The Journal of Human Resources*, vol. 50(2), pp. 464–83.
Schady, Norbert, Jere Behrman, Maria Caridad Araujo, Rodrigo Azuero, Raquel Bernal, David Bravo, Florencia Lopez-Boo, Karen Macours, Daniela Marshall, Christina Paxson, and Renos Vakis. (2015). 'Wealth gradients in early childhood cognitive development in five Latin American countries', *Journal of Human Resources*, vol. 50(2), pp. 446–63.
Singh, Abhijeet. (2014). 'Test score gaps between private and government sector students at school entry age in India', *Oxford Review of Education*, vol. 40(1), pp. 30–49.
Singh, Abhijeet. (2015). 'Private school effects in urban and rural India: Panel estimates at primary and secondary school ages', *Journal of Development Economics*, vol. 113(C), pp. 16–32.
Singh, Abhijeet. (2020). 'Learning more with every year: School year productivity and international learning divergence', *Journal of the European Economic Association*, vol. 18(4), pp. 1770–813.
Singh, Abhijeet, Mauricio Romero, and Karthik Muralidharan. (2024). 'COVID-19 learning loss and recovery: Panel data evidence from India', *Journal of Human Resources*, doi: 10.3368/jhr.0723-13025R2.
Todd, Petra E., and Kenneth I.. Wolpin (2003). 'On the specification and estimation of the production function for cognitive achievement', ECONOMIC JOURNAL, vol. 113(485), pp. F3–33.
Todd, Petra E., and Kenneth I. Wolpin. (2007). 'The production of cognitive achievement in children: Home, school, and racial test score gaps', *Journal of Human Capital*, vol. 1(1), pp. 91–136.
UNESCO. (2020a). 'Enrolment in pre-primary education, both sexes (number) [data]', http://data.uis.unesco.org/ (last accessed: 20 May 2025).
UNESCO. (2020b). 'Enrolment in primary education, both sexes (number) [data]', http://data.uis.unesco.org/ (last accessed: 20 May 2025).
UNESCO. (2020c). 'Initial government funding per pre-primary student as a percentage of GDP per capita [data]', http://data.uis.unesco.org/ (last accessed: 20 May 2025).
UNESCO. (2020d). 'Initial government funding per primary student as a percentage of GDP per capita [data]', http://data.uis.unesco.org/ (last accessed: 20 May 2025).
UNESCO. (2020e). 'Percentage of enrolment in pre-primary education in private institutions, both sexes (%) [data]', http://data.uis.unesco.org/ (last accessed: 20 May 2025).
UNESCO. (2020f). 'Percentage of enrolment in primary education in private institutions, both sexes (%) [data]', http://data.uis.unesco.org/ (last accessed: 20 May 2025).
UNESCO. (2021). *Global Education Monitoring Report 2021/2: Non-State Actors in Education: Who Chooses? Who Loses?*, Paris: UNESCO.
UNESCO. (2022). 'Distribution of enrollment by type of institution [data]', http://data.uis.unesco.org/ (last accessed: 18 March 2024).
Walters, Christopher. (2024). 'Empirical Bayes methods in labor economics', in (Christian Dustmann and Thomas Lemieux, eds.), *Handbook of Labor Economics*, vol. 5, pp. 183–260, Amsterdam: Elsevier.
World Bank. (2017). *World Development Report 2018: Learning to Realize Education's Promise (Overview)*, Washington DC: The World Bank.

The Economic Journal, **136** (*May*), 1254–1279 https://doi.org/10.1093/ej/ueaf089
Advance Access Publication Date: 27 September 2025