Crowding in Private Quality: The Equilibrium Effects of Public Spending in Education*

Tahir Andrabi† Natalie Bau‡ Jishnu Das§ Naureen Karachiwalla¶
Asim Ijaz Khwaja‖

January 18, 2023

Abstract

We estimate the equilibrium effects of a public school grant program administered through school councils in Pakistani villages with multiple public and private schools and clearly defined catchment boundaries. The program was randomized at the village-level, allowing us to estimate its causal impact on the market. Four years after the start of the program, test scores were 0.2 sd higher in public schools. We find evidence of an *education multiplier*: test scores in private schools were also 0.2 sd higher in treated markets. Consistent with standard models of product differentiation, the education multiplier is greater for those private schools that faced a greater threat to their market power. Accounting for private sector responses increases the program's cost-effectiveness by 85% and affects how a policymaker would target spending. Given that markets with several public and private schools are now pervasive in low- and middle-income countries, prudent policy requires us to account for private sector responses to public policy, both in their design and in their evaluation.

JEL Codes: O12, O15, H44, H52, I22, I25, I28

Keywords: Education Markets, School Grants, Spending Multiplier, Test Scores, Public and Private Schooling

---

*We thank seminar and conference participants at Columbia, Yale, Berkeley, Trinity College Dublin, Michigan, ASU, Minnesota, Melbourne, NBER Development/BREAD, IFPRI, PAC-DEV, NEUDC, IO+, and PSU for valuable comments. Catherine Michaud Leclerc provided exceptional research assistance. We acknowledge funding from the Research on Improving Systems in Education program, funded by UK Aid.

†Pomona College. Contact: tandrabi@pomona.edu

‡UCLA, NBER, and CEPR. Contact: nbau@ucla.edu

§Georgetown University and NBER. Contact: Jishnu.Das@georgetown.edu

¶IFPRI. Contact: N.Karachiwalla@cgiar.org

‖Harvard University, NBER, and CERP. Contact: khwaja@hks.harvard.edu

1 Introduction

Public interventions in service sectors, such as education and health, are commonplace and substantial. Yet such services increasingly feature both public and private providers locating and competing in the same markets. This is particularly true in low-income countries, where private providers enjoy a substantial market share.[1] An important implication of this fact is that the overall impact of investments, even when targeted to the public sector, will depend both on its direct effect in the public sector as well as the response of the private sector. In fact, a persistent concern in the literature is that the overall efficacy of public investments will be reduced to the extent that it crowds-out private sector participation (Cutler and Gruber, 1996; Dinerstein et al., 2015; Neilson et al., 2020). Conversely, improvements in the public sector that crowd-in private sector investments may be cost-effective even if the direct public sector impact is muted, precisely because high-powered incentives in the private sector can lead to a positive multiplier effect.

To assess whether such a (negative or positive) multiplier exists, we provide the first randomized evaluation of private school responses to spending on public schools. Using the randomized allocation of a public school grant program across villages with multiple public and private schools, we show that the program not only causally increased test scores in public schools but also in private schools. Furthermore, market structure played a key role in determining the size of the private sector's response, with private schools that were subject to greater competition improving more. Our results show that in the mixed public-private education systems that are now pervasive in LMICs, even policies specifically targeted to public schools will engender a wider market-level response that needs to be taken into account to determine the efficacy, targeting properties, and ex ante design of the policy.

The program we study was initiated by the Government of Punjab as a pilot program of grants to public schools administered through school-level bodies called school councils. The program sought to re-invigorate school councils with greater parental representation and the ability to fund school investments and expenditures through these school grants. In

[1]Private sector primary enrollment shares are 40% in countries such as India and Pakistan, and 28% in low and middle-income countries (LMICs) more generally, with significant penetration in rural areas (Baum et al., 2014; Andrabi et al., 2015, 2008; Kremer and Muralidharan, 2008). One consequence of this rise in private provision is the sheer density of schools, even in rural areas. For instance, the average village in this study had 7.2 schools for a population of 678 households (Andrabi et al., 2022b). Given that this study excludes large urban cities, this is an underestimate of overall market density. Beyond education, Das et al. (2020) show that the average Indian village has 3.2 healthcare providers, and every village with a public provider also has multiple private providers. Similarly, Bedoya et al. (2022) find that 79% of all health facilities in three Kenyan counties locate in markets with four or more providers.

two districts, the government agreed with our team to implement the program in randomly selected villages, drawn from a sample where we had collected data previously as part of the Learning and Educational Achievement in Pakistan Schools or LEAPS project (Andrabi et al., 2022a). The program was rolled out in 2007, and we collected endline data in 2011. This allowed us to evaluate the equilibrium effects of the program after schools had had time to improve and after substantial exit and entry in the private market.

We first verify that the program substantially increased resources in public schools. By 2011, the average public school in a treated village had received an additional PKR 122,000 or USD 1,740 in cumulative funding (using the nominal exchange rate of 70 PKR to 1 USD in 2007). This amount is equivalent to 29% of annual expenditures (inclusive of teacher salaries) at the beginning of our study period and represents a sizeable increase in discretionary resources. Also consistent with the stated aims of the program, in treated villages school councils met more frequently in 2007 and became more representative of parents in the school, although most of these effects had faded out by 2011. Our results therefore measure the effect of school grants when their use is determined through a decentralized decision-making process with oversight provided by the school council.[2]

To capture the equilibrium effects of the program, we first estimate the causal impact on test scores at the village-level. Four years after the funding started, average test scores among grade 4 children across the tested subjects of Urdu, mathematics, and English had increased by 0.18 to 0.19 sd in treated villages. This increase is substantial, lying between the 80th and 90th percentile of effect sizes in studies with large samples from low- and middle-income countries (Evans and Yuan, 2020). Importantly, our causal results at the village-level, with children tested in all schools, are uncontaminated by sorting across schools within the village. In the absence of cross-village migration, which we are able to rule out using our data, they *must* represent improvements in public and private schools rather than changes in the student body.

We next estimate treatment effects on public and private schools separately. Test scores were 0.2 sd higher in public schools in treated villages in 2011. This is an important finding given the perception that poor incentives reduce the efficacy of school grants in low-income countries (LMIC). Test scores in private schools also increased by a similar amount, with point estimates ranging from 0.16-0.32 sd. These increases do not appear to reflect changes

[2]Duflo et al. (2015), Pradhan et al. (2014), and Gertler et al. (2012) all report positive impacts of school based management in Kenya, Indonesia, and Mexico when combined with ancillary resources. One difference between previous studies and ours is the longer adjustment period, which has been shown to be important for these reforms to gather steam in the United States (Borman et al., 2003).

in the composition of children, as there is no evidence of differential sorting on the basis of parental education, wealth, or caste, all of which are associated with test scores in our setting, and no change in the private school enrollment share in response to the program. Nor do they appear to be driven by the exit of poorly-performing private schools. Instead, they appear to reflect quality improvements in existing schools.

The specific investments that led to those improvements were different in public and private schools. While there are some improvements in school infrastructure in both sectors, the more notable effects are on teachers. Public schools improved student teacher ratios by hiring more teachers on a contractual basis. While such teachers are typically less educated, and the average educational level of teachers in public schools in fact declined, the change may still lead to better overall quality outcomes. This is not only due to the direct effect of improving student-teacher ratios but also because the uncertain tenure of teachers hired on a contractual basis likely results in stronger incentives to exert effort. Contract teachers have been shown to improve test scores among children in India, Pakistan, and Kenya (Muralidharan and Sundararaman, 2013; Bau and Das, 2020; Duflo et al., 2015). In contrast, in private schools, teachers face steeper incentives but lack education and training. Consistent with this, we find that private schools responded to the program by hiring better educated and more trained teachers.

We next explore the role of market structure in determining the size of the private sector's response to the program. We examine heterogeneity along two dimensions that arguably affect the degree of competition that private schools face from public schools. As a measure of horizontal competition, we exploit the physical distance between public and private schools at baseline. Demand estimates in our previous work show that distance to a school is the strongest predictor of school choice in our context (Andrabi et al., 2022b; Carneiro et al., 2020; Bau, 2022). We therefore expect private schools that are physically closer to public schools will experience greater competitive pressure from public school improvements and improve more themselves. This is indeed the case: compared to control villages, test scores among private schools located at the 10th percentile of the distribution of distances from public schools are 0.28-0.36 sd higher with no impact on the test scores of private schools at the 90th percentile of the distance distribution.

As a measure of vertical competition, we compute the quality of public schools in the pre-treatment period using school value-added measures, as described and validated in Andrabi et al. (2022b). Consistent with better public schools exerting greater competitive pressure on private schools, test scores increase by an *additional* 0.28-0.32 sd among private schools in villages where the average public school SVA was 1 sd higher in the pre-treatment period. The heterogeneous effects of the program are thus consistent with public and private schools acting as strategic complements in our setting, as described by the theoretical model of Bulow et al. (1985).

Finally, we exploit non-experimental variation in the size of grants across villages to estimate the relationship between funding and test scores in public and private schools. We first confirm that there is zero correlation between baseline characteristics (or trends) and grant size in our data, lending credibility to the estimates. We then show that larger grants are associated with greater test score increases in both the public and private sector, but the marginal increase in the private sector is twice as high with suggestive evidence of diminishing returns to grant size in the public, but not the private sector.

Our full set of results have important implications for how government programs should be evaluated and targeted. First, the education multiplier that we uncover is large. Using the methods proposed by Dhaliwal et al. (2013), we show that accounting for private sector impacts increases the cost-effectiveness of the program by 85% from 1.18 to 2.18 test score standard deviations per 100 USD. The revised estimates place the program near the top-end for educational interventions in LMICs, and the differences are sufficiently large that they may plausibly affect the decision to implement the program in the first place (Dhaliwal et al., 2013). Second, our findings demonstrate that market structure affects the size of the multiplier and therefore, the optimal targeting of school grants. Regardless of whether the government is concerned only with cost-effectiveness or is also concerned about equity, we will show that taking private sector responses into account changes the optimal targeting of the program.

Our paper thus builds on and contributes to a growing literature on policy in markets where public and private schools interact (Muralidharan and Sundararaman, 2015; Dinerstein et al., 2015; Neilson et al., 2020; Bazzi et al., 2020; Estevan, 2015). Beyond education, the question of private sector responses when the public sector competes in the same *product market* has been investigated, for instance, in the case of health insurance where Medicaid expansion between 1987 and 1992 was shown to crowd-out private insurance (Cutler and Gruber, 1996). Given that such markets are increasingly the norm in low- and middle-income countries, ignoring private sector responses to government interventions is no longer a viable strategy; on the other hand, accounting for private sector responses can lead to better public policy.[3]

[3]As one example, Andrabi et al. (2020b) have shown previously that allocating grants to private schools

This is particularly important because the positive effect of public school investments on private school quality that we document may not apply in other settings. At the smaller levels of public investments that we study, private schools can respond by increasing quality and still make positive profits. But with very large public investments, the costs of quality improvements may be prohibitive and private schools will then exit the market. This was indeed the case for a program in New York City that increased funding to public schools (Dinerstein et al., 2015) and a public school construction program that cost 4% of GDP in the Dominican Republic (Neilson et al., 2020), both of which increased private school exits and reallocated students to public schools. Our study complements Dinerstein et al. (2015) and Neilson et al. (2020) by providing the first experimental evidence on this question and by focusing on a smaller increase in public resources. At these lower resource levels our results temper the concern that public investments in schooling will necessarily crowd-out private schooling. We are also able to inform the conditions under which public investments lead to beneficial private school responses and show that incorporating such responses changes both optimal targeting and measures of program cost-effectiveness.

Our paper also contributes to a recent revival in the literature on the direct impact of school funding on test scores in public schools (Jackson et al., 2016; Jackson, 2020; Hyman, 2017; Guryan, 2001; Neilson and Zimmerman, 2014; Lafortune et al., 2018; Card and Payne, 2002; Mbiti et al., 2019; Carneiro et al., 2020; Das et al., 2013). In the United States, Jackson shows that funding to public schools increases test scores in contrast to an older literature that argued for null effects, but has now been shown to be subject to several econometric concerns (Jackson et al., 2016; Jackson, 2020). In LMICs, the idea that public schools face poor incentives and therefore school grants do not improve test scores in the absence of additional accountability mechanisms is persuasive and has been demonstrated empirically by Mbiti et al. (2019). In contrast to these null results, we believe that multiple features of the design and use of our school grants may have helped schools in our study achieve the test score gains we observed.[4]

First, the grants were sufficiently large that parents may not have been able to fully offset them by reducing human capital investment, as has been documented in India and Zambia (Das et al., 2013). Indeed, we find that larger grant sizes were associated with

in a manner that exploits the competitive environment leads to greater improvements in test scores.

[4]Additionally, like in Jackson (2020), null results for the effect of school grants in LMICs may also reflect different statistical benchmarks in an earlier generation of studies. Studies arguing that school grants alone have no effect sometimes estimate economically meaningful but statistically insignificant effects (for example, see Pradhan et al. (2014)) or include moderately-sized effects in their confidence intervals (for example, see Mbiti et al. (2019)).

greater improvements in test scores. Second, the grants were subject to a clear accountability mechanism through the reconstituted school councils. This allowed schools to tailor their investment plans to their specific needs and reduced concerns over corruption, which paradoxically can hinder school investments.[5] Third, the grants were partly used to hire teachers on a contractual basis. Papers on India (Muralidharan and Sundararaman, 2013), Pakistan (Bau and Das, 2020), and Kenya (Duflo et al., 2015) all find a positive impact of contract teachers on test scores, even without the reduction in class sizes that we observe in our data. In Senegal, Carneiro et al. (2016) also find that grants increased test scores when schools focused on human resources rather than school materials.

The remainder of the paper is organized as follows. Section 2 describes the context and the intervention, while Section 3 describes the data and outlines the empirical strategy. Section 4 documents the program's effects on test scores, fees, and enrollment and explores potential channels through which spending may have affected school quality. Section 5 examines the role of market structure in determining the program's effects. Section 6 uses cost-effectiveness measures to quantify the size of the education multiplier and discusses implications for policy, and Section 7 concludes.

2 Context and Intervention

2.1 Context

Our study was conducted in Punjab, Pakistan's largest province with a population of 110 million. The province had 100,000 public schools and 60,000 private schools in 2016, the latter reflecting a remarkable doubling of the private sector over 20 years from 32,000 private schools in 1996 (National Center for Education Statistics, 2019). There are no restrictions on school attendance, so students can choose between sex-segregated public schools and secular, coeducational private schools, as long as they can afford the school fees. Some children also choose not to enroll in school or dropout at young ages.[6]

Public schools do not charge fees (other than a nominal administratively determined fee of PKR 10 Rupees annually in 2004) and teachers in public schools are centrally recruited at the provincial-level and then posted to schools. Teachers are required to have a Bachelor's degree

[5]School principals and head teachers often express concern that their books could be audited, and they may face disciplinary action for corruption if rules about legitimate grant expenditures are violated. As the rules are not clear, inaction (or a default action) may be an optimal strategy in many of these situations. See Bandiera et al. (2009) for an example from Italy.

[6]Fewer than 1% of students are enrolled in religious schools (Andrabi et al., 2005).

and some level of teacher training and are part of the regular civil service. Like in many LMICs, there are few accountability mechanisms for public schools in rural Pakistan. School funding is not tied to enrollment, public school teachers are never fired in our data, self-reported teacher absences are twice as high in public compared to private schools, and public schools do not appear to respond to competitive forces, even if their market shares decline (Bau, 2022; Michaud Leclerc, 2020). Recognizing the need to bolster accountability, school management committees or "school councils" were first adopted in public schools in 1994. However, multiple studies showed that they were largely ineffective due to a combination of elite capture and lack of clarity over their jurisdiction and functioning (Khan, 2007).

In contrast, private schools, while exposed to little *de facto* government regulation, face significant market pressures. They are almost exclusively dependent on revenues they raise from setting their own fees, manage their own expenses (including teacher wages), and contract with teachers independently.[7] The average school charged PKR 107 per month in 2004, which is less than 50% of average daily household income (Andrabi et al., 2008). Teachers are recruited by each school independently and are typically younger women who are local to the area with a secondary school degree. Strikingly, teacher salaries are 500% higher in public schools, and spending per student is correspondingly lower in private schools (Bau and Das, 2020; Andrabi et al., 2022b). One implication of the massive difference in salaries is that there is little overlap in the teacher labor market between the public and private sectors. Finally, despite the lower qualification of teachers in the private sector, the average private school's value-added for mean test scores is 0.15 sd greater than the average public school's (Andrabi et al., 2022b).

Our experiment was implemented in 80 villages of two districts in Punjab with an average of 2.6 private and 3.2 public schools in each village. At baseline, primary private school enrollment in these villages was 30%, although it ranged from 8% at the 5th percentile to 54% at the 95th percentile. Our ability to identify the program's effects on both the public and private sector follows from a feature of the setting where every village in our sample is a closed educational market with schools' potential competitors and household's potential choice sets clearly defined. This is because (a) households in villages are clustered and surrounded by farmland and (b) enrollment decisions are highly sensitive to distance, with more than 90% of children attending schools in their own village (Carneiro et al., 2016; Bau, 2022; Andrabi et al., 2022b).

[7] There is very little non-profit funding for such schools and limited public support programs at the time of our intervention.

A private sector that faces few regulatory requirements along with the availability of closed markets offers us a rare opportunity to identify the *equilibrium effects* of an intervention on the market as a whole, as private schools can, and will, respond to shocks to the local market. We exploit this opportunity by using a"market-level randomization" whereby our intervention is introduced in selected markets that are then tracked over time. We first used this experimental technique in Andrabi et al. (2017) and have developed it further in Andrabi et al. (2020a) and Bau (2022) to identify the equilibrium effects of providing better information, providing finance to private schools, and private school entry.

2.2 Intervention, Experimental Design, & Timeline

Intervention. The program that we evaluate had two components – providing school grants and revitalizing school councils. The school grant portion of the program provided public schools with a large, fungible infusion of cash. Under the program, schools created a list of their needs working with a well-established and highly reputable NGO, the National Rural Support Program (NRSP), and submitted funding requests to the district. From the start of the program to the end of our study period (2006-2011), 93% of public schools in treated villages received some funding, and the average school received PKR 317,348 during this period. The average yearly flow was equivalent to 14.3% of schools' operating budgets (including teacher salaries). The program was randomized at the village-level and within a village, grants were distributed to schools depending on local needs. The research team had no input on the size or distribution of grants either across or within villages. Our primary analysis therefore exploits the village-level randomization with additional non-experimental estimates of the impact of grant size on test scores.

The school councils component of the program reflected a revised 2007 policy that encouraged more frequent meetings and required greater inclusion of parents with children enrolled in the school, particularly those from disadvantaged backgrounds. School councils were required to work with NRSP and were also encouraged to discuss teacher attendance and performance, child attendance and dropout, as well as general problems faced by the school in regular meetings. Of particular interest is that the policy specifically allowed and encouraged councils to hire additional teachers on a temporary basis.[8] Our study thus cap-

[8]Specific examples of these different facets of the reform include: (a) parent members should constitute more than 50% of school council members; (b) the school council could monitor and report unexcused teacher absences to education officers; (c) the school council was responsible for planning and executing the school-based action plan and spending school council funds; (d) the school council could "appoint temporary teachers as per 'Form No.9' especially designed for 'Contract for Temporary Appointment' (Government of

tures the effects of increased public spending delivered through a mechanism that included strengthened school councils, who were responsible for both planning how these additional funds should be spent and overseeing the implementation of the plan.

Sampling. The multi-year Learning and Educational Achievements in Pakistani Schools (LEAPS) project was started in 2003 in 112 villages across three districts in Punjab to study education policy in markets with public and private schools (Andrabi et al., 2022a). All villages in the sample had at least one private school, but the sample was representative of 60% of the province's population in 2003. In 2004, the provincial government decided to pilot the intervention we study and agreed to randomize at the village-level in two districts (Attock and Faisalabad). Following an intention to treat protocol, they would implement the program in at least 80% of the villages we identified as "treatment" in these two districts and would not implement the program in the villages we identified as control. Treatment villages in the districts of Attock and Faisalabad were randomly selected from the existing LEAPS sample of 80 villages, stratified at the district-level.

Timeline. Figure 1 reports the timeline of our experiment and highlights several important features. The initial randomization of villages took place in 2003, but multiple delays including dealing with a large earthquake in Northern Pakistan in 2005 meant that initial funds were not disbursed to public schools until 2005-2006, and public schools did not receive substantial funding until 2006-2007. In 2006-2007, public schools had little time to spend the money, private schools had little time to respond, and students would have been unlikely to switch schools in the middle of the school year. Therefore, after a four-year hiatus during which there were no visits by the research team, we collected endline data in 2011 to observe the medium-term, equilibrium outcomes of the intervention. By the time the endline was conducted, respondents did not know that they had been part of a pilot program.

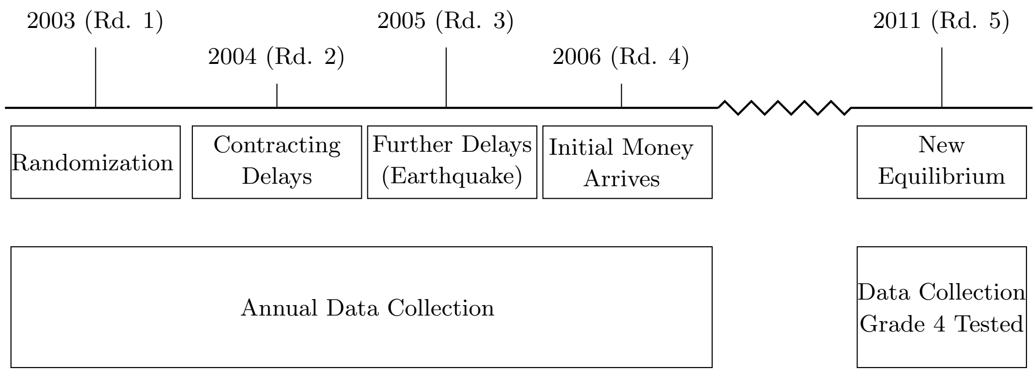
3 Data and Empirical Strategy

3.1 Data

We collected data each year for four years from 2003-2004 to 2006-2007, and then again in 2011 (Figure 1). The data collection frequency was dictated by the needs of the larger LEAPS project of which the current paper is one study (Andrabi et al., 2015, 2007, 2022b). Given the timing of the intervention, we view our data collection in 2003-2004 and 2004-2005

Punjab, 2007).

Figure 1: Timeline of Intervention and Data Collection



(rounds 1 and 2) as pre-treatment and use these data to create baseline control variables. We take a conservative approach to data collected in 2005-2006 and 2006-2007 (rounds 3 and 4). Although they were collected too early to observe equilibrium effects, we still allow for the possibility that they are impacted by the initial treatment. For our main analyses, we focus on the endline data collected in round 5 of the survey in 2011, though we exploit round 4 (2007) to document the initial roll out of the program.

Data collection in each round included school-based surveys from which all of the outcomes data used in this paper are drawn, as well as household surveys, which provide us with additional pre-treatment controls.[9]

School-Based Surveys. In every year of data collection, we compiled a list of all schools within a 15 minute walk from the perimeter of the village and accounted for any school entry or exit. Then, we collected data from each school using multiple surveys and tested children using low-stakes tests (our key outcome measure). We fielded three types of surveys in addition to testing students:

- A **child survey** on household demographics was administered to a random sample of children drawn from the test-taking sample. These data add to our pool of controls and provide us with potential covariates to study the heterogeneous effects of the intervention.

[9]In round 1, in every village, we sampled 16 households to administer surveys on demographics and educational investments, oversampling households with children enrolled in grade 3. We followed these households across all rounds of data collection.

- We also collected data from the **head teacher or owner** (in the case of private schools) on school-level covariates, including inputs and infrastructure, money received through the program, enrollment, and school fees.[10]

- Information on training, contract status, salary, and education was collected for all **teachers** from the head-teacher/owner. Additionally, teachers of tested cohorts were administered more detailed surveys on their characteristics and qualifications. Most of our non-test score outcomes come from these two sources of data on teachers, which are also used to supplement our pool of control variables.

Test Score Data. In 2003-2004 (round 1), we tested all third graders in the schools. We continued to test this cohort in the subsequent 3 rounds of data collection, as they continued through grade 6.[11] In 2011, for our endline outcomes, we tested a new cohort of fourth graders in each school. In each round, students were administered low-stakes, norm-referenced tests in math, English, and Urdu (the vernacular) that were created and validated by the research team. Following Das and Zajonc (2010), tests were scored and equated using item response theory, resulting in a distribution of test scores with a mean of zero and a standard deviation of (approximately) one. For our main results, we focus on mean test scores across the three subjects.

3.2 Empirical Strategy

Our empirical strategy allows us to measure the effects of the program on each of the key characteristics of educational markets: school quality (captured by test scores), quantity (measured by enrollment and private schools' exit and entry decisions), and prices (measured by private school fees). Our analysis first estimates the effect of the program on test scores in treated villages, allowing us to identify the net effects of the program on test scores following the strategy described in subsection 3.3 below. These estimates face the fewest threats to validity since villages are closed markets, and we test children in every school. We then separately estimate effects in the public and private sector using the strategy described in subsection 3.4 below, allowing us to better disentangle the drivers of the village-level test

[10] In 2011, recall questions were used to collect information on funding disbursed by the intervention and spending for the years between 2007 and 2011.

[11] A second cohort of third graders was also tested in 2005-2006 and followed in 2006-2007. We make only limited use of these data since, as described above, we do not consider 2005-2006 and 2006-2007 to be either part of the pre- or post-period.

score improvements. The village- and school-level estimates treat each village/school as a single treatment unit, and later in the paper, we complement these estimates with child-level estimates that are relevant for cost-effectiveness and welfare.

3.3 Village-Level Estimation Strategy

We use the following regression equation to estimate the net effect of being a randomly treated village on test scores:

$$y_{v,5} = \alpha_d + \beta_1 I_v^{Treatment} + \sum_{t=\{1,2\}} \mathbf{\Gamma}_{\boldsymbol{t}} \mathbf{X_{vt}} + \varepsilon_{v,5}, \tag{1}$$

where v denotes a village, t indexes rounds, and d denotes a district; $y_{v,5}$ is the average test score for village v in round 5, $I_v^{Treatment}$ is an indicator variable equal to 1 if village v was randomly selected for the program, α_d is a district fixed effect, and $\mathbf{X_{vt}}$ is a vector of controls from the pre-treatment periods ($t = 1, 2$). Therefore, β_1 identifies the causal effect of the program on test scores in treated markets.[12] By focusing on outcomes at the *village-level,* this specification ensures that we identify the net effects of the program on test scores and eliminates the possibility of changes in the composition of students in a school or sector biasing our results. The only possible source of contamination in our data is that households moved across villages in response to the program, a possibility that we are able to test and reject in Section 3.5 below.

Since the randomization was stratified by district, we always control for district fixed effects, but our most parsimonious regressions do not include other control variables (Glennerster and Takavarasha, 2013). In additional specifications, to improve precision and account for any chance imbalances, we (1) control for the baseline outcome variables at the village-level $y_{v,1}$ and $y_{v,2}$ from rounds 1 and 2 and (2) employ the double-lasso procedure of Urminsky et al. (2016) to select control variables from a large pool of pre-treatment measures. The double-lasso procedure selects the control variables that best predict the outcome variable (to improve precision) and best predict $I_v^{Treatment}$ (to improve balance).[13] Thus, it provides us with a principled method to select controls for inclusion and also provides re-

[12] Our results are thus intention to treat rather than local average treatment effects. This reflects the fact that virtually all public schools in our treatment sample received a non-zero grant, and the grants were accompanied with changes to their school councils. We return to the question of per-dollar spending effectiveness when we examine the impact of grant size on test scores and when we present cost-effectiveness estimates.

[13] In cases where the value of a control variable is missing, we code the missing value as 0 and include an indicator variable that is equal to 1 if the value is missing.

assurance that the results are robust to accounting for any chance imbalances. Appendix Table A1 lists the pool of 345 potential controls that the double-lasso procedure selects over and notes which data source each control variable is drawn from.[14]

3.4 Effects by Sector

In addition to estimating village-level effects, we also examine how the effects of the program vary by sector. Therefore, we separately estimate effects in the public and private sectors using school-level regressions that weight each school equally. The regression specification is analogous to equation (1), but an observation is at the school rather than the village-level. Our outcome variables consist of mean school-level test scores, enrollment, school composition measures, private school fees, and private school exit. As a village is the unit of randomization, we cluster our standard errors at the village-level.

Estimates of the treatment effects in regressions where the outcomes are school-level average test scores may reflect (1) intensive margin within-school quality changes, (2) extensive margin quality changes due to the entry and exit of schools, and (3) changes in school composition due to differential sorting of students across sectors. To disentangle these different drivers, in Section 4, we explore how the program directly affected these margins.

School entry and exit also complicates the interpretation of school-level effects on non-test score outcomes. For example, if low fee schools were more likely to exit in treatment villages, we will observe positive effects on fees in the school-level regressions even if there are no intensive margin, within-school changes in fees. Thus, by examining the effects of the program on exit and entry, we will also be able to determine if other treatment effects are driven by intensive margin changes within schools or changes in the composition of school types in the market.

Finally, one baseline variable of particular interest is our measure of school quality, or School Value Added (SVA). Following Andrabi et al. (2022b), we compute SVA in mean test scores for each school in round 2 and use Empirical Bayes to correct for estimation error. As lagged test scores are needed to estimate SVA, we cannot calculate value-added in round 1. We average across public/private schools in the village and then normalize the village-level average to have a mean of 0 and standard deviation of 1. The specific methods are detailed

[14]The additional, non-parsimonious specifications also include a control for whether a village took part in a report card experiment between 2003 and 2004 that provided parents with information on schools' average test scores and their own child's performance. This treatment was independently randomized at the village-level (Andrabi et al., 2017).

below in Appendix A. We use the SVA measures to examine heterogeneity in private school responses by public school quality.

3.5 Balance and Attrition

Unbiased estimates of the treatment effect require that randomization results in balanced characteristics across treatment and control villages, and that there is no differential attrition from the treatment. We assess the validity of these assumptions below.

Balance. Appendix Tables A2 and A3 verify that the randomization was balanced at the village-level, as well as the school-level within sectors. Across a rich set of pre-treatment covariates from round 2 (the last pre-treatment year), there are no significant differences between treatment and control villages. The p-value from an F-test to assess the joint significance of all the covariates in predicting treatment status is 0.56, further confirming that village-level characteristics are balanced.[15] At the sector-level, there is only one marginally significant difference between treatment and control villages in one sector, and it indicates that test scores in the public sector in treatment villages were *lower* than control villages. An F-test of the variables cannot reject that these covariates do not jointly predict treatment status in either the public or private sector (p-value = 0.22 and 0.68, respectively).

Attrition. Attrition may bias our village-level regressions if the treatment changed migration patterns, leading to differences in the tested populations in treatment and control villages by 2011. To assess whether this is the case, we used the household survey data and estimate the effect of the program on an indicator variable that equals one if a household moved away from the village by 2011. Appendix Table A4 shows that both the probability of migration and of completing the survey in 2011 are not affected by the treatment. This aligns with our understanding that, in Punjab, households are unlikely to migrate in response to a government education program. A second source of potential attrition is survey refusals among schools. In 2011, 6 schools out of 441 refused to participate, two private and two public schools in the treatment villages and two private and zero public schools in the control villages. The very small numbers of refusals precludes further statistical analysis but suggests that school refusals are unlikely to drive our results.

[15] We also confirmed that there is no imbalance in the pre-treatment characteristics from round 1 (results available on request).

4 Results

This section reports our main results. We first confirm that the treatment led to greater public school funding and changes in school councils. We then estimate the effect of the intervention on our key outcomes of interest in 2011: test scores, school composition, entry and exit, enrollment, and private school fees. Finally, we explore the channels for school improvement.

4.1 School Funding and Councils

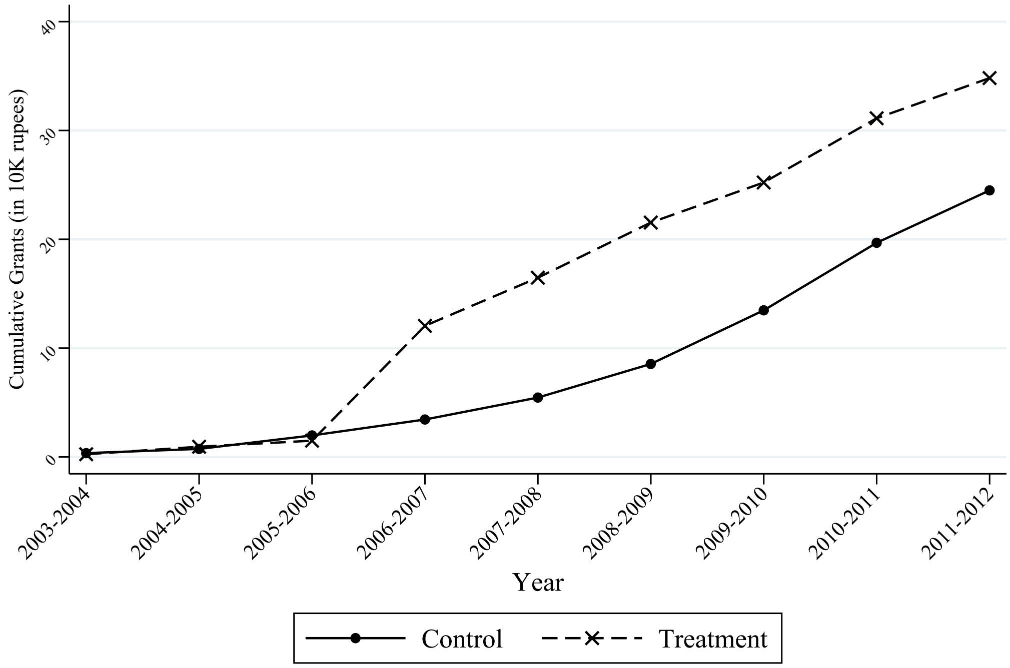
We first assess if the program resulted in differences in school funding and school council activity between treated and control villages, as well as the extent to which these differences persisted until 2011.

Differences in Funding. Figures 2 and 3 show the average cumulative funding by year to public schools in treatment and control villages and the annual flows by treatment status respectively. While there was a general trend of increased funding to public schools over time, funding flows to public schools in treated villages were substantially higher in 2006-2007 and remained elevated for the subsequent two years. This resulted in a persistent cumulative difference in funding from grants between the average treatment and control school of PKR 122,000, which is a 42% increase over the funding that control schools received in this period and amounts to 29% of median annual public school expenditures (inclusive of teacher salaries). Our causal estimates thus capture the equilibrium effects of this difference in cumulative funding.

Panel A of Table 1 reports the coefficients from regressions of village-level measures of cumulative funding in **2006-2007** (round 4) on treatment status, with coefficients reported in PKR 10,000. Treatment villages received an average of PKR 325,888 more than control villages (column 1). Public schools received an additional PKR 74,820 each (column 2), equivalent to an additional PKR 540 per student enrolled in the school (column 3) or PKR 770 per primary school student enrolled in the school (column 4).[16] Panel B shows that differences in cumulative funding between the treatment and control villages persisted into 2011. By this time, public schools in treatment villages had received PKR 492,020 more funding compared to those in control villages (column 1), which corresponds to PKR 121,700

[16]Most schools only offer primary school education, but some schools offer middle school classes as well.

Figure 2: Cumulative Amount of Funding Disbursed



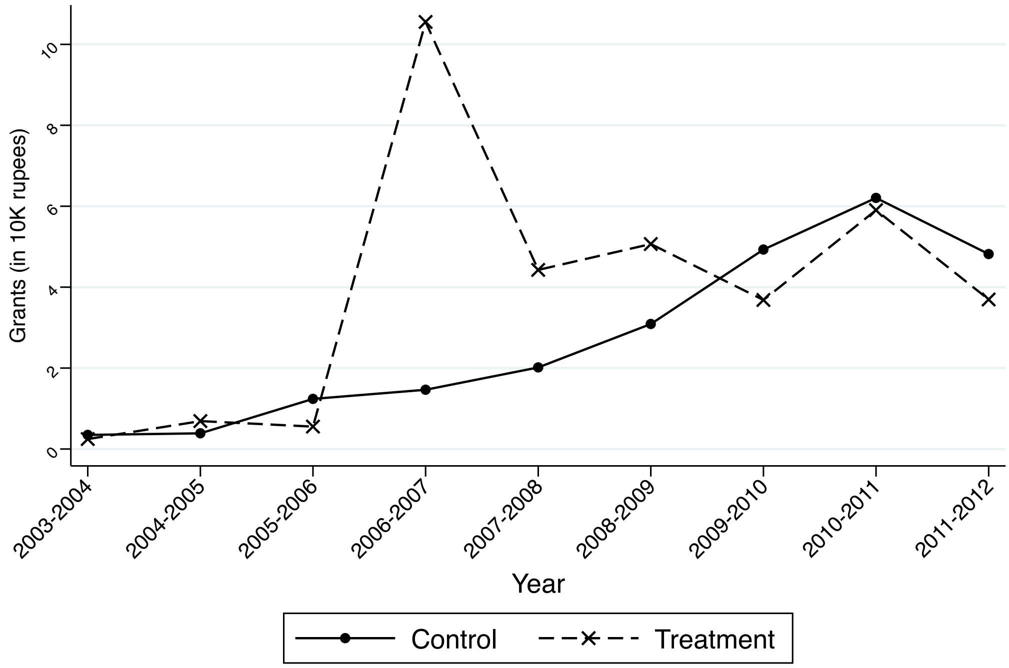
Notes: This figure plots the cumulative amount of funding received by public schools each year (in 10,000 PKR) by treatment arm, as reported by the school principal. For years between 2006-07 (round 4) and 2011-12 (round 5), recall data from the round 5 survey is used.

per school (column 2), PKR 1,070 per enrolled student (column 3), and PKR 1,480 per primary school student (column 4).

While the differences in Table 1 are generated by experimental variation, there is also non-experimental variation across and within treatment villages in the size of the grants. For instance, treatment villages at the 90th percentile of the grant distribution received PKR 737,601 compared to PKR 128,999 for villages at the 10th percentile. Schools at the 90th percentile of the grant distribution received PKR 589,722 compared to PKR 4,776 for schools at the 10th percentile. We will use this non-experimental variation to estimate the relationship between grant size and test score improvements.

Strengthening School Councils. Table 2 measures the effects of the intervention on school council characteristics. In 2007, school councils in treated villages met an additional 1.6 times per year compared to the mean in the control group of 7.4 meetings (column 1). Council members were (an insignificant) 6.4 percentage points (10%) less likely to own

Figure 3: Average Yearly Funding Flows



Notes: This figure plots the average amount of funding received by public schools each year (in 10,000 PKR) by treatment arm, as reported by the school principal. For years between 2006-07 (round 4) and 2011-12 (round 5), recall data from the round 5 survey was used.

land (column 2) and 7.4 percentage points (24%) more likely to have a member who was less educated (had at most a primary level education; column 3). This suggests that the school councils in treated villages diversified and became more socioeconomically inclusive, in line with the goals of the program.[17] School councils in treated villages also increased the share of parents whose children were enrolled in the school by 12.5 percentage points (38%) compared to the control mean of 32% (column 4), potentially improving accountability and giving parents a voice in the creation of the school investment plans.

Panel B shows that for most of these outcomes, the differences between schools in treatment and control villages have disappeared by round 5. School council members were still more likely to have a child enrolled in the school in treated villages, but the difference re-

[17] In 2007, 46% of fathers in public schools had not completed primary school (the share was even lower for mothers) but 71% of school council members in control villages were educated beyond primary school. Similarly, 49% of parents of children in public schools owned agricultural land relative to to 64% among school council members in control villages.

Table 1: Effect on Cumulative Funding to Public Schools

	Cumulative Grants (in 10K Rs) in the Public Sector			
	(1) Total	(2) Per School	(3) Per Student Enrolled	(4) Per Student Enrolled, Primary
Panel A: Round 4				
Treatment	32.588**	7.482**	0.054***	0.077***
	(16.278)	(3.309)	(0.020)	(0.028)
Control Mean	10.314	4.057	0.018	0.030
Adjusted R^2	0.032	0.058	0.061	0.068
Observations	80	80	80	80
Panel B: Round 5				
Treatment	49.202**	12.170**	0.107**	0.148**
	(21.716)	(5.131)	(0.045)	(0.060)
Control Mean	73.485	27.371	0.141	0.234
Adjusted R^2	0.035	0.106	0.040	0.044
Observations	80	80	80	80

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports treatment effects on the cumulative amount of funding that was disbursed to all public schools in the village by round 4 (immediately after the program began) in Panel A and by round 5 (4 years after the program began) in Panel B. Amounts are reported by school principals based on recall data in school surveys and are in 10,000 PKR. Column 1 shows effects on the total amount of funding received by all public schools in the village, column 2 shows the amount received per public school in the village, column 3 shows the amount received per student enrolled in public schools in the village, and column 4 shows the amount received per student enrolled in primary grades in public schools in the village. All regressions control for district fixed effects (the stratifying variable), and standard errors are clustered at the village-level.

duced to 6 percentage points on a lower base of 23.9%. The lack of a statistically significant treatment effect in 2011 is not due to a lack of statistical precision. For the number of meetings and the education of members, the signs are reversed, and for land ownership, the difference is close to 0. In the case of school council meetings, this could reflect an increase in the number of meetings in control villages, but the variables that capture council composition either remained the same or became less representative of the parent body in control villages.

Table 2: Effect on Public School Councils

		Proportion of members		
	(1)	(2)	(3)	(4)
	# Meetings	Own land	Prim. educ. or less	Has child enrolled
Panel A: Round 4				
Treatment	1.568***	-0.064	0.074**	0.125***
	(0.451)	(0.044)	(0.032)	(0.028)
Control Mean	7.433	0.643	0.290	0.321
Adjusted R^2	0.071	0.045	0.059	0.161
Observations	242	242	242	242
Clusters	80	80	80	80
Panel B: Round 5				
Treatment	-0.030	-0.029	0.020	0.061**
	(0.479)	(0.047)	(0.037)	(0.027)
Control Mean	9.902	0.661	0.311	0.239
Adjusted R^2	0.082	0.097	0.081	0.178
Observations	231	231	231	231
Clusters	80	80	80	80

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table shows the differences between treated and untreated villages' school council characteristics in public schools in round 4 (immediately after the program began) in Panel A and in round 5 (4 years after the program began) in Panel B. The outcome in column 1 is the number of school council meetings held in the past school year. columns $2-4$ report effects on demographic characteristics of school council members: the share that own land, the share with a primary school education or less, and the share with a child enrolled at the school. All regressions control for district fixed effects (the stratifying variable), and standard errors are clustered at the village-level.

4.2 Test Scores

We now turn to our main outcome of interest, which is how the public school grant program impacted student test scores for children tested in Grade 4.

Village-Level Estimates. Table 3 reports the treatment effects at the village-level for test scores across all students (averaging over maths, English, and Urdu), estimated using equation (1). Column 1 controls only for the randomization stratification (district fixed effects), column 2 includes round 1 and round 2 village-level test score measures, and column 3 additionally includes the controls selected by the double-lasso procedure.

The estimates are similar across all specifications and show that test scores were 0.15-0.19 sd higher in treated villages by 2011. Recall that Appendix Table A4 had shown there is no differential migration or survey non-response by treatment status, highlighting that the village level results are not affected by changes in student composition, but rather reflect higher performance across similarly aged children.

Figure 4 confirms that the timing of the divergence in test scores occurs only after the receipt of funding. There are no statistically significant differences in test scores between treated and control villages in each of the four rounds of testing through 2007 and then a sizeable positive effect in 2011. This difference is equivalent to almost 6 months of schooling in the LEAPS sample, so that by the time they are in 4th grade, children in treated villages are half-a-year ahead, placing this program between the 80th and 90th percentile of effect sizes (0.14 sd and 0.24 sd) among studies of educational interventions with more than 5000 children (Evans and Yuan, 2020).

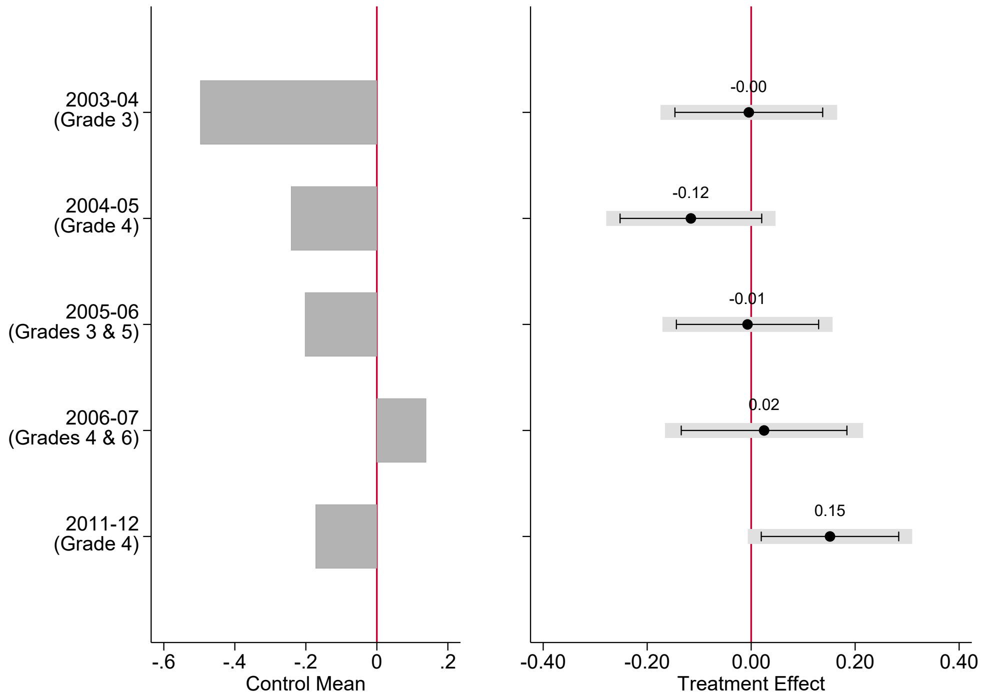
Table 3: Treatment Effects on Village-Level Mean Test Scores

	(1) OLS	(2) OLS	(3) Lasso
Treatment	0.152*	0.180**	0.191**
	(0.079)	(0.080)	(0.086)
Control Mean	-0.233	-0.233	-0.233
Baseline Controls	No	Yes	Yes
Adjusted R^2	0.529	0.560	0.626
Observations	80	80	80

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports village-level estimates of the effect of the program on average test scores (across tests in math, English, and Urdu and across all students in the village) in round 5 (2011-12). Tests are scored using item response theory (IRT), and test scores are measured in standard deviations. The first column controls only for district fixed effects (the stratifying variable), the second column additionally controls for the baseline values of the dependent variable from rounds 1 and 2, and the third column uses a post double-lasso procedure to select additional baseline controls. The second and third columns also include a control for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

Within-Sector Estimates. Table 4 reports school-level estimates for all (columns 1-3), public (columns 4-6), and private schools (columns 7-9) separately. As before, we report estimates from the parsimonious specification, the specification with controls for pre-treatment outcomes, and the double-lasso specification, treating each school as a single unit. Across all

Figure 4: Treatment Effects by Survey Round



Notes: This figure displays the treatment effects of the program on test scores in each of the survey rounds. Test scores are village-level averages (across tests in math, English, and Urdu and across all students in the village). Tests are scored using item response theory (IRT) with all tests on the same scale, and test scores are measured in standard deviations. Each estimate is derived from a separate regression with observations at the village-level. Only district fixed effects are included as controls in the regressions. Control group means are reported in the left panel, and the treatment effects are reported in the right panel. Dots represent the treatment effect, lines show the 90% confidence intervals, and the shaded area shows the 95% confidence interval.

schools, the intervention increased average test scores by 0.19-0.27 sd, with similarly-sized effects for both public and private schools.[18] These effects were larger for mathematics (0.30 sd) and Urdu (0.29 sd) in the public sector, and stronger for Urdu (0.20 sd) and English (0.23 sd) in the private sector (see subject-specific results in Appendix Table A5). This is consistent with a documented demand for and advantage in English instruction among private schools (Muralidharan and Sundararaman, 2015).

[18]The double-lasso estimate for mean test scores in private schools is somewhat larger, though not statistically significantly different, from the estimates with only district and baseline controls. Interestingly, this is not due to any imbalance in the randomization; the lasso does not select any covariates to predict the treatment variable.

Table 4: Effects on School-Level Mean Test Scores

	All Schools			Public Schools			Private Schools		
	(1) OLS	(2) OLS	(3) Lasso	(4) OLS	(5) OLS	(6) Lasso	(7) OLS	(8) OLS	(9) Lasso
Treatment	0.192**	0.216***	0.267***	0.194**	0.220**	0.209**	0.162	0.198**	0.324***
	(0.084)	(0.075)	(0.073)	(0.094)	(0.090)	(0.102)	(0.108)	(0.089)	(0.122)
Control Mean	-0.157	-0.157	-0.157	-0.550	-0.550	-0.550	0.310	0.310	0.310
Baseline Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Adjusted R^2	0.221	0.250	0.274	0.303	0.300	0.290	0.202	0.298	0.307
Observations	428	428	428	231	231	231	193	193	193
Clusters	80	80	80	80	80	80	74	74	74

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports treatment effects of the program on school-level average test scores (across tests in Math, English, and Urdu and across all students in the school) in round 5 (2011-12). Tests are scored using item response theory (IRT), and test scores are measured in standard deviations. All schools are included in columns 1-3, public schools in columns 4-6, and private schools in columns 7-9. Each set of three columns follows the same format. The first column only controls for district fixed effects (the stratifying variable), the second column additionally controls for the baseline values of the dependent variable from rounds 1 and 2, and the third column uses a post double-lasso procedure to select additional baseline controls. The second and third columns also include a control for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

We further explore heterogeneity in treatment effects by sex and assets using child-level regressions. In public schools, the points estimates are larger for girls, but the difference is not statistically significant, while in the private sector, the effects are statistically significantly larger for boys (Appendix Table A6). This is interesting because private schools are coeducational, and boys and girls are taught in the same classroom, making it harder to differentiate instructional quality by sex. One potential explanation is that private schools in treated villages re-targeted their instructional quality towards children with lower levels of learning, who were also the children most likely to improve in response to improvements in public schools, a mechanism discussed by Bau (2022). Since girls score higher than boys, this re-targeting would manifest as a gender difference although the true heterogeneity is by baseline test scores (which we do not observe at the child-level for the endline sample since we do not have a child-level panel in 2011). In contrast, we find no evidence of heterogeneous effects by wealth in either sector (Appendix Table A7).

Changes in Enrollment/Student Composition. With no cross-village migration in response to the intervention (Appendix Table A4), village-level improvements in aggregate test scores must reflect quality improvements within schools. However, improvements in test scores in public and private schools could also reflect student sorting. To test if this is the case, we estimate the effect of the treatment on school-level enrollment and composition in the public and private sectors.

To assess the impact of the intervention on enrollment, we first estimate the effect of the intervention on overall and sectoral primary enrollment (grades 1-5) in 2011 (Table 5). The null results in columns 1-3, which report the effects on overall enrollment, indicate that improving public school quality is insufficient to induce additional enrollment among children who do not attend school. The remaining columns show that the policy did not significantly affect enrollment in the public (columns 4-6) or private (columns 7-9) sectors either. Thus, the null effects on enrollment provide initial evidence that the school-level test score effects in both sectors are unlikely to be driven by changes in sorting.

Table 5: Effects on Primary Enrollment

	All Schools			Public Schools			Private Schools		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	OLS	Lasso	OLS	OLS	Lasso	OLS	OLS	Lasso
Treatment	-4.338	-5.174	-3.884	-1.912	-3.125	-0.105	-6.719	-6.092	-3.299
	(8.599)	(4.773)	(5.186)	(15.164)	(6.818)	(6.482)	(9.127)	(6.891)	(6.466)
Control Mean	114.476	114.476	114.476	130.206	130.206	130.206	94.965	94.965	94.965
Baseline Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Adjusted R^2	0.002	0.124	0.120	0.027	0.427	0.446	-0.004	0.107	0.120
Observations	439	439	439	232	232	232	202	202	202
Clusters	80	80	80	80	80	80	74	74	74

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports estimates of the program's effect on primary grade enrollment (grades 1-5) at the school-level in round 5 (2011-12). All schools are included in columns 1-3, public schools in columns 4-6, and private schools in columns 7-9. Each set of three columns follows the same format. The first column only controls for district fixed effects (the stratifying variable), the second column additionally controls for the baseline values of the dependent variable from rounds 1 and 2, and the third column uses a post double-lasso procedure to select additional baseline controls. The second and third columns also include a control for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

The lack of enrollment effects should not be interpreted to mean that parents are unaware of school quality or that schools do not experience competitive pressures. In previous work, we show that there are close links in the cross-section and over time between school quality as measured by value-added in test scores and market shares, particularly in the private sector (Andrabi et al., 2022b). We may observe null effects because enrollment is endogenous to the quality investments made by the schools, and we do not observe counterfactual enrollment decisions without quality improvements in the private sector.

To further rule out the possibility that the school-level effects are driven by sorting, we next test whether the treatment affects the composition of the student body even if enrollment is unchanged. In the absence of a student-level panel for our round 5 students, we cannot directly test whether specific students switch schools. Therefore, in the first 12 columns of Appendix Table A8, we estimate the effect of the intervention on four measures of socioeconomic status, each of which has been shown to be correlated with test scores in our context: (1) the share of students from low caste groups,[19](2) the share of students whose mothers have some education, (3) the share of students whose fathers have some education, and (4) the average asset index of students enrolled in the school, for all schools, public schools, and private schools.[20]

The point estimates are all small and insignificant, with only one marginally significant coefficient across 12 specifications and no consistent patterns in the direction of the coefficients. There is no evidence that the intervention changed the pattern of sorting into the public or private sector. To further ensure that our test score estimates in the private sector are not driven by certain types of students switching into the public sector, in column 13, we include all the (potentially endogenous) measures of school composition as controls in the school-level regression of average test scores on the treatment. Despite the inclusion of these endogenous controls, the treatment effect remains large and robust (0.2 sd) and is comparable to our preferred estimates in Table 4. Given the results in Table 5 and Appendix Table A8, we conclude that the within-sector test score estimates are unlikely to be driven by changes in school composition and instead reflect the causal effects of the program on learning in both sectors.

4.3 Exit and Entry

While the analysis in the previous subsection suggests that the program causally increased learning in the private as well as the public sector, this improvement could be due to either intensive margin changes in school quality or the exit of poorly-performing private schools. To disentangle these mechanisms, we first estimate the effect of the intervention on private

[19]In Pakistan caste is not a religious institution. As often done in the literature (Karachiwalla, 2019), we are using the term caste to refer to a social construct that reflects an individual's clan or tribe (zaat/biraderi). Since such (zaat/biraderi) groups often have persistent socioeconomic differences, one can classify them into "high and low caste" categories (Mohmand and Gazdar, 2007).

[20]Several papers show that these socioeconomic covariates are important correlates of learning in this context: Karachiwalla (2019) on caste, Andrabi et al. (2012) on mother's education, and Das et al. (2022) on assets. To reduce the table size, we only report the results from the double-lasso specification. The share of low caste students is from the school surveys, and the other measures are from the child surveys. Whether a child is from a low caste group is determined using the classifications from Karachiwalla (2019).

school entry and exit.[21]

Columns 1-3 of Table 6 show that there was no significant effect on the number of private schools in a village. Nevertheless, there is some evidence that the intervention increased churn in treated villages, with a marginally significant 0.12 increase in the number of entrants (normalized by the baseline number of private schools) in columns 4-5 and an insignificant 5-7 percentage point increase in likelihood of exit in columns 6-7. However, any additional churn did not affect many students. Appendix Table A9 estimates that the market share of entrants increased by an insignificant 1.3-1.5 percentage points, and exiting schools accounted for an insignificant additional 1.9-2.3 percent of the market in treated villages.

While this implies that exit and entry are unlikely to account for test score improvements in the private sector, to further assess this possibility, in Appendix Table A10, we also examine whether the treatment led to differential exit for lower quality schools as measured by pre-treatment school value-added and find no evidence that this is the case.[22] We conclude that improvements in test scores in the private sector are more likely to be driven by intensive margin quality improvements in existing private schools.[23]

4.4 Private School Fees

While the estimates in the previous subsections show that school quality increased in both the private and public sectors, this does not necessarily imply that households' welfare universally increased. If private schools paid for quality increases by charging higher fees, welfare for private sector consumers may have fallen. To examine whether this is the case, in Table

[21]Exits among public schools are too infrequent to allow for a meaningful differentiation by treatment status. In our data, 5 public schools enter (4 in treatment villages) and 19 exit between 2007 and 2011 (13 in treatment villages).

[22]Recall that we compute school value-added (SVA) in mean test scores – the predicted effect of attending a given school for one year on test scores – using data from rounds 1-2 and use empirical Bayes to correct for estimation error as described in Appendix A.

[23]We also evaluated and found little evidence for two additional routes to the test score improvements in private schools. One is that test score improvements among public school children induced similar improvements in private school children through interactions outside of school. The peer effect literature suggests that even when they are in the same classroom, effect sizes are much smaller than the one-for-one gains that we document here (Hoxby, 2000). Here, the possibility is even more remote since private school children are two grade-levels ahead in their test scores on average, and there is no evidence in the literature that marginal improvements among lower performing children can lead to a positive spillover for higher performing children. A second possibility is that there were spillovers across treatment and control villages through the teacher labor market. Perhaps talented teachers left public schools to join private schools or, alternatively, left control for treatment villages. As to the first, salaries are 3-5 times higher in the public sector, so any movement from the public to the private sector would lead to a substantial pay cut. The second possibility is also remote since our sample consists of 80 villages spread across two districts, each of which is 2,500 square miles in area, making it extremely unlikely that this kind of cross-market contamination drives our results.

Table 6: Effects on Private School Entry and Exit

	Number of Private Schools			Entry		Exit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	OLS	Lasso	OLS	Lasso	OLS	Lasso
Treatment	0.553	0.193	0.216	0.123*	0.123*	0.073	0.050
	(0.431)	(0.201)	(0.233)	(0.064)	(0.064)	(0.065)	(0.068)
Control Mean	2.289	2.289	2.289	0.182	0.182	0.300	0.300
Adjusted R^2	0.011	0.807	0.798	0.022	-0.004	-0.002	0.030
Observations	80	80	80	80	80	209	209
Clusters	80	80	80	80	80	78	78

Notes: $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports estimates of the program's effect on the number of private schools in a village (columns 1-3), the number of private entrants between rounds 2 and 5 (2004-05 and 2011-12) normalized by the number of private schools in the village in round 2 (columns 4-5), and the likelihood of a private school exiting the market between rounds 2 and 5 (columns 6-7). In columns 6-7, the sample is restricted to private schools open in round 2. Columns 1, 4, and 6 only control for district fixed effects (the stratifying variable), column 2 adds the baseline values of the outcome variable (note that it is not possible to include a "baseline" value of the private school exit and entry outcomes) and a control for the report card intervention, and columns 3, 5, and 7 use a post-double lasso procedure to select additional village-level baseline control variables and also include a control for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

7, we estimate the effect of the policy on (log) private school fees. We find no statistically significant impact though the point estimate is positive, consistent with some pass-through to parents.[24]

4.5 Channels for School Improvements

We next use the richness of the LEAPS data to explore *how* schools improved their quality in response to the intervention, focusing on personnel and infrastructure. In Table 8, we report the effect of the treatment on a range of school inputs from student-teacher ratios to teacher characteristics and infrastructure.Given the range of potential outcomes and the exploratory nature of this analysis, for the impact on infrastructure, we report average effect

[24]Taking the point estimate seriously, we can compare the price increase to the hedonic estimates in Andrabi et al. (2022b), who document that a 1 sd increase in test scores was associated with a 68% increase in private school fees in these villages. Since that relationship is linear and our test score impacts are 0.2 to 0.3 sd, price increases of 13.6% to 20.4% would be in line with the quality improvements we observe. The lower price increase we observe is consistent with standard oligopoly models where vertical and horizontal differentiation confer some degree of market power that is reduced when the quality of the outside option improves.

Table 7: Effects on Log Fees in the Private Sector

	(1) OLS	(2) OLS	(3) Lasso
Treatment	0.113	0.095	0.096
	(0.083)	(0.068)	(0.075)
Control Mean	7.937	7.937	7.937
Baseline Controls	No	Yes	Yes
Adjusted R^2	0.080	0.174	0.208
Observations	200	200	200
Clusters	74	74	74

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports treatment effect estimates for the program on private school fees in round 5 (2011). The outcome variable is the natural log of annual school fees charged to students, as reported by the school principal/owner. Column 1 controls only for district fixed effects (the stratifying variable), column 2 adds the round 1 and round 2 village-level (baseline) values of the dependent variable (if available), and column 3 uses a post double-lasso procedure to select additional baseline controls. Columns 2 and 3 also include a control for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

sizes to address the fact we have multiple outcome variables, with effects on the components of each infrastructure index reported in the appendix (Appendix Tables A11 and A12).[25] Panels A and B report results for public and private schools separately, and Panel C reports the difference between the two. We report differences here to highlight differences in the investment strategies of the two different sectors.

Personnel. The most important changes relate to teachers in public and private schools. For public schools in treated villages we find that there is a decline in the student teacher ratio of 0.17 log points or from 31 to 26 students per teacher (column 1). While there is a decline in the private sector as well it is smaller and not statistically significant. Consistent with the decline in the student teacher ratio, column 2 of Table 8 shows that public schools hired 0.18 more teachers on temporary contracts (doubling the total number of contract teachers), though this effect is not statistically significant. [26] This is in line with the stated

[25]Our calculation of average effect sizes follows Kling et al. (2007) and Clingingsmith et al. (2009). The average effect is calculated by first using the control group to standardize each outcome variable. Then, outcome-specific coefficients on treatment are generated using the standardized outcome variables with a seemingly unrelated regression, and these coefficients are linearly combined to arrive at an average effect size.

[26]We do not report this result for private schools because all teachers are hired on temporary contracts in the private sector.

Table 8: Evidence on Channels

	Log STR	Num. Contract Teachers	BA Plus	Some Training	Log Salary	Basic Fac.	Extra Fac.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Public Schools							
Treatment	-0.170**	0.183	-0.122***	-0.022	-0.002	0.084	0.092
	(0.072)	(0.120)	(0.034)	(0.020)	(0.042)	(0.084)	(0.062)
Control Mean	3.438	0.186	0.628	0.959	9.817		
Adjusted R^2	0.209	0.010	0.081	0.188	0.037		
Observations	232	232	232	232	232	232	232
Clusters	80	80	80	80	80	80	80
Panel B: Private Schools							
Treatment	-0.108		0.086*	0.066**	0.069	0.166*	0.065
	(0.072)		(0.045)	(0.030)	(0.062)	(0.090)	(0.061)
Control Mean	2.569		0.266	0.156	7.491		
Adjusted R^2	0.064		0.059	0.026	0.169		
Observations	202		202	202	200	202	202
Clusters	74		74	74	74	74	74
Panel C: Private-Public Difference							
Treatment	0.062		0.209***	0.088**	0.070	0.116	-0.010
	(0.099)		(0.056)	(0.034)	(0.077)	(0.126)	(0.077)
Control Mean	3.043		0.464	0.594	8.760		
Adjusted R^2	0.490		0.206	0.787	0.903		
Observations	434		434	434	432	434	434
Clusters	80		80	80	80	80	80

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports the effect of the program on the characteristics of teachers, student teacher ratios, and school facilities in public (Panel A) and private schools (Panel B) in round 5 (2011-12), as well as the differences in the treatment effects across sectors (Panel C). The outcomes are the log of the school's student-teacher ratio (column 1), the number of contract teachers in the school (column 2), the share of teachers with a bachelor's degree or higher (column 3), the share with some teaching-specific training (column 4), the log of the average teacher salary (column 5), and the average effect sizes for basic and extra facilities (columns 6 and 7). We do not report contract teacher effects for the private sector since all private school teachers are essentially contract teachers. All columns control for district fixed effects (the stratifying variable), use a post double-lasso procedure to select controls, include baseline village-sector level controls for the outcome when available, and include a control for a report card intervention (Andrabi et al., 2017). The average effect column uses the control selected by double-lasso when the outcome is the first principal component from a principal components analysis of the outcomes over which the average effect size is taken. Standard errors are clustered at the village-level.

policy aims of the grant. Column 3 shows that, in the public schools, the average level of education among teachers in treated villages was lower. This too is consistent with hiring more contract teachers, who are less qualified.

On the other hand, the education and qualification of teachers in private schools increased in treatment villages (columns 3 and 4). Teachers were 8.6 percentage points (32%) more likely to have at least a bachelor's degree and 6.6 percentage points (44%) more likely to have some teaching-specific training. Panel C shows that this difference in the impact of the treatment on teacher education and qualifications across public and private schools is significant.

These results are consistent with the observation that the significant constraint on teacher performance in public schools is incentives and in private schools is qualifications. In addition to the direct effect of a lower student-teacher ratio, multiple studies, including Bau and Das (2020) from Pakistan, show that contract teachers causally increase test scores, a result that is consistent with the higher incentives they face (Duflo et al., 2015; Muralidharan and Sundararaman, 2013). On the other hand, teachers in private schools have less training but face high-powered incentives since few have permanent contracts, and wages are responsive to their value-added (Bau and Das, 2020).

Physical Investments. We also evaluated whether physical investments changed in response to the intervention, categorizing school infrastructure as "basic" or "extra."[27] While all the estimates are positive, they are generally imprecise with the exception of a marginally significant improvement in basic facilities for private schools which is driven primarily by a significant increase in blackboards per student (Appendix Table A11). Public schools show an increase in semi-permanent classrooms per student (Appendix Table A11). Another interesting result is that virtually every public school that did not previously have a boundary wall built one under the program (see Appendix Table A12). Parents often demand walls to address safety concerns. Thus, even if more semi-permanent classroom and walls may not increase test scores, any public school investment that increases demand for the school will exert competitive pressure on private schools, who may then respond by increasing test scores. Indeed, Carneiro et al. (2016) show that parents in this context directly value school attributes such as infrastructure.

5 Role of Market Structure

In this section, we unpack the link between market structure and the size of the education multiplier to the private sector as well as how the impact varies by grant amount. In the first subsection, we use baseline variation in market structure to provide evidence that private schools improved because they responded to increased competition from the public sector. To do so, we note that if this is indeed the mechanism for private schools' improvement, we

[27] Basic infrastructure consists of per-student permanent and semi-permanent classrooms, toilets, and blackboards, as well as an indicator for whether students sit on chairs at desks rather than on the floor. Extra infrastructure consists of indicator variables for having a library, a computer, sports, a hall, a wall, fans, and electricity.

would expect the degree of improvement to depend on how much competitive pressure the program exerted on private schools, which itself will vary with a village's baseline market structure.

We exploit two sources of variation in market structure to identify the schools that likely faced more competitive pressure due to the policy. Motivated by the demand estimates presented in Bau (2022) and Carneiro et al. (2016), as a measure of horizontal differentiation, we examine differential responses by the physical distance of a private school to public schools. Intuitively, since students are highly distance sensitive, we expect increases in public school quality to have less effect on a private schools' market power – and therefore, to lead to smaller quality improvements – if the private school is very far from public schools. As a measure of vertical differentiation, we examine whether the intervention had larger effects on private schools in villages where pre-program public school quality was higher. Since private schools are higher quality than public schools on average (Andrabi et al., 2022b), we expect private schools to experience larger declines in their market power only when the public sector is sufficiently good to attract students who would otherwise enroll in the private sector.

In the second subsection, we turn to the idea that competitive pressures can also be generated through the size of the improvement in the public sector. We exploit the non-experimental variation in grant size across villages to examine whether larger grants were associated with greater test score improvements in the public sector and whether these in turn resulted in greater improvements in the private sector.

5.1 Role of Horizontal and Vertical Differentiation

Horizontal Differentiation: Heterogeneity by Distance. To allow private school-level treatment effects to vary with the distance to public schools, for the sample of private schools that were open in round 2, we estimate

$$y_{s,5} = \alpha_d + \beta_1 I_v^{Treatment} + \beta_2 D_{s,2} + \beta_3 I_v^{Treatment} \times D_{s,2} + \sum_{t=\{1,2\}} \mathbf{\Gamma_t X_{vt}} + \varepsilon_{s,5}, \qquad (2)$$

where s denotes a school, and $D_{s,2}$ is the average log distance between a private school s and all public schools in the village in round 2.[28] We focus on log distance since students

[28]Using GPS coordinates, for each private school, we calculate the distance to all open public schools within the village in round 2, replacing the value with 10 meters in the 1.2% of cases where the the distance is zero. We then take the log of this distance and average over all the log distances between a school s and

typically attend a school within 1 km of their households, and it is unlikely that their enrollment behavior would be affected by marginal differences in distance once a school is sufficiently far away. Then, β_3 identifies the differential treatment effect on private schools that are farther from public competitors.

Table 9 confirms that private schools located closer to public schools improve more. Because distances are typically lower than 1 and therefore, log distances are typically negative, the coefficients are difficult to interpret. In the bottom panel of Table 9, we use the coefficient estimates to calculate the predicted effect of the intervention for schools at the 10th, 50th, and 90th percentiles of the distribution of $D_{s,2}$, corresponding to 223, 466, and 1,057 meters, respectively. The intervention has a large and statistically significant treatment effect for schools at the 10th percentile (0.31-0.41 sd across specifications) but no effect on private schools at the 90th percentile of our distance metric, suggesting a radius of 1 km for the boundary of the competitive effects.

Vertical Differentiation: Heterogeneity by Ex-Ante Public School Quality. To estimate the effects of the intervention by baseline public school quality, we again estimate equation (2), replacing $D_{s,2}$ with average public school value-added (SVA) in a village. We report results for both the public and private sectors since treatment effects in the public sector may also depend on baseline value-added. For example, better managed public schools may have used the grants more effectively or better-resourced public schools may have benefited less from the additional resources.

Columns 1-3 in Table 10 show that treatment effects do not differ based on baseline SVA for public schools – public schools in villages with low and high-performing public schools were just as likely to benefit from the grant. In contrast, when private schools are located in villages with high levels of public SVA, their test scores increase more in response to the intervention. The effect sizes suggest that private schools located in villages with 1 sd higher public SVA (in the average public SVA distribution) increase their test scores by an *additional* 0.3 student test score sd. The bottom panel confirms this with the private school impact 0.46-0.50 in villages at the 90th percentile of the public SVA distribution compared to no significant effect for those at the 10th percentile.

Both the distance and baseline quality results are consistent with the idea that the response of the private school to a marginal improvement in public school quality is more aggressive when the public school is "closer" in product space to begin with, and the threat

all public schools in the same village.

Table 9: Effect on Private School Test Scores by Average Log Distance to Public Schools

	Private Schools		
	(1) OLS	(2) OLS	(3) Lasso
Treatment	0.009	0.001	-0.033
	(0.212)	(0.172)	(0.187)
Treatment × Avg Log Dist. Public Schools	-0.180	-0.241*	-0.249*
	(0.174)	(0.131)	(0.134)
Avg Log Dist. Public Schools	-0.030	0.029	0.060
	(0.141)	(0.091)	(0.089)
Effect at 90th perc. (1057m)	0.041	0.044	0.011
	(0.188)	(0.153)	(0.169)
Effect at 50th perc. (466m)	0.180	0.231**	0.204*
	(0.123)	(0.100)	(0.119)
Effect at 10th perc. (223m)	0.310*	0.406***	0.384***
	(0.170)	(0.127)	(0.143)
Control Mean	0.345	0.345	0.345
Baseline Controls	No	Yes	Yes
Adjusted R^2	0.215	0.354	0.364
Observations	134	134	134
Clusters	67	67	67

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table estimates heterogeneous effects by the average log distance from a private school to all public schools in the village using GPS data collected in round 1. The outcome variable is school-level average test scores (across tests in math, English, and Urdu and across all students in the school) in round 5 (2011-12). Tests are scored using item response theory (IRT), and test scores are measured in standard deviations. Column 1 controls only for district fixed effects (the stratifying variable), column 2 adds round 1 and round 2 (baseline) values of the dependent variable (if available), and column 3 uses a post double-lasso procedure to select additional baseline controls. Columns 2 and 3 also include a control for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level. The lower number of clusters in this table reflects the fact that exit and entry mean that some villages do not have any private schools where pre-treatment distance to public schools can be calculated. The bottom panel reports treatment effects at the 90th, 50th, and 10th percentiles of the average log distance distribution, calculated using the point estimates in the top panel.

to the private school's market power is therefore greater.

While our previous results showed that the public school grants did not increase exits among private schools exiting (Table 6), it could be that the aggregate impacts masked heterogeneity by baseline competition. If so, the measures of competition we have used

Table 10: Effect on Private School Test Scores by Village-Level Public School SVA

	Public Schools			Private Schools		
	(1) OLS	(2) OLS	(3) Lasso	(4) OLS	(5) OLS	(6) Lasso
Treatment	0.211**	0.219**	0.214**	0.190*	0.200**	0.192*
	(0.091)	(0.094)	(0.105)	(0.099)	(0.083)	(0.109)
Treatment × Avg. Village-Level Public School SVA	-0.026	-0.043	-0.049	0.281**	0.275***	0.320**
	(0.095)	(0.102)	(0.118)	(0.118)	(0.099)	(0.146)
Avg. Village-Level Public School SVA	0.069	0.044	0.084	0.005	-0.073	-0.104
	(0.073)	(0.119)	(0.138)	(0.097)	(0.076)	(0.110)
Effect at 90th perc.	0.185	0.177	0.165	0.456***	0.461***	0.496**
	(0.123)	(0.123)	(0.134)	(0.144)	(0.120)	(0.192)
Effect at 50th perc.	0.212**	0.220**	0.214**	0.186*	0.196**	0.187*
	(0.091)	(0.095)	(0.106)	(0.099)	(0.083)	(0.108)
Effect at 10th Perc.	0.248	0.280	0.283	-0.093	-0.076	-0.131
	(0.172)	(0.190)	(0.221)	(0.159)	(0.134)	(0.164)
Control Mean	-0.550	-0.550	-0.550	0.310	0.310	0.310
Baseline Controls	No	Yes	Yes	No	Yes	Yes
Adjusted R^2	0.301	0.294	0.286	0.279	0.334	0.369
N	231	231	231	193	193	193
Clusters	80	80	80	74	74	74

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table estimates heterogeneous effects by the average quality of public schools in the village, calculated using school value-added in mean test scores for each public school in round 2 and using empirical Bayes to correct for estimation error (Andrabi et al., 2022b). We normalize the village-level average school quality measure to have a mean of 0 and standard deviation of 1. The outcome variable is school-level average test scores (across tests in math, English, and Urdu and across all students in the school) in round 5 (2011-12). Tests are scored using item response theory (IRT), and test scores are measured in standard deviations. Public schools are included in Columns 1-3, and private schools in Columns 4-6. Each set of three columns follows the same format. The first column controls only for district fixed effects (the stratifying variable), the second column additionally controls for the baseline values of the dependent variable from rounds 1 and 2 (if available), and the third column uses a post double-lasso procedure to select additional baseline controls. The second and third columns also include a control for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level. The bottom panel reports treatment effects at the 90th, 50th, and 10th percentiles of the average quality of public schools in the village distribution, calculated using the point estimates in the top panel.

here present an ideal opportunity to reexamine the question of private school exits when competition is extremely intense. Appendix Table A13 therefore evaluates private school exits by baseline measures of physical distance (Columns 1-2) and average public school quality (Columns 3-4). In neither case do we find evidence that such schools faced greater exit as a result of the intervention.

5.2 Effects by Grant Size

We next turn to the effects of grant size on test scores in both sectors. Recall that the intervention was randomized at the village level, but the amount that each village or school received remained under the jurisdiction of the government. As we discussed previously in Section 4.1, there was considerable variation in grant size across schools and villages.

To assess the impact of this variation on test scores, we first use lasso regressions to test whether (and what) village and public school characteristics predict the level of funding a school received. Strikingly, Appendix B and Appendix Table A14 show that our lasso procedure does not select *any* variables in our experimental sample. Despite multiple years of pre-treatment variables, including test scores, we cannot predict the variation in grant amounts. This result is not simply due to lack of power; in a third district that explicitly stated they would not randomize the program, the lasso does select predictors of grant size. We are thus optimistic that estimating a relationship between village-level funding and test scores will not simply capture time-varying omitted variables.

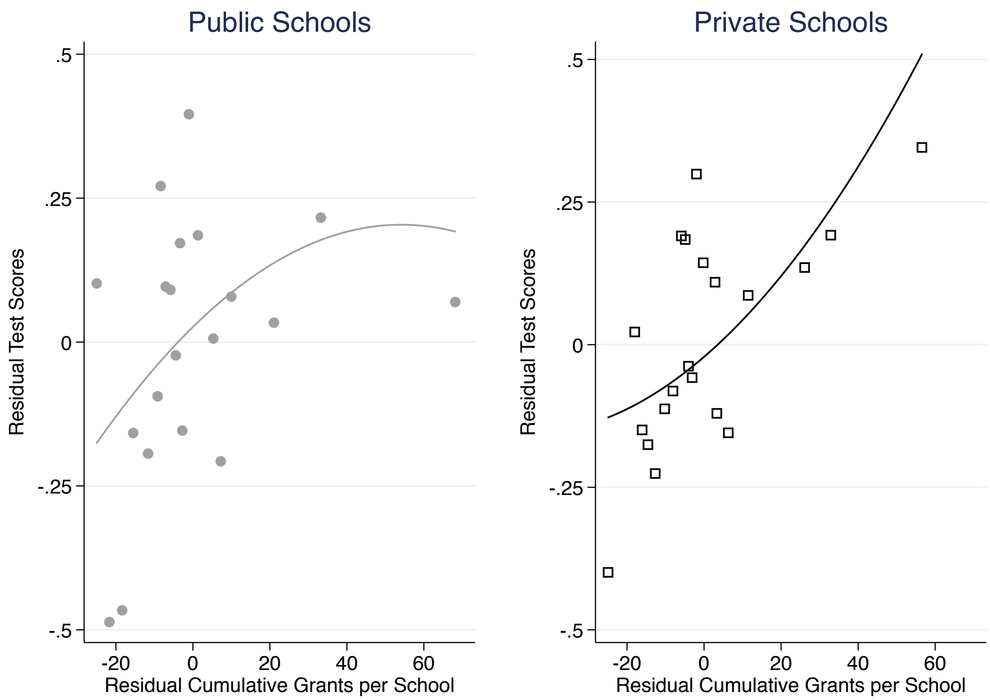
Having shown that neither village nor school characteristics predict grant amounts, we now examine the relationship between village-level funding and test scores. Figure 5 plots the relationship between average public school-level grants to a village and test scores in the public and private sectors.[29] Both the x- and y-axes variables are residualized by the baseline values (when available) and controls selected via the double-lasso procedure. Since we include controls for multiple pre-treatment test scores, our specification controls for level test score differences across villages, similar to a difference-in-differences, as well as the growth (trend) in test scores between the first and second baseline survey.

Figure 5 shows that larger grants to the village lead to greater test score improvements in public schools. For private schools, larger public grants lead to *greater* marginal improvements in test scores. Appendix Table A15 estimates a linear relationship between funding and test scores and confirms that larger grants lead to higher test scores in both public and private schools, with the marginal effect of an additional dollar in grant size twice as large in the private sector, consistent with Figure 5. Finally, the public sector improvements are consistent with diminishing returns in Figure 5 although we never have sufficient data to investigate this non-linear aspect of the relationship formally.

The positive relationship between grant size and test score improvements in the public sector is evidence that the effects of the program are driven, at least in part, by the grants

[29] We focus on the average grants to public schools in a village since there is no school-level grant measure for private schools.

Figure 5: Spending and Test Scores



Notes: The left figure plots school-level test scores in public schools against the average cumulative amount of funding received by public schools in the village. The right panel plots school-level test scores in private schools against the average amount of funding received by public schools in the village. In both panels, schools are divided into 20 bins, and both test scores and cumulative funding are residualized using controls selected by double-lasso from the pool listed in Appendix Table A1 and pre-treatment baseline values (when available).

themselves rather than simply changes in the school council. These results also show that larger grant sizes in the public sector are associated with a larger multiplier for the private sector.

6 Discussion: Cost-Effectiveness and Policy Design

We have demonstrated the existence of a public education spending multiplier due to competitive pressures on the private sector. Furthermore, we have shown that the multiplier's size is affected by market structure. We now show that accounting for this multiplier changes the cost-effectiveness and optimal targeting of the program we study and argue that accounting for market structure can improve the efficacy of school grant programs.

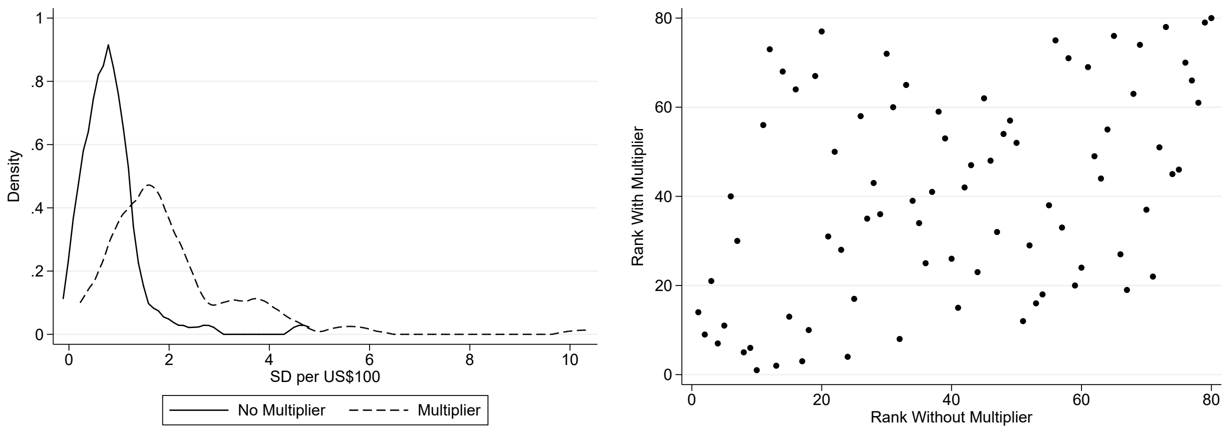
Cost-Effectiveness. Our data collection allows us to observe learning outcomes in both sectors. Thus, we can calculate the program's overall cost-effectiveness, taking into account the additional effects on the private sector, *and* quantify the bias from failing to account for them. Following Dhaliwal et al. (2013), we compute cost-effectiveness as the change in test scores per dollar spent on students ($CE = 100 \times \frac{\Delta Test\, Scores}{\Delta USD\, per\, student}$). Then, the change in test scores due to the intervention $\Delta Test\, Scores$ can be estimated as the treatment effect in a child-level regression of test scores on $I_v^{Treatment}$, while the change in dollars per student $\Delta USD\, per\, student$ is the coefficient on $I_v^{Treatment}$ in a regression where the total number of dollars spent per student is the outcome. To arrive at a cost-effectiveness measure that incorporates the additional effect on the private sector, we include children in both public and private schools in the calculation of $\Delta Test\, Scores$ (column 1 of Appendix Table A16) and calculate the dollars per student including students enrolled in both sectors (column 2 of Table A16). To arrive at a cost-effectiveness measure that ignores these effects, we only use public schools to estimate $\Delta Test\, Scores$ (column 3 of Table A16) and calculate the dollars per student enrolled in *public* school (column 4 of Table A16). Without accounting for private school responses, we estimate that test scores improved by 1.18 standard deviations per 100 USD. In contrast, accounting for the improvement in private schools increases the cost-effectiveness estimate by 85% to 2.18 standard deviations per 100 USD. At this higher cost-effectiveness, the program compares favorably to several highly regarded interventions such as girls' scholarships in Kenya (1.38), village-based schools in Afghanistan (2.13), and individually-paced computer assisted learning in India (1.55) (Kremer et al., 2013; Burde and Linden, 2013). Thus, failing to account for the multiplier effect severely underestimates the cost-effectiveness of the program in this context.[30]

Policy Design Implications. We now discuss the implications of our findings for how a policymaker should divide a grant budget among multiple villages once private sector

[30] This exercise assumes only children currently enrolled in the school are affected by the intervention and that they all experience the same benefits as the 4th graders we test. This may be an underestimate if the program effects are immediate and persistent, as we do not count positive effects on cohorts who have already graduated or will attend the school in the future. However, it may also be an overestimate to the extent that the effects are temporary and therefore effects on younger cohorts are smaller (though even in this case, we fail to account for potential positive effects on the cohorts who have graduated). While a full accounting of benefits that includes the test scores of all the students who were potentially affected by the intervention (including those who have left the school as well as future cohorts) would likely alter these estimates, the fact that accounting for private sector responses will result in meaningful and large changes to these estimates will still hold. Finally, the programs that we use for comparison do not account for private sector responses. If private schools are available, the cost-effectiveness of the other programs may be under- or over-estimated depending on their responses.

responses are taken into account.

Figure 6: Density and Ranking of Village-Level Cost-Effectiveness With and Without the Multiplier



(a) Density (b) Ranking

Notes: Panel (a) shows the density of the village-level cost-effectiveness estimates without (solid line) and with (dashed line) the inclusion of spillovers. Panel (b) plots the cost-effectiveness rank of each village without accounting for spillovers against the rank of each village with spillovers. Each dot represents a village. A higher rank indicates a village is more cost-effective.

First, consider a government that is only concerned about cost-effectiveness. Figure 6 shows the kernel densities of village-level cost-effectiveness measures with and without the private sector response (left panel) and how the ranking of villages' cost-effectiveness changes when we take into account these effects (right panel).[31] The left panel illustrates the large increase in cost-effectiveness measures from accounting for the private sector impact, while the right panel shows that there is little relationship between villages' cost-effectiveness ranks with and without these effects. Of the villages with the 10 highest cost-effectiveness measures using only public sector improvements, 6 are different once we take private sector responses into account. A government concerned with cost-effectiveness will target the program poorly if it fails to account for private sector responses.

[31]To compute village-level cost-effectiveness, we focus on the variation in the share of children enrolled in private schools. We denote s_{pr} as the share of students in private schools, s_{pu} the share in public schools, n_{pr} the number of children enrolled in private, n_{pu} the number enrolled in public, β_{pr} the treatment effect on the private schools' mean test scores, β_{pu} the treatment effect on public school test scores, and G the amount the program increased funding to the village's public sector estimated from the regression. Accounting for the private school impact, village-level cost-effectiveness is given by $\frac{s_{pu}\beta_{pu}+s_{pr}\beta_{pr}}{G/(n_{pu}+n_{pr})}$. Without spillovers, the village-level cost-effectiveness is $\frac{s_{pu}\beta_{pu}}{G/n_{pu}}$.

Next, suppose that the government is also concerned about equity in the sense that it wishes to equalize the gains from the grant program. If the government uses only public sector responses as the metric of improvement, either linear or diminishing returns to grant size combined with the fact that the gains in test scores do not vary by baseline levels of performance in public schools implies that an equal division of the grants among all villages will maximize test score gains. There is apparently no trade-off between equity and efficiency in terms of test score gains from the program.

That inference is incorrect once private school responses are factored in. Private schools improve more in villages where public school quality is higher. Now there is an equity-efficiency trade-off. Maximizing overall gains in test scores would require the government to give more money to villages where public schools are better off to begin with. However, such targeting or even an equal distribution of the grant will lead to unequal gains and an increase in test score inequality. Instead, a policy maker who wishes to equalize test score gains must distribute more cash to villages where public schools are poorly-performing. The variation in cash need not be extremely unequal as steeper responses to grant size in the private sector will lead to rapid increases in test scores. Regardless of the exact preferences of the policy maker in terms of equity and test score gains, policy will have to be very intentional in how to target grants across villages with considerable variation in the quality of public schools.

The specific calculations we have presented may change across contexts as they are sensitive to both the demand for quality and the cost of quality improvement. The general insight that remains is that accounting for the private sector multiplier alters the cost-effectiveness and targeting of government programs in fundamental ways. Understanding how this multiplier relates to underlying market structure and the size of the grant is key to effectively crafting policies that maximize the objective function of the policymaker.

7 Conclusion

Our market-level randomized evaluation produces three important results. First, grants to public schools–when administered through school councils–increase test scores in public schools. Second, since public schools exert competitive pressure on private schools, they also increase test scores in private schools through an education multiplier that is quantitatively and qualitatively significant. Third, the size of the education multiplier depends (in ways that are ex ante knowable) on the underlying nature of market competition.

In response to a government intervention, private schools will respond to protect their market power to the point that their profits are driven to zero and at this point, they will choose to exit the market. The same policy can therefore lead to crowd-in, crowd-out through exits or no impact on private schools at all depending on their cost structures and on parental preferences. In this regard it is interesting–and important–that even the private schools who were in very close competition with public schools chose to increase quality rather than exit the market. It seems that private schools enjoy sufficient market power in this context that public schools can increase quality substantially without forcing exits.

The broader point this paper makes is that we should be more intentional in how we design and evaluate public interventions and leverage policy tools in mixed systems, where parents can choose from both public and private options. Given the pervasiveness of such mixed systems in lower-income countries, optimal policy design must account for the private multiplier and how it varies with market structure. Perhaps counter-intuitively, a key insight from our work is that governments may still wish to invest in the public sector, even if weaker incentives limit the direct impact, precisely because it allows them to leverage the competitive pressures that private schools face. Designed appropriately, the presence of the private sector can make public interventions more effective and equitable, not less.

References

Andrabi, Tahir Das, Jishnu Khwaja, and Tristan Asim Ijaz Zajonc, "Madrassa metrics: the statistics and rhetoric of religious enrollment in Pakistan," 2005.

Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja, "A Dime a Day: The Possibilities and Limits of Private Schooling in Pakistan," *Comparative Education Review*, 2008, *52* (3), 329–355.

_ , _ , and _ , "What did you do all day? Maternal education and child outcomes," *Journal of Human Resources*, 2012, *47* (4), 873–912.

_ , _ , and _ , "Delivering education: a pragmatic framework for improving education in low-income countries," in "Handbook of International Development and Education," Edward Elgar Publishing, 2015.

_ , _ , and _ , "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets," *American Economic Review*, 2017, *107* (6), 1535–1563.

_ , _ , and Asim Khwaja, "The Learning and Educational Achievement in Pakistan Schools (LEAPS) Longitudinal Dataset, 2004-2011," 2022. Version: 11-2022.

_ , _ , Asim I Khwaja, Selcuk Ozyurt, and Niharika Singh, "Upping the ante: The equilibrium effects of unconditional grants to private schools," *American Economic Review*, 2020, *110* (10), 3315–49.

_ , _ , Asim Ijaz Khwaja, Tara Vishwanath, and Tristan Zajonc, "Learning and Educational Achievements in Punjab Schools (LEAPS): Insights to Inform the Education Policy Debate," *World Bank, Washington, DC*, 2007.

_ , Natalie Bau, Jishnu Das, and Asim Ijaz Khwaja, "The Equilibrium Effects of Unconditional Grants to Public Schools," *Working Paper*, 2020.

_ , _ , _ , and _ , "Heterogeneity in School-Value Added and the Private Premium," *Working Paper*, 2022.

Bandiera, Oriana, Andrea Prat, and Tommaso Valletti, "Active and passive waste in government spending: evidence from a policy experiment," *American Economic Review*, 2009, *99* (4), 1278–1308.

Bau, Natalie, "Estimating an equilibrium model of horizontal competition in education," *Journal of Political Economy*, 2022, *130* (7), 1717–1764.

_ **and Jishnu Das**, "Teacher Value-Added in a Low-Income Country," *American Economic Journal: Economic Policy*, 2020.

Baum, Donald, Laura Lewis, Oni Lusk-Stover, and Harry Patrinos, "What matters most for engaging the private sector in education," 2014.

Bazzi, Samuel, Masyhur Hilmy, and Benjamin Marx, "Islam and the state: Religious education in the age of mass schooling," Technical Report, National Bureau of Economic Research 2020.

Bedoya, Gaudalupe, Jishnu Das, and Amy Dolinger, "Randomized Regulation: The impact of minimum quality standards on health markets," *NBER Working Paper*, 2022.

Borman, Geoffrey D, Gina M Hewes, Laura T Overman, and Shelly Brown, "Comprehensive school reform and achievement: A meta-analysis," *Review of educational research*, 2003, *73* (2), 125–230.

Bulow, Jeremy I, John D Geanakoplos, and Paul D Klemperer, "Multimarket oligopoly: Strategic substitutes and complements," *Journal of Political economy*, 1985, *93* (3), 488–511.

Burde, Dana and Leigh L Linden, "Bringing education to Afghan girls: A randomized controlled trial of village-based schools," *American Economic Journal: Applied Economics*, 2013, *5* (3), 27–40.

Card, David and A Abigail Payne, "School finance reform, the distribution of school spending, and the distribution of student test scores," *Journal of public economics*, 2002, *83* (1), 49–82.

Carneiro, Pedro Manuel, Jishnu Das, and Hugo Reis, "The value of private schools: Evidence from Pakistan," 2016.

Carneiro, Pedro, Oswald Koussihouèdé, Nathalie Lahire, Costas Meghir, and Corina Mommaerts, "School Grants and Education Quality: Experimental Evidence from Senegal," *Economica*, 2020, *87* (345), 28–51.

Clingingsmith, David, Asim Ijaz Khwaja, and Michael Kremer, "Estimating the impact of the Hajj: religion and tolerance in Islam's global gathering," *The Quarterly Journal of Economics*, 2009, *124* (3), 1133–1170.

Cutler, David M and Jonathan Gruber, "Does public insurance crowd out private insurance?," *The Quarterly Journal of Economics*, 1996, *111* (2), 391–430.

Das, Jishnu, Abhijeet Singh, and Andres Yi Chang, "Test scores and educational opportunities: Panel evidence from five low-and middle-income countries," *Journal of Public Economics*, 2022, *206*, 104570.

_ **and Tristan Zajonc**, "India Shining and Bharat Drowning: Comparing Two Indian States to the Worldwide Distribution in Mathematics Achievement," *Journal of Development Economics*, 2010, *92* (2), 175–187.

_ **, Benjamin Daniels, Monisha Ashok, Eun-Young Shim, and Karthik Muralidharan**, "Two Indias: The structure of primary health care markets in rural Indian villages with implications for policy," *Social Science & Medicine*, 2020, p. 112799.

_ **, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman**, "School inputs, household substitution, and test scores," *American Economic Journal: Applied Economics*, 2013, *5* (2), 29–57.

Dhaliwal, Iqbal, Esther Duflo, Rachel Glennerster, and Caitlin Tulloch, "Comparative cost-effectiveness analysis to inform policy in developing countries: a general framework with applications for education," *Education policy in developing countries*, 2013, *17*, 285–338.

Dinerstein, Michael, Troy Smith et al., "Quantifying the supply response of private schools to public policies," 2015.

Duflo, Esther, Pascaline Dupas, and Michael Kremer, "School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from Kenyan primary schools," *Journal of public Economics*, 2015, *123*, 92–110.

Estevan, Fernanda, "Public education expenditures and private school enrollment," *Canadian Journal of Economics/Revue canadienne d'économique*, 2015, *48* (2), 561–584.

Evans, David K and Fei Yuan, "How Big Are Effect Sizes in International Education Studies?," *Center for Global Development, Working Paper*, 2020, *545.*

Gertler, Paul J, Harry Anthony Patrinos, and Marta Rubio-Codina, "Empowering parents to improve education: Evidence from rural Mexico," *Journal of Development Economics*, 2012, *99* (1), 68–79.

Glennerster, Rachel and Kudzai Takavarasha, *Running randomized evaluations: A practical guide*, Princeton University Press, 2013.

Government of Punjab, "Punjab Education Sector Reform Program," 2007.

Guryan, Jonathan, "Does money matter? Regression-discontinuity estimates from education finance reform in Massachusetts," 2001.

Hoxby, Caroline M, "Peer effects in the classroom: Learning from gender and race variation," *NBER Working Paper*, 2000.

Hyman, Joshua, "Does money matter in the long run? Effects of school spending on educational attainment," *American Economic Journal: Economic Policy*, 2017, *9* (4), 256–80.

Jackson, C Kirabo, *Does school spending matter? The new literature on an old question.*, American Psychological Association, 2020.

_ , Rucker C Johnson, and Claudia Persico, "The effects of school spending on educational and economic outcomes: Evidence from school finance reforms," *The Quarterly Journal of Economics*, 2016, *131* (1), 157–218.

Karachiwalla, Naureen, "A teacher unlike me: Social distance, learning, and intergenerational mobility in developing countries," *Economic Development and Cultural Change*, 2019, *67* (2), 225–271.

Khan, Faryal, "School Management Councils: a lever for mobilizing social capital in rural Punjab, Pakistan?," *Prospects*, 2007, *37* (1), 57–79.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental analysis of neighborhood effects," *Econometrica*, 2007, *75* (1), 83–119.

Kremer, Michael and Karthik Muralidharan, "Public and private schools in rural India," *School Choice International*, 2008, pp. 91–110.

— , **Conner Brannen, and Rachel Glennerster**, "The Challenge of Education and Learning in the Developing World," *Science*, 2013, *340* (6130), 297–300.

Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach, "School finance reform and the distribution of student achievement," *American Economic Journal: Applied Economics*, 2018, *10* (2), 1–26.

Leclerc, Catherine Michaud, "Private School Entry, Sorting, and Performance of Public Schools: Evidence from Pakistan," 2020.

Mbiti, Isaac, Karthik Muralidharan, Mauricio Romero, Youdi Schipper, Constantine Manda, and Rakesh Rajani, "Inputs, incentives, and complementarities in education: Experimental evidence from Tanzania," *The Quarterly Journal of Economics*, 2019, *134* (3), 1627–1673.

Mohmand, Shandana and Harris Gazdar, "Social Structures in Rural Pakistan," *Determinants and Drivers of Poverty Reduction and ADB's Contribution in Rural Pakistan, Asian Development Bank, Islamabad*, 2007.

Muralidharan, Karthik and Venkatesh Sundararaman, "Contract teachers: Experimental evidence from India," 2013.

— **and** — , "The aggregate effect of school choice: Evidence from a two-stage experiment in India," *The Quarterly Journal of Economics*, 2015, *130* (3), 1011–1066.

National Center for Education Statistics, "Digest of Education Statistics," https://nces.ed.gov/programs/digest/d19/tables/dt19_105.50.asp 2019. Accessed: 2022-11-1.

Neilson, Christopher A and Seth D Zimmerman, "The effect of school construction on test scores, school enrollment, and home prices," *Journal of Public Economics*, 2014, *120*, 18–31.

Neilson, Christopher, Michael Dinerstein, and Sebastián Otero, "The Equilibrium Effects of Public Provision in Education Markets: Evidence from a Public School Expansion Policy," 2020.

Pradhan, Menno, Daniel Suryadarma, Amanda Beatty, Maisy Wong, Arya Gaduh, Armida Alisjahbana, and Rima Prama Artha, "Improving educational quality through enhancing community participation: Results from a randomized field experiment in Indonesia," *American Economic Journal: Applied Economics*, 2014, *6* (2), 105–26.

Urminsky, Oleg, Christian Hansen, and Victor Chernozhukov, "Using double-lasso regression for principled variable selection," *Working Paper*, 2016.

Appendix A: Estimating School Value-Added (SVA)

The procedure for estimating school value-added closely follows Andrabi et al. (2022b). We first obtain a school fixed effect estimate of value-added for mean test scores from the regression

$$y_{igst} = \lambda_g y_{igs,t-1} + \theta_s + \alpha_g + \alpha_t + \epsilon_{igst}, \tag{3}$$

where i denotes a student, g a grade, s a school, and t a year. The outcome variable y_{igst} is student i's mean test score in year t, λ_g is a grade-specific coefficient that captures the effect of lagged performance, and α_g and α_t are grade and year fixed effects. Our estimate of a school's value-added is the estimate of θ_s, the school fixed effect. Because this specification controls for lagged test scores, value-added estimates require test score data from both round 1 and round 2.

Since estimation error will lead to attenuation bias in the coefficient on value-added in a regression with the fixed effect estimate of school value-added on right-hand side, we apply empirical Bayes shrinkage to the fixed effect estimates. To solve for the empirical Bayes estimates, we assume that mean test scores are given by

$$y_{ijst} = \beta X_{ijt} + \phi_s + \phi_j + \phi_{jt} + \mu_{ijt}, \tag{4}$$

where y_{ijst} is the test score, X_{ijt} is the set of controls, ϕ_s is a school effect (not including the teacher shock), ϕ_j is a teacher effect, ϕ_{jt} is a classroom effect, and μ_{ijt} is an idiosyncratic student-specific shock. The variances of these shocks are σ_S^2, σ_T^2, σ_C^2, and σ_ϵ^2 respectively, and they are assumed to be independent and homoskedastic.

The object of interest that our fixed effect estimates is the expected test score gains a child will experience in a school due to the combination of the school and teacher effect:

$$\theta_j = \phi_s + \sum_{j \in s} \frac{N_j}{N_s} \phi_j, \tag{5}$$

where N_j is the number of students taught by teacher j, and N_s is the number of students in school s. Note that this is just the independent school effect plus the weighted average of the teacher effects of the teachers who teach in a school. To calculate $Var(\theta_s)$, use the fact that $Var(\theta_s) = E(\theta_s^2) - E(\theta_s)^2$. Noting that $E(\theta_s)^2 = 0$ by construction, the variance of θ_s

is

$$Var(\theta_s) = E((\phi_s + \sum_{j \in s} \frac{N_j}{N_s} \phi_j)^2).$$

Recognizing that ϕ_j and ϕ_s are independent by assumption, this can be further simplified to

$$\begin{aligned} Var(\theta_s) =& E(\phi_s^2) + E(\sum_{j \in s} \sum_{j' \in s} \frac{N_j N_{j'}}{N_s^2} \phi_j \phi_{j'}). \\ =& \sigma_S^2 + E(\frac{\sum_j N_j^2}{N_s^2} \sigma_T^2). \end{aligned} \tag{6}$$

Our estimate of θ_s (the school fixed effect) is given by

$$\hat{\theta}_s = \theta_s + \frac{1}{N_s} \sum_{ijt \in s} \left(\phi_j + \phi_{jt} + \epsilon_{ijt}\right) \tag{7}$$

Then, the variance of $\hat{\theta}_s$ is

$$\begin{aligned} Var(\hat{\theta}_s) &= E\Big((\phi_s + \frac{1}{N_s} \sum_{ijt \in s} \left(\phi_j + \phi_{jt} + \mu_{ijt}\right))^2\Big) \\ &= \sigma_S^2 + E\Big(\sum_{j \in s} \sum_{k \in s} \frac{N_j N_k}{N_s^2} \phi_j \phi_k + \sum_{jt \in s} \sum_{kl \in s} \frac{N_{jt} N_{kl}}{N_s^2} \phi_{jt} \phi_{kl} + \sum_{ijt \in s} \sum_{i'j't' \in s} \frac{1}{N_s^2} \mu_{ijt} \mu_{i'j't'}\Big) \\ &= \sigma_S^2 + E\Big(\sum_j \frac{N_j^2}{N_s^2} \sigma_T^2 + \sum_{jt} \frac{N_C^2}{N_s^2} \sigma_C^2 + \frac{1}{N_s} \sigma_\epsilon^2\Big), \end{aligned} \tag{8}$$

Therefore, the variance of the school effects uncontaminated by estimation error is

$$Var(\theta_s) = Var(\hat{\theta}_s) - E\Big(\frac{\sum_{jt} N_{jt}^2}{N_s^2} \sigma_C^2 + \frac{1}{N_s} \sigma_\epsilon^2\Big). \tag{9}$$

For empirical Bayes, we should then scale $\hat{\theta}_s$ by

$$h_s = \frac{\sigma_S^2 + \frac{\sum_j N_j^2}{N_s^2} \sigma_T^2}{\sigma_S^2 + \sum_j \frac{N_j^2}{N_s^2} \sigma_T^2 + \sum_C \frac{N_{jt}^2}{N_s^2} \sigma_C^2 + \frac{1}{N_s} \sigma_\epsilon^2} \tag{10}$$

Note that σ_s^2, σ_{jt}^2, σ_j^2 and σ_ϵ^2 are all calculated in Bau and Das (2020) separately for private and public schools in the same data, so we can substitute these values into equation (9) to get the variances of school quality in the public and private sectors and into (10) to get the scaling value for calculating the empirical Bayes estimates of SVA.

Appendix B: Relationship between Grant Size and School/Village Characteristics

We use lasso regressions to identify the pre-treatment school and village characteristics from rounds 1 and 2 that are most predictive of the cumulative amount of funding received by a public school. The pool of school-level characteristics consists of test scores, parental education, household assets, total enrollment, primary enrollment, the share of low caste students, school facilities, and the student-teacher ratio. At the village-level, we use the same variables averaged over schools in the village (except enrollment, which is village-level total enrollment to capture village size), the number of public schools, the number of private schools, and the share of children enrolled in school.

Columns 1 and 2 of Appendix Table A14 report the lasso regression results for our experimental sample (Attock and Faisalabad) with and without controlling for village fixed effects. Strikingly, in both cases, the lasso procedure does not select *any* variables and the R-squared of the regression is zero. Despite multiple years of pre-treatment variables, we are not able to predict the variation in grant amounts.

One concern is that we may not observe key predictors or may be underpowered to identify predictors. In columns 3-5, we therefore report results for Rahim Yar Khan, a third district in the original LEAPS surveys, which chose not to randomize treatment and was thus excluded from our study.[32] The results indicate that, in Rahim Yar Khan, funding was directed to schools with more infrastructure (across villages) and more students (within villages). The results from Rahim Yar Khan indicate that when funding is not randomized, the lasso *does* have sufficient power to identify predictors of funding. Given these results we cautiously infer that selection may not play a major role in the allocation of funding across schools. Thus, the observed relationship between funding amounts and learning may provide us with information about the true, underlying relationship between funding and learning.

[32] We report more columns for Rahim Yar Khan since it is redundant to show a column for village predictors only for the experimental sample given that both village and school predictors have no predictive power in column 1.

Appendix Tables

Table A1: Variable Pool for Selection of Double Lasso Controls

Survey	Variable name	Round 1	Round 2	Village	Public	Private	Variable definition
Teacher Survey	Average teacher test scores	Yes	Yes	Yes	Yes	Yes	Average of teacher test scores in English, Urdu, Mathematics
	Teacher education level	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if teacher has a BA or higher level of education
	Teacher training	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if teacher has some formal teacher training
	Teacher absenteeism	Yes	Yes	Yes	Yes	Yes	Number of days the teacher has been absent in the past month (self-reported)
	Female teacher	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if teacher is female
	Experienced teacher	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if the teacher has more than 3 years of experience in teaching at any school
	Experienced teacher in this school	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if the teacher has more than 3 years of experience in teaching at this school
	Other source of income	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if the teacher has any source of income outside of the school
	Permanent contract (teachers of tested students)	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if the teacher has a permanent contract (only teachers of the students that were tested)
	Permanent contract (all teachers)	No	Yes	Yes	Yes	Yes	Indicator = 1 if the teacher has a permanent contract (all teachers in the school)
	Monthly salary	Yes	Yes	Yes	Yes	Yes	Monthly salary (in Rs)
	Log monthly salary	Yes	Yes	Yes	Yes	Yes	Log of monthly salary (in Rs)
	Teacher is from same village	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if the teacher is originally from the same village
	Teacher provides private tutoring	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if the teacher provides private tutoring outside school
Headteacher/ Owner Survey	Experience at this school	Yes	Yes	Yes	Yes	Yes	Number of years the headteacher/owner has been at their position at that school
	Experience teaching anywhere	Yes	Yes	Yes	Yes	Yes	Number of years the headteacher/owner has the in teaching sector at any school
	Currently teaches a class	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if the headteacher/owner is teaching any classes at present
	Female headteacher/owner	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if the headteacher/owner is female
School-Based Surveys	Log tuition fee	Yes	Yes	No	No	Yes	Log annual tuition fee in primary private schools (excluding admission fees)
	Log total fee	Yes	Yes	No	No	Yes	Log annual total fee (tuition and admission fees) in primary private schools
	School facilities index - basic facilities	Yes	Yes	Yes	Yes	Yes	Index of basic school facilities constructed using principal components analysis. Variables included: number of permanent classrooms per student, the number of semi-permanent classrooms per student, the number of toilets per student, the number of blackboards per student, and an indicator variable equal to one if students sit at desks and chairs (as opposed to on the floor or outside).
	School facilities index - additional facilities	Yes	Yes	Yes	Yes	Yes	Index of other facilities at the school, constructed using principal components analysis. Variables included: indicator variables = 1 if the school has a library, a computer, a sports area, a meeting hall, a boundary wall, any fans, and electricity.
	School age	Yes	Yes	Yes	Yes	Yes	Number of years since school was constructed
	Primary Enrollment	Yes	Yes	Yes	Yes	Yes	Number of students enrolled in grades 1 to 5

Survey	Variable name	Round 1	Round 2	Village	Public	Private	Variable definition
	Female Primary Enrollment	Yes	Yes	Yes	Yes	Yes	Number of female students enrolled in grades 1 to 5
	Total Enrollment	Yes	Yes	Yes	Yes	Yes	Number of students enrolled in grades 1 to 12
	Female Total Enrollment	Yes	Yes	Yes	Yes	Yes	Number of female students enrolled in grades 1 to 12
	Share of female students	Yes	Yes	Yes	Yes	Yes	Share of female students enrolled in grades 1 to 12
	Village Primary Enrollment	Yes	Yes	Yes	Yes	Yes	Total village-level primary enrollment (grades 1 to 5)
	Inspector has not visited in the past 6 months	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if the teacher reported that the last time an inspector visited the school was more than 6 months ago
	Number of different caste groups in the school	Yes	Yes	Yes	Yes	Yes	Number of different castes groups among students enrolled in the school
	Parents receive information	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if the school provides regular information to parents about the student
	Medium of instruction	Yes	Yes	Yes	Yes	Yes	Indicators: medium of instruction is English, Urdu, English and Urdu, Urdu and Punjabi, or other
	Teachers can get bonuses	Yes	Yes	Yes	Yes	Yes	Indicator =1 if teachers can receive bonuses or prizes in addition to their salary
	Receive funding from donors or charity	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if school received any funding from donors or charity
	Number of primary teachers	Yes	Yes	Yes	Yes	Yes	Number of primary level teachers
	Number of primary female teachers	Yes	Yes	Yes	Yes	Yes	Number of female teachers teaching at the primary level
	Log number of primary teachers	Yes	Yes	Yes	Yes	Yes	Log number of primary level teachers
	Log number of primary female teachers	Yes	Yes	Yes	Yes	Yes	Log number of primary level female teachers
	Student teacher ratio	Yes	Yes	Yes	Yes	Yes	Student teacher ratio at the primary level
	Log student teacher ratio	Yes	Yes	Yes	Yes	Yes	Log student teacher ratio at the primary level
	Number of schools	Yes	Yes	Yes	No	No	Number of schools in the village
	Number of public schools	Yes	Yes	Yes	No	No	Number of public schools in the village
	Number of private schools	Yes	Yes	Yes	No	No	Number of private schools in the village
Test Score Data/Child Survey	Average test scores	Yes	Yes	Yes	Yes	Yes	Average of test scores (math, English, and Urdu)
	English test score	Yes	Yes	Yes	Yes	Yes	English test score
	Urdu test score	Yes	Yes	Yes	Yes	Yes	Urdu test score
	Math test score	Yes	Yes	Yes	Yes	Yes	Math test scores
	Asset index	Yes	Yes	Yes	Yes	Yes	Index of household assets using principal components analysis. Variables: Whether the household has beds, radio, television, refrigerator, bicycle, plough, small agricultural tools, chairs, fans, tractor, cattle (horse, buffalo, cow), goats, chicken, watches, motor/rickshaw, car/taxi/van/pickup, telephone, tubewell.
	Mother lives in the household	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if mother lives in the household
	Father lives in the household	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if father lives in the household
	Mother has some education	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if mother has any formal education
	Father has some education	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if father has any formal education
	Mother has primary education	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if mother completed primary education
	Father has primary education	Yes	Yes	Yes	Yes	Yes	Indicator = 1 if father completed primary education

Survey	Variable name	Round 1	Round 2	Village	Public	Private	Variable definition
Household Survey	Asset index	Yes	Yes	Yes	No	No	Index of household assets using principal components analysis. Variables: Whether the household has beds, tables, chairs, fans, sewing machine, air cooler, air conditioner, refrigerator, radio/cassette recorder/CD player, television, VCR/VCD, watches, guns, plough, harvester, tractor, tubewell, other agricultural machinery, other agricultural hand-tools, motorcycle/scooter, car/taxi/vehicle, bicycle, cattle, goats, chicken.
	Mother lives in the household	Yes	No	Yes	No	No	Indicator = 1 if mother lives in the household
	Father lives in the household	Yes	No	Yes	No	No	Indicator = 1 if father lives in the household
	Mother has some education	Yes	No	Yes	No	No	Indicator = 1 if mother has any formal education
	Father has some education	Yes	No	Yes	No	No	Indicator = 1 if father has any formal education
	Mother has primary education	Yes	No	Yes	No	No	Indicator = 1 if mother completed primary education
	Father has primary education	Yes	No	Yes	No	No	Indicator = 1 if father completed primary education
	Household size	Yes	Yes	Yes	No	No	Number of household members
	Household owns land	Yes	No	Yes	No	No	Indicator = 1 if household owns any land
	Household has printed media	Yes	No	Yes	No	No	Indicator = 1 if the household has any printed media
	Student does not walk to school	Yes	No	Yes	No	No	Indicator = 1 if the student does not walk to school
	Student has help with homework	Yes	Yes	Yes	No	No	Indicator = 1 if the student can get help with their homework

Notes: This table lists potential baseline control variables used in post-double lasso regression models from the teacher survey, the headteacher or school owner survey, the school-based surveys, the test score data and child survey, and the household survey. All variables are constructed at the village-level combining both sectors and separately for the public and private sectors in that village (except in the case of school fees, which are only relevant for the private sector, and the number of schools, which pertains to the entire village). Variables from the teacher survey are first averaged across teachers within a school and then averaged across schools in the village. Each average is calculated separately for round 1 and for round 2.

Table A2: Test for Balance on Covariates Across Villages in Round 2

	(1) Control Mean	(2) Control SD	(3) Treatment Mean	(3) Treatment SD	(4) Mean Difference	(5) Difference P-Value
Share Low Caste	0.254	0.051	0.222	0.034	-0.031	0.472
Share of Female Enrolled	0.746	0.024	0.751	0.025	0.005	0.856
Test Scores	-0.455	0.075	-0.571	0.077	-0.116	0.160
Share Mothers with Some Education	0.276	0.030	0.285	0.023	0.009	0.758
Share Fathers with Some Education	0.608	0.028	0.635	0.025	0.027	0.291
Asset Index	0.088	0.089	0.144	0.071	0.056	0.495
Primary Enrollment in Public	137.324	9.722	133.259	13.459	-4.065	0.720
Primary Enrollment in Private	77.255	8.855	70.423	7.193	-6.832	0.420
Share of Enrollment in Private	0.259	0.030	0.264	0.027	0.005	0.880
Private School Annual Fees (PKR)	1,418.431	112.965	1,506.018	111.384	87.587	0.380
Private School Annual Fees (USD)	24.456	1.948	25.966	1.920	1.510	0.380
Basic Facility Index	0.395	0.166	0.391	0.121	-0.004	0.978
Extra Facility Index	0.328	0.172	0.225	0.167	-0.103	0.548
Teachers with BA plus	0.385	0.025	0.359	0.025	-0.025	0.400
Number of Public Schools	3.440	0.230	3.776	0.285	0.335	0.217
Number of Private Schools	2.509	0.349	2.967	0.551	0.458	0.326

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table tests balance between treated and untreated villages in round 2 (2004-05). Columns 1 and 2 report the mean and standard deviation of the variable in control villages, and Columns 3 and 4 report the mean and standard deviation in treated villages. Column 5 reports the difference between the treatment and control, and Column 6 provides the p-value of this difference.

Table A3: Test for Balance on Covariates Across Schools in Round 2

	(1) Control Mean	(2) Control SD	(3) Treatment Mean	(4) Treatment SD	(5) Mean Difference	(6) Difference P-Value
Panel A: Public Schools						
Share Low Caste	0.149	0.054	0.180	0.029	0.031	0.532
Test Scores	-0.676	0.065	-0.809	0.062	-0.133	0.088
Share Mothers with Some Education	0.229	0.029	0.242	0.020	0.012	0.654
Share Fathers with Some Education	0.583	0.025	0.594	0.025	0.012	0.656
Asset Index	-0.245	0.088	-0.209	0.075	0.035	0.707
Primary Enrollment	138.133	11.140	135.112	18.077	-3.020	0.839
Basic Facility Index	-0.213	0.180	-0.177	0.164	0.036	0.839
Extra Facility Index	-0.312	0.173	-0.591	0.162	-0.280	0.163
Teachers with BA plus	0.420	0.042	0.359	0.027	-0.060	0.167
Panel B: Private Schools						
Share Low Caste	0.225	0.056	0.188	0.032	-0.037	0.443
Test Scores	0.195	0.090	0.202	0.054	0.007	0.929
Share Mothers with Some Education	0.432	0.048	0.477	0.031	0.045	0.351
Share Fathers with Some Education	0.738	0.032	0.790	0.032	0.052	0.172
Asset Index	0.577	0.153	0.796	0.115	0.220	0.157
Primary Enrollment	79.084	7.608	78.587	6.188	-0.497	0.951
School Annual Fees (PKR)	1,518.331	102.871	1,660.663	122.867	142.332	0.160
School Annual Fees (USD)	26.178	1.774	28.632	2.118	2.454	0.160
Basic Facility Index	1.308	0.206	1.377	0.159	0.069	0.742
Extra Facility Index	1.562	0.167	1.662	0.108	0.100	0.573
Teachers with BA plus	0.157	0.023	0.186	0.023	0.029	0.259

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table examines balance for public (Panel A) and private (Panel B) schools across treated and untreated villages in round 2 (2004-05). Columns 1 and 2 report the mean and standard deviation of the variable in control villages, and columns 3 and 4 report the mean and standard deviation in treated villages. Column 5 reports the difference between the treated and control groups, and column 6 provides the p-value of this difference.

Table A4: Effect of Treatment on Household Migration and Attrition

	Household Moved		Survey Not Completed	
	(1)	(2)	(3)	(4)
	Round 1	Round 2	Round 1	Round 2
Treatment	-0.008	-0.009	-0.026	-0.025
	(0.021)	(0.019)	(0.020)	(0.019)
Control Mean	0.103	0.093	0.142	0.126
Adjusted R^2	0.004	0.001	0.013	0.002
Observations	1295	1269	1295	1269
Clusters	80	80	80	80

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table tests whether the treatment affected whether a household either migrated out of the village by or did not complete the survey in round 5 (2011). The data come from the household survey. Columns 1 and 3 and columns 2 and 4 include the sample of households interviewed in round 1 and round 2, respectively. All regressions control for district fixed effects (the stratifying variable). A post double-lasso procedure is used to select controls in the even columns, and even columns also include an indicator variable for the report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

Table A5: Effects on School-Level Test Scores by Subject

	Public Schools			Private Schools		
	(1)	(2)	(3)	(4)	(5)	(6)
	Math	Urdu	English	Math	Urdu	English
Treatment	0.295***	0.285**	0.109	0.133	0.201**	0.231**
	(0.106)	(0.126)	(0.099)	(0.112)	(0.100)	(0.103)
Control Mean	-0.715	-0.599	-0.336	0.100	0.249	0.583
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.288	0.167	0.360	0.288	0.279	0.235
Observations	231	231	231	193	193	193
Clusters	80	80	80	74	74	74

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports estimates of the program's effect on school-level test scores by subject in round 5 (2011-12). Tests are scored using item response theory (IRT), and test scores are measured in standard deviations. All columns use double-lasso to select the controls and include baseline controls for the outcomes from rounds 1 and 2, as well as a control for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

Table A6: Heterogeneous Effects on Mean Test Scores by Gender

	Public Schools			Private Schools		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	Lasso	OLS	OLS	Lasso
Treatment	0.070	0.051	0.063	0.260**	0.301***	0.448***
	(0.112)	(0.107)	(0.111)	(0.111)	(0.094)	(0.128)
Treatment × Female	0.113	0.110	0.117	-0.227***	-0.198***	-0.185**
	(0.143)	(0.145)	(0.146)	(0.072)	(0.072)	(0.071)
Female	0.301**	0.305**	0.298**	0.349***	0.320***	0.321***
	(0.122)	(0.124)	(0.124)	(0.045)	(0.044)	(0.043)
Female Effect	0.183*	0.162	0.179	0.034	0.102	0.263**
	(0.108)	(0.103)	(0.109)	(0.112)	(0.089)	(0.131)
Control Mean	-0.409	-0.409	-0.409	0.252	0.252	0.252
Baseline Controls	No	Yes	Yes	No	Yes	Yes
Adjusted R^2	0.174	0.178	0.184	0.133	0.189	0.206
Observations	4894	4894	4894	2932	2932	2932
Clusters	80	80	80	74	74	74

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table estimates heterogeneous effects on test scores by gender. The outcome variable is child-level average test scores (across tests in math, English, and Urdu) in round 5 (2011-12). Tests are scored using item response theory (IRT), and test scores are measured in standard deviations. Columns 1-3 include public schools, and columns 4-6 include private schools. Columns 1 and 4 control only for district fixed effects (the stratifying variable), columns 2 and 5 add the round 1 and round 2 (baseline) village-level value of the dependent variable, and columns 3 and 6 use a post double-lasso procedure to select additional baseline controls. Columns 2, 3, 5, and 6 also include a control for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

Table A7: Heterogeneous Effects on Mean Test Scores by Wealth

	Public Schools			Private Schools		
	(1) OLS	(2) OLS	(3) Lasso	(4) OLS	(5) OLS	(6) Lasso
Treatment	0.153*	0.162*	0.142	0.109	0.142	0.168
	(0.091)	(0.085)	(0.091)	(0.111)	(0.091)	(0.121)
Treatment × Assets	0.015	0.025	0.025	0.023	0.034	0.036
	(0.046)	(0.046)	(0.046)	(0.033)	(0.035)	(0.035)
Assets	0.018	0.012	0.017	-0.017	-0.007	-0.010
	(0.039)	(0.039)	(0.039)	(0.027)	(0.029)	(0.029)
Control Mean	-0.495	-0.495	-0.495	0.317	0.317	0.317
Baseline Controls	No	Yes	Yes	No	Yes	Yes
Adjusted R^2	0.172	0.176	0.185	0.126	0.187	0.191
Observations	2140	2140	2140	1645	1645	1645
Clusters	80	80	80	74	74	74

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table estimates heterogeneous effects on test scores by wealth. The outcome variable is child-level average test scores (across tests in math, English, and Urdu) in round 5 (2011). Tests are scored using item response theory (IRT), and test scores are measured in standard deviations. The assets measure is constructed using the first factor from a principal components analysis of asset ownership variables (see list in the description of the asset index in Table A1). Columns 1-3 include only public schools and columns 4-6 include only private schools. Columns 1 and 4 control only for district fixed effects (the stratifying variable), columns 2 and 5 add the round 1 and round 2 (baseline) village-level values of the dependent variable, and columns 3 and 6 use a post double-lasso procedure to select additional baseline controls. Columns 2, 3, 5, and 6 also include a treatment indicator for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

Table A8: Effects on Public and Private School Composition

	Share Low Caste			Mom Education			Dad Education			Assets			Test Scores
	(1) All	(2) Public	(3) Private	(4) All	(5) Public	(6) Private	(7) All	(8) Public	(9) Private	(10) All	(11) Public	(12) Private	(13) Private
Treatment	0.030	0.017	0.033	-0.026	-0.052	0.036	-0.031	-0.058*	0.006	-0.014	0.012	0.174	0.198**
	(0.025)	(0.033)	(0.028)	(0.028)	(0.035)	(0.042)	(0.024)	(0.034)	(0.030)	(0.093)	(0.165)	(0.123)	(0.089)
Control Mean	0.182	0.191	0.176	0.512	0.453	0.572	0.671	0.612	0.741	0.163	0.076	0.236	0.310
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.272	0.288	0.269	0.141	0.120	0.184	0.046	0.008	0.085	0.079	0.013	0.097	0.298
Observations	439	232	202	428	231	193	428	231	193	428	231	193	193
Clusters	80	80	74	80	80	74	80	80	74	80	80	74	74

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table tests whether public or private school composition changed due to the treatment. We examine the share of low caste students in the school (columns 1-3), the share of students for whom the mother or father, respectively, has some education (columns 4-6 and 7-9), and the average asset index of students' households (columns 10-12). Column 13 reports the treatment effect on average test scores (in math, English, and Urdu) in private schools, controlling for the endogenous school composition controls (share low caste, mother and father education, and mean assets). Tests are scored using item response theory (IRT), and test scores are measured in standard deviations. All columns use a post double-lasso procedure to select baseline controls and include a control for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

Table A9: Effects on Private School Entry and Exit Measured by Market Share

	Entrant Market Share		Exiter Market Share	
	(1) OLS	(2) Lasso	(3) OLS	(4) Lasso
Treatment	0.013	0.015	0.019	0.023
	(0.028)	(0.029)	(0.018)	(0.017)
Control Mean	0.074	0.074	0.055	0.055
Adjusted R^2	-0.019	-0.040	-0.008	0.094
Observations	80	80	80	80

Notes: $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports estimates of the program effect on the share of primary students attending private schools that entered the market between rounds 2 and 5 (columns 1-2) and belonged to schools that exited between rounds 2 and 5 (columns 3-4). Market shares in columns 1-2 are based on enrollments in round 5, and market shares in columns 3-4 are based on enrollments in round 2. All regressions control for district fixed effects (the stratifying variable), and even columns use a post-double lasso procedure to select additional village-level baseline control variables and also include a control for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

Table A10: Heterogeneity in Treatment Effects on Private School Exit by Ex-ante School Quality

	Private Schools	
	(1)	(2)
	OLS	Lasso
Treatment	-0.026	-0.009
	(0.134)	(0.136)
Treatment × SVA	0.176	0.156
	(0.233)	(0.241)
SVA	-0.169	-0.120
	(0.188)	(0.195)
Control Mean	0.306	0.306
Adjusted R^2	-0.009	-0.007
Observations	198	198
Clusters	76	76

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table estimates the effect of the treatment on school exit, allowing the treatment to have heterogeneous effects by ex-ante school quality. "SVA" is the school value-added for mean test scores, measured during the two pre-treatment periods and shrunk using empirical Bayes. The first column only controls for district fixed effects, while the second column includes an indicator variable for the report card treatment and controls selected via double-lasso. As, by definition, no schools in this regression exited in the pre-treatment period, it is not possible to control for the baseline value of the outcome. Standard errors are clustered at the village-level.

Table A11: Effect on School Investment in Basic Infrastructure

	Perm. Class. per Student	S-Perm. Class. per Student	Toilet per Student	Blackboard per Student	Sitting Arrange- ment	Avg. Effect
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Public Schools						
Treatment	-0.000	0.001**	-0.001	0.002	0.034	0.084
	(0.003)	(0.000)	(0.002)	(0.004)	(0.075)	(0.084)
Control Mean	0.0325	0.0002	0.0133	0.0456	0.4902	
Adjusted R^2	0.141	0.004	0.120	0.231	0.044	
Observations	232	232	232	232	232	232
Clusters	80	80	80	80	80	80
Panel B: Private Schools						
Treatment	0.005	0.001	-0.001	0.013***	-0.014	0.166*
	(0.004)	(0.001)	(0.002)	(0.004)	(0.061)	(0.090)
Control Mean	0.0472	0.0017	0.0090	0.0512	0.8235	
Adjusted R^2	0.034	0.030	0.141	0.101	0.054	
Observations	202	202	202	202	202	202
Clusters	74	74	74	74	74	74

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports the effect of the program on public (Panel A) and private school (Panel B) investments in basic infrastructure in round 5 (2011-12). Outcomes include the number of permanent classrooms per student (column 1), the number of semi-permanent classrooms per student (column 2), the number of toilets per student (column 3), the number of blackboards per student (column 4), and the share of students who sit at desks or chairs (column 5). Column 6 reports the average effect size across outcomes. All columns use a post double-lasso procedure to select the controls, control for district fixed effects (the stratifying variable), and include a control for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

Table A12: Effect on School Investment in Extra Infrastructure

	Library	Computer	Sports	Hall	Wall	Fan	Electricity	Avg. Effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Public Schools								
Treatment	0.014	-0.004	-0.009	0.047	0.096**	0.055	-0.004	0.092
	(0.048)	(0.047)	(0.044)	(0.049)	(0.037)	(0.042)	(0.036)	(0.062)
Control Mean	0.196	0.206	0.167	0.157	0.853	0.882	0.941	
Adjusted R^2	0.061	0.078	0.046	0.110	0.090	-0.006	0.028	
Observations	232	232	232	232	232	232	232	232
Clusters	80	80	80	80	80	80	80	80
Panel B: Private Schools								
Treatment	0.054	0.057	0.011	-0.034	-0.024	-0.001	0.003	0.065
	(0.074)	(0.052)	(0.062)	(0.079)	(0.020)	(0.027)	(0.028)	(0.061)
Control Mean	0.318	0.353	0.306	0.282	1.000	0.976	0.976	
Adjusted R^2	0.055	0.182	0.100	0.037	0.019	0.049	0.154	
Observations	202	202	202	202	202	202	202	202
Clusters	74	74	74	74	74	74	74	74

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports the effect of the program on public (Panel A) and private school (Panel B) investments in non-basic infrastructure in round 5 (2011-12). Outcome variables are all indicator variables equal to one if the school has: a library (column 1), a computer (column 2), a sports area (column 3), a meeting hall (column 4), a boundary wall (column 5), any fans (column 6), and electricity (column 7). Column 8 presents the average effect size across these outcomes. All columns use a post double-lasso procedure to select baseline controls, control for district fixed effects (the stratifying variable), and include a control for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

Table A13: Private School Exit by Distance to Public Schools and Value-Added of Public Schools in the Village

	Distance		SVA	
	(1) OLS	(2) Lasso	(3) OLS	(4) Lasso
Treatment	0.149	0.085	0.063	0.002
	(0.109)	(0.103)	(0.067)	(0.072)
Treatment × Avg Log Dist. Public Schools	0.088	0.018		
	(0.085)	(0.082)		
Avg Log Dist. Public Schools	0.002	0.067		
	(0.065)	(0.062)		
Treatment × Avg. Village-Level Public School SVA			0.071	-0.024
			(0.081)	(0.071)
Avg. Village-Level Public School SVA			-0.052	-0.011
			(0.068)	(0.057)
Effect at 90th perc. (1.057km)	0.135	0.072	–	–
	(0.105	(0.100)	–	–
Effect at 50th perc. (466m)	0.066	0.006	–	–
	(0.096)	(0.093)	–	–
Effect at 10th perc. (223m)	-0.001	-0.058	–	–
	(0.105)	(0.103)	–	–
Control Mean	0.300	0.300	0.300	0.300
Baseline Controls	No	Yes	No	Yes
Adjusted R^2	-0.002	0.030	-0.007	0.076
Observations	209	209	209	209
Clusters	78	78	78	78

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Columns 1 and 2 examine heterogeneity in the relationship between private school exit and the average distance from private schools to public schools in the village using GPS data from round 1. Columns 3 and 4 examine heterogeneity in the relationship between private school exit and the average quality of public schools in the village, calculated using school value-added in mean test scores for each public school in round 2 using empirical Bayes to correct for estimation error (Andrabi et al., 2022b). We normalize the village-level average to have a mean of 0 and standard deviation of 1. In all columns, the outcome variable is an indicator variable equal to one if a school closed down between rounds 2 and 5, and the sample consists of schools open in round 2. Columns 1 and 3 control only for district fixed effects (the stratifying variable), and columns 2 and 4 use a post double-lasso procedure to select controls and also include a treatment indicator for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level. The bottom panel reports treatment effects at the 90th, 50th, and 10th percentiles of the average log distance distribution.

Table A14: Lasso Regressions of Grant Amounts on School and Village Characteristics

	Att. Fais.		Rahim Yar Khan		
	(1)	(2)	(3)	(4)	(5)
Extra Facility Round 2 (Vill.)			6.406***		
			(1.384)		
Extra Facility Round 1				3.300***	
				(0.734)	
Primary Enroll. Round 1					0.089
					(0.077)
Primary Enroll. Round 2					0.030
					(0.066)
No Variables Selected	Yes	Yes	No	No	No
Mean Outcome	30.326	30.326	19.497	19.497	19.497
Potential School Controls	Yes	Yes	No	Yes	Yes
Potential Village Controls	Yes	No	Yes	Yes	No
Village Fixed Effects	No	Yes	No	No	Yes
Adjusted R^2	0.000	-0.045	0.025	0.074	0.144
Observations	262	262	253	253	253
Clusters	80	80	32	32	32

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table explores the relationship between the amount of funding received by a public school and the school and village's characteristics. The outcome variable is the school-level cumulative amount of funding received by round 5 (2011-12) in 10,000 PKR. Columns 1 and 2 present results for Attock and Faisalabad districts (the experimental sample), and columns 3-5 report results for the district of Rahim Yar Khan (which did not agree to randomize). All columns control for district fixed effects (the stratifying variable) and use a lasso specification to select village- and/or school-level baseline variables. The row "No Variables Selected" reports "Yes" if the lasso did not select any predictors and "No" otherwise. We do not include a column where we only include the village predictors for the experimental sample since no predictors are selected when we search over the village and school variables together. Village fixed effects are also included in columns 2 and 5. Standard errors are clustered at the village-level.

Table A15: Relationship Between Spending and Mean Test Scores

	All Schools			Public Schools			Private Schools		
	(1) OLS	(2) OLS	(3) Lasso	(4) OLS	(5) OLS	(6) Lasso	(7) OLS	(8) OLS	(9) Lasso
Average Pub. School Grant	0.006***	0.005***	0.006***	0.004**	0.004*	0.004*	0.009***	0.008***	0.008***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Control Mean	-0.157	-0.157	-0.157	-0.550	-0.550	-0.550	0.310	0.310	0.310
Baseline Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Adjusted R^2	0.232	0.249	0.274	0.300	0.293	0.293	0.266	0.327	0.327
Observations	428	428	428	231	231	231	193	193	193
Clusters	80	80	80	80	80	80	74	74	74

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table estimates the relationship between the average amount of funding received per public school in a village and learning. The outcome variable is school-level average test scores (across tests in math, English, and Urdu and across all students in the school) in round 5 (2011). Tests are scored using item response theory (IRT), and test scores are measured in standard deviations. The explanatory variable is the average total amount of funding received per public school in the village by round 5 (in 10,000 PKR). All schools are included in columns 1-3, public schools in columns 4-6, and private schools in columns 7-9. Each set of three columns follows the same format. The first column controls only for district fixed effects (the stratifying variable), the second column additionally controls for the baseline values of the dependent variable from rounds 1 and 2 (if available), and the third column uses a post double-lasso procedure to select additional baseline controls. The second and third columns also include a treatment indicator for a report card intervention (Andrabi et al., 2017). Standard errors are clustered at the village-level.

Table A16: Estimates Used for Cost-Effectiveness Calculations

	(1) Test Scores	(2) Per Student Enrolled Primary	(3) Test Scores Public	(4) Per Student Enrolled Public Primary
Treatment	0.145**	6.641*	0.138*	11.655**
	(0.071)	(3.674)	(0.082)	(5.641)
Control Mean	-0.159	18.137	-0.409	27.601
Adjusted R^2	0.114	0.017	0.143	0.028
Observations	7928	80	4894	80
Clusters	80	80	80	80

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This table reports the estimates used for the cost-effectiveness analysis. Columns 1 and 3 report the treatment effect of the program on child-level test scores in all schools and in public schools only, respectively. Columns 2 and 4 report the treatment effect on the cumulative amount of funding (in 2011 USD) per primary school student in the village and per public school student in the village, respectively. All regressions control for district fixed effects (the stratifying variable), and standard errors are clustered at the village-level.