Upping the Ante: The Equilibrium Effects of Unconditional Grants to Private Schools

Tahir Andrabi, Jishnu Das, Asim I Khwaja, Selcuk Ozyurt, and Niharika Singh [*]

February 16, 2020

Abstract

We assess whether financing can help private schools, which now account for one-third of primary school enrollment in low- and middle-income countries. Our experiment allocated unconditional cash grants to either one (L) or all (H) private schools in a village. In both arms, enrollment and revenues increased, leading to above market returns. However, test scores increased only in H schools, accompanied by higher fees, and a greater focus on teachers. We provide a model demonstrating that market forces can provide endogenous incentives to increase quality and increased financial saturation can be used to leverage competition, generating socially desirable outcomes.

JEL Codes: I25; I28; L22; L26; O16

Keywords: Private schools, Financial innovation, Educational Achievement, Education Markets, Return to Capital, SMEs

Rising global demand for education, coupled with an increasing recognition that addressing market failures in education does not always necessitate government *provision*, has led to the proliferation of schooling models, including private schooling. The experience with these models suggests that both design features and the underlying market structure mediate their impact.[1] However, establishing the causal impact of enabling policies for schools and understanding the link between impact, program design, and market structure remains challenging. The rise of private schooling in low- and middle-income (LMIC) countries offers an opportunity to map policies to school responses by designing interventions that uncover and address underlying market failures.[2] In previous work, we have leveraged "closed" education markets in rural Pakistan to evaluate interventions that address labor market and informational constraints in these settings (Andrabi et al., 2013, 2017).[3]

We now extend this approach to *school financing* among private schools in rural Pakistan. Our experiment (randomly) allocates an unconditional cash grant of Rs.50,000 ($500 or 15% of the median annual revenue for sample schools) to each treated (private) school from a sample of 855 private schools in 266 villages in the province of Punjab, Pakistan. We assign villages to a control group and one of two treatment arms: In the first treatment, referred to as the 'low-saturation' treatment or L arm, we offer the grant to a single, randomly assigned, private school within the village (from typically 3 private schools). We denote schools that receive (did not receive) the grants in L arm villages as L^t (L^u) schools. In the second treatment, the 'high-saturation' treatment or H arm, all private schools in the randomly assigned village are offered the Rs.50,000 grant. We refer to (all) schools that received the grants in the H arm as H schools. This saturation design is motivated by our previous research in education documenting the role of market competition in determining supply-side responses (Andrabi et al., 2017) as well work showing the return on funds may be smaller if all firms in a market receive financing (Rotemberg, 2019).

The experimental design first allows us to examine whether financial provision, regardless of saturation, can impact private school expansion and quality.[4] To do so, we

---

[1]Examples range from vouchers (Hsieh and Urquiola, 2006; Muralidharan et al., 2015; Barrera-Osorio et al., 2017; Neilson, 2017) to charter schools (Hoxby and Rockoff, 2004; Hoxby et al., 2009; Angrist et al., 2013; Abdulkadiroğlu et al., 2016) and, more recently, public–private partnership arrangements with private school chains (Romero et al., 2017).

[2]Private sector primary enrollment shares are 40% in countries like India and Pakistan and 28% in all LMIC combined with significant penetration in rural areas (Baum et al., 2013; Andrabi et al., 2015).

[3]Because villages are "closed"— children attend schools in the village and schools in the village are mostly attended by children in the village— it is easier to both define markets and to isolate the impact of interventions on a schooling market as a whole.

[4]Even if private schools lack access to finance, the results from the literature on small and medium enterprises (SME), discussed in Banerjee and Duflo (2012) and de Mel et al. (2012), may not extend to education. For instance, even if it is needed, more financing will not improve outcomes if parents are unable to discern and pay for quality improvements; school owners themselves do not know what innovations increase quality; alternate uses of such funds provide higher returns; or bargaining within the family limits how these funds can be used to improve schooling outcomes (de Mel et al., 2012). Alternatively, financial constraints may be exacerbated in the educational sector with fewer resources that can be used as collateral, social considerations that hinder fee collection and enforcement, and outcomes that are multi-dimensional and difficult to value for lenders.

analyze the overall impact of the grant by combining both treatments to estimate a single "pooled" treatment effect.[5] Our results show that schools receiving grants report higher fixed (but not variable) expenditures, with most of the additional spending occurring in the first year after grant receipt. They also report higher revenues driven by higher enrollment, leading to an internal rate of return (IRR) between 37-58% using our preferred approach. However, in the pooled treatment, we do not find any increase in test scores or school fees.

We next examine the experimental results separating the two treatments. In the L arm, where only one (randomly selected) school received the grant, treated (L^t) schools enrolled an additional 22 children, resulting in substantial revenue increases that persist for at least two years after the grant award. There are no increases in test scores or fees. L^t schools incur higher fixed expenditures, but no increase in variable expenditures. Therefore, the revenue increase translates directly into increased profits and we estimate that the IRR of the cash grant to be between 92-114% using our preferred approach, which is significantly above market lending rates. Finally, closure rates are also 9 percentage points lower among L^t schools, suggesting that the grant helped schools that were on the cusp of shutting down.

We then turn to results for the H arm where all private schools in the village received the grant. Enrollment also increases in H schools, though the impact is smaller (9 additional children per school) compared to L^t schools. Importantly, and unlike the L arm, test scores improve by 0.15 standard deviation in these schools, accompanied by an increase in tuition fees of Rs.19 or 8% of baseline fees. We confirm that the test score results are not driven by composition effects: Test scores gains are identical among children who were in the same school throughout the experiment. Additional checks and bounding exercises using data from a longitudinal study of learning in rural Pakistan lend further support against such concerns. The improvement in test scores from an unconditional grant is uncommon in the education literature; we discuss below how financial saturation can provide endogenous incentives for quality improvements in private schools.[6] Although revenue increases among H schools thus reflect an increase in both enrollment and fees, they still fall short relative to that in L^t schools. Moreover, H schools show an increase in fixed expenditures *and*, unlike L^t schools, a persistent increase in variable expenditures. Thus, the IRR, while still at or above market lending rates, is estimated to be lower for H relative to L^t schools.

A more detailed examination shows that spending patterns in L^t and H schools were also different. L^t schools invested primarily in desks, chairs and computers. H

---

[5]We pool H and L^t schools ('treatment schools') and compare them to control and L^u schools ('comparison' schools).

[6]The literature on school grants is based primarily on public schools. The lack of a quality response to school grants in public schools could reflect specific design features of these programs. For instance, if grant expenditures are restricted to items that can be purchased at home, they will crowd-out household spending (Das et al., 2013). Further, grants without performance incentives may even lower quality if the items purchased substitute for quality investments, such as teacher effort. Mbiti et al. (2019) demonstrate that grants to public schools when accompanied with explicit performance incentives indeed increase test scores.

schools invested in these items, but also spent money upgrading classrooms, libraries, and sporting facilities. More significantly, the wage bill in H schools increased, reflecting increased pay for both existing and new teachers. There was no corresponding change for teachers in L^t schools. A hypothesis consistent with the test score increases in H schools is that schools used higher salaries to retain and recruit higher value-added teachers as well as provide larger incentives for existing teachers.[7] Finally, we also show that while the (large) positive enrollment effect in L^t schools is partly due to fewer school closures, accounting for selection in school closures does not affect our estimates of expenditures, revenues, fees, and test scores.

In interpreting our results, it is useful to reiterate the differences between the two treatments: H schools show significant increases in test scores and fees compared to control schools, and furthermore, these increases as well as those in variable expenditures were statistically different from the estimated impacts in L^t schools, which are close to zero and never significant. While we discuss alternate explanations, we formally show in the appendix that these results arise naturally once we allow for vertically differentiated firms in the canonical model of Bertrand duopoly competition with capacity constraints due to Kreps and Scheinkman (1983).

Specifically, defining quality as any investment that existing users are willing to pay more for, we show that markets are more likely to generate endogenous incentives for quality improvements when grants are made available to all, rather than a single school in the market.[8] The key intuition is that when a (credit-constrained) school receives a grant, it faces a trade-off between (i) increasing revenue by bringing in additional children who pay the existing fee, or (ii) increasing quality, which also allows them to increase fees for existing students. To the extent that the school can increase market share without substantially poaching from other private schools, it will choose to expand capacity as it can increase enrollment without triggering a costly price war. However, when all schools in a market receive grants, (only) increasing capacity is less profitable as it intensifies competition for students and could induce a price war. Instead, investing in quality mitigates the adverse competitive effect by both increasing the overall size of the market and retaining some degree of market power through (vertical) product differentiation. As we show in the appendix, this basic intuition—that the incentives to increase revenues through investments that allow schools to charge higher fees among existing students is higher in H relative to L^t schools— is robust to a number of potential modifications that improve the fit of the model for the education market.

These differences between the L and H arms also highlight a potential tension between market-based and socially preferred outcomes. Although our data are inadequate for a full welfare comparison, we can use our experimental estimates to consider gains to

---

[7]Bau et al. (2020) show that a 1 standard deviation increase in teacher value-added increases student test scores by 0.15sd in a similar sample from Punjab, and, in the private sector, this higher value-added is associated with 41% higher wages.

[8]Here, 'more likely' implies that the parameter space under which quality improvements occur as an equilibrium response is larger in H relative to the L arm.

3

all market participants (school owners, teachers, parents, and children). We find school owners benefit from an increase in profits in L^t schools, whereas these gains are partially transferred to teachers in the H arm. We argue that the gains for parents are likely of similar magnitudes across the two arms, but the test score gains for children are higher in the H arm. As a consequence, if we value test scores gains over and above parental valuations or weigh teacher salaries more than the owner's profits, the H arm becomes more socially desirable. Although a (monopolist) private financier might prefer to finance a single school in each village, the H arm may be preferable for society. In fact, there is a case to be made for government subsidies that encourage lending to multiple schools in the same village: To the extent that a lender is primarily concerned with greater likelihood of default and, using the fact that school closures were 9 percentage points lower for L^t schools, a plausible form of this subsidy is a loan-loss guarantee for private investors. This suggests that the usual "priority sector" lending policies could be augmented with a "geographical targeting" subsidy that rewards the market for increasing financial saturation in a given area— the *density* of coverage matters.

Our paper contributes to the literature on education and SMEs, with a focus on how school financing impacts growth and innovation. As a complement to education research that focuses on enhancing inputs into the production function or seeks to improve allocative efficiency through school vouchers or matching, we focus on the impact of policies that alter the overall operating environments for schools, leaving school inputs and enrollment choices to be determined in equilibrium.[9] The rise of private schools provides an impetus for such policy experimentation, as their flexibility allows schools to respond endogenously to changes in the local policy regime.[10]

Closest to our approach of evaluating financing models for schools are Romero et al. (2017) and Barrera-Osorio et al. (2017). Romero et al. (2017) shows that a PPP arrangement in Liberia increased test scores, albeit at costs that were higher than business-as-usual approaches and with considerable variation across providers. In Pakistan, Barrera-Osorio et al. (2017) study a program where new schools were established by local private operators using public funding on a per-student basis. Again, test scores increased. Further, decentralized input optimization came close to what a social welfare maximizing government could achieve by tailoring school inputs to local demand. However, these interventions are not designed to exploit competitive forces *within* markets.

---

[9]McEwan (2015), Evans and Popova (2015), and JPAL (2017) provide reviews of the 'production function' approach, which changes specific schooling inputs to improve test scores; one successful approach tailors teaching to the level of the child rather than curricular standards— see Banerjee et al. (2017) and Muralidharan et al. (2016). Examples of approaches designed to increase allocative efficiency include a literature on vouchers (see Epple et al. (2015) for a critical review) and school matching algorithms (Abdulkadiroğlu et al., 2009; Ajayi, 2014; Kapor et al., 2017).

[10]Private schools in these markets face little (price/input) regulation, rarely receive public subsidies, and optimize based on local economic factors. While public schools can also change certain inputs that are locally controlled inputs, such as teacher effort, other inputs are governed through an administrative chain that starts at the province and includes the districts and are unlikely to respond to a local policy shock. In two previous papers, we show that these features permit greater understanding of the labor market for teachers (Andrabi et al., 2013) and the role of information on school quality for private school growth and test scores (Andrabi et al., 2017). In Andrabi et al. (2018), we examine the impact of similar grants to public schools, which addresses *government* rather than market failures.

Exploiting the 'closed' education markets in our setting allows us to study the nature of competition and to confirm that the specific design of subsidy schemes mediates impact (Epple et al., 2015). We are therefore able to directly isolate the link between policy and school level responses, with results that appear to be consistent with (an extension of) the theory of oligopolistic competition under credit constraints.

Our paper also contributes to an ongoing discussion in the SME literature on how best to use financial instruments to engender growth. Previous work from the SME literature consistently finds high returns to capital for SMEs in low-income countries (Banerjee and Duflo, 2012; de Mel et al., 2008, 2012; Udry and Anagol, 2006), but there is a concern that these returns reflect a movement of consumers from one firm to another, and therefore may be "crowded out" when credit becomes more widely available (Rotemberg, 2019). We extend this literature to education and simultaneously demonstrate a key trade-off between low and high-saturation approaches. While low-saturation infusions may lead SMEs to invest more in capacity and increase market share at the expense of other providers, high-saturation infusions can induce firms to offer better value to the consumer and effectively grow the size of the market by "crowding in" innovations and increasing quality. This underscores that the extent of crowd-out from a selective credit policy may not be predictive of what would happen when credit is extended to a large number of firms.

Finally, our experiment helps establish further parallels between the private school market and small enterprises. Like these enterprises, private schools cannot sustain negative profits; they obtain revenue from fee paying students and operate in a competitive environment with multiple public and private providers. We have shown previously that, with these features, the behavior of private schools can be approximated by standard economic models in the firm literature (Andrabi et al., 2017). The returns to financing private schools that we document are similar to those in the SME literature suggesting that much of the knowledge on financial design for SMEs may also be applicable to schools (Beck, 2007; de Mel et al., 2008; Banerjee and Duflo, 2012).

The remainder of the paper is structured as follows: section 1 outlines the context; section 2 describes the experiment, the data, and the empirical methodology; section 3 presents the results; section 4 discusses our results and their implications; and section 5 concludes.

1 Setting and Context

The private education market in Pakistan has grown rapidly over the last three decades. In Punjab, the largest province in the country and the site of our study, the number of private schools increased from 32,000 in 1990 to 60,000 in 2016 with the fastest growth in rural areas. In 2010-11, 38% of enrollment among children between the ages of 6 and 10 was in private schools (Nguyen and Raju, 2014). These schools operate in environments

with substantial school choice and competition; in our study district, 64% of villages have at least one private school, and within these villages there is a median of 5 (public and private) schools (NEC, 2005). They are not just for the wealthy, as 18% of the poorest third send their children to private schools in villages where they exist (Andrabi et al., 2009). One reason the demand for private schooling is high may be a relatively better learning environment: Test scores of children enrolled in private schools are 1 standard deviation higher than for those in public schools, amounting to 1.5 to 2.5 (additional) years of learning (depending on the subject) by Grade 3 (Andrabi et al., 2009). These differences remain large and significant after accounting for selection into schooling using the test score trajectories of children who switch schools (Andrabi et al., 2011).

These higher test scores are accompanied by relatively low private school fees. In our sample, the median private school reports a fee of Rs.201 or $2 per month, which is less than half the daily minimum wage in the province. We have argued previously that the 'business model' of these private schools relies on the local availability of secondary school educated women with low salaries and frequent churn (Andrabi et al., 2008). A typical teacher in our sample is female, young, and unmarried, and is likely to pause employment after marriage and her subsequent move to the marital home. In villages with a secondary school for girls, there is a steady supply of such potential teachers, but also frequent bargaining between teachers and school owners around wages. An important feature of this market is that the occupational choice for teachers is not between public and private schools: Becoming a teacher in the public sector requires a college degree, and an onerous and highly competitive selection process as earnings are 5-10 times as much as private school teachers and applicants far outweigh the intake. Accordingly, transitions from public to private school teaching and vice versa are extremely rare.

Despite their success in producing higher test scores (relative to the public sector) at fairly low costs, once a village has a private school, future quality improvements appear to be limited. We have collected data through the Learning and Educational Achievement in Pakistan Schools (LEAPS) panel for 112 villages in rural Punjab, each of which reported a private school in 2003. Over five rounds of surveys spanning 2003 to 2011, tests scores remain constant in "control" villages that were not exposed to any interventions from our team. Furthermore, there is no evidence of an increase in the enrollment share of private schools or greater allocative efficiency whereby more children attend higher quality schools. This could represent a (very) stable equilibrium, but could also be consistent with the presence of systematic constraints that impede the quality and growth potential of this sector.

Our focus on finance as one such constraint is driven, in part, by what school owners themselves tell us. In our survey of 800 school owners, two-thirds report that they want to borrow, but only 2% report any borrowing for school related loans.[11] School

---

[11]This is despite the fact that school owners are highly educated and integrated with the financial system: 65% have a college degree; 83% have at least high school education; and 73% have access to a

owners wish to make a range of investments to improve school performance as well as their revenues and profits. The most desired investments are in infrastructure, especially additional classrooms and furniture, which owners report as the primary means of increasing revenues. While also desirable, school owners find raising revenues through better test scores and therefore higher fees a more challenging proposition. Investments like teacher training that may directly impact learning are thought to be risky as they may not succeed (the training may not be effective or a trained teacher may leave) and even if they do, they may be harder to demonstrate and monetize. In this setting, alleviating financial constraints may have positive impacts on educational outcomes; whether these impacts arise due to infrastructure or pedagogical improvements depends on underlying features of the market and the competitive pressure schools face.

2   Experiment, Data and Empirical Methods

2.1   Experiment

Our intervention tests the impact of providing financing to schools on revenue, expenditures, enrollment, fees, and test scores and assesses whether this impact varies by the degree of financial saturation in the market. Our intervention has three features: (i) it is carried out only with private schools where all decisions are made at the level of the school;[12] (ii) we vary financial saturation in the market by comparing villages where only one (private) school receives a grant (L arm) versus villages where all (private) schools receive grants (H arm); and (iii) we never vary the grant amount at the school level, which remains fixed at Rs.50,000. We discuss in turn, the sample, the randomization and the experimental design.

2.1.1   Sample

Our sampling frame is defined as all villages in the district of Faisalabad in Punjab province with at least 2 private or NGO schools; 42% (334 out of 786) of villages in the district fall in this category. Based on power calculations using longitudinal LEAPS data, we sampled 266 villages out of the 334 eligible villages with a total of 880 schools, of which 855 (97%) agreed to participate in the study. Table 1 (Panel A) shows that the median village has 2 public schools, 3 private schools and 416 children enrolled in private schools. Table 1 (Panel B) shows that the median private school at baseline has 140 enrolled children, charges Rs.201 in monthly fees, and reports an annual revenue of Rs.317,820. Annual variable expenditures are Rs.194,400 and annual fixed expenditures

---

bank account.

[12]This excludes public schools, which cannot charge fees and lack control over hiring and pedagogic decisions. In Andrabi et al. (2018), we study the impact of a parallel experiment with public schools between 2004 and 2011. It also excludes 5 (out of 880) private schools that were part of a larger school chain with schooling decisions taken at the central office rather than within each school.

are Rs.33,000. The range of outcome variables is quite large. Relative to a mean of 164 students, the 5th percentile of enrollment is 45 compared to 353 at the 95th percentile of the distribution. Similarly, fees vary from Rs.81 (5th percentile) to Rs.503 (95th percentile), and revenues from Rs.59,316 to Rs.1,411,860. The kurtosis, a measure of the density at the tails, is 17 for annual fixed expenditures and 51 for revenues relative to a kurtosis of 3 for a standard normal distribution. Our decision to include all schools in the market provides external validity, but the resulting wide variation has implications for precision and mean imbalance, both of which we discuss below.

2.1.2 Randomization

We use a two-stage stratified randomization design where we first assign each village to one of three experimental groups and then schools within these villages to treatment. Stratification is based on village size and village average revenues, as both these variables are highly auto-correlated in our panel dataset (Bruhn and McKenzie, 2009). Based on power calculations, $\frac{3}{7}$ of the villages are assigned to the L arm, and $\frac{2}{7}$ to the H arm and the control group; a total of 342 schools across 189 villages receive grant offers (see Appendix Figure A1). In the second stage, for the L arm, we randomly select one school in the village to receive the grant offer; in the H arm, all schools receive offers; and, in the control group, no schools receive offers. The randomization was conducted through a public computerized ballot in Lahore on September 5, 2012, with third-party observers (funders, private school owners and local NGOs) in attendance.[13] Once the ballot was completed, schools received a text message informing them of their own ballot outcome. Given village structures, information on which schools received the grant in the L arm was not likely to have remained private, so we assume that the receipt of the grant was public information.

2.1.3 Experimental Design

Grant amount: We offer unconditional cash grants of Rs.50,000 (approximately $500 in 2012) to every treated school in both L and H arms. The size of the grant represents 5 months of operating profits for the median school and reflects both our overall budget constraint and our estimate of an amount that would allow for meaningful fixed and variable cost investments. For instance, the median wage for a private school teacher in our sample is Rs.24,000 per year; the grant thus would allow the school to hire 2 additional teachers for a year. Similarly, the costs of desks and chairs in the local markets range from Rs.500 to Rs.2,000, allowing the school to purchase 25-100 additional desks and chairs.

We do not impose any conditions on the use of the grant apart from the submission

---

[13]The public nature of the ballot and the presence of third-party observers ensured that there were no concerns about fairness; consequently, we did not receive any complaints from untreated schools regarding the assignment process.

of a (non-binding) business plan (see below). School owners retain complete flexibility over how and when they spend the grant and the amount they spend on schooling investments with no requirements of returning unused funds. As we show below, most schools choose not to spend the full amount in the first year and the total spending varies by the treatment arm. Our decision not to impose any conditions allows us to provide policy-relevant estimates for the simplest possible design; the returns we observe can be achieved through a relatively 'hands-off' approach to private school financing.

Grant Disbursement: All schools selected to receive grant offers are visited three times. In the first visit, schools choose to accept or reject the grant offer: 95% (325 out of 342) of schools accept.[14] School owners are informed that they must (a) complete an investment plan to gain access to the funds and may only spend these funds on items that would benefit the school and (b) be willing to open a one-time use bank account for cash deposits. Schools are given two weeks to fill out the plan and must specify a disbursement schedule with a minimum of two installments. In the second visit, investment plans are collected and installments are released according to desired disbursement schedules.[15] A third and final disbursement visit is conducted once at least half of the grant amount has been released. While schools are informed that failure to spend on items may result in a stoppage of payments, in practice, as long as schools provide an explanation of their spending or present a plausible account of why plans changed, the remainder of the grant is released. As a result, all 322 schools receive the full amount of the grant.

Design Confounders: If the investment plan or the temporary bank account affected decision making, our estimates will reflect an intervention that bundles cash with these additional features. We discuss the plausibility of these channels in section 4.2 and use additional variation in our experiment to evaluate the contribution of these mechanisms to our estimated treatment effects. The treatment unit in a saturation experiment is a design variable; in our case, this unit could have been either the village (total grants are equalized at the village level) or the school. We chose the latter to compare schools in different treatment arms that receive the same grant. Consequently, in the H arm, with a median of 3 private schools, the total grant to the village is 3 times as large as to the L arm. Observed differences between these arms could therefore reflect the equilibrium effects of the total inflow of resources into villages, rather than the degree of financial saturation. Using variation in village size, we show in section 4.2 that our results remain qualitatively the same when we compare villages with similar per capita grant inflows.

---

[14]Reasons for refusal include anticipated school closure; unwillingness to accept external funds; or a failure to reach owners despite multiple attempts.

[15]At this stage, 3 schools refused to complete the plans and hence do not receive any funds. Our final take-up is therefore 94% (322 out of 342 schools), with no systematic difference between the L and H arms.

2.2 Data Sources

Between July 2012 and November 2014, we conducted a baseline survey and five rounds of follow-up surveys. In each follow-up round, we survey all consenting schools in the original sample and any newly opened schools.[16]

Our data come from three different survey exercises, detailed in Appendix A. We conduct an extended school survey twice, once at baseline and again 8 months after treatment assignment in May 2013 (round 1 in Appendix Figure A2), collecting information on school characteristics, practices and management, as well as household information on school owners. In addition, there are 4 shorter follow-up rounds every 3-4 months that focus on enrollment, fees, revenues, and expenditures.[17]

Finally, children are tested at baseline and once more, 16 months after treatment (round 3). During the baseline, we did not have sufficient funds to test every school and therefore administered tests to a randomly selected half of the sample schools. We also never test children at their homes or in public schools; neither do we survey these schools. At baseline, this decision was driven by budgetary constraints and in later rounds we decided not to test children in public schools because our follow-up surveys showed enrollment increases of at most 30 children in treatment villages. Even if we were to assume that these children came exclusively from public schools, this suggests that public schools enrollment across all grades declined by less than 5%. This effect seemed too small to generate substantial impacts on public school quality, but a downside of our approach is that we cannot measure test score changes among children who left or (newly) entered private schools in our sample.[18] We discuss how this may affect the interpretation of our test score impacts in section 3.2.5.

2.3 Regression Specification

We estimate intent-to-treat (ITT) effects using the following school-level specification:[19]

$$Y_{ijt} = \alpha_s + \delta_t + \beta_1 T_{ijt} + \gamma Y_{ij0} + \epsilon_{ijt} \qquad (1)$$

---

[16]There were 31 new schools (3 public and 28 private) two years after baseline with 13 new private schools opening in H villages, 10 in L villages, and 5 in control villages. We omit these schools from our analysis, but note that H villages report a 2% higher fraction of new schools relative to control. Our main results remain qualitatively similar if we include these schools in our analyses with varying assumptions on their baseline value.

[17]Due to budgetary limitations, we varied the set of questions we asked in each of these rounds. See Appendix Figure A3 to see outcomes available by survey round.

[18]This is an important limitation of experiments such as ours. As there is regular churn between the public and private sector, identifying the marginal movers from any such experiment is fraught with difficulties (Dean and Jayachandran, 2019). Without being able to identify marginal movers, the gains from switching schools due to the treatment would have to be inferred from average movers in the population leading to very large home-based testing which was outside the scope of this project.

[19]We focus on ITT effects since take-up is near universal at 94%. To obtain the local average treatment effect (LATE), we can scale our effects by the fraction of compliers (0.94), under the assumption that our treatment effects are generated only through the receipt of the grant so that the exclusion restriction is not violated.

Y_{ijt} is the outcome of interest for a school i in village j at time t, which is measured in at least one of five follow-up rounds after treatment. T_{ijt} is a dummy variable taking a value of 1 for H and L^t schools and 0 for L^u and control schools.[20]

We use strata fixed effects, α_s, since randomization was stratified by village size and revenues, and δ_t are follow-up round dummies, which are included as necessary. Y_{ij0} is the baseline value of the dependent variable, and is used whenever available to increase precision and control for any potential baseline mean imbalance between the treated and control groups (see discussion in section 2.4). All regressions cluster standard errors at the village level. Our coefficient of interest is β_1, which provides the average ITT effect for the grant.

When we separate out the treatments, we estimate:

$$Y_{ijt} = \alpha_s + \delta_t + \beta_1 H_{ijt} + \beta_2 L_{ijt}^t + \beta_3 L_{ijt}^u + \gamma Y_{ij0} + \epsilon_{ijt} \tag{2}$$

H_{ijt}, L_{ijt}^t, and L_{ijt}^u are dummy variables for schools assigned to high-saturation villages, and treated and untreated schools in low-saturation villages respectively. Regressions are weighted to account for the differential probability of treatment selection in the L arm as unweighted regressions would assign disproportionate weight to treated (untreated) schools in smaller (larger) L villages relative to schools in the control or H arms (see Appendix A). Our coefficients of interest are β_1, β_2, and β_3, all of which identify the average ITT effect for their respective group.

One important consideration is whether to present test score results at the level of the child or the school (unweighted by enrollment). Child level results are the relevant welfare metric, but if demand responds to quality investments in the school (as it will be in the standard IO model), then the appropriate metric to understand *school* responses to the treatment is the unweighted regression at the school level.[21] Following the framework where schools first make quality investments and then demand is realized, we present test score results treating each school as a single unit. Then, when we look at welfare impacts, we return to child level test scores; the difference between the two reflects heterogeneity by school size. Although child level test score results tend to be larger, i.e. gains are greater among larger schools, the heterogeneity in the treatment effect by school size is not statistically significant.

A second important consideration is how we treat school closures. To the extent that closures differ by treatment status, they are endogenous to the treatment and a *channel* through which the treatment has altered the market for private schools. For this reason, when we present our main results, we always include closed schools in enrollment regressions as having zero enrollment and revenues, but exclude them from fee, test score, and expenditures regressions as these are, by definition, missing. When

---

[20]Excluding L^u schools from the comparison group does not alter our results.

[21]In cases where an experiment induces substantial movement, such a metric can be misleading, but as we will show, child movement is small relative to the size of the population that remains in the same school.

we discuss closure as a potential channel in section 3.3.3, we examine the extent to which our main findings are affected by closure. In doing so, we assess what the fees, test score, and expenditures for closed schools would have been had they remained open and show that impacts on these outcomes are not driven by closure.

2.4 Validity

2.4.1 Randomization Balance

To ensure the integrity of our randomization, we check for baseline differences in means and distributions and conduct joint tests of significance for key variables. We first consider balance tests at the village level in Appendix Table B1, Panel A. At the village level, the distributional tests are balanced across the three experimental groups (H, L and Control), and village level variables do not jointly predict village treatment status for either the H or the L arm. All, except 1 (out of 15), univariate comparisons are balanced as well.

Given our two-stage stratified randomization design, balance tests at the school level involve four experimental groups: L^t, L^u, H, and control schools. Panel B shows comparisons between control and each of the three treatment groups (columns 3-5) and between the H and L^t schools (column 6). Although our distributional tests are always balanced (Panel B, columns 7-9) and covariates do not jointly predict any treatment status, 5 out of 32 univariate comparisons (Panel B, columns 3-6) show mean imbalance at p-values lower than 0.10— a fraction slightly higher than what we may expect by random chance. The slight imbalance we observe however is largely a function of heavy(right)-tailed distributions arising from the inclusion of all schools in our sample, a fact we first documented in section 2.1.1. Nevertheless, if this imbalance leads to differential trends beyond what can be accounted for through the inclusion of the baseline value of the dependent variable in our specifications, our results for the L^t schools may be biased.

In order to allay concerns that our results from both specifications 1 and 2 may be driven by this imbalance, we conduct robustness of our main results by using the post-double selection lasso procedure to address imbalance. This procedure provides a principled way to select baseline controls in our regressions, beyond just the baseline of dependent variable. We discuss these checks in section 3.1.1 and 3.2.5, but note here that our results remain qualitatively similar after this adjustment.

2.4.2 Attrition Checks

Schools may exit from the study either due to closure, a treatment effect of interest that we examine in section 3.3.3, or due to survey refusals. Across all five rounds of follow-up surveys, our survey completion rates for open schools are uniformly high (95% for

rounds 1-4 and 90% for round 5). Whereas 79 unique schools refuse our survey at least once during the study period, only 14 schools refuse *all* follow-up surveys (7 control, 5 H, and 2 L^u). In addition, since round 5 was conducted 2 years after baseline, we implemented a randomized procedure for refusals, where we intensively tracked half of the schools who refused the survey in this round for an interview. We apply weights to the data from this round to account for this intensive tracking (see Appendix A for details).

Though survey completion rates are high in general, attrition does vary by treatment status (Appendix Table B2, Panel A). L^t schools are significantly less likely to attrit relative to control in every round. Attrition for the H and L^u schools, while generally lower relative to control, appears to be more idiosyncratic by round.

Despite this differential attrition, baseline characteristics of those who refuse surveying (at least once) do not in general vary by treatment status (see Panel B in Appendix Table B2).[22] There are a few idiosyncratic differences, but these could occur by chance: In Appendix Table B2, only 4 out of 24 comparisons show significant differences. Our results are similar when we adjust for attrition using inverse probability weights for both specifications 1 and 2; we discuss this further in sections 3.1.1 and 3.2.5.[23]

3 Results

In this section, we present results on the primary outcomes of interest and investigate channels of impact. We start by discussing results from the pooled treatment (specification 1) and then move to discuss results separating the H and L treatment arms (specification 2). A discussion of these results with the help of a conceptual framework followed by implications for welfare follows in section 4.

3.1 Pooled Treatment

Table 2 presents treatment effects from estimating specification 1, where we consider the impact of the pooled treatment. Panel A presents results from survey rounds during the first year; Panel B from surveys during the second year; and Panel C combines all rounds to present the average impact over the two years.

---

[22]Comparing characteristics for the at-least-once-refused set is a more conservative approach than looking at the always-refused set since the former includes idiosyncratic refusals. Since there are only 14 schools in the always-refused set, inference is imprecise; in this set of schools, only one significant difference emerges with lower enrollment in L^u relative to control schools.

[23]The procedure for re-weighting results accounting for attrition is as follows: We calculate the probability of refusal (in any follow-up round) given treatment variables and a set of covariates (fees, enrollment, revenues, test scores, fixed and variable expenditures, and infrastructure index) using a probit model, and use the predicted values to construct weights. In the probit model, only our treatment variables have any predictive power for attrition. The attrition weight is then the inverse probability of response $(1 - Pr(attrition))^{-1}$, giving greater weight to those observations that are more likely to refuse surveying.

Columns 1 and 2 first examine whether the grants lead to an increase in school expenditures.[24] We examine two types of expenditures: Fixed expenditures represent annual investments, usually before the start of the school year, for school infrastructure (furniture, fixtures, classroom upgrades) or educational materials (textbooks, school supplies); (annualized) variable expenditures are recurring monthly operational expenses on teacher salaries, the largest component of these expenses, and non-teaching staff salaries, utilities, and rent. Column 1 shows that treated schools increased their fixed expenditures by Rs.28,076 or 56% of the grant in the first year (Panel A) with no further increases in year 2 (Panel B), for an average increase of Rs.14,900 over the two years (Panel C). Since the grant size was Rs.50,000, schools spend a bit more than half the grant they receive, all within the first year. In contrast, column 2 shows that there is no increase in variable expenditures for the average treated school in either the first or the second year or on average across both years. As we show later, these results mask important differences between the two treatment arms.[25]

Columns 3 and 4 then examine the impact of receiving the grant on school revenues. Since schools may not always be able to fully collect fees from students, we use two revenue measures: (i) posted revenues based on posted fees and enrollment (column 3), calculated as the sum of revenues expected from each grade as given by the grade-specific monthly tuition fee multiplied by the grade-level enrollment; and (ii) collected revenues as reported by the school (column 4).[26] To obtain collected revenues, we inspected the school account books and computed revenues actually collected in the month prior to the survey.[27] Our results show large and persistent revenue gains for both measures. Posted revenues increase by Rs.101,189 in year 1 (Panel A) and are higher by year 2 at Rs.125,273 (Panel B) for an average increase of Rs.109,083 over two years (Panel C). This represents a 21%-26% increase over baseline posted revenues. Collected revenues increased by Rs.72,000 in year 1 (Panel A) and Rs.85,263 in year 2 (Panel B), for an average increase of Rs.78,616 on average across both years (Panel C).

Based on these impacts on revenues and expenditures, our preferred estimate of the internal rate of return (IRR) is between 37% and 58%, which is considerably higher than the prevailing market interest rates of 15-20%. The lower estimate is over two years, with an assumed resale value of 50% on assets purchased in the first year after treatment;

---

[24]As the grants were largely unconditional and spending on the household, school and other businesses is fungible, if schools were not constrained to begin with or have better alternative uses of their grant, school-related expenditures may not increase at all.

[25]In our primary specifications, we code expenditures as missing, but revenues as zero for closed schools. Revenues are coded as zero since by definition these schools have zero enrollment. In section 3.3.3, we present specifications where we instead predict expenditures and revenues for closed schools, with similar results.

[26]Posted revenues are available for rounds 1,2, and 4, and collected revenues are available from rounds 2 to 5. We do not have a baseline measure of collected revenues, so, in column 4, we instead use baseline posted revenues as our baseline control and show the follow-up control mean (across all rounds) for reference.

[27]Over 90% of schools have registers for fee payment collection, and for the remainder, we record self-reported fee collections. While this measure captures revenue shortfalls due to partial fee payment, discounts and reduced fees under exceptional circumstances, it may not adjust appropriately for delayed fee collection.

the higher number is the (extrapolated) estimate after five years with a zero asset resale value at the end of the period.[28] Our intervention of an unconditional grant with minimal supervision thus provides a directly policy actionable intervention for financial intermediaries who wish to invest in schooling, with realized returns substantially higher than the market lending rate. We discuss why such products may not already be widely available in the market and how such products may be brought to the market in the concluding section.

Columns 5 and 6 then decompose the revenue impact into its two components—changes in enrollment and changes in school fees. Column 5 shows that these additional revenues were driven primarily by higher enrollment, with 15 more children by year 1 (Panel A) and 19 more by year 2 (Panel B), for an average increase of 10% over baseline enrollment. In column 6, we see that although fees are higher, the increase is not statistically significant at p-values of 0.10 or below.

Column 7 then shows that there is no increase in the test scores of children for the pooled treatment, a result that is consistent with research on grants to public schools. Das et al. (2013) have suggested that that lack of quality increases may be because grants are often required to be spent on items that can be substituted by household spending; in particular, schools typically cannot use these grants to hire additional teachers, better teachers, or top-up teacher salaries. Mbiti et al. (2019) suggest that grants without performance incentives could even decrease test scores, if they are used to finance inputs that are substitutes for test scores. We return to these issues when we separate out the results by the H and L arms.

Finally, given the increased revenues and generally positive impact of the grant, we also ask whether these changes were significant enough to prevent school closure. Column 8 shows that this is indeed the case. Over the two year period, the closure rate for treated schools is 4 percentage points lower relative to a 13.7% closure rate in the control group. While these are small changes in absolute numbers and we do not know if the treatment simply delays inevitable closures, this effect does suggest that the grant had meaningful impacts. Moreover, it also suggests that preventing closure may be an important channel behind the enrollment impacts we observe. We will discuss this issue in more detail below.

---

[28]Our calculations in Appendix A account for closed schools using three different options, each of which can be justified under varying assumptions regarding what school owners do when a school closes and what an investor may have claims over. First, we consider only open schools, i.e. we treat both revenues and expenditures as missing for closed schools; second, we treat closed schools as having zero revenues and expenditures; and third, we predict both revenues and expenditures for closed schools. Our preferred approach is the first one, given in the text. Our estimates using the second approach for the 2 and 5-year calculations are 67% and 87%, and using the third approach are 45% and 67%, respectively.

3.1.1 Robustness

We now ensure that our main results are robust to two potential concerns, attrition due to survey non-response and baseline imbalance, discussed earlier in section 2.4.[29]

Appendix Table B3 repeats our pooled treatment effects from Table 2 with two modifications. We show attrition re-weighted estimates for years 1, 2, and averaged across the two years in Panels A, C, and E, respectively. We account for chance imbalance in our covariates by using the post double selection lasso methodology to select controls for our regressions in Panels B, D, and E for years 1, 2 and averaged across the two years, respectively.[30] Across both modifications, we continue to see large, significant average effects on fixed expenditures, posted revenues, and enrollment; collected revenues also see large gains under both modifications, but the attrition-reweighted results are noisier with p-values between 0.146-0.174, whereas the results accounting for imbalance are significant at p-values between 0.016-0.092. As before, we see no changes in test score. One consistent difference is that fee increases are slightly larger and closer to significance at traditional p-values in some specifications, suggesting that revenue gains arose both from new students and increasing fees among existing students. As we will show later, this is due to substantial heterogeneity in the impact of the intervention on fees across the two treatment arms. We also note that variable expenditures are larger in these specifications, though not statistically significantly so. As in Table 2, we find that fixed expenditures are the only variable that shows a large and significant effect in year 1, but no effect in year 2.

3.2 H and L Treatments

While the pooled treatment shows substantial impact on expenditures, enrollment, and revenues, we did not find impacts on test scores and fees. As alluded to previously, this lack of an overall impact on test scores and fees masks heterogeneity between the two saturation approaches. We turn to this next. We first present effects of the treatment in H and L arms for expenditures, revenues, enrollment, school closure, fees, and, finally, test scores. We then take a more detailed look as to how schools spent the money they received and whether closure affects our overall results in section 3.3. Given that we usually we do not see meaningful differences across the two years in the impact of the grant, we focus on the average impact across the two years, with separate results for year 1 and year 2 presented in Appendix B.

---

[29]Since closure is a channel of impact, we examine it separately later on in section 3.3.3, looking at both whether our results are partly driven due to impacts on closure and also how they would change had the differential closure not occurred. The latter also serves as a check on whether the coding choices in our primary specification— enrollment and revenues coded as zero for closed schools, and expenditures, fees, and test scores as missing— have any qualitative effect on our results.

[30]Since we always observe closure, it does not suffer from any attrition issues. We thus omit the closure regression from Panels A, C, and E.

3.2.1 Fixed and Variable Expenditures

Table 3 presents treatment effects on (annualized) fixed and variable expenditures. We separately estimate the impact on H, L^t and L^u schools, averaging our results across the two years after treatment.[31] Column 1 shows that *averaged* over the two years, fixed expenditures were higher in both H and L^t schools. To account for large right-tailed values in the expenditure data, we present two additional specifications where we "top-code" (assign the top 1% of observations the value at the 99th percentile) or "trim" (drop the top 1% of observations) the data, with broadly similar results (columns 2 and 3). Further, like for the pooled treatment, we show in Appendix Table B4 that fixed expenditures increased for H and L^t schools only during the first year as schools spend a majority of the grant money in the first year with no change in the second year.

Columns 4-6 present analogous results for variable expenditures. In contrast to our results on fixed expenditures, we find that variable expenditures increased in H (a 10-12% increase over baseline expenditures), but not L^t schools. This difference between H and L^t schools is statistically significant ($p<0.05$). Separating the results by both years shows that, unlike fixed expenditures, variable expenditures for H schools were higher in both years, with suggestive evidence of growth across the two years (Appendix Table B4, columns 4-6).

Note that we never find any significant change in fixed or variable expenditures among L^u schools, suggesting that these schools did not respond (to the L^t school) by trying to increase their own expenditures, perhaps because they were constrained in their ability to do so.

3.2.2 Revenues

Table 4 presents the impacts on annual school revenues averaged across the two years after treatment. Column 1 shows that on average schools in the H arm gained Rs.65,813 ($p=0.12$) while L^t schools gained Rs.127,987 ($p=0.03$) in posted revenues a year, both of which compare favorably to one-time disbursement of the Rs.50,000 grant. The impact on collected revenues is similar to posted revenues for H schools (Rs.52,696 with $p=0.22$), but is smaller (Rs.94,545 with $p=0.09$) for L^t schools (column 4). One explanation for the slightly lower rates of fee collection in L^t schools could be that the marginal new children pay less than the posted fees in L^t schools. These results are large but often imprecise due to the high variance and skew in the revenue distribution. Precision increases when we either top-code (columns 2 and 5) or trim the top 1% of data (columns 3 and 6), with p-values ranging from 0.02 to 0.068. While the points estimates are 1.5-2 times larger for L^t as compared to H schools, given the large standard errors we cannot reject equality of coefficients across the treatment arms of the intervention.

---

[31] Fixed costs include infrastructure-related investments, such as upgrading rooms or new furniture and fixtures; spending on these items is consistent with self-reported investment priorities in our baseline data. Variable expenditures include teaching and non-teaching staff salaries, utilities and rent.

Finally, we never find any significant change in revenues among L^u schools, with relatively small coefficients across all specifications. We show in Appendix Table B5 that these effects are similar when we separately look at each of the two years after the grant.

These revenue and expenditures estimates show that the IRR remains quite attractive in both treatment arms. With our preferred approach of treating closed schools as missing in our IRR calculations, we estimate IRR of 92-114% for L^t schools and 10%-30% for H schools for 2-year and 5-year scenarios (see Appendix A).[32] Moreover, as interest rates on loans to this sector range from 15-20%, the IRR almost always exceeds the market interest rate: L^t schools would be able to pay back a Rs.50,000 loan in 1 year whereas H schools would take 3 years.

3.2.3 Enrollment and Fees

Table 5 now considers the impact of the grant on the two main components of school revenues—enrollment and fees. Column 1 shows that enrollment is higher in L^t schools by an average of 21.8 children (p=0.005) over the two years, compared to 9 children for H schools (p=0.137). We can reject equality of these effects at a p-value of 0.101. These gains were similar across both years (Appendix Table B6) and were experienced across all grades (Appendix Table B7). Again, we do not observe an average impact on L^u schools. To the extent that there is typically more entry at lower grades and greater drop-out in higher grades, the fact that we see fairly similar increases across these grade levels suggests that both new student entry (in lower grades) and greater retention (in higher grades) are likely to have played a role.[33]

Unlike enrollment, which increased in both arms, fees increased only among H schools. Average monthly tuition fees across all grades in H schools is Rs.18.8 higher than in control schools, an increase of 8% relative to the baseline fee (column 2). Like with enrollment, these magnitudes are similar across the two years of the intervention (Appendix Table B6) and are observed in all grades (Appendix Table B9). In contrast, we are unable to detect any impact on school fees for either L^t or L^u schools. Consequently, we reject equality of coefficients between H and L^t at a p-value of 0.021 (Table 5, column 2). As an additional check, column 3 looks at collected fees which we compute as collected revenues divided by school enrollment. Since collected fees reflect both a school's sticker price and its efforts in collecting them from parents, they are a noisier

---

[32]Using the other approaches of either treating closed schools to have zero revenues and expenditures or imputing their values, we obtain IRR between 84%-166% for L^t schools and 26% to 48% for H schools.

[33]Further information on where this enrollment increase came from is difficult to identify for two reasons. First, we would have had to track all the children in these villages over time, which is very expensive without a school attendance system and a uniform student ID. Even with this tracking, it would not have been possible to separately identify the children who moved due to the experiment from regular churn. We can, however, partly track enrollment using data on the tested children. Appendix Table B8 shows that a higher fraction of tested children report being newly enrolled in L^t and H schools, where 'newly enrolled' is defined as 'attending their current endline school for fewer than 18 months from the date of treatment assignment' (column 2). Unfortunately, these data do not allow us to distinguish when (and where) these children were last enrolled.

measure with potential under-reporting if late fees are not adequately recorded. Nevertheless, we still find a collected fee increase of Rs.21 in H schools (p=0.028) with no significant change in L^t schools. This effect for H and L^t schools is again statistically different at a p-value of 0.099.

Finally, the closure rate two years after the intervention was 9 percentage points lower among L^t schools, with no statistically significant effect among H schools (column 4). This effect is sizable, as in the control group, 13.7% of schools had closed within the same time frame. Moreover, the difference in closure is significantly different between L^t and H schools (p=0.045), hinting that a decline in school closures may be one possible channel for greater enrollment increases among L^t schools. We return to a discussion of closure as a potential channel for our results in section 3.3.3.

These results also suggest that the main increase in revenues we found for L^t schools comes from marginal children who were newly enrolled or re-enrolled from other schools, whereas two-thirds of the revenue increase among H schools is from higher fees charged to children who were already in school. We now turn to an increase in test scores as a potential reason for the increased willingness to pay among H schools.

3.2.4 Test Scores

We examine whether increases in school revenues are accompanied by changes in school quality, as measured by test scores. To assess this, we use subject tests administered in Math, English, and the vernacular, Urdu, to children in all schools 16-18 months after the start of the intervention.[34] We graded the tests using item response theory, which allows us to equate tests across years and place them on a common scale (Das and Zajonc, 2010). See Appendix A for further details on testing, sample and procedures.

Table 6, column 1, shows that the average test score (across all subjects) increases for children in H schools by 0.153sd (p=0.074).[35] This represents a 39% additional gain relative to the (0.397sd) gain children in control schools experience over the same period. Columns 2-4 show that test score impacts were similar across the different subjects, with coefficients ranging from 0.157sd in Math (p=0.082) to 0.186sd in English (p=0.049) and 0.113sd in Urdu (p=0.175). In contrast, there are no detectable impacts on test scores for L^t schools relative to control. Given this pattern, we also reject a test of equality of coefficients between H and L^t schools at p-value of 0.073 (column 1).

Finally, we tested at most two grades per school. Therefore, we cannot directly

---

[34]Budgetary considerations precluded testing the full sample at baseline, leading us to randomly chose half our villages for testing. In the follow-up round, an average of 23 children from at least two grades were tested in every school, with the majority of tested children enrolled in grades 3-5; in a small number of cases, children from other grades were tested if enrollment in these grades was zero. In tested grades, all children were administered tests and surveys regardless of class size; the maximum enrollment in any single class was 78 children.

[35]We include baseline scores where available to increase precision. Since we randomly tested half our sample at baseline, we replace missing values with a constant and an additional dummy variable indicating the missing value.

examine whether children across all grades in the school have higher test scores due to our treatment. Instead, we make two points: (i) average fees are higher across all grades in H schools and insofar as fee increases are sustained through test score increases, this suggests that test score increases likely occurred across all grades; and (ii) if we examine test scores gains in the two tested grades separately, we observe test score improvements in H schools for each grade.[36]

3.2.5   Robustness

As a first robustness test, we demonstrate our results for the two treatments are robust to accounting for attrition and chance imbalance. Appendix Table B10 accounts for attrition using inverse probability weights (Panel A) and for imbalance in treatment assignment using the post double selection lasso procedure (Panel B).[37] Across these robustness checks, both the significance of the effects in the two treatment arms and the differences between L^t and H schools remains qualitatively similar.[38]

A second robustness exercise arises from the potential concern, specific to our results on tests scores, that the effect for H schools (or the lack of an effect for L^t schools) was due to changes in child composition. We undertake several additional tests to assess the plausibility of this hypothesis. First, in Appendix Table B11, we restrict the sample to only those children who were in the same school throughout our study, which includes 90% of all children observed in the follow-up round. School test score increases based only on these stayers are 0.132sd (p=0.086) for the H arm and the difference with L^t schools remains statistically significant.[39]

Second, we assess whether differential attrition of children across treatment arms could drive our test score results. However, we find no differential rates of exit between

---

[36]For H schools, the test score effect for grade 4 children is 0.15sd (p=0.117) and for children in the second tested grade (either 3 or 5) is 0.188sd (p=0.033).

[37]As discussed previously, attrition in our data was 5% in the first year and 10% in the second year of the study with similar baseline characteristics of attriters across groups (Appendix Table B2). We have also examined attrition corrected estimates in each round and again find the treatment effects to be similar.

[38]There are no notable changes in the results on fixed and variable expenditures, posted and collected revenues, fees or closure. For enrollment the gains in H schools becomes statistically somewhat weaker, as does the comparison between the two treatments when addressing attrition. For test scores, while we always find that test scores increased in H schools, the difference between H and L^t schools is less significant (p-value of 0.138) when we address potential imbalance concerns.

[39]While this restriction eliminates concerns that new children directly impact our results, stayers may still be affected if newly enrolled children generate peer effects. Given how few new children join a given class, even using the higher end of peer effects estimates in the literature, barely impacts our estimates. Hoxby (2000) finds that "*a credibly exogenous change of 1 point in peers' reading scores raises a student's own score between 0.15 and 0.4 points, depending on the specification.*" Even at the very high parts of that estimate and taking the point estimates of the endline test score of newly enrolled children at H (relative to control) schools (0.186sd) at face value, this would imply that the impact on H schools for existing children from peer effects would have been 0.003sd (0.001 at the lower end of the estimates). This is significantly smaller than the 0.132sd effects we see and therefore even if we were to "net out" possible peer effects our impact on H schools would be relatively unaffected. In terms of whether (adverse) peer effects could be actually masking a positive impact on L^t schools this is even less likely since newly enrolled children in such schools, while not significantly different from existing children, have a slightly larger point estimate for their endline test score (0.039sd) and so could only generate (very small) positive peer effects.

control and H or L^t schools and we do not find any significant difference in baseline test scores of children who leave across control and H or L^t schools. We also undertook a formal bounding exercise where we "fill in" the endline test scores of students who left by drawing on the multi-year test score data collected on over 12,000 children in the LEAPS project.[40] Our simulations show that, even after accounting for leavers, we obtain a mean H test score impact of 0.14sd with a 95% confidence interval (CI) bounded between 0.12sd and 0.16sd. Similarly, the mean L^t score effect is -0.004sd, with a 95% CI between -0.033sd and 0.026sd. Together these checks and bounds demonstrate that our test score results are unlikely to be driven by changes in the composition of children across schools.

3.3 Channels

We explore further the factors that could explain the impact of the two treatments through a more detailed examination of how the funds were used. In doing so, we also discuss the extent to which differential (lower) school closures for L^t schools can be behind some of our findings.

3.3.1 Infrastructure

In our earlier results, we showed that both H and L^t schools increased their fixed expenditures, largely in the first year of the treatment. Table 7 now considers the specific investments by schools during this period. Column 1 examines total spending on infrastructure-related items (e.g. school furniture, fixtures, or up-gradation of classroom facilities from a semi-permanent to a permanent structures) and shows increases in both H and L^t schools. Columns 2, 3, and 4 show that the spending increase reflects, in part, greater spending on desks, chairs, and computers for H and L^t schools.[41] In contrast, columns 5 and 6 show that only H schools are more likely to report having a library and a sports facility, respectively, and the difference between H and L^t schools on these measures is significant with p-values of less than of 0.01. Finally, column 7 shows that H schools upgraded more classrooms than control schools. While we do not find a statistically significant effect for L^t schools, we cannot reject equality of coefficients between the two treatment approaches. Consistent with most schools choosing to front-load their investments at the beginning of the school year immediately after they received the grant, there are no further effects for these investments in year 2 (Appendix Table

---

[40]We fill in test scores for leavers by assigning them actual gains experienced by leavers from the LEAPS study child (test score) panel data. This is a comparable sample since it carried out the same test for children from overlapping study areas and age groups. We then calculate school level average test scores (using our observed stayer children and simulated leaver children) and run our canonical regression specification to provide treatment estimates. Running this simulation 1,000 times provides us with bounds.

[41]A standard desk accommodates 2 students implying that 12 additional students can be seated in H schools, and 18 students in L^t school; these numbers are similar to magnitude to the enrollment gains documented in Table 5.

B12).

3.3.2 Teachers

Table 8 examines the increases in annual variable expenditures. Since teacher salaries are 75% of these expenditures, column 1 starts by looking at the teacher wage bill. We find that while there is no change in the total wage bill for L^t schools, H schools spend an average of Rs.32,983 a year more over the two years after treatment. This represents a 14% increase relative to the baseline wage bill and is significantly higher relative to L^t schools (p=0.056). We now examine whether this wage bill increase stems from more teachers or higher teacher wages. While there is no significant increase in the number of teachers employed at a school for either H or L^t school (column 2), there is a significant increase in the number of *new* teachers in H schools (column 3). Column 4 show significant increases in teacher wages in H relative to control and L^t schools. This pay differential emerges both for newly hired teachers (column 5) and for existing teachers (column 6), and range from an 18% to 22% increase over baseline pay. If teacher pay reflects teacher quality, either through retention and recruitment of higher quality teachers or through pay incentives, the combination of higher teacher pay and new teachers are a potential channel for the observed test score increases in H schools.[42] Not only do these salary changes persist for both years for H schools, the point estimates suggest that the impacts are somewhat larger in the second year (Appendix Table B13).

3.3.3 School Closures

In light of the fact that L^t schools were 9 percentage points less likely to close (relative to 13.7% closure in the control group), we now assess the extent to which closures can explain the results we obtain among L^t schools. Moreover, while H schools are not less likely to close (relative to control schools), we also assess whether the results observed for H schools (including the differences with L^t schools) could reflect differences in the types of schools that closed across these two treatment arms.[43]

We consider two distinct approaches to assessing closures as a potential channel. In Appendix Table B15, Panel A, we impute outcomes for (H, L^u, and control) schools after they closed down, using the trends and covariates for open schools in the control group.[44] In Panel B, we instead increase the number of closures in L^t schools to match

---

[42]Bau et al. (2020) show that there is a link between pay and teacher value-added in the private sector in our context as well.

[43]Appendix Table B14 shows that control schools that close (relative to those that remain open) tend to be significantly smaller, younger, employ fewer teachers, but have better infrastructure and test scores. They also spend less on fixed and variable expenditures, though not significantly so. However, we find little evidence that treatment changes the nature of closure. Closed H or L^t schools do not differ much from closed schools in control villages, except that closed L^t schools are a bit smaller and have lower fees.

[44]We regress each outcome on a set of baseline covariates (enrollment, fees, fixed and variable ex-

the closures in the control and H arms.[45]

For most of our main outcome measures— expenditures, revenues, fees, and test scores— these methods of adjusting for differential closure rates do not affect the estimated effect size, although they sometimes worsen the precision of the estimates. For instance, in the case of test scores, the estimates for L^t schools increases from a small negative to a small positive value (both insignificant) leading to a higher p-value in the test of differences between the two arms (p=0.119 in Panel A and 0.192 in Panel B). The only main outcome that does change is enrollment: In Panel A, the enrollment coefficient for L^t schools is halved (10.6 children; p-value = 0.151). However, our alternative correction in Panel B shows little noticeable change in this impact. Our H enrollment effects are similar in magnitude to the main specification across both panels, but statistically weaker. Together, these results suggest that (only) our enrollment results, especially for L^t schools, are partly driven by school closure.

4 Discussion And Implications

In this section, we start by offering our preferred explanation for the results drawing in particularly on the potential differences between the two treatment arms. We then consider alternative explanations and end with a discussion of the potential welfare implications of our findings.

4.1 Interpreting H versus L^t differences

Our results suggest that the reaction of schools to the grants is different across the two arms: L^t schools invested primarily in increasing capacity with no change in test scores or fees. On the other hand, H schools raised test scores and fees, with a smaller (but not significantly different) increase in capacity. These different strategies are reflected in schools' choice of fixed and variable investments, with H schools more focused on teacher hiring and remuneration.

Our preferred explanation for this difference is that it arises from the nature of competition in the market and how the variation in financial saturation between the two treatment arms impacts the relative attractiveness of different investment strategies.

To fix intuition, suppose there are two (capital/capacity constrained) private schools in the village and schools can invest in expanding capacity and/or increasing quality.[46]

_____

penditures, test score, school age, number of teachers, and infrastructure index) and strata and round fixed effects for open schools in the control group. We then use the coefficients from this regression to predict outcomes for schools in the sample after they close down.

[45]We generate predicted probabilities of closure for the full sample by using baseline covariates to predict closure in the control group. We then "force" shutdown of a fraction of schools in the L^t group with high predicted probabilities of closure at baseline such that we eliminate differential closures in the sample.

[46]One can think of capacity investments as those that allow schools to retain or increase enrollment

23

Whether schools invest (more) in capacity or quality when they receive grants depends on the trade-off between increasing market share and risking price competition. We argue that this trade-off implies that investing in quality is more likely when all schools receive the grant.

The intuition is as follows: When only one school receives the grant, it can increase capacity without risking a price-war as long as it does not poach from the other school. It can do so because capacity constraints imply that there are children who would like to attend school, but cannot. However, when all schools receive the grant simultaneously, if they both try to invest in capacity, they are more likely to draw on the same pool of children and, therefore, increase the risk of price competition. Since price competition hurts profitability, schools can alleviate this risk through quality improvements, which both increase the size of the market (measured by the total surplus generated in the market) and allow them to vertically differentiate.

The model in Appendix C formalizes this intuition by introducing credit constrained firms and quality in the canonical Kreps and Scheinkman (1983) framework (henceforth KS) of capacity pre-commitment.[47] Schools in the model are willing to increase their capacities or qualities, but are credit constrained beyond their initial endowments of capacity and quality; the unconditional grants alleviate these constraints. Equilibrium follows from a duopoly game where schools first choose capacity and quality, and then prices. Therefore, even capacity and quality constrained schools can react to other schools' investments by altering their own prices. Corollary 1 in Appendix C shows that schools are more likely to invest in quality when all schools receive the grant.[48]

Interestingly, our main result— that if a school invests in quality when it is the only one receiving the grant, it will always do so when all schools receive the grant, but not the other way around— is remarkably robust to a number of plausible modifications that improve the model's fit to the education market. Through a series of exercises, Appendix C shows, for instance, that this result will continue to hold if schools have initially heterogeneous capacity and/or are horizontally differentiated. We can also allow for more flexible (variable/fixed) nature of investment costs and allow owners to be risk-averse and insurance, rather than credit, constrained without changing our main result.

It is important to note that while the model links quality investments in both treatment arms to schooling demand and cost parameters, we unfortunately do not have empirical counterparts to these parameters and are therefore unable to "test" the

---

but without being able to increase fees from existing students (e.g. additional desks, chairs etc.), presumably because the per capita infrastructure availability remains unchanged. In contrast, quality-enhancing investments are those that enable them to charge higher fees to (existing) students. These could include investments that raise test scores such as enhanced teaching, but could also include specialty infrastructure such as upgraded classrooms, a library, or a sports facility.

[47]KS (1983) develop a model of firm behavior under binding capacity commitments. In their model, the Cournot equilibrium is recovered as the solution to a Bertrand game with capacity constraints.

[48]Formally, we show that for any parameter values where an L^t school invests in quality, at least one H school will also invest in quality. On the other hand, there are parameter values where H schools will invest in quality, but L^t schools will not.

model. Instead, the role of the model is primarily to offer a plausible explanation of our results. Encouragingly, the model is consistent with a number of empirical patterns we observe in the experiment. In the case where H schools invest more in quality than L^t schools, the model predicts that we would also expect fees to be higher in H schools and enrollment to be higher in L^t schools. This is indeed what we find. We should also expect H schools to make more "quality-enhancing" investments and we do see significantly higher investments in variables expenditures stemming from a higher teaching wage bill, which is arguably an important factor behind raising quality in private schools.

The model assumes that schools are credit constrained, and our results are also consistent with this assumption. Since increased investments in response to the grant could alternatively reflect the lower (zero) cost of financing, Banerjee and Duflo (2012) suggest the following additional test: If firms are not credit constrained, they should always use the cheaper credit (i.e. the grant) to pay off more expensive loans. In Appendix Table B16, we examine data on borrowing for school and household accounts of school owner households. While there is limited borrowing for investing in the school, over 20% of school owner households do borrow, presumably for personal reasons. Yet, we find no statistically significant declines in borrowing at the school or household level as a result of our intervention, either when we look at the pooled treatment or when we examine the separate arms. This leads us to believe that the grants likely solved a problem of credit constraints.

An extension of the model also shows that small schools will be less likely to close in the L^t arm, which is again consistent with our empirical findings. Finally, the model also suggests that profits in L^t schools should be higher than H schools and this is again consistent with estimates we provide below in section 4.3, although the precision of these estimates is quite poor.

This still leaves open the question of whether there are alternate models that would also be consistent with the data, and we turn to this next.

4.2 Alternative explanations

We discuss two classes of alternate explanations, both of which are tied to the design of our experiment.

Village level resources: Given our design preference for school level comparisons, the grant amount was the same for all schools regardless of treatment arm. Therefore, the grant per capita in an L village is always lower than in an H village, holding constant village size. This raises the concern that the differences between the two treatment arms may be less about differential equilibrium response by schools and instead simply a reflection of greater total funds in H villages.

A specific illustration of these concerns is the higher wage bill for teacher in H

schools. Our preferred explanation is that the nature of financing and competition led H schools to make greater quality-enhancing investments, and wage increases reflect changes in their recruitment and retention of high quality teachers as well as incentives to existing teachers. An important question is whether the teacher wage differential could arise even if the extent of quality-enhancing investments in H and L^t schools were the same. Consider a specific alternate model. Suppose the grant leads schools to invest in a capital input, such as computers, which is complementary to teaching investments. As long as the incentive to invest in computers is higher for H compared to L^t schools, the explanation is isomorphic to ours. Alternatively, suppose that there is no difference in the incentive to invest in computers, but that greater demand for computers in H schools leads to a greater (derived) demand for teachers at the village level since more schools have received the grant. If the supply of teachers is inelastic, this will increase their wages, which is something that we observe in the data, but this can be attributed entirely to differential total resources at the village level, rather than the degree of saturation.

While plausible, this explanation is not consistent with our results on teacher hiring. Under this alternative explanation, the shadow price of investments in computers must be *lower* in L^t compared to H schools, since schools will rationally anticipate the increase in teachers wages in the H arm. Because the price of investing in computers is now higher in H schools, we should see *less* investment in this arm *and* a lower demand for teachers. However, as Table 7 shows, we do not find lower investments in inputs that are complimentary to teachers (such as computers and libraries) in H schools and, if anything, there is weak evidence that H schools demand more teachers than L^t schools.

At a broader level, these explanations arise from adjustments to the model that generate *asymmetric* parameterization of the profit function in each treatment arm. For example, if school owners have the ability to collectively affect the market size or input prices (e.g. higher competition among schools may raise teacher salaries), then the return or cost of an investment would be different in each treatment arm, which may meaningfully change our results.

To further investigate such general village level resource-based explanations, we can also use baseline variation in village size to additionally control for the per capita grant size in each village. If per capita grant size is an omitted variable that is correlated with treatment saturation and driving our results, we should find that the additional inclusion of this variable drives the *difference* in our treatment coefficients to zero. We therefore replicate our base specifications including per capita grant size as an additional control in Appendix Table B17. We find that the qualitative pattern of our differential results between H and L^t schools does not change: L^t schools see higher enrollment on average, while H schools experience higher fees, test scores, and variable expenditures on average. While we lose some precision in the H arm, we cannot reject that these coefficients are identical to our base specification.

Additional Intervention Features: A second class of explanation concerns the specific additional features of our intervention. Like all financial interventions, ours is also an intervention that bundles money 'plus' additional requirements. Those additional requirements were designed to be less onerous in our intervention, but could still have driven part of our intervention impacts. Notably, we required (i) school owners to open a one-time use bank account with our banking partner in order to receive funds; and (ii) every treated school to submit an investment plan before any disbursement could take place. We consider each in turn.

In terms of the effects from opening a (one-time use) bank account, 73% of school owner households already had bank accounts at baseline and this fraction is balanced across treatment arms. Further, in Appendix Table B18, we use an interaction between treatment and baseline bank account availability to check whether our pattern of treatment effects is driven by previously unbanked households. We detect no statistically significant differential impact by baseline bank account status suggesting that this alternative explanation is unlikely.

Although the experimental literature on business plans seldom finds significant effects, and our process was designed to be minimally invasive and effectively non-binding, it could be that the plan forced school owners to consider new investments, or, that the act of submission itself notionally committed school owners to a course of action (McKenzie, 2017).[49] We examined this possibility by looking at three separate sources of (proposed and actual) school investments: (a) pre-treatment proposed investment questions from the baseline survey; (b) investment plan data; and (c) investments as reported in the follow-up surveys.

As the correlation in proposed investments between (a) and (b) is high, this suggests that simply asking schools about investment plans is unlikely to explain our treatment effects. This is because (a) is asked of all schools and (a) provides similar information to (b). So, if asking schools to formulate an investment plan is sufficient to induce investment, we should not see a differential impact between control and treatment arms. In addition, the investment plan itself was not particularly binding as the correlation between investments in (b) and (c) is low. Schools do not seem to have treated the business plan as a commitment device; instead, owners appeared to have finalized school investments after disbursement. Thus, it is unlikely that the submission of investment plans induced the kinds of large effects we document here or that these plans induced differential effects between the treatment arms.[50]

---

[49]Schools could propose investments with private value as long as they could argue it benefited the school or spend the money on previously planned investments, thereby effectively using the grant for personal uses. They could also propose changes to their plans at any time during the disbursement.

[50]While our checks show that these additional features are unlikely to have first order effects, there *is* a class of explanations where individuals are indifferent between various choices and small design changes can be quite instrumental. For instance, if school owners have multiple businesses, even with credit constraints, the marginal value of an additional dollar in credit will be equalized across these multiple businesses (and their home expenditures) and therefore small "nudges" may induce them to invest more in one business (the school) versus another. Our data are insufficient to fully rule out these cases, though it is unclear how these cases would generate *differential* effects across H and L^t arms.

4.3 Implications for Welfare

A complete welfare analysis is beyond the scope of the paper since we cannot adequately estimate the demand parameters required. Nevertheless, it is instructive to consider how the two treatments may impact welfare by considering gains for the four relevant groups of affected participants— school owners, teachers, parents, and children. In order to facilitate such a comparison consider two policies of unconditional grants given as either (a) PKR 50K to *one* school each in *three* villages (L policy), or (b) PKR 50K to *each* of the three schools in *one* village (H policy) (see Appendix A for details).

For schools owners, our previous analysis on revenues and expenditures reveals that the L policy will lead to substantially higher gains (Appendix Table B19 shows how this translates into school profits). In contrast, as our previous wage results show, gains for teachers are higher in the H policy.

For parents, estimating consumer surplus is challenging in the absence of household choice data that we would need to estimate parental demand for quality. While we can provide estimates under somewhat restrictive assumptions[51], they are quite sensitive to small changes in the estimated elasticity of demand. Back-of-the-envelope calculations based on these assumptions are presented in Appendix A and these suggest that the two treatment arms increase consumer surplus by similar amounts.

Finally, we consider overall test score gains for children. We find total gains of 115sd and (an upper bound of) 60-70sd under the H and L policies, respectively.[52] The increase for the L arm comes from the fact that children switched schools, and although we do not have baseline scores for these children, we optimistically estimate gains of 0.33sd for switchers based on our longitudinal study of child test score gains switching from a public to a private school (Andrabi et al., 2011).[53]

While admittedly speculative, this exercise highlights a tension between the two approaches. Relative to the L policy, the H policy transfers gains from school owners to school teachers. The gains to parents are of similar magnitudes across the two arms, but the H policy leads to larger increases in test scores. In general, whether the H or L policy is preferred would depend on how we weight gains to the different participants. If we believe that educational interventions should primarily focus on learning with limited weight on school owner profits, then the H approach seems preferable.[54]

---

[51] We need to assume that (i) the demand curve can be approximated as linear and (ii) regardless of quality, demand at zero price in the village is fixed at an upper-bound.

[52] These effect sizes compare favorably to gains observed in other successful educational experiments. The gain of 7.8sd for every $100 invested in H and of 4.1sd for L^t schools compared to the median test score gain of highly effective interventions at 2.3sd per $100 (JPAL, 2017). As Dhaliwal et al. (2013) discuss, our estimates assume that gains across students are perfectly substitutable and returns are linear.

[53] Since we see, at most, gains of 30 children in a village, a small fraction of the enrollment in public schools, we believe it is unlikely that public schools responded to the intervention by changing their own quality. Nevertheless, a limitation of this paper is that we are unable to examine whether the interventions had any (indirect) effects on public school quality.

[54] A related question that may have additional welfare implications is whether our treatments differentially impacted children based on their socioeconomic status and gender. Appendix Table B20

5 Conclusion

Unconditional grants to private schools generate large returns relative to the market lending rate and lead to higher enrollment. In addition, when all schools in a village receive grants, test scores and school fees also increase, as do wages for teachers. This suggests that designing interventions that take advantage of the underlying competitive structure of markets— such as financing that induces competitive actions— can be an important element of education policy. If so, this would reduce the need for additional supervision and monitoring, such as conditioning the grant on test scores directly, and allow for quality improvements even in scaled-up programs.

We make three further comments.

First, our approach, which selected a school *at random*, should be a lower-bound to what lenders can obtain in the market where (ex-ante) better schools can be selected through credit-screening. Yet, this market remains starved of finance, perhaps because lenders perceive the low-cost private schooling market as risky. We have now actively partnered with a a micro-finance provider and our initial results show high take-up of loans with low default rates, providing further impetus for designing and offering appropriate financial products in this space.

Second, even if such lending became more prevalent, left to the market, a lender would favor a low-saturation approach as long as the (village level) fixed costs of financing are not too large. If high-saturation lending is instead more socially desirable, loan-loss guarantees for lenders could encourage greater market saturation.[55] Interestingly, this policy also differs from standard "priority sector" lending policies in that the subsidy is not based on a sectoral preference but rather on the geographical "density/saturation" of the financial offering by a lender. Should our results from this paper also extend to lending, the socioeconomic impact of offering such loans can be large, benefiting both educational entrepreneurs and students.

Third, our results question the relative merits of addressing market needs in a way that allows *select* providers to flourish and grow rapidly, a strategy that, while used frequently, curtails rather than enhances competition (Smith and Baker, 2017). The relative success of either approach ultimately depends on the nature of market competition, market demand, and the production function facing providers. In our

---

shows that we do not find any difference in the average household assets of children attending school across the treatment arms or the fraction of students who are female. However, in both treatment arms there are more new children in the school who are particularly disadvantaged (those with family assets in the bottom 25th percentile). There is also suggestive evidence that more new children were girls, although these estimates are not significant. On balance therefore, there is an indication that while the treatments did not differ in this aspect, they may both have generated further benefits by helping enroll poorer children and girls.

[55]Recall that the L treatment arm reduced school closures relative to the H arm. Appendix A shows that a loan loss guarantee of Rs.15,675 over a two year period for a total loan value of Rs.150K would make banks indifferent between the two approaches. This calculation makes the conservative assumption that schools that shutdown will not pay back any of their loan. In practice, based on ongoing work, we find that schools continue to often repay after closure.

experiment, greater financial saturation increased quality and, possibly, welfare, more than an approach favoring a single school. There has been less consideration of market approaches such as ours in the field of education and educational finance, and our results can contribute to an overdue debate on this issue.

References

Abdulkadiroğlu, A., Angrist, J. D., Hull, P. D., and Pathak, P. A. (2016). Charters without lotteries: Testing takeovers in new orleans and boston. *American Economic Review*, 106(7):1878–1920.

Abdulkadiroğlu, A., Pathak, P. A., and Roth, A. E. (2009). Strategy-proofness versus efficiency in matching with indifferences: Redesigning the nyc high school match. *American Economic Review*, 99(5):1954–78.

Ajayi, K. F. (2014). Does school quality improve student performance? new evidence from ghana. *Unpublished working paper*.

Andrabi, T., Das, J., and Khwaja, A. I. (2008). A Dime a Day: The Possibilities and Limits of Private Schooling in Pakistan. *Comparative Education Review*, 52(3):329–355.

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

Andrabi, T., Das, J., and Khwaja, A. I. (2015). Delivering education a pragmatic framework for improving education in low-income countries. (May).

Andrabi, T., Das, J., and Khwaja, A. I. (2017). Report cards: The impact of providing school and child test scores on educational markets. *American Economic Review*, 107(6):1535–63.

Andrabi, T., Das, J., Khwaja, A. I., and Karachiwalla, N. (2018). The equilibrium effects of grants to public schools. *Working Paper*.

Andrabi, T., Das, J., Khwaja, A. I., Vishwanath, T., and Zajonc, T. (2009). Learning and Educational Achievements in Punjab Schools (LEAPS): Insights to inform the education policy debate.

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

Angrist, J. D., Pathak, P. A., and Walters, C. R. (2013). Explaining charter school effectiveness. *American Economic Journal: Applied Economics*, 5(4):1–27.

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

Banerjee, A. and Duflo, E. (2012). Do Firms Want to Borrow More: Testing Credit Constraints Using a Targeted Lending Program. *Review of Economic Studies*, page Forthcoming.

Barrera-Osorio, F., Blakeslee, D. S., Hoover, M., Linden, L., Raju, D., and Ryan, S. P. (2017). Delivering education to the underserved through a public-private partnership program in Pakistan. *Policy Research working paper; No. WPS 8177; Impact Evaluation series.*

Bau, N., Das, J., et al. (2020). Teacher value added in a low-income country. *American*

Economic Journal: Economic Policy, 12(1):62–96.

Baum, D., Lewis, L., and Patrinos, H. (2013). Engaging the private sector: What policies matter? a framework paper. *SABER Working Paper Series*.

Beck, T. (2007). Financing Constraints of SMEs in Developing Countries : Evidence , Determinants and Solutions. *Financing Innovation-Oriented Businesses to Promote Entrepreneurship*, (April):1–35.

Bruhn, M. and McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, 1(4):200–232.

Das, J., Dercon, S., Habyarimana, J., Krishnan, P., Muralidharan, K., and Sundararaman, V. (2013). School inputs, household substitution, and test scores. *American Economic Journal: Applied Economics*, 5(2):29–57.

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

de Mel, S., McKenzie, D., and Woodruff, C. (2008). Returns to capital in microenterprises: Evidence from a field experiment. *The Quarterly Journal of Economics*, 123(4):1329–1372.

de Mel, S., McKenzie, D., and Woodruff, C. (2012). One-time transfers of cash or capital have long-lasting effects on microenterprises in sri lanka. *Science*, 335(6071):962–966.

Dean, J. and Jayachandran, S. (2019). Attending kindergarten improves cognitive but not socio-emotional development in India. *Working Paper*.

Dhaliwal, I., Duflo, E., Glennerster, R., and Tulloch, C. (2013). Comparative cost-effectiveness analysis to inform policy in developing countries: a general framework with applications for education. *Education Policy in Developing Countries*, pages 285–338.

Epple, D., Romano, R. E., and Urquiola, M. (2015). School vouchers: A survey of the economics literature. Technical report, National Bureau of Economic Research.

Evans, D. and Popova, A. (2015). What really works to improve learning in developing countries? an analysis of divergent findings in systematic reviews.

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

Hoxby, C. M., Murarka, S., and Kang, J. (2009). How new york city's charter schools affect achievement. *Cambridge, MA: New York City Charter Schools Evaluation Project*, pages 1–85.

Hoxby, C. M. and Rockoff, J. E. (2004). *The impact of charter schools on student achievement*. Department of Economics, Harvard University Cambridge, MA.

Hsieh, C.-T. and Urquiola, M. (2006). The effects of generalized school choice on achievement and stratification: Evidence from chile's voucher program. *Journal of public Economics*, 90(8):1477–1503.

JPAL (2017). Increasing test score performance. *Retrieved from: https://www.povertyactionlab.org/policy-lessons/education/increasing-test-score-*

performance.

Kapor, A., Neilson, C. A., Zimmerman, S. D., et al. (2017). Heterogeneous beliefs and school choice mechanisms. Technical report.

Kreps, D. M. and Scheinkman, J. a. (1983). Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes. *The Bell Journal of Economics*, 14(2):326–337.

Mbiti, I., Muralidharan, K., Romero, M., Schipper, Y., Manda, C., and Rajani, R. (2019). Inputs, incentives, and complementarities in education: Experimental evidence from tanzania. *The Quarterly Journal of Economics*, 134(3):1627–1673.

McEwan, P. J. (2015). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of Educational Research*, 85(3):353–394.

McKenzie, D. (2017). Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition. *American Economic Review*, 107(8):2278–2307.

Muralidharan, K., Singh, A., and Ganimian, A. J. (2016). Disrupting education? experimental evidence on technology-aided instruction in india. Technical report, National Bureau of Economic Research.

Muralidharan, K., Sundararaman, V., and Sundararaman, K. M. V. (2015). The Aggregate Effect of School Choice: Evidence from a two-stage experiment in India. *The Quarterly Journal of Economics*, 130(3):1011–1066.

NEC (2005). National education census. Technical report, Pakistan Bureau of Statistics.

Neilson, C. (2017). Targeted vouchers, competition among schools, and the academic achievement of poor students. *Working Paper*.

Nguyen, Q. and Raju, D. (2014). Private school participation in pakistan.

Romero, M., Sandefur, J., and Sandholtz, W. A. (2017). Can Outsourcing Improve Liberia's Schools? *Working Paper 462*.

Rotemberg, M. (2019). Equilibrium effects of firm subsidies. *American Economic Review*, 109(10):3475–3513.

Smith, W. and Baker, T. (2017). From free to fee: Are for-profit, fee-charging private schools the solution for the world's poor.

Udry, C. and Anagol, S. (2006). The return to capital in ghana. *American Economic Review*, 96(2):388–393.

Table 1: Baseline Summary Statistics

	Mean	5th pctl	25th pctl	Median	75th pctl	95th pctl	SD	N
Panel A: Village level variables								
Private enrollment	523.5	149.0	281.0	415.5	637.0	1231.0	378.1	266
Number of private schools	3.3	2.0	2.0	3.0	4.0	7.0	1.6	266
Number of public schools	2.5	1.0	2.0	2.0	3.0	5.0	1.0	266
Panel B: School level variables								
Enrollment	163.64	45.00	88.00	140.00	205.00	353.00	115.98	851
Fees, monthly (PKR)	238.13	81.25	150.00	201.25	275.00	502.50	165.94	851
Revenues, annual (PKR)	482,173	59,316	163,200	317,820	532,800	1,411,860	653,986	850
Variable expenditures, annual (PKR)	304,644	46,800	112,800	194,400	326,400	948,000	371,533	848
Fixed expenditures, annual (PKR)	78,861	0	9,700	33,000	84,000	326,000	136,928	837
School age, in years	8.28	0.00	3.00	7.00	12.00	19.00	6.67	852
Number of teachers	8.16	3.00	5.00	7.00	10.00	17.00	4.85	851
Average teacher salary, monthly (PKR)	2,562.81	1,000.00	1,500.00	2,000.00	2,928.50	5,250.00	3,139.52	768
Number of children in tested grade	13.13	1.00	5.00	10.00	18.00	34.50	11.68	420
Number of tested children	11.74	1.00	4.00	9.00	16.00	31.50	10.56	420
Average test score	-0.21	-1.24	-0.59	-0.22	0.15	0.84	0.64	401

Notes:

a) This table displays summary statistics for the 266 villages (Panel A) and the 855 private schools (Panel B) in our sample.

b) These baseline data come from two sources: school surveys administered to the full sample (855 schools), and child tests administered to half of the sample (420 schools). Any missing data are due to school refusals, child absences, or zero enrollment in tested grades.

Table 2: Pooled Treatment Effects of a Grant

	Expenditures (annual)		Revenues (annual)					
	(1) Fixed	(2) Variable	(3) Posted	(4) Collected	(5) Enrollment	(6) Fees	(7) Test Score	(8) Closure†
Panel A: Year 1 Effects								
Treatment (T)	28,075.997	162.418	101,188.864	71,999.607	15.358	7.535	0.016	-0.040
	(8,079.319)	(18,788.233)	(39,988.114)	(39,286.198)	(4.697)	(5.773)	(0.068)	(0.016)
Baseline	0.165	0.901	0.955	0.763	0.778	0.833	0.345	
	(0.041)	(0.092)	(0.061)	(0.065)	(0.044)	(0.040)	(0.119)	
R-Squared	0.11	0.72	0.69	0.59	0.69	0.71	0.16	0.03
N Schools (Rounds)	767 (1)	788 (1)	825 (2)	822 (2)	827 (3)	796 (2)	725 (1)	855 (1)
Test pval (T=0)	0.001	0.993	0.012	0.068	0.001	0.193	0.818	0.014
Panel B: Year 2 Effects								
Treatment (T)	324.357	14,203.914	125,273.140	85,262.972	19.253	6.004		-0.044
	(7,298.252)	(19,514.688)	(52,634.747)	(45,399.519)	(6.699)	(6.688)		(0.022)
Baseline	0.041	0.872	0.953	0.758	0.714	0.831		
	(0.027)	(0.088)	(0.112)	(0.108)	(0.056)	(0.041)		
R-Squared	0.05	0.64	0.58	0.53	0.53	0.72		0.04
N Schools (Rounds)	676 (1)	682 (1)	821 (1)	825 (2)	826 (2)	749 (1)		855 (1)
Test pval (T=0)	0.965	0.467	0.018	0.061	0.004	0.370		0.043
Panel C: Average Effects Across Years 1 and 2								
Treatment (T)	14,900.355	6,996.023	109,083.133	78,615.846	16.930	7.038	0.016	-0.044
	(5,494.335)	(16,219.486)	(42,345.190)	(41,386.355)	(5.279)	(5.528)	(0.068)	(0.022)
Baseline	0.104	0.889	0.955	0.760	0.752	0.832	0.345	
	(0.025)	(0.068)	(0.068)	(0.084)	(0.047)	(0.038)	(0.119)	
R-Squared	0.10	0.69	0.65	0.56	0.62	0.71	0.16	0.04
N Schools (Rounds)	786 (2)	797 (2)	832 (3)	831 (4)	836 (5)	800 (3)	725 (1)	855 (1)
Test pval (T=0)	0.007	0.667	0.011	0.059	0.002	0.204	0.818	0.043
Depvar Mean	78,860.866	304,644.226	482,172.565	370,380.473	163.638	238.132	-0.210	0.137

Notes:
a) This table shows the pooled treatment effects of grants. Treatment refers to any H or L^t schools offered a grant; the comparison group includes L^u and control schools. The dependent variables are as follows: annual fixed expenditures, which include annual spending on infrastructure and educational supplies or materials (column 1); annual variable expenditures, which include expenses incurred on a monthly basis such as teaching and staff salaries, rent, or utilities (column 2); posted revenues, defined as the sum of revenues expected from each grade based on enrollment and posted fees (column 3); collected revenues, defined as revenues actually collected from students across all grades (column 4); enrollment (column 5); average monthly fees charged to students (column 6); average test score for all children in a given school across English, Math and Urdu (column 7); and closure (column 8). Panel A shows year 1 effects; Panel B shows year 2 effects; and Panel C looks at average effects across years 1 and 2. Data are pooled across rounds, except for column 8, as indicated by the † symbol, where closure is coded based on the last follow-up round in the year(s) under consideration. For columns 1-2 and 6-7, if a school closes down, the variable is coded is missing, and for columns 3-5, the variable is coded as 0. Closure is never missing in the data.
b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata and round fixed effects. Standard errors are clustered at the village level. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The number of schools may vary across columns due to attrition or because the variable was not collected in a given round. The mean of the dependent variable is its baseline value, or the follow-up control mean wherever the baseline value is unavailable as for columns 4 and 8. We also control for baseline value of dependent variable, wherever available; in the case of column 4, we just use the baseline posted revenues as a control.
c) For each panel, we show the p-values from the test of a zero average impact of the treatment (T=0).
† Data is based on the last follow-up round in the year(s) under consideration.

Table 3: Fixed and Variable Expenditures (Average Effects across Years 1 and 2)

	Fixed expenditures (annual)			Variable expenditures (annual)		
	(1) Raw	(2) Top Coded 1%	(3) Trim Top 1%	(4) Raw	(5) Top Coded 1%	(6) Trim Top 1%
High	20,238.1	19,728.5	17,838.4	37,741.0	35,840.6	30,321.9
	(6,046.2)	(5,873.8)	(5,325.7)	(22,719.1)	(22,203.9)	(19,767.9)
Low Treated	16,346.4	16,620.3	20,284.2	-13,592.6	-14,456.1	-18,716.3
	(7,310.6)	(7,248.4)	(6,911.8)	(20,662.2)	(18,586.5)	(17,062.5)
Low Untreated	4,598.5	2,843.1	1,667.0	-3,636.1	-402.5	1,353.9
	(6,797.2)	(6,381.5)	(5,531.7)	(16,496.6)	(16,208.7)	(15,469.4)
Baseline	0.1	0.1	0.1	0.9	0.9	0.8
	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)
R-Squared	0.10	0.11	0.12	0.69	0.67	0.58
N Schools (Rounds)	786 (2)	786 (2)	778 (2)	797 (2)	797 (2)	789 (2)
Depvar Mean	78,860.9	77,876.6	71,390.5	304,644.2	299,322.1	282,744.2
Test pval (H=0)	0.001	0.001	0.001	0.098	0.108	0.126
Test pval (L^t=0)	0.026	0.023	0.004	0.511	0.437	0.274
Test pval (L^t=H)	0.588	0.657	0.719	0.048	0.032	0.016

Notes:
a) This table looks at the average impacts on fixed and variable expenditures over two years. Fixed expenditures are measured on an annual basis and include spending on infrastructure or educational materials and supplies; variable expenditures include expenses incurred on a monthly basis—teaching and non-teaching staff salaries, utilities and rent— and we annualize the variable for ease of comparison. Data on expenditures are collected twice, once in each of the two years after treatment. Columns 1-3 focus on fixed expenditures and columns 4-6 on variable expenditures. Top coding of the data assigns the value at the 99th percentile to the top 1% of data (columns 2 and 5). Trimming top 1% of the data assigns a missing value to data above the 99th percentile (columns 3 and 6). Both the top coding and trimming procedures are applied to each round of the data separately. Variables are coded as missing once a school closes down.
b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata and round fixed effects. Standard errors are clustered at the village level. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The number of schools may vary across columns due to attrition. The mean of the dependent variable is always its baseline value across all columns.
c) The bottom panel shows p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.

Table 4: School Revenues (Average Effects across Years 1 and 2)

	Posted Revenues (annual)			Collected Revenues (annual)		
	(1) Raw	(2) Top Coded 1%	(3) Trim Top 1%	(4) Raw	(5) Top Coded 1%	(6) Trim Top 1%
High	65,812.8	60,053.4	57,259.4	52,695.7	55,625.2	42,653.4
	(42,388.5)	(31,223.5)	(26,439.6)	(43,173.6)	(29,061.8)	(23,244.2)
Low Treated	127,987.4	111,926.6	99,048.6	94,545.4	83,396.4	64,568.6
	(58,594.1)	(47,712.2)	(44,540.1)	(55,546.3)	(39,089.3)	(34,813.4)
Low Untreated	-6,598.1	-8,214.0	3,944.1	996.7	1,323.7	8,632.4
	(33,001.7)	(28,147.7)	(22,652.8)	(30,476.3)	(26,521.4)	(20,592.0)
Baseline	1.0	1.0	0.9	0.8	0.9	0.7
	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
R-Squared	0.65	0.65	0.58	0.56	0.62	0.53
N Schools (Rounds)	832 (3)	832 (3)	820 (3)	831 (4)	831 (4)	820 (4)
Depvar Mean	482,172.6	463,848.7	434,390.0	370,615.3	362,731.8	332,030.6
Test pval (H=0)	0.122	0.056	0.031	0.223	0.057	0.068
Test pval (L^t=0)	0.030	0.020	0.027	0.090	0.034	0.065
Test pval (L^t=H)	0.352	0.317	0.368	0.526	0.530	0.548

Notes:
a) This table examines the average impacts on annual school revenues over two years. Columns 1-3 consider posted revenues, defined as the sum of revenues expected from each grade based on enrollment and posted fees. Columns 4-6 consider collected revenues, defined as revenues actually collected from students at the school. Top coding of the data assigns the value at the 99th percentile to the top 1% of data (columns 2 and 5). Trimming top 1% of the data assigns a missing value to data above the 99th percentile (columns 3 and 6). Both the top coding and trimming procedures are applied to each round of the data separately. Across all columns, schools are coded to have zero revenues once they close down.

b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata and round fixed effects. Standard errors are clustered at the village level. Across all columns, we use baseline posted revenues as the baseline control as we did not measure collected revenues at baseline. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The number of schools may vary across columns due to attrition or because the variable was not collected in a given round. The mean of the dependent variable is its baseline value for the posted revenues and the follow-up control mean for the collected revenues.

c) The bottom panel shows p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.

Table 5: School Enrollment and Fees (Average Effects across Years 1 and 2)

	(1) Enrollment	(2) Posted Fees	(3) Collected Fees	(4) Closure†
High	9.01	18.83	20.98	-0.02
	(6.04)	(7.88)	(9.50)	(0.03)
Low Treated	21.80	0.51	2.61	-0.09
	(7.73)	(7.48)	(10.13)	(0.03)
Low Untreated	0.31	-0.00	10.37	-0.03
	(5.51)	(6.49)	(7.22)	(0.03)
Baseline	0.75	0.83	0.64	
	(0.05)	(0.04)	(0.04)	
R-Squared	0.62	0.72	0.53	0.05
N Schools (Rounds)	836 (5)	800 (3)	782 (4)	855 (1)
Depvar Mean	163.638	238.132	238.132	0.137
Test pval (H=0)	0.137	0.018	0.028	0.597
Test pval (L^t=0)	0.005	0.946	0.797	0.007
Test pval (L^t=H)	0.101	0.021	0.099	0.045

Notes:
a) This table examines average impacts on enrollment and tuition fees over two years. The dependent variables are as follows: enrollment (column 1); posted tuition fees, averaged across all grades taught at the school (column 2); collected tuition fees, generated by dividing collected revenues by enrollment in the given round (column 3); and closure (column 4). Whereas columns 1-3 pool data over the two years, closure is measured at the end of year 2, as indicated by †. Once a school closes down, enrollment is coded as 0 and posted and collected fees are coded as missing.
b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata and round fixed effects. Standard errors are clustered at the village level. Columns 1-2 use the baseline of the dependent variable as a control, and column 3 uses posted fees as the baseline control since collected revenues are not measured at baseline. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The number of schools may vary across columns due to attrition or because the variable was not collected in a given round. The mean of the dependent variable is its baseline value, except for columns 3 and 4 where the follow-up control mean is provided since the baseline is unavailable.
c) The bottom panel shows p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.
† Data is based on the last follow-up round in year 2.

Table 6: School Test Scores (Year 1)

	(1) Average	(2) Math	(3) English	(4) Urdu
High	0.153	0.157	0.186	0.113
	(0.09)	(0.09)	(0.09)	(0.08)
Low Treated	-0.027	-0.074	0.084	-0.079
	(0.10)	(0.11)	(0.11)	(0.11)
Low Untreated	0.033	0.034	0.061	0.009
	(0.07)	(0.08)	(0.08)	(0.07)
Baseline	0.357	0.274	0.429	0.249
	(0.12)	(0.11)	(0.08)	(0.12)
R-Squared	0.16	0.18	0.14	0.13
N Schools (Rounds)	725 (1)	725 (1)	725 (1)	725 (1)
Depvar Mean	-0.21	-0.21	-0.18	-0.24
Test pval (H=0)	0.074	0.082	0.049	0.175
Test pval (L^t=0)	0.786	0.499	0.428	0.454
Test pval (L^t=H)	0.073	0.034	0.334	0.071

Notes:

a) This table examines impacts on school test scores. Column 1 shows the average score (across all subjects) for all children tested at a school. Columns 2-4 show subject-specific school test scores, constructed by averaging across children for a given subject. Test scores are collected only in the first year after treatment. Variables are missing for closed schools.

b) Regressions are weighted to adjust for sampling and include strata fixed effects. Standard errors are clustered at the village level. We include a dummy variable for the untested sample at baseline across all columns and replace the baseline test score with a constant. Since the choice of the testing sample at baseline was random, this procedure allows us to control for baseline test scores wherever available. The number of schools is lower than the full sample due to: attrition; closure; and zero enrollment in the tested grades. The mean of dependent variable is given for those tested at random at baseline.

c) The bottom panel shows p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.

Table 7: School Infrastructure (Year 1)

	Infrastructure Spending	Number purchased		Facility present (Y/N)			Upgradation
	(1) Amount	(2) Desks	(3) Chairs	(4) Computers	(5) Library	(6) Sports	(7) # Rooms
High	25,591.25	6.09	3.70	0.21	0.11	0.10	0.64
	(9,049.23)	(1.69)	(1.42)	(0.05)	(0.04)	(0.04)	(0.26)
Low Treated	17,729.27	8.66	5.95	0.16	-0.03	-0.03	0.30
	(8,809.38)	(2.47)	(2.78)	(0.06)	(0.05)	(0.04)	(0.40)
Low Untreated	-2,201.51	1.33	0.83	0.05	-0.03	0.02	0.12
	(8,644.48)	(1.45)	(1.22)	(0.05)	(0.04)	(0.04)	(0.27)
Baseline	0.09	0.10	0.13	0.26	0.33	0.23	0.71
	(0.03)	(0.05)	(0.08)	(0.04)	(0.05)	(0.05)	(0.06)
R-squared	0.06	0.09	0.09	0.19	0.21	0.12	0.58
N Schools (Rounds)	771 (1)	781 (1)	782 (1)	793 (1)	793 (1)	793 (1)	793 (1)
Depvar Mean	57,258.483	14.592	10.923	0.394	0.351	0.195	6.363
Test pval (H=0)	0.005	0.000	0.010	0.000	0.006	0.020	0.016
Test pval (L^t=0)	0.045	0.001	0.033	0.009	0.564	0.442	0.465
Test pval (L^t=H)	0.405	0.349	0.473	0.477	0.008	0.007	0.428

Notes:
a) This table examines outcomes relating to school infrastructure in year 1. Column 1 reports annual infrastructure expenditures, which comprises the largest component of annual fixed expenditures and includes spending on furniture, fixtures, or upgradation of classroom facilities. Columns 2-3 consider the number of new desks and chairs purchased in the last year; columns 4-6 are dummy variables for the presence of particular school facilities; and column 7 measures the number of rooms upgraded from temporary to permanent or semi-permanent classrooms. Variables are coded as missing once a school closes down.

b) Regressions are weighted to adjust for sampling and include strata fixed effects. Standard errors are clustered at the village level. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The number of schools may vary across columns due to attrition. The mean of the dependent variable is its baseline value.

c) The bottom panel shows p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.

Table 8: Teacher Compensation and Composition (Average Effects across Years 1 and 2)

	School level			Teacher level salaries (monthly)		
	(1) Wage bill (Annual)	(2) Teachers (number)	(3) New Teachers (number)	(4) All	(5) New	(6) Existing
High	32,982.64 (18,134.01)	0.28 (0.29)	0.42 (0.19)	519.52 (257.94)	580.05 (265.80)	492.01 (284.29)
Low Treated	-9,867.52 (18,310.31)	-0.09 (0.32)	0.06 (0.25)	-175.63 (273.11)	-89.45 (406.49)	-223.10 (246.45)
Low Untreated	806.86 (13,285.63)	0.25 (0.27)	0.24 (0.19)	194.48 (202.53)	89.47 (236.07)	253.39 (201.69)
Baseline	0.85 (0.08)	0.74 (0.05)				
R-Squared	0.63	0.53	0.23	0.20	0.23	0.20
N Schools (Rounds)	797 (2)	802 (2)	802 (2)	802 (2)	723 (2)	793 (2)
N Teachers				11725	3903	7818
Depvar Mean	233,893.9	6.7	2.4	2,676.6	2,665.5	2,681.9
Test pval (H=0)	0.070	0.337	0.031	0.045	0.030	0.085
Test pval (L^t=0)	0.590	0.766	0.797	0.521	0.826	0.366
Test pval (L^t=H)	0.056	0.274	0.159	0.039	0.135	0.037

Notes:

a) This table looks at the average impacts on teacher compensation and composition over two years. The dependent variables are as follows: (annualized) wage bill, the largest component of variable expenditures (column 1); total number of teachers (column 2); number of new teachers (column 3); and monthly salaries for all, new, and existing teachers (columns 4-6). Column 1 data come from school surveys, whereas data for columns 2-6 come from teacher rosters collected at each school. In columns 2-3, we collapse data from these rosters at the school level to understand changes in teacher composition at the school level; columns 4-6 are measured at the teacher level. For columns 5 and 6, whether a teacher is new or existing is determined by their start date at the school relative to baseline. Variables are coded as missing once a school closes down.

b) Regressions are weighted to adjust for sampling and tracking where necessary and include strata and round fixed effects. Standard errors are clustered at the village level. For each regression, we report the unique number of schools and the number of rounds of data in parantheses. The mean of the dependent variable is its baseline value in columns 1-2 and the follow-up control mean for columns 3-6 where the baseline is unavailable. The number of schools may vary across columns due to attrition.

c) The bottom panel shows p-values from tests that either ask whether we can reject a zero average impact for high (H=0) and low treated (L^t=0) schools, or whether we can reject equality of coefficients between high and low treated (L^t=H) schools.