Delivering Education to the Underserved through a Public-Private Partnership Program in Pakistan

Felipe Barrera-Osorio[*] David S Blakeslee[†] Matthew Hoover[‡]
Leigh Linden[§] Dhushyanth Raju[¶] Stephen P. Ryan[‖]

Abstract: Governments are increasingly using the private sector to improve the delivery of public education. We contribute to this literature by evaluating a program that randomly assigned newly-created private schools to underserved villages in Pakistan. Private operators were given a per-student subsidy to provide tuition-free primary education in 100 villages, with half of them receiving a higher subsidy for female students. The program increased enrollment by 30 percentage points, and test scores by 0.63 standard deviations. The effects were similar across genders, and across the two subsidy treatments. Program schools were of higher quality than nearby government schools, despite their far lower unit costs. Entrepreneurs exercised wide latitude over school inputs, allowing us to investigate how private providers respond to household demand for education. A structural model for the supply and demand of school inputs indicates that program schools selected inputs similar to those of a social planner who internalizes all the educational benefits to society.

[*]Harvard Graduate School of Education

[†]New York University - Abu Dhabi

[‡]Gallup

[§]University of Texas at Austin, BREAD, J-PAL, IPA, IZA, NBER

[¶]The World Bank

[‖]Washington University in St. Louis, CESifo, NBER

This study is dedicated to the respectful memory of the late Anita Ghulam Ali, former Managing Director of the Sindh Education Foundation (SEF). The Government of Sindh's Education Sector Reform Program, which includes the intervention evaluated in this study, received financial and technical assistance from the World Bank and the European Commission. We have multiple organizations and several people to thank. First, the Government of Sindh's Planning and Development, Finance, and Education and Literacy Departments; the Sindh Education Sector Reform Program Support Unit; and SEF for partnering with the evaluation team. In this regard, the following SEF staff in particular (last names in alphabetical order): M. Abdullah Abbasi, Naheed Abbasi, Ambreena Ahmed, the late Anita Ghulam Ali, Imam Bux Arisar, Sadaf Bhojani, Mukhtiar Chandio, Sana Haidry, Abdul Fateh Jhokio, Aziz Kabani, Tauseef Latif, Adnan Mobin, Dilshad Pirzado, Shukri Rehman, Shahpara Rizvi, Rustam Samejo, Noman Siddique, and Sadaf Junaid Zuberi. Second, the following World Bank and European Commission staff for their support to the design, implementation, and evaluation of the intervention: Umbreen Arif, Salman Asim, Siddique Bhatti, Reema Nayar, Quynh Nguyen, Peter Portier, Uzma Sadaf, Benjamin Safran, and Sofia Shakil. Third, Mariam Adil and Aarij Bashir for their field-based support to the evaluation. We benefited from comments from Richard Murnane, Emmerich Daviesx' and from participants at presentations at Harvard University, the World Bank, Keywords: Primary education; gender disparity in education; public-private partnership; private schools; public subsidies; market competition; randomized controlled trial; Sindh; Pakistan

JEL codes: I20; I25; I28; 015

RISE Conference 2017, NBER Education Meeting 2013, and IZA Labor Conference 2011. Financial support for the study from the Australian Department of Foreign Affairs and Trade and the World Bank is gratefully acknowledged. The experimental project has IRB approval number AAAF4126, Columbia University. This trial has been registered at the American Economic Association RCT registry repository with the number AEARCTR-0002407.

1 Introduction

Despite dramatic improvements in recent years, school enrollment continues to be extremely low in many low-income countries (UNESCO, 2015a), while average student learning is often poor even when children are in school (UNESCO, 2014; Pritchett, 2013). In the face of these challenges, some governments are experimenting with educational models giving a greater role to private providers (Patrinos et al., 2009). The effectiveness of such an approach ultimately depends on whether private providers select appropriate school inputs. Though private entrepreneurs may have greater incentives to reduce cost and innovate than the public sector, their focus on profits may lead to the under-provision of other socially-valuable but non-contractible aspects of education (Hart et al., 1997). In this paper, we provide evidence from a field experiment in rural Pakistan on the effectiveness of private education. Our results indicate that private provision can be highly effective: entrepreneurs simultaneously reduced costs while providing high-quality school inputs nearly identical to those that the social planner would have chosen.

Estimating the effectiveness of private schools is complicated by the non-random placement of existing private schools and pervasive student sorting. We leverage a unique experiment to circumvent these problems, and provide a causally-identified assessment of privately operated schools. We evaluate the Promoting Low-Cost Private Schooling in Rural Sindh (PPRS) program, which was implemented in the Sindh province of Pakistan. In this program, publicly-subsidized private schools were randomly assigned to educationally underserved villages, with private entrepreneurs given responsibility for creating and managing these schools, and compensated according to enrollment on a per-child basis. Entrepreneurs exercised wide latitude in the inputs they provided, including the ability to hire teachers with lower formal qualifications than required for government teachers. In addition, a second treatment arm included a subsidy premium for girls' enrollment, incentivizing entrepreneurs to provide inputs that would be attractive to girls.

Since Friedman (1955), a lengthy literature has posited that giving a larger role to private education providers would increase school productivity. As discussed in Shleifer (1998), private schools run by entrepreneurs may have advantages over public schools due to their stronger incentives to reduce costs, innovate (e.g. tailoring school inputs more closely to the preferences and needs of their students), and their more direct accountability to the households they serve.[1] A substantial empirical literature has sought to assess the efficiency of private schools, generally through cross-sectional comparisons of student outcomes in public and private schools, or evaluations of voucher programs using pre-existing private schools (Friedman, 1955; Hoxby, 2003; MacLeod and Urquiola, 2012). The empirical liter-

[1] In turn, programs based on private schools, such as vouchers, may induce higher competition and general equilibrium effects (see Hoxby, 2003)

ature includes growing evidence from developing countries (Angrist et al., 2002; Newhouse and Beegle, 2006; Desai et al., 2009; French and Kingdon, 2010; Bold et al., 2014; Muralidharan and Sundararaman, 2015; Singh, 2015), and in particular evidence from Pakistan based on cross-sectional and panel observational data (Andrabi et al., 2010, 2011; Amjad and MacLeod, 2014).

The PPRS program provides a unique opportunity for assessing the quality of privately provided education, and the potential for improving educational outcomes in developing countries through greater reliance on the private sector. First, because the schools were allocated to villages through a randomized design, we are able to assess the effectiveness of such schools free from the biases usually present due to endogenous placement. Second, because a large share of villages in our sample had pre-existing government schools, we are able to assess the relative quality of the privately operated program schools against that of public schools, using detailed child- and household-level information to account for child sorting. Finally, because the schools were created at the beginning of the program, we are able to shed light on how private institutions choose their inputs. For this purpose, a structural model is used to show how school inputs are determined as an equilibrium outcome of the interaction of optimizing producers and consumers, and to disentangle the supply- and demand-side determinants of school inputs. This model also allows us to assess how closely program-school operators come to the configuration of school inputs that would have been selected by a social planner seeking to maximize total societal welfare.

The PPRS program was designed and administered by the Sindh Education Foundation (SEF), a semi-autonomous organization in the Sindh provincial government. The program offered qualified local entrepreneurs a set of benefits to establish and run tuition-free, co-educational primary schools in educationally underserved villages. The benefits included a per-student subsidy, school leadership and teacher training, and teaching and learning materials. The per-student subsidy amount was fixed at less than one-half the per-student cost for public primary and secondary education in the province. The program was randomly assigned to 200 out of 263 qualifying villages in eight districts selected for their poor education outcomes.[2]

To address the large gender disparity in primary school enrollment in rural Sindh, 100 of the 200 program villages were randomly assigned to a gender-differentiated subsidy scheme. Under the latter scheme, program school operators received a higher per-student subsidy

[2]After the random assignment was completed, the original evaluation sample was reduced from 263 to 199 villages to correct for errors made in determining whether a given village qualified for the program. The corrections were orthogonal to the assigned program status of the village. The effective evaluation sample consisted of 82 villages under the gender-uniform subsidy treatment, 79 under the gender-differentiated one, and 38 control villages. Mean household and child characteristics in the 199 villages were similar across the experimental groups in both baseline and follow-up. Follow-up measurement was conducted after program schools had been in operation for approximately 1.5 school years.

for girls than for boys, with the aim of more strongly incentivizing them to take steps to attract girls. Girls in developing countries are less likely to enroll in school than boys, especially among poor, rural, or socially disadvantaged households (UNESCO, 2015b). While these gender disparities are often attributed to lower household demand for girls' education, school supply factors—such as distance and time from home to school, school infrastructural features and environmental conditions, and teacher characteristics, attitudes, and behaviors—have also been documented to play an important role (Lloyd et al., 2005; Burde and Linden, 2013; Adukia, 2017; Muralidharan and Prakash, 2017).

The program was highly effective. Comparing treatment and control villages, after 1.5 school years the program increased school enrollment for children aged 6–10, the program's stated target age group, by 30 percentage points, and that for children aged 11–17 by 12 percentage points. The program also raised total test scores by 0.63 standard deviations, and by two standard deviations for children induced by the program to enroll in school. The gender-differentiated subsidy treatment had similar impacts on girls' enrollment and test scores as the gender-uniform one. Program-village households were more likely to voice aspirations for their boys to become doctors and engineers, rather than security personnel; and for their girls to become teachers, rather than housewives. Program-village households also expressed a desire for their boys and girls to attain higher levels of education.

Program schools appears to be more effective than nearby government schools, despite their far more modest resources. Children in program schools scored 0.21 standard deviations higher on the exam administered by surveyors than those in government schools in control villages. This finding is robust to a variety of sample restrictions based on the proximity of households to government and program schools (using the proposed sites of program school in control villages), as well as alternative specifications including only children for whom enrollment was verified by their observed attendance. An education production function is estimated to shed light on the reasons for program school success. Consistent with previous research, teacher characteristics are found to be the most important factor for child achievement. Significantly, teachers are an input over which individual government schools have little discretion, whereas program schools were given control over teacher recruitment.

Because the comparison of the effectiveness of government and program schools may be confounded by child sorting, we compare the characteristics of children across school types, and show that children enrolled in program schools have *lower* socioeconomic status than those in government schools (in control villages), and that the characteristics of children in program schools in fact more closely resemble those of unenrolled children in control villages. This is likely due to the enrollment in program schools of children who otherwise would not have been in school, who generally come from households of lower socio-economic status, as proxied by the education and occupation of the head of household and the building material of their house. In addition, the characteristics of students in government schools was virtually identical across control and treatment villages, with no evidence for student sorting in program villages.

We also examine the efficiency of the input choices in program schools vis-à-vis the social planner's solution, based on structural model estimations of schooling demand and education production. Using information about household choices, we first estimate a demand model for school inputs. We then use these estimates to bound the costs of school inputs. The intuition is that for schools which provide a given input, the benefit must have exceeded the cost of the input, in terms of additional enrollment; while for schools without that input the opposite must be true. Finally, we estimate an education-production function relating test scores to school inputs and student characteristics. We compute the optimal set of school inputs that a social planner would have chosen, combining the input costs incurred by program-school operators, the deadweight loss from taxes for providing program subsidies, the surplus accruing to students, and the social benefit of education.

We find that SEF and program-school operators did remarkably well in choosing school inputs, capturing approximately 90 percent of the total amount of possible surplus. The main differences between the program schools' and social planner's solutions are that the latter requires schools to have toilets and/or drinking water, and employs less-experienced, but more-educated, teachers. The social planner also better matches the gender ratio of teachers with that of children in the village.

In sum, we find that private providers can deliver high-quality education at relatively low cost. For policy makers seeking to improve education service delivery in developing countries, government support for local private providers may be a viable alternative to pure public provision. The challenging context in which the program was implemented suggests the potential for such an approach to be effective in many other parts of the developing world.

2 Literature review

Low- and middle-income countries continue to struggle with the related problems of low enrollment rates and low student achievement (World Bank, 2018). Because public education is generally seen to be failing in these countries, governments have experimented with new models giving a greater role to private providers, in order to take advantages of the purported efficiencies of the private sector. One common model involves providing vouchers to students for enrollment in pre-existing private schools (for example, Angrist et al., 2002; Muralidharan and Sundararaman, 2015). Another type of program aims to convert existing public schools to non-government management (Romero et al., 2017), akin to charter schools in the US.

A third type of program, known broadly as Public-Private Partnerships (PPPs), uses a contract between the public and the private sectors for service delivery by the latter, generally specifying aspects of the service to be delivered (Patrinos et al., 2009).[3]

Most research on PPP models has focused on programs which use existing private schools to serve students (Kim et al., 1999; Alderman et al., 2001, 2003; Barrera-Osorio and Raju, 2015; Barrera-Osorio et al., 2016; Romero et al., 2017). An alternative approach entails the creation of new private schools, and is motivated either by the absence of pre-existing private schools, or the belief that new institutions may be more adaptable to government reform efforts than incumbent schools. Our paper contributes to the limited evidence on this alternative approach, with Alderman et al. (2003) being, to the best of our knowledge, the only existing evaluation of a similar program.[4]

The program studied here was designed to increase both enrollment and student achievement. The theoretical motivation for the program comes from the human capital model, with households investing in education if the present and future benefits are larger than the costs (Becker, 1962). The most important costs of primary schooling in low-income countries tend to be school fees and transportation costs, with many children living far from school. International evidence on the effects of distance from home to school on enrollment is strong: when schools are introduced into underserved areas, household demand for schooling increases, often substantially (Burde and Linden, 2013; Foster and Rosenzweig, 1996; Duflo, 2001; Berlinski et al., 2009; Carniero et al., 2016).[5] The international evidence for the effects of user fees on enrollment is also strong, with enrollment fees associated with lower enrollment rates (Deininger, 2003; Barrera-Osorio et al., 2007; Borkum, 2012; Lucas and Mbiti, 2012).

Evidence from demand-side interventions that lower the costs of education—for example, conditional cash transfer (CCT) programs—shows that bringing children into school does not guarantee higher student achievement.[6] In order to improve student outcomes, the pro-

[3] PPPs may enable higher-quality education for beneficiary students through the terms of the contract between the government and the private sector, with some programs regulating school inputs (including teachers), and others selecting program schools based on measures of quality. As such, this paper evaluates a particular policy that shares several characteristics with the generic PPP program: no-user fees; a contract based on number of enrolled students; external vetting; large freedom in input selection by the provider; among other characteristics.

[4] Alderman et al. (2003) evaluates a similar program conducted in the Balochistan province of Pakistan in the 1980s. This program was largely unsuccessful in rural areas, due in part to the low supply of qualified teachers. In contrast, the PPRS program was able to tap into a fairly large supply of educated females due to recent advances in rural education.

[5] Burde and Linden (2013) use experimental evidence to show the relationship between distance and enrollment. Foster and Rosenzweig (1996), Duflo (2001), Berlinski et al. (2009), and Carniero et al. (2016) use quasi-experimental evidence to show this relationship.

[6] Early rigorous evaluations of CCT in different contexts shows zero results in test scores (see Fitzbein and Schady, 2009; Saavedra and Garcia, 2012). However, more recent evidence of long-term effects shows more promising results (Barham et al., 2013; Barrera-Osorio and Filmer, 2016).

gram therefore sought to make use of the potential gains from private schooling (Friedman, 1955; Hoxby, 2003; MacLeod and Urquiola, 2012). In addition to a wider literature on the advantages of private education in developing countries,[7] this approach was motivated by research from Pakistan using quasi-experimental methods to show the relative effectiveness of private schools (Andrabi et al., 2010, 2011; Amjad and MacLeod, 2014).

From a theoretical perspective, Shleifer (1998) describes several economic forces which may drive private enterprises to be more efficient than their public counterparts: the incentives to reduce costs and innovate are generally higher for private enterprises, while reputation concerns and local competition induced by school choice may lead to higher incentives to provide high-quality services. These forces map to our setting along several dimensions. First, the entrepreneurs running the program schools are paid on the basis of enrollment, giving them high-powered incentives to reduce costs while maximizing enrollment. Second, program schools may enjoy greater local information and ability to adjust to local conditions. For example, research from India has shown that private schools tailor their curricula more closely to the preferences of local students (Muralidharan and Sundararaman, 2015). Third, program schools are potentially more accountable to the households they serve—in contrast to the indirect accountability of government schools through political processes, which has proven far less effective (World Bank, 2004). Finally, program schools may face different factor prices—especially for teachers—and higher flexibility in choosing inputs than public schools (Hoxby, 2003; MacLeod and Urquiola, 2012). The combination of these forces has the potential to lower costs while providing higher-quality educational services.

3 Background

3.1 Schooling in Pakistan

School enrollment is low in Pakistan, even when compared to countries with a similar income level (Andrabi et al., 2008b). At the time the PPRS program was initiated in 2009, the primary-school net enrollment rate (NER) for children aged 6–10 in Pakistan was 67 percent (72 percent for boys and 62 percent for girls) (Government of Pakistan, 2009).[8] In rural Sindh, where the PPRS program was implemented, the primary-school NER was 56 percent for children aged 6–10. The gender disparity was even wider, with a primary-school NER of 65 percent for boys and 46 percent for girls (Government of Pakistan, 2009).

[7]The relative effectiveness of government versus private schools on student learning has been the subject of significant interest in the education-economics literature. The empirical evidence in developing countries includes (Angrist et al., 2002; Newhouse and Beegle, 2006; Desai et al., 2009; French and Kingdon, 2010; Bold et al., 2014; Muralidharan and Sundararaman, 2015; Singh, 2015).

[8]The primary school net enrollment rate is defined as the number of children aged 6–10 attending school in grades 1–5 divided by the number of children aged 6–10.

Pakistan has witnessed a dramatic growth in private schools.[9] The number of private schools increased from around 4,000 in the early 1980s to 36,000 in 2000 to 47,000 in 2005; much of the expansion in the 1990s and 2000s occurred in villages and poorer urban neighborhoods (Andrabi et al., 2008a). By 2010–11, one-fifth of children aged 6–15 in Pakistan were enrolled in private school (or one-third of students, given the large share of unenrolled children) (Nguyen and Raju, 2015). These schools have succeeded both in terms of cost and quality. At less than $20 in 2000, the mean annual cost of private primary school fees represented about two percent of mean total household spending (Andrabi et al., 2008a). Low fees were enabled by few fixed costs, and low operational costs, specifically low teacher wages. Discussed in section 3, low-cost private schools are also found to produce higher test scores than government schools in rural Punjab province (Andrabi et al., 2010, 2011).

However, large spatial differences exist in private-school enrollment rates across Pakistan. Thirty percent of primary-school students were in private schools in the country, compared to only five percent in rural Sindh in 2008–09 (Government of Sindh, 2009). Andrabi et al. (2008a) find that, as in the rest of the country, private schools in rural Sindh tended to be found in larger villages with better infrastructure. They argue that the main constraint to the further expansion of low-cost private schools is the lack of a local supply of women with secondary education who can be hired as teachers.[10]

3.2 PPRS Program

To address the education access, equity, and quality challenges in Sindh, the provincial government in 2007 initiated the Sindh Education Sector Reform Program (SERP), a multifaceted reform of public spending and provision in primary and secondary education. Public-private partnerships in education, entailing public financing and private provision, was a key component of SERP, aimed at increasing access to schooling and the quality of education for socioeconomically disadvantaged children.

Funded by the provincial government, the Promoting Private Schooling in Rural Sindh (PPRS) program was designed and administered by the Sindh Education Foundation (SEF), a semi-autonomous organization established in 1992 by the provincial government to undertake education initiatives targeting less-developed areas and marginalized populations in Sindh province. The main, stated objectives of the PPRS program were to increase access to schooling in marginalized areas, to reduce the gender disparity in school enrollment, and

[9] These schools are for-profit, fee-based, and secular. They are unregulated in practice, and do not receive any direct government assistance.

[10] Andrabi et al. (2013) find that villages with government secondary schools for girls were much more likely to see low-cost private primary schools arise in later years. They argue that the main channel is the local supply of women with secondary education. These women can be hired as teachers at low wages, as they face limited alternative employment opportunities and restrictions on their geographic mobility.

to increase student learning, in a cost-effective manner.

The first phase of the program, which we evaluate in this study, was implemented in eight (out of, at that time, 23) districts in the province. SEF selected the districts based on how they ranked in terms of the size of the out-of-school child population, the gender disparity in school enrollment, and the percentage of households located at least 15 minutes away from the nearest primary school. The eight poorest-ranked districts were selected, excluding those that were viewed by the provincial government and SEF as experiencing heightened law-and-order concerns.[11]

Based on a budgetary assessment, SEF supported coeducational, private primary schools in 200 villages in the selected districts. The main benefits that program-school operators received included a per-student cash subsidy; free school leadership and teacher training; and free textbooks, other teaching and learning materials, stationery, and bookbags.[12]

Two types of monthly per-student subsidies were provided: a gender-uniform subsidy, where the school received 350 rupees (approximately $5 in annualized 2008 US dollars) for each student; and a gender-differentiated subsidy, where the school received 350 rupees for each male student and 450 rupees ($6.4) for each female student. A total of 100 schools received the gender-uniform subsidy, and another 100 schools, the gender-differentiated subsidy. All schools received the other benefits.

The subsidy amounts were set at less than one-half of the per-student cost in public primary and secondary government schools in the province, and were not adjusted for price inflation over the evaluation period.[13] Provided on a quarterly basis, a school's total subsidy was linked to the number of children in attendance multiplied by 1.25 to reflect an expected 20-percent student-absence rate. SEF gathered attendance information through periodic, unannounced monitoring visits to program schools.

Local private entrepreneurs were invited to apply to the program through an open call in newspapers, and to propose educationally underserved villages in the selected districts to establish and operate schools. SEF vetted the applications (ultimately, through visits to shortlisted villages) based on several criteria, including written assent from the parents of at least 75 children of primary-school ages that they would enroll their children in the school, should it be established; a school site in the village that was located at least 1.5 kilometers from the nearest other school; a building of sufficient size; and the identification of teachers with a minimum of eight years of schooling (middle school completion), with at least two

[11] The district rankings were determined using district-representative data from the 2006–07 Pakistan Social and Living Standards Measurement Survey (Government of Pakistan, 2007).

[12] The subsidies were transferred electronically to the bank accounts of the private entrepreneurs.

[13] Per-student costs in government schools were based on information from the provincial Education and Literacy Department's annual census of government schools and the provincial Finance Department's records on recurrent budgets and expenditures toward primary and secondary education.

being female.[14]

Once in the program, school operators would continue to receive the subsidy and other benefits as long as they provided free schooling to children and provided and maintained school infrastructure and the schooling environment consistent with SEF guidelines. SEF strictly enforced the free-schooling condition, but was more lenient in enforcing the school infrastructural features and environmental conditions.

4 Data

SEF administered a vetting survey to determine whether proposed villages qualified for the program. This survey, which we refer to as the baseline survey, was conducted in February 2009. Following the baseline survey, 263 villages that qualified for the program were randomly assigned to the two subsidy treatments, or to the control group. However, after random assignment, SEF scaled down the original evaluation sample of 263 villages to 199 villages to correct for errors in determining whether a village qualified for the program. The decisions taken by SEF were orthogonal to the assigned program status of the village. The effective evaluation sample consisted of 82 villages under the gender-uniform subsidy treatment, 79 under the gender-differentiated one, and 38 as controls.

Schools were established in summer 2009. Because the new school year normally commences in the spring, program-school students had an abbreviated first school year. A follow-up survey was conducted in April/May 2011, after the conclusion of the second school year under the program.

The baseline survey consisted of a village survey answered by village leaders, a school survey of all schools in the general vicinity of the village, and a household survey of 12 households randomly selected from the list submitted by the entrepreneur of 75 households that had agreed to send their children to the proposed program school. The household survey collected information on the household, the household head, and on each child aged 5–9. In each village, the baseline survey also consisted of a survey of the entrepreneur and proposed teachers, as well as physical checks by the survey interviewers of the proposed school site and building. GPS data were collected from all schools, the proposed program-school site, and surveyed households.

The sampling frame for the followup survey was based on all the households within the village. To determine which households belonged to the village, at the time of the first followup (6 months after the commencement of the program) enumerators asked to meet with a prominent member of the community. They then had this individual take them on a

[14] SEF viewed eight years of schooling as sufficiently high that teachers would have the competency to teach primary school content, but low enough that qualified individuals could be found locally.

tour of the village and indicate every household that belonged to the village. The full set of households in the village was then used as the sampling frame from which households were (randomly) selected for inclusion in the survey.

The follow-up survey consisted of three instruments: a school survey; a household survey; and a child survey, which included a test. The household and child surveys, and child tests, were administered at the child's home. The household survey was administered to households with at least one child aged 5–9. In large villages, up to 42 randomly sampled households in the village were interviewed; in villages with fewer than 42 households, which comprised the majority, all households in the village were interviewed. The survey was multi-topic, and had extensive modules on past and current schooling and other activities for children aged 5–17, answered by the household head or another primary adult household member. A child survey was administered to each child aged 5–9. It asked questions mainly on work activities performed inside and outside the home, past and current schooling, and aspirations. Each child was then administered tests on language (either Urdu or Sindhi, as preferred) and mathematics.

The school survey gathered information from interviews of head teachers and all other teachers, and visual checks by the survey interviewers of school infrastructural and environmental conditions. The survey also collected student attendance information through a headcount, with the attendance lists used during the household survey to verify the child's enrollment status reported by the household. GPS data were gathered from all surveyed households and schools.

Table 1 reports sample sizes of the baseline and follow-up surveys, by treatment status. The baseline survey interviewed 2,089 households and 5,556 children aged 5–9, and the follow-up survey interviewed 5,966 households and 17,720 children aged 5–17.

5 Empirical strategy

5.1 Conceptual Framework

Our broad research agenda is to understand how effective private entrepreneurs can be in setting up schools to provide educational services. In principle, a multitude of interventions might help to shed light on the various issues relevant to the private provision of education, among which include: the contractually-specified payoff incentives faced by entrepreneurs, which educational requirements and school inputs are regulated, where interventions were held, and so on. While logistical constraints require us to focus on a relatively narrow set of experimental interventions, our setting allows a particularly clean look at two specific questions of interest: how do educational enrollment and outcomes change with the presence of a private school, and how well do private schools do in providing education services?

To answer the first question, we provide estimates from an intention-to-treat analysis to show how the random presence of private schools changed equilibrium child enrollment and test scores. Due to the isolation of our rural subject villages, which prevented spillover effects across treatment and control villages, our interventions approach the ideal experiment for answering this question. We also show results for additional important outcomes, including: the influence of program schools on vocational aspirations, how treatment effects varied with distance, whether the gender-differentiated interventions had different outcomes, and how cost-effective the interventions were.

To address the second question, we pose and estimate a structural model, allowing us to provide a counterfactual evaluation of how closely the school inputs selected by private entrepreneurs aligned with those of a benevolent social planner. The intervention studied provides a uniquely clean setting for answering this question: survey data covering both treatment and control villages allows us to estimate demand for schooling with and without the randomly-assigned private schools, controlling for the presence of competing schools using a discrete choice model. A key aspect of the program—namely, the latitude enjoyed by entrepreneurs in making school input decisions—allows us to use a revealed preference approach to infer input costs. Survey data on enrollment and test scores allows us to then link school choice and school inputs to educational outcomes. Combining these elements with estimates from the literature on the social value of education, we explore how demand, cost, and the social value of educational outcomes would change with different school inputs in each treatment village. To our knowledge, this type of welfare analysis is novel in the literature, and would be much more difficult to evaluate without the leverage provided by the experimental intervention.

5.2 Intent-to-treat Estimates

Two of our primary outcomes of interest are child enrollment and test scores. We estimate the intention-to-treat (ITT) based on the following specification:

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 X_i + \varepsilon_{ij}, \tag{1}$$

where Y_i is the outcome of interest for child i, T_i is an indicator variable indicating whether child i resides in a village assigned a program school, and X_i is a vector of child and household controls. In other specifications, we examine the differential impacts of the program by gender, by the two subsidy treatments, and by the two subsidy treatments interacted with gender. Standard errors are clustered at the village level, j.

Internal validity: The validity of our results depends upon the comparability of populations across the experimental groups. Because the program was randomly assigned across villages, treatment status should be orthogonal to household and child characteristics that might be correlated with the outcomes. Insofar as this holds, it will be sufficient to compare outcomes across the treatment and control groups to evaluate the impacts of the program.

To assess comparability, we estimate the differences in mean household and child characteristics between program and control villages at baseline and follow-up. In Table 2, columns (1) and (3) report mean characteristics in control villages, at baseline and follow-up, respectively. Columns (2) and (4) report the differences in mean characteristics between program and control villages, at baseline and follow-up, respectively.

Differences in means in virtually all household and child characteristics between program and control villages were small and statistically insignificant. As an exception, the percentage of girls in program villages was slightly higher (4.2 percentage points and 3.0 percentage points at baseline and follow-up, respectively).

Analogously structured to Table 2, Appendix Table A.1 reports the differences in mean household and child characteristics between villages under the gender-uniform and -differentiated subsidy treatments. Differences in mean characteristics between the two subsidy treatments were small, and always statistically insignificant.

6 Program impacts

6.1 Enrollment

School enrollment information was collected in two ways. First, the household survey respondent was asked whether the child was enrolled during the just-concluded school term. Second, the enrollment of the child aged 5–10 in a given school was verified using a student attendance list compiled through a headcount conducted during the school survey.[15] We discuss results based on both reported and verified enrollment measures.

Table 3 reports the pooled-treatment impacts on school enrollment and grade attainment. Panels A and B report results for young children (aged 5–10) and older children (aged 11–17), respectively. Columns (1) through (4) report impacts on reported enrollment, with different sets of controls. Columns (5) and (6) report impacts on verified enrollment and grade attained, respectively, with the full set of controls. Based on the model with the full set of controls, the program increased reported enrollment among young children by 31.5 percentage points, and verified enrollment by 29.4 percentage points. In addition, the

[15]The school surveys were conducted first, so that the school-attendance decision would not be influenced by the presence of survey interviewers in the communities. Using the attendance sheets collected during the school survey, the survey interviewers verified the child's enrollment status reported by the household-survey respondent.

program increased attainment by 0.38 grades.

While older children were not the expressed target population for the program, we nonetheless find significant increases in reported enrollment for them. Based on the model with the full set of controls, we find that the program increased reported school enrollment among older children by 11.0 percentage points. We do not find an impact on grade attainment for these children. The reason for this is a combination of the smaller impact on enrollment, as well as the fact that the older children were enrolling in the grades offered in program schools, which were at the primary level.

6.2 Test Scores

Table 4 reports the impacts of the pooled treatment on test scores. Test scores are standardized by subtracting the mean score for all children aged 5–9 in control villages and dividing by the standard deviation (47% and 31%, respectively). Columns (1) through (4) report impacts, with various sets of controls. Based on the model with the full set of controls, the program increased total test scores by 0.63 standard deviations. Program impacts were similar for both subject test scores.

We also estimate the treatment-on-the-treated (TOT) impact of reported and verified enrollment on test scores (reported in columns (5) and (6), respectively). For these estimates, we regress the respective enrollment measure on pooled-treatment status in the first stage, and regress test scores on predicted enrollment in the second stage. The program increased total test scores by two standard deviations among children induced by the program to enroll in school. The impacts were similar for both subject test scores. The results suggest that program schools were highly effective in imparting basic numeracy and literacy skills to students.

We also report program impacts on test scores taking as the outcome the percent of test questions answered correctly. Appendix Table B1 reports the results. Columns (2)–(5) report the ITT estimates, with various sets of controls; and Columns (6) and (7) report the TOT estimated effects of reported and verified enrollment on test scores, respectively. The ITT impact on the total test in the model with the full set of controls was 19.7 percentage points, and the TOT estimate was approximately 60 percentage points.

We also examine program impacts on test scores by child age (Appendix Table B2) and the competency being tested (Appendix Table B3). For child age, we report test-score effects using as the outcome variables both the percent of questions answered correctly (Columns (3) and (4)) as well as the standardized test score measure (Columns (5) and (6)). For the results on tested competency, the outcome is percent of questions answered correctly. The program effects were largely similar across age groups, and were also similar across the various tested competencies.

6.3 Differential impacts on school enrollment and test scores

We examine the impacts of the two subsidy treatments (Table 5, panel A), the impacts of the pooled treatment by gender (Table 5, panel B), and the impacts of the two subsidy treatments by gender (Table 6), based on models with the full set of controls. We do not find differential impacts by subsidy treatment, by gender, or by subsidy treatment and gender.

6.4 Distance to school and educational outcomes

We first describe the relationship between educational outcomes and distance to the proposed program school site, shown in Figure 1. As is apparent, the treatment increased enrollment rates and test scores at all distances from the proposed site. The impacts of the program schools, however, were only in part due to the decline caused in the distance to the nearest school. For villages lacking any proximate non-program school, the creation of a the program schools was decisive. Even for villages that did possess a non-program school, however, there was a sharp increase in enrollment in treatment villages.

In Figure 2 we plot the relationship between educational outcomes and the distance to the nearest school regardless of school type. For control villages we plot this relationship up to a distance of 1.5 kms. In program villages we plot the relationships up to a distance of 1 km, as there are very few households in these villages located further away. Remarkably, there is no relationship between the educational outcomes and school distance in treatment villages. In control villages, in contrast, there is clear gradient between distance and the two educational outcomes. In addition, even at relatively close distances children in control villages are less likely to be enrolled, and receive a lower test score, than children in treatment villages.

There are two likely explanations for the disparity between the importance of distance in control and treatment villages. First, it may be the case that program schools are perceived to be relatively high quality, and that distance is less important to households under such circumstances. Alternatively, because the payment scheme is based on the number of students enrolled, entrepreneurs may have taken measures to maximize enrollment.

Appendix Figures A1 and A2 show the relationship between educational outcomes and distance to the nearest school, disaggregated by village treatment status and child gender. In treatment villages, boys and girls have virtually identical enrollment rates and test scores at all distances. In contrast, in control villages, a substantial gap opens up between boys and girls in both enrollment rates and test scores above distances of 0.6 kilometers. This finding resembles that of Burde and Linden (2013), who show that school proximity is a significant determinant of the educational gender gap.

6.5 Aspirations

Given the impacts on school enrollment and test scores, it is unsurprising that households may have adjusted their aspirations. In Table 7, panel A reports impacts on aspirations for each child aged 5–17 conveyed by the household, and panel B, on those conveyed by each child aged 5–9. Column (1) reports means in control villages, and column (2) reports the differences in means between program and control villages. We also examine gender-differential impacts on aspirations. Columns (3), (4), and (5) report regression coefficients for girls, the program, and the interaction of the two, respectively.

Relative to their counterparts in control villages, program-village households were more likely to desire that their boys become doctors (+5.8 percentage points) and engineers (+2.6 percentage points), and less likely to desire that they become security personnel (-5.0 percentage points). They were also more likely to desire that their girls become teachers (+6.5 percentage points), and less likely to desire that they become housewives (-14.9 percentage points). Program-village households desired higher attainment for their boys and girls (+1.5 and +1.7 years, respectively). In terms of age of marriage, program-village households had similar preferences to control-village households, for boys and girls.

Program-village boys were more likely to desire public sector employment (+12.2 percentage points). Program-village children did not desire more years of education than control-village children. However, children in both program and control villages desired more years of education than was desired by households (11.3 years for children versus 7.4 years for households, in control villages).

7 Program cost-effectiveness

SEF maintained records of all program costs under detailed accounting heads. Figure 3 depicts the distribution of program cost components in fiscal years 2008–09, 2009–10, and 2010–11. The fiscal year runs from July 1 to June 30. In fiscal year 2008–09, the program was launched. However, subsidy payments to phase-1 program schools, those under this evaluation, were only provided in the last quarter of that fiscal year. Consequently, subsidies represented a small percentage of total program costs in fiscal year 2008–09, while fixed costs and other variable costs such as those related to administering the first phase of entry into the program represented large percentages. Costs in fiscal year 2010–11 are until April 2011, when the follow-up survey was administered, but two months short of the end of the fiscal year.

Over the evaluation period, SEF incrementally scaled up the program in phases, which affected the level and composition of costs. In fiscal year 2009–10, SEF administered a second phase of entry, with 97 schools added to the program. By the time of the follow-up survey, SEF had provided eight subsidy payments to phase-1 program schools. As school operators could not charge any fees, subsidy payments represented the sole source of school revenues. Subsidy costs for phase-1 (all) program schools evolved from 30 percent (30 percent) of total costs in fiscal year 2008–09 to 67 percent (72 percent) in 2009–10 to 48 percent (73 percent) in 2010–11.

The scale-up of the program during the evaluation period also affected how cleanly we could assign costs to phase-1 program schools. The cost data allow us to distinguish between subsidy costs for phase-1 and phase-2 program schools, but we could not separate out other types of costs in the same way. The cost data for fiscal year 2010–11 include early expenses for administering a third phase of entry into the program, which we also could not separate out. Given this, for the cost-effectiveness calculation, we simply treat non-subsidy costs to be fully assigned to phase-1 program schools. In addition, in July and August 2010, Sindh experienced major floods, and some schools were damaged or their operations were disrupted. SEF incurred costs helping to rehabilitate schools and restore school operations. The assignment of total non-subsidy costs to phase-1 schools raises costs used in the cost-effectiveness calculation. The natural disaster also raises costs relative to what could be expected in more normal times. These two factors would work to bias downwards the cost-effectiveness of the program.

All program costs are calculated in present value terms in 2011 US dollars following the method proposed by Dhaliwal et al. (2013). SEF conducted its last unannounced monitoring activity before the follow-up survey in February 2011. In that activity, for phase-1 program schools, SEF found 28,827 children enrolled based on school registers, and 18,820 children in attendance based on a head count. Enrollment counts obtained from school registers may not be reliable if, for example, the registers are not updated regularly or schools perceive it is in their interest to inflate their enrollment counts. Assuming a 20-percent student-absence rate in rural, remote Sindh, we estimate an enrollment of 23,525 children, which we presume to be more accurate. Although the evaluation period runs over three fiscal years, program schools operated for 1.5 school years over the period. Depending on the year type (fiscal, school) and child (enrolled, attending), the annual program cost per student ranges from a low of \$77 to a high of \$184.

Program impacts on school enrollment and total test scores were roughly 30 percentage points and 0.6 standard deviations, respectively. Using the low and high values of annual cost per student, we estimate cost-effectiveness values of 16 to 39 percentage points in school enrollment and 0.3 to 0.8 standard deviations in total test scores, both per \$100 spent.

Program cost-effectiveness values associated with test scores appear to be at the lower end of the range of similarly estimated cost-effectiveness values for 14 education interventions reported by Evans and Popova (2016), only superior to a conditional cash transfer program in Africa.

Since the program impacts are measured with imprecision, following Evans and Popova (2016), we also estimate cost-effectiveness values at the lower and upper bounds of the 90-percent confidence intervals around the impacts. At the lower bound, we estimate cost-effectiveness values (associated with the alternative annual cost per student values) of 11 to 27 percentage points in school enrollment and 0.2 to 0.5 standard deviations in total test scores, per $100 spent. At the upper bound, we estimate cost-effectiveness values of 22 to 52 percentage points in school enrollment and 0.5 to 1.1 standard deviations in total test scores, per $100 spent.

While the program had large impacts on school enrollment and test scores, these impacts were accompanied by relatively large expenditures. Both the large impacts and expenditures are arguably due to the type of intervention: introducing new schools. Most of the other interventions with comparable cost-effectiveness analyses—and with superior cost-effectiveness results—were those introduced into (communities with) preexisting schools (Evans and Popova, 2016). SEF has continued to scale up the program, adding more schools and upgrading some primary schools to middle schools (up to grade 8), which has contributed to falling annual costs per student, as operating costs associated with such things as program administration and teacher-training workshops are spread over increasing numbers of schools and students. However, we do not know how program impacts have evolved in tandem with the scale-up.

Our program cost-effectiveness results only account for the costs borne by SEF which subsume all expenditures made by program-school operators. The results do not include the net costs—including opportunity costs—borne by households in choosing to send their children to program schools.

8 Program schools

8.1 Characteristics

A variety of schools were present in the program villages other than the program schools. Appendix Table A.3 reports statistics on school availability and enrollment according to school type. Control villages were slightly more likely to have a government school at the time of the survey than were treatment villages (55% and 46%, respectively), though these imbalances are absent when limiting the sample to schools founded prior to the study period.[16] Despite the similar availability of government schools, children in treatment villages are far less likely to be enrolled in government schools.

We next examine some of the differences in mean characteristics between program and government schools, and between program and private schools (Table 8), as well as between program schools under the gender-uniform and -differentiated subsidy treatments (Appendix Table A.2), at follow-up. The means are estimated based on student-school observations.

In Table 8, columns (1) and (4) report mean characteristics for program schools, and columns (2) and (5) report the differences in mean characteristics between program and government schools. Columns (3) and (6) report differences in mean characteristics between program and private schools. We find that program schools were open 0.5 more days per week than government schools, indicating that they were generally open six days per week. Program schools were more likely to use English as the medium of instruction (+31.0 percentage points), and less likely to use Sindhi (-37.0 percentage points). The extent of physical infrastructure was higher in program schools than in government schools, with more having an adequate number of desks (+14.4 percentage points), drinking water (+30.7 percentage points), and toilets (+29.1 percentage points).

Based on information from head teachers, we find that program schools were staffed with more teachers than government schools (+1 teachers), with a larger number of teachers being female (+1.5 teachers). A greater number of teachers at program schools had either less than five years of teaching experience (+2.5 teachers) or 5–10 years of teaching experience (+0.4 teachers), and fewer had more than 10 years of teaching experience (-2 teachers).

Based on information from individual teachers, we find that program-school teachers were more likely to be female (+24.3 percentage points); they were younger (-13.6 years), and received lower monthly salaries (-11,454 rupees). In addition, program-school teachers had fewer years of teaching experience (-11.3 years). Program-school teachers spent a similar amount of time engaged in various classroom activities as government-school teachers, save for an additional 0.9 hours per week testing children, and an additional 0.6 hours per week teaching small groups.

In Appendix Table A.2, columns (1) and (3) report the mean characteristics in program schools under the gender-uniform subsidy treatment, and columns (2) and (4) report the differences in mean characteristics between program schools under the gender-uniform and -differentiated subsidy treatments. We do not find that program-school operators under the gender-differentiated subsidy treatment had structured their schools differently from their counterparts under the gender-uniform one. The absence of differential impacts on girls' school enrollment and test scores across the two subsidy treatments discussed earlier is in

[16] There was a push during these years to (re-)establish schools in villages without operational government schools.

line with the findings here.

8.2 Quality

One reason for the program's public-private partnership design was to take advantage of the potential private-school effect on test scores, for which evidence is found for low-cost private schools in larger, more-developed villages in Pakistan (Andrabi et al., 2010, 2011). The operator had flexibility in how to structure and run the program school around guidelines provided—but applied leniently—by SEF.

To determine whether the private-school effect on test scores exists for program schools, we compare mean test scores of children enrolled in program schools to those of children enrolled in proximate government (and private) schools in control villages. In Table 9, column (1) reports mean test scores for program schools, columns (2) and (3) report differences in mean test scores between program schools and children enrolled in the indicated school type in control villages, and column (4) reports p-values from tests of differences in mean test scores between government and private schools. Children in program schools scored 0.21 standard deviations higher on the total test than those in government schools (0.24 standard deviations higher on the mathematics test, and 0.16 standard deviations higher on the language test). In contrast, differences in mean test scores between program and private schools were small (about 0.02–0.05 standard deviations) and statistically insignificant.

These comparisons do not causally identify differences in quality between school types, as student-composition effects are likely to bias these estimates. If program schools attract students who would not otherwise have been enrolled, and if these students come from more socioeconomically disadvantaged backgrounds, the program-school effect on test scores may be biased downwards. In contrast, if the most talented students in government schools switch to program schools in treatment villages, the test scores of children in program schools would overstate their quality.

The evidence is strongly supportive of the former hypothesis, with program schools appearing to have encouraged the enrollment of socioeconomically disadvantaged students. Appendix Table A.4 reports differences in household and child characteristics across unenrolled and government-school students in control villages (columns 1, 2, 5, and 6); and across government-school and program-school students (columns 3 and 7). Government-school students came from households where household heads had more years of school (+1.7 years) and were less likely to be farmers (-10.9 percentage points). Government-school students were also less likely to come from households residing in poor-quality dwellings made of mud or thatch (-19.4 percentage points). These differences are almost perfectly off-set by program schools, so that program-school students more closely resemble unenrolled children in control villages.

To test for student sorting across school types, Appendix Table A.4 also shows the differences across treatment and control villages for students who were enrolled in government schools (columns 4 and 8). We find little evidence for sorting, with mean characteristics of government-school students largely similar across control and program villages, with the exception that government-school students in program villages were slightly older (+0.3 years) than their counterparts in control villages. This is presumably because some share of the younger children who would have otherwise enrolled in government schools absent the program selected program schools instead, skewing the age distribution slightly upwards.

As an alternative approach, we compare test scores under a variety of sampling restrictions. We first restrict the sample to households within 0.5 kms of a non-program school (columns 3 and 4), and then to households within 0.5 kms of both a non-program school as well the proposed site of the the propgram school (columns 5 and 6). The aim of these sample restrictions is to compare children who would have been likely to enrol even absent the treatment, and therefore to focus on children who were induced to switch schools because the program school was introduced. This approach allows us to reduce biases coming from household characteristics that may be correlated with the household's location within the village. As seen in Panel A of Appendix Table A.5, the results are robust to these samplerestrictions, with virtually no change in the coefficients. In Panel B of the table we include a control for the highest grade that the child has completed, and in Panel C we restrict the sample of enrolled students to those for whom attendance was verified. The results are again highly robust when using these alternative specifications.

To understand the determinants of the program-school effect on test scores, we estimate an education-production function, by regressing test scores (for students in all schools) on a set of school inputs, controlling for child characteristics. Table 10 reports estimates for total test scores; Appendix Table A.6 reports estimates for subject test scores. The results support the findings discussed above. Specifically, school inputs significantly associated with test scores—such as the teacher's education and teaching experience—are those over which program schools exercise greater independent control do than government schools. This finding holds even when include an indicator variable for program schools in the regression, indicating that the estimates of the education-production function are not capturing some unobserved feature of program schools with which they are correlated.

8.3 Efficiency

We assess the efficiency of the input choices made by SEF and program-school operators by asking whether the social planner could have improved on the program solution, and if so, by how much and by what mechanism. A simple model shows why SEF- and program-school operators (hereafter, simply the program-school operator) may have incentives that are not perfectly aligned with those of the social planner. Consider the following model of a program-school operator deciding which school inputs to provide. As the program-school operator is provided a subsidy based on enrollment, let child demand for the school be denoted by $q(x) > 0$, where x is a vector of inputs and $q'(x) < 0$. The cost of providing the inputs is given by a positive increasing function, $c(x)$. The social value of providing the inputs is given by a positive increasing function, $h(x)$; this function captures both consumer surplus and broader societal benefits from children receiving an education. The first-order condition for the program-school operator is:

$$pq'(x) - c'(x) = 0, \tag{2}$$

while the corresponding first-order condition for the social planner is:

$$pq'(x) - c'(x) + h'(x) = 0. \tag{3}$$

The difference between these two first-order conditions is the inclusion of the marginal social benefit. In our setting, that term is consumer surplus plus the social value of higher test scores. In general, the program-school operator will fail to provide the socially optimal level of inputs because it does not capture the complete rents generated by their provision. In contrast, the social planner will provide inputs if their marginal social benefit exceeds their cost.

Our exercise consists of four steps. First, we estimate a discrete choice model of household demand for schools (referred to as "child" demand); it allows us to compute both the expected distribution of school enrollment, which in turn determines school-operator revenues, and consumer surplus under both observed and counterfactual school-input configurations. Second, we estimate the costs of providing school inputs using a simple revealed preference argument; it allows us to calculate the cost of providing such input configurations. Third, we estimate an education-production function relating student test scores to school inputs; it allows us to calculate the counterfactual distribution of student test scores. Finally, we tie it all together with a calculation of the social value of school-input configurations that accounts for surplus accruing to students, school-operator input costs, and the broader societal value of education.

We begin by estimating the demand for schooling by children in the villages. In turn, this allows us to evaluate how that surplus changes with changes in school inputs. We model child demand for schools using a standard logit random utility framework. Each child makes a single choice from a set of schools, J, where the utility of choice $j \in J$ to child i is given by:

$$u_{ij} = X_{ij}\beta + \epsilon_{ij}, \tag{4}$$

where X_{ij} is a vector of child characteristics and school inputs, β is a vector of marginal utilities, and ϵ_{ij} is an idiosyncratic preference shock distributed as Type I Extreme Value. We normalize the utility of not going to school to zero.

For the demand function estimation, we include a variety of school inputs and child characteristics shown to be important in the education-production literature (Todd and Wolpin, 2003). Child characteristics consist of gender, age, distance from home to the school, and interactions between the child's gender and school inputs. School inputs consist of toilets and/or drinking water (a single indicator variable); as well as teacher characteristics, such as gender, teaching experience, frequency of absence from school, and time spent teaching. We also include interactions of school inputs with an indicator variable for female students, as a substantial body of research shows the importance of school-supply factors for girls' enrollment and learning—such as distance and time from home to school, school infrastructural features and environmental conditions, and teacher characteristics and behaviors (Lloyd et al., 2005; Burde and Linden, 2013; Muralidharan and Sheth, 2016; Adukia, 2017; Muralidharan and Prakash, 2017).

Table 11 reports the demand estimates for schools. Column (1) reports results for a parsimonious specification; column (4) reports those from our richest specification, which includes interactions and an indicator variable for government schools. The coefficients have the expected signs. Looking at column (4), our preferred specification, boys and older children were more likely to enroll in school. Children were more likely to enroll in school if it had toilets and/or drinking water, had lower fees, had teachers who had fewer absences, and was not a government school. The percentage of female teachers has a large, negative effect on enrollment; however, female students have a positive and statistically significant demand for female teachers. This is interesting since it implies both boys and girls would prefer to attend schools with teachers of the same gender.

The demand estimates capture children's willingness to pay for various school inputs. To understand the role of program schools in producing learning, we regress student test scores on the same school inputs to estimate an education-production function. Table 10 reports the results. Column (1) does not control for school type; column (2) includes an indicator variable for government schools; and column (3) includes an indicator variable for program schools. The key school inputs that influenced test scores were the percentage of teachers with less than five years of teaching experience, and the percentage of teachers with post-secondary education.

Next, we use the demand curve to estimate bounds on school-input costs. We focus on those inputs that are most relevant to the education-production function and that were under the control of the school operator: provision of toilets and/or drinking water, the percentage of female teachers, the percentage of more-educated teachers, and whether teachers were frequently absent. We assume that schools will provide an input, such as toilets and/or drinking water, if its cost did not exceed the additional revenue that it generates through increased enrollment. Likewise, for schools that did not provide the input, the opposite must be true. These two inequalities bound the cost of the input. This exercise requires the use of a structural model, since we need to recalculate the expected distribution of students across schools under a counterfactual set of inputs not observed in the data. Our demand model also corrects for the fact that in areas with competing schools, providing an additional input may not be as profitable as in other areas.

Table 12 presents the results. As expected, the first input, toilets and/or drinking water, has a positive cost, as this amenity generates positive demand for all students, although it is differentially demanded more strongly by females. The next four inputs change the composition of teachers at the school. The first estimate reflects the cost of replacing a male teacher with a female teacher. Male students reacted negatively to the presence of a female teacher, while the opposite was true for female students. In combination with the number of boys and girls in each village, the sum of these forces implies that enrollment was lower when program schools substituted a female teacher for a male teacher. This, in turn, implies that female teachers must have been less costly than their male counterparts. Evidence for Pakistan is consistent with this finding. Andrabi et al. (2008a) find that female teachers in private schools earn 33-percent less than their male counterparts, after controlling for other characteristics. As might be expected, we also find that adding teachers who were frequently absent or had less than five years of teaching experience were less costly compared to more reliable and experienced teachers. Surprisingly, our model suggests that teachers with post-secondary education were also less expensive to hire than less-educated teachers, which apparently contradicts the higher salaries they receive, although this result is not statistically significant. This may indicate that these teachers were more efficient, or provided additional services to the school. In results not shown, we find that post-secondary educated teachers indeed worked longer hours.

Combining these pieces allows us to address a key question: are program-school operators providing inputs that maximize child outcomes? To answer this question, we first parameterize the social welfare function:

$$W(x) = CV(x) - TC(x) + \tau g(x), \tag{5}$$

where $CV(x) = \sum_{i=1}^{N} CV_i(x)$ is the sum of the consumer surplus over all children in the village, $TC(x)$ is the (total) cost incurred in providing inputs and subsidies, and $\tau g(x)$ is the social value of higher test scores.[17] We assume that the social value of education is related to overall test scores, given by $g(x)$, and a scalar multiplier, τ.

The logit model provides a basis for computing the consumer welfare generated by the school. Following Small and Rosen (1981), the compensating variation of choice set under the logit model is:

$$CV_i = \frac{(\gamma + \ln \exp \sum(\delta_{ij}(x)))}{\alpha}, \tag{6}$$

where $\delta_{ij}(x)$ is the deterministic component of utility of student i choosing school j, α is the disutility of school fees, and γ is Euler's constant. Our estimates above give the cost of each input, x.

The last component of our welfare analysis is the social benefits of education that are not internalized in the demand function. Since we do not know exactly the social benefits of education, we choose to parameterize the social benefit function as $h(x) = \tau g(x)$, where $g(x)$ is the estimated education production function. This specification assumes that the social benefits of education are only a function of test scores, and τ captures the marginal (social) utility of higher test scores. This approach allows us to: first, solve the social planner's solution, as total benefits of providing inputs can be consistently compared with their costs; second, show how the social planner's solution to providing inputs changes with τ, a parameter for which we do not have measurements; and, third, compute the efficiency of the observed allocation relative to what the social planner would have done.

The costs incurred in providing education are twofold. First, there is the direct cost of the inputs provided by program-school operators. Second, there is a deadweight loss due to the taxes necessary for providing subsidies to program-school operators.[18] To estimate the latter, we assume a deadweight loss of 30 percent, and multiply this by the total annual subsidy, which is a fixed per-student amount of 4,200 rupees per year under the gender-uniform subsidy treatment, which rises to 5,400 rupees per year for girls under the gender-differentiated one.

We define the social value of education as the product of the student's annual adult income and a social externality multiplier. To estimate the effect of education on labor earnings, based on estimates for Pakistan by Montenegro and Patrinos (2014), we fix upper- and lower-bound wage gains from an additional year of education at 10.8 percent and 6.8 percent, respectively. In addition, Bau and Das (2017) find that an additional year of education in rural Punjab, Pakistan is associated with a test-score gain of 0.4 standard deviations. Combining these two findings, we assume a test-score gain of 0.4 standard deviations to

[17] The profit of the program-school operator has been omitted from the social welfare function, as the income earned by the operator is a transfer. Although in this case the funds came from international donors, we compute the social planner's solution treating these funds as if they had been raised from domestic sources.

[18] See footnote 13.

be equivalent to an additional year of education, and, therefore, to produce wage gains of 10.8 percent and 6.8 percent at the upper and lower bounds, respectively. The wage gain is calculated as a function of the baseline wage and the labor force participation rate:

$$\triangle\ wage_{gb} = blwage_g \times (\frac{zscore(test)}{0.40}) \times percent\triangle wage_b \times participationrate_g, \tag{7}$$

where the subscript g indicates the gender of the child, and b the upper and lower bound estimates of wage gains. In rural Sindh, the baseline monthly wage ($blwage$) for men aged 15–34 is 6,600 rupees, and that for women in the same age group is 2,000 rupees; and labor force participation rates for the two are 80 percent and 36 percent, respectively.[19] We inflate the term with the multiplier above to account for social externalities.

For each program school in our sample, we solve the following social planner's problem:

$$\max_x W(x). \tag{8}$$

This problem is non-convex, due to the presence of discrete variables. We solve this problem by exhaustively computing all outcomes for all possible school-input configurations. This is computationally feasible since, by construction, there is only one program school in each village, and our structural model allows us to solve for enrollment, test scores, and costs for every possible input configuration in program schools. We assume that the inputs of other schools remained constant as the program school's inputs adjusted. We think this is reasonable, as the primary competition for most program schools were government schools, which were centrally regulated by provincial and district education administrations, and did not adjust inputs across program and control villages.

Table 13 reports the levels and changes in school inputs across the solutions offered by program-school operators and the social planner. Program-school operators have proven remarkably successful at establishing and operating schools that generated most of the possible surplus in the environment. Assuming a social value of education equal to one, the social planner's solution generates gains of slightly less than 10 percent relative to that of program-school operators. Large variation exists across villages, from a lower bound of a zero-percent increase (i.e., the program-school operator selected the same set of school inputs as the social planner would have) to an upper bound of a 33-percent increase. The social planner achieves these gains through various changes to program schools. First, under the social planner, all program schools have toilets and/or drinking water (+10 percentage points relative to the program-school operator solution). The social planner employs teachers with post-secondary education only (+52 percentage points), and with less than

[19] Estimates are based on data from the 2010–11 Pakistan Social and Living Standards Measurement Survey (Government of Pakistan, 2011).

five years of teaching experience (+16 percentage points); and requires all teachers to be absent fewer than two days per month (-37 percentage points). The composition of female teachers is relatively unchanged, the social planner employing 47 percent female teachers (-3 percentage points), with substantial variation across villages. Differences are driven by the gender composition of the child population. Specifically, in villages with a relatively large number of boys, enrollment, and consequently, test scores, will suffer if the school employs female teachers, while the opposite is true in villages with a relatively large number of girls.

To understand why the social planner chooses these inputs, Table 13 also reports the changes in consumer surplus, enrollment, input costs, and test scores. On average, the social planner chooses inputs that lower costs. While total costs decrease, test scores increase dramatically. This results from higher school enrollment under the social planner, averaging 48 more students; and from higher test scores resulting from the interactions among teachers, other school inputs, and students. The better match quality between students and schools is reflected in the gains to consumer surplus, which are large and uniformly positive across all villages. Finally, higher test scores have substantial income effects, which translate into higher social welfare.

One of the key parameters in our social planner solution is the social value of education. This parameter does not come from any empirical or model-based foundation. Therefore, we are interested in understanding how robust our results are when we vary the social value of education. In Table 13, columns (3) through (6) report the results when the social planner places weights of 0, 0.5, 1.5, and 2, respectively, on test scores. The optimal education and teaching experience of teachers are invariant to the social value of education. In contrast, the optimal provision of toilets and/or drinking water falls, while the optimal levels of teacher absence and percentage of teachers who are female increase, as the social value of education falls to zero. Interestingly, the program configuration generates a social surplus closest to that achieved by the social planner when a weight of one is assigned by the latter to the social value of education.

Two aspects of the above calculations deserve emphasis. First, because men have higher labor force participation and labor earnings than women, factors improving enrollment and test scores for boys are given greater weight than those that raise them for girls. This can be seen with respect to female teachers, where increases in the social return to education lead to a decline in the optimal percentage of female teachers, which is driven by the lower preference for female teachers by boys. Because the model being used is static, it does not account for the possibility that female labor earnings and labor force participation may increase over time, potentially due to the very increases in enrollment and test scores found in this study. Second, the social planner's solution is village-specific. This means that, while the statistics given in Table 13 ostensibly indicate that program-school operators have provided inputs similar to those arrived at in the social planner's solution, the similarity in mean inputs does not necessarily imply that the village-specific solutions are similarly close.

9 Conclusion

The program evaluated in this study has proven remarkably effective in increasing school enrollment and test scores, measured after 1.5 school years. Introduced into educationally underserved villages, the program increased school enrollment by 30 percentage points, and total test scores by 0.63 standard deviations. For children induced by the program to enroll in school, the impact on total test scores was two standard deviations. Program impacts on school enrollment and test scores did not differ by gender, or by the subsidy treatment. We do not find that the gender-differentiated subsidy treatment had larger impacts on girls' enrollment or test scores than the gender-uniform one. Program-village households were more likely to express aspirations that their boys become doctors and engineers, rather than security personnel; and that their girls become teachers, rather than housewives. Program-village households also voiced a desire for their boys and girls to attain higher levels of education.

The study also assesses the effectiveness and efficiency of program schools. We find that program-school students had higher test scores than government-school students, despite coming from more socioeconomically disadvantaged households. With respect to efficiency, while program-school operators only captured profits through enrollment, the equilibrium social surplus is within 10 percentage points of the social planner. Compared to program-school operators, the social planner adjusts the gender ratio of teachers to better match the gender ratio of children in the village, ensures toilets and/or drinking water in the school, and hires less-experienced, but more-educated, teachers. It is remarkable and reassuring that program-school operators have proven so successful in selecting the most essential inputs for their schools. Our results contribute to the growing literature on the private provision of public goods by demonstrating that it is possible for profit-oriented local entrepreneurs to provide high-quality, low-cost educational solutions.

References

Adukia, A. (2017). Sanitation and education. *American Economic Journal: Applied Economics*, 9(2):23–59.

Alderman, H., Kim, J., and Orazem, P. F. (2003). Design, evaluation, and sustainability of private schools for the poor: the pakistan urban and rural fellowship school experiments. *Economics of Education review*, 22(3):265–274.

Alderman, H., Orazem, P. F., and Paterno, E. M. (2001). School quality, school cost, and the public/private school choices of low-income households in pakistan. *Journal of Human Resources*, pages 304–326.

Amjad, R. and MacLeod, G. (2014). Academic effectiveness of private, public and private–public partnership schools in pakistan. *International Journal of Educational Development*, 37:22–31.

Andrabi, T., Bau, N., Das, J., and Khwaja, A. I. (2010). Are bad public schools public 'bads'? test scores and civic values in public and private schools. *(Working Paper No. 11904)*.

Andrabi, T., Das, J., and Khwaja, A. I. (2008a). 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., Khwaja, A. I., Vishwanath, T., Zajonc, T., and the LEAPS Team (2008b). *Pakistan: Learning and Educational Achievements in Punjab Schools (LEAPS): Insights to Inform the Education Policy Debate.* JSTOR.

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., Bettinger, E., Bloom, E., King, E., and Kremer, M. (2002). Vouchers for private schooling in colombia: Evidence from a randomized natural experiment. *The American Economic Review*, 92(5):1535–1558.

Barham, T., Macours, K., and Maluccio, J. A. (2013). More schooling and more learning? effects of a three-year conditional cash transfer program in nicaragua after 10 years. Technical report, IDB Working Paper Series.

Barrera-Osorio, F. and Filmer, D. (2016). Incentivizing schooling for learning: Evidence on the impact of alternative targeting approaches. *Journal of Human Resources*, 51(2).

Barrera-Osorio, F., Galbert, D., Gaspard, P., Habyarimana, J. P., and Sabarwal, S. (2016). Impact of public-private partnerships on private school performance: evidence from a randomized controlled trial in uganda.

Barrera-Osorio, F., Linden, L. L., and Urquiola, M. (2007). The effects of user fee reductions on enrollment: Evidence from a quasi-experiment. *Manuscript. Department of Economics. The University of Texas at Austin*.

Barrera-Osorio, F. and Raju, D. (2015). Evaluating the impact of public student subsidies on low-cost private schools in pakistan. *The Journal of Development Studies*, 51(7):808–825.

Bau, N. and Das, J. (2017). The misallocation of pay and productivity in the public sector: Evidence from the labor market for teachers.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy*, 70(5, Part 2):9–49.

Berlinski, S., Galiani, S., and Gertler, P. (2009). The effect of pre-primary education on primary school performance. *Journal of Public Economics*, 93(1):219–234.

Bold, T., Kimenyi, M., Mwabu, G., and Sandefur, J. (2014). Can free provision reduce demand for public services? evidence from kenyan education. *The World Bank Economic Review*, 29(2):293–326.

Borkum, E. (2012). Can eliminating school fees in poor districts boost enrollment? evidence from south africa. *Economic Development and Cultural Change*, 60(2):359–398.

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

Carniero, P., Das, J., and Reis, H. R. (2016). [problem with this paper]the value of private schools: Evidence from pakistan. *(Discussion Paper 9960)*.

Deininger, K. (2003). Does cost of schooling affect enrollment by the poor? universal primary education in uganda. *Economics of Education review*, 22(3):291–305.

Desai, S., Dubey, A., Vanneman, R., Banerji, R., et al. (2009). Private schooling in india: A new educational landscape. In *India Policy Forum*, volume 5, pages 1–38. National Council of Applied Economic Research.

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.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *The American Economic Review*, 91(4):795–813.

Evans, D. K. and Popova, A. (2016). Cost-effectiveness analysis in development: Accounting for local costs and noisy impacts. *World Development*, 77:262–276.

Fiszbein, A. and Schady, N. R. (2009). *Conditional cash transfers: reducing present and future poverty*. World Bank Publications.

Foster, A. D. and Rosenzweig, M. R. (1996). Technical change and human-capital returns and investments: evidence from the green revolution. *The American Economic Review*, pages 931–953.

French, R. and Kingdon, G. (2010). The relative effectiveness of private and government schools in rural india: Evidence from aser data. *London: Institute of Education*.

Friedman, M. (1955). *The role of government in education*. Rutgers University Press.

Government of Pakistan (2011). *Pakistan Social and Living Standards Measurement Survey 2010-11*. Pakistan Bureau of Statistics, Government of Pakistan.

Hart, O., Shleifer, A., and Vishny, R. W. (1997). The proper scope of government: theory and an application to prisons. *The Quarterly Journal of Economics*, 112(4):1127–1161.

Hoxby, C. M. (2003). School choice and school productivity. could school choice be a tide that lifts all boats? In *The economics of school choice*, pages 287–342. University of Chicago Press.

Kim, J., Alderman, H., and Orazem, P. F. (1999). Can private school subsidies increase enrollment for the poor? the quetta urban fellowship program. *The World Bank Economic Review*, 13(3):443–465.

Lloyd, C. B., Mete, C., and Sathar, Z. A. (2005). The effect of gender differences in primary school access, type, and quality on the decision to enroll in rural pakistan. *Economic Development and Cultural Change*, 53(3):685–710.

Lucas, A. M. and Mbiti, I. M. (2012). Access, sorting, and achievement: the short-run effects of free primary education in kenya. *American Economic Journal: Applied Economics*, 4(4):226–253.

MacLeod, W. B. and Urquiola, M. (2012). Competition and educational productivity: incentives writ large. In Glewwe, P., editor, *Education Policy in Developing Countries*. The University of Chicago Press.

Montenegro, C. E. and Patrinos, H. A. (2014). Comparable estimates of returns to schooling around the world. *World Bank policy research working paper*, (7020).

Muralidharan, K. and Prakash, N. (2017). Cycling to school: Increasing secondary school enrollment for girls in india. *American Economic Journal: Applied Economics*, 9(3):321–350.

Muralidharan, K. and Sheth, K. (2016). Bridging education gender gaps in developing countries: The role of female teachers. *Journal of Human Resources*, 51(2):269–297.

Muralidharan, K. and Sundararaman, 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.

Newhouse, D. and Beegle, K. (2006). The effect of school type on academic achievement evidence from indonesia. *Journal of Human Resources*, 41(3):529–557.

Nguyen, Q. T. and Raju, D. (2015). Private school participation in pakistan. *Lahore Journal of Economics*, 20(1):1–46.

Patrinos, H. A., Barrera-Osorio, F., and Guáqueta, J. (2009). *The role and impact of public-private partnerships in education*. World Bank Publications.

Pritchett, L. (2013). *The rebirth of education: Schooling ain't learning*. CGD Books.

Romero, M., Sandefur, J., Sandholtz, W. A., et al. (2017). Can outsourcing improve liberia?s schools? preliminary results from year one of a three-year randomized evaluation of partnership schools for liberia. Technical report.

Saavedra, J. E. and García, S. (2012). Impacts of conditional cash transfer programs on educational outcomes in developing countries: a meta-analysis. *RAND Labor and Population Working Paper Series, WR-921-1.*

Shleifer, A. (1998). State versus private ownership. *Journal of economic perspectives*, 12(4):133–150.

Singh, A. (2015). The private school premium: Size and sources of the private school advantage in test scores in india. *Journal of Development Economics*, 113:16–32.

Small, K. A. and Rosen, H. S. (1981). Applied welfare economics with discrete choice models. *Econometrica: Journal of the Econometric Society*, pages 105–130.

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

United Nations Educational, S. and (UNESCO), C. O. (2014). *Teaching and Learning: Achieving Quality for All*. UNESCO, Paris, France.

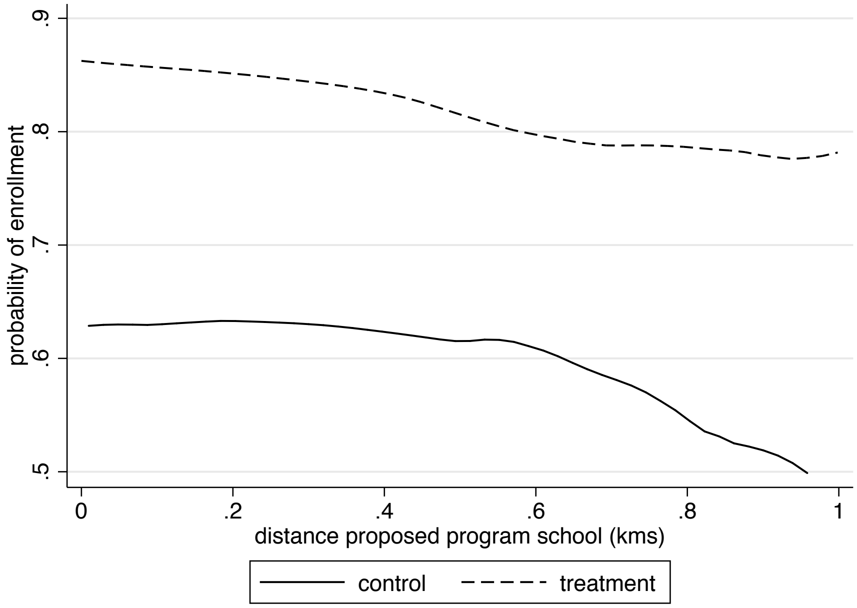
United Nations Educational, S. and (UNESCO), C. O. (2015a). *Achievements and Challenges: Gender Summary*. UNESCO, Paris, France.

United Nations Educational, S. and (UNESCO), C. O. (2015b). *Education for All: Achievements and Challenges*. UNESCO, Paris, France.

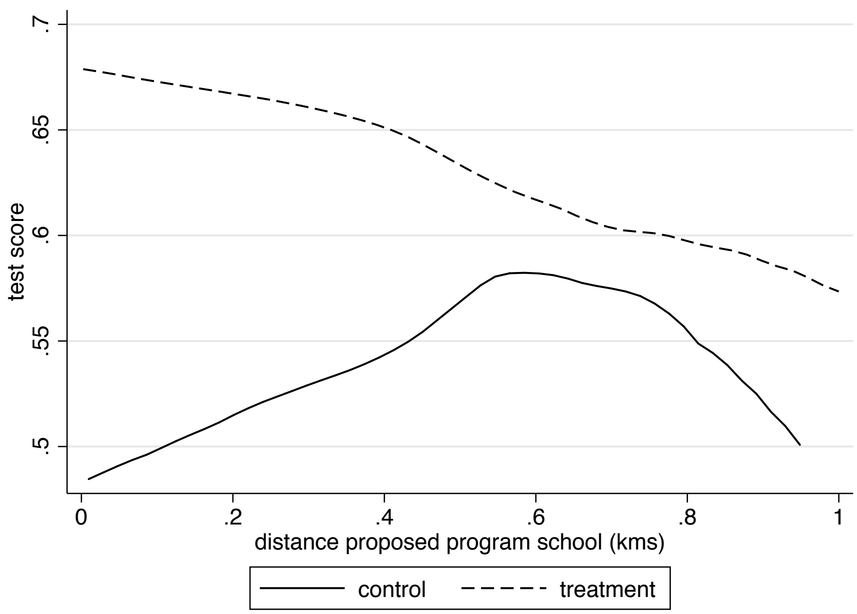
World Bank (2004). *World Development Report 2004: Making Services Work for Poor People*. World Bank Publications.

World Bank (2018). *World Development Report 2018: LEARNING to Realize Educations Promise*. World Bank Publications.

Figure 1: School proximity and enrollment



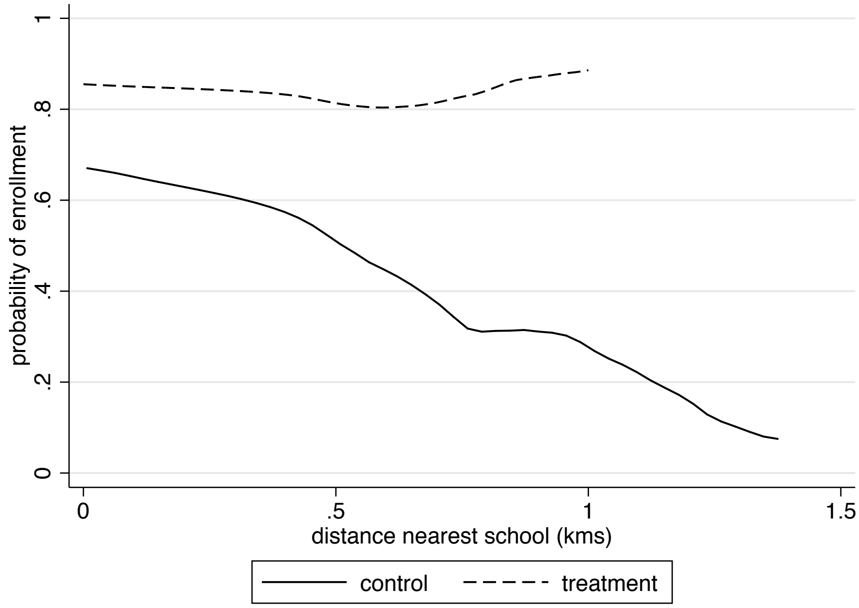
1.1: Enrollment



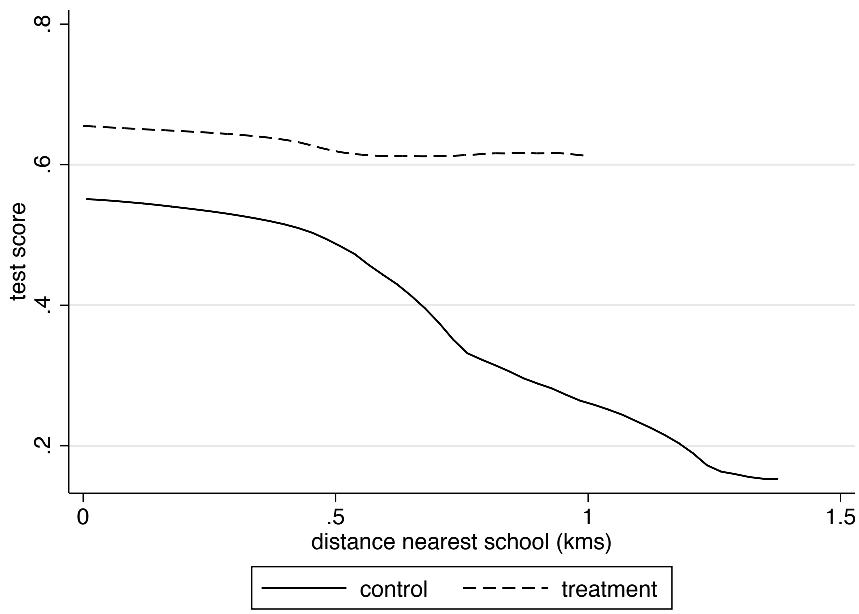
1.2: Test Score

Notes: Figure 1.1 plots the probability of enrollment against the distance to the proposed program school site, disaggregated by treatment status. Figure 1.2 plots the test score against the distance to the proposed program school site, disaggregated by treatment status.

Figure 2: School proximity and educational outcomes



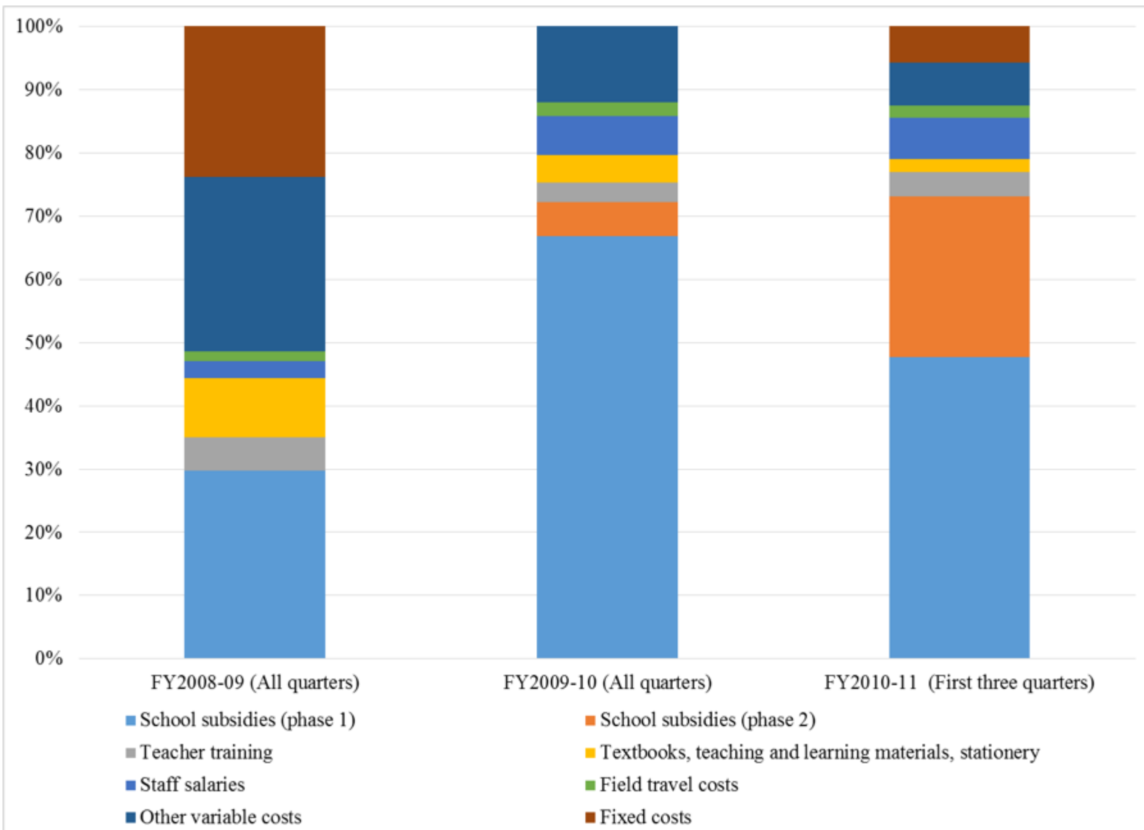
2.1: Enrollment



2.2: Test Score

Notes: Figure 2.1 plots the probability of enrollment against the distance to the nearest school, disaggregated by treatment status. Figure 2.2 plots the test score against the distance to the nearest school, disaggregated by treatment status.

Figure 3: Distribution of program costs over the evaluation period



Note: FY denotes fiscal year, which runs from July 1 to June 30. School subsidies (phase-1) reflects per-student public subsidies offered to program schools under this evaluation.

Table 1: Evaluation sample sizes

	control	treat_p	treat_gu	treat_gd	total
	(1)	(2)	(3)	(4)	(5)
num villages	38	161	82	79	199
num baseline households	445	1644	823	821	2089
num baseline young children	1141	4415	2261	2154	5556
num followup households	1069	4897	2594	2303	5966
num followup young children	3093	14627	7717	6910	17720

Note: This table reports sample sizes by treatment status. treat_p denotes pooled treatment; treat_gu, the gender-uniform subsidy treatment; and treat_gd, the gender-differentiated subsidy treatment.

Table 2: Balance across program and control villages

	baseline		followup	
		treat_p -		treat_p -
	control	control	control	control
	(1)	(2)	(5)	(6)
child age	6.859	-0.023	7.359	0.075
		(0.071)		(0.055)
child female	0.379	0.042*	0.425	0.030*
		(0.024)		(0.017)
child enrolled at baseline	0.261	0.008	0.284	-0.028
		(0.046)		(0.085)
child of hh head			0.857	0.022
				(0.026)
household size	9.858	-0.833	7.221	-0.088
		(0.563)		(0.290)
number children	3.018	-0.257	4.757	-0.133
		(0.166)		(0.189)
hh head education	2.571	0.252	2.650	0.111
		(0.398)		(0.315)
hh head farmer	0.613	0.030	0.562	-0.017
		(0.062)		(0.067)
total land			4.254	0.890
				(1.124)
pukka house			0.057	-0.005
				(0.024)
semi-pukka house			0.193	-0.016
				(0.065)
kaccha house			0.511	0.084
				(0.076)
thatched huts			0.240	-0.064
				(0.071)
goats			3.916	-0.052
				(0.793)
sunni			0.877	0.034
				(0.060)
urdu			0.114	0.044
				(0.043)
sindhi			0.664	0.062
				(0.071)
				(0.071)

Note: This table reports balance in characteristics across program and control villages. Columns (1) and (3) report mean child and household characteristics in control villages at baseline and follow-up, respectively. Columns (2) and (4) report differences in mean child and household characteristics in program villages at baseline and follow-up, respectively. treat_p denotes pooled treatment. Standard errors, reported in parentheses, are clustered at the village level. Statistical significance at the one-, five-, and ten-percent levels denoted by ***, **, and *, respectively.

Table 3: Program impacts on enrollment

	reported enrollment				verified enrollment	highest grade attained
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: children aged 6-10						
treat_p	0.315***	0.315***	0.312***	0.315***	0.294***	0.381***
	(0.066)	(0.066)	(0.064)	(0.065)	(0.041)	(0.120)
N	11571	11571	11571	11571	10285	11116
R-squared	0.086	0.087	0.103	0.108	0.103	0.224
Panel B: children aged 11-17						
treat_p	0.109*	0.112*	0.108**	0.110**		-0.016
	(0.057)	(0.058)	(0.049)	(0.052)		(0.319)
N	5583	5583	5583	5583		5360
R-squared	0.006	0.039	0.097	0.148		0.133
child controls	no	yes	yes	yes	yes	yes
HH controls	no	no	yes	yes	yes	yes
district fixed effects	no	no	no	yes	yes	yes

Note: This table reports program impacts on enrollment and highest grade attained at follow-up measurement. treat_p denotes pooled treatment. Standard errors, reported in parentheses, are clustered at the village level. Statistical significance at the one-, five-, and ten-percent levels denoted by ***, **, and *, respectively.

Table 4: Program impacts on test scores

	ITT				TOT	
					reported	verified
	(1)	(2)	(3)	(4)	(5)	(6)
math	0.528***	0.518***	0.517***	0.625***	1.945***	2.015***
	(0.153)	(0.156)	(0.154)	(0.123)	(0.282)	(0.451)
language	0.495***	0.487***	0.485***	0.587***	1.795***	1.943***
	(0.170)	(0.173)	(0.170)	(0.128)	(0.227)	(0.435)
total	0.530***	0.520***	0.519***	0.626***	1.930***	2.046***
	(0.164)	(0.167)	(0.165)	(0.127)	(0.259)	(0.457)
N	10327	10327	10327	10327	10071	9014
child controls	no	yes	yes	yes	yes	yes
HH controls	no	no	yes	yes	yes	yes
district fixed effects	no	no	no	yes	yes	yes

Note: This table reports program impacts on standardized test scores. Columns (1) through (4) report the intention-to-treat (ITT) impacts, with various sets of controls. Columns (5) and (6) report the treatment-on-the-treated (TOT) impacts on test scores, based on reported and verified enrollment, respectively. Standard errors, reported in parentheses, are clustered at the village level. Statistical significance at the one-, five-, and ten-percent levels denoted by ***, **, and *, respectively.

Table 5: Differential impacts by subsidy treatment and by gender

	enrollment		highest grade	test
	reported	verified	attained	scores
	(1)	(2)	(3)	(4)
Panel A: subsidy treatments				
treat_gu	0.316***	0.266***	0.369***	0.609***
	(0.065)	(0.045)	(0.124)	(0.132)
treat_gd - treat_gd	-0.001	0.058	0.027	0.038
	(0.024)	(0.039)	(0.067)	(0.062)
N	11571	10285	11116	10323
R-squared	0.108	0.106	0.224	0.202
Panel B: gender				
treat_p	0.324***	0.293***	0.393***	0.601***
	(0.066)	(0.042)	(0.130)	(0.133)
treat_p X female	-0.018	0.005	-0.024	0.060
	(0.027)	(0.025)	(0.065)	(0.053)
N	11521	10240	11066	10282
R-squared	0.108	0.103	0.225	0.201

Note: This table reports program impacts on outcomes by subsidy treatment (panel A), and by gender (panel B), with the full set of child and household controls and district fixed effects. treat_p denotes pooled treatment; treat_gu, the gender-uniform subsidy treatment; and treat_gd, the gender-differentiated subsidy treatment. Standard errors, reported in parentheses, are clustered at the village level. Statistical significance at the one-, five-, and ten-percent levels denoted by ***, **, and *, respectively.

Table 6: Gender differential impacts by subsidy treatment

		enrollment		highest grade	test
		reported	verified	attained	scores
		(1)	(2)	(3)	(4)
treat_gu		0.332***	0.261***	0.410***	0.574***
		(0.066)	(0.047)	(0.136)	(0.136)
treat_gu X female		-0.035	0.013	-0.091	0.081
		(0.031)	(0.028)	(0.081)	(0.054)
treat_gd		0.315***	0.329***	0.374***	0.633***
		(0.068)	(0.046)	(0.135)	(0.137)
treat_gd X female		0.000	-0.006	0.049	0.034
		(0.027)	(0.031)	(0.063)	(0.059)
N		11521	10240	11066	10282
R-squared		0.109	0.106	0.225	0.202
H0: treat_gu = treat_gd	F-stat	0.544	2.639	0.249	0.895
	p-value	0.462	0.106	0.618	0.345
H0: treat_gu + treat_gu X female =	F-stat	0.397	1.205	1.779	0.033
treat_gd + treat_gd X female	p-value	0.529	0.274	0.184	0.857
H0: treat_gu X female =	F-stat	2.831	0.355	4.123	1.276
treat_gd X female	p-value	0.094	0.552	0.044	0.260

Note: This table reports gender-differential impacts on outcomes by subsidy treatment, with the full set of child and household controls and district fixed effects. treat_gu denotes the gender-uniform subsidy treatment; and treat_gd, the gender-differentiated subsidy treatment. Standard errors, reported in parentheses, are clustered at the village level. Statistical significance at the one-, five-, and ten-percent levels denoted by ***, **, and *, respectively.

Table 7: Program impacts on aspirations

	control	treat_p - control	female	treat_p	treat_p X female
	(1)	(2)	(3)	(4)	(5)
Panel A: household aspirations					
civil servant	0.127	0.030	-0.060	0.050	-0.027
		(0.036)	(0.047)	(0.048)	(0.049)
doctor	0.082	0.047***	-0.005	0.058***	-0.025
		(0.018)	(0.022)	(0.019)	(0.025)
private enterprise	0.024	-0.005	-0.019**	-0.009	0.012
		(0.012)	(0.009)	(0.015)	(0.011)
engineer	0.013	0.025***	-0.016**	0.026***	0.006
		(0.007)	(0.007)	(0.009)	(0.010)
farmer	0.105	-0.044*	-0.144***	-0.061	0.056
		(0.025)	(0.031)	(0.038)	(0.035)
housewife	0.179	-0.048**	0.409***	-0.003	-0.146***
		(0.023)	(0.043)	(0.010)	(0.049)
laborer	0.028	-0.010	-0.023**	-0.004	-0.001
		(0.008)	(0.010)	(0.010)	(0.011)
landlord	0.013	0.004	-0.017*	0.004	0.000
		(0.006)	(0.009)	(0.010)	(0.010)
lawyer	0.004	0.009**	-0.007*	0.009*	0.002
		(0.003)	(0.003)	(0.005)	(0.005)
police/army/security	0.098	-0.031	-0.101***	-0.050*	0.042*
		(0.020)	(0.022)	(0.026)	(0.023)
raise livestock	0.018	-0.009	0.002	-0.007	-0.008
		(0.011)	(0.012)	(0.010)	(0.012)
teacher	0.248	0.026	0.027	-0.012	0.077**
		(0.028)	(0.029)	(0.025)	(0.035)
marriage age	18.496	0.254	-1.019**	0.331	-0.160
		(0.439)	(0.413)	(0.456)	(0.448)
education attainment (in years)	7.428	1.537**	-0.829**	1.466**	0.242
		(0.606)	(0.396)	(0.682)	(0.458)
Panel B: child aspirations					
army	0.083	-0.031	-0.085	-0.068	0.054
		(0.044)	(0.060)	(0.098)	(0.066)
doctor	0.224	0.030	-0.027	0.094	0.066
		(0.055)	(0.093)	(0.074)	(0.108)
farmer	0.019	-0.019	0.011	-0.032	-0.011
		(0.013)	(0.054)	(0.033)	(0.054)
government	0.028	0.041**	0.000	0.122***	-0.112***
		(0.021)	(0.000)	(0.034)	(0.036)
other	0.068	-0.008	-0.093	0.002	0.064
		(0.052)	(0.079)	(0.084)	(0.084)
private	0.169	-0.003	-0.007	-0.063	0.083
		(0.068)	(0.131)	(0.099)	(0.146)
teacher	0.379	-0.002	0.301**	0.036	-0.241
		(0.085)	(0.149)	(0.128)	(0.165)
education attainment (in years)	11.258	-0.203	-0.381	-0.262	0.496
		(0.376)	(0.440)	(0.588)	(0.514)

Note: This table reports program impacts on household-reported aspirations for the child (panel A) and child-reported aspirations (panel B), with the full set of child and household controls and district fixed effects. Column (1) reports mean aspirations in control villages, and column (2) reports differences in mean aspirations between program and control villages. Columns (3), (4), and (5) report coefficients from a regressions of an indicator variable for girls, program status, and the interaction of the two. treat_p denotes pooled-treatment. Standard errors, reported in parentheses, are clustered at the village level. Statistical significance at the one-, five-, and ten-percent levels denoted by ***, **, and *, respectively.

Table 8: Characteristics by school type

	program	program - govt	program - private		program	program - govt	program - private
	(1)	(2)	(3)		(4)	(5)	(6)
Characteristics from school survey				number boys	88.711	18.695	-42.460
days operational	5.118	0.481*	0.326			(11.366)	(55.158)
		(0.285)	(0.572)	number girls	71.294	31.715***	-18.743
open admission	0.858	-0.025	-0.075			(5.714)	(29.671)
		(0.050)	(0.074)	percent female students	0.448	0.046	0.010
uniform required	0.024	0.024	-0.312*			(0.048)	(0.051)
		(0.017)	(0.181)	student-teacher ratio	44.250	0.274	5.935
tuition required	0.000	0.000	-0.441**			(3.967)	(7.623)
		(0.000)	(0.180)				
medium: sindhi	0.612	-0.370***	-0.028	**Characteristics from teacher survey**			
		(0.050)	(0.188)	days absent/month	0.838	-0.184	0.324
medium: english	0.310	0.310***	-0.005			(0.313)	(0.274)
		(0.045)	(0.187)	female	0.492	0.243***	-0.001
number teachers						(0.072)	(0.182)
total	3.782	0.967***	-2.501	age	24.196	-13.622***	-0.473
		(0.332)	(2.009)			(1.452)	(1.354)
female	1.986	1.462***	-3.407**	education	10.642	-0.538	-1.157***
		(0.198)	(1.654)			(0.337)	(0.305)
postsecondary	1.898	-0.426	-1.476*	salary (1000's rupees)	4.066	-11.454***	0.254
		(0.460)	(0.845)			(1.017)	(0.551)
<5yrs exp	3.132	2.474***	0.945	years teaching	2.784	-11.320***	-0.676
		(0.181)	(0.682)			(1.310)	(0.773)
5-10yrs exp	0.603	0.405***	-3.092	years teaching same school	1.774	-4.897***	-0.927
		(0.122)	(2.355)			(0.966)	(0.733)
>10 yrs exp	0.047	-1.954***	-0.354	hours teaching			
		(0.301)	(0.396)	total	25.253	-0.249	-1.630
avg teacher absent $\geq$ 2 days/month	0.396	-0.039	0.129			(2.123)	(1.205)
		(0.101)	(0.168)	teaching whole class	5.164	0.132	1.088
						(0.776)	(0.753)
building	0.961	0.024	-0.039*	teaching small group	3.925	0.562*	0.125
		(0.039)	(0.020)			(0.321)	(0.728)
number classrooms	3.229	0.501	0.115	teaching individual	3.738	-0.100	0.064
		(0.371)	(0.925)			(0.382)	(0.661)
sufficient desks	0.756	0.144*	0.173	blackboard/dictation	3.640	0.322	0.828
		(0.084)	(0.180)			(0.495)	(0.517)
drinking water	0.846	0.307***	-0.154***	classroom management	2.239	-0.124	-0.814**
		(0.105)	(0.037)			(0.187)	(0.327)
electricity	0.724	0.063	-0.005	testing	2.438	0.938***	0.642*
		(0.071)	(0.152)			(0.336)	(0.383)
toilet	0.788	0.291***	0.162	administrative	2.023	-0.316	0.419
		(0.109)	(0.178)			(0.392)	(0.300)

Note: This table reports differences in mean characteristics between program and government schools, and between program and private schools. The unit of observation is child-school. Columns (1) and (4) report means for program schools; columns (2) and (5), differences in means between program and government schools; and columns (3) and (6), differences in means between program and private schools. Standard errors, reported in parentheses, are clustered at the village level. *, **, and *** denote statistical significance at the ten-, five-, and one-percent levels, respectively.

Table 9: Test Scores by school type

	program	program– govt	program– priv	p-value govt = priv	govt-enrolled treatment– control
	(1)	(2)	(3)	(4)	(5)
math score	0.715	0.236***	0.049	-0.193	-0.053
		(0.081)	(0.226)	(0.180)	(0.087)
urdu score	0.709	0.160***	0.025	-0.189	-0.015
		(0.061)	(0.125)	(0.117)	(0.075)
total score	0.734	0.211***	0.040	-0.205	-0.040
		(0.073)	(0.188)	(0.151)	(0.080)

Note: This table reports differences in mean standardized test scores according to the type of school children are enrolled in, controlling for student characteristics and district fixed effects. Column (1) reports mean test scores for children enrolled in program schools; column (2), differences in means between program- and government-enrolled children; column (3), differences in means between program- and private-enrolled children; and column (4) the p-value for a test of equality of government and private school coefficients. Column (5) reports the difference in test score across control and treatment villages for government-enrolled children. Standard errors, reported in parentheses, are clustered at the village level. *, **, and *** denote statistical significance at the ten-, five-, and one-percent levels, respectively.

Table 10: Education production estimates, total test scores

	(1)	(2)	(3)
toilets and/or drinking water	0.138	0.148	0.152
	(0.102)	(0.106)	(0.110)
female	0.029	0.030	0.039
	(0.046)	(0.046)	(0.047)
age	0.103***	0.103***	0.103***
	(0.011)	(0.011)	(0.011)
tuition required	0.001	0.001	0.000
	(0.000)	(0.000)	(0.001)
distance from home to school	-0.028	-0.030	-0.027
	(0.058)	(0.058)	(0.057)
pct teachers $<$5yrs exp	0.215**	0.243**	0.237**
	(0.100)	(0.109)	(0.103)
pct teachers postsecondary	0.197*	0.190*	0.187*
	(0.102)	(0.102)	(0.106)
pct teachers female	0.088	0.093	0.093
	(0.087)	(0.088)	(0.089)
pct time teaching	-0.173	-0.178	-0.171
	(0.300)	(0.297)	(0.318)
avg teacher absent $\geq$2 days/month	-0.045	-0.043	-0.044
	(0.068)	(0.067)	(0.068)
pct teachers female X female student	0.037	0.037	0.030
	(0.045)	(0.045)	(0.045)
distance X female student	0.005	0.003	0.002
	(0.023)	(0.023)	(0.023)
toilets and/or drinking water X female student	-0.062	-0.063	-0.066
	(0.051)	(0.051)	(0.052)
government school		0.058	
		(0.110)	
program school			-0.047
			(0.099)
R-squared	0.075	0.076	0.075
N	7182	7182	7098

Note: This table reports education production estimates, relating standardized total test scores to school inputs and student characteristics, controlling for district fixed effects. Column (1) reports estimates without indicator variables for program or government schools; column (2), with an indicator variable for government schools; and column (3), with an indicator variable for program schools. Standard errors, reported in parentheses, are clustered at the village level. *, **, and *** denote statistical significance at the ten-, five-, and one-percent levels, respectively.

Table 11: Schooling Demand Estimates

	(1)	(2)	(3)	(4)
constant	0.306***	-0.141**	-0.013	1.365***
	(0.091)	(0.080)	(0.067)	(0.141)
toilets and/or drinking water	0.841***	0.904***	0.882***	0.567***
	(0.062)	(0.065)	(0.087)	(0.095)
student female	0.012	0.030	-0.213**	-0.232***
	(0.048)	(0.055)	(0.111)	(0.115)
student age	0.037***	0.041***	0.040***	0.035***
	(0.012)	(0.012)	(0.012)	(0.013)
distance from home to school	-0.131***	-0.136***	-0.102***	-0.051**
	(0.033)	(0.030)	(0.041)	(0.031)
pct teachers with $<$5yrs exp		0.792***	0.795***	-0.133**
		(0.067)	(0.073)	(0.074)
pct teachers post-secondary		-0.252***	-0.252***	-0.044
		(0.054)	(0.060)	(0.064)
pct teachers female		-0.461***	-0.703***	-0.859***
		(0.055)	(0.054)	(0.068)
pct time teaching		0.263*	0.255***	-0.047
		(0.178)	(0.061)	(0.144)
avg teacher absent $\geq$2 days/month		-0.113***	-0.113***	-0.148***
		(0.041)	(0.048)	(0.041)
pct female teachers X female student			0.528***	0.562***
			(0.070)	(0.105)
distance X female student			-0.061	-0.020
			(0.057)	(0.050)
toilets and/or drinking water X female student			0.022	0.002
			(0.133)	(0.118)
tuition cost per year	-0.008***	-0.006***	-0.006***	-0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
govt school				-1.473***
				(0.062)

Note: This table reports schooling demand estimates. Columns (1) and (2) exclude and include an indicator variable for government schools, respectively. Standard errors, reported in parentheses, are clustered at the village level. *, **, and *** denote statistical significance at the ten-, five-, and one-percent levels, respectively.

Table 12: Cost estimates

toilets and/or drinking water	3.604***
	(0.265)
teacher female	-4.249***
	(0.243)
post-secondary	-0.404
	(0.377)
$<$5 yrs experience	-1.379***
	(0.483)
avg teacher absent $\geq$2 days/month	-1.332***
	(0.234)

Note: This table reports cost estimates. Standard errors, reported in parentheses, are clustered at the village level. *, **, and *** denote statistical significance at the ten-, five-, and one-percent levels, respectively

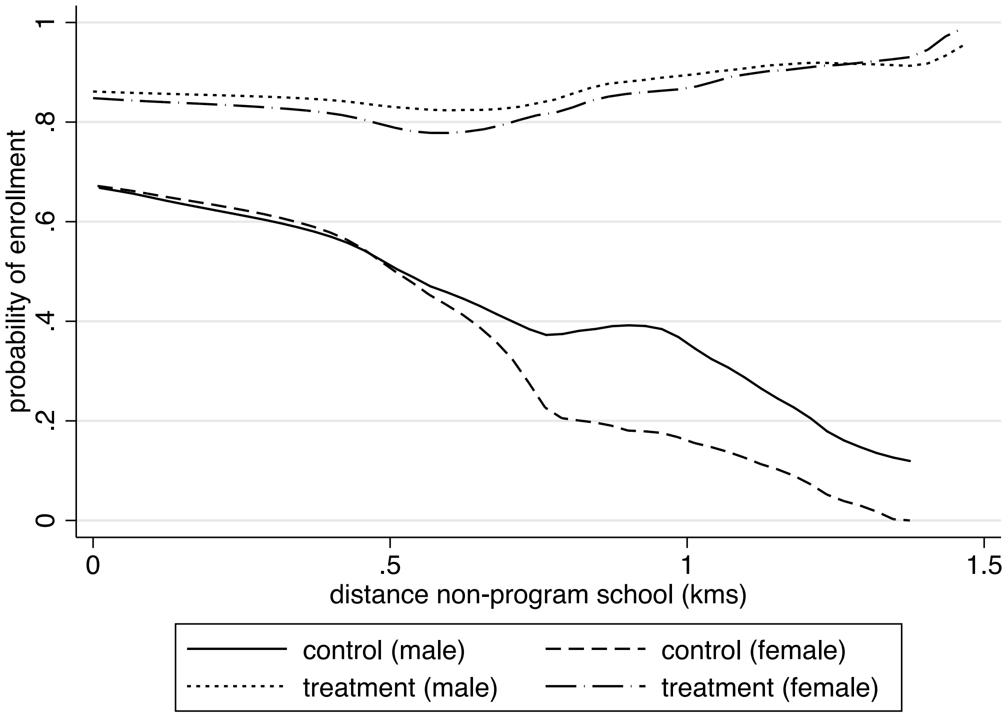
Table 13: Estimated social planner solution

	Program Solution	Social Planner Solution				
		externality				
		1	0	0.5	1.5	2
	(1)	(2)	(3)	(4)	(5)	(6)
toilets and/or drinking water	0.90	1.00	0.00	0.96	1.00	1.00
	(0.31)	(0.00)	(0.00)	(0.20)	(0.00)	(0.00)
pct teachers female	0.50	0.47	1.00	0.84	0.26	0.15
	(0.41)	(0.39)	(0.00)	(0.29)	(0.36)	(0.31)
pct teachers post-secondary	0.48	1.00	1.00	1.00	1.00	1.00
	(0.35)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
pct teachers <5yrs experience	0.84	1.00	1.00	1.00	1.00	1.00
	(0.25)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
avg teacher absent ≥ 2 days/month	0.37	0.00	1.00	0.07	0.00	0.00
	(0.48)	(0.00)	(0.00)	(0.26)	(0.00)	(0.00)
change in test scores		1458.17	1161.07	1433.57	1466.45	1469.45
change in cost		-662.63	-1463.81	-1044.79	-437.87	-319.49
change in consumer surplus		15345.22	10295.39	14026.50	15965.18	16251.96
change in enrollment		47.94	35.75	45.53	49.00	49.46
change in income (upper bound)		1240438.42	952332.50	1202761.21	1254648.59	1260248.59
change in income (lower bound)		781016.76	599616.74	757294.11	789963.88	793489.84
total surplus (upper bound)	1142581.63	1253427.58	69172.34	637142.02	1882273.07	2502540.44
total surplus (lower bound)	719306.00	790793.19	69172.34	412850.88	1180371.78	1577271.70

Note: This table presents the social planner's solution and the observed program solution.

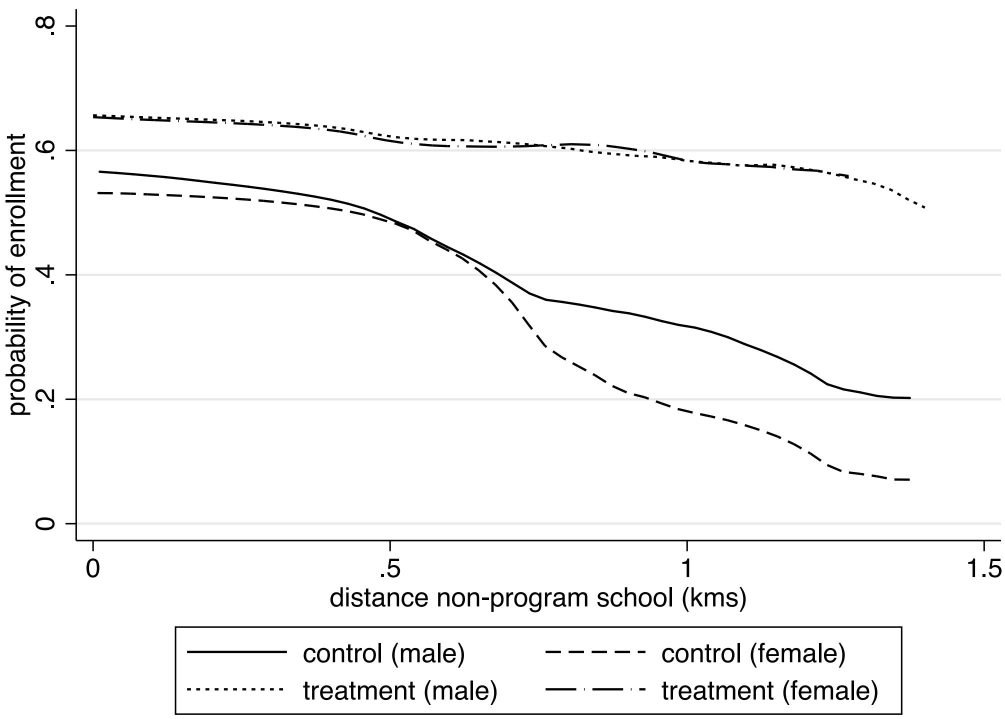
Appendix

Figure A1: Distance and enrollment, by gender



Notes: Figure A1 plots the probability of enrollment against the distance to the nearest school, disaggregated by treatment status and gender.

Figure A2: Distance and test score, by gender



Notes: Figure A2 plots the test score against the distance to the nearest school, disaggregated by treatment status and gender.

Table A.1: Balance Across Treatment Groups

	Baseline		Followup	
	treat_gu mean (1)	treat_gd-treat_gu (2)	treat_gu mean (5)	treat_gd-treat_gu (6)
child age	6.858	-0.044	9.421	-0.064
		(0.062)		(0.121)
female	0.413	0.015	0.436	0.010
		(0.018)		(0.012)
child in school	0.275	-0.013	0.292	0.001
		(0.042)		(0.062)
child of hh head			0.881	0.020
				(0.021)
household size	9.202	-0.364	7.294	0.107
		(0.438)		(0.228)
number children	2.760	0.001	4.793	-0.002
		(0.133)		(0.140)
hh head education	2.906	-0.169	2.690	0.093
		(0.342)		(0.291)
hh head farmer	0.648	-0.010	0.556	-0.044
		(0.047)		(0.047)
total land			5.656	-1.114
				(1.366)
pukka house			0.046	0.016
				(0.026)
semi-pukka house			0.194	-0.023
				(0.056)
kaccha house			0.604	-0.023
				(0.065)
thatched hut			0.156	0.030
				(0.068)
goats			3.878	0.256
				(0.834)
sunni			0.910	-0.012
				(0.047)
urdu			0.152	-0.004
				(0.046)
sindhi			0.710	0.060
				(0.059)

Note: This table reports balance in characteristics across villages under the gender-uniform and gender-differentiated subsidy treatments. Columns (1) and (3) report mean child and household characteristics in villages under the gender-uniform subsidy treatment at baseline and follow-up, respectively. Columns (2) and (4) report differences in mean child and household characteristics between villages under the gender-uniform and -differentiated subsidy treatments at baseline and follow-up, respectively. treat_gu denotes the gender-uniform subsidy treatment; and treat_gd, the gender-differentiated one. Standard errors, reported in parentheses, are clustered at the village level. Statistical significance at the one-, five-, and ten-percent levels denoted by ***, **, and *, respectively.

Table A.2: Program-school characteristics by subsidy treatment

	treat_gu mean (1)	treat_gd-treat_gu (2)		treat_gu mean (4)	treat_gd-treat_gu (5)
Characteristics from school survey			number boys	90.594	-3.786
days operational	5.088	0.069			(9.397)
		(0.246)	number girls	71.019	0.555
open admission	0.881	-0.046			(7.567)
		(0.064)	pct female students	0.445	0.007
uniform required	0.047	-0.047			(0.031)
		(0.033)	student-teacher ratio	44.752	-1.010
tuition required	0.000	0.000			(3.072)
		(0.000)			
medium: sindhi	0.669	-0.116	**Characteristics from teacher survey**		
		(0.096)	days absent/month	0.864	-0.053
medium: english	0.257	0.108			(0.222)
		(0.090)	female	0.502	-0.020
number of teachers					(0.087)
total	3.654	0.261	age	25.210	-0.081
		(0.324)			(0.838)
female	2.050	-0.132	education	11.051	-0.171
		(0.344)			(0.160)
post-secondary	1.954	-0.115	salary (1000's rupees)	4.027	0.079
		(0.461)			(0.223)
<5yrs exp	2.963	0.345	years teaching	2.604	0.368
		(0.290)			(0.247)
5yrs≤ exp<10yrs	0.648	-0.091	years teaching same school	1.825	-0.106
		(0.185)			(0.175)
≥10 yrs exp	0.043	0.008	hours teaching		
		(0.036)	total	25.669	-0.908
avg teacher absent ≥ 2 days/month	0.455	-0.121			(1.445)
		(0.095)	teaching whole class	5.379	-0.464
					(0.598)
building	0.993	-0.067	teaching small group	4.013	-0.188
		(0.040)			(0.444)
number classrooms	3.167	0.127	teaching individual	4.077	-0.736
		(0.286)			(0.521)
sufficient desks	0.812	-0.113	blackboard/dictation	3.891	-0.542
		(0.086)			(0.439)
drinking water	0.825	0.043	classroom management	2.208	0.065
		(0.073)			(0.239)
electricity	0.741	-0.034	testing	2.180	0.557
		(0.091)			(0.479)
toilet	0.766	0.045	administrative	1.824	0.431
		(0.081)			(0.358)

Note: This table reports differences in mean characteristics between program schools under the gender-uniform and gender-differentiated subsidy treatments. The unit of observation is child-school. Columns (1) and (4) report means for program schools under the gender-uniform subsidy treatment; and columns (2) and (5), differences in means between program schools under the two subsidy treatments. Standard errors, reported in parentheses, are clustered at the village level. *, **, and *** denote statistical significance at the ten-, five-, and one-percent levels, respectively.

Table A.3: School proximity and enrollment by school type

	pct villages w/school < 1.5km				pct villages w/ any child enrolled		number enrolled per village	
	all		founded pre-2007					
	control	treatment	control	treatment	control	treatment	control	treatment
school type:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Program	0.000	0.944			0.026	0.944	0.105	49.506
Government	0.553	0.463	0.474	0.463	0.632	0.253	19.211	3.389
Private	0.158	0.056	0.105	0.049	0.184	0.037	6.579	0.549
NGO	0.053	0.043	0.000	0.043	0.053	0.006	1.500	0.025

Note: This table reports the presence of schools and enrollment by school type, disaggregated by treatment status. Columns (1) and (2) give the percentage of villages with a school of the indicated type within 1.5 kilometers of the mean sample household; and columns (3) and (4) give the percentage of villages with schools of the indicated type which were founded prior to 2007. Columns (5) and (6) give the percentage of villages with any sample child enrolled in a school of the indicated type. Columns (7) and (8) give the number of sample children enrolled per village in a school of the indicated type.

Table A.4: Balance by school enrollment type

	control		enrolled	govt-enr
	unenrolled	govt-enr – unenrolled	program – govt	treat – control
	(1)	(2)	(3)	(4)
child/hh characteristics				
age	7.295	0.093	0.024	0.320***
		(0.184)	(0.098)	(0.118)
female	0.427	-0.019	0.032	0.010
		(0.027)	(0.025)	(0.052)
child of hh head	0.849	-0.021	0.038	0.001
		(0.037)	(0.027)	(0.042)
household size	7.343	0.309	-0.456	-0.634
		(0.460)	(0.361)	(0.506)
num children	4.871	0.135	-0.312	-0.553*
		(0.256)	(0.260)	(0.289)
hh head edu	1.632	1.671***	-0.808**	0.347
		(0.357)	(0.379)	(0.374)
total land (acres)	3.599	0.710	-0.510	-2.972
		(1.228)	(1.312)	(2.474)
house mud/thatch	0.821	-0.194***	0.149*	0.064
		(0.061)	(0.083)	(0.117)
num goats	3.695	0.098	-0.521	-1.047
		(0.477)	(0.436)	(0.875)
sunni	0.870	-0.077	0.080	0.084
		(0.105)	(0.089)	(0.138)
sindhi	0.618	0.023	0.043	0.121
		(0.059)	(0.068)	(0.104)

	control		enrolled	govt-enr
	unenrolled	govt-enr – unenrolled	program – govt	treat – control
	(5)	(6)	(7)	(8)
hh head occuptions				
landlord	0.051	0.004	-0.012	-0.009
		(0.017)	(0.016)	(0.028)
farmer	0.638	-0.109*	0.132**	0.019
		(0.062)	(0.058)	(0.073)
laborer	0.133	0.029	-0.023	-0.024
		(0.037)	(0.034)	(0.050)
teacher/civil servant	0.036	-0.013	0.017	0.044*
		(0.018)	(0.024)	(0.023)
other	0.174	0.074*	-0.097***	0.011
		(0.044)	(0.036)	(0.054)

Note: This table reports balance across program- and government-enrolled children. Columns (1) and (5) give the mean of the indicated variables for unenrolled children in control villages. Columns (2) and (6) give the difference between unenrolled children and government-enrolled children in control villages. Columns (3) and (7) give the difference between program-enrolled children in treatment villages and government-enrolled children in control villages. Columns (4) and (8) give the difference between treatment and control villages for children enrolled in government schools. Standard errors, reported in parentheses, are clustered at the village level. *, **, and *** denote statistical significance at the ten-, five-, and one-percent levels, respectively.

Table A.5: Test scores by school type

	program – govt	program – priv	program – govt	program – priv	program – govt	program – priv
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: baseline						
math score	0.236***	0.049	0.176**	0.058	0.266*	0.148
	(0.081)	(0.226)	(0.081)	(0.239)	(0.146)	(0.191)
urdu score	0.160***	0.025	0.091	0.046	0.121	0.061
	(0.061)	(0.125)	(0.060)	(0.132)	(0.094)	(0.102)
total score	0.211***	0.040	0.145**	0.054	0.215*	0.116
	(0.073)	(0.188)	(0.073)	(0.198)	(0.129)	(0.154)
Panel B: controling for highest grade						
math score	0.259***	0.031	0.196**	0.044	0.306**	0.127
	(0.086)	(0.230)	(0.085)	(0.239)	(0.149)	(0.194)
urdu score	0.174***	-0.007	0.103*	0.019	0.144	0.032
	(0.064)	(0.133)	(0.062)	(0.135)	(0.096)	(0.108)
total score	0.230***	0.015	0.162**	0.033	0.247*	0.090
	(0.078)	(0.193)	(0.076)	(0.199)	(0.131)	(0.159)
Panel C: verified attendance						
math score	0.251**	0.187	0.241	0.200	0.490***	0.287*
	(0.121)	(0.208)	(0.149)	(0.225)	(0.183)	(0.164)
urdu score	0.142	0.070	0.081	0.101	0.232*	0.124
	(0.092)	(0.131)	(0.107)	(0.144)	(0.135)	(0.096)
total score	0.213*	0.138	0.181	0.158	0.394**	0.221
	(0.111)	(0.177)	(0.136)	(0.192)	(0.168)	(0.135)
sample restrictions						
non-program school within 0.5 km	no	no	yes	yes	yes	yes
proposed program school within 0.5km	no	no	no	no	yes	yes

Note: This table reports differences in mean standardized test scores across school types, controlling for student characteristics and district fixed effects. Columns (1), (3), and (5) report differences in means between program and government schools; columns (2), (4), and (6) differences in means between program and private schools. Columns (3)–(6) restrict the sample to households within 0.5kms of a non-program school; and columns (5) and (6) further restrict the sample to households located within 0.5kms of the proposed program school site. Standard errors, reported in parentheses, are clustered at the village level. *, **, and *** denote statistical significance at the ten-, five-, and one-percent levels, respectively.

Table A.6: Education production estimates, subject test scores

	language test scores			mathematics test scores		
	(1)	(2)	(3)	(4)	(5)	(6)
toilets and/or drinking water	0.129	0.141	0.158	0.132	0.140	0.158
	(0.098)	(0.099)	(0.111)	(0.103)	(0.108)	(0.111)
female	0.021	0.022	0.034	0.021	0.022	0.034
	(0.042)	(0.041)	(0.054)	(0.053)	(0.053)	(0.054)
age	0.093***	0.093***	0.105***	0.105***	0.104***	0.105***
	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)
tuition required	0.001	0.001	0.000	0.001	0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
distance from home to school	-0.022	-0.024	-0.029	-0.031	-0.033	-0.029
	(0.053)	(0.053)	(0.058)	(0.059)	(0.060)	(0.058)
pct teachers $<$5yrs exp	0.114	0.148	0.306***	0.276**	0.301**	0.306***
	(0.088)	(0.099)	(0.110)	(0.108)	(0.118)	(0.110)
pct teachers post-secondary	0.230***	0.221**	0.152	0.167	0.160	0.152
	(0.087)	(0.087)	(0.115)	(0.111)	(0.111)	(0.115)
pct teachers female	0.061	0.067	0.107	0.099	0.104	0.107
	(0.079)	(0.079)	(0.094)	(0.092)	(0.094)	(0.094)
pct time teaching	-0.266	-0.272	-0.085	-0.102	-0.106	-0.085
	(0.256)	(0.253)	(0.342)	(0.325)	(0.322)	(0.342)
avg teacher absent $\geq$2 days/month	-0.007	-0.005	-0.071	-0.072	-0.070	-0.071
	(0.059)	(0.059)	(0.074)	(0.073)	(0.073)	(0.074)
pct teachers female X female student	0.050	0.050	0.018	0.026	0.026	0.018
	(0.043)	(0.043)	(0.051)	(0.051)	(0.051)	(0.051)
distance X female student	-0.000	-0.002	0.005	0.008	0.007	0.005
	(0.022)	(0.022)	(0.024)	(0.024)	(0.024)	(0.024)
toilets and/or drinking water X female student	-0.046	-0.047	-0.061	-0.056	-0.056	-0.061
	(0.046)	(0.046)	(0.059)	(0.058)	(0.059)	(0.059)
government school		0.070			0.052	
		(0.092)			(0.125)	
program school			-0.062			-0.062
			(0.108)			(0.108)
R-squared	0.068	0.068	0.069	0.069	0.069	0.069
N	7246	7246	7112	7196	7196	7112

Note: This table reports education production estimates, relating standardized subject test scores to school inputs and student characteristics, controlling for district fixed effects. Columns (1) through (3) report estimates for language test scores, and columns (4) through (6), estimates for mathematics test scores. Columns (1) and (4) report estimates without indicator variables for program or government schools; columns (2) and (5) with an indicator variable for government schools, and columns (3) and (6) with an indicator variable for program schools. Standard errors, reported in parentheses, are clustered at the village level. *, **, and *** denote statistical significance at the ten-, five-, and one-percent levels, respectively.

Appendix

Table B1: Treatment and test scores (pct)

	control mean(sd)	ITT				IV	
						reported	verified
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
math pct correct	0.461	0.163***	0.160***	0.160***	0.194***	0.603***	0.624***
	(0.307)	(0.047)	(0.048)	(0.048)	(0.038)	(0.087)	(0.140)
urdu pct correct	0.487	0.171***	0.168***	0.167***	0.202***	0.619***	0.670***
	(0.342)	(0.058)	(0.060)	(0.058)	(0.044)	(0.078)	(0.150)
total pct correct	0.471	0.167***	0.163***	0.163***	0.197***	0.607***	0.643***
	(0.310)	(0.051)	(0.052)	(0.052)	(0.040)	(0.081)	(0.144)
child controls		no	yes	yes	yes	yes	yes
HH controls		no	no	yes	yes	yes	yes
district fixed effects		no	no	no	yes	yes	yes

Note: This table reports program impacts on test scores measured as the percentage of questions answered correctly. Column (1) gives the mean percent of correct answers for children aged 5–9 in control villages, with the standard deviation reported in parentheses. Columns (2) through (5) report the intention-to-treat (ITT) impacts, with various sets of controls. Columns (6) and (7) report the treatment-on-the-treated (TOT) impacts on test scores, based on reported and verified enrollment, respectively. Standard errors, reported in parentheses, are clustered at the village level. Statistical significance at the one-, five-, and ten-percent levels denoted by ***, **, and *, respectively.

Table B2: Treatment and test scores, disaggregated by age

child	control village		test score, ITT				first	
			percent		deviation			
age	enrolment	test score					stage	IV (dev)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
5	0.440	0.425	0.114**	0.180***	0.413**	0.650***	0.335***	1.831***
		(0.276)	(0.055)	(0.031)	(0.198)	(0.113)	(0.064)	(0.368)
6	0.551	0.439	0.150***	0.174***	0.489***	0.568***	0.298***	1.938***
		(0.307)	(0.054)	(0.043)	(0.176)	(0.139)	(0.073)	(0.358)
7	0.555	0.461	0.172***	0.212***	0.557***	0.686***	0.307***	2.095***
		(0.309)	(0.059)	(0.043)	(0.190)	(0.139)	(0.074)	(0.333)
8	0.545	0.493	0.175***	0.207***	0.556***	0.659***	0.335***	1.961***
		(0.314)	(0.054)	(0.046)	(0.172)	(0.146)	(0.072)	(0.260)
9	0.511	0.474	0.207***	0.232***	0.636***	0.711***	0.337***	2.098***
		(0.326)	(0.053)	(0.050)	(0.164)	(0.155)	(0.074)	(0.294)
10	0.557	0.553	0.149***	0.185***	0.469***	0.583***	0.265***	2.013***
		(0.318)	(0.055)	(0.048)	(0.172)	(0.153)	(0.065)	(0.358)
child controls			yes	yes	yes	yes	yes	yes
HH controls			yes	yes	yes	yes	yes	yes
district fixed effects			no	yes	no	yes	yes	yes

Note: This table reports program impacts on test scores and enrolment disaggregated by age. Column (1) gives the mean enrolment rate in control villages, and column (2) the mean and standard deviation of the test score. Columns (2) through (6) report the intention-to-treat (ITT) impacts on test scores, measured as percent (columns 3 and 4) and standard deviations (columns 5 and 6). Column (7) gives the effect of treatment on enrolment; and column (8) reports the treatment-on-the-treated (TOT) impacts on test scores measured using the normalized test score. Standard errors, reported in parentheses, are clustered at the village level. Statistical significance at the one-, five-, and ten-percent levels denoted by ***, **, and *, respectively.

Table B3: Treatment and test scores, disaggregated by question type

	control score	ITT				TOT reported	TOT verified
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Language							
full letter identification	0.576	0.185***	0.182***	0.181***	0.222***	0.681***	0.748***
		(0.064)	(0.065)	(0.064)	(0.052)	(0.092)	(0.167)
half letter identification	0.458	0.174***	0.172***	0.171***	0.202***	0.616***	0.666***
		(0.058)	(0.059)	(0.058)	(0.044)	(0.085)	(0.152)
first letter of objects	0.475	0.168***	0.165***	0.165***	0.195***	0.596***	0.644***
		(0.053)	(0.055)	(0.053)	(0.043)	(0.080)	(0.144)
reading full word	0.434	0.166***	0.162***	0.162***	0.201***	0.616***	0.652***
		(0.060)	(0.061)	(0.060)	(0.042)	(0.085)	(0.151)
reading full sentence	0.407	0.121	0.118	0.118	0.156***	0.477***	0.514***
		(0.077)	(0.078)	(0.077)	(0.044)	(0.091)	(0.161)
Numeracy							
size comaparison of objects	0.730	0.107***	0.105**	0.105***	0.138***	0.428***	0.443***
		(0.041)	(0.041)	(0.040)	(0.032)	(0.070)	(0.111)
number awareness	0.528	0.187***	0.185***	0.184***	0.217***	0.681***	0.711***
		(0.053)	(0.054)	(0.052)	(0.044)	(0.094)	(0.144)
number identification	0.457	0.191***	0.188***	0.188***	0.211***	0.657***	0.695***
		(0.050)	(0.051)	(0.050)	(0.042)	(0.086)	(0.151)
number ordering	0.429	0.199***	0.196***	0.195***	0.224***	0.698***	0.718***
		(0.044)	(0.045)	(0.045)	(0.038)	(0.104)	(0.140)
object counting	0.449	0.186***	0.182***	0.182***	0.213***	0.664***	0.698***
		(0.051)	(0.052)	(0.051)	(0.044)	(0.095)	(0.153)
object counting & comparison	0.424	0.186***	0.183***	0.183***	0.216***	0.664***	0.674***
		(0.048)	(0.049)	(0.048)	(0.041)	(0.099)	(0.147)
shape identification	0.418	0.176***	0.172***	0.172***	0.212***	0.660***	0.670***
		(0.056)	(0.057)	(0.056)	(0.044)	(0.095)	(0.161)
addition	0.429	0.158**	0.155**	0.155**	0.200***	0.628***	0.654***
		(0.063)	(0.064)	(0.064)	(0.045)	(0.124)	(0.171)
subtraction	0.348	0.144***	0.141***	0.142***	0.174***	0.541***	0.562***
		(0.054)	(0.054)	(0.054)	(0.039)	(0.110)	(0.147)
telling time	0.371	0.127**	0.123**	0.123**	0.159***	0.496***	0.499***
		(0.057)	(0.058)	(0.057)	(0.043)	(0.112)	(0.155)
child controls		no	yes	yes	yes	yes	yes
HH controls		no	no	yes	yes	yes	yes
district fixed effects		no	no	no	yes	yes	yes

Note: This table reports program impacts on test scores measured as the percentage of questions answered correctly disaggregated by the type of question. Column (1) gives the mean percent of correct answers for children aged 5–9 in control villages. Columns (2) through (5) report the intention-to-treat (ITT) impacts, with various sets of controls. Columns (6) and (7) report the treatment-on-the-treated (TOT) impacts on test scores, based on reported and verified enrollment, respectively. Standard errors, reported in parentheses, are clustered at the village level. Statistical significance at the one-, five-, and ten-percent levels denoted by ***, **, and *, respectively.