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 partnering with the private sector to improve the delivery of education. We evaluate an innovative program that enlisted local entrepreneurs to open and operate new schools in 200 randomly selected, underserved villages in Sindh, Pakistan. School operators received a per-student subsidy from the local government to provide tuition-free primary education, and in half of the treated villages operators received a higher subsidy for female students. Management of these schools was highly decentralized, with school operators permitted to tailor inputs to local demand. The program increased enrollment in treatment villages by 31 percentage points, and test scores by 0.64 standard deviations, with no difference by gender or across the two subsidy schemes. Treatment effects are driven primarily by the establishment of schools in villages where they were previously absent, though program schools also improve educational outcomes even when nearby government schools are available. To gain greater insight into program school efficiency, we estimate a structural model of the demand and supply for school inputs. This exercise reveals that program schools selected inputs similar to those of a social planner who internalizes all the education 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, 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

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 Davies, and participants at presentations at Harvard University, the World Bank, 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

Low- and middle-income countries continue to struggle with 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 increasingly experimented with models giving a greater role to private education providers. Existing research on the effectiveness of this approach has largely focused on programs in which governments subsidize enrollment in existing private schools (Patrinos et al., 2009). Because many of the most educationally deprived areas often lack pre-existing private schools with which to partner, governments have also experimented with policies involving the creation of new private schools. Whether such an approach can be successful, however, is far less certain, as the absence of pre-existing private schools may be driven by unfavorable local conditions.[1]

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. In addition, a second treatment arm incentivize girls' enrollment by providing entrepreneurs with a subsidy premium. Entrepreneurs exercised wide latitude in the inputs they provided, including the ability to hire teachers with lower formal qualifications than required for government teachers.

A lengthy literature has argued that private schools have advantages over public schools due to their stronger incentives to reduce costs and innovate, and that they more closely tailor school inputs to the preferences and needs of their students (Friedman, 1955; Shleifer, 1998).[2] A number of papers have tested this thesis empirically using experiments with vouchers, and have generally found either that private schools deliver better educational outcomes than government schools, or that they produce similar educational outcomes but at a significantly lower cost (Kim et al., 1999; Angrist et al., 2002; Alderman et al., 2001, 2003; Barrera-Osorio and Raju, 2015; Barrera-Osorio et al., 2016; Muralidharan and Sundararaman, 2015; Romero et al., 2019).[3] Within Pakistan, an influential literature has shown that, conditional on child

[1] To the best of our knowledge, Alderman et al. (2003) is the only other paper to evaluate such a program. That paper evaluates a similar program conducted in the Balochistan province of Pakistan in the 1980s. The 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.

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

[3] Angrist et al. (2002) show that voucher winners in Colombia had higher test scores and school progression. Muralidharan and Sundararaman (2015) use a voucher scheme to show that private schools in Andhra Pradesh, India deliver similar levels of instruction in most subjects as public schools, though at a fraction of the cost and time; and have a large positive impact on Hindi (a non-local language) skills.

characteristics, children enrolled in low-cost private schools have higher test scores than government-enrolled children, though this finding is not based on experimental variation (Andrabi et al., 2010, 2011).

The purported advantages of private education, coupled with often limited state capacity, has led developing country governments to increasingly make use of Public Private Partnerships (PPPs) in order to meet their education objectives (Patrinos et al., 2009). Among the most common types of PPPs are schemes in which governments provide funding for children to enroll in existing private schools.[4] To mitigate the possibility that privately operated schools will pursue objectives different than those of the government, PPPs generally include extensive contractual obligations for the provision of specific services (Patrinos et al., 2009). For example, contracts may regulate school inputs, or stipulate some level of school quality in order to participate. However, even where such contracts are in place, the focus of private entrepreneurs on profits may lead to the under-provision of socially-valuable but non-contractible aspects of education (Hart et al., 1997). For this reason, it is theoretically ambiguous whether market incentives will yield the intended benefits from relying on private providers.

An important question regarding PPPs is whether greater centralization of decision-making improves service delivery. While centralized control may facilitate the implementation of contractual terms specified by the government, decentralization has the potential to make schools more responsive to local demand. Recent research from Liberia studies the effects of a program in which the management of failing government schools was handed over to large companies operating chains of private schools, in which decision-making was highly centralized (Romero et al., 2019). The authors find that these schools were successful in improving educational outcomes, though followup research has somewhat tempered these findings (Romero and Sandefur, 2019).

The program studied in this paper extends the existing research in two important ways. First, the management of these schools was highly decentralized, with schools being operated by local entrepreneurs, who exercised wide discretion in the inputs they provided. Second, the PPRS program involved the establishment of new, privately operated schools. In contrast, most previously studied programs have involved schools that had already existed for some time, so that inclusion in such programs implicitly selects on the prior success of participating schools.

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 local entrepreneurs a set of benefits to establish and run tuition-free, coeducational primary schools in educationally underserved villages. The benefits included a per-student

[4] See Patrinos et al. (2009) for a comprehensive survey of the types of PPPs.

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.[5] To address the large gender disparity in primary school enrollment in rural Sindh, half of the program villages were randomly assigned to a gender-differentiated subsidy scheme. Under the latter scheme, program school operators received a higher per-student subsidy for girls than for boys, with the aim of more strongly incentivizing them to take steps to attract girls.[6]

For the purpose of assessing the performance of program schools, we explore three counterfactuals. First, we compare educational outcomes of children in treatment villages to those in control villages in order to determine the effect of gaining access to program schools on village-wide educational outcomes. Second, we compare the test scores of children enrolled in program schools to those in government schools (in control villages), in order to assess whether program schools yield the quality advantages often ascribed to private provision. Finally, we undertake an absolute assessment of the efficiency of program schools (given local resources) by comparing program school inputs to those of a social planner who maximizes social surplus.

The program was highly effective. Comparing treatment and control villages nearly 2 years after the schools were opened, the program increased school enrollment for children aged 5–10, the program's stated target age group, by 31 percentage points, and that for children aged 11–17 by 12 percentage points. The program also raised total test scores by 0.64 standard deviations, with mean test scores increasing from 47.1% correctly answered questions in control villages to 67% in treatment villages. For children induced by the program to enroll in school, the increase in test scores was approximately 2 standard deviations. The overall treatment effect was the same for boys and girls; and the gender-differentiated subsidy treatment had similar impacts on girls' enrollment and test scores as the gender-uniform

[5]After the random assignment was completed, we reduced the original evaluation sample 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 nearly two years after the start of the program, at which time the schools had been operational for approximately 1.5 school years.

[6]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).

one.

Improvements in educational outcomes were primarily driven by making schools available in villages where they had previously been absent. However, program schools yielded additional gains by both increasing enrollment in villages where government schools were present, as well as through the higher quality of program schools relative to government schools. Evidence for the quality of program schools can be seen in the fact that virtually all government-enrolled children in treatment villages switch to program schools, as well as the higher test score received by children enrolled in program schools. Though the latter finding is not based on experimental variation, we show that it is unlikely to be due to selection effects.

In order to gain greater insight into the quality of program schools, we examine the efficiency of input choices in program schools vis-à-vis the social planner's solution based on structural model estimations of schooling demand and education production. The experimental design provides a unique opportunity for conducting this exercise in a credible manner. In non-experimental settings, one would be concerned that there are correlated unobservables (such as village-level preferences for education) that are driving both the educational outcomes of interest, as well as the presence of schools and the inputs they select. Because the experiment exogenously varies the set of schools across villages, it allows us to estimate a structural model without the biases that would otherwise be introduced by endogenous school placement.

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 compute the optimal set of school inputs that a social planner would have chosen for each village, 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 very well in choosing school inputs, capturing approximately 94 percent of the total amount of potential social surplus. Both the program schools' and the social planner's solutions show substantial variation in provided inputs across villages. The principal difference between the two is that program school operators hire teachers that attract slightly fewer students, but who are cheaper and increase profits. Specifically, program school operators hire more female teachers, and hire teachers with less teaching experience and a higher rate of absenteeism than those hired by the social planner. The program schools are quite similar to pre-existing private schools in survey villages, consistent with their solving a similar optimization problem.[7] Most significantly, teacher characteristics are virtually identical across program and private schools, and contrast sharply with teacher characteristics in government schools.

In sum, we find that private providers 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. Indeed, since its inception, the PPRS program, and the related SEF Assisted Schools (SAS) program, have been expanded to cover more than 550,000 students across more than 2,000 schools, speaking to the importance and potential of this model.

2 Background

2.1 Schooling in Pakistan

School enrollment is low in Pakistan, even in comparison to countries with a similar income level (Andrabi et al., 2008b). At the time the PPRS program was initiated in 2008/09, 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).

Pakistan has witnessed a dramatic growth in private schools in the last three decades, increasing from 4,000 in the early 1980s to 47,000 in 2005, with thirty percent of primary-school students in 2009 being enrolled in private schools.[9] Much of this expansion occurred in villages and poorer urban neighborhoods (Andrabi et al., 2008a). These schools have succeeded in terms of both cost and quality. At less than $20 per annum in 2000, the cost of private primary school fees represented about two percent of mean total household spending (Andrabi et al., 2008a). Low-cost private schools are nonetheless found to produce higher test scores than government schools in rural Punjab province (Andrabi et al., 2010, 2011).

The affordability of these schools has been made possible by both low fixed costs, as well

[7] The principal difference between program and private schools is that private schools are larger, both in terms of the number of students and the number of teachers. This is likely because program schools were only allowed to enroll children located within the villages to which they are assigned.

[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.

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

as low operational costs, the latter driven primarily by the low wages paid to teachers. Low wages are in turn made possible by the generally lower educational qualifications of private school teachers, as well as the fact that many teachers are female, for whom there are fewer alternative labor market opportunities. Teachers in government schools, in contrast, are part of the civil service and are required to have certain minimal educational qualifications; and their salaries are determined by seniority and formal educational qualifications.[10] As a consequence, teacher salaries in government schools are 5 times higher than those in private schools, and constitute 80% of expenditures in public institutions (Bau and Das, 2019).

Another advantage enjoyed by private schools is the autonomy they enjoy in the design of their curriculum. This contrasts sharply with public schools, where the curriculum is set by the central government, and each school must implement it with little room for variation.

Despite their documented advantages, only five percent of primary-school students in rural Sindh were enrolled in private schools during the 2008–09 term (Government of Sindh, 2009). One of the most important constraints on the presence of low-cost private schools appears to be the supply of local women with secondary education, as this labor force is crucial to the cost structure that makes these schools viable (Andrabi et al., 2013).[11]

The location of government schools, on contrast, depends primarily on budget constraints. While the typical village will have one or two public schools, in many remote areas schools will often be absent or insufficiently staffed, leading to high student-teacher ratios and poor service delivery. Where government schools are present, access is open to all children free of charge, though some contributions may be required for school supplies.

2.2 PPRS Program

In 2007, the provincial government initiated the Sindh Education Sector Reform Program (SERP), a multifaceted reform of public spending and provision in primary and secondary education. A key component of SERP was the use of PPPs, entailing public financing and private provision, with the objective of simultaneously 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. The principal objectives of the PPRS program were to increase access to schooling in marginalized areas, to reduce the gender

[10] Teachers in the public sector must have, at a minimum, a Primary Teachers Certificate and Certificate of Teaching. The previous requirement that they have a B.A or a BS.C. has been phased out.

[11] 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.

disparity in school enrollment, and to increase student learning, in a cost-effective manner.

We evaluate the first phase of this program, which was implemented in eight (out of, at that time, 23) districts in the province. SEF selected the districts based on 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 lowest-ranked districts were selected, excluding those that were experiencing heightened law-and-order concerns.[12]

Based on a budgetary assessment, SEF approved the creation of primary schools in 200 villages. These schools were to be established and operated by private providers, and were required to admit all children within the village free of charge. 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.[13]

Two types of subsidies were provided: a gender-uniform subsidy, in which entrepreneurs received 350 rupees per-student per-month (approximately $5 in annualized 2008 US dollars); and a gender-differentiated subsidy, in which entrepreneurs received an additional 100 rupees per month for each female student (i.e., 450 rupees). 100 villages were assigned to each of the two subsidy treatments. The subsidy amounts were set at less than one-half the per-student cost of public primary and secondary government schools in the province.[14] The subsidies were provided to entrepreneurs on a quarterly basis, and were based on a formula which multiplied the number of children in attendance by 1.25 to reflect an expected 20-percent absence rate. Attendance was assessed by SEF during periodic, unannounced monitoring visits.

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 age that they intended to enroll their children in the school should it be established; an available building in the village that was located at least 1.5 kilometers from the nearest school, and which was of sufficient size; and the identification of potential teachers with a minimum of eight years of schooling (middle school completion), with at least two being female.[15]

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

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

[14]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.

[15]SEF viewed eight years of schooling as sufficiently high that teachers would have the competency to

Once in the program, school operators would continue to receive the subsidy and other benefits as long as they adhered to certain basic conditions. The SEF strictly enforced the condition that families not be charged for enrollment, but was more lenient in enforcing the school infrastructural features and environmental conditions. In addition, the contract stipulated that compensation would be based on a formula using verified attendance, as described above. No contract was terminated in any of the sample schools due to contract breach.

3 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 this survey, the 263 qualifying villages were randomly assigned to the two subsidy treatments and the control. After random assignment, the original evaluation sample was reduced to 199 villages through the exclusion of sites that were situated in large towns with numerous existing schools. The effective evaluation sample consisted of 82 villages under the gender-uniform subsidy treatment, 79 under the gender-differentiated treatment, and 38 in the control group.

Schools were established in the summer of 2009. Because the new school year normally commences in the spring, program-school students had an abbreviated first school year. An initial followup survey was conducted nearly one year after the program started (May/June 2010), during which a full census of the village was taken. A second followup survey was conducted in April/May 2011, after the conclusion of the second school year under the program.

The baseline survey consisted of, first, a community survey answered by village leaders; second, a school survey of all schools in the general vicinity of the village;and, finally, 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 should it be established. The household survey collected information on the household, the household head, and on each child aged 5–9. There was also a survey of the entrepreneur and proposed teachers, as well as physical checks of the proposed school site and building. GPS data was collected for surveyed households, the proposed program-school site, and all nearby schools.

The first followup survey was implemented as a full village census, and included only a small number of questions on household, household head, and child characteristics. For this activity, enumerators had a prominent member of the community guide them through the

teach primary school content, but low enough that qualified individuals could be found locally.

village and indicate every household that belonged to the village. The full list of households was then used as the sampling frame for the second followup survey. The second followup survey was longer and more comprehensive than the first, but included only a subset of households.

It is important to note that the sampling frame used for the followup surveys may differ from that used for the baseline survey, as the latter was based on the *entrepreneur*'s assessment of which children belonged to the village. For the same reason, the catchment area from which children were admitted into the program schools was likely also different from the village boundaries used for the followup survey. This can be seen most clearly in the enrollment figures from the school surveys, which often exceeded the total number of children within the village. This is likely due to both entrepreneurs' admitting children from outside the village, as well as ambiguities in the definition of village boundaries. Reassuringly, where control and treatment villages were located close to one another (i.e., within 1–2 kms), there is no evidence that children in the control villages enrolled in nearby program schools.[16]

The second followup survey, conducted nearly 2 years after the start of the program, and with schools having been in operation for 1.5 years, consisted of three instruments: a school survey; a household survey; and a child survey, which included a test administered and supervised by the surveyors. The household survey was administered to households with at least one child aged 5–9. In large villages, up to 42 randomly sampled households (with qualifying children) in the village were interviewed; in villages with fewer than 42 qualifying households, which comprised the majority, all households in the village were interviewed. The household survey included 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 household and child surveys were administered at the child's home.

The school survey was conducted for all schools located within the village. The survey included: interviews of head teachers and all other teachers; and visual inspections by enumerators of school infrastructural and environmental conditions. GPS data were gathered from all surveyed households and schools. Where possible, we also surveyed schools outside the village, but located within 3 kms. Due to time constraints, we generally administered an abbreviated school survey for these proximate schools, which included a small number of

[16] In addition, the distribution of the distances of households from program school sites (proposed and actual), as well as visual inspections of GIS maps of villages, indicates that the village boundaries determined by the census included all households within the primary clusters of settlements.

questions on basic school characteristics.

Table 1 reports sample sizes of the baseline and follow-up surveys, by treatment status. The census conducted during the first followup survey indicates that there were 8,639 households with children aged 5–10, and 25,157 children in this age range, in the 199 sample villages. The baseline survey included 2,089 households with 5,556 young children (in this case, ages 5–9); and the second followup survey included 5,966 households and 17,720 young children.

4 Empirical strategy

We assess the effectiveness of program schools along two dimensions. First, we ask how successfully program schools meet their objective of increasing enrollment and test scores in treatment villages relative to control villages. Second, we seek to assess the efficiency with which program schools meet this objective.

To answer the first question, we estimate the intention-to-treat effect of program schools comparing child enrollment and test scores across control and treatment villages. We also present estimates for: the influence of program schools on vocational aspirations; how treatment effects varied with distance; whether the gender-differentiated treatment differentially affected female enrollment; and the cost-effective the intervention. In addition, we seek to disentangle the mechanisms driving treatment effects by assessing the respective roles played by school proximity and school quality, using existing government schools as the counterfactual.

To address the second question, pose and estimate a structural model, which we use to assess how closely the school inputs selected by private entrepreneurs aligned with those of a benevolent social planner. We use a discrete choice model to estimate demand for schooling with and without the randomly-assigned program schools, controlling for the presence of competing schools. Because entrepreneurs enjoyed wide latitude in making school input decisions, we are able to use a revealed preference approach to infer input costs. Combining these elements with estimates from the literature on the social value of education, we explore how demand, cost, and the social value of education would change with different school inputs in each treatment village. To our knowledge, this type of welfare analysis is novel in the literature, and depends crucially on the experimental intervention for generating random variation in the location of newly created schools.

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 reduced form 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, based on a regression of the indicated variable on an indicator variable for treatment status.

Differences across control and treatment villages were small and statistically insignificant for virtually all household and child characteristics. One notable exception is gender, with children in treatment villages being slightly more likely to be female than those in control villages (4.2 percentage points and 3.0 percentage points at baseline and follow-up, respectively). Appendix Table A.1 reports differences in household and child characteristics across villages under the gender-uniform and -differentiated subsidy treatments. Differences between the two subsidy treatments were small, and always statistically insignificant.

5 Program impacts

The ITT effect of the program is based on the following specification:

$$Y_{ij} = \beta_0 + \beta_1 T_j + \beta_2 X_i + \delta_i + \varepsilon_{ij}, \tag{1}$$

where Y_{ij} is the outcome of interest for child i in village j, T_j is an indicator variable indicating whether village j was assigned a program school, X_i is a vector of child and household characteristics, and δ_i are district fixed effects. Household characteristics include the education of the household head, whether the household head is a farmer, total land holdings, the number of children, and the number of adults. Child characteristics are child age and child gender. 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.

5.1 Enrollment

The first outcome we measure is the effect of the treatment on child enrollment. Enrollment is based on the respondent-reported enrollment status of the child in the just-concluded school term. We also estimate the effect of the treatment on the highest grade attained. Because these measures may be subject to misreporting, we also administered tests to the children to give a better assessment of the true educational outcomes. As we show subsequently, the treatment effects for test scores are consistent with those for self-reported enrollment.

Table 3 reports the impacts of the treatment on school enrollment and grade attainment. Because treatment effects are similar across the two treatment arms (as shown in subsequent analysis), we use the pooled treatment in our baseline specification. 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 enrollment with different sets of controls. Column (5) reports impacts on highest grade attained with the full set of controls. Mean reported enrollment of young children in control villages was 52.9%, and the average child had 0.81 years of education. Based on the model with the full set of controls, the program increased enrollment among young children by 31.6 percentage points, and increased grade attainment by 0.38 grades.

While older children were not the target population for the program, those who were in the qualifying age range (5–10) when they first enrolled in program schools were allowed to continue attending even after they aged out of this group. Mean enrollment of older children in control villages was 37.5%. Based on the model with the full set of controls, we find that the program increased reported school enrollment among older children by 11 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 grades offered in program schools, which were at the primary level.

5.2 Test Scores

Table 4 reports the impact of the pooled treatment on test scores. Test scores are standardized by subtracting the mean score for all children aged 5–10 in control villages and dividing by the standard deviation (47% and 31%, respectively). Columns (2)–(5) report treatment effects with various sets of controls. The outcomes are math score, language score, and the total score. Based on the model with the full set of controls, the program increased total test scores by 0.64 standard deviations. Program impacts were similar for both subject test scores.

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

In Appendix Table B1 we report results using as the outcome the percent of test questions answered correctly. Columns (2)–(5) report the ITT estimates with various sets of controls;

and Column (6) reports the TOT estimates with the full set of controls. The ITT effect on total test score was a 19.7 percentage points increase, and the TOT effect was 61 percentage points.

Appendix Figure 2 shows the full distribution of test scores across the control and treatment groups. As is apparent, there is a mass of students answering 0% of questions correctly in control villages, and 100% of questions in treatment villages, which may lead us to underestimate the treatment effects. We therefore estimate an Item Response Theory (IRT) model, using both MLE and Bayesian (EAP) methods. The intuition for this method is that a latent skill parameter (θ) can be estimated for each child based on the difficulty of correctly answered questions, where the difficulty of a question is based on the correlation between answering that question correctly and the overall test score.[17] The results of this exercise are given in Appendix Table B2. The results of both EAP and MLE procedures are similar to those observed using the standardized test score as the outcome, indicating that floor and ceiling effects are not systematically biasing our results.

We also examine program impacts on test scores by the competency being tested (Table 5), and by child age (Appendix Table B3). The treatment effect is generally stable across different competencies. There were two exceptions to this: on the language exam, students showed a smaller improvement on full sentence comprehension, the most challenging competency; and on the math exam students showed a somewhat smaller improvement on the two practical competencies, size comparison and telling time. For child age, we report the ITT and TOT test-score effects using as the outcome variables both the percent of questions answered correctly (columns (4) and (6), respectively), as well as the standardized test score measure (columns (5) and (7), respectively). We include the full set of controls, excluding child age. The program effects were generally similar across age groups.

5.3 Differential impacts on school enrollment and test scores

We also examine the impacts of the two subsidy treatments (Table 6, panel A), the impacts of the pooled treatment by gender (Table 6, panel B), and the impacts of the two subsidy treatments by gender (Table 7). We do not find differential effects by subsidy treatment, by gender, or by subsidy treatment and gender.

5.4 School proximity and educational outcomes

The principal mechanism driving improvements in educational outcomes is that the treatment leads to a dramatic reduction in the distance to school, thereby reducing the costs of enrollment. This can be seen most clearly in the fact that, in results not shown, we find that

[17]See van der Linden and Hambleton (2013) for a fuller discussion of this method.

treatment villages which lacked a nearby government school witnessed a 58 percentage point increase in enrollment, whereas the presence of a government school reduced the treatment effect to a 20 percentage point increase in enrollment.

To better understand the role of school proximity, we next present figures displaying the relationship between school proximity and educational outcomes. Figure 1 shows the relationship between educational outcomes and the distance to the nearest *proposed* program school site. The treatment causes an upward shift in both enrollment and test score at all distance from the proposed program school site. This relationship is relatively similar across genders (Appendix Figure A1).

In Appendix Figure A2, we plot the relationship between educational outcomes and the distance to the nearest operational primary school of *any type*. We plot this relationship in control villages up to a distance of 1.5 kms; and in program villages up to a distance of 1 km, due to the small number of households in these villages located farther away. Remarkably, there is virtually no relationship between educational outcomes and school distance in treatment villages, while in control villages there is clear gradient between distance and both educational outcomes. In addition, even when located within very small distances of the nearest school, children in control villages are less likely to be enrolled, and receive a lower test score, than children in treatment villages.

Appendix Figure A3 shows 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, have better educational outcomes than girls at distances more than 0.6 kilometers from the nearest school.[18] This suggests that program schools either provide inputs (such as female teachers) more attractive to female students than those of non-program schools, or that entrepreneurs have taken alternative measures to recruit female students.

There are two likely explanations for the disparity across control and treatment villages in the relationship between distance and educational outcomes. First, because the payment scheme is based on the number of students enrolled, entrepreneurs may have taken measures to maximize enrollment. Alternatively, it may be the case that program schools are perceived to be relatively high quality, and that the returns to education therefore overwhelm the costs incurred in traveling greater distances.

[18]This finding resembles that of Burde and Linden (2013), who show that school proximity is a significant determinant of the educational gender gap.

5.5 School quality and educational outcomes

We find strong evidence that attributes other than the proximity of program schools also contributed to program-school enrollment and improvements in educational outcomes. As noted above, approximately half of the villages had a nearby government school at the time of the survey, and a smaller number had other types of primary schools (Appendix Table A.4).[19] However, not only do we find a substantial increase in enrollment even in villages with a proximate government school, we also find that children generally switched from government to program schools when given the option. Specifically, whereas an average of 19 children were enrolled in government schools in control villages, only 3 were enrolled in each treatment village (Appendix Table A.4), constituting 89% and 8% of enrolled children in the control and treatment groups, respectively.

One likely reason for the preference for program schools is their perceived quality. Indeed, a central motivation for the use of a PPP design was the evidence found in earlier research indicating that low-cost private schools in Pakistan deliver better educational outcomes than government schools (Andrabi et al., 2010, 2011). We therefore test whether the advantages observed with private schools carry over to program schools.

For this exercise, 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 8, column (1) reports mean test scores for program schools; columns (2) and (3) report differences in mean test scores between program schools and and government and private schools (in control villages), respectively; and column (4) reports p-values from tests of differences in mean test scores between government and private schools. These regressions include the full vector of child and household controls, as well as district fixed effects. 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 (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 may bias the estimates. For example, 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

[19] 55% and 46% of control and treatment villages, respectively. This difference is not statistically significant, and represents just 2.5 additional villages with government schools across the entire sample of 38 control villages.

overstate their quality.

The evidence is strongly supportive of the former hypothesis. First, as previously noted, program schools attract nearly all the children that would have otherwise been enrolled in government schools, making it unlikely that the differential is due to cream-skimming. As further evidence against child sorting, in Appendix Table A.5 (columns 4 and 8) we find that mean characteristics of government-school students are largely similar across control and program villages.[20]

In addition, program schools have encouraged the enrollment of socioeconomically disadvantaged students. Appendix Table A.5 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.

5.6 Aspirations

The program has substantial impacts on aspirations for children. In Table 9, panel A reports impacts on aspirations for each child aged 5–17 expressed by the adult respondent; and in panel B the impacts on aspirations expressed by the children themselves (aged 5–10). Column (1) reports means in control villages, and column (2) reports the differences in means between program and control villages. We also examine gender-differentiated impacts on aspirations: columns (3), (4), and (5) report regression coefficients for girls, the treatment group, 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 (+7.7 percentage points), and less likely to desire that they become housewives (-14.6 percentage points). Program-village households desired higher educational attainment for their boys and girls (+1.5 and +1.7 years, respectively). There was no effect of the treatment on the

[20] The principal exception is 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.

desired age of marriage for either boys or 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 for them by the adult respondent (11.3 years when asking children versus 7.4 years when asking adult respondents, in control villages).

6 Program cost-effectiveness

The cost-effectiveness calculation has three steps. First, we calculate total costs by fiscal year. Second, we determine the number of students attending program schools, as well as mean test score improvements for these children. Finally, we calculate the cost-effectiveness ratio, using the impact of the program on either enrollment or test score.

All program costs, which were provided by SEF, are calculated in present value terms in 2011 US dollars following the method proposed by Dhaliwal et al. (2013). We use the cost in fiscal years 2008–09, 2009–10, and 2010–11, and use the cost variation across these years to put upper and lower bounds on program costs. One challenge in calculating program costs is that the program included two rounds—"phase-1" and "phase-2"—but the impact evaluation only covers the phase-1 schools. While the cost data allows us to distinguish between subsidy costs for phase-1 and phase-2 program schools, we could not separate out other types of costs in the same way. In order to be conservative, we impute non-subsidy costs fully to phase-1 program schools. Additionally, Sindh experienced major floods in 2010, and the costs SEF incurred in helping to rehabilitate damaged program schools are included in the total costs. These two factors–assignment of total non-subsidy costs to phase-1 schools and costs for helping schools due to the floods–would inflate the estimated cost of the program.

In order to calculate the number of beneficiaries, we use data collected by SEF during an unannounced monitoring activity in February, 2011. During this visit, SEF found 28,827 children enrolled based on school registers, and 18,820 children in attendance based on a head count. Assuming a 20-percent student absence rate, as is typical in rural Sindh, we estimate an enrollment of 23,525. Depending on the fiscal year and number of students, the annual program cost per student therefore 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), and was only superior to a conditional cash transfer program in Africa. In Appendix C we present the details of the data and also lower and upper bounds of the cost-effectiveness numbers under different assumptions.

7 Structural estimation of program school efficiency

Having established the reduced form effects of the program and its cost effectiveness, we next turn to understanding the underlying mechanisms. We first present descriptive statistics of program school characteristics, and compare these to those of government schools in the study villages. This exercise helps us understand how the choice of school inputs by program-school operators fits in with the results discussed above.

We then extend the analysis to 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 means. We first estimate the supply and demand for school inputs, and the social surplus generated by program schools. Using the parameters of this model, we then determine the inputs chosen by the social planner, and estimate the share of the potential social surplus captured by program schools.

The experiment provides a unique opportunity for conducting this exercise in a credible manner. In non-experimental settings, endogenous school placement could bias the estimated parameters of the model. For example, unobservable village-level educational preferences may be correlated with school presence, selected inputs, and enrollment outcomes. Because the experiment exogenously varies the placement of schools across villages, it allows us to estimate the demand- and supply-side parameters necessary for conducting the structural exercise.

7.1 Program school inputs

Table 10 compares mean school characteristics across program and government schools, and Appendix Table A.2 compares mean school characteristics across program and private schools. Appendix Table A.3 compares school characteristics across the gender-uniform and gender-differentiated subsidy treatments. All observations come from the second followup survey, and means are estimated at the child-school level.

In Table 10, columns (1) and (3) report means and standard deviations for characteristics of program schools. Columns (2) and (4) report the differences in mean characteristics between program and government schools. Differences are estimated using a seemingly unrelated regressions (SUR) specification to account for within-school correlations in school characteristics. Program schools were open 5 days per week, which was 0.5 more days per week than government schools. Program schools were more likely to use English as the medium of instruction (+31.0 percentage points), and less likely to use Sindhi (-37.1 percentage points).[21] Physical infrastructure was generally better in program schools than in government schools: they were more likely to have an adequate number of desks (+14.9 percentage points), drinking water (+30.4 percentage points), and toilets (+28.7 percentage points).

According to responses by head masters, program schools had a larger number of teachers than government schools (+1 teacher), and a larger number of female teachers (+1.5 teachers). There were also more teachers with either less than five years of teaching experience (+2.5 teachers) or 5–10 years of teaching experience (+0.4 teachers), and fewer with more than 10 years of teaching experience (-2 teachers).

Based on information collected from individual teachers, we find that program-school teachers were more likely to be female (+23.8 percentage points), were younger (-13.8 years), and received lower monthly salaries (-11,492 rupees). In addition, program-school teachers had fewer years of teaching experience (-11.4 years). Program-school teachers spent a similar amount of time engaged in various classroom activities as government-school teachers.

It is important to note that there is substantial variation in school inputs *across* program schools. This marks a key difference from other PPP models, which generally involve greater centralized control over the amenities offered by publicly funded private schools. As we show subsequently in the structural exercise estimating the social planner's solution, this variation in inputs is key to identifying parameters for the supply and demand of school inputs. This autonomy also allows us to assess how effectively entrepreneurs chose inputs relative to what a social planner would have chosen.

In contrast to the sharp differences between program and government schools, program schools were relatively similar to private schools (Appendix Table A.3).[22] Most importantly, teacher characteristics were virtually identical across the program and private schools, with teachers who were relatively young, more likely to be female, and receiving a lower salary. The most salient differences are that private schools were more likely to charge tuition, require a uniform, and have drinking water; and were also substantially larger, with on average 43 more students and 2.5 more teachers, though these differences were not statistically significant.[23]

In Appendix Table A.3, columns (1) and (3) report the mean characteristics in program

[21] English as a medium of instruction in private schools has been been found in other studies (e.g., Muralidharan and Sundararaman, 2015), and is an example of private schools' tailoring supply to the demands of families.

[22] For this exercise, differences are estimated using OLS regressions, as the sample of private schools was too small to permit SUR estimation.

[23] This is likely because the program schools were only allowed to enroll children within the villages to which they were assigned.

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. There is no evidence that program-school operators structured school inputs differently across the gender-uniform and gender-differentiates subsidy schemes. This findings is consistent with the lack of a differential effect on girls' enrollment and test scores across the two subsidy treatments discussed earlier.

7.2 Program school efficiency

Because program-school operators may have incentives that are not perfectly aligned with those of the social planner, it is unclear that market incentives will lead them to choose optimal school inputs. 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 an input 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 educational attainment. 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 three steps. First, we estimate a discrete choice model of household demand for schools (referred to as "child" demand). This allows us to compute both the expected distribution of school enrollment—which in turn determines school-operator revenues—as well as consumer surplus, under both observed and counterfactual school-input configurations. Second, we estimate the opportunity costs of providing school inputs using a simple revealed preference argument, which allows us to calculate the cost of providing input configurations. Third, we calculate the social value of school-input configurations based on surplus accruing to students, school-operator input costs, and the social value of education.

As noted above, the experiment removes endogeneity in school placement, which would otherwise bias the estimated parameters of the model. In order to solve the model, we impose the additional assumption that both student demand and input costs are homogeneous across villages. The latter assumption is necessary due to a lack of variation in program school characteristics within a given village, as there was only a single program school per village and their characteristics were fixed during the sample period. Therefore, while some parameters of the demand model are identified through student-school interactions within a village, all of the cost parameters are identified only via cross-village variation. Second, we also assume that program-school operators and the social planner both have full information about demand and costs.

We begin by estimating the demand for schooling by children in the villages. This allows us to estimate consumer surplus, compute how that surplus changes with changes in school inputs, and predict enrollments under counterfactual school configurations. We estimate demand using a standard logit random utility framework. Each child makes a single choice from a set of schools, J, where the utility for child i of choice $j \in J$ is given by:

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

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 deterministic utility of not going to school to zero.

For the demand 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, and distance from home to the school. School inputs consist of toilets and/or drinking water (a single indicator variable); as well as mean 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 supply-side factors for girls' enrollment and learning, such as distance 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. Each column includes successively more controls, and in column (4) we include an indicator for government schools. As expected, we find that boys and older children are more likely to be enrolled. We also find that 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 for boys, but that this effect is sharply attenuated for female students. While distance has a negative effect on enrollment, there is only weak evidence that this effect is stronger for girls than boys.

We next use the demand curve to estimate bounds on school-input opportunity 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, the percentage of less experienced teachers, and the teacher absenteeism rate. We assume that schools will provide an input if its cost did not exceed the additional revenue that it generates through increased enrollment; and that for schools that did not provide the input, the opposite must be true. These two inequalities bound the opportunity cost of the input. This exercise requires the use of a structural model, since we must 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 effect was far smaller 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. This result is consistent with existing research from Pakistan, which finds that female teachers in private schools earn 33 percent less than their male counterparts, after controlling for other characteristics (Andrabi et al., 2008a). 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 also estimates a negative cost for post-secondary teachers, suggesting that they were less expensive than less-educated teachers, which seems to contradict the higher salaries they receive. However, this is likely because post-secondary teachers were more productive than those with lower educational qualifications. In results not shown, we find that post-secondary educated teachers worked longer hours than other teachers, and were more likely to have the rank of head teachers.

Finally, we combine these pieces to determine whether program-school operators provide inputs that maximize social welfare. 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 child enrollment (i.e., an additional year of schooling).[24] We assume that the social value of education is related to overall enrollment, 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 a 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 piece of the welfare analysis deals with 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 enrollment, and τ captures the marginal (social) utility of higher enrollment. 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.[25] To account for both of these costs, we assume a deadweight loss of 30 percent and multiply this by the total cost of each school. This specification implicitly assumes that all profits are returned to the government (as would be the case on a cost-plus zero contract) and distortions are only incurred in raising tax revenues.

We define the social value of education as the product of the student's annual adult income

[24]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.

[25]See footnote 13.

and a social externality multiplier. Using the estimates from Montenegro and Patrinos (2014) for the returns to education in Pakistan, we fix upper- and lower-bound wage gains from an additional year of education at 10.8 percent and 6.8 percent, respectively. The wage gain is calculated as a function of the baseline wage and the labor force participation rate:

$$\triangle\, wage_{gb} = blwage_g \times \triangle enrolled_g \times participationrate_g, \tag{7}$$

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.[26] 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 therefore solve this by exhaustively computing 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, wages, and costs for every possible input configuration in program schools. We assume that the inputs of other schools remain constant as the program school's inputs are adjusted. This assumption 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 of school inputs across the solutions of program-school operators and the social planner, and the estimated social surplus associated with each solution. Column (1) shows the actual inputs provided in program schools, and the associated upper and lower bound for the social surplus (which depends on the wage return to education), with the social externality parameter τ assumed to be equal to 1. Columns (2)–(6) show the social planner solutions for various levels of τ. In addition, we show the change in cost, consumer surplus, enrollment, and income associated with each school-input configuration relative to the values for the program schools.[27]

The primary result is that program-school operators have proven relatively successful at establishing and operating schools that generated a substantial share of the possible surplus

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

[27] Social surplus, cost, consumer surplus, and income are winsorized at the 5th and 95th percentiles.

in the environment. Assuming a social value of education equal to one, the program schools generate a social surplus of approximately 94 percent the potential surplus of the social planner. Large variation exists across villages, from a lower bound of ten percent to an upper bound of 100 percent (i.e., the program school inputs are identical to those of the social planner).

The social planner achieves these gains through various changes to program schools. First, under the social planner, a larger share of schools have toilets and/or drinking water (+8 percentage points relative to the program-school operator solution). The social planner employs more teachers with post-secondary education (+23 percentage points), and fewer teachers with less than five years of teaching experience (-22 percentage points); and allows a smaller share of teachers to be absent 2 or more days per month (-32 percentage points). The composition of female teachers is also substantially lower (-34 percentage points) in the social planner's solution.

To understand why the social planner chooses these inputs, Table 13 also reports the changes in consumer surplus, enrollment, input costs, and income. On average, the social planner chooses inputs that increase costs, but which also increase consumer surplus, enrollment, and income. In other words, program school operators appear to be foregoing some socially beneficial increase in enrollment in order to reduce costs and thereby increase their private profit.

One of the key parameters of the social planner's solution is the social value of education. Because this parameter does not come from any empirical or model-based foundation, it is important to understand how robust our results are to different assumptions of its value. In Table 13, columns (3)–(6) report the results when the social planner places weights of 0, 0.5, 1.5, and 2, respectively, on the income effects of enrollment. The optimal education, teaching experience, and absenteeism of teachers are declining in the social value of education, as is the share of female teachers. In contrast, the optimal provision of toilets and/or drinking water is increasing in the social value of education. These changes are being driven by the greater importance of increasing enrollment when the social returns to education are large, even at the expense of some loss of profitability for school operators.

A natural question in this setting is whether welfare could have been improved by relocating the program schools from some program villages to some control villages. In Appendix Table A.6, we use the parameters of our structural model to ask whether the social planner could have increased the social surplus by reallocating some of the program schools to treatment villages. For this exercise, we added a synthetic program school to each village in the control group. Since there were no baseline program school characteristics for control villages, we simulated student outcomes, entrepreneur profits, and costs for all possible combinations of school characteristics. We then took the optimal school configuration within each village and ranked the social welfare gains against the optimal program schools. We find that 31 of the schools should have been allocated to control villages, and that this would have yielded an approximately 13% increase in total social surplus compared to the baseline social planner's solution.

Two aspects of the above calculations deserve emphasis. First, because men have higher labor force participation and labor earnings than women, factors improving boys' enrollment are given greater weight than those that increase girls' enrollment. This can be seen with respect to female teachers, where increases in the social return to education leads to a decline in the optimal percentage of female teachers, which is driven by the lower preference for female teachers by boys. However, this result does not account for the possibility that female labor earnings and labor force participation may increase over time. In addition, while the model assumes an identical social value of education across genders, if women provide additional services—such as caretakers for children, or potential future teachers in local schools—then the social value of female education may be higher than that of men. Second, the social planner's solution is village-specific. This means that, while the statistics given in Table 13 ostensibly show program-school operators to have provided inputs similar to those in the social planner's solution, the correspondence in mean inputs does not necessarily imply that the village-specific solutions are similarly close.

One important caveat to the foregoing analysis is that the social planner's solution does not account for supply constraints that may face the entrepreneur. For example, based on the sub-sample used in the followup survey, in 53% of the villages there were no women with an eighth grade education or better; and in 48% of villages there were no adults with a post-secondary education.[28] It is unclear how large a role this constraint plays: a regression of the share of female teachers on the number of village women with an eighth grade education or better shows only a small relationship between the two, and a similarly small relationship holds for the share of post-secondary teachers, indicating that entrepreneurs are relatively successful at recruiting teachers from surrounding areas.

8 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

[28]Indeed, in later rounds of the program, SEF was forced to relax the constraint that the entrepreneurs identify potential female teachers with at least an eighth grade education due to the lack of qualifying individuals.

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. 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 remarkably close to the social planner, and enrollment is higher than would have been achieved in the social planner's solution. Compared to program-school operators, the social planner hires more female teachers, and also hires more post-secondary-educated teachers. Interestingly, the greater reliance on female teachers in the social planner's solution is driven by cost considerations, as female teachers are in fact associated with a decline in enrollment, driven primarily by the response of male students. 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 governments to set contracts with private, 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.

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.

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. 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. (2019). Teacher value-added in a low-income country. *American Economic Journal: Economic Policy (forthcoming).*

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.

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

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.

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.

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

Romero, M. and Sandefur, J. (2019). Beyond short-term learning gains: The impact of outsourcing schools in liberia after three years.

Romero, M., Sandefur, J., Sandholtz, W. A., et al. (2019). Outsourcing education: Experimental evidence from liberia. *American Economic Review (forthcoming)*.

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

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

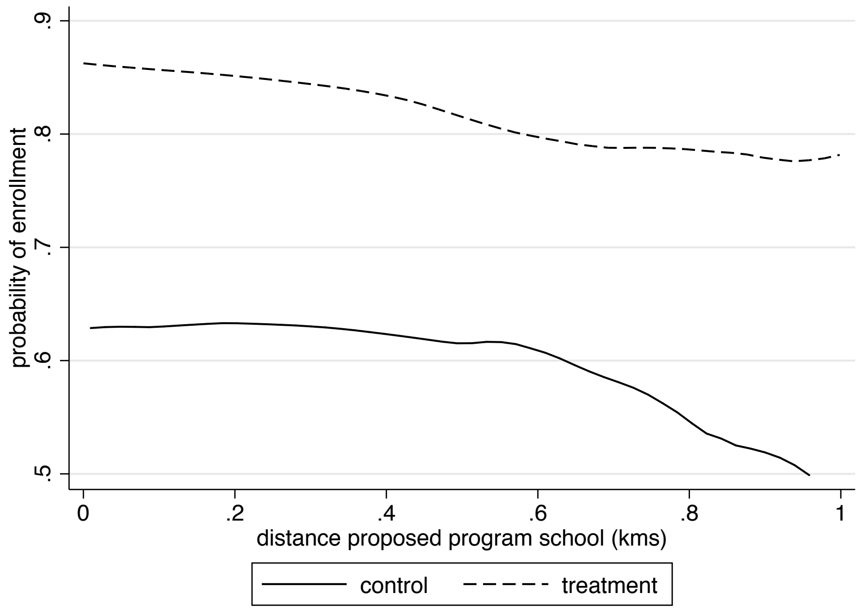
Todd, P. E. and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485).

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

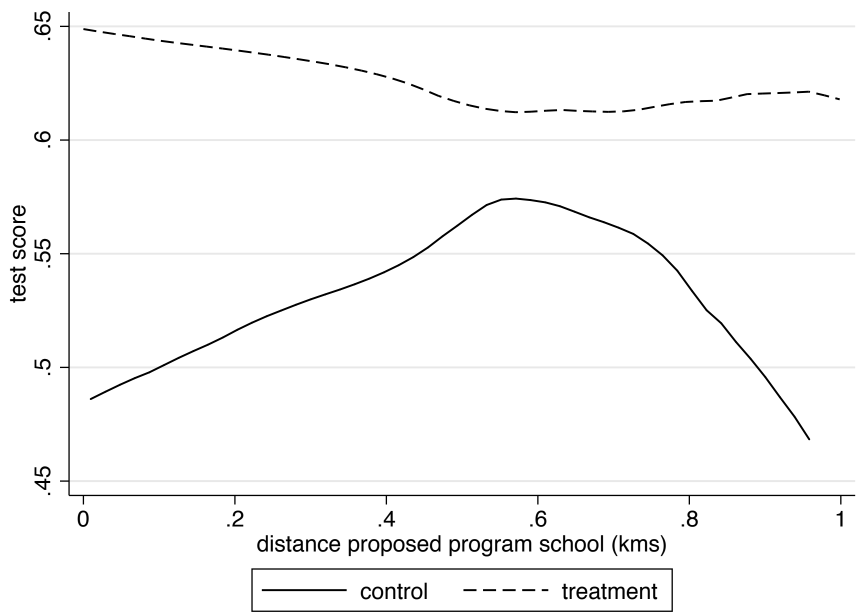
van der Linden, W. J. and Hambleton, R. K. (2013). *Handbook of modern item response theory*. Springer Science & Business Media.

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

Figure 1: School Proximity and Educational Outcomes



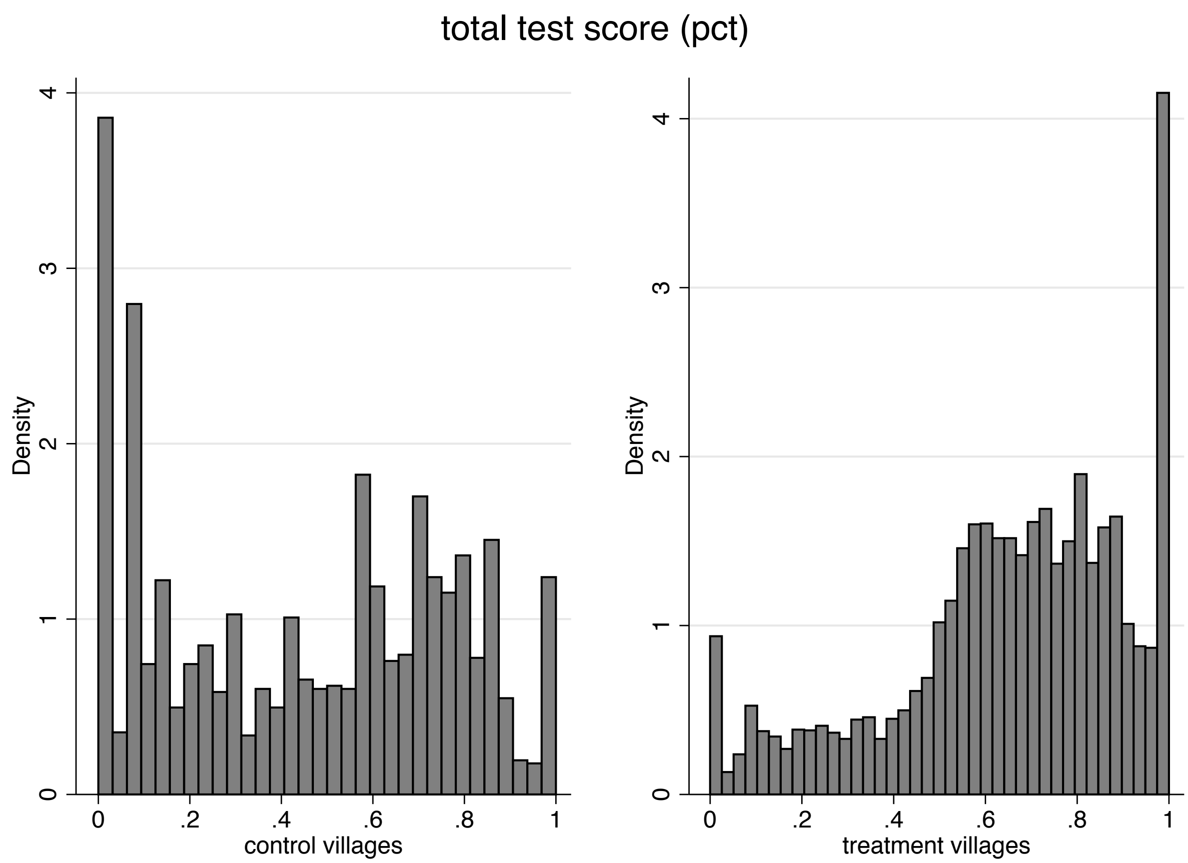
1.1: Enrollment



1.2: Test Score

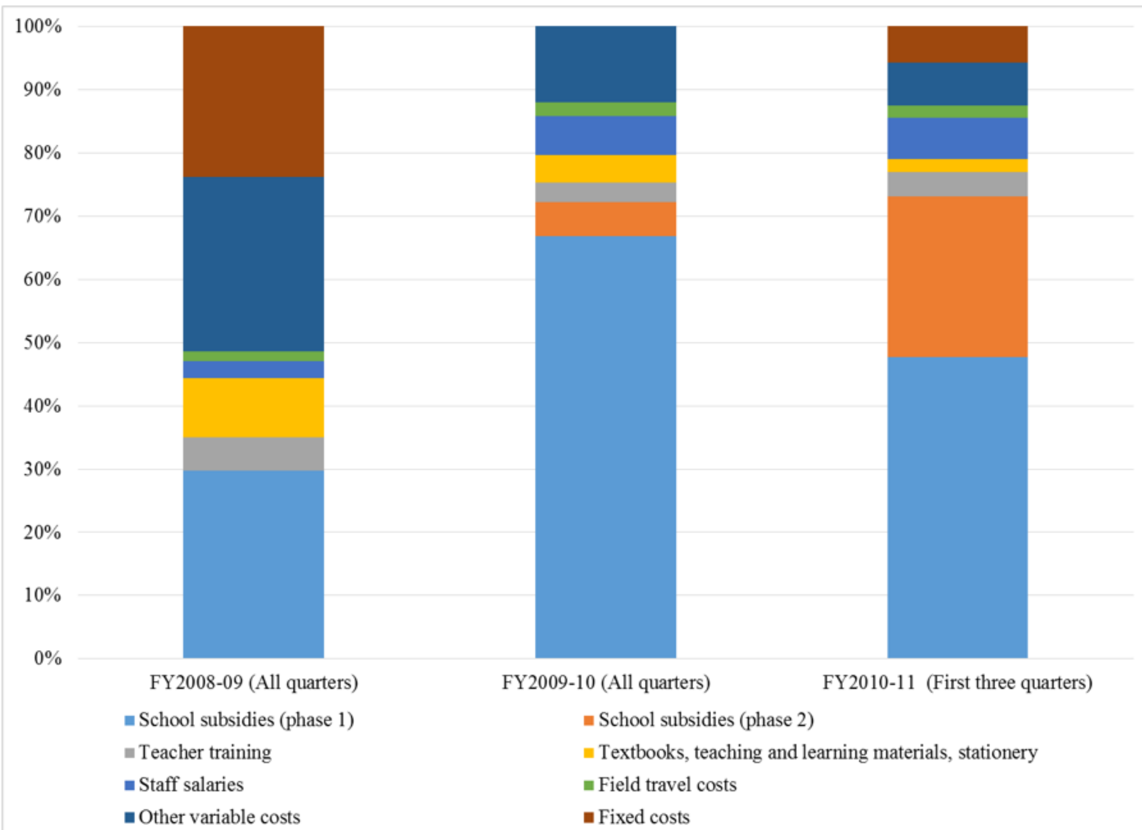
Notes: Figure 1.1 plots the probability of enrollment for children aged 5–10 against the distance to the proposed program school site, disaggregated by treatment status. Figure 1.2 plots the test score against the distance to the proposed program school site, disaggregated by treatment status.

Figure 2: Treatment and Test Scores



Notes: Figure 2 shows the distribution of test scores (for children aged 5–10) disaggregated by treatment status. Test scores are measured as the percentage of correct answers.

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	Treatment	Uniform	Differentiated	Total
	(1)	(2)	(3)	(4)	(5)
Number Villages	38	161	82	79	199
Full Census					
Number Households w/Young Children	1530	7109	3795	3314	8639
Number Young Children	4567	20590	11231	9359	25157
Followup Sample					
Number Households w/Young Children	1069	4897	2594	2303	5966
Number Young Children	3093	14627	7718	6909	17720

Note: This table reports sample sizes by treatment status. Treatment denotes pooled treatment; Uniform, the gender-Uniform subsidy treatment; and Differentiated, the gender-differentiated subsidy treatment.

Table 2: Balance Across Program and Control Villages

	Baseline		Followup	
	Control	Treatment - Control	Control	Treatment - Control
	(1)	(2)	(3)	(4)
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)
Brick House			0.057	-0.005
				(0.024)
Semi-brick House			0.193	-0.016
				(0.065)
Mud House			0.511	0.084
				(0.076)
Thatched Hut			0.240	-0.064
				(0.071)
Goats			3.916	-0.052
				(0.793)
Sect: Sunni			0.877	0.034
				(0.060)
Language: Urdu			0.114	0.044
				(0.043)
Language: Sindhi			0.664	0.062
				(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. Treatment 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

	Enrollment				Highest Grade Attained
	(1)	(2)	(3)	(4)	(5)
Panel A: Children Aged 5–10					
Treatment	0.316***	0.315***	0.313***	0.316***	0.381***
	(0.066)	(0.066)	(0.064)	(0.065)	(0.120)
Control Mean	0.529				0.806
N	11572	11572	11572	11572	11118
R-squared	0.086	0.087	0.103	0.109	0.225
Panel B: Children Aged 11–17					
Treatment	0.109*	0.112*	0.109**	0.111**	-0.009
	(0.057)	(0.058)	(0.049)	(0.052)	(0.319)
Control Mean	0.375				2.129
N	5584	5584	5584	5584	5361
R-squared	0.006	0.039	0.097	0.148	0.132
Child Controls	No	Yes	Yes	Yes	Yes
HH Controls	No	No	Yes	Yes	Yes
District F.E.s	No	No	No	Yes	Yes

Note: This table reports program impacts on enrollment and highest grade attained at follow-up measurement. Treatment denotes pooled treatment. Panel A limits the sample to children aged 5–10. Panel B limits the sample to children aged 11–17. Means of outcome variables in control villages are reported in the second row of each panel. 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

	Control Mean	Treatment Effects, Z-Score				
		ITT				TOT
	(1)	(2)	(3)	(4)	(5)	(6)
Math Score	0.460	0.534***	0.525***	0.524***	0.631***	1.962***
	(0.308)	(0.155)	(0.158)	(0.156)	(0.124)	(0.285)
Language Score	0.487	0.502***	0.495***	0.493***	0.594***	1.821***
	(0.341)	(0.171)	(0.174)	(0.171)	(0.129)	(0.230)
Total Score	0.470	0.540***	0.531***	0.529***	0.637***	1.964***
	(0.310)	(0.166)	(0.170)	(0.167)	(0.130)	(0.263)
Child Controls		No	Yes	Yes	Yes	Yes
HH Controls		No	No	Yes	Yes	Yes
District F.E.s		No	No	No	Yes	Yes

Note: This table reports program impacts on standardized test scores (for children aged 5–10). Column (1) gives the mean percent of correct answers in control villages, with the standard deviation reported in parentheses. Columns (2) through (5) report the intention-to-treat (ITT) impacts on test scores, with various sets of controls. Test scores are standardized using the mean and standard deviation from control villages. Column (6) reports the treatment-on-the-treated (TOT) impacts on test scores for enrolled children. 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: Treatment and Test Scores, Disaggregated by Question Type

	Control Mean	Treatment Effect, Pct Correct				
		ITT				TOT
	(1)	(2)	(3)	(4)	(5)	(6)
Language						
Full Letter Identification	0.576	0.185***	0.183***	0.181***	0.222***	0.683***
	(0.401)	(0.064)	(0.065)	(0.064)	(0.052)	(0.092)
Half Letter Identification	0.458	0.174***	0.172***	0.171***	0.202***	0.617***
	(0.377)	(0.058)	(0.059)	(0.058)	(0.044)	(0.085)
First Letter of Objects	0.474	0.170***	0.167***	0.167***	0.197***	0.601***
	(0.372)	(0.053)	(0.054)	(0.053)	(0.043)	(0.080)
Reading Full Word	0.432	0.168***	0.164***	0.164***	0.202***	0.622***
	(0.418)	(0.060)	(0.061)	(0.060)	(0.042)	(0.086)
Reading Full Sentence	0.406	0.122	0.119	0.119	0.158***	0.482***
	(0.491)	(0.077)	(0.077)	(0.077)	(0.044)	(0.091)
Numeracy						
Size Comaparison of Objects	0.729	0.108***	0.107**	0.106***	0.139***	0.432***
	(0.377)	(0.041)	(0.041)	(0.040)	(0.032)	(0.070)
Number Awareness	0.526	0.189***	0.187***	0.186***	0.220***	0.687***
	(0.499)	(0.053)	(0.054)	(0.052)	(0.044)	(0.093)
Number Identification	0.457	0.191***	0.188***	0.188***	0.211***	0.656***
	(0.367)	(0.050)	(0.051)	(0.050)	(0.042)	(0.086)
Number Ordering	0.429	0.199***	0.196***	0.195***	0.224***	0.696***
	(0.417)	(0.045)	(0.046)	(0.046)	(0.038)	(0.104)
Object Counting	0.449	0.185***	0.182***	0.182***	0.212***	0.661***
	(0.427)	(0.051)	(0.052)	(0.052)	(0.044)	(0.095)
Object Counting & Comparison	0.423	0.187***	0.185***	0.184***	0.217***	0.666***
	(0.420)	(0.049)	(0.049)	(0.048)	(0.041)	(0.099)
Shape Identification	0.419	0.175***	0.172***	0.172***	0.212***	0.657***
	(0.432)	(0.057)	(0.058)	(0.057)	(0.044)	(0.094)
Addition	0.428	0.160**	0.156**	0.157**	0.201***	0.631***
	(0.406)	(0.064)	(0.065)	(0.064)	(0.045)	(0.125)
Subtraction	0.347	0.145***	0.142***	0.143***	0.175***	0.542***
	(0.365)	(0.054)	(0.054)	(0.054)	(0.039)	(0.111)
Telling Time	0.370	0.128**	0.124**	0.124**	0.160***	0.498***
	(0.406)	(0.057)	(0.058)	(0.057)	(0.043)	(0.112)
Child Controls		No	Yes	Yes	Yes	Yes
HH Controls		No	No	Yes	Yes	Yes
District F.E.s		No	No	No	Yes	Yes

Note: This table reports program impacts on test scores (for children aged 5–10), measured as the percentage of questions answered correctly, with test questions disaggregated by the competency being measured. 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. Column (6) reports the treatment-on-the-treated (TOT) impacts on test scores for enrolled children. 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: Differential Impacts by Subsidy Treatment and by Gender

	Enrollment	Highest Grade Attained	Test Score
	(1)	(2)	(3)
Panel A: Subsidy Treatments			
Uniform	0.316***	0.368***	0.620***
	(0.065)	(0.124)	(0.134)
Differentiated – Uniform	-0.001	0.026	0.037
	(0.024)	(0.067)	(0.063)
N	11572	11118	10327
R-squared	0.109	0.225	0.203
Panel B: Gender			
Treatment	0.325***	0.393***	0.613***
	(0.066)	(0.130)	(0.136)
Treatment X Female	-0.019	-0.024	0.059
	(0.027)	(0.065)	(0.053)
Control Mean:			
Male	0.529	0.828	0.480
Female	0.528	0.773	0.457
N	11521	11067	10286
R-squared	0.109	0.225	0.202

Note: This table reports program impacts on outcomes (for children aged 5–10) by subsidy treatment (panel A), and by gender (panel B), with the full set of child and household controls and district fixed effects. Treatment denotes pooled treatment; Uniform, the gender-Uniform subsidy treatment; and Differentiated, 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: Gender Differential Impacts by Subsidy Treatment

		Enrollment	Highest Grade Attained	Test Score
		(1)	(2)	(3)
Uniform		0.333***	0.410***	0.586***
		(0.066)	(0.136)	(0.139)
Uniform X Female		-0.036	-0.092	0.079
		(0.030)	(0.080)	(0.055)
Differentiated		0.316***	0.373***	0.645***
		(0.068)	(0.135)	(0.140)
Differentiated X Female		-0.000	0.050	0.032
		(0.027)	(0.063)	(0.060)
N		11521	11067	10286
R-squared		0.109	0.226	0.203
H0: Uniform = Differentiated	F-stat	0.553	0.271	0.864
	p-value	0.458	0.603	0.354
H0: Uniform + Uniform X Female =	F-stat	0.384	1.773	0.029
Differentiated + Differentiated X Female	p-value	0.536	0.184	0.866
H0: Uniform X Female =	F-stat	2.826	4.201	1.256
Differentiated X Female	p-value	0.094	0.042	0.264

Note: This table reports gender-differential impacts on outcomes (for children aged 5–10) by subsidy treatment, with the full set of child and household controls and district fixed effects. Uniform denotes the gender-Uniform subsidy treatment; and Differentiated, 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 8: Test Scores by School Type

	Program	Govt – Program	Priv - Program	p-value Govt = Priv	Govt-Enrolled Treatment – Control
	(1)	(2)	(3)	(4)	(5)
Math Score	0.715	-0.236***	-0.049	0.433	-0.053
		(0.081)	(0.226)		(0.087)
Urdu Score	0.709	-0.160***	-0.025	0.314	-0.015
		(0.061)	(0.125)		(0.075)
Total Score	0.734	-0.211***	-0.040	0.391	-0.040
		(0.073)	(0.188)		(0.080)

Note: This table reports differences in mean standardized test scores (for children aged 5–10) according to the type of school children are enrolled in, with the full set of child and household controls and district fixed effects. Column (1) reports mean test scores for children enrolled in program schools. Columns (2) and (3) give the coefficients for an indicator variables denoting whether children are enrolled in government or private schools, respectively. Column (4) gives the p-value for a test of equality of government and private school coefficients. The sample is limited to children who either reside in a treatment village and are enrolled in a program school, or who reside in a control village and are enrolled in either a government or private school. Column (5) limits the sample to children enrolled in a government school in either a treatment or control village, and reports the difference in test score across control and treatment villages. 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: Program Impacts on Aspirations

	Control	Treatment - Control	Female	Treatment	Treatment X Female
	(1)	(2)	(3)	(4)	(5)
Panel A: Parental 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 parental-reported aspirations for the child (panel A) and child-reported aspirations (panel B) (for children aged 5–10), with the full set of child and household controls and district fixed effects. Column (1) reports mean aspirations in control villages for the indicated variables, and column (2) reports differences in mean aspirations between program and control villages. Columns (3), (4), and (5) report coefficients from a regressions of the indicated variable on indicator variables for girls, program status, and the interaction of the two. Treatment 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 10: Characteristics by School Type

	Program Mean/sd	Govt – Program		Program Mean/sd	Govt – Program
	(1)	(2)		(3)	(4)
Characteristics from School Survey			Students		
Days Operational	5.122	-0.513*	Number Boys	88.711	-18.766
	(1.371)	(0.285)		(45.290)	(11.652)
Open Admission	0.858	0.025	Number Girls	71.294	-30.787***
	(0.349)	(0.050)		(34.721)	(5.818)
Uniform Required	0.024	-0.024	Percent Female Students	0.448	-0.039
	(0.152)	(0.017)		(0.138)	(0.049)
Tuition Required	0.000	0.000	Student-teacher Ratio	44.250	-0.221
	(0.000)	(0.000)		(14.269)	(3.966)
Medium: Sindhi	0.612	0.371***	**Characteristics from Teacher Survey**		
	(0.487)	(0.050)	Days Absent/Month	0.838	0.172
Medium: English	0.310	-0.310***		(1.121)	(0.312)
	(0.463)	(0.045)	Female	0.493	-0.238***
Number Teachers				(0.424)	(0.073)
Total	3.782	-0.948***	Age	25.170	13.826***
	(1.592)	(0.340)		(4.171)	(1.460)
Female	1.986	-1.449***	Education	10.967	0.919***
	(1.806)	(0.199)		(0.762)	(0.169)
Post-secondary	1.898	0.437	Salary (1000s Rs)	4.066	11.492***
	(1.942)	(0.459)		(1.255)	(1.012)
<5 years	3.132	-2.479***	Years Teaching	2.784	11.404***
	(1.486)	(0.182)		(1.270)	(1.331)
$5-10$ years	0.603	-0.406***	Years Teaching this School	1.773	4.909***
	(0.965)	(0.122)		(0.848)	(0.968)
>10 years	0.047	1.980***	Hours Teaching (per week)		
	(0.211)	(0.314)	Total	26.916	1.209
Average Teacher Absent ≥ 2	0.396	0.035		(6.724)	(1.876)
Days/Month	(0.489)	(0.101)	Teaching Whole Class	6.207	0.488
Amenities				(4.677)	(0.807)
Building	0.961	-0.024	Teaching Small Group	5.415	-0.293
	(0.194)	(0.038)		(4.367)	(0.678)
Number Classrooms	3.229	-0.466	Teaching Individual	5.467	0.351
	(1.413)	(0.398)		(4.832)	(0.731)
Sufficient Desks	0.756	-0.149*	Blackboard/Dictation	5.177	-0.252
	(0.429)	(0.084)		(4.042)	(0.729)
Drinking Water	0.846	-0.304***	Classroom Management	3.225	0.282
	(0.361)	(0.105)		(2.637)	(0.464)
Electricity	0.724	-0.067	Testing	3.217	-0.448
	(0.447)	(0.070)		(2.769)	(0.503)
Toilet	0.788	-0.287***	Administrative	2.497	0.343
	(0.409)	(0.108)		(1.895)	(0.335)

Note: This table reports differences in mean characteristics between program and government schools. The unit of observation is child-school. Columns (1) and (3) report means and standard deviations for program schools; and columns (2) and (4) differences in means between program and government schools. Differences are estimated using a seemingly unrelated regressions (SUR) specification to account for within-school correlations in school characteristics. 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	0.344***	-0.285	-0.122	1.639***
	(0.138)	(0.262)	(0.176)	(0.193)
Toilets and/or Drinking Water	0.841***	0.919***	0.855***	0.543***
	(0.075)	(0.076)	(0.078)	(0.104)
Student Female	-0.002	0.016	-0.295***	-0.312**
	(0.048)	(0.054)	(0.097)	(0.159)
Student Age	0.031**	0.035***	0.034***	0.032***
	(0.015)	(0.014)	(0.014)	(0.011)
Distance from Home to School	-0.193***	-0.155***	-0.151***	-0.073*
	(0.032)	(0.035)	(0.050)	(0.049)
Pct Teachers with $<$ 5 Yrs Teaching		0.0617***	0.623***	-0.413***
		(0.080)	(0.075)	(0.091)
Pct Teachers Post-secondary		-0.577***	-0.581***	-0.383***
		(0.080)	(0.065)	(0.068)
Pct Teachers Female		-0.235***	-0.458***	-0.664***
		(0.056)	(0.069)	(0.081)
Pct Time Teaching		0.816***	0.799***	0.094
		(0.232)	(0.149)	(0.179)
Avg Teacher Absent $\geq$ 2 Day/month		-0.040	-0.038	-0.207***
		(0.049)	(0.051)	(0.051)
Pct Female Teachers X Female Student			0.480***	0.515***
			(0.104)	(0.113)
Distance X Female Student			-0.002	0.034
			(0.072)	(0.069)
Toilets and/or Drinking Water X Female Student			0.118	0.089
			(0.096)	(0.167)
Tuition Cost per Year	-0.007***	0.005***	-0.005***	-0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
Govt School				-1.526***
				(0.077)

Note: This table reports schooling demand estimates (for children aged 5–10). 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	3.358***
	(0.154)
Pct Teachers Female	-2.618***
	(0.231)
Pct Teachers Post-secondary	-3.413***
	(0.242)
Pct Teachers < 5 Yrs Experience	-3.681***
	(0.253)
Avg Teacher Absent $\geq$ 2 Days/month	-1.658***
	(0.109)

Note: This table reports cost estimates (for children aged 5–10). Standard errors are reported in parentheses. *, **, and *** denote statistical significance at the ten-, five-, and one-percent levels, respectively

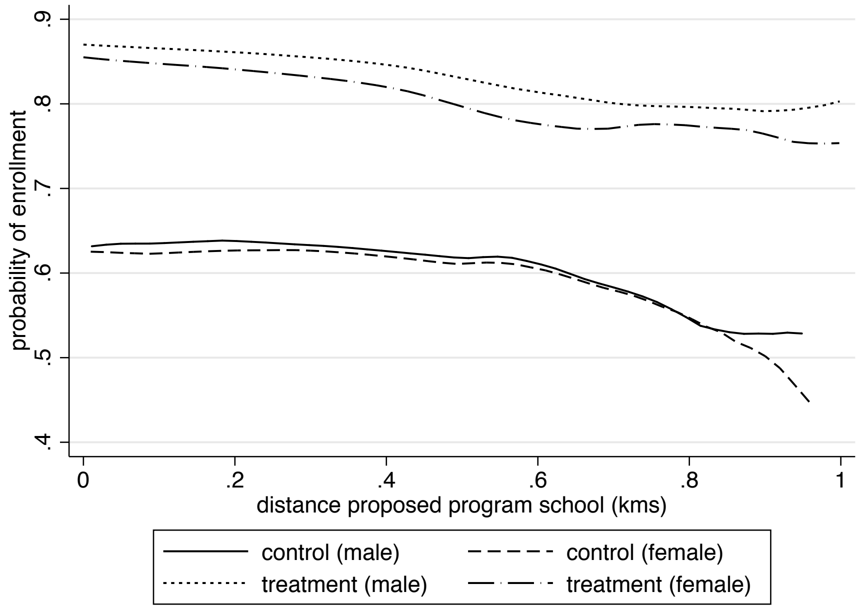
Table 13: Estimated Social Planner Solution

	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	0.98	0.03	0.92	0.99	0.99
	(0.30)	(0.14)	(0.16)	(0.27)	(0.12)	(0.12)
Pct Teachers Female	0.49	0.15	1.00	0.39	0.08	0.04
	(0.41)	(0.34)	(0.00)	(0.46)	(0.26)	(0.19)
Pct Teachers Post-secondary	0.48	0.71	1.00	0.94	0.52	0.35
	(0.34)	(0.42)	(0.00)	(0.23)	(0.48)	(0.45)
Pct Teachers <5 Yrs Experience	0.85	0.63	1.00	0.92	0.43	0.28
	(0.25)	(0.46)	(0.00)	(0.26)	(0.47)	(0.41)
Avg Teacher Absent $\geq$2 Days/month	0.37	0.05	1.00	0.11	0.02	0.01
	(0.48)	(0.21)	(0.00)	(0.32)	(0.14)	(0.12)
Change in Cost (1,000)		9.26	-76.36	-28.67	35.41	57.99
Change in Consumer Surplus (1,000)		2.07	-5.03	0.16	3.12	3.85
Change in Enrollment		3.14	-13.35	0.21	4.49	5.37
Change in Income (Upper Bound, 1,000)		15.52	-59.66	2.24	21.10	24.62
Change in Income (Lower Bound, 1,000)		9.77	-37.56	1.41	13.28	15.50
Total Surplus (Upper Bound, 1,000)	232.24	247.05	59.21	146.66	351.83	453.48
Total Surplus (Lower Bound, 1,000)	161.53	170.59	59.21	110.89	234.05	300.57

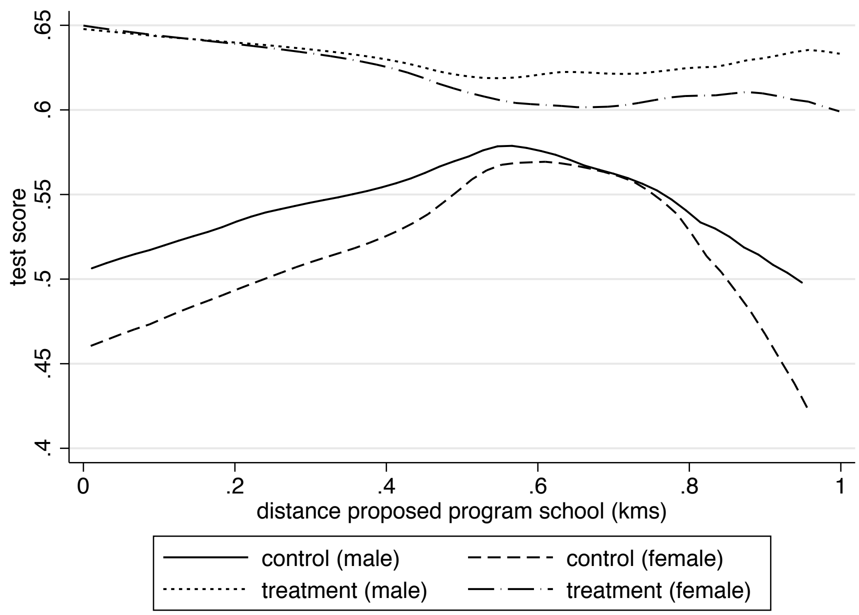
Note: This table presents the social planner's solution and the observed program solution.

Appendix A

Figure A1: School Proximity and Educational Outcomes



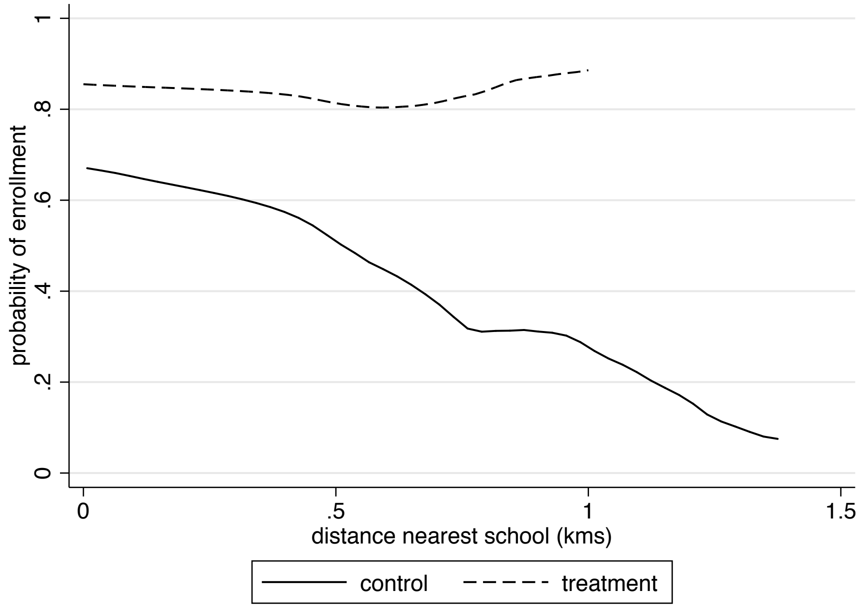
A1.1: Enrollment



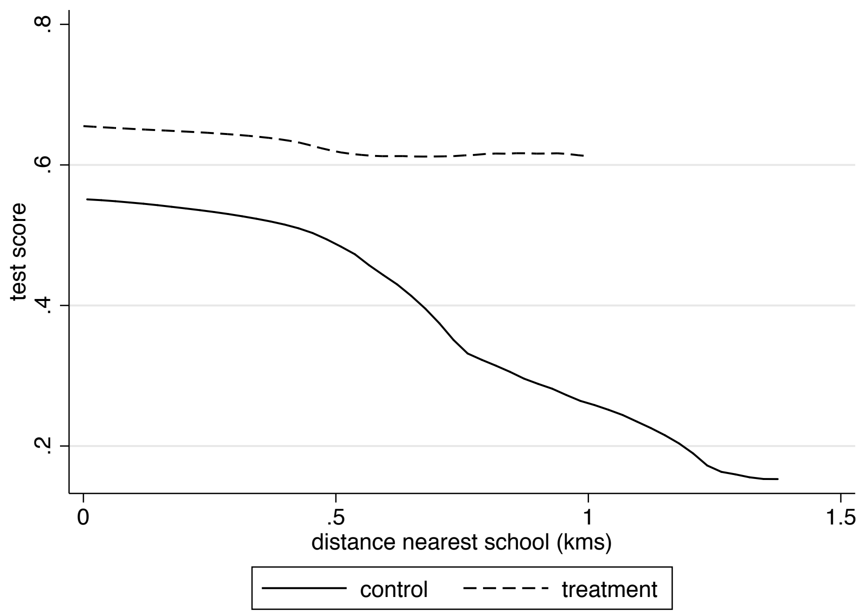
A1.2: Test Score

Notes: Figure A1.1 plots the probability of enrollment for children aged 5–10 against the distance to the proposed program school site, disaggregated by treatment status and gender. Figure A1.2 plots the test score against the distance to the proposed program school site, disaggregated by treatment status and gender.

Figure A2: School Proximity and Educational Outcomes



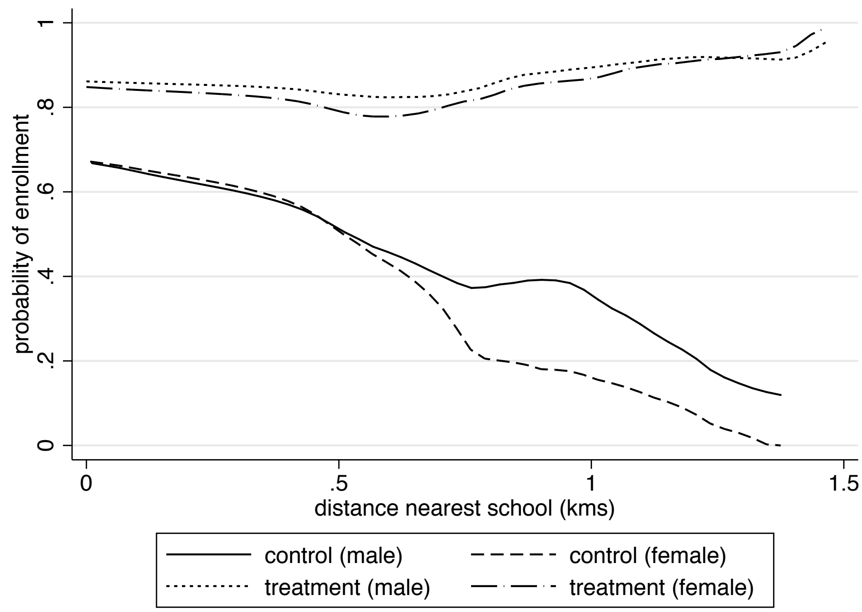
A2.1: Enrollment



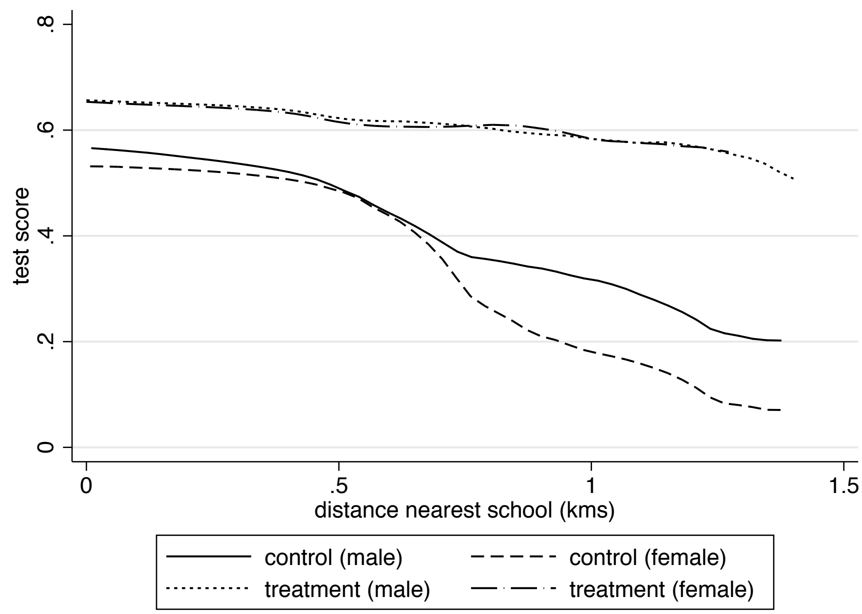
A2.2: Test Score

Notes: Figure A2.1 plots the probability of enrollment for children aged 5–10 against the distance to the proposed program school site, disaggregated by treatment status. Figure A2.2 plots the test score against the distance to the proposed program school site, disaggregated by treatment status.

Figure A3: School Proximity and Educational Outcomes



A3.1: Enrollment



A3.2: Test Score

Notes: Figure A3.1 plots the probability of enrollment for children aged 5–10 against the distance to the nearest school, disaggregated by treatment status. Figure A3.2 plots the test score against the distance to the nearest school, disaggregated by treatment status.

Table A.1: Balance Across Treatment Groups

	Baseline		Followup	
	Treat_GU Mean (1)	Treat_GD-Treat_GU (2)	Treat_GU Mean (3)	Treat_GD-Treat_GU (4)
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)
Brick House			0.046	0.016
				(0.026)
Semi-brick House			0.194	-0.023
				(0.056)
Mud House			0.604	-0.023
				(0.065)
Thatched Hut			0.156	0.030
				(0.068)
Goats			3.878	0.256
				(0.834)
Sect: Sunni			0.910	-0.012
				(0.047)
Language: Urdu			0.152	-0.004
				(0.046)
Language: 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: Characteristics by School Type

	Program	Private - Program		Program	Private - Program
	(1)	(2)		(3)	(4)
Characteristics from School Survey			Students		
Days Operational	5.122	-0.331	Number Boys	88.711	42.460
	(1.371)	(0.572)		(45.290)	(55.158)
Open Admission	0.858	0.075	Number Girls	71.294	18.743
	(0.349)	(0.074)		(34.721)	(29.671)
Uniform Required	0.024	0.312*	Percent Female Students	0.448	-0.010
	(0.152)	(0.181)		(0.138)	(0.051)
Tuition Required	0.000	0.441**	Student-teacher Ratio	44.250	-5.935
	(0.000)	(0.180)		(14.269)	(7.623)
Medium: Sindhi	0.612	0.028	**Characteristics from Teacher Survey**		
	(0.487)	(0.188)	Days Absent/Month	0.838	-0.324
Medium: English	0.310	0.050		(1.121)	(0.274)
	(0.463)	(0.187)	Female	0.493	0.001
Number Teachers				(0.424)	(0.182)
Total	3.782	2.501	Age	25.170	-0.501
	(1.592)	(2.009)		(4.171)	(1.306)
Female	1.986	3.407**	Education	10.967	0.832***
	(1.806)	(1.654)		(0.762)	(0.278)
Post-secondary	1.898	1.476*	Salary (1000s Rs)	4.066	-0.254
	(1.942)	(0.845)		(1.255)	(0.551)
< 5 years	3.132	-0.945	Years Teaching	2.784	0.676
	(1.486)	(0.682)		(1.270)	(0.773)
$5 - 10$ years	0.603	3.092	Years Teaching this School	1.773	0.927
	(0.965)	(2.355)		(0.848)	(0.733)
> 10 years	0.047	0.354	Hours Teaching (per week)		
	(0.211)	(0.396)	Total	26.916	0.787
Average Teacher Absent ≥ 2	0.396	-0.129		(6.724)	(0.943)
Days/Month	(0.489)	(0.168)	Teaching Whole Class	6.207	0.543
Amenities				(4.677)	(2.473)
Building	0.961	0.039*	Teaching Small Group	5.415	0.633
	(0.194)	(0.020)		(4.367)	(1.991)
Number Classrooms	3.229	-0.115	Teaching Individual	5.467	0.769
	(1.413)	(0.925)		(4.832)	(2.305)
Sufficient Desks	0.756	-0.173	Blackboard/Dictation	5.177	0.072
	(0.429)	(0.180)		(4.042)	(2.136)
Drinking Water	0.846	0.154***	Classroom Management	3.225	-0.039
	(0.361)	(0.037)		(2.637)	(0.343)
Electricity	0.724	0.050	Testing	3.217	-1.142***
	(0.447)	(0.152)		(2.769)	(0.428)
Toilet	0.788	-0.162	Administrative	2.497	-0.865***
	(0.409)	(0.178)		(1.895)	(0.278)

Note: This table reports differences in mean characteristics between program and and private schools. The unit of observation is child-school. Columns (1) and (3) report means and standard deviations for program schools; and columns (2) and (4) differences in means between program and 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.3: Program-school Characteristics by Subsidy Treatment

	Uniform Mean/sd (1)	Differentiated – Uniform (2)		Uniform Mean/sd (3)	Differentiated – Uniform (4)
Characteristics from School Survey			Students		
Days Operational	5.088	0.069	Number Boys	90.594	-3.786
	(1.497)	(0.246)		(45.847)	(9.396)
Open Admission	0.881	-0.046	Number Girls	71.019	0.555
	(0.324)	(0.064)		(34.524)	(7.566)
Uniform Required	0.047	-0.047	Percent Female Students	0.445	0.007
	(0.211)	(0.033)		(0.132)	(0.031)
Tuition Required	0.000	0.000	Student-teacher Ratio	44.752	-1.010
	(0.000)	(0.000)		(13.184)	(3.071)
Medium: Sindhi	0.669	-0.116	**Characteristics from Teacher Survey**		
	(0.471)	(0.096)	Days Absent/Month	0.864	-0.053
Medium: English	0.257	0.108		(1.235)	(0.222)
	(0.437)	(0.090)	Female	0.502	-0.020
Number Teachers				(0.428)	(0.087)
Total	3.654	0.261	Age	25.210	-0.081
	(1.489)	(0.324)		(4.229)	(0.838)
Female	2.050	-0.132	Education	11.051	-0.171
	(1.851)	(0.344)		(0.855)	(0.160)
Post-secondary	1.954	-0.115	Salary (1000s Rs)	4.027	0.079
	(2.225)	(0.461)		(0.764)	(0.223)
< 5 years	2.963	0.345	Years Teaching	2.604	0.368
	(1.425)	(0.290)		(1.145)	(0.247)
$5 - 10$ years	0.648	-0.091	Years Teaching this School	1.825	-0.106
	(1.050)	(0.185)		(0.773)	(0.175)
> 10 years	0.043	0.008	Hours Teaching (per week)		
	(0.203)	(0.036)	Total	26.778	0.281
Average Teacher Absent ≥ 2	0.455	-0.121		(6.564)	(1.309)
Days/Month	(0.498)	(0.095)	Teaching Whole Class	5.861	0.699
Amenities				(4.208)	(0.975)
Building	0.993	-0.067*	Teaching Small Group	5.087	0.660
	(0.082)	(0.040)		(3.925)	(0.863)
Number Classrooms	3.167	0.127	Teaching Individual	5.296	0.343
	(1.278)	(0.286)		(4.508)	(0.925)
Sufficient Desks	0.812	-0.113	Blackboard/Dictation	5.219	-0.085
	(0.391)	(0.086)		(3.746)	(0.842)
Drinking Water	0.825	0.043	Classroom Management	3.090	0.271
	(0.380)	(0.073)		(2.546)	(0.498)
Electricity	0.741	-0.034	Testing	2.811	0.819
	(0.438)	(0.091)		(2.548)	(0.564)
Toilet	0.766	0.045	Administrative	2.307	0.382
	(0.424)	(0.081)		(1.668)	(0.387)

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 (3) report means and standard deviations for program schools under the gender-uniform subsidy treatment; and columns (2) and (4) differences in means between program schools under the two subsidy treatments. Differences are estimated using a seemingly unrelated regressions (SUR) specification to account for within-school correlations in school characteristics. 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.4: School Proximity and Enrollment by School Type

	Pct Villages with School < 1.5km		Pct Villages with Any Child Enrolled		Number Enrolled Per Village	
	Control	Teatment	Control	Treatment	Control	Teatment
School Type:	(1)	(2)	(3)	(4)	(5)	(6)
Program	0.000	0.944	0.026	0.944	0.105	49.506
Government	0.553	0.463	0.632	0.259	19.211	3.451
Private	0.158	0.056	0.184	0.037	6.579	0.549
NGO	0.053	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. Columns (3) and (4) give the percentage of villages with any sample child enrolled in a school of the indicated type. Columns (5) and (6) give the number of sample children enrolled per village in a school of the indicated type.

Table A.5: Balance by School Enrollment Type

	Control		Enrolled	Govt-enr		Control		Enrolled	Govt-enr
	Unenrolled	Govt-enr – Unenrolled	Program – Govt	Treat – Control		Unenrolled	Govt-enr – Unenrolled	Program – Govt	Treat – Control
	(1)	(2)	(3)	(4)		(5)	(6)	(7)	(8)
Child/HH Characteristics					HH Head Occuptions				
Age	7.295	0.093	0.024	0.320***	Landlord	0.051	0.004	-0.012	-0.009
		(0.184)	(0.098)	(0.118)			(0.017)	(0.016)	(0.028)
Female	0.427	-0.019	0.032	0.010	Farmer	0.638	-0.109*	0.132**	0.019
		(0.027)	(0.025)	(0.052)			(0.062)	(0.058)	(0.073)
Child of HH Head	0.849	-0.021	0.038	0.001	Laborer	0.133	0.029	-0.023	-0.024
		(0.037)	(0.027)	(0.042)			(0.037)	(0.034)	(0.050)
Household Size	7.343	0.309	-0.456	-0.634	Teacher/Civil Servant	0.036	-0.013	0.017	0.044*
		(0.460)	(0.361)	(0.506)			(0.018)	(0.024)	(0.023)
Num Children	4.871	0.135	-0.312	-0.553*	Other	0.174	0.074*	-0.097***	0.011
		(0.256)	(0.260)	(0.289)			(0.044)	(0.036)	(0.054)
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)					
Sect: Sunni	0.870	-0.077	0.080	0.084					
		(0.105)	(0.089)	(0.138)					
Language: Sindhi	0.618	0.023	0.043	0.121					
		(0.059)	(0.068)	(0.104)					

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.6: Estimated Social Planner Solution, Counterfactual

	Program Solution	Social Planner Solution	
		Sample:	
		Treatment	All Villages
	(1)	(2)	(3)
Toilets and/or Drinking Water	0.90	0.98	1.00
	(0.30)	(0.14)	(0.00)
Pct Teachers Female	0.49	0.15	0.04
	(0.41)	(0.34)	(0.18)
Pct Teachers Post-secondary	0.48	0.71	0.63
	(0.34)	(0.42)	(0.45)
Pct Teachers $<$5 Yrs Experience	0.85	0.63	0.51
	(0.25)	(0.46)	(0.48)
Avg Teacher Absent $\geq$2 Days/month	0.37	0.05	0.01
	(0.48)	(0.21)	(0.08)
Total Surplus (Upper Bound, 1,000)	232.24	247.05	280.03
Total Surplus (Lower Bound, 1,000)	161.53	170.59	191.55
Sample:			
Number Treatment Villages	150	150	119
Number Control Villages	0	0	31

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

Appendix B

Table B1: Treatment and Test Scores (Pct)

	Control Mean	Treatment Effect, Pct Correct				
		ITT				TOT
	(1)	(2)	(3)	(4)	(5)	(6)
Math Score	0.461	0.163***	0.160***	0.160***	0.194***	0.603***
	(0.307)	(0.047)	(0.048)	(0.048)	(0.038)	(0.087)
Language Score	0.487	0.171***	0.168***	0.167***	0.202***	0.619***
	(0.342)	(0.058)	(0.060)	(0.058)	(0.044)	(0.078)
Total Score	0.471	0.167***	0.163***	0.163***	0.197***	0.607***
	(0.310)	(0.051)	(0.052)	(0.052)	(0.040)	(0.081)
Child Controls		No	Yes	Yes	yes	yes
HH Controls		No	No	Yes	yes	yes
District F.E.s		No	No	No	yes	yes

Note: This table reports program impacts on test scores (for children aged 5–10) measured as the percentage of questions answered correctly. Column (1) gives the mean percent of correct answers (for children aged 5–10) 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. Column (6) reports the treatment-on-the-treated (TOT) impacts on test scores for enrolled children. 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, IRT Model

	Treatment Effect, Skill (θ) Parameter, Z-Score				
	ITT				TOT
	(1)	(2)	(3)	(4)	(5)
EAP (Bayesian)	0.520***	0.512***	0.510***	0.622***	1.928***
	(0.164)	(0.167)	(0.165)	(0.132)	(0.289)
MLE	0.461***	0.454***	0.452***	0.564***	1.745***
	(0.149)	(0.151)	(0.150)	(0.126)	(0.305)
Child Controls	No	Yes	Yes	Yes	Yes
HH Controls	No	No	Yes	Yes	Yes
District F.E.s	No	No	No	Yes	Yes

Note: This table reports program impacts on the latent skill (θ) parameter (for children aged 5–10) using an IRT model using EAP and MLE methods (rows (1) and (2), respectively). Columns (1) through (4) report the intention-to-treat (ITT) impacts, with various sets of controls. The skill (θ) parameter is standardized using the mean and standard deviation from control villages. Column (5) reports the treatment-on-the-treated (TOT) impacts on the skill (θ) parameter for enrolled children. 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 Age

Child Age	Control Village Mean		Treatment Effect				
				Test Score, ITT		Test Score, TOT	
	Enrollment	Test Score	Enrolled	Pct	Z-Score	Pct	Z-Score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
5	0.437	0.425	0.337***	0.179***	0.647***	0.494***	1.788***
		(0.276)	(0.065)	(0.031)	(0.113)	(0.098)	(0.355)
6	0.552	0.438	0.297***	0.176***	0.575***	0.600***	1.963***
		(0.305)	(0.073)	(0.043)	(0.139)	(0.111)	(0.362)
7	0.554	0.461	0.308***	0.212***	0.686***	0.649***	2.099***
		(0.309)	(0.074)	(0.043)	(0.139)	(0.103)	(0.334)
8	0.542	0.489	0.337***	0.208***	0.656***	0.620***	1.959***
		(0.316)	(0.071)	(0.047)	(0.148)	(0.083)	(0.262)
9	0.514	0.475	0.334***	0.231***	0.712***	0.695***	2.142***
		(0.324)	(0.074)	(0.050)	(0.154)	(0.098)	(0.303)
10	0.556	0.550	0.268***	0.190***	0.598***	0.648***	2.041***
		(0.317)	(0.065)	(0.049)	(0.155)	(0.113)	(0.355)
Child Controls			Yes	Yes	Yes	Yes	Yes
HH Controls			Yes	Yes	Yes	Yes	Yes
District F.E.s			Yes	Yes	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. Column (3) gives the effect of treatment on enrolment. Columns (4) and (5) report the intention-to-treat (ITT) impacts on test scores, measured as percent and standard deviations, respectively. Columns (6) and (7) report the treatment-on-the-treated (TOT) impacts on test scores for enrolled children, measured as percent and standard deviations, 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.

Appendix C: For Online-Publication

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 the schools in this evaluation (phase-1 schools), 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 policy 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 the program schools in this evaluation. As school operators could not charge any fees, subsidy payments represented the sole source of school revenues.

Subsidy costs for phase-1 program schools evolved from 30 percent (30 percent for all schools) of total costs in fiscal year 2008-09 to 67 percent (72 percent for all schools) in 2009-10 to 48 percent (73 percent for all schools) 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 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 and disrupted. SEF incurred costs helping to rehabilitate schools and restore school operations. These two factorsassignment of total non-subsidy costs to phase-1 schools and costs for helping schools due to the floodswould inflate the cost 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. Under the assumption that school registers may not reliable (for example, for incentives to report higher enrollment), and assuming a 20-percent student-absence typical of rural, remote Sindh, we estimate an enrollment of 23,525.

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.

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 analysesand with superior cost-effectiveness resultswere 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.