Public Disclosure Authorized

Public Disclosure Authorized

Public Disclosure Authorized

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

Felipe Barrera-Osorio
David S. Blakeslee
Matthew Hoover
Leigh L. Linden
Dhushyanth Raju
Stephen P. Ryan

Abstract

This study experimentally evaluates the short-term impacts of public per-student subsidies to partnering local entrepreneurs to establish and operate tuition-free, coeducational, private primary schools in educationally underserved villages in Sindh province, Pakistan. Two subsidy structures were tested, one in which the subsidy amount did not differ by student gender, and the other in which the subsidy amount was higher for female students. The program administrator introduced the latter structure with the aim of correcting for the gender disparity in school enrollment in the general program setting. The program increased school enrollment by 30 percentage points in treated villages, for boys and girls.

It increased test scores by 0.63 standard deviations in treated villages. The gender-differentiated subsidy structure did not have larger impacts on girls' enrollment or test scores than the gender-uniform one. Program schools proved more effective in raising test scores than government schools located near the villages, with program-school students scoring 0.16 standard deviations higher, despite coming from more socioeconomically disadvantaged backgrounds. Estimations of the demand for schooling and education production suggest nearly efficient choices on school inputs by the program administrator and partnering entrepreneurs.

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 Ryan[‖]

[*]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, NBER

1 Introduction

Although many countries have achieved considerable progress in increasing school enrollment in recent decades, Sub-Saharan Africa and South Asia continue to suffer from low student enrollment rates (UNESCO, 2015a). In addition, average student learning remains dismally low in many developing countries (UNESCO, 2014; Pritchett, 2013). In the face of these challenges, governments are increasingly resorting to supporting private schools to more efficiently achieve their education aims (Patrinos, Barrera-Osorio, and Guáqueta, 2009).

This study evaluates, through a randomized-controlled trial, the short-term impacts of a public-private partnership (PPP) program in Sindh province, Pakistan, called the Promoting Low-Cost Private Schooling in Rural Sindh (PPRS) program. The program was funded by the provincial government, and designed and administered by the Sindh Education Foundation (SEF), a semi-autonomous organization. The main, stated program objectives were to increase access to schooling in marginalized areas, to reduce the gender disparity in school enrollment, and to increase student learning, in a cost-effective manner.

The program offered local private entrepreneurs who were qualified to participate a set of benefits to establish and run tuition-free, coeducational primary schools in educationally underserved villages. The benefits included a per-student subsidy, school leadership and teacher training, and teaching and learning materials. The per-student subsidy amount was fixed at less than one-half the per-student cost for public primary and secondary education in the province. The provincial government provided SEF with full discretion over the regulation of program schools, and program-school operators enjoyed some flexibility to decide how to structure and run schools around the guidelines provided by SEF. The program was randomly assigned to 200 out of 263 qualifying villages in selected districts with poor education outcomes in the province.

A central challenge in the pursuit of universal enrollment is the gender disparity. Girls are less likely to enroll in school than boys, especially among poor, rural, or socially disadvantaged households in many developing countries (UNESCO, 2015b). While gender disparities in school enrollment in these countries 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, Mete, and Sathar, 2005; Burde and Linden, 2013; Adukia, 2017; Muralidharan and Prakash, 2017). To address the large gender disparity in primary school enrollment in rural Sindh, 100 of the 200 program villages were randomly assigned to a gender-differentiated subsidy scheme. Under this scheme, program-school operators in these 100 villages received a higher per-student subsidy for girls than for boys, with the aim of more strongly incentivizing the schools to take steps to attract girls.

After the random assignment was completed, SEF scaled down the original evaluation sample from 263 to

199 villages to correct for errors made in determining whether a given village qualified for the program. SEF's 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 as controls. Mean household and child characteristics, at baseline and follow-up measurement in the 199 villages, were similar across the experimental groups. Follow-up measurement was conducted after program schools had completed about 1.5 school years.

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

The relative effectiveness of government versus private schools on student learning has been the subject of significant interest in the education-economics literature. This literature includes growing empirical evidence from developing countries (Angrist et al., 2002; Newhouse and Beegle, 2006; Desai et al., 2009; French and Kingdon, 2010; Bold et al., 2014; Muralidharan and Sundararaman, 2015; Singh, 2015), including from Pakistan based on cross-sectional and panel observational data (Andrabi et al., 2010, 2011; Amjad and MacLeod, 2014). The main general concern, however, is the extent to which the estimates of private-school effectiveness are due to household selection—that is, by the types of children that enroll in private schools—rather than actual private-school productivity.

We contribute to the literature on private-school effectiveness by comparing the test-score performance of program schools and proximate government schools, and assessing the degree to which changes in student composition are behind the result. We do this by comparing government-school students in program and control villages at follow-up, which serves as our measure of potential sorting. We find that program-school students performed better than government-school students, scoring 0.16 standard deviations higher in the total test, despite coming from poorer socioeconomic backgrounds. We also find that the composition of government-school students remained largely unchanged in program villages. Education-production estimates show that student test scores were driven by precisely those school inputs that SEF and individual program-school operators can, and do, adjust – such as the characteristics of teachers. These inputs lie outside the control of individual government schools, given their centralized administration by provincial and district education officials.

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

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

2 Pakistan and the PPRS Program

2.1 Schooling in Pakistan

School enrollment is low in Pakistan, even when compared to countries with a similar income level (Andrabi et al., 2008). At the time the PPRS program was initiated in 2009, the primary-school net enrollment rate (NER) for children aged 6–10 in Pakistan was 67 percent (72 percent for boys and 62 percent for girls) (Government of Pakistan, 2009).[1] 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.[2] The number of private schools increased from around 4,000 in the early 1980s to 36,000 in 2000 to 47,000 in 2005; much of the expansion in the 1990s and 2000s occurred in villages and poorer urban neighborhoods (Andrabi, Das, and Khwaja, 2008). By 2010–11, one-fifth of children aged 6–15 in Pakistan were enrolled in private school (or one-third of students, given the large share of unenrolled children) (Nguyen and Raju, 2015). These schools have succeeded both in terms of cost and quality. At less than $20 in 2000, the mean annual cost of private primary school fees

---

[1]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.

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

represented about two percent of mean total household spending (Andrabi, Das, and Khwaja, 2008). Low fees were enabled by few fixed costs, and low operational costs, specifically low teacher wages. Discussed in section 3, low-cost private schools are also found to produce higher test scores than government schools in rural Punjab province (Andrabi et al., 2010, 2011).

However, large spatial differences exist in private-school enrollment rates across Pakistan. Thirty percent of primary-school students were in private schools in the country, compared to only five percent in rural Sindh in 2008–09 (Government of Sindh, 2009). Andrabi, Das, and Khwaja (2008) find that, as in the rest of the country, private schools in rural Sindh tended to be found in larger villages with better infrastructure. They argue that the main constraint to the further expansion of low-cost private schools is the lack of a local supply of women with secondary education who can be hired as teachers.[3]

2.2 Program

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

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

The first phase of the program, which we evaluate in this study, was implemented in eight (out of, at that time, 23) districts in the province. SEF selected the districts based on how they ranked in terms of the size of the out-of-school child population, the gender disparity in school enrollment, and the percentage of households located at least 15 minutes away from the nearest primary school. The eight poorest-ranked districts were selected, excluding those that were viewed by the provincial government and SEF as experiencing heightened law-and-order concerns.[4]

---

[3]Andrabi, Das, and Khwaja (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.

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

Based on a budgetary assessment, SEF supported coeducational, private primary schools in 200 villages in the selected districts. The main benefits that program-school operators received included a per-student cash subsidy; free school leadership and teacher training; and free textbooks, other teaching and learning materials, stationery, and bookbags.[5]

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

The subsidy amounts were set at less than one-half of the per-student cost in public primary and secondary government schools in the province, and were not adjusted for price inflation over the evaluation period.[6] Provided on a quarterly basis, a school's total subsidy was linked to the number of children in attendance multiplied by 1.25 to reflect an expected 20-percent student-absence rate. SEF gathered attendance information through periodic, unannounced monitoring visits to program schools.

Local private entrepreneurs were invited to apply to the program through an open call in newspapers, and to propose educationally underserved villages in the selected districts to establish and operate schools. SEF vetted the applications (ultimately, through visits to shortlisted villages) based on several criteria, including written assent from the parents of at least 75 children of primary-school ages that they would enroll their children in the school, should it be established; a school site in the village that was located at least 1.5 kilometers from the nearest other school; a building of sufficient size; and the identification of teachers with a minimum of eight years of schooling (middle school completion), with at least two being female.[7] Once in the program, school operators would continue to receive the subsidy and other benefits, as long as they provided free schooling to children, and provided and maintained school infrastructure and the schooling environment consistent with SEF guidelines. SEF strictly enforced the free-schooling condition, but was more lenient in enforcing the school infrastructural features and environmental conditions.

3   Potential mechanisms

In the basic human-capital acquisition model (Becker, 1962), households compare the present and future costs and benefits of educating their children. School fees and transportation costs, both in terms of direct

---

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

[6]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.

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

payments and time, tend to be the largest costs of primary schooling for households. The PPRS program reduces the cost of transportation by situating schools in underserved villages, and reduces school fees by providing free schooling. The immediate effect of these two changes should be an increase in household demand for schooling, and, thus, higher school enrollment.

International evidence on the effects of distance from home to school on enrollment is strong, from experimental evaluations (Burde and Linden, 2013) as well as from quasi-experimental evaluations (Foster and Rosenzweig, 1996; Duflo, 2001; Berlinski, Galiani, and Gertler, 2009). These studies find that when schools are introduced into underserved areas, household demand for schooling increases, often substantially. Consistent with this evidence, Carneiro, Das, and Reis (2016) estimate household demand for alternative primary schooling choices (private and government schools) in rural Punjab, Pakistan, and find that the strongest determinant is distance from home to school. International evidence for the effects of free schooling on household demand for schooling is also strong, suggesting a negative association between price and enrollment (Deininger, 2003; Barrera-Osorio, Linden, and Urquiola, 2007; Holla and Kremer, 2009; Borkum, 2012; Lucas and Mbiti, 2012).

As another supply-side element, the enrollment-related subsidy structure incentivizes program-school operators to take measures to attract children to enroll in school. This feature is strengthened for girls in the gender-differentiated subsidy structure, which potentially incentivizes private-school operators to take stronger measures to draw girls, for example, through employing female teachers, providing safe transportation and a safe schooling environment, or even offering a small, conditional cash transfer to girls.

The positive effects on enrollment from interventions that reduce the cost of enrolling in school contrast with the mixed effects of the interventions on student learning, measured through test scores. For example, early, international evidence from conditional cash-transfer programs suggests that exposure to (more) schooling does not necessarily produce higher test scores (Fiszbein and Schady, 2009; Saavedra and García, 2012). More recently, international evidence points to positive effects on test scores, although the results depend crucially on the specific design components of each individual program (Barham, Macours, and Maluccio, 2013; Barrera-Osorio and Filmer, 2015). To address the challenge of raising test scores, the program offered training on school leadership and teaching, and supplied teaching and learning materials, both free of cost. The program also supported local private entrepreneurs to establish and operate schools (rather than SEF doing so itself). Program-school operators had greater flexibility than government schools to determine their input mix and to account for specific local conditions of the communities and demands from households (Hoxby, 2003; MacLeod and Urquiola, 2012), and were potentially more accountable to the households that they served (World Bank, 2004).

An influential literature argues for the potential gains from private schooling (Friedman, 1955; Hoxby,

2003; MacLeod and Urquiola, 2012). Evidence exists for Pakistan on the relative effectiveness of private versus government schools. Using an instrumental-variables approach, Andrabi et al. (2010) find that low-cost private schools in rural Punjab have large, positive effects on test scores and civic knowledge among primary-school students. They attribute this to the fact that private schools adjust their inputs to meet local conditions, whereas government schools do not. In the same setting, Andrabi et al. (2011) also examine the effects of enrollment in low-cost private schools on test scores, based on data for students who switch from government to private schools, and vice versa. They find that students who move from government to private schools see a large gain in test scores shortly after moving, while those who move from private to government schools see a large decline. The estimated effects from these studies may, however, be biased due to demand-side selection into private schools by households and supply-side selection of students by schools. These forms of selection make it difficult to disentangle the effect of school effectiveness from that of the composition of students in terms of ability.

An important source of rigorous evidence on private-school effectiveness comes from studies of private-school voucher programs in developing countries. Based on vouchers offered through lotteries to low-income, government-school students for private secondary schooling in Colombia, Angrist et al. (2002) find positive effects of private-school enrollment on several outcomes, including grade progression and test scores. However, these effects potentially reflect both private-school effectiveness and incentivized student effort.[8] Randomly assigning vouchers at the village and household levels to government-school students for private primary schooling in India, Muralidharan and Sundararaman (2015) find that private schools are highly cost-effective in raising test scores. They also find that private schools are more responsive in terms of matching what's taught to household demand, reallocating time away from subjects such as reading and mathematics, towards English, Hindi, science, and social studies.[9] In addition, the study did not find spillover effects on government-school students who were not assigned vouchers, nor students initially enrolled in private schools, in program villages.

This study is also linked to the literature on PPP programs in education (Patrinos, Barrera-Osorio, and Guáqueta, 2009). Although these programs share many of the theoretical justifications discussed in the private-school voucher literature (for example, private-school effectiveness, market competition), they differ in important ways. Most conspicuously, PPPs enable higher-quality education for beneficiary students through the terms of the contract between the program school and the government, with some programs regulating school inputs (including teachers), and others selecting program schools based on measures of

---

[8]In addition to enabling private-school enrollment, the program stipulated that students would lose their vouchers if their grades fell below a certain level.

[9]Test scores for Telugu and mathematics remained the same, despite significantly less time being spent on these topics; while time allocated towards other topics yielded test-score improvements.

quality.

However, PPP programs can be structured such that they lose the market-based accountability mechanisms at play in the private-school market. Schools under PPP programs may be less effective than non-program private schools in such cases. For example, some PPP programs require that schools offer free education, as is the case with the PPRS program. This condition cancels out the role that prices play in informing market participants of the relative efficiencies of producers and the preferences of consumers (Hayek, 1945).[10] The effect of PPP programs on learning then is theoretically ambiguous. Notwithstanding, the available evidence from developing countries, including from Pakistan, generally indicates that PPP programs have positive effects on school enrollment and test scores (Kim, Alderman, and Orazem, 1999; Alderman, Orazem, and Paterno, 2001; Alderman, Kim, and Orazem, 2003; Barrera-Osorio and Raju, 2015; Barrera-Osorio et al., 2016).

4   Data

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

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

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

---

[10]Prices may, however, obscure broader problems of limited information in the education market (MacLeod and Urquiola, 2012).

and building. GPS data were collected from all schools, the proposed program-school site, and surveyed households.

The follow-up survey consisted of three instruments: a school survey; a household survey; and a child survey, which included a test. The household and child surveys, and child tests, were administered at the child's home. The household survey was administered to households with at least one child aged 5–9. In large villages, up to 42 randomly sampled households in the village were interviewed; in villages with fewer than 42 households, which comprised the majority, all households in the village were interviewed. The survey was multi-topic, and had extensive modules on past and current schooling and other activities for children aged 5–17, answered by the household head or another primary adult household member.

A child survey was administered to each child aged 5–9. It asked questions mainly on work activities performed inside and outside the home, past and current schooling, and aspirations. Each child was then administered tests on language (either Urdu or Sindhi, as preferred) and mathematics.

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

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

5 Empirical strategy

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

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

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

Internal validity: The validity of our results depends upon the comparability of populations across the

experimental groups. Because the program was randomly assigned across villages, treatment status should be orthogonal to household and child characteristics that might be correlated with the outcomes. Insofar as this holds, it will be sufficient to compare outcomes across the treatment and control groups to evaluate the impacts of the program.

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

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

Analogously structured to table 2, appendix table A1 reports the differences in mean household and child characteristics between villages under the gender-uniform and -differentiated subsidy treatments. Differences in mean characteristics between the two subsidy treatments were small, and always statistically insignificant.

6 Program impacts

6.1 Enrollment

School enrollment information was collected in two ways. First, the household survey respondent was asked whether the child was enrolled during the just-concluded school term. Second, the enrollment of the child aged 5–10 in a given school was verified using a student attendance list compiled through a headcount conducted during the school survey.[11] We discuss results based on both reported and verified enrollment measures.

Table 3 reports the pooled-treatment impacts on school enrollment and grade attainment. Panels A and B report results for young children (aged 5–10) and older children (aged 11–17), respectively. Columns (1) through (4) report impacts on reported enrollment, with different sets of controls. Columns (5) and (6) report impacts on verified enrollment and grade attained, respectively, with the full set of controls. Based on the model with the full set of controls, the program increased reported enrollment among young children by 31.5 percentage points, and verified enrollment by 29.4 percentage points. In addition, the program increased attainment by 0.4 grades.

---

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

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

6.2 Test Scores

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

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

6.3 Differential impacts on school enrollment and test scores

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

6.4 Aspirations

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

for girls, the program, and the interaction of the two, respectively.

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

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

7 Program cost-effectiveness

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

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

The scale-up of the program during the evaluation period also affected how cleanly we could assign costs to phase-1 program schools. The cost data allow us to distinguish between subsidy costs for phase-1 and phase-2 program schools, but we could not separate out other types of costs in the same way. The cost data

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

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

Program impacts on school enrollment and total test scores were 30 percent and 0.6 standard deviations, respectively. Using the low and high values of annual cost per student, we estimate cost-effectiveness values of 16 percent to 39 percent in school enrollment and 0.3 to 0.8 standard deviations in total test scores, both per $100 spent. Program cost-effectiveness values associated with test scores appear to be at the lower end of the range of similarly estimated cost-effectiveness values for 14 education interventions reported by Evans and Popova (2016), only superior to a conditional cash transfer program in Africa.

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

While the program had large impacts on school enrollment and test scores, these impacts were accompanied by relatively large expenditures. Both the large impacts and expenditures are arguably due to the type of intervention: introducing new schools. Most of the other interventions with comparable cost-effectiveness

analyses—and with superior cost-effectiveness results—were those introduced into (communities with) pre-existing schools (Evans and Popova, 2016). SEF has continued to scale up the program, adding more schools and upgrading some primary schools to middle schools (up to grade 8), which has contributed to falling annual costs per student, as operating costs associated with such things as program administration and teacher-training workshops are spread over increasing numbers of schools and students. However, we do not know how program impacts have evolved in tandem with the scale-up.

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

8 Program schools

8.1 Characteristics

We examine differences in mean characteristics between program and government schools, and between program and private schools (table 8), as well as between program schools under the gender-uniform and -differentiated subsidy treatments (appendix table A2), at follow-up. The means are estimated based on student-school observations.

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

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

Based on information from individual teachers, we find that program-school teachers were more likely

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

In appendix table A2, columns (1) and (3) report the mean characteristics in program schools under the gender-uniform subsidy treatment, and columns (2) and (4) report the differences in mean characteristics between program schools under the gender-uniform and -differentiated subsidy treatments. We do not find that program-school operators under the gender-differentiated subsidy treatment had structured their schools differently from their counterparts under the gender-uniform one. The absence of differential impacts on girls' school enrollment and test scores across the two subsidy treatments discussed earlier is in line with the findings here.

8.2 Quality

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

In table 9, column (1) reports mean test scores for program schools, columns (2) and (3) report differences in mean test scores between program schools and the indicated school type, and column (4) reports p-values from tests of differences in mean test scores between government and private schools. Program schools scored 0.17 standard deviations higher on the total test than government schools (0.20 standard deviations higher on the mathematics test, and 0.12 standard deviations higher on the language test). In contrast, differences in mean test scores between program and private schools were small (about 0.01–0.02 standard deviations) and statistically insignificant.

These comparisons do not causally identify differences in quality between school types, as student-composition effects are likely to bias these estimates. If program schools attract students who would not otherwise have been enrolled, and because these students come from more socioeconomically disadvantaged backgrounds, the program-school effect on test scores may be biased downwards. If, however, government-school students are enrolling in program schools, and these students are more advantaged than other children

in the village, the program-school effect may be biased upwards. Although it cannot be ruled out, we do not find clear evidence supporting the latter possibility. Appendix table A3 reports differences in mean household and child characteristics by school type. As column (2) shows, program-school students were younger, enrolled in a lower grade, and had first enrolled at a later age. They also had fathers who were less educated and were more likely to be farmers, and resided in poorer-built dwellings.

In addition, we find little evidence that program schools induced student sorting in proximate government schools. As column (5) shows, mean characteristics of government-school students were largely similar across control and program villages, with the exception that government-school students in program villages were slightly older than their counterparts in control villages (and therefore in a slightly higher grade). 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.

To understand the determinants of the program-school effect on test scores, we estimate an education-production function, by regressing test scores (for students in all schools) on a set of school inputs, controlling for child characteristics. Table 10 reports estimates for total test scores; appendix table A4 reports estimates for subject test scores. The results support the findings discussed above. Specifically, school inputs significantly associated with test scores–such as the teacher's education and teaching experience–are those which program schools exercise greater, independent control over than government schools. This finding holds even when we control for program schools in the estimations, indicating that the estimates of the education-production function are not capturing some unobserved feature of program schools with which they are correlated.

8.3   Efficiency

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

The first-order condition for the program-school operator is:

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

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

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

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

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

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

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

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

For the demand function estimation, we include a variety of school inputs and child characteristics shown to be important in the education-production literature (Todd and Wolpin, 2003). Child characteristics

consist of gender, age, distance from home to the school, and interactions between the child's gender and school inputs. School inputs consist of toilets and/or drinking water (a single indicator variable); as well as teacher characteristics, such as gender, teaching experience, frequency of absence from school, and time spent teaching. We also include interactions of school inputs with an indicator variable for female students, as a substantial body of research shows the importance of school-supply factors for girls' enrollment and learning – such as distance and time from home to school, school infrastructural features and environmental conditions, and teacher characteristics and behaviors (Lloyd, Mete, and Sathar, 2005; Burde and Linden, 2013; Muralidharan and Sheth, 2016; Adukia, 2017; Muralidharan and Prakash, 2017).

Table 11 reports the schooling-demand estimates. Column (1) reports results for a parsimonious specification; column (4) reports those from our richest specification, which includes interactions and an indicator variable for government schools. The coefficients have the expected signs. Looking at column (4), boys and older children were more likely to enroll in school. Children were more likely to enroll in school if it had toilets and/or drinking water, had lower fees, was closer to their homes, had teachers who spent more time teaching, and was not a government school. The coefficient on girls' interaction with the percentage of teachers who were female is large and statistically significant, indicating that the enrollment decision for girls was more sensitive to this dimension than that for boys.

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

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

Table 12 presents the results. The first input is toilets and/or drinking water. Demand is apparent in the costs incurred in providing these inputs. The next four inputs change the composition of teachers at the school. The first estimate reflects the cost of replacing a male teacher with a female teacher. Male students reacted negatively to the presence of a female teacher, while the opposite was true for female students. In combination with the number of boys and girls in each village, the sum of these forces implies that enrollment was lower when program schools substituted a female teacher for a male teacher. This, in turn, implies that female teachers must have been less costly than their male counterparts. Evidence for Pakistan is consistent with this finding. Andrabi, Das, and Khwaja (2008) find that female teachers in private schools earn 33-percent less than their male counterparts, after controlling for other characteristics. As might be expected, we also find that adding a teacher with post-secondary education was costly compared to adding one who was less educated, as was adding a teacher who was frequently absent; and adding a teacher with less than five years of teaching experience was less costly compared to adding one with more experience.

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

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

where $CV(x) = \sum_{i=1}^{N} CV_i(x)$ is the sum of the consumer surplus over all children in the village, $TC(x)$ is the (total) cost incurred in providing inputs and subsidies, and $\tau g(x)$ is the social value of higher test scores.[12] We assume that the social value of education is related to overall test scores, given by $g(x)$, and a scalar multiplier, τ.

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

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

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

The last component of our welfare analysis is the social benefits of education that are not internalized in the demand function. Since we do not know how exactly the social benefits of education, we choose to parameterize the social benefit function as $h(x) = \tau g(x)$, where $g(x)$ is the estimated education production function. This specification assumes that the social benefits of education are only a function of test scores,

---

[12]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. Though 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.

and captures the marginal (social) utility of higher test scores. This approach allows us to: first, solve the social planner's solution, as total benefits of providing inputs can be consistently compared with their costs; second, show how the social planner's solution to providing inputs changes with τ, a parameter that we do not have measurements of; and, third, compute the relative 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.[13] To estimate the latter, we assume a deadweight loss of 30 percent, and multiply this by the total annual subsidy, which is a fixed per-student amount of 4,200 rupees per year under the gender-uniform subsidy treatment, which rises to 5,400 rupees per year for girls under the gender-differentiated one.

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

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

where the subscript g indicates the gender of the child, and b the upper and lower bound estimates of wage gains. In rural Sindh, the baseline monthly wage (*blwage*) for men aged 15–34 is 6,600 rupees, and that for women in the same age group is 2,000 rupees; and labor force participation rates for the two are 80 percent and 36 percent, respectively.[14] 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:

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

---

[13]See footnote 12.

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

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

Table 13 reports the levels and changes in school inputs across the solutions offered by program-school operators and the social planner. Program-school operators have proven remarkably successful at establishing and operating schools that generated most of the possible surplus in the environment. Assuming a social value of education equal to 1, the social planner's solution generates gains of slightly less than 10 percent relative to that of private-school operators. Large variation exists across villages, from a lower bound of a zero-percent increase (i.e., the program-school operator selected the same set of school inputs as the social planner would have) to an upper bound of a 33-percent increase. The social planner achieves these gains through various changes to program schools. First, under the social planner, all program schools have toilets and/or drinking water (+10 percentage points relative to the program-school operator solution). The social planner employs teachers with post-secondary education only (+52 percentage points), and with less than five years of teaching experience (+16 percentage points); and requires all teachers to be absent fewer than two days per month (-37 percentage points). The composition of female teachers is relatively unchanged, the social planner employing 47 percent female teachers (-3 percentage points), with substantial variation across villages. Differences are driven by the gender composition of the child population. Specifically, in villages with a relatively large number of boys, enrollment, and consequently, test scores, will suffer if the school employs female teachers, while the opposite is true in villages with a relatively large number of girls.

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

One of the key parameters in our social planner solution is the social value of education. This parameter

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

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

9 Conclusion

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

The study also assesses the effectiveness and efficiency of program schools. Consistent with existing

evidence for Pakistan from larger, more-developed villages, we find that program-school students had higher test scores than government-school students, despite coming from more socioeconomically disadvantaged households. While program-school operators only captured profits through enrollment, the equilibrium social surplus is within 10 percentage points of the social planner. Compared to program-school operators, the social planner adjusts the gender ratio of teachers to better match the gender ratio of children in the village, ensures toilets and/or drinking water in the school, and hires less-experienced, but more-educated, teachers. It is remarkable and reassuring that program-school operators have proven so successful in selecting the most-essential inputs for their schools. The results suggest that, when the government provides adequate support, enormous potential exists for local actors to find appropriate solutions to their challenges.

References

[1] Adukia, Anjali. 2017. "Sanitation and Education." *American Economic Journal: Applied Economics*, 9(2): 23-59.

[2] Alderman, Harold, Jooseop Kim, and Peter F. Orazem. 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.

[3] Alderman, Harold, Peter F. Orazem, and Elizabeth M. Paterno. 2001. "School Quality, School Cost, and the Public/Private School Choices of Low-Income Households in Pakistan." *Journal of Human Resources*, 36(2): 304–326.

[4] Amjad, Ravish, and Gordon MacLeod. 2014. "Academic Effectiveness of Private, Public, and Private-Public Partnership Schools in Pakistan." *International Journal of Educational Development*, 37: 22–31.

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

[6] ———. 2013. "Students Today, Teachers Tomorrow: Identifying Constraints on the Provision of Education." *Journal of Public Economics*, 100: 1–14.

[7] Andrabi, Tahir, Jishnu Das, Asim Ijaz Khwaja, and Tristan Zajonc. 2011. "Do Value-Added Estimates Add Value? Accounting for Learning Dynamics." *American Economic Journal: Applied Economics*, 3: 29–54.

[8] Andrabi, Tahir, Jishnu Das, Asim Ijaz Khwaja, Tara Vishwanath, Tristan Zajonc, and the LEAPS Team. 2008. *Pakistan: Learning and Educational Achievements in Punjab Schools (LEAPS): Insights to Inform the Education Policy Debate*. Washington, DC: World Bank.

[9] Andrabi, Tahir, Natalie Bau, Jishnu Das, and Asim Ijaz Khwaja. 2010. "Are Bad Public Schools Public "Bads"? Test Scores and Civic Values in Public and Private Schools." Manuscript.

[10] Angrist, Joshua, Eric Bettinger, Eric Bloom, Elizabeth King, and Michael Kremer. 2002. "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment." *American Economic Review*, 92(5): 1535–1558.

[11] Barham, Tania, Karen Macours, and John A. Maluccio. 2013. "More Schooling and More Learning? Effects of a Three-Year Conditional Cash Transfer Program in Nicaragua After 10 Years." Working Paper IDB-WP-432. Washington, DC: Inter-American Development Bank.

[12] Barrera-Osorio, Felipe, and Deon Filmer. 2015. "Incentivizing Schooling for Learning: Evidence on the Impact of Alternative Targeting Approaches." *Journal of Human Resources*, 51(2): 461–499.

[13] Barrera-Osorio, Felipe, and Dhushyanth Raju. 2015. "Evaluating the Impact of Public Student Subsidies on Low-Cost Private Schools in Pakistan." *Journal of Development Studies*, 51(7): 808–825.

[14] Barrera-Osorio, Felipe, Leigh L. Linden, and Miguel Urquiola. 2007. "The Effects of User Fee Reductions on Enrollment: Evidence from a Quasi-Experiment." Brief 81437. Washington, DC: World Bank.

[15] Barrera-Osorio, Felipe, Pierre de Galbert, James P. Habyarimana, and Shwetlena Sabarwal. 2016. "The Impact of Public-Private Partnerships on Private School Performance: Evidence from a Randomized Controlled Trial in Uganda." Policy Research Working Paper 7905. Washington, DC: World Bank.

[16] Bau, Natalie, and Jishnu Das. 2017. "The Misallocation of Pay and Productivity in the Public Sector: Evidence from the Labor Market for Teachers." Policy Research Working Paper 8050. Washington, DC: World Bank.

[17] Becker, Gary S. 1962. "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy*, 70(5): 9–49.

[18] Berlinski, Samuel, Sebastian Galiani, and Paul Gertler. 2009. "The Effect of Pre-Primary Education on Primary School Performance." *Journal of Public Economics*, 93(1–2): 219–234.

[19] Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, and Justin Sandefur. 2011. "The High Return to Private Schooling in a Low-Income Country." Working Paper 279. Washington, DC: Center for Global Development.

[20] Borkum, Evan. 2012. "Can Eliminating School Fees in Poor Districts Boost Enrollment? Evidence from South Africa." *Economic Development and Cultural Change*, 60(2): 359–398.

[21] Burde, Dana, and Leigh L. Linden. 2013. "Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools." *American Economic Journal: Applied Economics*, 5(3): 27–40.

[22] Carniero, Pedro, Jishnu Das, and Hugo Reis. 2016. "The Value of Private Schools: Evidence from Pakistan." Discussion Paper 9960. Bonn: Institute for the Study of Labor (IZA).

[23] Deininger, Klaus. 2003. "Does Cost of Schooling Affect Enrollment by the Poor? Universal Primary Education in Uganda." *Economics of Education Review*, 22(3): 291–305.

[24] Desai, Sonalde, Amaresh Dubey, Reeve Vanneman, and Rukmini Banerji. 2009. "Private Schooling in India: A New Educational Landscape," in *India Policy Forum*, Suman Bery, Barry Bosworth, and Arvind Panagariya, eds. New Delhi: Sage Publications.

[25] Dhaliwal, Iqbal, Esther Duflo, Rachel Glennerster, and Caitlin Tulloch. 2013. "Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications for Education," in *Education Policy in Developing Countries*, Paul Glewwe, ed. Chicago: University of Chicago Press.

[26] Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review*, 91(4): 795–813.

[27] Evans, David K., and Anna Popova. 2016. "Cost-Effectiveness Analysis in Development: Accounting for Local Costs and Noisy Impacts." *World Development*, 77: 262–276.

[28] Fiszbein, Ariel, and Norbert Schady. 2009. *Conditional Cash Transfers: Reducing Present and Future Poverty. Policy Research Report*. Washington, DC: World Bank.

[29] Foster, Andrew D., and Mark R. Rosenzweig. 1996. "Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution." *American Economic Review*, 86(4): 931–953.

[30] French, Rob J., and Geeta Kingdon. 2010. "The Relative Effectiveness of Private and Government Schools in Rural India: Evidence from ASER Data." Working Paper 10-03. London: Department of Quantitative Social Science - UCL Institute of Education, University College London.

[31] Friedman, Milton. 1955. "The Role of Government in Education," in *Capitalism and Freedom*. New Brunswick, New Jersey: Rutgers University Press.

[32] Glewwe, Paul, and Hanan Jacoby. 1994. "Student Achievement and Schooling Choice in Low-Income Countries: Evidence from Ghana." *Journal of Human Resources*, 29(3): 843–864.

[33] Government of Pakistan. 2007. *Pakistan Social and Living Standards Measurement Survey 2006-07*. Pakistan Bureau of Statistics, Government of Pakistan.

[34] Government of Pakistan. 2009. *Pakistan Social and Living Standards Measurement Survey 2008-09*. Pakistan Bureau of Statistics, Government of Pakistan.

[35] Government of Pakistan. 2011. *Pakistan Social and Living Standards Measurement Survey 2010-11*. Pakistan Bureau of Statistics, Government of Pakistan.

[36] Handa, Sudhanshu. 2001. "Raising Primary School Enrollment in Developing Countries: The Relative Importance of Supply and Demand." *Journal of Development Economics* 69: 103–128.

[37] Hayek, Friedrich. A. 1945. "The Use of Knowledge in Society." *American Economic Review*, 35(4): 519–530.

[38] Holla, Alaka, and Michael Kremer. 2009. "Pricing and Access: Lessons from Randomized Evaluation in Education and Health." William Easterly and Jessica Cohen, eds., *What Works in Development? Thinking Big and Thinking Small.* Washington, DC: Brookings Institution Press.

[39] Hoxby, Caroline M. 2003. *Economics of School Choice.* Chicago: University of Chicago Press.

[40] Kazianga, Harounan, Dan Levy, Leigh L. Linden, and Matt Sloan. 2013. "The Effects of 'Girl-Friendly' Schools: Evidence from the BRIGHT School Construction Program in Burkina Faso." *American Economic Journal: Applied Economics*, 5(3): 41–62.

[41] Kim, Jooseop, Harold Alderman, and Peter F. Orazem. 1999. "Can Private School Subsidies Increase Schooling for the Poor? The Quetta Urban Fellowship Program." *World Bank Economic Review*, 13(3): 443.

[42] Lloyd, Cynthia B., Cem Mete, and Zeba A. Sathar. 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.

[43] Lucas, Adrienne M. and Isaac M. Mbiti. 2012. "Access, Sorting, and Achievement: The Short-Run Effects of Free Primary Education in Kenya." *American Economic Journal: Applied Economics*, 4(4): 226–253.

[44] MacLeod, W. Bentley, and Miguel Urquiola. 2012. "Competition and Educational Productivity: Incentives Writ Large," in *Education Policy in Developing Countries*, ed. by P. Glewwe. Chicago: University of Chicago Press.

[45] Montenegro, Claudio E., and Harry Anthony Patrinos. 2014. "Comparable Estimates of Returns to Schooling Around the World." Policy Research Working Paper 7020. Washington, DC: World Bank.

[46] Muralidharan, Karthik, and Ketki Sheth. 2016. "Bridging Education Gender Gaps in Developing Countries: The Role of Female Teachers." *Journal of Human Resources*, 51(2): 269–297.

[47] Muralidharan, Karthik, and Nishith Prakash. 2017. "Cycling to School: Increasing Secondary School Enrollment for Girls in India." *American Economic Journal: Applied Economics*, 9(3): 321–350.

[48] Muralidharan, Karthik, and Venkatesh Sundararaman. 2015. "The Aggregate Effect of School Choice: Evidence from a Two-Stage Experiment in India." *Quarterly Journal of Economics*, 130(3): 1011–1066.

[49] Newhouse, David, and Kathleen Beegle. 2006. "The Effect of School Type on Academic Achievement: Evidence from Indonesia." *Journal of Human Resources*, 41(3): 529–557.

[50] Nguyen, Quynh T., and Dhushyanth Raju. 2015. "Private School Participation in Pakistan." *Lahore Journal of Economics*, 20(1): 1–46.

[51] Patrinos, Harry, Felipe Barrera-Osorio, and Juliana Guáqueta. 2009. *The Role and Impact of Public-Private Partnerships in Education*. Washington, DC: World Bank.

[52] Pritchett, Lant. 2013. *The Rebirth of Education: Schooling Ain't Learning*. Washington, DC: Center for Global Development.

[53] Saavedra, J. Eesteban, and Sandra García. 2012. "Impacts of Conditional Cash Transfer Programs on Educational Outcomes in Developing Countries: A Meta-Analysis. RAND Labor and Population Working Paper Series, WR-921-1.

[54] Schultz, Paul T. 2004. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program." *Journal of Development Economics*, 74(1): 199–250.

[55] Singh, Abhijeet. 2015. "The Private School Premium: Size and Sources of the Private School Advantage in Test Scores in India." *Journal of Development Economics*, 113: 16–32.

[56] Small, Kenneth, and Harvey S. Rosen. 1981. "Applied Welfare Economics with Discrete Choice Models." *Econometrica*, 49(1): 105–130.

[57] Todd, Petra E., and Kenneth I. Wolpin. 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement." *Economic Journal*, 113(485): F3–F33.

[58] UNESCO (United Nations Educational, Scientific, and Cultural Organization). 2015b. *EFA Global Monitoring Report 2015: Gender and EFA 2010-2015: Achievements and Challenges: Gender Summary*. Paris: UNESCO.

[59] ——— 2015a. *EFA Global Monitoring Report 2015: Education for All: Achievements and Challenges*. Paris: UNESCO.

[60] ———. 2014. *EFA Global Monitoring Report 2013/14: Teaching and Learning: Achieving Quality for All*. Paris: UNESCO.

[61] World Bank. 2004. *World Development Report 2004: Making Services Work for Poor People.* Washington, DC: World Bank.

Figure 1: Distribution of program costs over the evaluation period

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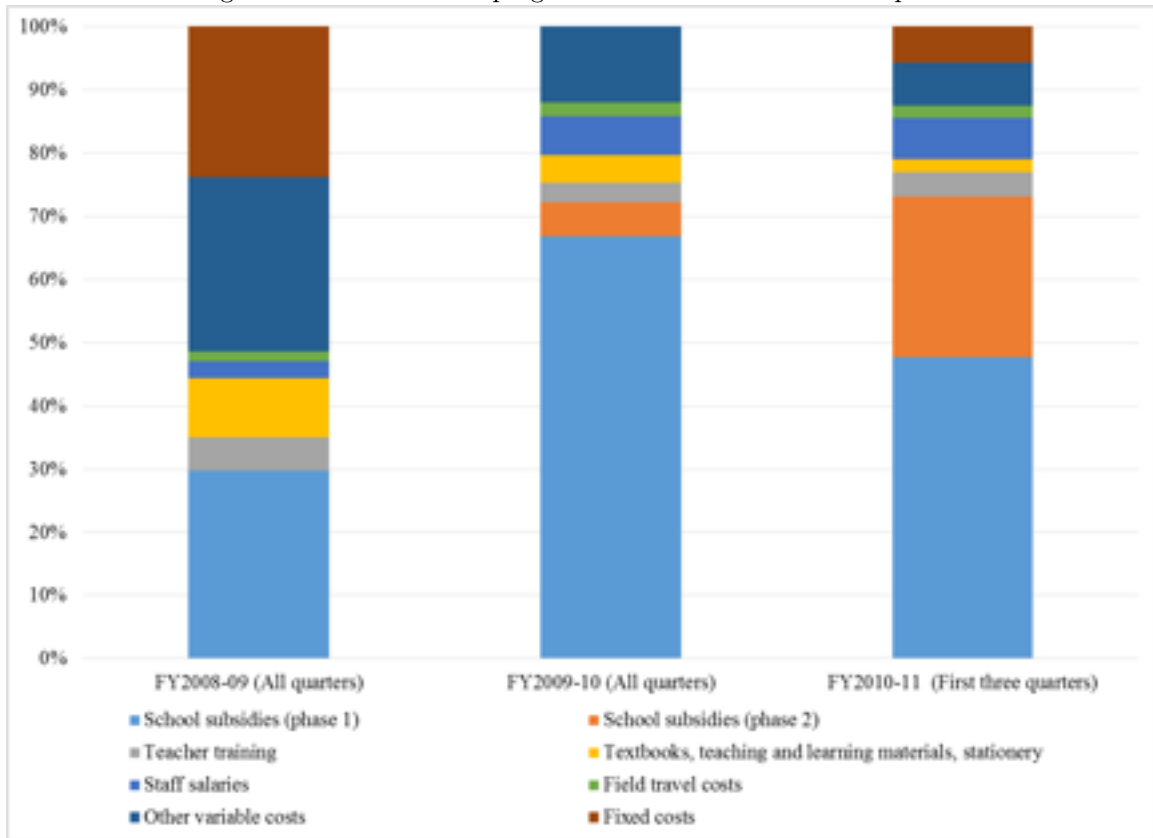
Note: FY denotes fiscal year, which runs from July 1 to June 30. School subsidies (phase-1) reflects per-student public subsidies offered to program schools under this evaluation.

Table 1: Evaluation sample sizes

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

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

Table 2: Balance across program and control villages

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

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

Table 3: Program impacts on enrollment

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

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

Table 4: Program impacts on test scores

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

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

Table 5: Differential impacts by subsidy treatment and by gender

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

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

Table 6: Gender differential impacts by subsidy treatment

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

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

Table 7: Program impacts on aspirations

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

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

Table 8: Characteristics by school type

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

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

Table 9: Test Scores by school type

	program (1)	program - govt (2)	program - private (3)	p-value govt=priv (4)	govt program - control (5)
math	0.717	0.200***	0.012	0.442	0.025
		(0.065)	(0.240)		(0.102)
language	0.708	0.118***	0.019	0.485	0.009
		(0.045)	(0.138)		(0.078)
total	0.735	0.170***	0.016	0.456	0.018
		(0.057)	(0.202)		(0.092)

Note: This table reports differences in mean standardized test scores across school types, controlling for student characteristics and district fixed effects. Column (1) reports means for program schools; column (2), differences in means between program and government schools; column (3), differences in means between program and private schools; column (4), differences in means between government and private schools; and column (5), differences in means between government schools in program and control 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 10: Education production estimates, total test scores

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

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

Table 11: Schooling Demand Estimates

	(1)	(2)	(3)	(4)
constant	0.306***	-0.141**	-0.013	1.365***
	(0.091)	(0.080)	(0.067)	(0.117)
toilets and/or drinking water	0.841***	0.904***	0.882***	0.567***
	(0.062)	(0.065)	(0.087)	(0.078)
student female	0.012	0.030	-0.213**	-0.232*
	(0.048)	(0.055)	(0.111)	(0.147)
student age	0.037***	0.041***	0.040***	0.035**
	(0.012)	(0.012)	(0.012)	(0.016)
distance from home to school	-0.131***	-0.136***	-0.102***	-0.051
	(0.033)	(0.030)	(0.041)	(0.055)
pct teachers with <5yrs exp		0.792***	0.795***	-0.133**
		(0.067)	(0.073)	(0.076)
pct teachers post-secondary		-0.252***	-0.252***	-0.044
		(0.054)	(0.060)	(0.061)
pct teachers female		-0.461***	-0.703***	-0.859***
		(0.055)	(0.054)	(0.091)
pct time teaching		0.263*	0.255***	-0.047
		(0.178)	(0.061)	(0.088)
avg teacher absent ≥2 days/month		-0.113***	-0.113***	-0.148***
		(0.041)	(0.048)	(0.054)
pct female teachers X female student			0.528***	0.562***
			(0.070)	(0.119)
distance X female student			-0.061	-0.020
			(0.057)	(0.080)
toilets and/or drinking water X female student			0.022	0.002
			(0.133)	(0.122)
tuition cost per year	-0.008***	-0.006***	-0.006***	-0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
govt school				-1.473***
				(0.075)

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

Table 12: Cost estimates

toilets and/or drinking water	3.604***
	(0.307)
teacher female	-4.249***
	(0.942)
post-secondary	-0.404***
	(0.184)
<5 yrs experience	-1.379***
	(0.315)
avg teacher absent ≥2 days/month	-1.332***
	(0.125)

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

Table 13: Estimated social planner solution

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

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

38

Appendix

Table A.1: Balance Across Treatment Groups

	Baseline		Followup	
	treat_gu mean (1)	treat_gd-treat_gu (2)	treat_gu mean (5)	treat_gd-treat_gu (6)
child age	6.858	-0.044	9.421	-0.064
		(0.062)		(0.121)
female	0.413	0.015	0.436	0.010
		(0.018)		(0.012)
child in school	0.275	-0.013	0.292	0.001
		(0.042)		(0.062)
child of hh head			0.881	0.020
				(0.021)
household size	9.202	-0.364	7.294	0.107
		(0.438)		(0.228)
number children	2.760	0.001	4.793	-0.002
		(0.133)		(0.140)
hh head education	2.906	-0.169	2.690	0.093
		(0.342)		(0.291)
hh head farmer	0.648	-0.010	0.556	-0.044
		(0.047)		(0.047)
total land			5.656	-1.114
				(1.366)
pukka house			0.046	0.016
				(0.026)
semi-pukka house			0.194	-0.023
				(0.056)
kaccha house			0.604	-0.023
				(0.065)
thatched hut			0.156	0.030
				(0.068)
goats			3.878	0.256
				(0.834)
sunni			0.910	-0.012
				(0.047)
urdu			0.152	-0.004
				(0.046)
sindhi			0.710	0.060
				(0.059)
Estimated bias				
estimate		0.003		-0.010
p-value		0.777		0.195

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

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

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

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

Table A.3: Child and household characteristics by school type

	program	program - govt	program - priv	p-value govt=priv	govt treat - control
	(1)	(2)	(3)	(4)	(5)
child age	7.427	-0.135*	-0.069	0.645	0.331**
		(0.080)	(0.119)		(0.116)
female	0.459	0.036	0.057	0.672	0.004
		(0.024)	(0.043)		(0.050)
current grade	1.402	-0.525***	0.308	0.026	0.390**
		(0.070)	(0.367)		(0.149)
age first enrolled	5.520	0.261***	-0.681**	0.000	0.060
		(0.062)	(0.254)		(0.093)
child of hh head	0.883	0.040*	0.014	0.697	0.002
		(0.021)	(0.064)		(0.042)
household size	7.022	-0.212	0.685*	0.047	-0.608
		(0.257)	(0.410)		(0.502)
number children	4.577	-0.107	0.303	0.240	-0.529*
		(0.176)	(0.322)		(0.287)
hh head education	2.761	-0.823***	-0.565	0.749	0.451
		(0.297)	(0.781)		(0.388)
hh head farmer	0.560	0.098**	-0.162*	0.008	0.012
		(0.043)	(0.095)		(0.073)
total land	5.400	-1.533	-1.256	0.234	-2.585
		(1.393)	(2.239)		(2.300)
pukka house	0.053	0.019	-0.079	0.117	0.001
		(0.020)	(0.061)		(0.027)
semi-pukka house	0.162	-0.131***	0.027	0.018	-0.066
		(0.049)	(0.061)		(0.121)
kaccha house	0.624	0.101**	0.006	0.358	0.038
		(0.055)	(0.098)		(0.078)
thatched hut	0.162	0.011	0.046	0.547	0.018
		(0.039)	(0.048)		(0.113)
goats	3.776	0.124	-0.205	0.617	-1.052
		(0.460)	(0.532)		(0.866)
sunni	0.914	0.070	-0.084	0.024	0.083
		(0.038)	(0.052)		(0.136)
urdu	0.165	0.007	-0.024	0.803	-0.014
		(0.038)	(0.123)		(0.055)
sindhi	0.720	0.029	0.043	0.924	0.121
		(0.046)	(0.139)		(0.103)

Note: This table reports differences in mean child and household characteristics between school types. Column (1) reports means for program schools; column (2), differences in means between program and government schools; column (3), differences in means between program and private schools; column (4), differences in means between government and private schools; and column (5), differences in means between government schools in program and control 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 A.4: Education production estimates, subject test scores

	language test scores			mathematics test scores		
	(1)	(2)	(3)	(4)	(5)	(6)
toilets and/or drinking water	0.129	0.141	0.158	0.132	0.140	0.124
	(0.098)	(0.099)	(0.111)	(0.103)	(0.108)	(0.104)
female	0.021	0.022	0.034	0.021	0.022	0.024
	(0.042)	(0.041)	(0.054)	(0.053)	(0.053)	(0.042)
age	0.093***	0.093***	0.105***	0.105***	0.104***	0.093***
	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)	(0.010)
tuition required	0.001	0.001	0.000	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
distance from home to school	-0.022	-0.024	-0.029	-0.031	-0.033	-0.022
	(0.053)	(0.053)	(0.058)	(0.059)	(0.060)	(0.052)
pct teachers <5yrs exp	0.114	0.148	0.306***	0.276**	0.301**	0.126
	(0.088)	(0.099)	(0.110)	(0.108)	(0.118)	(0.093)
pct teachers post-secondary	0.230***	0.221**	0.152	0.167	0.160	0.226**
	(0.087)	(0.087)	(0.115)	(0.111)	(0.111)	(0.091)
pct teachers female	0.061	0.067	0.107	0.099	0.104	0.063
	(0.079)	(0.079)	(0.094)	(0.092)	(0.094)	(0.081)
pct time teaching	-0.266	-0.272	-0.085	-0.102	-0.106	-0.286
	(0.256)	(0.253)	(0.342)	(0.325)	(0.322)	(0.272)
avg teacher absent ≥2 days/month	-0.007	-0.005	-0.071	-0.072	-0.070	-0.007
	(0.059)	(0.059)	(0.074)	(0.073)	(0.073)	(0.060)
pct teachers female X female student	0.050	0.050	0.018	0.026	0.026	0.045
	(0.043)	(0.043)	(0.051)	(0.051)	(0.051)	(0.043)
distance X female student	-0.000	-0.002	0.005	0.008	0.007	-0.003
	(0.022)	(0.022)	(0.024)	(0.024)	(0.024)	(0.022)
toilets and/or drinking water X female student	-0.046	-0.047	-0.061	-0.056	-0.056	-0.046
	(0.046)	(0.046)	(0.059)	(0.058)	(0.059)	(0.047)
government school		0.070			0.052	
		(0.092)			(0.125)	
program school			-0.062			-0.031
			(0.108)			(0.083)
R-squared	0.068	0.068	0.069	0.069	0.069	0.067
N	7246	7246	7112	7196	7196	7162

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