THE EFFECT OF VILLAGE-BASED SCHOOLS:
EVIDENCE FROM A RANDOMIZED CONTROLLED TRIAL IN AFGHANISTAN

Dana Burde
Leigh L. Linden

The Effect of Village-Based Schools: Evidence from a Randomized Controlled Trial in Afghanistan
Dana Burde and Leigh L. Linden
NBER Working Paper No. 18039
May 2012
JEL No. I25,I28,O12,O22,O38

ABSTRACT

We conduct a randomized evaluation of the effect of village-based schools on children's academic performance using a sample of 31 villages and 1,490 children in rural northwestern Afghanistan. The program significantly increases enrollment and test scores among all children, eliminates the 21 percentage point gender disparity in enrollment, and dramatically reduces the disparity in test scores. The intervention increases formal school enrollment by 42 percentage points among all children and increases test scores by 0.51 standard deviations (1.2 standard deviations for children that enroll in school). While all students benefit, the effects accrue disproportionately to girls. Evidence suggests that the village-based schools provide a comparable education to traditional schools. Estimating the effects of distance on academic outcomes, children prove very sensitive: enrollment and test scores fall by 16 percentage points and 0.19 standard deviations per mile. Distance affects girls more than boys—girls' enrollment falls by 6 percentage points more per mile (19 percentage points total per mile) and their test scores fall by an additional 0.09 standard deviations (0.24 standard deviations total per mile).

Dana Burde
Steinhardt School of Culture, Education,
  and Human Development
New York University
246 Greene Street
New York, NY 10003
dana.burde@nyu.edu

Leigh L. Linden
Department of Economics
The Univesrity of Texas at Austin
1 University Station
BRB 1.116, C3100
Austin, Texas 78712
and NBER
leigh.linden@austin.utexas.edu

I. Introduction

Despite calls for universal primary education and an end to gender disparities in education, developing countries have not yet succeeded in either goal. Too few children go to school and of those who do not go, girls constitute the majority. In 2006, 73 million school-age children were not enrolled in primary school. This is down from 103 million in 1999, but still far from the millennium development goal of full primary school enrollment by 2015 (UN, 2008a). Unfortunately, over half of these children are girls (55 percent), and different countries' progress toward the goal of eliminating the gender disparity in primary education is mixed. Of the 113 countries that failed to meet the Millennium Development Goals' preliminary target for gender equity in 2005, only 18 are considered by the UN to be "on track" to meet the goal by 2015. The worst performers in terms of both overall enrollment and gender disparity are countries in Sub-Saharan Africa, Oceania, and Western Asia (UN, 2008b). These challenges are particularly acute in rural areas. Globally, 25 percent of children in rural areas are out of school compared with 16 percent of children in urban areas (UN, 2008a). The same is true for gender equality. In urban areas, boys and girls attend school at similar rates on average. However, in rural areas gender disparities persist (UN, 2008c).

The major challenge is to determine whether these figures reflect a supply-side or a demand-side challenge. On the supply side, educational institutions are often scarce in rural areas, making them inaccessible to many families. And even when they are accessible, children may have to travel long distances to attend (Al-Qusdi, 2003; Kristiansena and Pratiknob, 2006; Adele, 2008). Attending school in such conditions requires significant investments in time, transportation, or alternative housing strategies like arranging for children to stay with relatives. On the demand side, rural areas usually offer few opportunities for skilled employment, possibly

reducing the returns to education. The opportunity cost of children's labor may also be high, particularly in agricultural communities (Jafarey and Lahiri, 2005; Lloyd et al., 2005; Schultz, 2004; ICLP, 2000).

Both of these causes could also explain the existing gender disparities. On the supply side, the costs of girls traveling long distances may be greater than for boys because families may believe girls face unique risks to both safety and chastity (Lloyd et al., 2005; Sutton, 1998). In addition, existing educational infrastructure may be better suited to meet the needs of boys than girls. For example, lack of separate sanitation facilities, female teachers, and gender-segregated classrooms are considered to be a greater deterrence to girls' enrollment than boys' (UN, 2008c; Al-Qudsi, 2003; Kristiansena and Pratiknob, 2006; Adele, 2008). Families also have gender-specific preferences for their children's behavior that generate potential demand-side differences. Girls are much more likely than boys to perform domestic chores like cooking and childcare, possibly making the opportunity cost of their time higher than their brothers' (Sutton, 1998). Additionally, the return to girls' education may not be as high as that of boys'. Girls may marry early and be engaged solely in managing the family, limiting the returns to education to the rearing of children. Even if women do eventually enter the labor force, the industries in which most work pay lower wages than men typically earn (UN, 2008c), and the income they generate usually benefits the family of the husband. This reduces parents' incentive to invest in their daughter's human capital by sending her to school.

In this study, we focus on a single major supply-side challenge—location of schools. Governments do not randomly choose the location of schools, and any deliberate placement of schools is likely to be correlated with outcomes of interest such as school participation and test

scores. As a result, simple correlations between the distance children must travel to school and their participation rates are likely biased.[2]

Complicating matters further, schools are structured differently based on the number of villages they are designed to serve. More remote schools, designed to serve larger populations of students, will necessarily be large, multi-room structures capable of exploiting economies of scale such as age-graded classes, direct monitoring, and support of teachers by on-site headmasters. In contrast, more proximate schools that are designed to serve fewer students may have to make accommodations not required in larger schools, such as teaching multi-grade classes and working from more remote locations. The latter also makes the schools much more difficult to monitor and support. Traditional public schools in developing countries are often of poor quality. If these factors make smaller schools even worse, then estimates that rely only on the relationship between student performance and the distance to traditional large schools may not fully represent the effects of a policy shift from less to more densely located schools.

We take advantage of a program being implemented in rural northwestern Afghanistan that places schools in villages underserved by traditional public schools. We use a randomized controlled trial to ensure the independence of distance and other household characteristics. And since the schools are designed to serve children only in the selected villages, the treatment captures the necessary infrastructural differences. Specifically, within a sample of 31 villages, we randomly assign village-based schools[3] to 13 villages to estimate the one-year effects of these

---

[2] For example, governments could place schools in areas with high demand for education in an effort to meet that demand, generating a positive bias. Alternatively, in areas with low school participation, governments might choose to increase the supply of educational resources in the hopes of increasing the number of enrolled children, potentially creating a negative bias. Or governments might simply place schools in areas with the densest populations, which would have an ambiguous effect. Similarly, families may also choose to locate endogenously, with families that value education more choosing to live near educational institutions or other geographic amenities correlated with greater access to educational resources.

[3] The existing fields of study use different terms for these schools. Educators working in international development and humanitarian aid, for example, refer to these schools as "community-based schools".

schools on the enrollment and academic performance (math and language skills) of 1,490 primary school-age children. The remaining villages receive schools the subsequent year.

Placing a school in a village dramatically improves academic participation and performance among all children. The presence of a village-based school increases overall enrollment by 42 percentage points and increases test scores by 0.51 standard deviations for all children in the village. Those children caused to attend a formal school as a result of the intervention experience an increase in test scores of 1.2 standard deviations.

While we find consistent gains for all children, these benefits accrue disproportionately to girls. The treatment virtually eliminates the gender disparity in enrollment and improves the disparity in test scores by a third in a single year. Specifically, girls' enrollment increases by 17 percentage points more than boys', reducing the gender gap in treatment villages to a statistically insignificant four percentage points from 21 percentage points in control villages. This suggests that, at least in villages with schools, girls are given the same educational opportunities as boys at the primary school level. The net effect of the participation changes results in a net increase in test scores of 0.25 standard deviations more than boys, reducing the gender gap in achievement (-0.685 standard deviations in the control villages) by over a third in only a single academic year. Additionally, we find that the intervention disproportionately increases enrollment among students engaged in several household activities hypothesized to interfere with school attendance. However, these effects do not explain the larger effects on girls overall.

Using the fact that 15 percent of the children who would have attended a traditional school attend a village-based school due to the treatment, we test for differences in the performance of these children as a measure of the relative quality of the two types of schools.

Using a two-stage estimator that controls for the projected probability of counter-factual enrollment in each research group, we find no evidence of a difference in quality.

We then estimate the effect of distance[4] itself of enrollment and test scores using two strategies. First, the similarity in the quality of the two types of schools allows us to use an instrumental variables strategy in which we use the assignment of a school to a village as an instrument for the distance between a household and the nearest formal school. As a check, we exploit the random assignment of the treatment schools by estimating the relationship between both outcomes and the distance to the nearest village-based school. The results are similar.

We find that, unlike in previous cross-sectional studies described below, enrollment rates and achievement are very sensitive to differences in distance. Overall, enrollment rates of children fall by 16 percentage points per mile and test scores fall by 0.19 standard deviations per mile, suggesting that a change in distance of a single mile or two can replicate the learning effects of some of the most effective educational interventions. Girls are also more sensitive than boys to distance. The enrollment rate for girls falls by 19 percentage points per mile, while boys' enrollment falls by only 13. Girls' test scores fall by 0.24 standard deviations per mile, while boys' fall by 0.15.

Unlike the conflicting results in the existing literature, these results suggest that village-based schools can be an effective tool for improving primary education in rural areas. Existing estimates of the effects of school access follow two different strategies. Those that estimate the effect of a school existing within a rural village generally find strong positive effects of the

---

[4] While the main difference is the distance needed to travel to the schools, this does not mean that the primary barrier to enrollment is the travel costs faced by the child. In fact, the variation in treatment effects we observe in Section V.B suggests that the primary issue may be safety. However, we are unable to provide more than suggestive evidence. Given that proximity is a sufficient condition for significantly improving school participation, identifying the underlying factors within this relationship is an important area for future research. This is particularly true for secondary schools where economies of scale may be a more important factor in the quality of schools due to overall lower enrollment levels and the types of subjects covered (e.g. the availability of science labs and other relatively expensive but sharable educational resources), making village schools less cost-effective than at the primary level.

presence of a school as long as the endogeneity of school placement is taken into account (Andrabi, Das, and Khwaja, 2010; Duflo, 2001; Foster and Rosenzweig, 1996; Kazianga et al., 2010; Kim et al., 1999; and Pitt, Rosenzweig, and Gibbons, 1993). Estimates that fail to account for endogenous school placement typically find little effect because governments seem to place schools in areas with low enrollment rates (Pitt, Rosenzweig, and Gibbons, 1993).

The second strand of literature attempts to use information on the distance between households and schools to estimate the elasticity of enrollment in distance or travel time using cross-sectional data. While there are exceptions (Lavy, 1996), most of these studies seem to suggest that child enrollment levels are relatively inelastic to distance even over surprisingly long distances (Filmer, 2007; Handa, 2002). Filmer (2007), for example, estimates the cross-sectional relationship between enrollment and distance for children aged six to 14 in 21 rural areas of low-income countries around the world. He finds the largest decline in enrollment per mile in Niger using the 1998 round of the Demographic and Health Survey, but even then, it is only 4.8 percentage points per mile, leading to the conclusion that "expectations for large overall increases in enrolment as a result of school construction should be tempered"(p103).

Our results build on this literature in several ways. First, we provide experimental confirmation of the estimates of the effects of village-based schools from natural experiments and instrumental variables strategies. Methodologically, our approach is closest to Thornton (2008) who randomly locates HIV testing centers to experimentally manipulate individuals' knowledge of their HIV status.[5] Second, analogous to Andrabi, Das, and Khwaja's (2010) exploration of the role of public schools in Pakistan, we introduce variation in both school location and school structure. Finally, by providing experimentally estimated effects of the

---

[5] Our methodology is also similar to Kim et al. (1999) who conduct a randomized evaluation of community-organized schools based on a girls' stipend program in urban Quetta in the Balochistan province of Pakistan.

relationship between distance and enrollment, we present estimates that are in line with the large effects observed from the placement of schools within villages.

The paper is organized as follows. Section II provides a description of the state of education in Afghanistan and the intervention. Section III outlines the research design, including the data and models used in the analysis. Section IV provides an overview of the sample and verifies the internal validity of the study. In Section V, we analyze the effects of the program, estimating first the effects of the intervention, disaggregating the average results by gender, comparing the quality of the schools, and then directly estimating the effects of distance on children's school participation. We conclude in Section VI.

II. Afghanistan and the PACE-A Program

A. Education in Afghanistan

Afghanistan's educational infrastructure was never comprehensive, but the system that existed in the 1970s has been crippled by decades of war. As of 2007, half of school-age children were not enrolled in school, and of primary school-age children, only 37 percent attended school (UNDP, 2007). Girls fare especially poorly. Of children with any formal schooling, boys average eleven years of education, while girls average four (UN, 2009). Girls make up only a third of the student population (MoE, 2007).[6] The vast rural areas of the country are disproportionately underserved by the existing infrastructure. Despite the fact that 83 percent of schools are in rural areas, urban schools serve a disproportionate share of the student population—35 percent of students are in rural schools as opposed to the 65 percent in urban ones (MoE, 2007).

---

[6] Even these numbers are an improvement, however. In 2001, only 900,000 children were enrolled in school, but by 2007, this number had reached 5.4 million (UNDP, 2007). In 2007 alone, 800,000 new children were enrolled and girls made up 40 percent of these students (MoE, 2007).

Our sample comprises villages from Ghor province in northwestern Afghanistan. Compared to the south and southeast parts of the country, the area, during the time of the study, was largely unaffected by the ongoing war. It does, however, face the same security challenges as other remote rural areas, including lawlessness and tribal conflict that are often the result of weak or inaccessible judicial and police services. The enrollment rates in this province are similar to those in other rural areas of Afghanistan. Of children aged six to 13, only 28 percent are enrolled in school. The gender gap in enrollment is 17 percentage points—with 35 percent of boys and only 18 percent of girls participating in school. The majority of villages have little access to educational resources—only 29 percent of the population lives within five kilometers of a primary school (MRRD, 2007).

B. Existing Educational System

Afghanistan has an official national curriculum which is delivered to students through three types of institutions: traditional government schools, village-based schools, and government-sponsored madrassas. We refer to these institutions collectively as "formal schools".[7] Traditional government schools are typical large-scale public schools, designed to serve numerous children from multiple villages. These are currently the main providers of education, with 95 percent of enrolled students attending. They are typically age-graded and organized around a headmaster responsible for directly overseeing a team of teachers, most of whom have received formal education training.

---

[7] The typical language used to describe these institutions varies. Since the goal of this paper is to explore the degree to which village-based schools can effectively teach the official national curriculum, we use the term formal school to refer to all institutions that teach the national curriculum, distinguishing these institutions from those that do not, such as local mosque schools (which provide informal religious instruction, much like Christian Sunday schools), madrassas which exclusively provide formal religious instruction, and other similar programs.

Village-based schools are public schools that are designed to serve only children who are living in close geographic proximity.[8]   They have been one of the major educational interventions supported by international aid agencies, with the goal of increasing exposure to the official government curriculum in rural areas—particularly among girls. While these schools are currently managed by local staff employed by international development organizations, they will eventually be integrated into the national system (Guyot, 2007; MoE, 2007).

Compared to traditional government schools, the major challenge for village-based schools is quality. While the traditional schools face all of the usual challenges—including teacher shortages, high rates of teacher absenteeism, insufficient resources, and poor infrastructure (Adele, 2008; Alcazar et al., 2006)—village-based schools face additional hurdles. They are almost always remotely located, serve a much smaller number of students than traditional schools, and rely on the existing infrastructure within villages. While the schools are usually provided with basic equipment such as floor mats, writing material, and textbooks, they are housed in spaces provided by the village. Classes are not age-graded, although boys and girls are often taught separately.

Finally, madrassas are religious schools that provide religious instruction in addition to teaching the official curriculum.[9] They account for only 1.5 percent of total enrollment. In the analysis that follows, our sample includes one such school, but only nine children attend it. As a result, we classify the school as a formal school, but focus on the comparison between the traditional public schools and the village-based schools.

---

[8] Note that in this study we refer to these schools as village-based schools to distinguish them from the traditional schools designed to serve multiple villages. However, in the international education literature as noted above, they are commonly referred to as "community-based" schools.

[9] Madrassas also exist outside of the government system. These madrassas may rely exclusively on religious texts and are used primarily to prepare young men for a role as a professional religious leader.

C. Village-Based Schools Intervention

The evaluated intervention is part of a larger country-wide effort called the Partnership for Advancing Community-based Education in Afghanistan (PACE-A). It is a five-year, USAID-funded program to expand educational opportunities to children, especially girls, in areas of Afghanistan that lack immediate access to formal governmental schools. The part of the program involved in our study is run by Catholic Relief Services (CRS) through local staff.

For each program village, CRS provides educational materials as well as training for teachers using the official government training programs. Educational materials include writing utensils, notebooks, books, and teacher materials. Teachers receive training on topics such as monitoring and evaluation, teaching methods, and instruction using both specially designed programs and the Afghan Ministry of Education Teacher Education Program (TEP) materials— the same materials used to train teachers in traditional schools.[10] Shortly after our experiment concluded, the Ministry of Education decided to include PACE-A teachers as formal government employees in the same category as official contract teachers.

Within the village-based schools, students are taught the official national curriculum. They study for a minimum of 2.5 hours a day, 6 days a week (excluding Fridays). Rather than the formally trained teachers who staff traditional schools, the teachers in village-based schools are locally recruited, educated individuals, typically with less than twelve years of education. CRS staff aims to visit each village-based school at least monthly to assist teachers, but during the period of our study, visits were made less frequently (once every few months) due to difficulties traveling to these remote locations.

---

[10] Use of the TEP as a training course also streamlines the certification of teachers into the Ministry of Education system of educators.

III. Research Design

A. Experimental Framework

A randomized controlled trial generates exogenous variation in treatment assignment, allowing for a direct estimation of the causal effect of the intervention on the chosen outcomes. Our initial sample consisted of 34 villages chosen by CRS to receive schools over a two-year period starting in the summer of 2007. We randomly assigned a subset to receive a school a year before the rest. The treatment implementation exactly followed the random assignment. All treatment villages received a school and no control village did. However, we did face three challenges.

The main threat to validity in this framework is that the future receipt of the treatment might change the behavior of control households. The most likely avenue would be for families either to withdraw their children from a traditional school or postpone the enrollment of new students. Phase-in designs have been successfully used to evaluate many educational programs in developing countries (Miguel and Kremer, 2004 is one of the first examples), and the available data corroborate the use of the design in the current context. As we show below, the control group enrollment rates and gender differential match the province-level averages very closely. Furthermore, during qualitative interviews, we received no reports of individuals changing enrollment decisions prior to the receipt of the treatment.[11]

Second, we were unable to randomize at the village level. Individual villages are connected to surrounding villages by political and cultural alliances that would be strained if one village were to receive an intervention that was withheld from the others. As a result, we randomized the villages in 12 equally-sized village groups. Statistically, this requires us to take

---

[11] In fact, the primary concern in the control villages was the opposite—they feared that they would not receive the schools as promised, a prescient concern given that many of the schools in control villages were not started on time.

into account not just village-level, but village-group-level, correlation in outcomes in the statistical models described in Section III.C.

Finally, as part of the lawlessness and tribal conflicts described in Section II.A, households in remote rural areas use armed conflict to resolve disputes. Unfortunately, a conflict that included one group of three villages broke out shortly after the randomization. While we monitored the situation for the duration of the study, we were unable to conduct any surveys in these villages and ultimately had to drop them from the sample. As a result, the final sample consisted of 11 village groups (five treatment and six control) and 31 villages (13 treatment and 18 control).

B. Surveys and Available Data

To assess the effects of the intervention, we conducted two waves of surveys in the fall of 2007 and then again in the spring of 2008. In what follows, we treat the fall 2007 survey as our primary outcome measure because it was conducted just prior to the end of the traditional school year. However, the school year ends because by the end of the fall, it becomes impossible for anyone to move easily between villages due to heavy snowfall and limited infrastructure. Since students in village-based schools do not face such limitations, the schools operated into the winter period. While our estimates belie the argument, we conducted the second survey to assess whether this extra period of instruction would prove additionally beneficial.

We designed both surveys to meet three objectives. First, in the absence of a baseline survey, we collected information on socio-demographic characteristics that would not change as a result of the treatment, providing variables that we could use to compare the treatment and control groups to assess whether the randomization did indeed create comparable research groups. Second, we matched the data over time to create a panel data set. This allows us to

compare attrition patterns between the treatment and control groups to ensure that differential migration patterns or other factors affecting the availability of households to complete a survey did not differentially affect the treatment and control groups. Finally, as our primary outcome variables, we asked parents about their children's school enrollment and directly tested children on their math and language skills.[12]

The socio-demographic information collected in the first module of the survey was chosen such that each variable was unlikely to be affected by the treatment. This provides characteristics to assess the comparability of the treatment and control groups and to use as controls in regressions comparing the outcomes of children in the research groups. This included the length of time that the family had lived in the village, the family's ethnic identity, the occupation of the primary earner, the primary earner's level of education, the family size, land holdings, and other similar characteristics. We collected the same information in both surveys to ensure that the same control variables could be used when analyzing data from either survey.

We also collected information that would allow us to match across survey rounds and to construct the various distance measures. This included the name of the head of the household and the longitude and latitude of each home. Geographic information was also collected on every government school and madrassa that a family reported their child attending, as well as the village-based schools in the treatment villages.

---

[12] The surveys were designed with the input of both CRS staff (both national and international) as well as our own survey team. We employed the following process to develop the survey questions. First, from our prior experience and research in Afghanistan, we drafted possible measures independently or based on instruments that had been tested in other settings. (Some of the questions were adapted from a pilot study in Panjshir Province by Burde (2008).) These were reviewed by the CRS staff and our surveyors, who provided invaluable input in further adapting the measures to the local context (for example, in identifying how families counted livestock and possible child activities within the household). The measures were then field tested. Finally, because we carried out multiple rounds of data collection, the responses of families to the fall 2007 survey were used to update the structure of the spring 2008 survey.

The surveys contained two outcome modules with specific questions for each child between the ages of six and eleven within the household. The surveyor asked the household respondent for a list of all children in the age range targeted by the program. Then for each child, the surveyor collected information on whether the child attended school, the type of school attended, and the frequency of attendance.[13] The child's age, gender, and relationship to the head of the household were also collected.

Finally, for each child that was available to question directly, the surveyor administered a short test covering math and language skills. The questions were taken directly from the first grade government textbooks to ensure that the test covered material from the official Afghan curriculum. The math section included questions on number identification, counting, greater than or less than, addition, and subtraction. The language section covered Dari—the children's primary language as well as the medium of instruction in all government schools in the region— and included questions on letter identification, reading words of varying difficulty, basic grammar (subject-verb agreement), and simple reading comprehension. The administered tests differed in that the spring 2008 survey covered a larger number of questions than the fall 2007 survey, though the same topics were covered in each.

Since the school participation information is self-reported, we are careful to assess the accuracy of this information (Barrera-Osorio, Bertrand, Linden, and Perez, 2011). First, the information itself is not obviously fabricated—for example, all parents did not report sending all of their children to school every day. Parents do not report that their children attend school during times when the schools are closed—such as during the winter. And the levels of reported

---

[13] We also attempted to collect retrospective information on children's school enrollment in the spring 2008 survey. Unfortunately, these estimates seem to suffer from relatively severe hindsight bias. For example, the village-based schools were started in the summer of 2007, but a significant number of treatment families report sending their children to these schools in the spring of 2007 when the schools did not exist. As a result, we restrict our attention to questions that relate to the school participation at the time the surveys were actually administered.

school enrollment are consistent across survey rounds, seem reasonable for the context, and follow expected patterns in the data such as higher enrollment levels for boys than girls. In fact, the results are identical to government estimates of the average enrollment rates of boys and girls within Ghor Province (MRRD, 2007).

Second, we also conducted qualitative, semi-structured interviews with parents after the final survey was conducted. Lying in this format is much more difficult due to the sequence of detailed follow-up questions. The information provided in those interviews was consistent with the information provided in the surveys.

Finally, the test scores in our data cannot be fabricated, providing two additional checks. We compare the relationship between demographic characteristics of the family and the probability of enrollment with the relationship between those variables and test scores. We find that both measures follow the same patterns: for example, boys and older children attend school more often and also score higher on the exams. And second, for all of the models, we estimate the effect on both enrollment and children's test scores. In all instances, both measures provide consistent outcome estimates, showing both large effects on enrollment and learning levels.

C. Econometric Models

Our analysis proceeds in five steps. First, we assess the comparability of the treatment and control groups using the available demographic data to assess the internal validity of the study. Having established that the groups appear comparable, we then estimate the net effects of the program by directly comparing the enrollment rates and test scores in the treatment and control groups. Third, we specifically assess the relative effects of the program on girls. Fourth, we assess the quality of the village-based schools by estimating the fraction of the treatment effect

that can be explained by differences in formal school enrollment between the treatment and control groups. Finally, because the schools appear to be of net similar quality, the remaining primary difference between the village-based schools and the traditional schools is distance. This allows us to re-parameterize the effects of the treatment on enrollment and test scores in terms of distance and to estimate the marginal effects of distance to school on each.

In the first two parts of the analysis, our preferred estimate of the treatment effect is a direct comparison of the treatment and control groups controlling for socio-demographic characteristics and estimated by ordinary least squares using the following model: [14]

$$Y_i = \beta_0 + \beta_1 T_k + \beta_2 X_{ij} + \varepsilon_{ijk} \tag{1}$$

The variable Y_i in this specification is the child-level outcome variable for child i, in household j, and village group k. The variable T_k is a dummy variable for whether or not village group k was selected for treatment in the randomization process, and the vector of variables X_{ij} contains the socio-demographic controls (the complete list is the same as those included in Table 2). The coefficient β_1 then provides the estimated treatment effect. We also estimate a version of equation (1) that does not include socio-demographic characteristics to estimate the simple difference in outcomes and to establish internal validity.

To estimate the difference in quality of the two types of schools, we cannot, as we show, simply compare the scores of children enrolled in each institution because the characteristics of students in each type of institution differ significantly. If we knew the counter-factual enrollment patterns for the treatment group, we could identify which children in the treatment group would have been enrolled in traditional schools absent the treatment. These would be the

---

[14] Many of the outcome variables that we consider are binary. In the presented results, we exclusively use linear probability models in order to use a consistent statistical models for all of the treatment effect estimates. For estimates of the effects of the program on enrollment, we have also estimated the main treatment effects using probit models with consistent results.

same children who we observed to be enrolled in the control group. Then, as long as some of these students switch from the traditional schools to the village-based schools in the treatment group (which we show to be true), we could directly compare the test scores of these students through the following equation:

$$Score_i = \phi_0 + \phi_1 T_k + \phi_2 T_k * Not\_Enrolled_i + \phi_3 Not\_Enrolled_i + \phi_4 X_{ij} + \varepsilon_{ijk} \qquad (2)$$

where $Not\_Enrolled_i$ is an indicator variable equal to one if a child would not have been enrolled absent the treatment, and zero otherwise. The coefficient ϕ_1 is then the estimated difference in test scores for children who would have been enrolled absent the treatment.

While we obviously do not have this counter-factual information, we can estimate it from our estimate of the counter-factual, the control group. To create a variable similar to the indicator variable in equation (2), we estimate a model of a child's propensity not to be enrolled in a formal school in the control group by regressing an indicator for whether a child is not enrolled on the set of socio-demographic characteristics in Table 2 using ordinary least squares.[15] We then predict, for each child, the probability of not being enrolled absent the treatment using these estimated coefficients and include that estimated regressor in the following second-stage equation:[16]

$$Score_i = \varphi_0 + \varphi_1 T_k + \varphi_2 T_k * Prob\_Not\_Enrolled_i + \varphi_3 Prob\_Not\_Enrolled_i + \varepsilon_{ijk} \qquad (3)$$

The specification is the same as that for equation (2), but replaces the indicator for counter-factual non-enrollment with the probability of not being enrolled. The interpretation of the coefficient on the treatment indicator, however, remains the same—this is the estimated

---

[15] As in Footnote 14, a probit model yields similar results.
[16] Note that because the estimated probability of not being enrolled is a linear combination of the vector of control variables, it cannot be included in the regression along with the full set of controls. The two are perfectly collinear. Below, we estimate the model both with the full set of controls and controlling for the estimated probability of not being enrolled separately.

difference in test scores for a child who would have certainly gone to a traditional school absent the treatment. There are, of course, few children with estimated probabilities of non-enrollment in this range. So, the test does not directly compare the scores of students with characteristics that make them very likely to have attended school. Rather, it projects the behavior of those children out of sample, testing whether the ratio of the difference in treatment effects by probability of non-enrollment diminishes at a rate that would be consistent with such children not scoring differently. Finally, because the predicted variable is an estimated regressor, we need to account for the uncertainty in the first-stage model when testing the significance of the coefficients in the second stage. We do this through the wild-cluster bootstrap procedure described below.

The results suggest that the schools are overall equivalent in quality. If the schools are similar, then the main difference between the schools is the distance the children have to travel. This allows us to then estimate the relationship between enrollment, test scores, and distance by estimating the following equations in an instrumental variables framework:

$$Enrolled_i = \alpha_0 + \alpha_1 Distance_j + \alpha_2 X_{ij} + \varepsilon_{ijk} \qquad (4)$$

$$Distance_j = \delta_0 + \delta_1 T_k + \delta_2 X_{ij} + \upsilon_{ijk} \qquad (5)$$

The variables $Enrolled_i$ and $Distance_j$ are, respectively, an indicator variable for whether or not a child is enrolled and the distance to the nearest school. The remaining variables are defined as in the previous equations. Equation (5) provides the first stage regression for equation (4), and the coefficient α_1 provides an estimate of the causal effect of distance on enrollment. Because the instrument is an indicator variable, the estimated effect is simply a re-parameterization of the overall average treatment effect for these outcomes. As a check, we also estimate this

relationship using only the distance to the village-based schools to leverage the explicit random assignment of their locations.

In all of these models, it is important to take into account the correlation between children's performance and behavior. Not taking this into account would cause us to underestimate the variance of the estimated treatment effect, resulting in an over-rejection of the null hypotheses at any significance level (Bertrand, Duflo, and Mullainathan, 2004). To be conservative, we cluster the standard errors at the level of the village group, the unit of randomization. However, the estimates are generally not sensitive to the level of clustering. Estimates of random effects models at various levels suggest that the majority of the variation occurs at the household level.[17]

The main constraint we face when accounting for this correlation is the small number of clusters at the village group level. Since there are only 11 village groups, the standard clustering strategy with standard errors calculated relative to the limiting t-distribution are likely to over-reject the null hypothesis (Cameron, Gelbach, and Miller, 2008). To correct for this, we do three things. First, for all tables except Table 7, we calculate statistical significance relative to the small sample t-distribution with 10 degrees of freedom. Second, for the estimates which can be estimated using only the inter-village group variation in the outcomes (test score and enrollment), we aggregate the data up to the village group level and estimate the differences using just the 11 observations (Angrist and Pischke, 2008). We then test for differences in the average outcomes by calculating the Fisher Exact P-values for the Wilcoxen Rank Test.[18]

---

[17] To investigate the correlation in outcomes among children at each level, we estimated the treatment effect on our main outcomes of enrollment and tests scores using equation (1) while including random effects at each level. Within this framework, allowing random effects at the village group and village level only explains between one and six percent of the residual variance while random effects at the household level explains about 20 percent.

[18] The Wilcoxon Rank Test tests for the difference in sum of the ordinal rank of village groups for each research group using the ordinal rank of the village group when ranking the entire sample by the average outcome (Wilcoxon, 1945). Testing for the difference in rank yields higher power than simply testing for differences in the mean

Finally, for the estimates which rely on both inter- and intra-village group variation in outcomes (interactions with gender, distance, and probability of non-enrollment), we check the estimates from the cluster regressions by bootstrapping the distribution of the test statistics using the wild-cluster bootstrap (Cameron, Gelbach, and Miller, 2008). These estimates are consistent with those presented in the tables.

IV. Research Sample

A. Sample Size and Coverage

The survey is a census of all households in the villages available to be surveyed. Table 1 provides a tabulation of the responses from the survey comparing the coverage rates between the treatment and control groups. The information for the fall 2007 survey is provided in the first four columns and the information for the spring 2008 survey is provided in the last four columns. In the latter survey, households are listed if they were identified as being physically present in the village at the time of the survey regardless of whether they were surveyed in the previous round. In each case, we provide the total number of respondents in each category for each research group, followed by the difference between the treatment and control group and then the total number of respondents.

Along every category, the treatment and control groups are similar. Panel A provides the total number of households identified and the number surveyed. In both surveys, about 93 to 95 percent of households were surveyed, and coverage rates among households identified as being available to survey were similar across the two research groups. Only about two-thirds of surveyed households had children whose ages made them eligible to attend the village-based

---

outcomes. We calculate the Fisher Exact P-values by calculating all possible outcomes of the random assignment process and then constructing the resulting distribution of the difference in the sum of ranks for the research groups under the null hypothesis of no treatment effect.

schools, and again the fraction of families meeting these criteria was the same across research groups in both rounds of the survey. In total, this provides a sample of 805 households in the fall 2007 survey and 794 in the spring 2008 survey. Finally, 1,490 and 1,477 eligible children from the fall 2007 and spring 2008 surveys, respectively, were identified in these households and had enrollment information provided to the surveyors. Of these, 1,374 children from the fall 2007 survey and 1,401 children from the spring 2008 survey were available to be tested.[19] As with the total number of households, almost all of the identified children were tested (92 and 95 percent in the fall 2007 and spring 2008 surveys, respectively), and the coverage rates are balanced across the research groups.

B. Internal Validity

The purpose of a randomized evaluation is to ensure that the assignment of the treatment is orthogonal to other characteristics of the sample that may be correlated with school participation and test scores. Such correlations could compromise the internal validity of our study in two ways. First, it is possible that the randomization simply created treatment and control groups with large differences in the characteristics of children that are also correlated with the outcomes of interest. Second, even if the research groups are initially similar, it is possible that over time the sample may be affected by processes that differentially change the composition of the two groups. For example, if the treatment group proved more mobile than the control group, we may lose more families due to migration. The net effect would be that while the groups were initially similar, differences would emerge over time that would compromise the study's internal validity.

---

[19] Our initial sample included a small number of extremely large and wealthy households that we exclude as outliers. These included families with more than 20 household members, 10 units (jeribs) of land, or over 50 head of sheep. In each case, these families constituted the top one to two percent of households along each measure, and in total represented 3.3 percent of households in the fall 2007 survey and 2.9 percent in the spring 2008 survey. None of the point estimates are sensitive to the exclusion of these households.

We measure the differences both in composition and changes in composition between the two surveys and find both that the randomization succeeded in creating comparable groups of children and that the groups did not change differentially between the two surveys.

To check the comparability of the two research groups, we directly compare the average children in the treatment and control groups using socio-demographic characteristics that would not have been affected by the presence of a closer school. These comparisons are presented in Table 2. Using the data from the fall 2007 survey, the first three columns contain estimates for all of the children in the sample while the second set of three columns contains estimates for only those children who took the exam. The first column provides the average characteristic of the treatment group. The second column contains the control average, and the third column contains the difference estimated using equation (1). Panel A contains child demographic variables and Panel B contains the household characteristics. Columns four through six contain the same estimates, but using only the subsample of children that were available to test at the time of the survey. Columns seven and eight contain the coefficients from a regression of the respective outcomes on the provided control variables using only the children from the control group.

On average, all of the differences are practically small, and none of the differences are statistically significant. Consider, for example, the number of sheep owned by a household. On average, treatment families own about 7.6 sheep per household when considering the sample with all children, while the control families own 5.6 sheep, yielding a difference of 1.9 sheep per household. Two sheep is a relatively small difference, especially when we consider the relationship between the number of sheep and our outcomes presented in columns seven and eight. The correlation with probability of enrollment is positive, which may reflect the small difference in family wealth, but the estimated coefficient is only 0.9 percentage points per sheep.

Therefore, the difference of 2 sheep reflects a possible difference in enrollment rates between the two groups (absent the treatment) of 1.8 percentage points. Given that we estimate treatment effects in enrollment of 40 to 60 percentage points, it is unlikely that these small differences in the composition of the groups could significantly affect the estimated treatment effects.[20]

Given the nature of the study, a particularly important variable to consider is the distance each child would have had to travel to attend the nearest formal school, had we not implemented the treatment. Using the geographic information we collected, we estimate the distance between each house and every formal school, excluding the village-based schools. This allows us to calculate the shortest distance children would have to travel absent the treatment for every child in the sample. The averages are presented in the last row of Table 2. On average, children live about 3 miles from the nearest formal school, and the average difference in the distances between the children in each group is only a quarter of a mile. Figure 1 shows these differences graphically by plotting a non-parametric estimate of the density of distances for the treatment and control children. The distributions are very similar, especially when compared to the distribution after the treatment, presented in Figure 2.

Columns four through six of Table 2 show the comparisons for just the subsample of children who were available to be tested at the time of the survey, and the results are very similar to those in the first three columns. None of the differences are statistically significant and they are all small in magnitude. This confirms that, among all of the children for whom we were able to collect information and the subset that we were also able to test, the children in each of the

---

[20] The relative nature of the magnitudes of the observed differences is an important general point that is relevant to all of the comparisons estimated in this manuscript. Many successful interventions designed to change enrollment levels have effects of only a few percentage points, and for these studies, a possible bias of two percentage points would be important. Similarly, in Section V.C., we argue that the different types of schools seem to deliver similar levels of test scores despite significant differences in structure. Again, this does not mean that the individual differences do not affect students' test scores, but rather that, in the context of an intervention which causes individual students' test scores to increase by 1.2 standard deviations (the effect on the compliers), these resulting differences are small.

research groups are comparable, suggesting that the randomization did indeed succeed in creating similar treatment and control groups.

Table 3 investigates whether the sample of children we observe changes significantly over time. Panel A contains the raw attrition rates between the fall 2007 and spring 2008 surveys for each group. On average, the attrition rate is only about 16 to 17 percentage points with a difference of only one percentage point between the two groups. This similarity suggests that the attrition patterns are similar between the two groups. However, even with similar rates, differences in the composition of attritors and non-attriors could cause difference to emerge between the two groups. This is assessed in Panel B. The first three columns provide the net effect of the attrition process by comparing the relative characteristics of the children whose families were surveyed in both rounds of the survey. The final three columns show the relative characteristics of attriting and non-attriting children in the two groups.

Looking at the first three columns, column one contains the average characteristics of the non-attriting treatment children from the fall 2007 survey, column two contains the average characteristics of the non-attriting control children, and column three contains the differences across research groups estimated using equation (1) without controls. The results confirm that the attrition process did not generate significant differences in the characteristics of the two groups. As in Table 2, all of the differences are small and, with the exception of the duration of the family's tenure in the village, statistically insignificant. In fact, the differences between the non-attriting children are almost identical to the differences in the entire sample presented in Table 2.

The last three columns in Table 3 show the reason for these similarities. On average, there are almost no differences between attriting and non-attriting children in either of the

research groups. Column four provides the average difference from treatment villages in characteristics between children who attrited from the fall 2007 sample and those that did not. On average, the children are very similar. The largest differences are in the duration of the family's tenure in the village (attritors have lived in the village 5.3 years less than non-attritors) and the age of the head of the household (the heads of attritors' households are 1.4 years younger), but these differences are small given the marginal correlations presented in Table 2, and all are statistically insignificant. Column five provides the same estimates for the control group and finds similar patterns. The differences in attrition patterns are compared directly in column six. The results confirm that, in fact, the patterns are similar across the two groups—the differences are small and all are statistically insignificant.

These results are consistent with the observations of our survey team. They reported that the main reason for failing to observe a family was not migration or other causes that might have a clear relationship with wealth or some of the other characteristics correlated with academic performance. Rather, attrition was caused by factors that would be common to almost all of the families in the sample, such as traveling out of the village temporarily to visit family members or going to a local market. As a result, the net differences in the characteristics between attritors and non-attritors of the treatment and control groups are small.

Finally, in results not presented in this draft, we perform similar comparisons to those in Table 2 using the data from the spring 2008 survey. We also compare the characteristics of children for whom we had only enrollment information to those of children for whom we were able to obtain both enrollment information and test scores. And we compare the attrition patterns using just the sample of children who provided a test in both surveys. In all instances, the observed differences were as small as those presented in Tables 2 and 3.

V. Outcomes

Given that the randomization created comparable treatment and control groups, we can attribute any difference between the groups in enrollment and test scores to the receipt of the treatment. We assess these differences in four steps. First, we estimate the overall average effects of the program. Second, we estimate the relative effects of the program on girls. Third, we check for differences in quality. Finally, finding no difference in quality, we estimate the effects of distance on enrollment and test scores. Overall, we find the intervention to have a very large impact on children's academic outcomes and to ameliorate significantly the existing gender disparities in our sample.

A. Overall Treatment Effects

We are primarily interested in the effects of the program on two outcomes: enrollment and test scores. Since the main purpose of establishing the new schools is to expose more children to the official Afghan curriculum, we first analyze the effects of the program on children's enrollment in schools teaching this curriculum. Then, we assess the differences in children's test scores using the tests administered in both survey rounds. Along all dimensions, the program proves extremely effective.

Table 4 contains the main outcome of interest—enrollment in formal schools. Panel A includes all children while Panel B includes only those children taking a test in the respective survey round. For each panel, we first estimate the enrollment levels in formal schools during the fall of 2007, formal schools in the winter of 2008, and the number of days that families report their children attend a formal school each week for those children who are attending. The first column contains the average enrollment rates for the treatment group followed by the same

estimate for the control group in the second column. We present three different difference estimates: the simple difference in these rates estimated using equation (1) without controls, the average differences controlling for socio-demographic characteristics using equation (1) with controls, and the estimated difference calculated at the village group level using only average estimates per group and calculating statistical significance based on the Fischer Exact P-values from the Wilcoxon Rank Test. As described in Section III.C, this latter estimate confirms that the statistical significance of the results does not rely on within-village-group variation in the outcomes.

The program has a large impact on child enrollment. Turning to the first row of Panel A, the overall increase in enrollment rates in formal schools for treatment villages is about 42 percentage points—a very large increase in enrollment over the control enrollment rate of 27 percentage points. The next row shows that formal school enrollment continues into the winter for the treatment group, emphasizing that an advantage of the village-based schools is a more flexible academic year. Another possible advantage of the village-based schools is that students might be able to attend more frequently. The third row shows, however, that children who are enrolled attend at equal rates. While the differences are statistically significant, they are practically very small—less than a tenth of a day a week. This suggests that school participation changes mostly on the extensive margin.

Panel B shows the same comparisons, but only for children who were tested at the time of the survey. Consistent with the fact that there are no systematic differences between tested and non-tested children, there is no difference in the results for children who also took the test when compared to the results estimated using the entire sample.

Comparing the estimates in column three with those in column four, the estimates from the simple difference estimator and those from the estimator that controls for demographic characteristics are very similar. This similarity reinforces the conclusions of Section IV.B that the treatment and control groups are comparable. Had they been significantly different along dimensions correlated with child enrollment, the point estimates for the two estimators would differ significantly. Finally, the significance levels of the estimates presented in column five confirm that the results are independent of the use of within-village-group variation.

The ultimate effect of the village-based schools is a high level of primary school enrollment given the control average of 27 percentage points. While many children still do not attend school, an impressive 74 percent of children in the treatment group do. This is still below the overall Western Asia[21] average (for 2006), but it is equal to the global rural average enrollment rate (UN, 2008c). This suggests that village-based schools could be an important tool for achieving universal primary education in rural areas.

Table 5 presents the difference in average test scores between the treatment and control groups for both survey rounds. Columns one through five are organized as in Table 4, with Panel A providing the results for the fall 2007 exam and Panel B providing the results for the spring 2008 exam. Starting with Panel A, the results demonstrate that the schools generated large changes in children's test scores. On average, the program caused an overall change in test scores of 0.51 standard deviations, a result that is statistically significant at the one percent level. The change in math scores was larger than the improvement in language scores (0.55 standard deviations versus 0.35 standard deviations respectively), but both are quite large. The relative pattern of results is consistent with those of other interventions, which generally find larger

---

[21] Formally, Afghanistan is usually placed in the region of South Asia. However, given that it has stronger cultural similarities to the Islamic countries in Western Asia than countries like India, Sri Lanka, and Nepal, it is natural to compare it to Western Asian countries. Some scholars do this by referring to the region as "Southwestern Asia".

effects of treatments on math scores than languages scores (see for example Banerjee, Cole, Duflo, and Linden, 2007). The results for the spring 2008 exam in Panel B show the same patterns. As in Table 4, the results do not change with the addition of controls, and the significance levels do not depend on the use of within-village group variation in scores.

These estimates in columns one through five are the overall average estimates for all children in the villages. However, as shown in Table 4, not all children in the treatment villages attend a formal school (either a village-based school or otherwise) and some of the control children attend formal schools. Both of these mean that the overall average treatment effect, or the intent-to-treat estimates, underestimates the actual change in test scores on children caused to attend school due to the treatment. To estimate the treatment effect on those children, we use an instrumental variables procedure similar to the one using equations (4) and (5), in which we use the relationship between formal school enrollment and the treatment assignment as a first stage for a regression of children's test scores on formal school enrollment (Imbens and Angrist, 1994). These estimated effects on the "compliers" are presented in column six. As expected, they are much larger than the average treatment effects. One year of formal school causes an increase in test scores of 1.2 standard deviations—an extremely large increase in scores. As with the raw difference, the effect for language (0.82 and 0.85 standard deviations) is smaller than the effect for math (1.3 and 1.4 standard deviations).

When evaluating an intervention, it also is important to judge the estimated treatment effects relative to the costs of implementation. One should note, however, that comparing this intervention to other educational programs is complicated by the fact that most interventions are implemented from an existing school, while in this intervention, we estimate the effects of the schools themselves. That said, we can ask how cost-effective establishing village-based schools

is as a strategy relative to interventions that may, for example, be implemented in the existing public schools in our sample. Using established methodology (Iqbal et al., 2011), we estimate the additional child-academic-years caused by the intervention per $100 to be 1.48. This is relatively much less cost-effective than de-worming interventions which can generate anywhere from 3-40 years per $100, but it is in the same range as girls scholarship programs at 1.4 student-years, school meals at 2.8 student-years, and subsidized school uniforms at one student-year (Iqbal et al., 2011). For test scores, the program costs $4.80 per child per tenth of a standard deviation. This is more cost-effective than most computer-assisted learning programs, but less cost-effective than cheaper interventions such as remedial education programs (see Banerjee et al., 2007 for examples of both).

B. Relative Effects for Girls

Given the strong overall impact of the program, the existing gender disparities, and the presumed relationship between gender and the effects of distance, we next assess the differential effects of the program on boys and girls. These results are presented in Table 6, where we replicate all of the preceding outcome estimates allowing for interaction with the gender of the child. All of the estimates demonstrate that the program has a dramatic effect on existing gender disparities. While boys benefit significantly from the program, the benefits accrue much more strongly to girls—so much so, that placing schools in the villages virtually eliminates the gender gap in enrollment and reduces the test score gap by a third after only one academic year of treatment.

Columns one and two of Table 6 replicate the estimates of the primary treatment effects on enrollment and test scores, respectively, from Tables 4 and 5. The first row provides the existing gender gaps in the control villages. Row four provides the treatment effect for boys and

row five provides the estimated difference in treatment effects for girls over boys. In both cases, the differences are large and statistically significant. Although the enrollment rate for boys increases by 34.4 percentage points, it increases by a further 16.7 percentage points for girls. Compared to the control group gap of 20.5 percentage points, the resulting gender disparity in the treatment villages is only four percentage points—a difference that is not statistically significant (p-value is 0.192). The estimates for test scores show a similar pattern, although probably because test scores are more of a stock-type variable, the treatment does not eliminate the gap. However, at 0.44 standard deviations, it is reduced by over a third within one year.

Column three replicates the estimated effects of the treatment on the compliers, those children who are caused to attend school by the treatment. The estimated effect of attending formal school on boys is provided in row eight and the differential effect for girls in row nine. The effect for boys is almost as large as the effect for all children reported in Table 5, and the differential effect for girls is comparatively smaller and statistically insignificant, suggesting that, of the compliers, girls benefit at the same level from school as boys do.

Finally, because these estimates depend on within-village-group variation, we cannot check the importance of the within-village-group variation by estimating the differences after aggregating to the village-group level as in the previous outcome tables. Instead we follow Cameron, Gelbach, and Miller (2008) and bootstrap the distribution of the test statistic for the test of the null hypothesis that the interaction between the treatment effect and gender is zero. We use the wild-cluster bootstrap with 10,000 iterations. The resulting significance levels are identical to the cluster-estimated ones based on the small sample t-distribution.

Why do girls benefit disproportionately from the intervention? And does this suggest a reason for the existing large gender disparities in enrollment? The explanations for the gender

disparity in primary education presented in Section I generally fall into three categories: differential returns to education either due perceived returns in the labor market or institutional features such as the fact that daughters ultimately leave their parents' households; differential costs such as the fact that it is considered less acceptable for girls to travel; and differential opportunity costs due to the fact that girls often supply household services that may be incompatible with school attendance, such as caring for younger siblings.

The results are most compatible with the differential costs models. First, the very fact that enrollment rates equalize in the treatment group suggests that at this education level, demand for school is the same for both genders as long as the school is placed in the village. Second, the fact that there is no differential treatment effect in test scores for the compliers belies the claim that girls may learn less than boys given existing investments in complements to school. Third, as we show in Section V.D below, once we properly estimate the relationship between distance and school, the effect of distance explains the entire gender gap in enrollment.

Fourth, at least through the primary grades, parents express a desire for both girls and boys to attend school. In qualitative interviews with members of the villages, the main explanation provided for not sending girls to school was security. The following quote from a member of the village council (or "shura") is succinctly indicative, "…about girls it's a little different, because the way is long, so there should be one or two people in the family to take the girls to school to the center and bring them back which is the main reason why they don't let them go. Second, because the younger children can't cope the long way, so they can't go." The issue is a combination of security and propriety. Generally, it would be considered inappropriate to allow a girl or woman to travel the distances required to attend school outside of the village without a male escort. Local culture places a premium on protecting women, and given the

existing state of lawlessness and local rivalries in such remote rural areas, unaccompanied women would be considered particularly vulnerable.

This practice is so uniform that there is simply no variation in households' willingness to allow their older girls to travel alone. Instead, we can first verify the age-gender interaction implied by this statement. Second, we can then directly test to see if a few alternative explanations account for the differential effect by gender.

Columns four and five of Table 6 present a triple interaction of age, gender, and the treatment effect to test this hypothesis. To facilitate the interpretation of the coefficients, we have rescaled the age variables in this regression relative to six years—so, children who are six years of age are measured as having an age of zero, seven year old children an age of one, etc. Rows one through three show the effects of the interactions in the control group. Starting with formal enrollment, the coefficients are consistent with the hypothesis. First, the coefficient on gender is effectively zero, suggesting that differences in school enrollment by age completely explain the observed gender gap in enrollment. In other words, at age six and independent of subsequent age, there is no constant component to the gender gap. Looking to rows two and three, we see that for each additional year a boy gains in age, he is 7.7 percentage points more likely to be enrolled. However, relative to boys, girls are 6.9 percentage points less likely to be enrolled with each additional year. This means that initially boys and girls are treated similarly, but as the children age, the gender gap emerges because boys are more likely to be allowed to go to school while girls are not.

Rows four through seven then provide estimates of the relationships in the treatment group relative to those in the control group. As the coefficient in row four indicates, the treatment causes a net increase in the probability of enrollment of 40.5 percentage points for all

children, independent of age and gender. Turning to the interaction effects, the coefficients on the interaction between treatment and gender and then treatment and age are both effectively zero. However, the coefficient on the triple interaction is 8.2 percentage points, a level that almost exactly offsets the -6.9 percentage point per year increment in the gender gap in the control group. Thus, in the treatment group, the treatment causes all children to attend school more frequently—even the youngest, consistent with the shura member's statement. However, the interaction between age and gender changes relative to the control group. Older children are still more likely to be enrolled in school, but in the treatment group, the degree to which older children are likely to attend school is the same for boys and girls. In short, families treat girls and boys the same with respect to primary school enrollment if a school is placed in the village.

Estimating the same dynamics with respect to test scores in column five is difficult because older children are simply more likely to score better on the test than younger children due to age. However, as the estimates clearly show, the same patterns exist in the data. Additionally, for girls, who have little school experience in both groups, we do find that the treatment (the triple interaction coefficient in row seven) still completely offsets the per-age gender gap (row three).

Finally, in results not presented in this manuscript, we consider other child and household characteristics as potential explanations for the observed change in the gender gap. First, we have explored the possibility that differences in household responsibilities might explain the gap. We first test whether children's engagement in a range of activities is affected by the treatment, either directly or differentially by gender, and find no effects. Since these characteristics do not change as a result of the treatment, we can then use them as explanatory variables, using a triple interaction between treatment, gender, and engagement in each activity. We find that, consistent

with the hypothesis that some activities may interfere with school enrollment, the treatment effects are larger for children who engage in several activities including sibling care, fetching water, and washing clothes. However, none of these activities explains the observed differential in the treatment effect by gender—the partial derivative with respect to gender and treatment remains stable. Thus, we conclude that families are willing to send their girls to primary school, but with the existing infrastructure, the main barriers to attendance are the risks associated with traveling the necessary distances.[22,23]

C. Quality

Some of the children in the treatment group who would have gone to traditional schools absent the treatment instead attend village-based schools, allowing us to assess whether the change of institution causes a difference in test scores. This allows us to understand better how the village-based schools function, and it allows us to answer one of the major policy concerns with these types of institutions—their relative quality compared to existing institutions.

The strategy hinges on there being children in the treatment group who attend the village-based schools but who would have otherwise attended a traditional public school. Table 7 demonstrates that such children exist. Columns two and three provide estimates of the probability that a child attends either a village-based school or a traditional government school

---

[22] Note that the caveat raised in Footnote 4 applies here. While the experimental design allows us to conclude that village-based schools are sufficient for equal participation, we cannot experimentally identify the underlying causes. The observed pattern of treatment effects rules out a number of alternative explanations and provides strong suggestive evidence that safety may be the primary issue. As a result, directly testing these underlying hypotheses is an important area for future work. If village schools are not a necessary condition for making such gains in school participation, then other more cost-effective strategies may be possible.
[23] It is important to note that these results are probably only true for primary education. Responses to some of the survey questions foreshadow an eventual disparity for older girls even in the treatment groups. Parents generally reported that the purpose of educating girls was so that they could better educate their own future children, while the purpose of boys' education was generally to provide a better livelihood for the family. Similarly, when asked for how long children should ideally be educated, parents were mostly likely to want girls to stop around 13-14 years of age and for boys to continue their education to 18 and often beyond.

respectively using equation (1), while column one replicates the estimated difference in overall

formal school enrollment for reference. The increase in the enrollment in village-based schools

is larger than the overall increase in formal school enrollment—55.9 percentage points compared

to 42.2 percentage points. This difference is due to the fact that children are 14.7 percentage

points less likely to attend a traditional government school—suggesting that some children did

indeed attend a village-based school who would have otherwise attended a traditional school.

Given these differences, we employ the strategy outlined in Section III.C, estimating a

model of non-enrollment from the control group and using it to estimate the probability that

children in the sample are not enrolled absent the treatment.[24] We then interact this with the

treatment effect using equation (3). The results are presented in Table 8. We present estimates

from the version of the model that includes the probability of not being enrolled as an

explanatory variable in Panel A and the model that includes the full set of control variables

instead in Panel B. Since the significance of each coefficient is determined using the wild-

cluster bootstrap procedure to estimate the distribution of the resulting t-statistics while

accounting for the estimated regressors and the small number of clusters, we provide the

individual p-values for a test of each coefficient individually in parentheses, rather than the

standard errors as in the other tables.

Starting with Panel A, we first disaggregate the changes in the probability of attending a

formal school. As expected, the small and statistically insignificant coefficient on the treatment

variable indicates that the entire difference in the change of formal school enrollment occurs in

---

[24] As explained in Section III.C, we use the linear probability model whose estimates are presented in column seven of Table 2. We chose this specification because it employs the same set of control variables as every other estimate, and this allows us to keep the specification of the control function constant across the various estimates. Like our other estimates, however, the results are relatively robust to changes in specification. The results are, for example, robust to the use of a probit model and to variations in the set of included variables. We have also checked specifications that include higher level polynomials (through a quartic) in the probability of not being enrolled. The higher order terms are never statistically significant.

children who would, with some probability, have not been enrolled absent the treatment. For those children who would have certainly not been enrolled, the increase is 58.2 percentage points. By construction, the probability of not being enrolled fits the control group perfectly.

Figure 3 provides a graphical description of these results. The lines with long dashes represent the results from the control group, while those with the short dashes represent the results from the treatment group. For both sets of estimates, the functions represented with dots and dashes are plots of the estimates presented in Table 8 while those with no dots are the result of non-parametric estimates using a local linear polynomial estimator with a bandwidth of 0.075. As the non-parametric plots show, the vast majority of children (97.2 percent) have an estimated probability of not being enrolled of between 35 and 100 percent. For a fixed probability of not being enrolled absent the treatment, the treatment effect guarantees that children in the treatment villages will be more likely to be enrolled than those in the control group. However, as the probability of not being enrolled decreases, the difference should shrink if those children who would have been enrolled in the treatment group without the treatment experience no change in behavior or performance. The test simply asks whether or not the rate of decline of the treatment effect is consistent with a prediction of no difference between those children who would definitely attended a formal school absent the treatment. In the models presented in Table 7, this is represented by a test of whether the coefficients on the treatment variables are equal to zero, and graphically, we ask whether or not the estimates from the treatment and control groups would intersect at the point where the probability of not being enrolled is zero, the y-axis. This is clearly the case for overall formal school enrollment.

Enrollment in traditional schools and enrollment in village-based schools, presented in columns two and three respectively, show how the estimates diverge. Starting with column two

in rows three and four, the estimated probability of not being enrolled again almost perfectly fits the control group. However, in the treatment group, the estimates indicate that there is a 37 percentage point difference in the probability that a child who would have certainly been enrolled in a traditional school in the control group would be enrolled in the same type of school in the treatment group. When we look at the probability of being enrolled in a village-based school, we see an increase in the probability of almost the same magnitude, which is consistent with the absence of a net difference in column one. Graphically, the estimates for traditional school enrollment are presented in Figure 4. Consistent with the regression results, there is a large gap in the points at which the plots are projected to intersect the y-axis.

Finally, in column four we present the results using our preferred measure of child performance—the fall 2007 normalized test scores. The results suggest that all of the treatment effect in test scores is indeed accruing to children who would not otherwise be enrolled in schools absent the treatment. The coefficient on the treatment variable is small. The results are depicted graphically in Figure 5. The linear estimates fit the non-parametric plots reasonably well and are consistent with the results in column four. As shown in column five, the results are similar if we use the test administered in the spring of 2008 as well.[25]

To check the robustness of these results, we also estimate the specifications using the full set of control variables rather than including the estimated probability of not being enrolled. With the exception of the coefficient on the treatment variable for traditional school enrollment (column two), all of the same variables are statistically significant, and more importantly, the values of the coefficients are effectively the same.

---

[25] Because the predicted probabilities of formal school enrollment are based on the correlations in the fall 2007 survey, we use only those 1,181 students surveyed in both the fall and spring surveys to estimate the test score differences in the spring of 2008.

It is important to note a major limitation of this approach. We can only test the difference for those children who switch between traditional and village-based formal schools. It is possible that those children who switch, do so because the differences are not significant enough to affect their learning. We cannot rule out the possibility that those children who remain at the traditional schools in the treatment group would suffer academically if forced to switch, nor can we rule out the possibility that those children who would not have attended school at all absent the treatment might be better served by facilitating their enrollment in traditional schools rather than village-based schools. However, to the extent that we can measure, the schools are of similar quality.

D. The Effect of Distance

We estimate the effects of distance using two complementary strategies. First, if traditional and village-based schools offer similar educational experiences as the previous section suggests, then the main difference between them is the distance that children have to travel to attend them. By placing schools in the treatment villages, the treatment generates an exogenous change in the proximity of treatment households to the nearest school, and we can therefore estimate the causal relationship between enrollment, test scores, and distance using instrumental variables as described in equations (4) and (5).[26] This strategy has the advantage of using variation in distance from all of the schools in our sample. However, if, as feared, the village-based schools had been of lower quality than the traditional schools, the net effect would cause us to under-estimate the importance of distance, and our estimates would serve as a lower bound on the

---

[26] We use the straight-line distance, ignoring topography, based on GPS coordinates for all households and schools.

relationship.[27]   Additionally, since the instrument is an indicator variable, we are limited to estimating a linear relationship.

As a check on these estimates, we also estimate the relationship between enrollment and distance to the nearest village-based school. Since the placement of these schools was deliberately random and the schools are similar in structure, the OLS estimate should be unbiased and will allow for the estimation of non-linear relationships. This strategy, however, ignores all of the other schools in our sample, but it does provide a check for the previous strategy. If there are quality differences that reduce the elasticity of enrollment in distance, then the two sets of estimates will differ.

We present our estimates in Table 9. Column one presents the estimated average change in distance due to the treatment (equation (4))—a reduction of 2.70 miles—that is graphically depicted in Figure 2. Columns two and three provide estimates of distance on enrollment using the IV and OLS strategies respectively. Columns four and five then provide the same estimates for test scores. We find that the effect of distance on enrollment is to reduce it by 16.3 percentage points per mile in the preferred specification, an effect that is statistically significant at the one-percent level. The effect on test scores is a reduction of 0.19 standard deviations per mile, also statistically significant at the one-percent level in the preferred specification. The estimates using just the village-based schools are very close. For enrollment, the relationship is non-linear, but the linear model provides a good approximation in range of distances under consideration.

---

[27] As shown in Figure 2, the village-based schools are always the closest schools, and the traditional schools are the farthest schools. If the village-based schools are of lower quality, families living close to a school would be less likely to send their children than if the schools were of similar quality. The net effect is that enrollment rates for families who are close to a school would reflect both the closer distance and the lower quality. The estimated enrollment levels for close schools would be too low, and our estimates would underestimate the effect of distance on enrollment.

Figure 6 depicts the relationships graphically, including also non-parametric estimates of the relationships. While there is some variation, they all show a large reduction in enrollment between zero and four miles. The non-parametric estimates seem to suggest that enrollment is stable between zero and 1.5 miles, suggesting a possible village-boundary effect, but the data is insufficient to test for such detailed differences.

Table 10 presents the same estimates by gender. As before, columns one and two provide the estimates for enrollment while columns three and four provide the effects on test scores for the IV and OLS models respectively. All of the estimates show that the effect of distance is greater for girls. In the IV specification, the effect on boys is 13.2 percentage points per mile while girls' enrollment falls by an additional 5.9 percentage points. For tests scores, the effect for boys is 0.15 standard deviations per mile with girls' scores falling by an additional 0.09 standard deviations. Both differences are statistically significant at the ten-percent level. The estimates using only the village-based schools show similar large, statistically significant effects, although the estimated effects for boys are a smaller and the effects for girls are larger. This is true for both outcomes, but largest for test scores. Additionally, in both of the enrollment regressions, the coefficient on female falls to zero, suggesting that once properly estimated, the relationship between distance and enrollment explains the observed gender gap.[28]

---

[28] While not shown in the table, the estimated gender gap in enrollment from the IV and OLS specifications in Table 9 are 11.6 and 11.2 percentage points respectively, both statistically significant at the five-percent level.

VI. Conclusion

We show that a program designed to place formal schools within villages has a significant effect on children's school participation and learning, and significantly reduces the existing gender disparities in educational outcomes. Overall, these results document that village-based schools should be a considered an important tool for improving primary education in rural areas.

The program increases enrollment in formal schools by 42 percentage points. Average test scores among all children increase by 0.51 standard deviations, with improvements in both math and language skills. The scores of the children that actually enroll in school as a result of the program improve by 1.2 standard deviations. The benefits, however, accrue disproportionately to girls. Placing a village-based school in a village increases girls' formal school enrollment by 17 percentage points more than boys'. On average, the test scores of all girls in the village also increase by 0.25 standard deviations more than boys, though this is primarily due to higher enrollment rates. Girls caused to enroll as a result of the program experience the same increase in test scores as boys. The net effect of these differences is that placing a school in each village dramatically reduces the existing gender disparities. The village-based schools virtually eliminate the enrollment gap between boys and girls, leaving a difference of only 4 percentage points, compared to a deficit of 21 percentage points in the control villages. The test score gap also closes significantly—the difference in treatment villages is over one-third less than in control villages after only a year of treatment. The limited available evidence suggests that the village-based schools deliver similar experiences to those of traditional public schools, despite fears of their relative lower quality, and given their differences in structure.

Estimates of the marginal effects of distance on enrollment and test scores suggest that, unlike in existing cross-sectional studies, children are very sensitive to changes in distance. On average, enrollment and test scores increase by 16 percentage points and 0.19 standard

deviations respectively per mile. Girls also respond much more strongly than boys to changes in distance. Girls' enrollment increases by 19 percentage points per mile, compared to 13 for boys, and girls' test scores increase by 0.24 standard deviations per mile, 0.09 standard deviations more than boys'.

VII. Bibliography

Adele, M. E. (2008) "Afghanistan on the Educational Road to Access and Equity," *Asia Pacific Journal of Education.* 28(3): 277-290.

Alcazar, L., Rogers, F.H., Chaudhury, N., Hammer, J., Kremer, M., and Muralidharan, K. (2006) Why are teachers absent? Probing service delivery in Peruvian primary schools. *International Journal of Educational Research*. 45, p. 117-136.

Al-Qudsi, S. S. (2003) "Family Background, School Enrollments and Wastage: Evidence from Arab Countries," *Economics of Education Review.* 22: 567-580.

Andrabi, Tahir, Jisnu Das, and Asim Khwaja (2010) "Students Today, Teachers Tomorrow? Identifying Constraints on the Provision of Education," Harvard Kennedy School of Government Working Paper.

Angrist, Joshua and Jorn-Steffen Pischke (2008) *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press: Princeton, NJ.

Banerjee, Abhijit, Shawn Cole, Esther Duflo, and Leigh Linden (2007) "Remedying Education: Evidence from Two Randomized Experiments in India," *Quarterly Journal of Economics.* 122(3): 1235-1264.

Barrera-Osorio, Felipe, Marianne Bertrand, Leigh L. Linden, and Francisco Perez-Calle (2010) "Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia," *American Economic Journal: Applied Economics.* 3(2): 167-95.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004) "How Much Should We Trust Difference in Differences Estimates?" *Quarterly Journal of Economics.* 119(1): 249-275.

Burde, Dana (2008) "Protecting Children from War and Ensuring their Prospects for the Future, Pilot Study Preliminary Findings," Report to the Spencer Foundation. New York: Steinhardt School of Culture, Education, and Human Development.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2008) "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*. 90(3): 414-427.

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.

Filmer, Deon (2007) "If You Build It, Will They Come? School Availability and School Enrolment in 21 Poor Countries," *Journal of Development Studies*. 43(5): 901-928.

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

Guyot, Wendy (2007) *Transition from Community Based Education to the Ministry of Education System in Afghanistan: An Investigation of the Integration Process for Students and Teachers in Selected Districts of the PACE-A Partnership.* Kabul, Afghanistan: PACE-A Publication.

Handa, S. (2002) "Raising Primary School Enrolment in Developing Countries: The Relative Importance of Supply and Demand," *Journal of Development Economics*. 69(1): 103-128.

He, Fang, Leigh Linden, and Margaret MacLeod (2008) "How to Teach English in India: Testing the Relative Productivity of Instruction Methods within the Pratham English Language Education Program," Mimeo. Columbia University, Department of Economics.

Imbens, Guido and Joshua Angrist (1994) "Identification and Estimation of Local Average Treatment Effects," *Econometrica*. 62(2): 467-476.

International Child Labor Program (ICLP) (2000) "Chapter IV: Access to Primary Education," *By the Sweat & Toil of Children: An Economic Consideration of Child Labor, Volume VI*. United States Department of Labor, Bureau of International Labor Affairs.

Iqbal, Dhaliwal, Esther Duflo, Rachel Glennerster, and Caitlin Tulloch (2011) "Comparative Cost-Effectiveness Analysis to Inform Policy in Developing Countries: A General Framework with Applications for Education" Working Paper. Abdul Latif Jameel Povery Action Lab, MIT.

Jafarey, Saqib and Sajal Lahiri (2005) "Food for Education Versus School Quality: A Comparison of Policy Options to Reduce Child Labor," *The Canadian Journal of Economics*. 38(2): 394-419.

Kazianga, Harounan, Daniel Levy, Leigh Linden, Matthew Sloan (2010) "The Effect of Girl-Friendly School Construction: Evaluation of the BRIGHT School Program in Burkina Faso," Columbia University Working Paper.

Kim, Jooseop, Harold Alderman, and Peter F. Orazem (1999) "Can Private School Subsidies Increase Enrollment for the Poor? The Quetta Urban Fellowship Program," *The World Bank Economic Review*. 13(3): 443-465.

Kristiansen, Stein and Pratiknob (2006) "Decentralizing Education in Indonesia," *International Journal of Educational Development*. 26: 513-531.

Lavy, Victor (1996) "School Supply Constraints and Children's Educational Outcomes in Rural Ghana," *Journal of Development Economics*. 51(2): 291-314.

Lloyd, Cynthia, 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.

Miguel, Edward and Michael Kremer (2004) "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*. 72(1): 159-217.

Ministry of Education, Islamic Republic of Afghanistan (MoE) (2007). Schools Survey Summary Report. EMIS Department.

Ministry of Rural Rehabilitation and Development, Islamic Republic of Afghanistan (MRRD) (2007). Ghor Provincial Profile. Retrieved 2/13/09 from http://www.mrrd.gov.af/NABDP/Provincial%20Profiles/Ghor%20PDP%20Provincial%20profile.pdf.

Muralidharan, Karthik and Venkatesh Sundararaman (2008) "Teacher Performance Pay: Experiment Evidence from India," Mimeo, Department of Economics, University of California at San Diego.

Pitt, M., Mark Rosenzweig, and D. Gibbons (1993) "Determinants and Consequences of the Placement of Government Programs in Indonesia," *World Bank Economic Review*. 16(2): 297-319.

Rosenzweig, Mark and Kenneth Wolpin (1986) "Evaluating the Effects of Optimally Distributed Public Programs: Child Health and Family Planning Interventions," *American Economic Review*. 76(3): 470-482.

Schultz, T. P. (2004) "School subsidies for the poor: Evaluating the Mexican PROGRESA poverty program," Journal of Development Economics, 74(1), 199-250.

Sutton, Margaret (1998) "Girls Educational Access and Attainment" In *Women in the Third World: And Encyclopedia of Contemporary Issues*. Ed. Nelly P. Stromquist. 381-396. New York: Garland Publishing.

Thornton, Rebecca L. (2008) "The Demand for, and Impact of, Learning HIV Status," *American Economic Review*. 98(5): 1829-1863.

UNDP (2007). Afghanistan Human Development Report, 2007. Center for Policy and Human Development, Kabul University.

United Nations (2008a) *Fact Sheet: Goal 2 Achieve Universal Primary Education*. United Nations Department of Public Information, Publication Number DPI/2517 H, September.

United Nations (2008b) *Fact Sheet: Goal 3 Promote Gender Equality and Empower Women,* United Nations Department of Public Information. Publication Number DPI/2517I, September.

United Nations (2008c). *The Millennium Development Goals Report 2008*. United Nations Department of Public Information.

United Nations (2009). Afghanistan Humanitarian Action Plan. United Nations.

Wilcoxon, Frank (1945) "Individual Comparisons by Ranking Methods," *Biometrics Bulletin.* 1(6): 80-83.

Figure 1: Distribution of Distance from Households to Nearest Non-Village Formal School
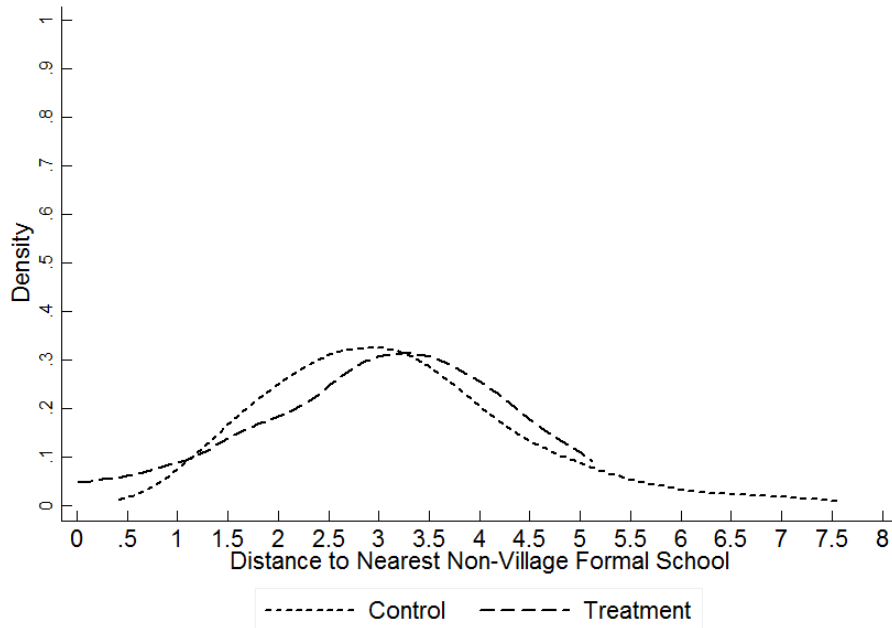
Note: This figure presents the distribution of the distances between households and the nearest non-village-based formal school. This is the distribution of distances that would have existed without the implementation of the treatment. All distributions are estimated using a kernel density estimator with a bandwidth of 0.75 miles and an Epanechnikov kernel.

Figure 2: Distribution of Distance between Household and Nearest Formal School
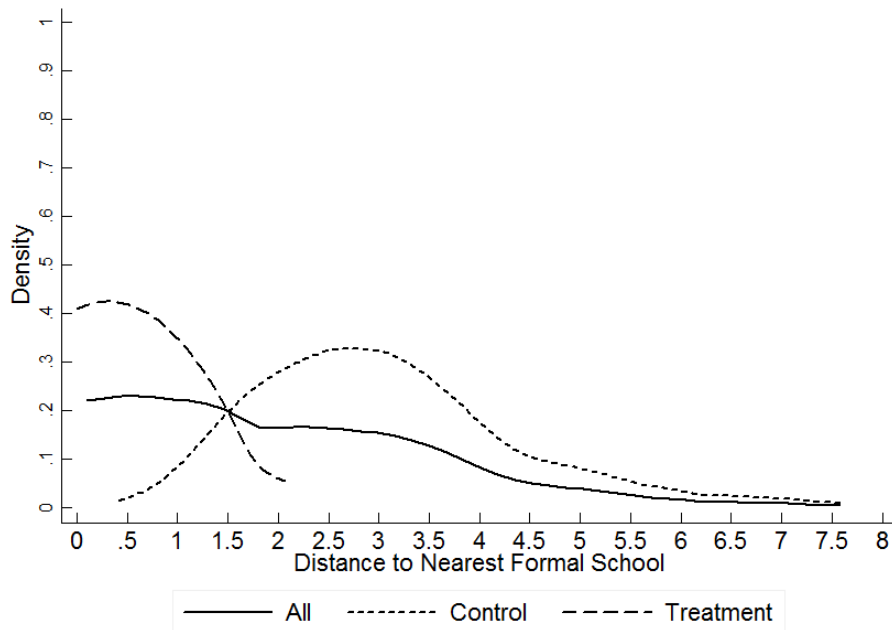
Note: This figure presents the distribution of the distances between households and the nearest formal school. This is the distribution of distances that exists after implementation of the treatment. All distributions are estimated using a kernel density estimator with a bandwidth of 0.75 miles and an Epanechnikov kernel.

Figure 3: Formal School Enrollment by Estimated Probability of Not Being Enrolled
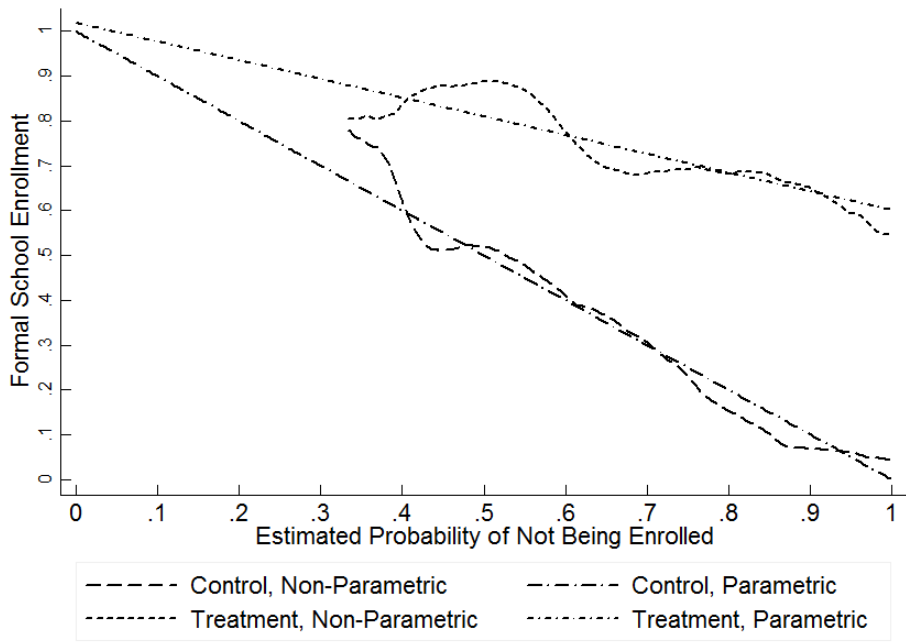
Note: This figure presents estimates of formal school enrollment as a function of the estimated probability of not being enrolled in a formal school absent the treatment. Parametric estimates are plotted using estimates from the regression presented in column one of Panel A of Table 7. The non-parametric estimates are generated using a local linear polynomial estimator with a bandwidth of 0.075 and an Epanechnikov kernel. Estimated probabilities of not being enrolled are projected estimates created from a model relating formal school enrollment in the control group to the available socio-demographic characteristics from Table 2 using a linear probability model.

Figure 4: Traditional School Enrollment by Estimated Probability of Not Being Enrolled
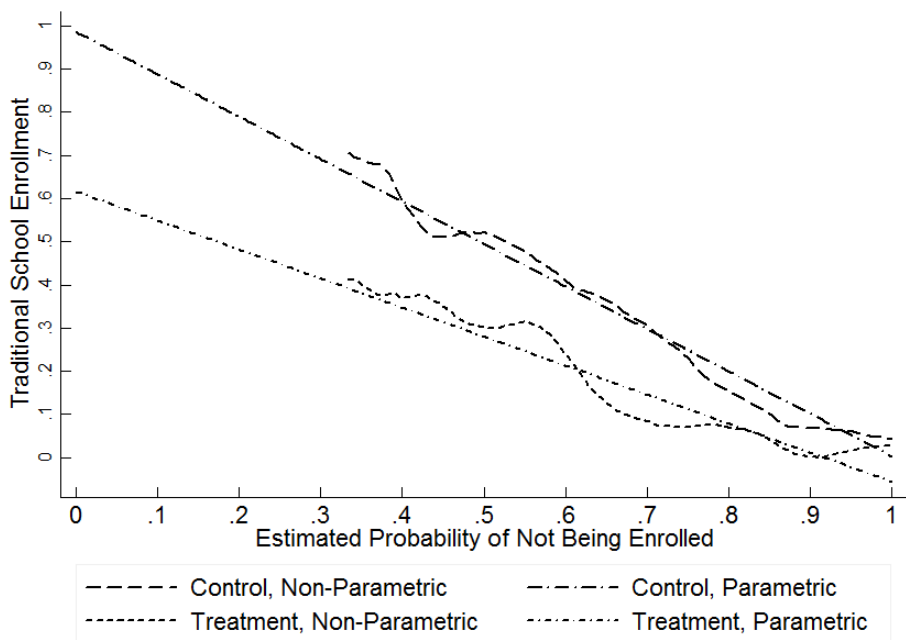
Note: This figure presents estimates of traditional school enrollment as a function of the estimated probability of not being enrolled in a formal school absent the treatment. Parametric estimates are plotted using estimates from the regression presented in column two of Panel A of Table 7. The non-parametric estimates are generated using a local linear polynomial estimator with a bandwidth of 0.075 and an Epanechnikov kernel. Estimated probabilities of not being enrolled are projected estimates created from a model relating formal school enrollment in the control group to the available socio-demographic characteristics from Table 2 using a linear probability model.

Figure 5: Total Score, Fall 2007, by Estimated Probability of Not Being Enrolled
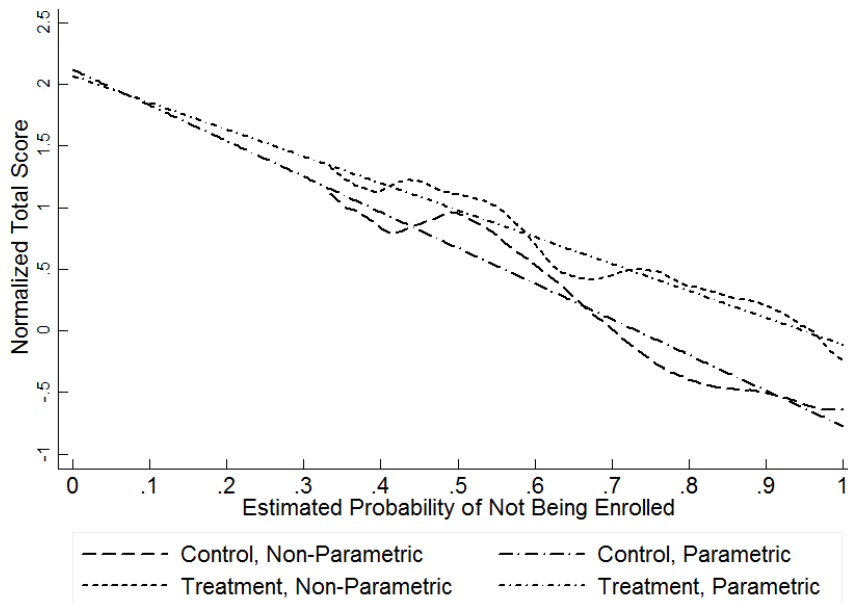
Note: This figure presents estimates of total normalize score on the fall 2007 survey as a function of the estimated probability of not being enrolled in a formal school absent the treatment. Parametric estimates are plotted using estimates from the regression presented in column four of Panel A of Table 7. The non-parametric estimates are generated using a local linear polynomial estimator with a bandwidth of 0.075 and an Epanechnikov kernel. Estimated probabilities of not being enrolled are projected estimates created from a model relating formal school enrollment in the control group to the available socio-demographic characteristics from Table 2 using a linear probability model.

Figure 6: Formal School Enrollment as a Function of Distance to Nearest School
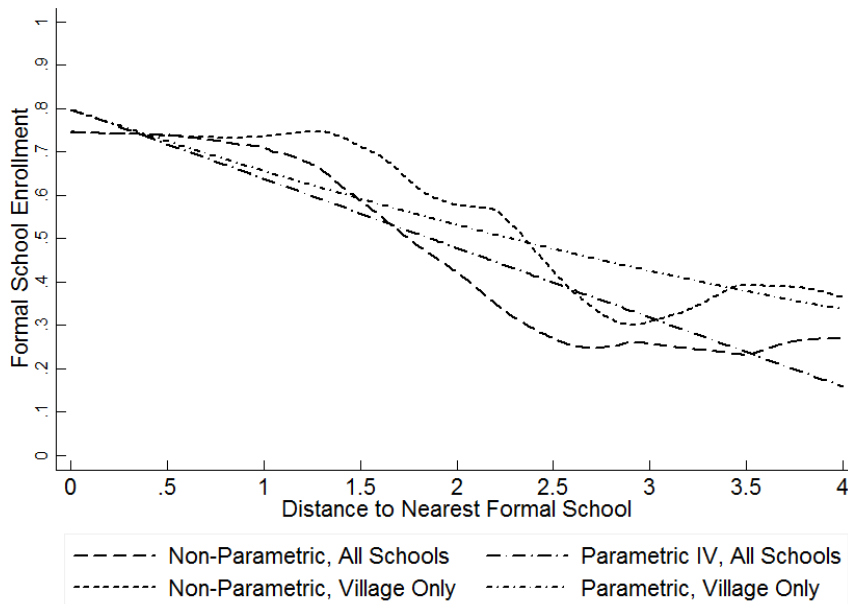
Note: This figure presents four estimates of the relationship between enrollment and the distance to the nearest school. First, the parametric instrumental variables estimate is plotted using the estimates from column two of Table 9 (Parametric IV, All Schools). The non-parametric estimates are constructed using a local linear polynomial estimator with a bandwidth of 0.75 miles and an Epanechnikov kernel. We present both an estimate using all schools (Non-Parametric, All Schools) and one using only the village-based schools (Non-Parametric, Village Only). Finally, we present the parametric estimates of the relationship between enrollment and distance to the nearest village-based school presented in bcolumn three of Table 9 (Parametric, Village Only).

Table 1: Sample Size and Coverage Rates by Research Group

	Fall 2007 Survey				Spring 2008 Survey			
	Treatment Group	Control Group	Estimated Difference	Total	Treatment Group	Control Group	Estimated Difference	Total
Panel A: Households Surveyed								
Identified	680	663	17	1,343	637	616	21	1,253
Surveyed	635	628	7	1,263	603	582	21	1,185
Percent of Households Surveyed	0.934	0.947	-0.013 (0.025)	0.94	0.947	0.945	0.002 (0.014)	0.946
Panel B: Households with Eligible Children								
Households with Children	414	391	23	805	399	395	4	794
Percentage with Children	0.65	0.618	0.033 (0.037)	0.634	0.662	0.679	-0.017 (0.026)	0.67
Panel C: Children Tested								
Identified	782	708	74	1,490	756	721	35	1,477
Tested	721	653	68	1,374	722	679	43	1,401
Percent of Children Tested	0.922	0.922	< 0.001 (0.020)	0.922	0.955	0.942	0.013 (0.012)	0.949

Note: This table contains the tabulation of the sample used for the study, divided by survey round and research group. The differences are estimated using equation (1) without controls and with standard errors clustered at the village-group level. Statistical significance at the one-, five-, and ten-percent levels is indicated by ***, **, and * respectively and evaluated relative to the small sample t-distribution to account for the small number of clusters.

Table 2: Demographic Characteristics by Research Group

	All Children			Tested Children			Control Correlations	
	Treatment Average	Control Average	Estimated Difference	Treatment Average	Control Average	Estimated Difference	Formal Enrollment	Total Score
Panel A: Child Level Variables								
Head of Household's Child	0.935	0.911	0.024	0.939	0.917	0.022	0.038	-0.083
			(0.015)			(0.017)	(0.061)	(0.110)
Girl	0.474	0.455	0.02	0.495	0.475	0.02	-0.209**	-0.687***
			(0.020)			(0.021)	(0.079)	(0.100)
Age	8.321	8.312	0.009	8.323	8.303	0.02	0.046**	0.288***
			(0.040)			(0.051)	(0.017)	(0.019)
Panel B: Household Level Variables								
Years Family in Village	30.302	27.594	2.709	30.239	27.852	2.387	-0.001	-0.005**
			(1.605)			(1.626)	(0.002)	(0.002)
Farsi	0.208	0.209	-0.001	0.209	0.202	0.007	-0.032	0.087
			(0.054)			(0.057)	(0.077)	(0.106)
Tajik	0.243	0.208	0.035	0.245	0.214	0.031	0.02	0.08
			(0.049)			(0.052)	(0.064)	(0.066)
Famers	0.717	0.727	-0.01	0.709	0.721	-0.013	-0.061	-0.033
			(0.034)			(0.033)	(0.074)	(0.122)
Age of Household Head	40.142	39.97	0.172	40.268	39.839	0.428	-0.004	-0.004
			(1.101)			(1.045)	(0.002)	(0.002)
Years of Ed of Household Head	3.315	3.076	0.239	3.296	3.085	0.211	0.002	0.039***
			(0.442)			(0.446)	(0.007)	(0.006)
Number of People in Household	8.399	7.818	0.581	8.462	7.779	0.682*	0.001	-0.008
			(0.340)			(0.329)	(0.004)	(0.011)
Jeribs of Land	1.345	1.274	0.071	1.345	1.239	0.106	0.024*	0.059***
			(0.107)			(0.116)	(0.011)	(0.012)
Number of Sheep	7.552	5.631	1.921	7.408	5.755	1.653	0.009	0.014*
			(1.504)			(1.486)	(0.004)	(0.006)
Distance to Nearest Formal School (Non-Village-Based School)	2.91	3.163	-0.253	2.923	3.161	-0.238	-0.049	-0.074*
			(0.349)			(0.355)	(0.033)	(0.035)

Note: This table contains average demographic characteristics from the fall 2007 survey, divided by research group. The first three columns include all 1,490 identified children while the second three columns include only the 1,374 children that were directly tested. All differences are estimated using equation (1) without controls and with standard errors clustered at the village-group level. Estimates in columns seven and eight are the coefficients of a regression of the respective outcome variable on the controls using only data from the control group, including data from 708 and 653 children respectively. Statistical significance at the one-, five-, and ten-percent levels is indicated by ***, **, and * respectively and evaluated relative to the small sample t-distribution to account for the small number of clusters.

Table 3: Attrition Patterns by Research Group

	Non-Attritors			Attritors less Non-Attritors		
	Treatment Average	Control Average	Estimated Difference	Treatment Difference	Control Difference	Difference in Difference
Panel A: Attrition Rates	0.174	0.162	0.011			
	(0.014)	(0.014)	(0.033)			
Panel B: Child Characteristics						
Head of Household's Child	0.935	0.919	0.016	-0.001	-0.049	0.048
	(0.020)		(0.023)	(0.029)	(0.042)	
Girl	0.481	0.459	0.023	-0.04	-0.024	-0.016
	(0.022)		(0.047)	(0.051)	(0.055)	
Age	8.229	8.275	-0.046	0.528	0.229	0.299
	(0.066)		(0.155)	(0.167)	(0.220)	
Panel B: Household Characteristics						
Years Family in Village	31.224	28.028	3.197*	-5.302	-2.671	-2.63
	(1.635)		(1.452)	(1.593)	(2.191)	
Farsi	0.209	0.204	0.005	-0.003	0.031	-0.034
	(0.055)		(0.038)	(0.041)	(0.051)	
Tajik	0.252	0.216	0.036	-0.054	-0.051	-0.003
	(0.054)		(0.040)	(0.041)	(0.098)	
Famers	0.723	0.722	0.001	-0.032	0.035	-0.067
	(0.033)		(0.043)	(0.045)	(0.089)	
Age of Household Head	40.382	39.791	0.591	-1.382	1.105	-2.487
	(1.060)		(1.055)	(1.162)	(1.668)	
Years of Ed of Household Head	3.379	3.084	0.295	-0.372	-0.054	-0.318
	(0.451)		(0.333)	(0.358)	(0.569)	
Number of People in Household	8.497	7.862	0.635	-0.563	-0.27	-0.293
	(0.377)		(0.275)	(0.261)	(0.554)	
Jeribs of Land	1.3	1.264	0.036	0.259	0.062	0.197
	(0.116)		(0.147)	(0.166)	(0.255)	
Number of Sheep	7.599	5.909	1.69	-0.268	-1.709	1.441
	(1.584)		(0.763)	(0.710)	(0.837)	
Distance to Nearest Formal School	2.955	3.137	-0.182	-0.258	0.161	-0.418
(Non-Village-Based School School)			(0.325)	(0.107)	(0.111)	(0.284)

Note: This table contains average demographic characteristics of families from the fall 2007 survey divided by research group for children attriting and non-attriting at the spring 2008 survey. The first three columns report the average characteristics of non-attriting children from the fall 2007 survey. The differences in column three are estimated using equation (1) without controls and with standard errors clustered at the village-group level. The second three columns compare the average characteristics of attriting and non-attriting children. Columns four and five contain the average differences between attriting and non-attring children in the treatment and control group respectively. The difference in the difference between research groups is provided in column six. Statistical significance at the one-, five-, and ten-percent levels is indicated by ***, **, and * respectively and evaluated relative to the small sample t-distribution to account for the small number of clusters.

Table 4: Estimated Average Treatment Effects, Formal School Enrollment

	Treatment Average	Control Average	Difference	Difference w/ Controls	Difference Group Avg
Panel A: All Children					
Formal School, Fall 2007	0.735	0.268	0.467***	0.422***	0.421***
			(0.085)	(0.082)	(0.109)
Formal School, Winter 2008	0.388	0.007	0.381***	0.399***	0.327***
			(0.040)	(0.029)	(0.055)
Days Attended per 6-Day Week	5.922	5.831	0.091*	0.073**	0.049
Conditional on Enrollment			(0.041)	(0.028)	(0.038)
Panel B: Tested Children					
Formal School, Fall 2007	0.739	0.27	0.470***	0.425***	0.429***
			(0.084)	(0.079)	(0.109)
Formal School, Winter 2008	0.384	0.007	0.376***	0.396***	0.370***
			(0.040)	(0.028)	(0.045)
Days Attended per 6-Day Week	5.92	5.831	0.089*	0.070*	0.057
Conditional on Enrollment			(0.045)	(0.032)	(0.035)

Note: This table contains estimates of the effect of the treatment on children's participation in formal schools. The first two columns contain the average enrollment rates in the treatment and control groups. The second two columns contain the estimated differences using equation (1) with and without socio-demographic controls respectively clustering at the village-group level. Column five contains the estimated differences after aggregating the observations to the village-group level. Statistical significance at the one-, five-, and ten-percent levels is indicated by ***, **, and * respectively and is evaluated relative to the small sample t-distribution to account for the small number of clusters.

Table 5: Effects of Treatments on Test Scores

	Treatment Average	Control Average	Difference	Difference w/ Controls	Difference Group Avg	Compliers w/ Controls
Panel A: Fall 2007 Survey						
Total Score	0.58	-0.007	0.587***	0.511***	0.588***	1.205***
			(0.130)	(0.092)	(0.146)	(0.245)
Math Score	0.614	-0.005	0.620***	0.554***	0.631***	1.306***
			(0.124)	(0.100)	(0.144)	(0.224)
Dari Score	0.418	-0.008	0.426**	0.349***	0.411**	0.823**
			(0.136)	(0.093)	(0.145)	(0.272)
Panel B: Spring 2008 Survey						
Total Score	0.598	0.003	0.596***	0.527***	0.655***	1.187***
			(0.121)	(0.082)	(0.182)	(0.137)
Math Score	0.671	0.004	0.667***	0.612***	0.738***	1.379***
			(0.111)	(0.088)	(0.191)	(0.113)
Dari Score	0.456	0.001	0.455***	0.376***	0.495***	0.848***
			(0.135)	(0.077)	(0.161)	(0.162)

Note: This table contains estimates of the effect of the treatment on children's test scores. The first two columns contain the average test scores in the treatment and control groups. The second two columns contain the estimated differences using equation (1) with and without socio-demographic controls respectively clustering at the village-group level. Column five contains the estimated differences after aggregating the data to the village-group level with the significance levels calculated from the Fisher Exact P-values of the Wilcoxen Rank Test. Finally, the estimated effect of attending a formal school on the compliers is presented in column six. Statistical significance at the one-, five-, and ten-percent levels is indicated by ***, **, and * respectively and is evaluated relative to the small sample t-distribution to account for the small number of clusters.

Table 6: Treatment Effects by Gender

Dependent Variable	Formal Enrollment (1)	Total Score (2)	Total Score (3)	Formal Enrollment (4)	Total Score (5)
Girl	-0.205**	-0.684***	-0.474***	-0.045	-0.153
	(0.073)	(0.094)	(0.076)	(0.029)	(0.092)
Age				0.077**	0.392***
				(0.025)	(0.013)
Girl * Age				-0.069**	-0.231***
				(0.030)	(0.016)
Treatment	0.344***	0.390***		0.405***	0.496***
	(0.093)	(0.092)		(0.123)	(0.153)
Treatment * Girl	0.167**	0.248**		-0.023	-0.208
	(0.074)	(0.108)		(0.054)	(0.140)
Treatment * Age				-0.027	-0.047
				(0.034)	(0.034)
Treatment * Girl * Age				0.082**	0.196***
				(0.034)	(0.056)
Formal School			1.132***		
			(0.285)		
Formal School * Girl			0.126		
			(0.140)		
Observations	1,490	1,374	1,374	1,490	1,374
R^2	0.30	0.40	0.47	0.30	0.42
Socio-Demographic Controls	Yes	Yes	Yes	Yes	Yes
Model	OLS	OLS	IV	OLS	OLS

Note: This table contains estimates of the effect of the program by gender. Columns one and two show the net disaggregated treatment effect on formal school enrollment and the fall 2007 test using equation (1). Column three contains the effect of attending a formal school on the compliers. Finally, columns four and five contain estimates of the effect of the treatment disaggregated by age and gender. Note that for ease of interpretation, age is rescaled relative to the age of the youngest students in the data set. So, children who are six years of age are measured as having an age of zero. All standard errors are clustered at the village-group level. Statistical significance at the one-, five-, and ten-percent levels is indicated by ***, **, and * respectively and evaluated relative to the small sample t-distribution to account for the small number of clusters.

Table 7: Disaggregated Changes in Enrollment and Child Characteristics by Institution Type

	Formal School (1)	Village School (2)	Traditional School (3)
Treatment	0.422***	0.559***	-0.147*
	(0.082)	(0.027)	(0.080)
Observations	1,490	1,490	1,490
R^2	0.29	0.40	0.17
Socio-Demographic Controls	Yes	Yes	Yes

Note: This table contains estimates of the effect of the treatment on children's participation in individual types of schools. All coefficients are estimated using equation (1). All standard errors are clustered at the village-group level. Statistical significance at the one-, five-, and ten-percent levels is indicated by ***, **, and * respectively and evaluated relative to the small sample t-distribution to account for the small number of clusters.

Table 8: Treatment Effects Disaggregated by Predicted Non-Enrollment

	Formal School (1)	Traditional School (2)	Village School (3)	Total Score Fall 2007 (4)	Total Score Spr 2008 (5)
Panel A: Controlling for Probability of Being Unenrolled					
Treatment	0.019	-0.37*	0.367**	-0.05	-0.051
	(0.670)	(0.091)	(0.012)	(0.654)	(0.203)
Treatment *	0.582**	0.311	0.291**	0.712*	0.880***
Probability of Not Being Enrolled	(0.038)	(0.214)	(0.019)	(0.060)	(0.007)
Probability of Not Being Enrolled	-1.000***	-0.983***	< 0.001***	-2.894***	-2.993***
	(0.002)	(0.003)	(< 0.001)	(0.001)	(< 0.001)
Constant	1.000***	0.986***	< 0.001***	2.119***	2.204***
	(0.003)	(0.003)	(< 0.001)	(0.003)	(< 0.001)
Observations	1,490	1,490	1,490	1,374	1,181
R^2	0.29	0.15	0.39	0.26	0.27
Panel B: Controlling for Socio-Demographic Characteristics					
Treatment	0.028	-0.347	0.347**	0.017	0.07
	(0.727)	(0.152)	(0.062)	(0.777)	(0.409)
Treatment *	0.561**	0.284	0.302**	0.701***	0.775**
Probability of Not Being Enrolled	(0.041)	(0.242)	(0.039)	(0.007)	(0.012)
Observations	1,490	1,490	1,490	1,374	1,181
R^2	0.30	0.18	0.41	0.40	0.38

Note: This table contains estimates of the effect of the treatment interacted with the probability that a child would not have attended a formal school absent the treatment. Estimates are made using a two-step process. First, the probability of non-enrollment is estimated by regressing formal school enrollment on socio-demographic characteristics using the control group. Second, this estimated regressor is interacted with the treatment variable and estimated using equation (3). P-values for the significance tests of individual coefficients using the distribution of bootstrapped t-statistics from 10,000 samples are provided in parentheses. Statistical significance at the one-, five-, and ten-percent levels is indicated by ***, **, and * respectively.

Table 9: Effect of Distance to Nearest School

	Distance to School	Formal Enrollment		Test Scores	
		All	Village Only	All	Village Only
	(1)	(2)	(3)	(4)	(5)
Treatment	-2.696***				
	(0.271)				
Distance (Miles)		-0.163***	-0.150***	-0.193***	-0.158**
		(0.034)	(0.027)	(0.039)	(0.053)
Distance^2 (Miles)			0.009***		0.009
			(0.003)		(0.007)
Observations	1,490	1,490	1,490	1,374	1,374
R^2	0.75	0.27	0.31	0.39	0.39
Model	OLS	IV	OLS	IV	OLS

Note: This table contains the estimated effect of distance on enrollment and test scores. Column one contains the first stage estimate of the relationship between receipt of the treatment and distance to the nearest formal school (equation (4)). Column two estimates the relationship between enrollment and distance to the nearest formal government school of any type using treatment assignment as an instrument using equations (4) and (5). Column three provides OLS estimates of the relationship between formal school enrollment and the distance to the nearest village-based school. Column four estimates the effect of distance to the nearest formal government school on children's test scores using equations (4) and (5) while column five provides the estimates of an OLS regression of test scores on the distance to the nearest village-based school. All standard errors are clustered at the village group level. Statistical significance at the one-, five-, and ten-percent levels is indicated by ***, **, and * respectively and evaluated relative to the small sample t-distribution to account for the small number of clusters.

Table 10: Effects of Distance by Gender

	Formal Enrollment		Test Scores	
	All	Village Only	All	Village Only
	(1)	(2)	(3)	(4)
Distance (miles)	-0.132***	-0.103***	-0.150***	-0.082*
	(0.033)	(0.026)	(0.033)	(0.043)
Distance^2		0.003		< 0.001
		(0.002)		(0.005)
Distance * Female	-0.059*	-0.101**	-0.088*	-0.159**
	(0.027)	(0.037)	(0.042)	(0.054)
Distance^2 * Female		0.012**		0.019**
		(0.004)		(0.007)
Female	-0.022	-0.017	-0.414***	-0.400***
	(0.018)	(0.024)	(0.060)	(0.051)
Observations	1,490	1,490	1,374	1,374
R^2	0.26	0.32	0.39	0.40
Model	IV	OLS	IV	OLS

Note: This table displays estimates of the effects of distance disaggregated by gender. Columns one and three contain IV estimates of the effects of the distance to any formal school on enrollment and test scores respectively using treatment assignment as an instrument. Columns two and four contain OLS estimates of the relationship between distance to the nearest village-based schools and enrollment and test scores respectively. All regressions contain child and demographic characteristics, and the estimates are clustered at the village group level. Statistical significance at the one-, five-, and ten-percent levels is indicated by ***, **, and * respectively and evaluated relative to the small sample t-distribution to account for the small number of clusters.