Incentivizing schooling for learning: Evidence on the impact of alternative targeting approaches[1]

Felipe Barrera-Osorio
Harvard Graduate School of Education

and

Deon Filmer
World Bank

Draft version: February, 2013

Abstract. Although demand-side incentive programs (such as monetary incentives, scholarships, and conditional cash transfer programs) have been shown to increase school participation in a number of countries, impacts on achievement (as measured by students' test results) have been harder to identify. We study a large scale primary school scholarship pilot program in Cambodia with the objective to evaluate educational effects of two alternative targeting approaches: in some randomly selected schools recipients were targeted on the basis of poverty, and in others, the recipients were targeted on the basis of merit. While we show positive impacts on enrollment and school progression from both targeting approaches, achievement impacts are only detectable among merit-based recipients. We find differences in student effort and in the household's investment decision on education that are compatible with the asymmetry in achievement impacts. While there are some equity implications of merit-based approaches to targeting, the tradeoff is not particularly stark when the program reaches the very poor households. In this case, scaling up by targeting students with high academic potential—while ensuring that the poorest students are included—is likely to maximize both equity and effectiveness objectives.

JEL classification codes: I21; I24; I28; O10
Keywords: education; Cambodia; randomization; scholarships; merit-based targeting; poverty-based targeting.

[1] We thank Luis Benveniste, Norbert Schady, Beng Simeth, and Tsuyoshi Fukoaka and the members of the Primary School Scholarship Team of the Royal Government of Cambodia's Ministry of Education for valuable input and assistance in carrying out this work. Adela Soliz provided able research assistance. The paper has also benefitted from comments by Muna Meky, Richard Murnane, Halsey Rogers, Shwetlena Sabarwal, and Katja Vinha. The authors are responsible for any errors. This work benefited from funding from the World Bank as well as from the EPDF Trust Fund (TF095245). The findings, interpretations, and conclusions expressed in this paper are those of the authors and do not necessarily represent the views of the World Bank, its Executive Directors, or the governments they represent.

1. Introduction

Several recent rigorous evaluations in varied contexts and with diverse designs have improved our understanding of the effects of cash incentive programs —such as, conditional cash transfers, monetary incentives, and scholarships—on educational outcomes. The emerging consensus from the evidence appears to be that direct and indirect costs are important determinants of school participation, and therefore, programs that reduce these costs are quite effective in inducing higher enrollment and attendance rates (Fiszbein and Schady, 2009). However, once the students are enrolled in school, the effects of these programs on achievement outcomes are unclear. For example, the effects of conditional cash transfers on test scores are quite mixed, with negative or null results in Mexico (Behrman, Parker and Todd, 2005; Behrman, Sengupta, and Todd, 2000) but with positive results in Malawi (Baird, McIntosh and Ozler, 2011). Similarly, the effects of cash incentives are also mixed with positive results in Israel (Angrist and Lavy, 2009) and with null results in USA (Bettinger, 2010; Fryer, 2011). Finally, Kremer, Miguel and Thornton (2009) find positive effects on tests scores of a merit-based scholarship in only one of two districts in Kenya.

To further our understanding of the effects of scholarships in the educational decisions of individuals, this paper presents results from an evaluation of a scholarship program in Cambodia. The scholarship program randomly allocated scholarships to students in grade 4 in 209 schools in rural Cambodia. Depending on the school, the scholarship was either offered based on the student's household's economic status, poverty-based, or the student's performance on a test, merit-based. We address the question, what is the measurable impact of these primary school scholarships on measures of school participation and achievement. Our evaluation shows that both targeting approaches augment enrollment and attendance rates, but only the merit-based scholarship shows a positive impact on achievement. We find evidence of distinct responses in effort by the two treatment groups suggesting that the labeling of the recipients as either "poor" or "high achievers" matters.

There are five salient aspects that frame this evaluation. First, an important feature of the program we evaluate is that students were selected on the basis of two alternative criteria. Past evaluations have only looked at programs with either poverty or merit targeting. Targeting students through poverty-based criteria help students with the greatest costs to access education. However, targeting students through the merit-based criteria may be desirable if it maximizes the impact on learning as well as on enrollment—effectively incentivizing schooling for learning. If there is a positive relationship between academic success and household economic status (Orfield, 2002), then the cost of merit-based targeting may be the exclusion of the poorest students from the program. Using the parallel application of the two targeting

mechanism, we can address the question of whether the efficiency gain (defined as getting more achievement and enrollment per dollar transferred) comes at too great a cost of lower financial benefits to the poorest students.

Second, we collect follow-up data three years after the students received the scholarship, allowing us to capture longer run effects. In general, follow-up studies are carried out one to two years after the student was admitted into the scholarship program (for instance, Angrist and Lavy 2009, and Kremer, Miguel, and Thornton 2009). The longer lag allows us to overcome two general concerns with these types of evaluations: first, one year effects can disappear in the medium and long run; and second, it may take time for the effects to accrue. Ideally, any (positive) intervention will have long lasting effects.

Third, the program recipients were in grade 4 in primary school when they were accepted into the program. At this stage, the discount rate for the future is very high and problems of information present (Bettinger, 2010). As such, it is critical to incentivized students to stay in school early on in the educational ladder.

Fourth, in our evaluation we use follow-up information collected from home visits, in contrast to data collected at the school. This strategy allows us to avoid the common problems of student composition present when follow-up information is collected at school. Following up at home, all students are included, not only those who attend school regularly.

Finally, the program is in a low-income country, in a context where enrollment rates are low and drop-out rates are high. Given the scarcity of studies in this context, establishing the extent to which cash transfers are effective is an important contribution of this evaluation.

Before addressing the evaluation it is important to emphasize that we focus on a narrow set of objectives, namely schooling and achievement as measured by test scores. The program transfers cash to poor households, which in itself is potentially welfare enhancing. We do not address this aspect directly. Moreover, academic achievement is but one objective—albeit an important one—of schooling. There are additional social, as well as personal, (for example, better health, delayed marriage) impacts of schooling that the paper does not address or evaluate.[2]

[2] As discussed below, we focus below on the fact that the merit scholarship showed impacts on achievement, while the poverty scholarships did not. It is possible that poverty scholarships nevertheless have salutary impacts on other outcomes through their impact on schooling.

The paper is organized as follows. In the next section we discuss relevant literature. In section 3 we describe the setting, the program and the evaluation. In section 4 we present the empirical strategy and the validation of the identification strategy. In section 5 we present the main results. Finally, in section 6 we discuss the results and offer conclusions.

2. Literature review

There are three common underlying theoretical assumptions for the use of monetary incentive programs to induce higher investment in education. First, the direct and indirect costs of attending school may be an important deterrent to the optimal educational decisions of families. Therefore, reducing these costs may induce a higher investment in education. Second, students and families may discount very heavily the future, and as a result, they may not exercise the optimal effort in education. Monetary incentives will change the short run benefit of such investments. Third, families and students may not have complete information on the returns to education. Monetary incentives may serve as a signal and as a devise to highlight the importance of education.

Three different approaches to provide monetary incentives to students have been studied in the literature: Conditional Cash Transfers programs, direct monetary incentives to students, and scholarships. Conditional Cash Transfer (CCT) programs transfer cash to families on the condition that they comply with a set of conditions such as that children enroll and regularly attend school, pregnant women make regular prenatal visits, and that young children get regular health checkups. They have been extensively studied and have gained popularity in much of the developing world, becoming the largest form of social assistance in several countries (see the review in Fiszbein and Schady, 2009). Scholarship programs, such as the one we evaluate here, can be thought of as a single-child (i.e. individual recipient), single-sector (i.e. education) CCT.

Much of the rigorous evidence on the impact of CCTs has come from middle-income countries, mainly from Latin America. As such, the majority of the evidence comes from countries where baseline enrollment rates are high and impacts subsequently small (in part because there is little room for an increase). In contrast, this paper presents evidence from a scholarship program in Cambodia where enrollment rates are low, joining a much smaller set of evidence focused on low-income countries, such as Bangladesh (Chaudhury and Parajuli, 2008) and Malawi (Bair, McIntosh and Ozler, 2009).

Despite this large body of evidence on the impact of CCTs on enrollment and attendance, there is limited evidence to date on the impact of these programs on test scores, and the few available studies show results ranging from positive effects to no effects. For Mexico, Behrman, Parker and Todd (2005) and Behrman, Sengupta, and Todd (2000) do not find any impacts on achievement tests. However, Baird, McIntosh and Özler (2011) find positive effects in Malawi. The question of how to turn incentives for schooling into incentives for learning remains an open one.

Monetary incentive programs are slightly different in nature than CCTs, since they aim to encourage performance of students already in the educational system. The underlining rationale for such programs is to alter the motivation of individuals assuming their high discount rates of the future and imperfect information on the returns to education. Two recent evaluations of these programs in the USA show either mixed (Fryer, 2011) or no results (Bettinger, 2010) on standardized tests. However, an evaluation of a program in Israel shows positive results in exit-exams, take-up rates and college matriculations (Angrist and Lavy, 2009).

Rigorous evaluations of scholarships are rare due to lack of the exogenous variation in scholarship programs. Usually, scholarship beneficiaries are chosen based on observables and unobservable characteristics that are correlated with the educational outcomes of interest. Studies that have successfully evaluated scholarship programs include Filmer and Schady (2008, 2009), who evaluate a program in Cambodia, and Kremer, Miguel and Thornton (2009), who evaluate one in Kenya. These three papers concur in scholarships improving both enrollment and attendance, but they find mixed results on effects on standardized test scores.[3]

The mixed effects on achievement, as measured by students' test results, can be explained by three main theoretical considerations. The positive results may reflect monetary incentives being an extrinsic motivator (for both, the student and the family), especially among individuals from low-income families. Such short term motivation may be important when information on the returns to education is imperfect and when the discount rate is very high. Bettinger (2010) argues that these considerations may be highly relevant for primary school students. On the contrary, the program can induce a negative impact if it lowers intrinsic motivation. A person's interest in performing well can be negatively impacted if the program induces people explicitly to engage in that activity to achieve a certain goal by converting the activity in an obligation (for example, see Lepper et al., 1973). Finally, the program may not lead to any

[3] Besides giving scholarships concurrent to studies, there are programs where students are "promised" a scholarship if they perform well on a test administer in the future. For example, Kremer, Miguel and Thronton (2009) evaluate one such program. In theory, the promise of a reward elicits increased effort from students with abilities within reach of a scholarship.

impact on academic achievement if students are not able to respond to the incentives. It is possible that students do not know how to convert the incentive into actions that impact achievement. In short, they may not know the production function of test scores (Fryer, 2011). Or there may be strong complementarities with other inputs that are out of their control, and that are not similarly altered by the program (Fryer, 2011). Potentially the differences in the degree in which a student can respond to the incentives can explain the success of CCTs in terms of enrollment but their failure in terms of achievement. Possibly the individual and the family knows how to arrive to school and they can therefore alter their behavior to improve attendance, but the student may not know how to do better in tests (i.e. how to study well) translating into a lack of effects on test scores.

3. Country setting, program design, evaluation design, and data

Country setting

Cambodia has a tradition of using demand-side incentives to raise school enrollment and attendance rates. While some of these programs operate at the primary level—such as school feeding programs, or small scale programs that incentivize attendance for primary school children—most of the programs are targeted at the lower secondary school level, with the largest of them being the PB Scholarships (Filmer and Schady, 2009). The programs do not operate as simple "fee-waivers"; rather, the families of children selected for a "scholarship" receive a small cash transfer, conditional on school enrollment, regular attendance, and satisfactory grades.

Two rigorous evaluations of the impact of these lower secondary school programs have shown substantial increases in school enrollment and attendance as a direct consequence of the programs (Filmer and Schady, 2008 and 2009). Recipients are in the order of 20 to 30 percentage points more likely to be enrolled and attending school as a result of the scholarships. Filmer and Schady (2009) also show that the scholarships targeted to lower secondary school students led to more expenditures on education, and to less work for pay among recipients. There were no negative spillover effects— neither to non-recipients in schools or to ineligible siblings in the households. Impacts on learning outcomes were limited, suggesting issues in the quality and the match between student skill levels and instruction.

However, one important finding from the evaluations of the previous programs is that their targeting was only mildly pro-poor. For example, Filmer and Schady (2009) show that, despite the fact that the program was able to reach the poorest children who *applied* for the scholarships, the poorest of the poor

had already dropped out of school before grade 6—when they would apply for secondary school scholarships. Figure 1 presents the proportion of children aged 15 to 19 nationally who have completed each grade. As is evident, children from the poorest quintiles are much less likely to reach 6th grade. This finding suggests that a program that targets children at the end of grade 6 is not likely to be pro-poor—and that a program which targets poor students earlier in the schooling cycle is needed when the goal is to reach the poorest of the poor.

Program and evaluation design

Based in part on these findings, and the desire to assess the viability, effectiveness, and optimal design of such a program, the Royal Government of Cambodia implemented a new scholarship pilot program. The stated goal of the program was to offset the direct and opportunity costs of education, and as a result, increase schooling.[4] Implicitly, the goal was also to improve achievement outcomes through the additional schooling. This paper reports the results of the impact evaluation of that pilot program.

The basic design of the primary scholarship pilot was to select participating schools; and then, within schools identify scholarship recipients according to clear and transparent criteria. Once selected, recipients were required to stay enrolled, attend school regularly, and maintain passing grades in order to keep the scholarship until they graduated from primary school.[5] The program targeted students in selected schools entering the upper-primary level (Grades 4, 5 and 6). The scholarship amount was set to US$20 per student annually.[6] This represented approximately 3% of the typical 2009 annual income.[7] The scholarships were intended to be disbursed in two tranches of US$10: once in the beginning of the year, and once in the middle of the school year. However, in the first year of the program, scholarships were distributed in one lump sum due to delays in implementation.[8]

The pilot program was implemented in three provinces—Mondulkiri, Ratanakiri, and Preah Vihear—where average dropout rates between grades 3 and 6 were the highest, as determined by an analysis of Cambodia's Education Information Management System (EMIS). In order to narrow the geographic

[4] Primary schools are officially non-fee based. Opportunity costs include various forms of child labor which are relatively prevalent in the study area—although typically labor is combined with schooling at the primary school ages.

[5] There requirements are moderately enforced. Students absent for many days are followed up by school officials and if they return to school they remain eligible for the scholarship. After a student is absent for too many days they are classified as having dropped out and no longer are eligible for the scholarship.

[6] The previous lower-secondary scholarships were in the amounts of $45 and $60 per year—however, the evaluation found little impact on enrollment and attendance of $60 over and above that of $45 (Filmer and Schady, 2011)

[7] The 2009 GDP per capital in Cambodia was US $703; International Monetary Fund database.

[8] Scholarships distributions for the cohort of recipients analyzed here (Phase 1) took place in July 2009 (US$ 20), November 2009 (US$ 10), April 2010 (US$ 10), November 2010 (US$ 10); and April 2011 (US$ 10).

scope of the program, in Ratanakiri only seven districts with the highest dropout rates were selected for participation (out of a total of nine districts). In the other two provinces all districts were included. Within these selected districts, all primary schools which offered classes through grade 6 participated in the program.

The program offered scholarships to approximately 3,200 students from a total pool of 12,078 individuals in the 209 schools. Half of the 209 schools were randomly assigned to join the program in Phase 1 (during first year of the program, 2008/2009) and half in Phase 2 (or during the second year of the program, 2009/2010).[9] Figure 2 depicts the assignment scheme. The identification of impacts relies on the fact that among the cohort of students studied, 4th graders in Phase 2 schools were not eligible for scholarships in the 2008/2009 school year or later. These students therefore serve as a valid counterfactual group—a group that differs, on average, from the treatment group only in that it did not receive scholarships. Since these students in the control schools were never exposed to the program (even after the *subsequent* cohorts in Phase 2 schools became eligible for scholarships), the two groups of students can be tracked over time and their enrollment, attendance, and other outcomes compared.[10]

In order to evaluate the effectiveness of different targeting criteria, schools were randomly assigned to either "poverty-based" or "merit-based" targeting. Schools using "poverty-based" targeting for the scholarships selected students based on a poverty index similar to one used in the secondary school programs. All targeted students filled out a simple form with questions regarding their household and family socio-economic characteristics.[11] These forms were scored according to a strict formula based on weights derived from an analysis of household survey data, and as a result, a poverty index for each student was created.[12] The poverty index ranged from 0 (richest household) to 292 (poorest household). The application forms were scored centrally by a firm contracted specifically for this purpose, thereby reducing the probability of any manipulation of program eligibility. Within each school, the applicants with the highest scores (i.e. "the highest poverty") were offered a scholarship. As such, each school had a

[9] There were 104 Phase 1 schools and 105 Phase 2 schools.

[10] Because the scholarship offers are made according to strict criteria, applicants "just above" and "just below" the cutoff for eligibility could be compared using a regression discontinuity design (RDD) approach to evaluate program impact. In future work we will use this approach, and we will be able to use data from the first and second cohorts of students who applied to the program.

[11] Table 1 reports the full set of variables included in the calculation of the score.

[12] The weights were determined by estimating a model predicting the probability that a student would drop out of school during grades 4 to 6—since addressing this dropout was the stated goal of the program. Strictly speaking, the score should be referred to as a "dropout-risk score." However, the risk is essentially a set of household characteristics that capture the socioeconomic status of a household—weighted to capture those elements that predict dropout best. For convenience and ease of exposition, the score is referred to in this paper, as well as in program documents, as a "poverty" score.

different cut-off point. There were 52 "poverty-based" schools in Phase 1 and 53 in Phase 2—see Figure 2.

Schools using "merit-based" targeting, selected scholarship recipients based on a test score, proxying achievement. The test was adapted from the Grade 3 National Learning Assessment developed under the Cambodia Education Sector Support Project (Filmer and Schady, 2009).[13] All eligible students took the test and, within each school, the applicants with the highest test scores were offered scholarships. Again, scoring the tests was done centrally in order to minimize the risk of program manipulation. The number of students in each school type was fixed exogenously, and set to half the number of registered students in the year prior to the program (as determined by an analysis of EMIS data).[14] There were 52 "merit-based" schools in Phase 1 and 52 schools in Phase 2.

Implementation and data

This study evaluates the impact of the program on the first cohort of 4th graders—those who filled out the application forms in the first year of implementation (December 2008/January 2009).[15] All 4th grade students in program schools, Phase 1 and Phase 2 schools, filled out the application forms as well as took the assessment test (see Figure 2). Recipients received scholarship disbursements (conditional on remaining in school, attending regularly, and maintaining passing grades) during the 2008/2009, 2009/2010 and 2010/2011 school years.

We use three main data sources to evaluate program impacts. First, we use the full set of data collected at the time the students applied for the scholarships. That is, we have baseline household characteristics as well as math and Khmer language test scores for all applicants. Second, we use the official list of students who were offered a scholarship. Third, we use endline data that were collected specifically for this evaluation. These data are derived from a survey administered at the end of the 2010/2011 school year, three years after the implementation of the program, to a random subsample of students from each program school. The students who began participation in Phase 1 and stayed in school were finishing (or just finished) grade 6. The survey was administered at home (i.e. not in schools) to the child who applied for the scholarship, and included a household module administered to

[13] The National Assessment was implemented in a sample of schools nationwide in Grade 3 during the 2005/2006 school year (Royal Government of Cambodia, 2006).

[14] The number of scholarships is not equal to half the number of applicants because the rule was to allocate scholarships to all applicants who had scores above *or tie* the cutoff score.

[15] These students were "supposed" to have filled out application forms prior to the beginning of the school year, i.e. when they were still in 3rd grade. Because of delays in effectiveness and implementation of the overall project, the application process could only be implemented once the students had begun 4th grade.

their mother, father or other caregiver. In total 3,618 applicants were interviewed.[16] For the majority of the analyses we use data from 1,377 students who were in grade 4 at baseline and were offered a poverty- or merit-based scholarship, or would have been offered a poverty- or merit-based scholarship had they attended a Phase 1 school.

The survey covers a broad range of issues, both in terms of school participation (for example, the "intensity" of school participation though questions related to time spent in school), other activities such as labor market participation, as well as various measures of cognitive development, and achievement. Importantly, given that these data were collected at the home of the student, both recipients and non-recipients were interviewed regardless of their attendance in school, thus avoiding problems arising from only interviewing those children enrolled and attending school at follow-up (Kremer, Miguel and Thornton, 2009).

A Mathematics and Digitspan tests were applied to students at follow-up. The items on the Mathematics test were drawn from a variety of sources including the baseline mathematics test, questions from the national grade 6 assessments, and publicly released questions from the Trends in International Maths and Science (TIMSS) Grade 4 Assessment. Items were tested during a pretest and only those with adequate properties were retained for the final test. The final multiple choice test measured both knowledge and the capacity to use this knowledge for problem solving. Performance in the math test is a measure of the most immediate academic impact of the intervention. Since part of the requirement for the scholarship is to receive passing grades, we would expect the program to incentivize the student to study more and which in turn should affect the student's ability to solve mathematical problems.

The other test that we evaluate is the Digitspan test. In a Digitspan test a series of numbers are read to a respondent who is then asked to repeat the numbers back to the enumerator. The series increase from 2 digits to a larger and larger number, up to 9-digit numbers. Respondents are also asked to repeat the numbers back in reverse order. The test is typically interpreted as a measure of short term memory and working memory capacity.

4. Empirical strategy

Empirical Strategy

[16] Attrition was 15%. Analyses of attrition patterns show that attrition is not different in Phase 1 and Phase 2 schools, nor is it related to poverty/merit status of the school.

We estimate a reduced-form equation of the program impact on enrollment and attendance outcomes, on test scores, and on potential mechanisms of transmission (school / teacher effort and student / household effort). A separate estimation is carried out for each targeting mechanism. For the merit-based schools, the estimation is based on the equation

$$Y_{i,t1} = \beta_0 + \beta_1 T_{i,s}^{M} + X'_{i,t0}\mathrm{B} + \eta_{i,s,t1} \quad (1)$$

$Y_{i,t1}$ denotes the value of the outcome variable for individual *i* at follow-up (*t1*); $X'_{i,t0}$is a vector of controls measured at baseline; and $\eta_{i,t1}$ captures unobserved student characteristics and idiosyncratic shocks. T_i^M indicates treatment status for student *i* in the merit-based school *s*. It is equal one if the student was offered the merit-based scholarship, and zero for untreated students in the control schools who would have been eligible for merit-based scholarships —based on their test scores— had they attended a treatment school:

$$T_{i,s}^{M} = \begin{cases} 1 \; if \; Test\; Score_i \geq Test\; Score_{50\; Perc,Merit\; School,} \\ 0 \; if \; Test\; Score_i \geq Test\; Score_{50\; Perc,Control\; School,} \end{cases}$$

The controls include an indicator for gender, the number of minors in the household, indicators for whether the household owns a motorcycle, a car/truck, an oxen/buffalo, a pig, an ox or buffalo cart; indicators for whether the house has a hard roof, a hard wall, a hard floor, an automatic toilet, a pit toilet, electricity, piped water; as well as the poverty index and test scores. [17]

An analogous estimation is done for the poverty-based schools:

$$Y_{i,t1} = \beta_0 + \beta_1 T_{i,s}^{P} + X'_{i,t0}\mathrm{B} + \eta_{i,s,t1} \quad (2)$$

where $T_{i,s}^{P}$ is equal one if the student was offered the poverty-based scholarship, and zero for untreated students in the control schools who would have been eligible for poverty based scholarships had they attended a treatment school:

$$T_{i,s}^{P} = \begin{cases} 1 \; if \; Poverty\; Index_i \geq Poverty\; Index_{50\; Perc,Poverty\; School} \\ 0 \; if \; Poverty\; Index_i \geq Poverty\; Index_{50\; Perc,Control\; School} \end{cases}$$

[17] We impute values of (few) missing observations in the control variables, and include dummy indicative variables for the imputation. Since the estimation is based on a random assignment, the imputation does not make any difference in the estimation.

Given that the treatment variable identifies at the baseline all individuals who were offered the scholarship, the estimator is in practice an intention-to-treat estimator (ITT). Errors are clustered at school level, and each model includes district-level fixed effects.

The design of the intervention also allows for directly estimating the effects of the program on non-treated students, or spillover effects from the treatment. In each treatment school, approximately half of the students were not treated. Therefore, given the random assignment of schools into treatment, we can compare the non-poor (and non-treated) students in the poverty treatment schools with non-poor (and non-treated) students in the control schools. Similarly, we can compare students who did not receive scholarships in the merit-based schools with students in control schools who would not have qualified for a merit scholarship. In order to estimate the effects on the non-treated students, we used Equation (1), but replace T_i with the appropriate group of students. In this program, the scholarship is offered after the baseline test, and during the duration of the program, non-scholarship recipients cannot change their status. Therefore, any effects on non-treated students emanate from complementarities and interactions between treated and non-treated students in the treatment schools during the academic year.[18]

Baseline balance and characterization of the study sample

This section presents the general characteristics of the study sample and the validation of the random assignment by comparing treatment and control students at the baseline. In Table 1, columns (1) and (2) are based on grade 4 students in the control and the treatment schools, whereas columns (3)-(6) only use information of the treated students in the treatment schools and untreated students in the control schools who would have been eligible for treatment —based on their poverty index score or on the baseline test score—had they attended a treatment school. The sample is restricted to Grade 4 students at baseline, and as described above, the control group of students were not beneficiaries of the program expansion. Columns (1), (3), (5) show the means and standard deviations of household and individual characteristics prior to the intervention. Columns (2), (4) and (6) show the differences and standard errors at baseline between the treatment and control groups. That is, the values are the coefficient estimates from a regression of each characteristic on a dummy variable equal to 1 for a treatment school.

Two main features emerge from Table 1. First, treatment and control groups are similar in observed characteristics. Only a few coefficients in columns (2), (4) and (6) are statistically significant (of the 48

[18] These peer effects are different in nature from the externality effects estimated in Kremer et al. (2009). There the authors estimate the effect of the "promise" of a scholarship for students with low pre-intervention scores. Those effects presumably emanate from, first, the effort that all students may exercise in order to get the scholarship and second, the pressure that parents may exercise on teachers.

differences reported, 5 are significant at the 10% level, and among them, 1 is significant at the 1% level). The results in Table 1 confirm the validity of the random assignment since both control and treatment groups are similar in their observed characteristics. More importantly the poverty index and the test score are statistically equal between the treatment and the control groups. Figures 3 and 4 present the density of poverty index at baseline for treatment and control schools. There is a clear overlap throughout the whole distribution. We cannot reject equality of both distributions using a Kolmogorov-Smirnov test.

Second, on average, the recipients who are offered scholarships on merit basis have more assets, a lower poverty index score, and performed better on the baseline test than students with a poverty-based scholarship. For instance, the poverty index, which ranges from 0 (wealthiest family in the sample) to 292 (poorest family in the sample), has a mean of 245.13 for poverty based students, and 218.2 for merit-based students. Likewise, the baseline test (ranging from 0 to 25) has a mean of 19.77 for merit-based students and 17.74 for poverty-based students. However, the correlation coefficient between the poverty index and the test score for the entire sample of individuals at baseline is remarkably low (for example, the spearman correlation is 0.1593, and the pairwise correlation is 0.17). We return to this issue when we discuss the tradeoffs between the two targeting approaches.

The main finding from Table 1, however, is that the samples are balanced at baseline, allowing us to identify casual effects.

5. Results

Impacts on enrollment and attendance

The intervention is aimed at directly incentivizing higher enrollment and attendance rates. In order to keep the scholarship, the selected students must stay enrolled, attend school regularly, and maintain passing grades until they graduate from primary school (sixth grade). We focus on three enrollment and attendance proxies: the proportion of students reaching 6th grade, the highest grade completed, and the hours of school attended in the past seven days.

In order to provide a baseline for assessing the relative magnitude of impacts, Table 2 reports average outcome variables of the students at the follow-up in the control groups (Columns denoted with (1)) and their difference with respect to the treatment students (Columns (2)). For the poverty and merit controls, 61% and 64% of the students reported reaching at least sixth grade, respectively. The average grade

completion for the students in the control groups was around 5.4 grades. The third outcome variable was constructed from a question on how many hours the students attended school the past seven days, conditional on being enrolled. Depending on the control group, students reported an average having attended school for about 8.83 hours (poverty) and 9.27 hours (merit) in the past week.

Columns (2) show simple differences between treatment and control groups. In the poverty-targeted treatment, there were significant and positive differences in the enrollment and attendance variables. For example, treated students with the poverty scholarship are 18.6 percentage points more likely to reach grade 6 than control students; they complete 0.349 more grades and they attend 3.46 hours more than the control group. However, there are not effects on any of the achievement measures in this treatment. In contrast, the merit-targeted treatment shows effects both in enrollment/attendance and achievement. Students with the merit-targeted treatment are 13.1 percentage points more likely to reach grade 6 than control students; they complete 0.234 more grades than controls; and they achieve around 0.15 sd more in the math and digtspan tests than the controls.

Table 3 reports the program impacts on the enrollment and attendance proxies—coefficient β_1, Equations (1) and (2). Columns (1) and (2) present the effects of treatment of the poverty and merit-based interventions on the outcomes, when controlling for baseline student characteristics (i.e. the variables reported in Table 1), poverty index, test scores at baseline, and province fixed effects. As discussed above, the randomized assignment was successful in that it produced a balanced sample. Controlling for additional variables should therefore only affect the precision of the estimates, not the magnitude of the estimated effects. As expected, the results are very similar to the simple differences without controls, results in Columns (2), Table 2.

Overall, Table 3 shows consistent evidence of positive impacts from the interventions on enrollment and attendance. The proportion of students reaching grade 6 increased with both treatments, and the effects are similar in magnitude. The estimated impacts range between 12 and 17 percentage points from the control mean of 61% and 64%, for the poverty and merit treatments respectively. Similarly, the intervention increased the average highest grade completed in both treatment samples. It increased by 0.332 grades in the poverty-based treatment and 0.182 grades in the merit-based treatment, from a control group mean of about 5.4 years. These impacts are similar to those found in the context of the Secondary School scholarships program where impacts on enrollment were in the order of 20 to 25 percentage point increases (Filmer and Schady, 2008). These impacts are larger than the majority of impacts documented in countries elsewhere in the world (Fiszbein and Schady, 2011), and should be assessed against the very small size of the transfer (i.e. US$ 20 per year).

Although attendance (as measured by the number of hours in school in the past seven days) shows positive impacts, in neither case are the impacts statistically significant. The point estimates at face value suggest that the intervention increases attendance on the order of 2.9 hours per week for the poverty-based treatment and 0.64 hours for merit-based treatment.

In sum, there is strong evidence that the program increased enrollment and suggestive evidence that it also improved attendance rates, regardless of whether the scholarship is based on merit or on poverty status.

Results on test scores

We can expect impacts from the program on test scores for both targeting mechanisms for two main reasons. First, the program incentivized more enrollment and school attendance and therefore students are more exposed to school—and potentially through that additional schooling acquire more learning. Second, the program requires that all treated students —both with merit and poverty scholarships— maintain passing grades, and therefore, it gives students an incentive to study more.

Table 4 presents impacts of the interventions on a Mathematics test and a Digitspan test, our two measures of academic and cognitive achievement. Table 4 has the same structure as Table 3; it presents impact coefficients from models that control for baseline characteristics and district fixed effects. All achievement measures are standardized such that the control group has a mean of zero and standard deviation one. As such, impacts can therefore be interpreted as changes in a standard deviation of the achievement measure.

In contrast to the results on enrollment and attendance, Table 4 reveals different impacts on test scores for the poverty and the merit treatments. While there are no effects from treatment on test scores for students treated based on their poverty status, there is a clear pattern of positive effects from the intervention for students treated based on merit. All the point estimates for the merit treatment are positive and statistically significant. For the merit sample when using additional controls, the effect of the intervention on the math test is 0.170 standard deviations and the effect on the Digitspan test is 0.149 standard deviations. The results on math are of particular interest, since math ability is potentially the measure that is the most sensitive to exposure to additional schooling. Moreover, households have potentially less ability to substitute for teaching math than vocabulary or other types of learning. These effects on test scores are similar in magnitude to the merit-based scholarship program evaluated in Busia, Kenya (Kremer, Miguel and Thornton, 2009).

In sum, the results suggest that the program incentivized both types of students – those from poorer households and those with higher academic merit—to enroll in additional years of schooling. However, only students who received treatment based on merit show any gains in academic achievement from the intervention. That is, additional schooling (e.g. enrollment and attendance) results in better achievement only for those students who had better test scores at the baseline.

In light of this result, it is important to explore the heterogeneity of effects by baseline test score and poverty index. For each poverty school (treated and control) we identify students that were above and below each school's median baseline test (M_i), where M_i is one for those who scored higher than the median and zero for those who scored lower. We run the following model:

$$Y_{i,t1} = \beta_0 + \beta_2 T_{i,s}^{P} + \beta_2 M_i + \beta_3 (T_{i,s}^{P} * M_i) + X'_{i,t0}\mathrm{B} + \eta_{i,s,t1} \quad (3)$$

The interaction coefficient estimates the effect of the program for students with the poverty scholarship who score high in the baseline test.[19]

In an analogous way, for merit-based schools (treatment and controls) we identify students above and below each school's median poverty index (P_i), and then, we run the model

$$Y_{i,t1} = \beta_0 + \beta_2 T_{i,s}^{M} + \beta_2 P_i + \beta_3 (T_{i,s}^{M} * P_i) + X'_{i,t0}\mathrm{B} + \eta_{i,s,t1} \quad (4)$$

In Table 5, Column (1) presents the impact of the poverty-based treatment, and Column (2) presents the impact of merit-based treatment, for each follow-up test. As Table 5 shows, the merit-based treatment has similar or larger positive impacts on the poor population than in the non-poor population. The interaction coefficient for the math test is positive, with a p-value of 0.115: poor students in the merit-based treatment score 0.216 sd higher than the control students. In terms of the Digitspan test, both poor and non-poor, score 0.192 sd higher than the control students. In contrast, the poverty-based treatment does not elicit positive results from either the high baseline achievers or the low baseline achievers. The treatment coefficient, as well as the interaction term, is still close to zero, and statistically insignificant. In other words, the asymmetry in test results for poverty-based and merit-based targeting mechanisms persists. More importantly, the poverty-based treatment does not induce better test results among high baseline performers, whereas the merit-based treatment does induce better test scores among baseline poor individuals. This result is significant since the conditions for both scholarships—merit and

[19] An alternative specification is to run the separate regressions for each subgroup. The results are very similar to the ones presented in Table 5.

poverty—are the same. Given the results for poor individuals in the merit-based treatment, similar results were expected for the high achievers in the poverty-based treatment.

Effort

Both types of scholarships provide incentives for students to increase effort. For instance, it is plausible that, due to the scholarship, students increase the hours studied outside of school. Also, families may be motivated to invest more in education expenditures such as textbooks, and as a result, help the student to keep the scholarship. Likewise, the program can impact directly the behavior of the school and teachers. For instance, under an altruistic model, teachers can increase attention to students with scholarships with the hope that they can keep them. Also, it is possible that pressure from scholarship winners' parents to their child's teacher may elicit higher effort by the teacher. As such, the school may change its behavior with the introduction of the scholarships. Banerjee and Duflo (2006), while presenting the results of a previous version of Kremer et al. (2009), discuss changes in teacher motivation from greater accountability expected by families. As mentioned before, the follow-up data come from a household interview and therefore we do not have school level information. However, household interviews allow us to gather evidence regarding potential non-school related channels that may explain the results, such as, student effort outside of school and the household's investment in education.

Students' effort is captured by the amount of time spent studying, doing homework, and taking private lessons outside of school. The household's effort is measured by the total amount of education expenditure by the household. The effects of treatment on each of these two variables are presented in Table 6. On average, control students spent 3.5 hours per week doing school tasks outside of school. While merit-treated students spent more time doing academic work outside of school (an increment of 0.7 hours), poverty-based students did not increase the amount of time spent on academic work. The household response to scholarships in terms of education spending and the nature of that spending differs across the scholarship types. On average, household's education expenditures are approximately US$17. Households with a merit scholarship recipient spent US$5.4 more on education than control students' households. In contrast, there were no impacts on these outcomes in households with poverty targeted recipients. The results suggest that only household with merit-based students increased effort.

All in all, it seems that families in the merit-based treatment invest more in the education of their students; in contrast to what happens in families with the poverty-based treatment. Also, the students themselves in the merit-based treatment put more effort outside school than the poverty-based students.

Peer effects on enrollment, attendance and achievement

As discussed above, the design of the intervention allows us to test directly for peer effects. These peer effects are different in nature from the externality effects estimated in Kremer et al. (2009). Kremer et al. (2009) estimate the effect of the "promise" of a scholarship for students with pre-intervention low scores. Those effects presumably emanate from, first, the effort that all students may exercise in order to get the scholarship and second, the pressure that parents may exercise on teachers.

In contrast, this program has two different potential sources of peer effects. If increased enrollment and attendance leads to classroom overcrowding, then this may hurt the learning opportunities afforded to other students. At the same time, there might also be positive spillover effects if the scholarships create a positive environment in favor of schooling and learning which could affect all children in a classroom.[20] Moreover, peer effects might differ by the targeting approach. For instance, applicants denied a merit-based scholarship might become discouraged, and eventually perform worse. Since there will always be limited scope for containing spillover effects, they should be estimated directly, and factored into the "equity/efficiency" tradeoff in program design.

Table 7 presents the coefficient estimates where the treatment indicator variable that is one for those students not treated in the treated schools, and zero for students in control schools with analogous baseline test scores or poverty index scores. The dependent variables are the same attendance and achievement variables; the proportion of students reaching 6th grade, the highest grade completed, and the hours of school attended. The estimates can be read similarly to those reported in Table 3, but the sample consists of of the non-recipients in the treatment schools and their counterfactual in the control schools. There seems to be a positive peer effect in poverty-treated schools: the three coefficients are positive; and one of them is significant at the 5% level. Moreover, as expected, the point estimates are lower than the direct effect of the program (Table 3). Regarding the merit-based treatment, we do not find any evidence of peer effects. [21] Table 8 is the analogous table for the two measures of achievement. None of the coefficient estimates are statistically significant, and all the coefficient estimates are very close to zero.

In sum, it seems that the poverty-based treatment may induce more attendance from non-treated peers. No other peer effects are detected, neither positive nor negative, from the program.

[20] An additional positive spillover might be on younger cohorts who stay in school longer in order to potentially benefit from scholarships. We do not have the data available to address this issue.
[21] The results therefore dispel the notion that there was discouragement among merit scholarship applicants who did not receive a scholarship.

6. Conclusions

The fact that some students were able to take better academic advantage than others from additional school exposure highlights an issue rarely addressed in previous evaluations of conditional cash transfer programs. Recent evidence on monetary incentives for schooling shows that students are able to change their behavior on the margins that are under their control—for example, enrollment and attendance. However, these positive effects do not necessarily translate into test score gains. For example, despite the fact that Mexico's *Oportunidades* program—a rigorously evaluated conditional cash transfer program—induced students to enroll and attend additional years of school, the program did not induce better test scores. A recent set of papers has argued that education systems in developing countries are typically tailored towards better-off and better-skilled students. Specifically, Glewwe, Kremer and Moulin (2009) show that only the strongest students at baseline were able to take advantage of textbooks that were provided to schools in Kenya. Furthermore, Duflo, Dupas and Kremer (2011) show that teachers who were assigned to students at the bottom of the achievement distribution were less likely to teach.

Our findings based on the Cambodian Primary Scholarships Pilot add to this discussion. On one hand, poorer students who are not academically prepared are not able to gain significantly in learning, as measured by scores, from the additional schooling. Clearly more work is needed—in Cambodia and elsewhere—to establish how to best prepare and incentivize teachers to reach these students. On the other hand, this evaluation suggests that for students who are better academically prepared—including poor students—incentivizing school attendance can pay off in measurable learning outcomes. Therefore, remedial lessons for students in the early grades, or increasing school readiness among poorer students, for example through early child development programs, might be complementary approaches to increase the impact of schooling, and of programs that incentivize schooling. Indeed, data from Cambodia suggest that children suffer from substantial delays in cognitive development hampering their school readiness (Naudeau and others, 2011).

Furthermore, this paper's findings are compatible with motivation being an important factor in the success of the program in incentivizing students. Some studies in psychology stress the role of stereotypes. A particular treatment can label individuals in certain way (stereotype), and the label can reinforce, or detract, the treatment (Hoff and Pandey, 2006). By the merit of having been labeled as succeeding academically, merit-based students may be motivated to work more, and their families to invest more. In contrast, students in the poverty treatment group are not recognized for their achievements

and may not get similar encouragement to excel in their studies. As such, we posit a possible explanation to explain the asymmetry in effects. The merit-based treatment may additionally motivate the recipients of the scholarship and their labeling for academic achievement may reinforce the treatment effects by concurrent behavior changes in other inputs of achievement, mainly those provided by parents and teachers. On the contrary, recipients of a poverty-based scholarship do not benefit from the positive label and in fact may suffer discouragement. To test whether such labeling effects are indeed present, an ideal experiment would randomize the labeling of students, orthogonal to achievement at baseline. Clearly, it is also important to disentangle the motivation of students and from the motivation of teachers.

These findings open new avenues of research, especially in light of the results in the USA (Fryer, 2011). Fryer (2011) shows that monetary incentives change student behavior on margins that are under the direct control of the individual. That is, monetary incentives induce changes in classroom behavior and in the number of books read, but they do not induce (significant) changes in test scores. Following the same logic, a conditional cash transfer program can be successful in motivating individuals to increase school attendance since they are in (almost perfect) control of attendance, but not be successful in changing test scores, which depend on additional inputs such as the quality of teaching. However, we find that students can actually change their performance in tests, probably because the program triggers increased effort from students and households, and potentially also from induces changes in teacher behavior. To further understand whose changes in behavior have an impact, an ideal experiment would randomize the information teachers receive on scholarship recipients, and in that way, isolate the effect of changes in effort by individuals (and households) from changes in teacher behavior toward recipients.

Given two targeting approaches, the Cambodia Pilot program lends itself to shed light on the potential tradeoff between efficiency—defined as achieving more achievement per dollar transferred—versus equity—defined as reaching the poorest population. Analyses of the socio-economic profile of program applicants and recipients under the two targeting schemes, and comparisons to the national distribution of socio-economic characteristics, suggest that both targeting approaches are heavily skewed to the poor. The first panel in Figure 5 shows that 50% of those who applied to the program are in the poorest nationally-benchmarked quintile; fewer than 3% of applicants were from the richest quintile. Clearly the program was targeted to poor areas and poor schools in Cambodia. Unsurprisingly, scholarships targeted to the poorest from each school yields a greater pro-poor distribution of benefits than when targeting by merit. In the poverty-targeted schools 85% of applicants were from the poorest two quintiles of the population, and 63% were in the poorest quintile. Merit-based targeting is less pro-poor but is still able to reach the poorest groups in the population. In the merit-targeted schools 76% of applicants were from the poorest two quintiles of the population and 54% were in the poorest quintile.

A complementary analysis of the within-school correspondence between high/low poverty applicants and high/low test scoring applicants yields a similar conclusion: the tradeoff between efficiency and equity was not particularly stark because wealthier applicants were not necessarily the higher-scoring ones. Given the relatively effective geographic targeting of the program it is unclear whether this result is generalizable. In other settings (e.g. where there is greater heterogeneity in student poverty levels) the result may not hold. Scaling up an approach that targets students with high academic potential—while ensuring that the poorest student are among that set—is likely to be the approach that maximizes both the equity and effectiveness objectives of the program.

References

Angrist J. and V. Lavy. 2009. "The Effects of High Stakes High School Achievement Awards: Evidence from a Randomized Trial" American Economic Review, 99:4, pp. 1384-1414

Baird, Sarah, Craig McIntosh and Berk Özler, 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment," The Quarterly Journal of Economics, Oxford University Press, vol. 126(4), pages 1709-1753.

Banerjee, Abhijit, and Esther Duflo. 2006. "Addressing Absence." Journal of Economic Perspectives, 20(1): 117–132.

Behrman, Jere R., Susan W. Parker, and Petra E. Todd. 2005. "Long-TermImpacts of the Oportunidades Conditional Cash Transfer Program onRural Youth in Mexico." Discussion Paper 122, Ibero-America Institute for Economic Research, Göttingen, Germany.

Behrman, Jere R., Piyali Sengupta, and Petra Todd. 2000. "The Impact of PROGRESA on Achievement Test Scores in the First Year." Unpublished manuscript, International Food Policy Research Institute, Washington, DC.

Bettinger, Eric. 2010 "Paying to Learn: The Effect of Financial Incentives on Elementary School Test Scores,"NBER Working Paper16333

Chaudhury, Nazmul and Dilip Parajuli. 2008. "Conditional Cash Transfers and Female Schooling: The Impact of the Female School Stipend Programme on Public School Enrolments in Punjab, Pakistan." *Applied Economics.*

Duflo, Esther, Pascaline Dupas and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review.* 101(5): 1739-74.

Filmer, Deon, and Norbert Schady. 2008. "Getting Girls into School: Evidence from a Scholarship Program in Cambodia." Economic Development and Cultural Change 56(2): 581–617

Filmer, Deon and Norbert Schady. 2009. "School Enrollment, Selection and Test Scores." World Bank Policy Research Working Paper No. 4998. The World Bank, Washington, DC.

Filmer, Deon and Norbert Schady. 2011. "Does more cash in conditional cash transfer programs always lead to larger impacts on school attendance?" *Journal of Development Economics*. 96(1): 150–157.

Fiszbein, Ariel and Norbert Schady. 2009. Conditional Cash Transfers: Reducing Present and Future Poverty. The World Bank. Washington, DC

Fryer Jr., R. G. 2011. "Financial Incentives and student Achievement: Evidence from randomized trials." The Quarterly Journal of Economics (2011) 126, 1755–1798.

Glewwe, Paul, Michael Kremer and Sylvie Moulin. 2009. "Many Children Left Behind? Textbooks and Test Scores in Kenya." *American Economic Journal: Applied Economics*. 1(1): 112-135.

Hoff, K. and P. Pandey. 2006. "Discrimination, Social Identity, and Durable Inequalities" American Economic Review, Vol 96, No. 2, pp 206-211

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. "Incentives to Learn." *Review of Economics and Statistics*. 91(3): 437-456.

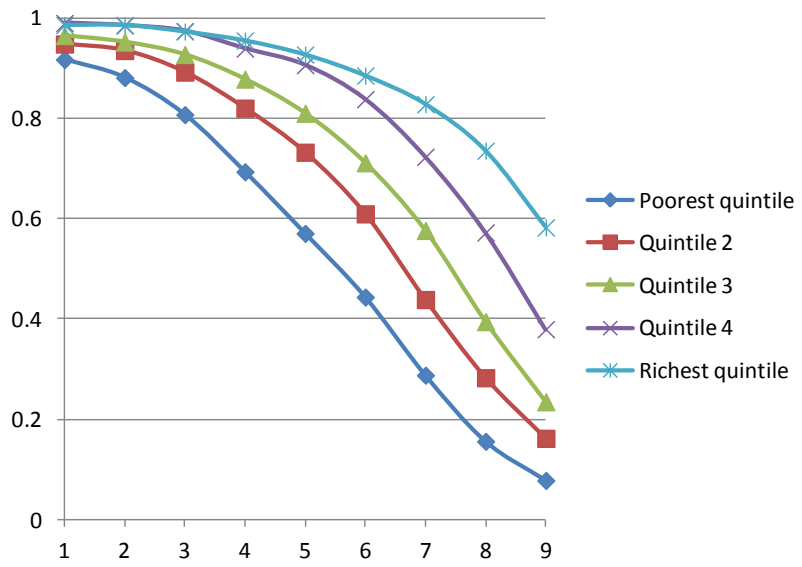
Lepper, M. R., D. Greene, and R. Nisbett. 1973. "Undermining Children's Intrinsic Interest with Extrinsic Rewards: A Test of the "Overhustification" Hypothesis" Journal of Personality and Social Psychology, Vol. 28, No. 1, pp. 129-137

Naudeau, Sophie, Sebastian Martinez, Patrick Premand, and Deon Filmer. 2011. "Cognitive Development among Young Children in Low-Income Countries" in Alderman, Harold ed. *No Small Matter: The Impact of Poverty, Shocks, and Human Capital Investments in Early Childhood Development*. The World Bank. Washington, DC.

Orfield, Gary, "Foreword," in Donald E. Heller and Patricia Marin (Eds.), Who Should We Help? The Negative Social Consequences of Merit Aid Scholarships (2002) (Papers presented at the conference "State Merit Aid Programs: College Access and Equity" at Harvard University). Document found in http://civilrightsproject.ucla.edu/research/college-access/financing/who-should-we-help-the-negative-social-consequences-of-merit-scholarships/

Royal Government of Cambodia. 2006. Student Achievement and Education Policy: Results from the Grade Three Assessment—Final Report. Cambodia Education Sector Support Project—National Assessment Component. Phnom Penh, Cambodia.

Figure 1: Proportion of 15 to 19 year olds who have completed each grade, by quintile.



Source: DHS 2010

Figure 2. Design of the intervention

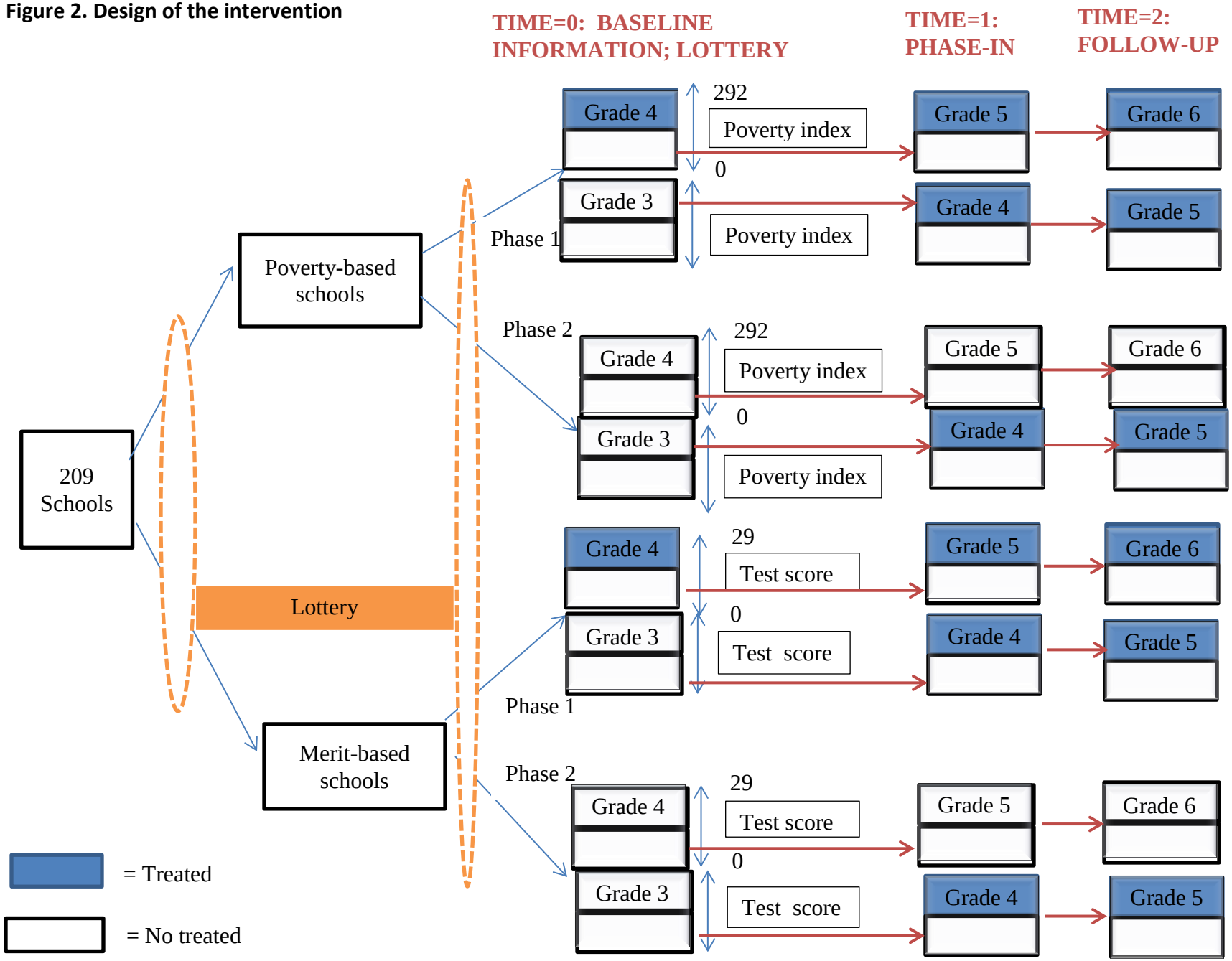
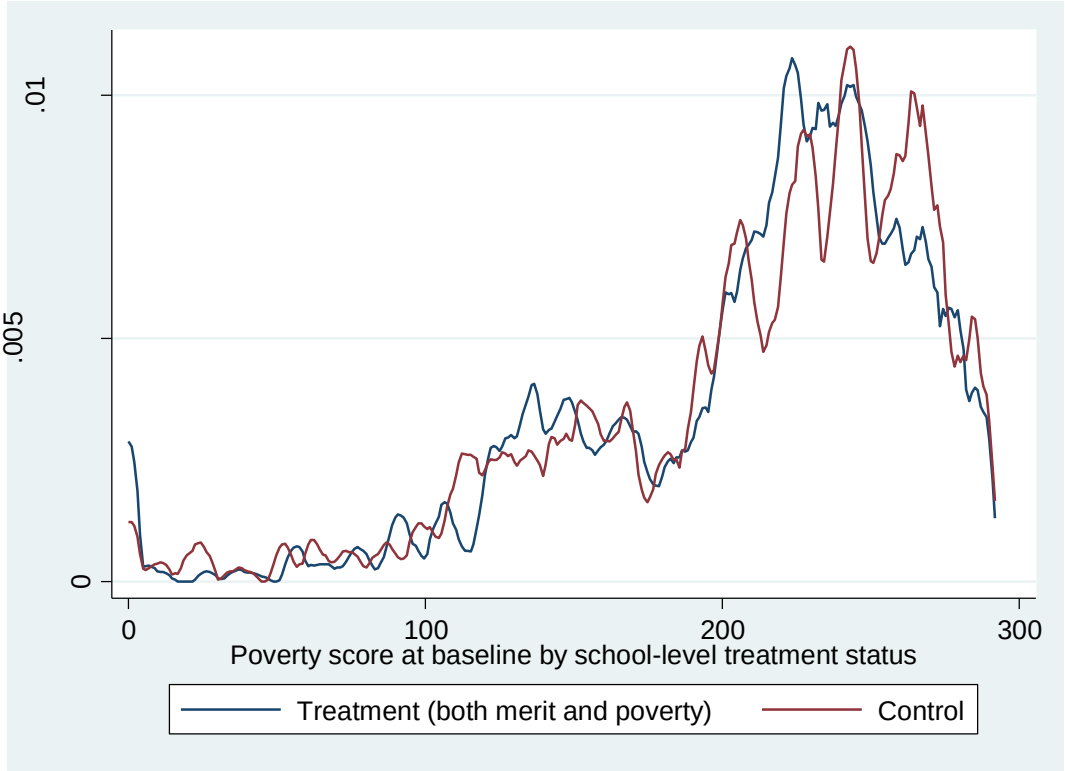
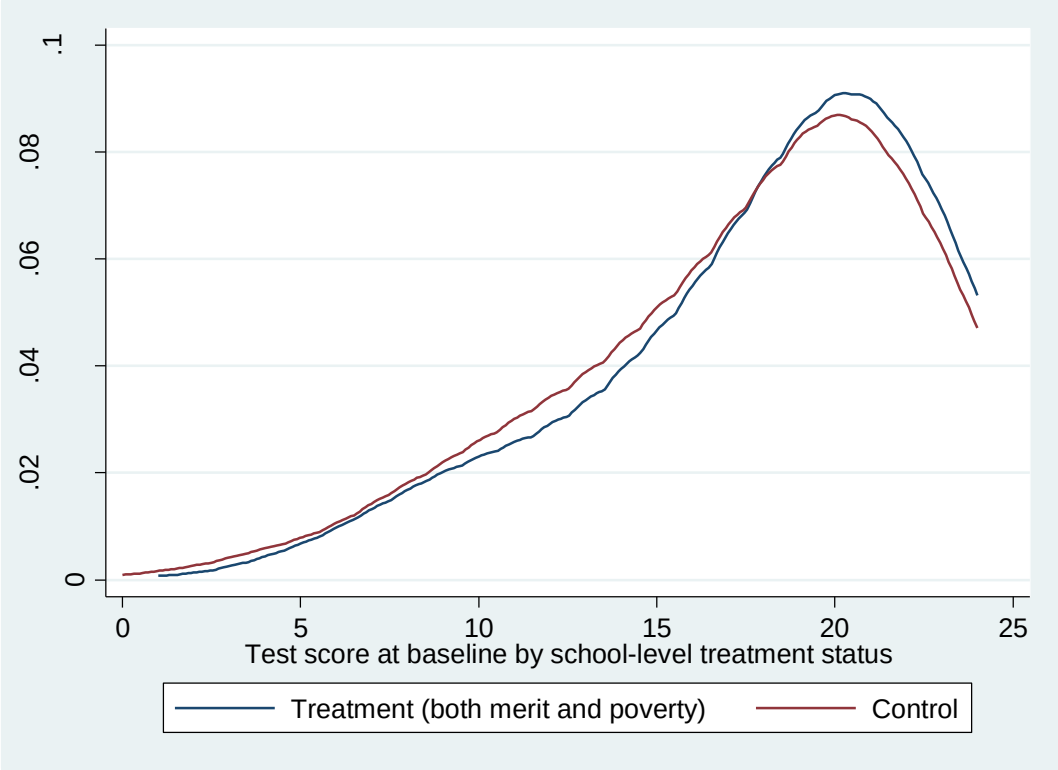


Figure 3: Poverty score at baseline, Treatment versus Control



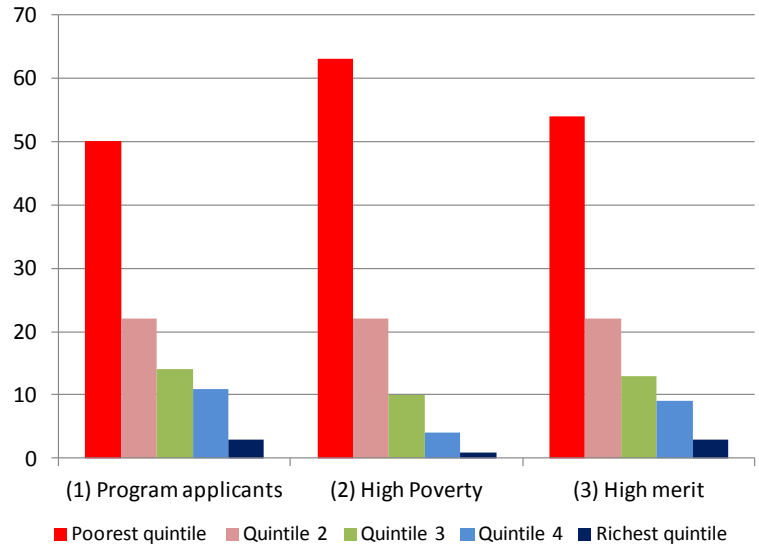
Source: Students at follow-up, using baseline information

Figure 4: Test scores at baseline, Treatment versus Control



Source: Students at follow-up, using baseline information

Figure 5: Distribution of selected populations across nationally benchmarked quintiles



Source: Analysis of Cambodia DHS 2010 and Primary Scholarship Application forms. Quintiles are defined on the basis of an index of household wealth-related variables that are collected in both the DHS 2010 as well as on the scholarship program application forms.

Table 1. Baseline Balance and mean and standard deviation of baseline characteristics

	School Level		Student Level			
	Control	Difference with treatment	Control Poverty	Difference with treatment	Control Merit	Difference with treatment
	(1)	(2)	(3)	(4)	(5)	(6)
Gender	0.49	0.037	0.52	0.100***	0.49	-0.03
	(0.50)	(0.02)	(0.50)	(0.03)	(0.50)	(0.04)
No of minors	1.69	-0.026	1.79	0.075	1.73	-0.097
	(1.11)	(0.09)	(1.12)	(0.12)	(1.12)	(0.12)
Own motorcycle	0.42	0.008	0.28	-0.035	0.42	0.003
	(0.49)	(0.04)	(0.45)	(0.05)	(0.49)	(0.05)
Own car/truck	0.16	0.017	0.04	0.01	0.13	0.028
	(0.37)	(0.03)	(0.19)	(0.03)	(0.34)	(0.04)
Own oxen/buffalo	0.55	0.032	0.39	0.109*	0.53	0.033
	(0.50)	(0.05)	(0.49)	(0.06)	(0.50)	(0.06)
Own pig	0.56	0.028	0.43	0.117**	0.55	0.029
	(0.50)	(0.04)	(0.50)	(0.06)	(0.50)	(0.05)
Own ox or buffalo cart	0.31	0.02	0.19	0.058	0.29	0.009
	(0.46)	(0.04)	(0.40)	(0.05)	(0.45)	(0.05)
Hard roof	0.49	0.064	0.32	0.047	0.48	0.102**
	(0.50)	(0.04)	(0.47)	(0.05)	(0.50)	(0.05)
Hard wall	0.54	0.032	0.38	0.045	0.55	0.018
	(0.50)	(0.04)	(0.49)	(0.06)	(0.50)	(0.05)
Hard floor	0.85	0.039	0.79	0.037	0.84	0.068*
	(0.36)	(0.03)	(0.41)	(0.05)	(0.37)	(0.04)
Have automatic toilet	0.07	-0.02	0.02	-0.01	0.05	0.005
	(0.25)	(0.02)	(0.13)	(0.01)	(0.22)	(0.02)
Have pit toilet	0.12	0.018	0.11	0.02	0.13	0.001
	(0.32)	(0.03)	(0.32)	(0.03)	(0.34)	(0.04)
Electricity	0.25	0.011	0.16	-0.01	0.23	-0.002
	(0.43)	(0.04)	(0.37)	(0.04)	(0.42)	(0.05)
Pipe water	0.06	-0.001	0.03	-0.012	0.06	-0.013
	(0.24)	(0.02)	(0.17)	(0.01)	(0.23)	(0.02)
Poverty Index (o to 292)	210.16	-1.609	245.13	-2.924	218.2	-11.771
	(60.18)	(5.43)	(32.73)	(5.14)	(51.66)	(8.76)
Test score (0 to 25)	17.47	0.534	17.74	0.888	19.77	0.028
	(4.81)	(0.52)	(4.71)	(0.68)	(3.22)	(0.48)
Number of students	940	2448	431	883	474	940
Number of schools	101	204	67	119	67	118

Columns (1), (3), (5): means and standard deviation, control group. Columns (2), (4), (6): difference with treatment, estimated by regressing each variable against corresponding treatment variable; standard error in parenthesis

Table 2. Outcome variables at follow-up and simple differences

	Student Level			
	Poverty-targeting		Merit-targeting	
	Control (1)	Treatment-Control (2)	Control (1)	Treatment-Control (2)
Reach grade 6	0.61	0.186***	0.64	0.131***
	(0.49)	(0.03)	(0.48)	(0.03)
Completed grades	5.38	0.349***	5.45	0.234***
	(1.22)	(0.08)	(1.23)	(0.07)
Number of hours	8.83	3.466***	9.27	1.374
	(12.97)	(1.09)	(13.22)	(1.04)
Math test	0.02	-0.035	0.16	0.150**
	(1.01)	(0.07)	(1.04)	(0.07)
Digitspan	0.02	-0.019	0.08	0.147**
	(0.98)	(0.07)	(0.99)	(0.06)
Number of observations	431	883	474	940

(1) Mean and () std deviation; (2) Simple mean differences and () std error

Table 3. Impact on Enrollment and Attendance

	Reach Grade Six		Highest Grade Completed		Num. of hours in school, last 7 days (conditional on enrollment)	
	(1)	(2)	(1)	(2)	(1)	(2)
Poverty-targeting treatment	0.170***		0.332***		2.865	
	(0.04)		(0.11)		(1.87)	
Merit-targeting treatment		0.120***		0.182*		0.635
		(0.04)		(0.10)		(1.55)
Constant	1.764	0.514	8.161**	5.285**	-100.065**	0.575
	(1.42)	(0.83)	(3.35)	(2.33)	(45.16)	(28.25)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
No. Obs	883	940	831	897	665	713
F()	6.435	4.872	2.271	1.759	1.246	2.026
R2 Adj	0.18	0.155	0.145	0.122	0.131	0.199

Regression coefficient of dependent variable against treatment indicator controlling for Table 1 baseline characteristics and district fixed effects. All estimators using clusters at school level.

Table 4. Impact on test scores

	Mathematics		Digitspan Test	
	(1)	(2)	(1)	(2)
Poverty-targeting treatment	-0.041		-0.059	
	(0.08)		(0.07)	
Merit-targeting treatment		0.170*		0.149**
		(0.09)		(0.08)
Constant	-2.394	1.331	-2.055	-0.139
	(2.38)	(2.03)	(3.11)	(2.18)
Control	Yes	Yes	Yes	Yes
No. Obs	883	940	883	940
Chi_ 2	3.297	3.213	2.727	2.186
R2 Adj	0.167	0.16	0.113	0.093

Regression coefficient of dependent variable against treatment indicator controlling for Table 1 baseline characteristics and district fixed effects. All estimators using clusters at school level.

Table 5. Heterogeneity impacts on test scores by baseline test score and poverty index

	Mathematics		Digitspan Test	
	(1)	(2)	(1)	(2)
Poverty-targeting treatment	-0.044		0.003	
	(0.10)		(0.12)	
Above median baseline test score	0.232**		0.153	
	(0.10)		(0.11)	
Poverty Treatment * Above median test score	-0.002		-0.105	
	(0.12)		(0.14)	
Merit-targeting treatment		0.052		0.192*
		(0.12)		(0.11)
Above median baseline poverty index		-0.079		0.031
		(0.11)		(0.12)
Merit Treatment * Above median poverty index		0.216		-0.079
		(0.14)		(0.14)
Constant	-2.196	1.218	-1.977	-0.098
	(2.45)	(2.01)	(3.14)	(2.20)
Control	Yes	Yes	Yes	Yes
No. Obs	883	940	883	940
Chi_ 2	3.405	3.326	2.709	2.044
R2 Adj	0.175	0.162	0.115	0.094

Regression coefficient of dependent variable against treatment indicator controlling for Table 1 baseline characteristics and district fixed effects. All estimators using clusters at school level.

Table 6. Potential mechanisms for impact

	Student and Household (hh) effort			
	Time doing homework, studying or private lessons		Total yearly expenditure in education, hh (US $)	
	(5)	(6)	(7)	(8)
Poverty-targeting treatment	0.42		-0.726	
	(0.38)		(2.86)	
Merit-targeting treatment		0.700**		5.388*
		(0.33)		(3.11)
Constant	-16.841	-2.764	25.004	360.193*
	(13.14)	(9.96)	(51.78)	(201.67)
Control Variables	Yes	Yes	Yes	Yes
No. Obs	599	657	665	713
F()	7.282	3.401	4.895	2.585
R2 Adj	0.145	0.176	0.138	0.157

Regression coefficient of dependent variable against treatment indicator controlling for Table 1 baseline characteristics and district fixed effects. All estimators using clusters at school level.

Table 7. Impact on Enrollment and Attendance on Non-treated peers

	Reach Grade Six		Highest Grade Completed		Num. of hours in school, last 7 days (conditional on enrollment)	
	(1)	(2)	(1)	(2)	(1)	(2)
Poverty-targeting treatment	0.082**		0.058		0.529	
	(0.04)		(0.11)		(1.82)	
Merit-targeting treatment		-0.009		-0.099		0.291
		(0.05)		(0.12)		(1.67)
Constant	0.168	2.201**	4.030**	8.599***	12.924	14.213
	(0.83)	(0.94)	(1.92)	(1.81)	(29.16)	(28.46)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
No. Obs	785	678	732	633	576	486
F()	7.528	6.94	1.603	1.765	2.166	3.195
R2 Adj	0.172	0.183	0.125	0.118	0.191	0.21

Regression coefficient of dependent variable against treatment indicator controlling for Table 1 baseline characteristics and district fixed effects. All estimators using clusters at school level.

Table 8. Impact on test scores on Non-treated Peers

	Mathematics		Digitspan Test	
	(1)	(2)	(1)	(2)
Poverty-targeting treatment	-0.105		-0.009	
	(0.08)		(0.08)	
Merit-targeting treatment		-0.061		-0.048
		(0.08)		(0.08)
Constant	-1.172	-0.705	-0.947	-0.67
	(1.40)	(1.29)	(1.40)	(1.76)
Control Variables	Yes	Yes	Yes	Yes
No. Obs	785	678	785	678
F()	2.119	3.808	2.233	2.691
R2 Adj	0.153	0.129	0.115	0.131

Regression coefficient of dependent variable against treatment indicator controlling for Table 1 baseline characteristics and district fixed effects. All estimators using clusters at school level.