Preliminary and Incomplete: Please do not cite or circulate without authors' permission

The Aggregate Effect of School Choice: Evidence from a two-stage experiment in India

Karthik Muralidharan[†]
Michael Kremer
Venkatesh Sundararaman

November 11, 2012

Abstract: We present experimental evidence on the impact of a program that provided economically disadvantaged children in the Indian state of Andhra Pradesh (AP) with a scholarship that allowed them to attend a private school of their choice. The design featured a unique two-stage lottery that creates both a student-level as well as a market-level experiment that allows us to study both the individual and aggregate effects of school choice. Compared to teachers in government-run schools, private school teachers are paid much lower salaries and have lower levels of formal education and training. The mean annual cost per student in the private schools in our sample is less than a third of the costs in public schools. On the other hand, private schools have a longer school day, a longer school year, smaller class sizes, lower teacher absence, higher teaching activity, and better school hygiene. Private schools spend significantly less instructional time on Telugu (native language of AP) and Math, and instead spend more time on English, Science, Social Studies, and especially Hindi. At the end of four years of the school choice program, we find that lottery winners do not have higher test scores than lottery losers on tests of Telugu, Math, English, Science, and Social Studies, but score significantly higher in Hindi. There is evidence of heterogeneity of impact by medium of instruction in the private school, but not by most other demographic characteristics. We find some evidence suggesting that the impact of the voucher may have been greater in areas with more choice and competition. We find no evidence of significant spillovers on students who do not apply for the voucher and remain in the government schools or on students who start out in private schools to begin with, suggesting that the program had no adverse effects on these groups.

JEL Classification: C93, I21, M52, O15

Keywords: school vouchers, school choice, private schools, India, education, field experiments

---

[†] Karthik Muralidharan: UC San Diego, NBER, BREAD, J-PAL; E-mail: kamurali@ucsd.edu
Michael Kremer: Harvard, NBER, BREAD, J-PAL; E-mail: mkremer@fas.harvard.edu
Venkatesh Sundararaman: World Bank; E-mail: vsundararaman@worldbank.org

We thank Julian Betts, Prashant Bharadwaj, Gordon Dahl, Jishnu Das, Esther Duflo, Mark Jacobsen and several seminar participants for comments.

This paper is based on the Andhra Pradesh School Choice Project implemented by the Azim Premji Foundation as part of a Memorandum of Understanding with the Government of Andhra Pradesh. The majority of the funding for the project was provided by the Legatum Institute, with additional financial support from the UK Department for International Development (DFID), and the World Bank's Strategic Impact Evaluation Fund (SIEF). Muralidharan also acknowledges financial support from a National Academy of Education/Spencer Foundation post-doctoral fellowship. We are especially grateful to M Srinivasa Rao, B Srinivasulu, S Ramamurthy, and staff of the Azim Premji Foundation for their outstanding efforts in implementing the project in Andhra Pradesh. We thank Gautam Bastian, Vikram Jambulapati, Naveen Mandava, Jayash Paudel, and Arman Rezaae for excellent research assistance at various stages of the project. The findings, interpretations, and conclusions expressed in this paper are those of the authors and do not necessarily represent the views of the Government of Andhra Pradesh, the Azim Premji Foundation, or the World Bank.

1. Introduction

Economists have long hypothesized that improved choice for parents among schools and competition between schools to attract students would improve the performance of education systems (Friedman 1955, 1962). The theoretical promise of increased choice and competition has in turn generated a large empirical literature trying to measure the impacts of school choice on education outcomes, with the best-identified studies typically comparing winners and losers of lotteries that are used to determine access to over-subscribed programs. Such lottery-based designs have been used to study school voucher programs (Angrist et al. 2002, 2006; Mayer et al. 2002; Krueger and Zhu 2004; Howell and Peterson 2004), the impact of more selective schools (Cullen et al 2005; Zhang 2009; Lucas and Mbiti 2012; Pop-Eleches and Urquiola 2012; Angrist et al 2012), and more recently charter school programs (Hoxby et al 2009; Angrist et al 2011).[1] The results to date are quite mixed with most studies typically finding no or only modest positive effects of receiving a voucher or attending a more selective school on test scores (Rouse and Barrow 2009), though the results on charter schools are considerably more promising.

Nevertheless, the existing empirical literature on school choice has some important limitations. First, the comparison group in most experimental studies of school choice (consisting of losers of lotteries to attend more selective schools) is usually contaminated because school resources typically do not adjust proportionally with the departure of lottery winners, and there may be behavioral responses of schools and teachers to the possible and actual departure of lottery winners. Second, the research to date on the effects of both voucher programs and charter schools typically does not account for the fact that instructional programs and time allocation across subjects vary considerably across schools. Third, existing experimental studies are unable to study the effects of school competition on education outcomes since they are typically student-level as opposed to market-level experiments.[2] Finally, a key

[1] It is important to note that the likely mechanisms of impact are different across these three types of programs. In particular studies that evaluate the impact of going to a "better" school (typically defined in terms of observed outcomes) are typically not evaluations of school choice. But these studies are still relevant to the school choice literature, because one of the key mechanisms by which school choice is posited to work is that students can transfer from low-performing to high-performing schools thereby contributing to an expansion of market-share of good schools and a reduction in that of weak schools. However, if the observed cross-sectional differences in outcomes between 'good' and 'poor' schools are mostly driven by selection and unobservables (as opposed to school effectiveness), then the empirical case for school choice is less compelling.

[2] Andrabi, Das, and Khwaja (2012) is a recent exception that experimentally studies the market impact of providing more information on school performance in villages in Pakistan. Empirical studies of the impacts of competition have typically relied on natural experiments (Hoxby 2000; Lavy 2010).

limitation in the literature is the inability to study the spillover effects of voucher programs on students who do not apply for vouchers, and students who are in private schools to begin with.

In this paper, we present results from an experimental evaluation of a school choice program in the Indian state of Andhra Pradesh (AP) that was designed to address these gaps in the global literature on school choice. The Andhra Pradesh School Choice Project provided economically disadvantaged children who were enrolled in free government-run primary schools with a voucher that allowed them to attend a private school of their choice. The project design featured a unique two-stage lottery that creates both a student-level as well as a market-level experiment that allows us to study both the individual and aggregate effects of school choice. The two-stage randomization of the offer of a voucher (across villages and students) allows us to estimate the impact of the voucher on lottery-winners, relative to lottery losers in *control villages* who do not experience any changes in their schooling experiences as a result of the voucher program (because there are no lottery winners in control villages), thus creating an uncontaminated comparison group. Second, our detailed data on school time tables (and collecting test score outcomes to reflect the patterns of time allocation) highlight the importance of incorporating this data into studies of school choice and charter schools. Third, the market-level experiment allows us to study heterogeneous effects of school choice as a function of initial levels of school competition. Finally, the market-level experimental design also allows us to study the aggregate effects of such a program by comparing outcomes for non-applicants as well as students who start out in private schools across treatment and control villages.

The typical private school in our sample is a low-cost private school – with per student spending that is around 30% of the per-student spending in the government schooling system. The value of the voucher was set near the 90th percentile of the distribution of private school fees in the sampled villages, but was still only around 40% of the per-child spending in the government schools. The main operating difference between private and public schools in rural AP (and rural India in general) is that private schools pay substantially lower teacher salaries (less than a sixth of that paid to government school teachers), and hire teachers who are younger, less educated, and much less likely to have professional teaching credentials. However, they hire more teachers and have smaller class sizes and less multi-grade teaching than public schools.[3]

[3] This is true both in our sample used for this paper, as well as other representative all-India samples such as those described in Muralidharan and Kremer 2008, and Desai et al (2009). Descriptive data on private versus

Using official data as well as data collected from direct observations conducted during unannounced visits to schools, we find that private schools have a longer school day, a longer school year, lower teacher absence, higher teaching activity, and better school hygiene. However, in spite of the superior performance of the private schools on most measures of school processes, we find at the end of two and four years of the school choice program that lottery winners do no better than lottery losers on tests of Telugu (native language of AP) and Math, which are the two main subjects in the primary school curriculum of AP.

Our data from school time tables suggest that a possible explanation for our results is that private schools spend significantly less instructional time on Telugu and Math, and instead spend more time on English, Science, Social Studies, and Hindi. Private schools spend around 200 minutes less on Telugu and 160 minutes less on Math per week than government schools (40% and 32% less instructional time respectively), but spend 60 to 100 minutes extra per week on Science, Social Studies, and English. A striking finding is that they also spend 215 minutes per week extra on Hindi (the national language which is not taught in government schools in AP, where the native language is Telugu).

We conduct tests in these subjects at the end of four years of the program and find positive (but insignificant) effects of winning the voucher on test scores in English, Science, and Social Studies (of around 0.1 standard deviation each), and positive (and highly significant) effects on test scores in Hindi (of 0.5 standard deviations). There is considerable evidence to suggest that labor market returns to knowing English in India are positive (holding education constant),[4] but in the absence of long-term follow up data on wages and labor market outcomes, there is no obvious way to weight the outcomes on these different subjects. However, even without a basis for weighting across subjects, our results imply that private schools are more productive than government-run schools because they are able to achieve similar Telugu and Math test scores for

---

government-run schools in specific locations in India are also provided by Tooley and Dixon (2007), Kingdon (2008), Tooley (2009), and Rangaraju et al. (2012).

[4] Munshi and Rosenzweig (2006), Kapur and Chakraborty (2008), Azam, Chin, and Prakash (2011), and Shastry (2012) all find significant positive labor market returns in India to knowledge of English. We are not aware of any corresponding study of the returns to speaking Hindi for non-native speakers. Nevertheless, there may be reasons to believe that these returns are positive since they facilitate interactions with a much larger fraction of the Indian population, and provide access to a greater set of labor market opportunities. It is likely therefore, that the private schools are being more responsive to market demand and shifting instructional time towards subjects with greater market returns (Clingingsmith 2011 shows that increases in industrialization in India led to increases in investments in bilingualism, suggesting positive returns to bilingualism in a growing economy).

the lottery winners as compared to the lottery losers with substantially less instructional time, and use the additional time to improve outcomes on other subjects - especially Hindi.

We find no significant heterogeneity in program impact by student gender, socio-economic status, or baseline test scores, though Muslim recipients of the voucher do significantly better. An important dimension of heterogeneity among private schools in AP is medium of instruction with around a third of voucher winners attending an English medium private school and two-thirds attending a Telugu medium private school. We find using instrumental variable estimates that students attending Telugu medium private schools do better than those attending government schools in *all* subjects, whereas those attending English medium private schools do worse than those in government schools in Telugu, Math, Science and Social Studies, while doing sharply better in English and Hindi.[5] These results suggest that private schools were probably even more effective across the board when students did not also experience a change in the medium of instruction, and that switching to English medium private schools may have important costs for literacy in the native language and also in numeracy and the learning of other subjects.[6]

We find that voucher winners do significantly better when there are four or more schools within a half kilometer distance from their homes or when there are six or more schools within a one kilometer radius. One caveat is that we find no significant effect of competition when estimated with a linear interaction between voucher receipt and number of schools, and that the estimated effects are only significant above the 90th percentile of the size of the choice set in this experiment. However, a limitation of conducting the study in a rural sample is that over 60% of voucher applicants have only 0 (40%) or 1 (21%) private school within a half kilometer radius and nearly 50% have only 0 (27%) or 1 (21%) private school within a kilometer radius. Urban India however has much greater population and school density and a recent census of schools (with geo-coding) in the city of Patna found that there are between 9 and 93 private schools within a one kilometer radius of every government school, with the median being greater than 50

[5] Since the medium of instruction of the school attended is a choice variable, we instrument for medium of instruction of the school attended with the interaction of receiving the voucher and the medium of instruction of the nearest private school.

[6] Abadzi (2006) summarizes research from the cognitive neuroscience literature and strongly recommends that instruction for the poor in developing countries be conducted in the native language. Of course, the results above are the composite effect of English medium instruction in the school and sharply lower instructional time in Telugu as a subject, and so the effects should be interpreted mainly in contrast to the Telugu medium private schools (whose instructional time allocations do not differ much from those of English medium private schools).

(Rangaraju et al 2012). Our results therefore suggest that the effects of choice and competition may be considerably larger in such a context.

A key limitation in interpreting existing studies on school choice is the concern that students who are left behind in public schools will be worse off after the departure of their more motivated peers (who apply for the lotteries and transfer to private schools). A related concern is that measured gains (if any) in outcomes for voucher winners may be coming at the cost of students who started out in private schools to begin with and are now exposed to lower-achieving peers in their classrooms. We find no evidence of significant spillovers on students who do not apply for the voucher or students who start out in private schools to begin with, suggesting that there were no adverse peer effects on these groups. We also do not find any significant difference in the test scores of students who applied for the lottery and were not awarded a voucher in the comparison villages (the uncontaminated control group), and those who lost the lottery in the program villages. Thus, even though we use the "correct" control group for our estimates, using the typical comparison group as used in most experimental studies would not have significantly altered our results.

The combination of test score results and school time table data already suggest that private schools are more productive than private schools, but the comparison is rendered stark by the fact that the annual cost per student in the government-school system is over three times the mean cost per student in the private schools in our sample. Thus, students who win a lottery to attend private schools do as well on some subjects and better on others even though the private school is spending substantially lower amounts per student.

These results have direct policy implications given the recent Right to Education (RtE) Act passed by the Indian parliament, which includes a provision mandating that private schools reserve up to 25% of the seats in their school for students from disadvantaged backgrounds, with a reimbursement of fees by the government (subject to a maximum of the per-child spending in the public schools). This provision was motivated by a desire to reduce social stratification across economic lines between schools in India, but if implemented as per the letter of the law, the RtE Act could lead to India having the world's largest number of children attending private schools with public funding. While the law was passed before any evidence on its likely impacts was available, our results suggest that this provision could lead to significant gains in the cost effectiveness of human capital acquisition in India, with limited negative spillovers to children

who are already attending private schools, or on children who choose not to avail of the opportunity to apply to transfer to a private school. The main caveat suggested by our results is the need for more research on and advice to parents regarding the trade-offs associated with switching first generation learners into English medium schools.

More broadly, our results are consistent with those in Bloom & Van Reenen (2010), who find that privately-managed firms in developing countries are considerably more productive than publicly-managed ones in the same sector, and highlight the potential to leverage private management in education production for human capital formation in India and other developing countries. Our results also highlight the importance of recognizing that schools are vectors of attributes and show that evaluating school choice and charter school programs on a limited set of test scores (typically in math and reading) may provide an incomplete picture of the impact of such programs.

The rest of this paper is structured as follows: Section 2 describes the AP School Choice Experiment; section 3 discusses the data and attrition; section 4 presents results on summary statistics of school, teacher, and household inputs into education; section 5 presents the main results and section 6 concludes.

2. The Andhra Pradesh (AP) School Choice Experiment

2.1 Background and Context

India has the largest school education system in the world comprising around 200 million children. Primary school enrollments have steadily increased over the past two decades and over 96% of primary-school aged children are now enrolled in school (ASER 2011). Nevertheless education quality is low with less than 40% of children aged 6 to 14 being able to read at the second grade level.

A prominent trend in India has been that parents are enrolling their children in fee-charging private schools in increasing numbers. Over 25% of children between the ages of 6 and 14 in *rural* India attend private schools (ASER 2011), with the corresponding fraction in urban India being over 50% (Desai et al. 2009).[7] The majority of these private schools are low-cost or 'budget' private schools that cater to non-affluent sections of the population, and have per-student spending that is significantly lower than that in public schools (Tooley 2009). However,

[7] The annual time-series data provided by the ASER reports show an increasing private school share in rural India (the urban trends are likely to be similar though there is no corresponding annual time-series available).

since private schools charge fees and public schools are free, students attending private schools on average come from more affluent households with higher levels of parental education (Muralidharan and Kremer 2008; see also Table 1). Cross-sectional evidence finds that students in private schools significantly outperform their counterparts in public schools, even after correcting for observable differences between the characteristics of students attending the two types of schools (Muralidharan and Kremer 2008; Desai et al. 2009; French and Kingdon 2010). Nevertheless, these studies cannot fully address selection and omitted variable concerns with respect to identifying the causal impact of attending a private school.

The growing popularity of private schools has led to concerns about increasing economic and social stratification in education, leading to calls for expanding access to private schools for all children, regardless of socioeconomic background – including experimenting with voucher-based school choice programs. The recent Right to Education (RtE) Act passed by the Indian parliament includes a provision mandating that private schools reserve up to 25% of the seats in their school for students from disadvantaged backgrounds, with a reimbursement of fees by the government (subject to a maximum of the per-child spending in the public schools). While the specific implementation details have not yet been fully specified, the allocation of these places is likely to be based on a combination of location of residence and a lottery.[8] If implemented as per the letter of the law, this provision in the RtE Act could lead to India having the world's largest number of children attending private schools with public funding.

2.2 Conceptual Overview of Experiment Design

Figure 1 (Panel A) shows the typical design used in experimental evaluations of voucher programs around the world. The key feature of this design is that a limited number of vouchers are offered that enable students currently enrolled in public schools to defray the costs of attending a private school. The program is typically oversubscribed and the limited slots are allocated by lottery. Such a program design creates four groups of students: those who do not apply for the voucher (group 1), those who apply and lose the lottery (group 2), those who apply and win the lottery (group 3) and those who were in private schools to begin with (group 4). The

[8] The initial draft of the RtE Act that was distributed for comments in 2005 (on the basis of which this study was designed) envisaged an allocation mechanism based purely on a lottery. The final draft that was passed in 2009 introduced location as a criterion for the allocation of places in private schools under the "Economically Weaker Sections (EWS)" category. The specific rules under which the 25% reservation provision will be implemented have been left up to individual states to determine, and while there is uncertainty with respect to the final allocation rules that will be adopted, it is likely to involve a combination of residential location and a lottery.

best studies to date on school choice estimate the impact of winning the lottery conditional on applying for it (i.e. they compare groups 3 and 2). The lottery allows researchers to estimate both the impact of winning the lottery (the 'intention to treat' effect) and the impact of attending a private school (using the lottery as an instrumental variable for attending a private school - the 'treatment on treated' estimate).

However, even an experimental design of this sort (while better than the alternatives) ignores the potential spillover effects of the voucher program on the losers of the lottery. Thus, the departure of group 3 students may have additional effects on students in group 2. Some possible mechanisms include changing of the peer group (because motivated students may have left),[9] changes in per-student resources (for instance, class sizes may be smaller after some students leave because the teacher allocation is not proportionately reduced), and changes in behavior by public school teachers in response to the voucher program (such as a competitive response to improve quality and keep children from leaving the government schools). These confounding factors may bias a simple comparison between groups 2 and 3. In other words, the "control" group even in experimental studies is not truly a "business as usual" control group because of potentially unobserved spillover effects, and even the internal validity of the estimates from the literature to date can be questioned on this basis.

Moreover, existing studies typically cannot estimate the program's impact on students in group 1 (who did not apply for the voucher and who are subject to similar spillovers as group 2) or students in group 4 (who may be worse off because of an influx of low-performing students from public schools). Thus, even if group 3 is doing better than group 2 (which is what the traditional experimental studies focus on), this may have come at the cost of poorer performance for groups 1 and 4. Thus, a critical open question in the global literature on vouchers and school choice is that of the "aggregate impact" of such programs (Hsieh and Urquiola 2005).

The AP School Choice Experiment aims to address both these issues by employing a two-stage randomization design, where we first use a lottery to assign entire villages into control and treatment groups (where treatment villages participate in the scholarship[10] program while control

[9] While the vouchers are offered by lottery, not all winners will typically accept it and move to a private school. It is possible that the most motivated students may be the ones who accept the voucher to go to a private school.

[10] We use the term scholarship and voucher interchangeably in this paper. The program as implemented used the term “scholarship” because the term voucher is not well known in India, whereas the idea of a scholarship as an instrument that defrays the costs of education is well understood. However, there was no component of the

villages have no students who receive the scholarship), and then conduct a second lottery to assign scholarships to applicants in the treatment villages. Since children typically do not travel far beyond their own villages for primary school, villages in rural India can be thought of as approximating 'closed economies' for school choice.[11] Thus, comparing the aggregate outcomes by treatment and control villages allows us to estimate the aggregate impact of school choice programs in rural India.

Figure 1 (Panel B) presents the conceptual overview of the experiment design. The key innovation in this design is that the control *villages* provide a 'system-level' counterfactual to the scholarship program and hence provide the kind of control group that has not typically been found in the literature (villages are randomized into treatment and control status after baseline tests are conducted and after parents apply for the voucher). Now, comparing the recipients of the voucher (3T) with applicants in control *villages* (2C) will provide an experimental estimate of the impact of the choice program *without being contaminated by the spillovers*. In other words, group 2C represents the "true" control group because they have applied for the scholarship and lost the lottery (at the village level), but nothing else has changed for them because there is no scholarship programs in these villages.

The design also lets us to do 3 additional comparisons, which have not been possible in the literature to date. First, comparing groups 2T ('control' students with spillovers) and 2C ('control' students without spillovers), will provide a sense of the extent to which ignoring spillovers may bias the estimates existing studies. Second, the comparison between groups 1T and 1C will let us estimate the impact of school choice programs on the children 'left behind' (who for reasons of limited information or motivation choose to not apply for the voucher). Third and finally, comparing outcomes between groups 4T and 4C will provide an estimate of whether students in private schools are adversely affected by an influx of students from the government school (which is exactly what will happen if the provision in the RtE Act regarding reserving 25% of places in private schools for disadvantaged students is implemented).

---

scholarship that depended on student performance, and so the 'scholarship' had no merit component and was equivalent to a voucher (the specific features of the program are explained in the next section).

[11] Econometrically, it is not a problem even if this is not fully true. As long as children are much more likely to go to school in their own village, the design provides a strong 'first stage' for the village-level lottery to be used as a valid instrumental variable for village-level school choice.

2.3 The AP School Choice Experiment

Andhra Pradesh (AP) is the 5th most populous state in India, with a population of over 80 million (70% rural). Recent estimates suggest that over 30% of students in rural AP are enrolled in private schools (ASER 2011), compared to an all India average of around 25%. The project that this paper is based on is called the Andhra Pradesh School Choice (APSC) Project and was implemented in the Indian state of Andhra Pradesh by the Azim Premji Foundation (one of India's leading non-profits working on education).[12] The academic year in AP runs from mid-June to mid-April. The AP School Choice project started in the academic year 2008-09, with preparatory work starting in early 2008.

The project was carried out in five districts across AP over a universe of 180 villages that had at least one recognized private school.[13] Baseline tests were conducted for *all* students in 2 cohorts of *all schools* (public and private) in these villages in March-April 2008.[14] This was followed by an invitation to apply for a voucher to parents of students in government schools (who had taken the baseline test) in all 180 villages. The application specified the full terms of the voucher including the fact that it would be allocated by lottery and that applying did not guarantee receipt of the voucher. The communication regarding the scholarship program and the application process was done by field staff of the Azim Premji Foundation during the summer break in May 2008.

Participation of both households and schools was completely voluntary. Households were told that they could go back to the government school at any time and there were no terms and conditions for participation beyond consent for answering surveys and taking tests. The scholarship covered all school fees, textbooks, workbooks, notebooks and stationery, and school uniforms and shoes, but did not cover transport costs to attend a private school outside the village and did not provide any supplement in lieu of the free mid-day meals that the government

[12] The AP School Choice Project was carried out under the larger program of the "Andhra Pradesh Randomized Evaluation Studies (AP RESt)" which was set up as a research partnership between the Government of Andhra Pradesh, the Azim Premji Foundation, and the World Bank.

[13] These were the same districts as in the overall AP RESt project (Muralidharan and Sundararaman 2010, 2011, 2012), but the AP School Choice Project was conducted in different sub-districts and so there was no overlap in the schools/villages across these studies.

[14] The cohorts covered were students attending kindergarten and grade 1 in the school year 2007-08, and the voucher covered the entire primary education of recipients (from grade 1 to 5 for the younger cohort and from grade 2 to 5 for the older cohort).

schools provide. The value of the voucher was paid directly to the school, and the materials were provided directly to the voucher households by the schools.[15]

At the same time as the baseline tests, the Azim Premji Foundation (the Foundation) also solicited participation in the project from private schools in the sample villages, and school participation was completely voluntary. The value of the voucher was set at the 90th percentile of the distribution of the all-inclusive private school fees in the sampled villages, and schools were asked to indicate if (a) they wanted to participate in the program by being willing to admit economically disadvantaged students who would be awarded a voucher by the Foundation, and (b) if so, how many seats they could make available to scholarship students in each of the two cohorts.[16] The terms and conditions specified that the Foundation would directly pay the value of the voucher to the school's bank account (in three annual installments – which was the typical fee cycle of the schools). The only condition imposed on the schools was that they were not allowed to cherry pick students. If there was greater demand for a school than the number of places offered, then the school could either admit all voucher recipients who wanted to attend the concerned school or the Foundation would conduct a lottery to allocate the places among the applicants.

All communications with schools (and elicitation of willingness to participate) was conducted before the village-level randomization took place.[17] Once the applications were completed, 90 villages (stratified by district) were assigned by lottery to be voucher villages (Figure 1 - Panel A), while the other 90 villages continued "as usual" with no voucher program (Figure 1 – Panel B). Conditional on being a "voucher village", a second lottery was conducted to offer the vouchers to a subset of applicants. The design therefore creates two lottery-based comparison groups – those who did not get the voucher due their village not being selected for

[15] This was consistent with standard practice we observed in the field. The private schools had a recommended set of books, uniforms etc. which they procured in bulk and supplied to parents for a fixed fee. It was therefore easiest to have the voucher cover these payments directly as opposed to making cash payments to parents for other incidental education expenses.

[16] At the time of starting the project, the 2005 draft of the Right to Education (RtE) Act was already in circulation and private schools knew that the stipulation regarding reserving 25% of seats for economically disadvantaged children was likely to be implemented

[17] The initial frame for the project was 200 villages, which was reduced to 180 after dropping villages where there was no private school willing to participate, or where the private schools did not obtain recognition at the start of the 2008-09 school year (the sample initially included villages with unrecognized schools that said that they were in the process of getting recognized, but villages where there was no school that had obtained recognition were dropped from the study universe). This was done because the Foundation did not want to put voucher winning children in a situation where the school they went to would be shut down by the government (as the law entitles them to do).

the program (group 2C in Figure 1), and those who did not get the voucher due to losing the individual level lottery conducted within voucher villages (group 2T in Figure 1).

Out of 10,935 eligible households, a total of 6,433 households applied for the voucher (59%). A total of 3,097 households had applied in the treatment villages, from which 1,980 were selected by lottery to receive the voucher (64%). 1,210 of these 1,980 households accepted the voucher and enrolled in a private school at the start of the project (61%). At the end of four years of the project, a total of 1,005 students continued to avail of the voucher. The intent to treat estimates will therefore be based on a net take up rate of 51% (1,005/1,980). Figure 2 shows the program design with the actual number of students in each of the cells.

Appendix Table 1 shows the correlates of application for the voucher, acceptance conditional on being awarded one, and Application rates are not correlated with observable demographic characteristics like parental income, education, or caste (Table A1). The only observables that are correlated with application are having a sibling in the government school (negative) and having a private school within a radius of half a kilometer (positive), which are as expected. The same patterns are observed in acceptance conditional on being awarded the voucher. Thus, while it is possible that the decision to apply and/or to accept may be driven by unobserved household characteristics, we do not see any correlation between household socio-economic characteristics and voucher application or acceptance.

The allocation of villages and students to the voucher program by lottery ensured that the treatment groups and the corresponding comparison groups are not significantly different on observable characteristics including baseline test scores, parental education, assets, and caste. Table 2 - Panel A shows the balance between lottery winners and losers – first showing the comparison with lottery losers in the treatment villages and then showing it with lottery losers in control villages. Panel B shows the balance for the groups of students who will be used for the spillover analysis – first showing the comparison between non-applicants across treatment and control villages, and then showing it between students who start out in private schools across these villages.

To be Completed

Figure 1: Design of AP School Choice Program

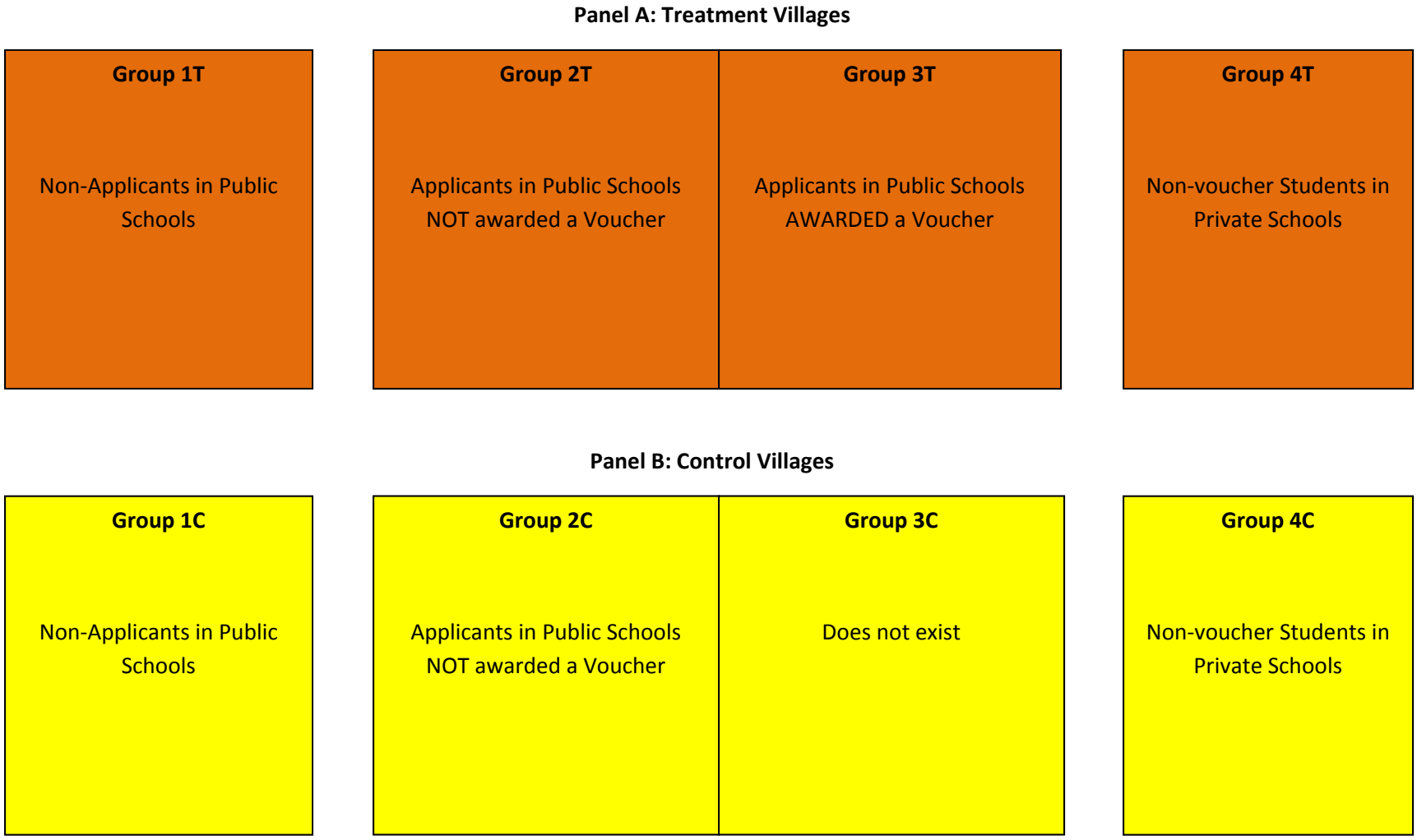
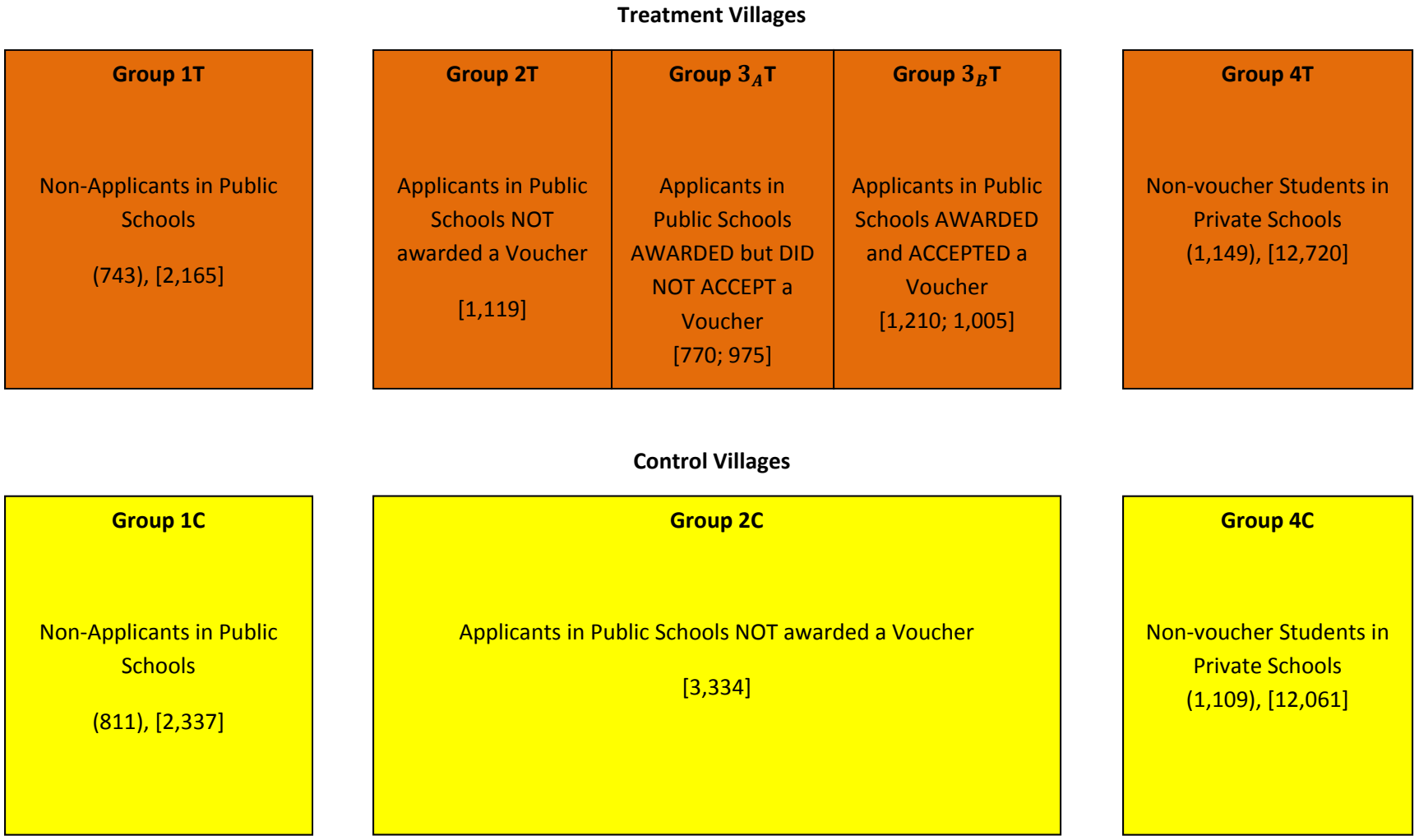


Figure 2: Design of AP School Choice Program with Student Counts



Notes: All of groups 2T, 2T, and 2C were sampled. For other groups, numbers in parentheses are the sample size that was tracked (with the total population in brackets). The two numbers under group 3_BT represent those who first accepted and started in a private school (1210) and those who were still in a private school at the end of 4 years (1,005). Conversely in group 3_AT, 770 initially rejected the offer, while 975 were no longer availing the voucher at the end of 4 years

Figure 3 (a): Heterogeneous Y4 Test Score Impacts by Baseline Percentile - Telugu

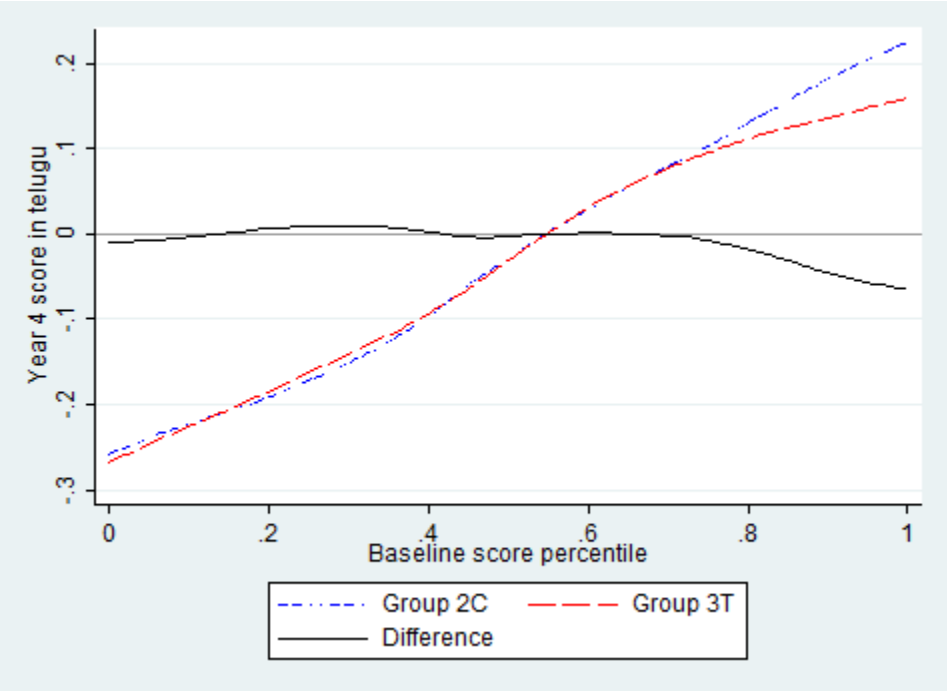


Figure 3 (b): Heterogeneous Y4 Test Score Impacts by Baseline Percentile - Math

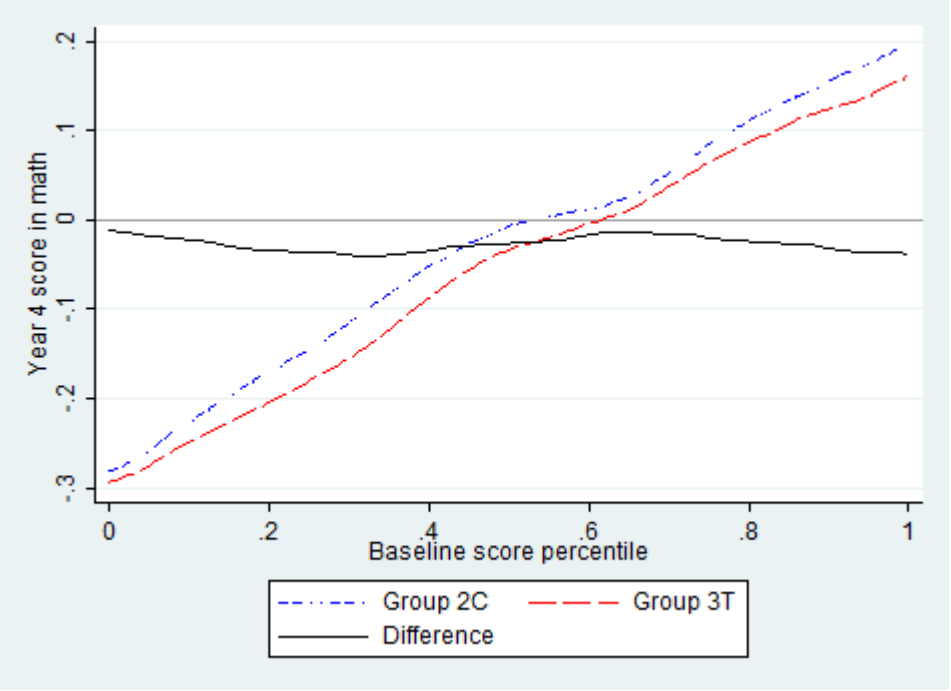


Figure 3 (c): Heterogeneous Y4 Test Score Impacts by Baseline Percentile - English

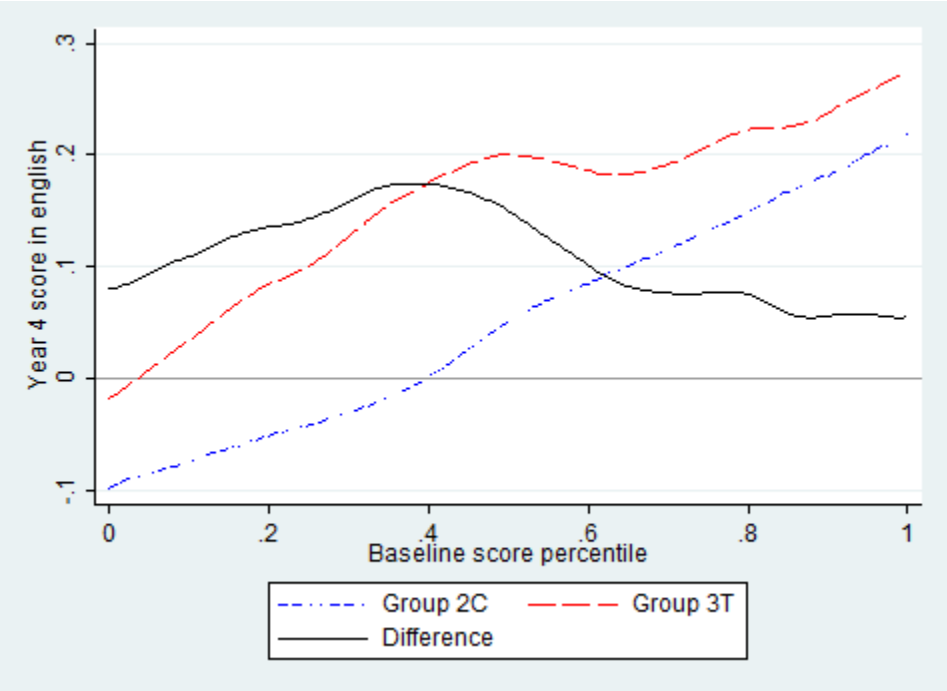


Figure 3 (d): Heterogeneous Y4 Test Score Impacts by Baseline Percentile - Hindi

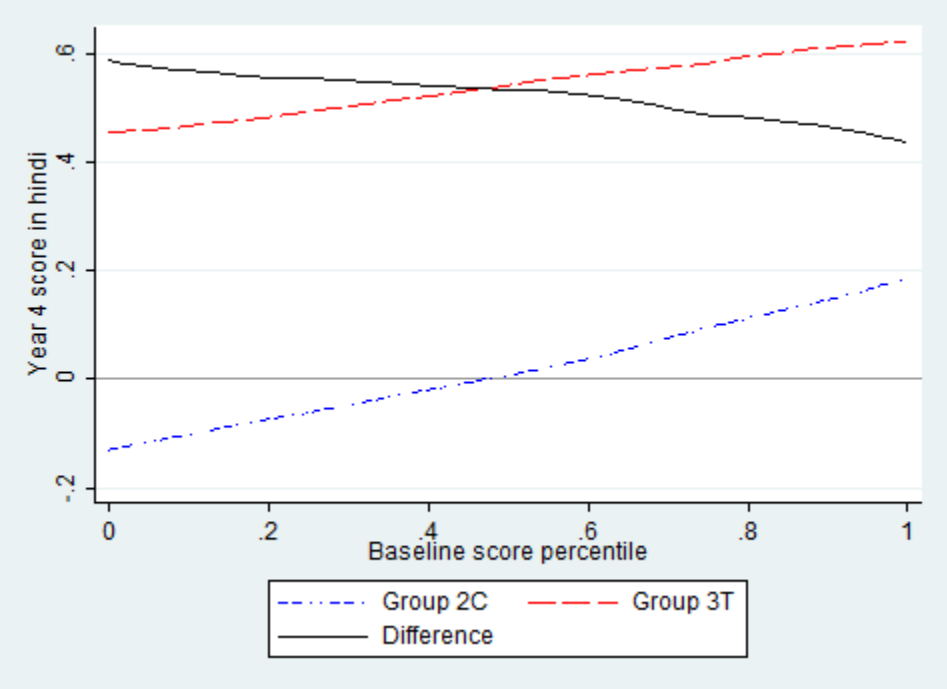


Table 1: Baseline Test Scores and Socio-Economic Characteristics

	Private schools	Government schools	Difference	Difference with village fixed effects
	[1]	[2]	[3]	[4]
Normalized baseline Telugu score	0.639	0.004	0.635***	0.677***
Normalized baseline math score	0.661	0.015	0.646***	0.678***
Both parents have completed at least primary school	0.558	0.285	0.273***	0.308***
At least one parent has completed grade 10	0.547	0.352	0.195***	0.219***
Scheduled caste	0.128	0.329	-0.201***	-0.193***
Household asset index	3.846	3.193	0.653***	0.646***
Annual school fees paid (Rs./month)	1330.37	3.79	1326.57***	1326.92***
Observations	14,541	8,538		

Notes:

* p<0.05; ** p<0.01; *** p<0.001. All standard errors are clustered at the school level. The sample for this table is restricted to students in control villages at the baseline (2008). Telugu and math scores are normalized across treatment and control students with respect to students in control villages by subject and grade. The household asset index reported is a sum of five household indicators, including whether a household owns its own home, has a proper house, has at least one covered room, has working water facilities, and has a toilet available.

Table 2: Validity of Design

Panel A: Treatment and Control Students

	Lottery winners [treatment villages]	Lottery losers [treatment villages]	P-value for difference	Lottery winners [treatment villages]	Lottery losers [control villages]	P-value for difference
	[1]	[2]	[3]	[4]	[5]	[6]
Normalized baseline Telugu score	0.01	0.03	0.44	0.01	-0.05	0.31
Normalized baseline math score	-0.02	-0.03	0.75	-0.02	-0.05	0.74
Both parents have completed at least primary school	0.3	0.28	0.29	0.3	0.28	0.68
At least one parent has completed grade 10	0.34	0.33	0.76	0.34	0.36	0.14
Scheduled caste	0.35	0.33	0.3	0.35	0.32	0.43
Household asset index	3.17	3.14	0.41	3.17	3.19	0.68
Observations	1,980	1,119		1,980	3,334	

Panel B: Students for Spillover Analysis

	Non-applicants in treatment villages	Non-applicants control villages	P-value for difference	Students initially in private schools in treatment villages	Students initially in private schools incontrol villages	P-value for difference
	[1]	[2]	[3]	[4]	[5]	[6]
Normalized baseline Telugu score	-0.03	0.09	0.12	0.57	0.64	0.21
Normalized baseline math score	-0.01	0.09	0.28	0.68	0.66	0.75
Both parents have completed at least primary school	0.28	0.29	0.69	0.51	0.56	0.1
At least one parent has completed grade 10	0.31	0.34	0.19	0.52	0.55	0.11
Scheduled caste	0.34	0.34	0.98	0.12	0.13	0.64
Household asset index	3.16	3.2	0.49	3.85	3.85	0.96
Observations	2,165	2,337		12,720	12,061	

Notes:

All standard errors are clustered at the school level. This table reports responses to baseline household surveys and assessments (2008). Telugu and math scores are normalized across treatment and control students with respect to students in control villages by subject and grade. The household asset index reported is a sum of five household indicators, including whether a household owns its own home, has a proper house, has at least one covered room, has working water facilities, and has a toilet available.

Table 3: Attrition

Panel A: Treatment and Control Students

	Year 4 assessments						Year 2.5 assessments					
	Lottery winners [treatment villages]	Lottery losers [treatment villages]	P-value for difference	Lottery winners [treatment villages]	Lottery losers [control villages]	P-value for difference	Lottery winners [treatment villages]	Lottery losers [treatment villages]	P-value for difference	Lottery winners [treatment villages]	Lottery losers [control villages]	P-value for difference
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Characteristics												
Present during the test	0.85	0.80	0.00	0.85	0.81	0.02	0.90	0.84	0.00	0.90	0.85	0.00
Comparison of attritors												
Normalized baseline telugu score	0.03	0.14	0.21	0.02	0.10	0.27	0.05	0.14	0.37	0.04	0.09	0.70
Normalized baseline math score	0.05	0.01	0.68	0.05	0.10	0.61	0.06	0.03	0.79	0.06	0.10	0.82
Both parents have completed at least primary school	0.28	0.24	0.24	0.28	0.30	0.49	0.29	0.25	0.43	0.29	0.25	0.44
Scheduled caste	0.38	0.41	0.42	0.38	0.34	0.29	0.39	0.34	0.38	0.39	0.32	0.15
Household asset index	3.15	3.10	0.54	3.15	3.21	0.48	3.15	3.01	0.15	3.15	3.09	0.59
Observations	1,980	1,117		1,980	3,338		1,980	1,117		1,980	3,336	

Panel B: Students for Spillover Analysis

	Year 4 assessments						Year 2.5 assessments					
	Non-applicants in treatment villages	Non-applicants control villages	P-value for difference	Students initially in private schools in treatment villages	Students initially in private schools incontrol villages	P-value for difference	Non-applicants in treatment villages	Non-applicants control villages	P-value for difference	Students initially in private schools in treatment villages	Students initially in private schools incontrol villages	P-value for difference
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Characteristics												
Present during the test	0.66	0.67	0.84	0.70	0.65	0.11	0.66	0.67	0.84	0.62	0.60	0.48
Comparison of attritors												
Normalized baseline telugu score	-0.03	0.07	0.41	0.51	0.75	0.00	-0.03	0.07	0.41	0.58	0.72	0.10
Normalized baseline math score	-0.03	0.11	0.36	0.75	0.72	0.81	-0.03	0.11	0.36	0.76	0.69	0.53
Both parents have completed at least primary school	0.26	0.25	0.81	0.57	0.63	0.13	0.26	0.25	0.81	0.57	0.61	0.30
Scheduled caste	0.35	0.29	0.28	0.16	0.15	0.75	0.38	0.28	0.06	0.14	0.13	0.83
Household asset index	3.13	3.21	0.49	3.94	3.94	1.00	3.13	3.29	0.20	3.88	3.90	0.80
Observations	743	811		1,152	1,106		743	811		1,149	1,109	

Notes:

All standard errors are clustered at the school level. Telugu and math scores are normalized across treatment and control students with respect to students in control villages by subject and grade. The household asset index reported is a sum of five household indicators, including whether a household owns its own home, has a proper house, has at least one covered room, has working water facilities, and has a toilet available.

Table 4: School and Teacher Characteristics by School Type

Panel A: School Characteristics

	Private schools	Government schools	Difference
	[1]	[2]	[3]
Total enrollment	301.71	83.31	218.4***
Total working days	229.42	218.40	11.02***
Pupil-teacher ratio	16.86	26.37	-9.514***
Drinking water available	0.99	0.92	0.0730***
Functional toilets	0.89	0.68	0.205***
Separate functional toilets for girls	0.79	0.43	0.364***
Functional electricity	0.90	0.59	0.305***
Functional computers	0.53	0.04	0.484***
Functional library	0.81	0.98	-0.169***
Functional radio	0.14	0.80	-0.660***
Observations	926	1,183	

Panel B: Teacher Characteristics

	[1]	[2]	[3]
Male	0.25	0.44	-0.19***
Age	35.47	47.04	-11.57***
Years of teaching	5.61	14.82	-9.21***
Completed at least college or masters	0.69	0.86	-0.16***
Teacher training completed	0.34	0.98	-0.64***
Come from the same village	0.46	0.14	0.32***
Current gross salary per month (Rs.)	2310.10	13720.90	-11410.85***
Observations	2,868	2,370	

Panel C: School Expenditures

	[1]	[2]	[3]
Annual cost per child (Rs./child)	2334.03	8390.00	-6055.97***
Observations	695	1,052	

Notes:

* p<0.1; ** p<0.05; *** p<0.01. All regressions include district fixed effects. All standard errors are clustered at the school level. The sample for this table is restricted to schools and teachers in control villages across years 2008 through 2012. All expenditures are measured in Rupees per student per year.

Table 5: Measures of Teaching Activity

Panel A: Measures of Classroom Activity

	Private schools	Government schools	Difference
	[1]	[2]	[3]
Class is engaged in active teaching	0.51	0.34	0.17***
A teacher is present in class	0.97	0.92	0.048***
Effective in teaching and maintaining discipline	0.50	0.36	0.14***
Teacher has complete control over class	0.69	0.41	0.28***
Teachers teaching mutliple classes at the same time	0.24	0.79	-0.55***
Observations	2,738	2,784	

Panel B: Measures of Teacher Activity

	[1]	[2]	[3]
Cannot find the teacher (absent) before the class starts	0.09	0.24	-0.15***
Teacher is actively teaching	0.50	0.35	0.15***
Teacher is in school and not teaching	0.01	0.03	-0.02***
Observations	6,577	5,552	

Panel C: Measures of School Hygiene

	[1]	[2]	[3]
Flies heavily present on premises of the school	0.14	0.19	-0.05**
Stagnant water present on premises of the school	0.18	0.28	-0.10***
Garbage dumped on premises of the school	0.33	0.44	-0.11***
Observations	426	614	

Notes:

* p<0.05; ** p<0.01; *** p<0.001. All regressions include district fixed effects. All standard errors are clustered at the school level. The sample for this table is restricted to classrooms, teachers, and schools in control villages. Measures of classroom activity are from school short surveys administered in years 2008 and 2009. Measures of teacher activitiy are from teacher short surveys administered across years 2008 through 2012. Measures of school hygiene are from school short surveys administered across years 2010 through 2012. Actual observations for each regression vary in small amounts within panels based on the dependent variable.

Table 6: Changes in Household Inputs

Panel A: Student Time Diaries (Minutes per Day)

	Private schools	Government schools	Difference	Applicants offered scholarship	Applicants in control villages	Intention to treat estimate	Treatment on the treated estimate
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Activity							
Time spent in school	423.53	380.25	43.28***	409.34	383.38	25.96***	46.93***
Studying and doing homework at home	75.99	52.72	23.27***	59.83	56.86	2.97	5.38
Private Tuition	25.15	16.62	8.53**	21.95	17.43	4.52	8.17
Bathing/Toilet/Getting ready	55.11	61.7	-6.59***	57.82	61.24	-3.42	-6.19
Time traveling to school	23.5	20.92	2.58*	23.51	21.43	2.08	3.75
Working (outside and inside the household)	1.51	11.05	-9.54**	5.46	9.36	-3.90	-7.14
Chores	16.82	31.18	-14.36***	21.62	34.45	-12.83**	-23.51**
Watching TV	75.88	83.38	-7.50**	80.57	84.04	-3.47	-6.28
Playing with friends	82.34	101.99	-19.65***	100.88	99.73	1.15	2.08
Eating	43.57	44.69	-1.12	43.78	44.12	-0.34	-0.61
Free time	53.38	64.38	-11.00**	56.69	62.13	-5.44	-9.96
Observations	652	1839		885	1212		2097

Panel B: Household Student Expenditure (Rupees per year)

	Private schools	Government schools	Difference	Applicants offered scholarship	Applicants in control villages	Intention to treat estimate	Treatment on the treated estimate
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Household student expenditure	2910.36	566.73	2343.64***	774.94	892.69	-117.75***	-215.95***
Observations							

Notes:

* p<0.05; ** p<0.01; *** p<0.001. All regressions include district fixed effects. All standard errors are clustered at the village level. In all columns, the sample is restricted to those students who reported activities and processes from a normal, non-sick school day. The sample for columns [1] through [3] is restricted to students and schools in control villages. The samples for columns [4] through [7] is applicants offered scholarships in treatment villages and applicants not offered scholarships in control villages. Data for both panels come from the parent child surveys administered between 2008 and 2012. The chores activity consists of preparing meals, caring for other children, and caring for the elderly. Household student expenditures includes expenditures on admissions exams, uniforms, notebooks, textbooks, special events, transportation, and private tuition. Actual observations for each regression vary in small amounts within panels based on the dependent variable.

Table 7: Test Score Impacts

Panel A: Intention to Treat Effects

	Year 2 assessments			Year 4 assessments				
	Telugu score	Math score	English score	Telugu score	Math score	English score	Science and social studies score	Hindi score
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Offered scholarship	-0.079	-0.053	0.179**	-0.017	-0.031	0.114	0.084	0.526***
	(0.055)	(0.065)	(0.079)	(0.052)	(0.053)	(0.072)	(0.061)	(0.068)
Total observations	4,620	4,620	4,525	4,385	4,385	4,217	4,243	1,691
Treatment observations	1,778	1,778	1,738	1,674	1,675	1,607	1,628	867
Control observations	2,842	2,842	2,787	2,711	2,710	2,610	2,615	824

Panel B: Treatment on the Treated Effects

	Year 2 assessments			Year 4 assessments				
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Scholarship recipient in private school	-0.140	-0.094	0.317**	-0.030	-0.055	0.201	0.149	0.891***
	(0.098)	(0.115)	(0.139)	(0.092)	(0.093)	(0.127)	(0.108)	(0.103)
Total observations	4,620	4,620	4,525	4,385	4,385	4,217	4,243	1,691
Scholarship recipients	997	997	982	945	946	911	920	510
Non-recipients	3,623	3,623	3,543	3,440	3,439	3,306	3,323	1,181

Notes:

* p<0.1; ** p<0.05; *** p<0.01. All regressions control for baseline normalized test scores and include a constant and district fixed effects. All standard errors are clustered at the village level. All test scores are normalized across treatment and control students with respect to students in control villages by subject and grade. Telugu, math, English and social studies test scores are from endline exams, Hindi test scores are from a special assessment. Combined scores are obtained by running a pooled regression across all test scores in each year, with each observation weighted by the inverse of its probability of being observed in the sample.

Table 8: School Time Tables

	Instructional Time by Subject (Minutes per week)		
	Private schools	Government schools	Difference
	[1]	[2]	[3]
Telugu	307.46	511.49	-204.03***
	(6.36)	(3.63)	(6.98)
Math	339.59	500.62	-161.02***
	(7.53)	(3.38)	(8.58)
English	322.60	235.41	87.19***
	(8.10)	(5.37)	(9.82)
Social studies	239.18	173.57	65.61***
	(6.30)	(6.88)	(9.92)
General science	205.53	104.39	101.14***
	(9.28)	(5.83)	(9.55)
Hindi	215.97	0.02	215.96***
	(6.13)	(0.90)	(6.46)
Moral science	16.75	20.30	-3.55
	(4.86)	(3.19)	(5.60)
Computer use	46.57	0.38	46.19***
	(6.53)	(1.02)	(6.85)
Other	311.95***	250.11***	61.84***
	(14.56)	(6.75)	(16.24)
Break	461.51	473.10	-11.60
	(9.12)	(3.07)	(10.59)
Total	2467.10***	2269.38***	197.72***
	(17.43)	(8.33)	(19.79)
Observations	323	200	

Notes:

* p<0.1; ** p<0.05; *** p<0.01. All regressions include a constant and district fixed effects. All standard errors are clustered at the school level. The sample for this table is restricted to schools in control villages. All numbers in minutes per week. Other includes sports, arts and crafts, and other subjects.

Table 9: Robustness to Attrition

Panel A: Inverse Probability Weighting

	Year 2 assessments			Year 4 assessments				
	Telugu score	Math score	English score	Telugu score	Math score	English score	Science and social studies score	Hindi score
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Offered scholarship	-0.081	-0.055	0.177**	-0.017	-0.031	0.113	0.083	0.522***
	(0.055)	(0.065)	(0.079)	(0.052)	(0.053)	(0.072)	(0.061)	(0.068)
Total observations	4,620	4,620	4,525	4,385	4,385	4,217	4,243	1,696
Treatment observations	1,778	1,778	1,738	1,674	1,675	1,607	1,628	867
Control observations	2,842	2,842	2,787	2,711	2,710	2,610	2,615	829

Notes:

Note that the same model using the set of covariates from table 9 predicts attrition for both students offered the scholarship and those who applied in control villages once scholarship offer is controlled for.

Panel B: Lee Bounds

	Year 2 assessments			Year 4 assessments				
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Lower bound estimate	-0.148	-0.142	0.076	-0.076	-0.112	0.001	-0.010	0.479
	(0.031)	(0.031)	(0.039)	(0.036)	(0.035)	(0.046)	(0.040)	(0.062)
Upper bound estimate	0.028	0.041	0.262	0.059	0.045	0.199	0.200	0.575
	(0.030)	(0.030)	(0.039)	(0.034)	(0.033)	(0.042)	(0.041)	(0.060)
Confidence interval low	-0.208	-0.202	-0.001	-0.145	-0.180	-0.090	-0.089	0.358
Confidence interval high	0.087	0.099	0.337	0.125	0.109	0.281	0.280	0.694

Notes:

* p<0.1; ** p<0.05; *** p<0.01. All regressions in Panel A control for baseline normalized test scores and include a constant and district fixed effects. All standard errors are clustered at the village level. All test scores are normalized across treatment and control students with respect to students in control villages by subject and grade. Telugu, math, English and social studies test scores are from endline exams, Hindi test scores are from a special assessment. Inverse probabilities calculated using a probit regression of whether a student attrited on whether they were offered a scholarship and their baseline test scores, clustered at the school level. Combined scores are obtained by running a pooled regression across all test scores in each year, with each observation weighted by the inverse of its probability of being observed in the sample.

Table 10: Heterogeneous Test Score Impacts

	Year 2 assessments			Year 4 assessments				
	Telugu score	Math score	English score	Telugu score	Math score	English score	Science and social studies score	Hindi score
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Offered scholarship * covariate								
Baseline test score	-0.056	-0.054	-0.074	0.004	0.000	-0.022	-0.022	-0.105*
	(0.042)	(0.043)	(0.052)	(0.037)	(0.037)	(0.047)	(0.043)	(0.060)
Female indicator	0.013	0.069	0.117	0.010	-0.037	0.013	0.017	0.169*
	(0.055)	(0.054)	(0.076)	(0.064)	(0.064)	(0.071)	(0.071)	(0.098)
Scheduled caste indicator	-0.012	0.042	-0.025	0.029	0.099	0.006	0.056	-0.108
	(0.069)	(0.066)	(0.082)	(0.070)	(0.070)	(0.090)	(0.083)	(0.121)
Parents literate indicator	0.043	-0.022	-0.011	-0.031	-0.006	0.132	-0.138*	-0.234**
	(0.065)	(0.060)	(0.089)	(0.068)	(0.070)	(0.120)	(0.077)	(0.117)
Parents laborers indicator	-0.006	0.042	0.018	0.050	0.144**	0.148	0.019	-0.123
	(0.069)	(0.069)	(0.083)	(0.069)	(0.071)	(0.102)	(0.081)	(0.115)
Household asset index	0.014	0.031	-0.018	-0.028	-0.001	0.009	-0.019	0.017
	(0.031)	(0.033)	(0.045)	(0.033)	(0.031)	(0.038)	(0.035)	(0.062)
Muslim indicator	0.232**	0.258**	0.404**	0.364***	0.290**	0.151	0.288**	0.111
	(0.116)	(0.137)	(0.177)	(0.112)	(0.128)	(0.147)	(0.140)	(0.168)
Christian indicator	-0.099	-0.085	-0.208	-0.154	-0.232**	-0.113	-0.109	-0.193
	(0.122)	(0.126)	(0.157)	(0.130)	(0.111)	(0.121)	(0.159)	(0.248)
Indicator for older cohort at baseline	0.020	-0.062	0.107	-0.045	-0.055	0.101	-0.051	0.116
	(0.072)	(0.083)	(0.089)	(0.082)	(0.087)	(0.116)	(0.097)	(0.104)
Observations	4,620	4,620	4,525	4,385	4,385	4,217	4,243	1,691

Notes:

* p<0.1; ** p<0.05; *** p<0.01. Each cell reports the coefficient from a separate regression that also includes controls for whether a student was offered a scholarship and the student's baseline test scores. All regressions include district fixed effects and standard errors are clustered at the village level. All test scores are normalized across treatment and control students with respect to students in control villages by subject and grade. Telugu, math, English and social studies test scores are from endline exams, Hindi test scores are from a special assessment. The household asset index reported is a sum of five household indicators, including whether a household owns its own home, has a proper house, has at least one covered room, has working water facilities, and has a toilet available. Actual observations for each regression vary in small amounts within columns based on the dependent variable. Combined scores are obtained by running a pooled regression across all test scores in each year, with each observation weighted by the inverse of its probability of being observed in the sample.

Table 11: Test Score Impacts by Medium of Instruction

Panel A: Variance decomposition by medium of instruction

	Year 2 assessments			Year 4 assessments				
	Telugu score	Math score	English score	Telugu score	Math score	English score	Science and social studies score	Hindi score
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Scholarship recipients attending private English medium schools	-0.168**	-0.135*	0.495***	-0.065	-0.050	0.424***	0.107	1.150***
	(0.065)	(0.076)	(0.088)	(0.076)	(0.076)	(0.135)	(0.084)	(0.085)
Scholarship recipients attending private Telugu medium schools	-0.051	-0.053	0.105	0.067	0.005	0.126	0.145*	0.954***
	(0.084)	(0.085)	(0.112)	(0.071)	(0.072)	(0.084)	(0.076)	(0.084)
Students who rejected the scholarship	-0.048	-0.003	0.006	-0.020	-0.031	-0.068	0.047	-0.078
	(0.062)	(0.069)	(0.086)	(0.059)	(0.060)	(0.066)	(0.074)	(0.075)
Total observations	4,620	4,620	4,525	4,385	4,385	4,217	4,243	1,696
Treatment observations	1,778	1,778	1,738	1,674	1,675	1,607	1,628	867
Control observations	2,842	2,842	2,787	2,711	2,710	2,610	2,615	829

Panel B: Treatment on the Treated Effects by Medium

	Year 2 assessments			Year 4 assessments				
	Telugu score	Math score	English score	Telugu score	Math score	English score	Social studies score	Hindi score
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Scholarship recipients attending private English medium schools	-0.143	-0.130	0.637***	-0.290	-0.434*	0.739**	-0.187	1.439***
	(0.137)	(0.163)	(0.194)	(0.220)	(0.260)	(0.334)	(0.273)	(0.280)
Scholarship recipients attending private Telugu medium schools	-0.088	0.021	0.386*	0.189	0.291	0.104	0.630**	1.148***
	(0.170)	(0.190)	(0.217)	(0.198)	(0.259)	(0.221)	(0.279)	(0.280)
Total observations	3,520	3,520	3,474	3,536	3,536	3,423	3,419	1,359
Treatment observations	1,407	1,407	1,383	1,396	1,397	1,349	1,358	704
Control observations	2,113	2,113	2,091	2,140	2,139	2,074	2,061	655
First-stage F-stat on first regressor	24.0	24.0	23.9	10.9	11.0	10.9	10.9	8.8
First-stage F-stat on second regressc	20.7	20.7	21.1	15.3	15.3	14.9	15.5	15.4
P-value of equality by medium	0.78	0.50	0.32	0.16	0.11	0.16	0.08	0.54

Panel C: Overall Treatment on the Treated Effects in Distance Sample

	Year 2 assessments			Year 4 assessments				
	Telugu score	Math score	English score	Telugu score	Math score	English score	Social studies score	Hindi score
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Scholarship recipient in private schoo	-0.115	-0.075	0.386***	-0.032	-0.050	0.267**	0.168	0.864***
	(0.098)	(0.115)	(0.138)	(0.086)	(0.091)	(0.124)	(0.105)	(0.109)
Total observations	3,520	3,520	3,474	3,536	3,536	3,423	3,419	1,359
Scholarship recipients	842	842	831	828	829	800	805	436
Non-recipients	2,678	2,678	2,643	2,708	2,707	2,623	2,614	923

Notes:
* p<0.1; ** p<0.05; *** p<0.01. All regressions control for baseline normalized test scores and include a constant and district fixed effects. All standard errors are clustered at the village level. The sample in all panels restricted to those students for which location data is available and students with at least one private school in their village. All test scores are normalized across treatment and control students with respect to students in control villages by subject and grade. Telugu, math, English and social studies test scores are from endline exams, Hindi test scores are from a special assessment. Panel D instruments for the medium of of instruction using scholarship offer and the interaction between a scholarship offer and the medium of the nearest private school.

Table 12: Heterogeneous Impacts by Market Competition

	Year 2 assessments			Year 4 assessments				
	Telugu score	Math score	English score	Telugu score	Math score	English score	Science and social studies score	Hindi score
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Offered scholarship * covariate								
Panel A: Heteroegeneous Effects as a Function of Number of Schools within 0.5 km								
Number of Private Schools (Linear)	0.057	-0.009	-0.021	0.095	0.006	-0.054	-0.015	-0.084
	(0.077)	(0.084)	(0.103)	(0.076)	(0.078)	(0.113)	(0.084)	(0.106)
3 or more Private Schools	0.160	0.025	-0.031	0.189*	0.100	-0.063	0.058	-0.011
	(0.111)	(0.125)	(0.147)	(0.103)	(0.103)	(0.161)	(0.113)	(0.157)
4 or more Private Schools	0.203	0.159	0.222	0.264**	0.198	0.214	0.232*	0.081
	(0.123)	(0.123)	(0.166)	(0.115)	(0.129)	(0.245)	(0.121)	(0.167)
5 or more Private Schools	0.206	0.182	0.306	0.284*	0.198	-0.006	0.272*	-0.243
	(0.186)	(0.168)	(0.192)	(0.162)	(0.153)	(0.183)	(0.163)	(0.271)
6 or more Private Schools	0.247	0.142	0.319	0.388***	0.329**	0.267**	0.386**	0.076
	(0.193)	(0.220)	(0.228)	(0.143)	(0.152)	(0.117)	(0.183)	(0.334)
Panel B: Heteroegeneous Effects as a Function of Number of Schools within 1 km								
3 or more Private Schools	0.038	-0.066	-0.128	0.046	-0.048	-0.063	-0.082	-0.127
	(0.111)	(0.126)	(0.152)	(0.104)	(0.106)	(0.167)	(0.115)	(0.136)
4 or more Private Schools	0.161	0.142	0.167	0.168	0.060	0.040	0.029	0.027
	(0.116)	(0.131)	(0.169)	(0.119)	(0.126)	(0.243)	(0.140)	(0.150)
5 or more Private Schools	0.188	0.250*	0.278	0.228*	0.127	-0.142	0.175	-0.030
	(0.118)	(0.135)	(0.197)	(0.126)	(0.146)	(0.174)	(0.157)	(0.163)
6 or more Private Schools	0.170	0.283	0.427	0.457***	0.410***	0.170	0.449***	-0.156
	(0.143)	(0.202)	(0.323)	(0.116)	(0.130)	(0.113)	(0.136)	(0.228)
Observations	4612	4612	4518	4378	4378	4215	4237	1689

Notes:

* p<0.1; ** p<0.05; *** p<0.01. Each cell reports the coefficient from a separate regression that also includes controls for whether a student was offered a scholarship and the student's baseline test scores. All regressions include district fixed effects and standard errors are clustered at the village level. All test scores are normalized across treatment and control students with respect to students in control villages by subject and grade. Telugu, math, English and social studies test scores are from endline exams, Hindi test scores are from a special assessment. The household asset index reported is a sum of five household indicators, including whether a household owns its own home, has a proper house, has at least one covered room, has working water facilities, and has a toilet available. Actual observations for each regression vary in small amounts within columns based on the dependent variable. Combined scores are obtained by running a pooled regression across all test scores in each year, with each observation weighted by the inverse of its probability of being observed in the sample.

Table 13: Estimating Spillover Effects

	Year 2 assessments			Year 4 assessments			
	Telugu score	Math score	English score	Telugu score	Math score	English score	Science and social studies score
Panel A: Comparing the Within-Village to Across-Village Controls							
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Lottery Loser in Treatment Village	0.009	0.010	0.033	0.014	0.001	-0.049	0.095*
	(0.041)	(0.044)	(0.056)	(0.044)	(0.045)	(0.059)	(0.051)
Panel B: Impact on Non-applicants from Government Schools							
Treatment village	-0.025	0.046	0.119	0.049	-0.002	0.024	-0.023
	(0.072)	(0.068)	(0.089)	(0.063)	(0.069)	(0.071)	(0.073)
Panel C: Impact on Non-scholarship Students from Private Schools							
Treatment village	0.065	0.025	-0.114	0.040	0.037	-0.026	0.029
	(0.062)	(0.074)	(0.076)	(0.062)	(0.059)	(0.104)	(0.073)

Notes:

* p<0.1; ** p<0.05; *** p<0.01. All regressions control for baseline normalized test scores and include a constant and district fixed effects. All standard errors are clustered at the village level. All test scores are normalized across treatment and control students with respect to students in control villages by subject and grade. Telugu, math, English and science and social studies test scores are from endline exams, Hindi test scores are from a special assessment.

Table 14: Aggregate effects

	Year 2 assessments			Year 4 assessments			
	Telugu score	Math score	English score	Telugu score	Math score	English score	Science and social studies score
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Treatment effect on:							
Non-applicants from government schools	-0.025	0.046	0.119	0.049	-0.002	0.024	-0.023
	(0.069)	(0.063)	(0.076)	(0.063)	(0.063)	(0.068)	(0.069)
Applicants that were not offered scholarship	0.009	0.010	0.033	0.014	0.001	-0.049	0.095*
	(0.041)	(0.044)	(0.056)	(0.044)	(0.045)	(0.059)	(0.051)
Scholarship students	-0.079**	-0.053	0.179***	-0.017	-0.031	0.114*	0.084*
	(0.038)	(0.040)	(0.051)	(0.040)	(0.041)	(0.064)	(0.044)
Non-scholarship students from private schools	0.065	0.025	-0.114*	0.040	0.037	-0.026	0.029
	(0.049)	(0.054)	(0.062)	(0.053)	(0.052)	(0.091)	(0.059)
All students in treatment villages (sample weighted)	**0.053**	**0.050**	**0.036**	**0.063**	**0.050**	**0.057**	**0.057**
	(0.040)	**(0.043)**	**(0.057)**	**(0.040)**	**(0.040)**	**(0.072)**	**(0.045)**

Notes:

* p<0.1; ** p<0.05; *** p<0.01. All regressions control for baseline normalized test scores and include a constant and district fixed effects. All standard errors are clustered at the village level. All test scores are normalized across treatment and control students with respect to students in control villages by subject and grade. Telugu, math, English and social studies test scores are from endline exams, Hindi test scores are from a special assessment. Scholarship students are compared to applicants not offered scholarships in control villages. Sample weights for village-wide treatment effect are the the inverse of the ratio of sampled students to the entire population: Non-applicants---6.61, rejected applicants---1.00, scholarship students---1.00, and non-scholarship students---13.31. Actual observations for each regression vary in small amounts within rows based on the dependent variable and year of assessment.