

NBER WORKING PAPER SERIES

MONITORING WORKS:
GETTING TEACHERS TO COME TO SCHOOL

Esther Duflo
Rema Hanna

Working Paper 11880
<http://www.nber.org/papers/w11880>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2005

This project is a collaborative exercise involving many people. Foremost, we are deeply indebted to Seva Mandir, and especially to Neelima Khetan and Priyanka Singh, who made this evaluation possible. We thank Ritwik Sakar and Ashwin Vasani for their excellent work coordinating the fieldwork. Greg Fischer, Shehla Imran, Callie Scott and Kudzai Takavarasha provided superb research assistance. For their helpful comments, we thank Abhijit Banerjee, Rachel Glennerster, Michael Kremer and Sendhil Mullainathan. For financial support, we thank the John D. and Catherine T. MacArthur Foundation. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

©2005 by Esther Duflo and Rema Hanna. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Monitoring Works: Getting Teachers to Come to School
Esther Duflo and Rema Hanna
NBER Working Paper No. 11880
December 2005
JEL No. I20, I21, J13, J30, O10

ABSTRACT

In the rural areas of developing countries, teacher absence is a widespread problem. This paper tests whether a simple incentive program based on teacher presence can reduce teacher absence, and whether it has the potential to lead to more teaching activities and better learning. In 60 informal one-teacher schools in rural India, randomly chosen out of 120 (the treatment schools), a financial incentive program was initiated to reduce absenteeism. Teachers were given a camera with a tamper-proof date and time function, along with instructions to have one of the children photograph the teacher and other students at the beginning and end of the school day. The time and date stamps on the photographs were used to track teacher attendance. A teacher's salary was a direct function of his attendance. The remaining 60 schools served as comparison schools. The introduction of the program resulted in an immediate decline in teacher absence. The absence rate (measured using unannounced visits both in treatment and comparison schools) changed from an average of 42 percent in the comparison schools to 22 percent in the treatment schools. When the schools were open, teachers were as likely to be teaching in both types of schools, and the number of students present was roughly the same. The program positively affected child achievement levels: a year after the start of the program, test scores in program schools were 0.17 standard deviations higher than in the comparison schools and children were 40 percent more likely to be admitted into regular schools.

Esther Duflo
Department of Economics
MIT, E52-252G
50 Memorial Drive
Cambridge, MA 02142
and NBER
eduflo@mit.edu

Rema Hanna
Department of Economics
Wagner Graduate School of Public Service
New York University
295 Lafayette Street
New York, NY 10012
rnh4@nyu.edu

I. INTRODUCTION

The United Nations Millennium Development Goals call for achieving universal primary education by 2015. In response, many developing countries, including India, are rapidly improving access to primary schooling. However, improved access often is not matched by improvements in school quality. As a result, while more children start primary school, many leave after just a few years, after having learned very little in the process. For example, in Uttar Pradesh, India, half of the students enrolled in primary school can not even read a simple sentence (Banerjee et al., 2005). Such poor learning outcomes may be an artifact of high absence among teachers. Using unannounced visits to measure teacher attendance, a nationally representative survey found that 24 percent of teachers in India were absent from the classroom during normal school hours (Chaudhury, et al., 2005a, b).¹ The situation in India is particularly bleak. In terms of absence rates, India ranked seventh among the eight countries for which comparable data was collected. Getting teachers to attend school may help India achieve the improvements in school quality needed to make “universal primary education” a meaningful term.

It has been argued that teachers fail to attend school because neither their principal nor the beneficiary has the capacity to both effectively monitor and penalize absence. The principals, usually governments (but also NGOs), have the power to penalize absences, but, being far removed, may not be able to effectively monitor attendance. As such, they often lack the information needed to enforce attendance rules. The community can effectively monitor attendance, but it often lacks the power to penalize absence. One solution—championed by many, including the 2004 World Development Report—is to expand community control by improving community-based monitoring; strengthening the flow of information between the community and the principal; involving the community in decisions to hire, fire, and pay teachers; or transferring wholesome control of teachers to the community.

However, evidence from a variety of contexts suggests that community control interventions have not been particularly effective at reducing absence (Banerjee and Duflo, 2005). Banerjee, Deaton, and

¹ Although teachers do have some official non-teaching duties, this absence rate is much too high to be fully explained by this particular story.

Duflo (2004) found that community-based monitoring, even when robustly structured, did not reduce absenteeism among service providers at government health facilities in rural India. Information sharing and auxiliary rewards for teachers fare no better. Kremer and Vermeersch (2005) found no effect for a program in rural Kenya that empowered school committees to monitor teachers, share performance information with officers in the Ministry of Education, and to give prizes to the best-performing teachers. Ensuring that teachers have closer ties to the community, which may entail high community pressure, had no effect on absence either. Chaudhury et al. (2005b) found that locally hired teachers, teachers in schools with a Parents-Teachers' Association, teachers with longer local tenure, and contract teachers and teachers at non-formal schools run by NGOs (who, in addition, faced a greater risk of dismissal) all had absence rates significantly higher than those of government school teachers.² Finally, Olken (2004) found that increasing community participation in meetings where public officials accounted for expenditure of public funds did not reduce corruption in local development projects in Indonesia.

There is limited evidence that external control, coupled with a clear and credible threat of punishment, may be more effective at inducing “good” behavior. Contrary to his findings on community participation, in the same study, Olken (2004) found that the threat of a top-down audit resulted in a significant decline in corruption. Chaudhury et al. (2005b) report that teachers at schools that were inspected more often tended to have lower absentee rates.

In this paper, we formally test whether direct monitoring, coupled with high-powered incentives, results in higher quality schooling. In particular, we study a scheme aimed at reducing truancy among teachers at NGO-run non-formal education centers (NFEs) in rural India. Seva Mandir, the NGO running the NFEs, used cameras with tamper-proof date and time functions to monitor daily teacher absence. Then, they provided teachers with financial incentives that were based on the detailed attendance data recorded by the cameras. We take advantage of Seva Mandir's program to answer three main questions: If teachers are given high-powered incentives to attend school based on external monitoring, will they

² It may, however, be said that the difficult conditions (remote areas, part-time teachers and students, etc.) under which these schools operate may counterbalance the effects of community pressure.

attend school more? If teachers attend school more, will they teach more? Finally, if teacher absenteeism is reduced, will children learn more as a result?

While there are many good reasons to believe that high-powered incentives based on presence may reduce absenteeism, the incentives may fail if teachers face constraints that do not allow them to take advantage of the incentive scheme. For example, some argue that teachers do not come to school because they must participate in meetings, training sessions, election or census duty, etc. Others add that teachers live so far from school and are subject to so many pressures that attending school regularly is not possible, especially in informal schools.

Even if incentives increase teacher attendance, it is unclear whether or not they actually increase child learning levels. Teachers may be subject to *multitasking*, where the agent concentrates on the easiest way to increase the rewarded measure with little or no gains in the measure the principal ultimately wants to improve (Holstrom and Milgrom, 1991). Under this particular type of incentive scheme, teachers may focus on being present (or even on being present just during the few minutes when they are to be monitored), but reduce their efforts in other dimensions. The fear of multitasking is not completely unfounded; there is evidence that other incentive programs have been prone to multitasking. Glewwe, Ilias and Kremer (2003) estimated the effect of a teacher incentive program based on child test scores in Kenya. They found that the program did increase test scores in the short run, but that the gains in learning were only temporary and were not accompanied by increases in teacher attendance or effort. Teachers, they concluded, may have just gamed the system by teaching to the test.³ Studies in the United States provide further evidence of similar gaming behavior among educators facing high-powered incentives, including altering what was served at lunch on the day of the test (Figlio and Winicki, 2002), manipulating who took the test (Figlio and Getzler, 2002), and outright cheating (Jacob and Levitt, 2003).

³ Lavy (2004) provides a more optimistic assessment of a teacher incentive program in Israel.

Several other theories also suggest that providing financial incentives to attend may cause teachers to teach less even as they attend school more.⁴ First, such schemes may demoralize teachers, resulting in less effort. In laboratory experiments, Fehr and Schmidt (2004) found that individuals under high-powered incentive systems may lose their motivation and, thus, work less than under a flat wage regime. Second, financial incentives may harm a teacher's intrinsic motivation, that is, the sense of duty or enjoyment of the job that motivates them to come to work (Kreps, 1997). This threat is particularly real for teachers, who as a group may have strong intrinsic motivation because of the value they place on interacting with children and in seeing the children succeed. (Despite the high absence rate, this can be said of teachers in developing countries. After all, despite very difficult circumstances and the lack of any sanctions if they do not come, most of them do come to school and do teach on most days.) If, given incentives based on presence, teachers may come to believe that just attending class is enough and that their behavior in the classroom is not important. Finally, some teachers, who previously believed that they were required to work every day in the month, might decide to stop working once they have reached their target income for the month (Fehr and Gotte, 2002).

On the other hand, incentives can improve child learning levels if the main cost of working for a teacher is the opportunity cost of going to school, rather than carrying out other income generating activities (such as tending the field or working in another job for a day). Once a teacher has come to school, the marginal cost of actually teaching may be quite low. Under these circumstances, an incentive system that directly rewards presence would have the best chance of increasing child learning. Thus, whether or not an incentive program based on absence can improve teaching effectiveness and learning outcomes is ultimately an empirical question.

In this study, we examine the impact of Seva Mandir's teacher incentive program on teacher presence, teaching activities, and child learning. Seva Mandir runs single-teacher non-formal education centers (NFEs). NFEs are informal schools, run either by local governments (Panchayat) or by NGOs,

⁴ Chaudhuury (2005a), for example, found that only 45 percent of teachers in India were actually teaching at the time of the unannounced visits.

which cater to children for whom the government school is too far or has too rigid a schedule. NFEs have played a crucial role in increasing access to education in India, with over 21 million Indian children, mostly poor and rural, enrolled in NFEs. Seva Mandir's NFEs serve the populations in the tribal villages of Udaipur, Rajasthan. Udaipur is a sparsely populated, hilly region where villages are often remote and hard to reach, making regular monitoring of the NFEs difficult. As a result, absence rates among teachers are high, despite the threat of dismissal for repeated absence. Banerjee et al. (2004) found an absence rate of 40 percent in 1995. In this study, at the baseline taken in 2003, we find an absence rate of 44 percent.

Faced with such high absenteeism, Seva Mandir implemented an innovative incentive program in September, 2003. In 60 randomly selected program schools, Seva Mandir gave teachers a camera, along with instructions to have one of the students take a picture of the teacher and the other students at the start and close of each school day. The cameras had tamper-proof date and time functions, allowing for the collection of accurate information on teacher attendance that was then used to calculate teachers' salaries (see Figure 1 for a sample picture, with the date in the bottom right corner). Each teacher was paid according to the number of "valid" school days for which they were actually present. A "valid" day was defined as one for which the opening and closing photos were separated by at least five hours and both photos showed a minimum number of children. In the 60 comparison schools, teachers were paid a fixed rate for the month, and were told (as usual) that they could be dismissed for repeated, unexcused absences.

The introduction of the program resulted in an immediate and long lasting improvement in teacher attendance rates in treatment schools (as measured through one unannounced visit per month in both treatment and comparison schools). Over the 18 months of the program, teachers at program schools had an absence rate of 22 percent, roughly half of the 44 percent baseline and the 42 percent at comparison schools. Some 36 percent of program teachers had better than 90 percent presence compared to only 1 percent of comparison teachers. Extreme delinquency, over 50 percent absence, was eradicated in program schools. That absence rates stayed low after the end of the (proper) evaluation phase implies

that the effect was not due to a Hawthorne Effect—namely, teachers did not change their behavior simply for the experiment.

We see no evidence that effort declined in other dimensions. When school was open, teachers were as likely to be teaching in treatment as in comparison schools, confirming our intuition that the marginal costs of teaching are low once the teacher is present. However, because they had better attendance records than their comparison school counterparts, teachers at treatment schools taught for the equivalent of 54 more child days (or a third more) per month. Student attendance was the same in both groups, but more teaching meant more learning for children in treatment schools. A year after the start of the program, their test scores were 0.17 standard deviations higher than those of children in comparison schools. The program impact and cost compares favorably with other successful education programs in developing countries.

The findings clearly demonstrate the link between simple, straightforward, well-enforced incentives and teacher presence, as well as the link between teacher presence and student achievement. In theory, this type of incentive scheme is already in place. Teachers are paid to come to work every day, and most school systems, both private and public, have provisions to penalize unexplained absences. In developing countries, however, teachers are typically not punished (much less dismissed) for poor attendance.⁵ Seva Mandir's program did not require elaborate reengineering of school institutions; it instead provided the means to enforce existing rules and strengthen existing incentives. The implication for policy is that one way to rapidly achieve the much-needed improvements in school quality may be to find ways of enforcing the existing rules.⁶ In addition, the findings suggest that external monitoring coupled with simple, direct incentives may also be used to reduce absence among providers of other services essential to development, such as health, in rural areas where the need is the greatest and absence is most prevalent.

⁵ Chaudhury et al. (2005a) report that, the 25 percent absence rate notwithstanding, only one principal in their sample of 3,000 government schools reported a case in which a teacher was fired for repeated absence.

⁶ Of course, it is not necessarily easy to enforce such rules: teachers are an important political constituency and may be able to successfully oppose even a drive to enforce existing rules.

The remainder of the paper is organized as follows. Section II provides a detailed description of the incentive program and the evaluation techniques. The results are presented in Section III. Section IV concludes.

II. THE PROGRAM AND THE EVALUATION

1. Non-formal Education Centers

Non-formal education centers (NFEs) are an integral component of India's education system. Since the National Policy on Education of 1986, they have played an increasingly important role in India's drive towards universal primary education. The NFEs serve two main purposes. First, since they are easier to establish and cheaper to run, they have been the primary instrument for rapidly expanding access to schooling to children in poor, remote rural areas where there are no government schools or where schools are far away. The government of Madhya Pradesh, for example, mandated that NFEs be established for all communities where there were no schools within a kilometer. Second, the NFEs have been used to ease children, who may otherwise not attend school, to join a government school at the age-appropriate grade level. In particular, since NFEs are subject to fewer regulations than government schools, they can tailor their hours and curricula to meet the diverse needs of the children. As of 1997, 21 million children were enrolled in NFEs across India (Education for All Forum, 2000).

Children of all ages may attend, though most were between 7-10 years in our sample. Nearly all the children are illiterate when they first join the centers. In the setting of our project, the NFEs are open six hours a day and have 20 students, all taught in one classroom by one teacher, who is recruited from the local community and has, on average, completed up to a 10th grade education. Instruction focuses on teaching children basic Hindi and math skills.

2. The Program

Seva Mandir administers about 150 non-formal primary education centers (NFEs) in the tribal villages of Udaipur, Rajasthan. Udaipur is a sparsely populated, arid and hilly region, where villages are remote and access is difficult. As a result, it is often difficult for Seva Mandir to regularly monitor the NFEs. Absenteeism is high, despite the organization's policy calling for dismissal of absent teachers. A 1995 study (Banerjee et al., 2005) found that the absence rate was 40 percent, while the baseline of this study (in August 2003) found that the rate was 44 percent.

Seva Mandir was, therefore, motivated to identify ways to reduce absenteeism among its teachers. To this end, they implemented an innovative external monitoring program in September 2003. They chose 120 schools to participate in the study, with 60 randomly selected schools for the program serving as the treatment group and the remaining 60 as the comparison group. In the 60 treatment schools, Seva Mandir gave each teacher a camera, along with instructions for one of the students to take a photograph of the teacher and the other students at the start and end of each school day. The cameras had a tamper-proof date and time function, which made it possible to precisely track each school's openings and closings.⁷ As Figure 1 demonstrates, the day of the month and the time of day appear in the right corner (the month does not appear, but there is no ambiguity about that since the rolls were changed every month). Camera upkeep (replacing batteries, changing and collecting film) was done monthly at regularly scheduled teacher meetings. If a camera malfunctioned, teachers were instructed to call the program hotline within 48 hours. Someone was then dispatched to replace the camera, and teachers were credited for the day in which the camera was broken.⁸

The monthly base salary for teachers was set at Rs1000 for 21 days of work in a month. In the treatment schools, teachers received a Rs50 bonus for each additional day they attended in excess of the 21 days. Similarly, they received a Rs50 fine for each day of the 21 days they did not attend work. A "valid" day was defined as a day in which the opening and closing photographs were separated by at least

⁷ The time and data buttons on the cameras were covered with heavy tape, and each had a seal that would indicate if it had been tampered with. Fines would have been imposed if cameras had been tampered with (this did not happen) or if they had been used for another purpose (this happened in one case, when a teacher photographed his family).

⁸ Teachers were given the 48-hour leeway to report malfunctioning cameras because not all villages have a working phone and phone services are not always reliable.

five hours and enough children (at least eight) were present in both photos to indicate that the school was actually functioning. Due to ethical and political concerns, Seva Mandir capped the fine at Rs500; hence, a teacher's salary ranged from Rs 500 to Rs 1,300. In the 60 comparison schools, teachers were paid the flat rate of Rs 1,000, and were told that they could be dismissed for poor attendance (though this happens very rarely, and did not happen during the span of the evaluation).

Seva Mandir pays its teachers every two months. In each two-month period, they collected the last roll of film a few days before the salary payment, so that the bonus or the fine was paid immediately after the end of the relevant time period. Moreover, after the first payment, teachers in the treatment schools were shown a detailed breakdown of how their payment was calculated, in order to reinforce the understanding of the program.

3. Evaluation

In this paper, we evaluate the effectiveness of Seva Mandir's incentive program in improving school quality. To do so, an independent evaluation team led by Vidhya Bhawan (a Udaipur-based consortium of schools and teacher training institutes) and MIT's Poverty Action Lab collected regular data on the functioning of the program. Data was collected to answer three basic questions: If teachers are provided with high-powered incentives to attend school that are based on external monitoring, will they attend more? If they do attend school more, will teaching time increase? Finally, will children learn more as a result?

The Poverty Action Lab collected data on teacher attendance through one random unannounced visit per month in both treatment and comparison schools. By comparing the absence rates obtained from the random checks across the two types of schools, we can determine the incentive program's effect on absenteeism. In addition, Seva Mandir provided access to all of the camera and payment data for the treatment schools, allowing us to compare absence rates measured by the random checks against those measured by the cameras. In addition to verifying whether the random checks provide a good estimate of

actual attendance rates, this comparison also allows us to verify whether teachers were simply coming to school in the mornings and afternoons for the photos, rather than attending the entire school day.

Data collected on teacher and student activity at the time of the random unannounced visit allow us to determine whether the teacher taught more as a result of the program. For schools that were open during the visit, the enumerator noted what the teachers and students were doing: if the children were sitting in class, if anything was written on the blackboard, and if the teacher was talking to the children. Since the schools have only one teacher and one classroom, these activities could be observed before the teacher and students could adjust their behavior.

Since teaching time is also a function of child attendance, student attendance data was collected at the time of the random check. After completing the observation sheet, the enumerator conducted a roll call to document which children on the evaluation roster were present.⁹ Enumerators also noted whether any of the absent children had dropped out of school or had enrolled in a government school, and then updated the evaluation roster to include new children.

To determine whether child learning increased as a result of the incentive program, in collaboration with Seva Mandir, the evaluation team administered three basic competency exams: a pre-test in August 2003, a mid-test in April 2004, and a post-test in September 2004. The pre-test followed Seva Mandir's usual testing protocol: children were given *either* a written exam (for those who could write) *or* an oral exam (for those who could not). For the mid-test and post-test, all children were given the oral exam and an opportunity to try the written exam. Those unable to write got a zero on the written section. The oral exam tested simple math skills (counting, one-digit addition, simple division) and basic Hindi vocabulary skills, while the written exam tested for these competencies plus more complex math skills (two-digit addition and subtraction, multiplication and division), the ability to construct sentences, and reading comprehension. Thus, the written exam tested both a child's ability to write and his ability to handle material requiring higher levels of competency relative to the oral exam.

⁹ Evaluation rosters were different from the school roster in that they included all children enrolled at the beginning of the experiment and all children enrolled subsequently.

Finally, detailed data were collected on teachers' characteristics to determine the extent to which the program impact on child learning varied with teacher characteristics. First, to determine whether the effect on learning depended upon a teacher's academic ability, Seva Mandir administered a competency exam to each teacher prior to the program. Second, after the program had been in place for two months, the evaluation team observed each school for a whole day, in order to assess whether the impact of the program depended on the pedagogy employed by the teachers.

III. Results

In this section, we begin by reporting the results of the baseline survey and assessing the integrity of the randomized framework (Section 1). Then, we discuss the impact of the program on teacher attendance (Section 2), child attendance (Section 3) and child learning (Section 4). Finally, in Section 5, we provide a cost-benefit analysis of the program.

1. Baseline and Experiment Integrity

Given that schools were randomly allocated to the treatment and comparison groups, we expected the quality of schooling measures before the program onset to be similar across the groups. Before the program was announced in August 2003, the evaluators were able to randomly visit 44 schools in the treatment group and 41 in the comparison. Panel A of Table 1 shows that the attendance rates were 66 percent and 63 percent, respectively. The difference is not significant. Other measures of school quality were also similar prior to the program: in all dimensions shown in Table 1 (number of students present in school at the time of visit, infrastructure, teacher qualification and performance), the treatment schools appear to be slightly better than comparison schools, but the differences are always small and never significant. The last row in the table shows the F-statistic for the joint significance of the treatment variable in all of the equations in Panel B through E. The F-statistic is 1.13, with a p-value of 0.25,

implying that the comparison and treatment schools were indistinguishable from one another at the program's inception.

Baseline academic achievement and preparedness were the same for students across the two types of schools. Table 2 presents the results of the pre-test (administered in August 2003). Panel A shows the percentage of children who could write. Panels B and C show the results from the oral and written tests, respectively. On average, students in both groups were at the same level of preparedness before the program, though there seems to be greater dispersion in the treatment schools. In the pre-test, 17 percent of children in the treatment schools and 19 percent in the comparison schools took the written exam. The difference is not significant. Those who took the oral exam were somewhat worse in treatment schools, and those who took the written exam were somewhat better in treatment schools. Again, the differences are not significant.

2. Teacher Absence

The effect on teacher absence was both immediate and long-lasting. Figure 2 shows the fraction of schools found open on the day of the random visit, by month. Between August and September, teacher attendance increased in treatment schools relative to the comparison schools. For the remainder of the program, the attendance rates in treatment and comparison schools followed similar seasonal fluctuations, with treatment school attendance systematically higher than comparison school attendance.

As Figure 2 shows, the effect of the program remained strong even after the administration of the post-test, which marked the end of the evaluation. Since the program had been so effective, Seva Mandir maintained it, but only had enough resources to keep the program running at the 60 treatment schools (expansion to all the schools is planned in the coming months). Random checks conducted after the post-test showed that higher attendance rates persisted at treatment schools even after the teachers knew that the experiment was over and that the program had become permanent. This implies that teachers did not

change their behavior simply for the duration of the experiment. In other words, there is no evidence that the program's effect is due to a Hawthorne effect.

Table 3 presents a detailed breakdown of the effect of the program on absentee rates. Columns 1 and 2 report the means for the treatment and comparison schools, respectively, over the entire period for which random checks were conducted (September 2003 to March 2005). Column 3 presents the difference between the treatment and comparison schools for this entire period, while Columns 4 through 6 present the difference for three time periods: until the mid-test, between the mid-test and post-test, and after the post-test. On average, teacher absence was 20 percentage points lower in the treatment schools than in the comparison schools. Thus, the program almost halved absence rates in treatment schools. The treatment effect was smaller for the period between the mid-test and post-test, largely because comparison school teachers attended class more often, and then rose 26 percentage points after the post-test. The reduction in the number of instances where the school was closed was much larger than that of a previous program which tried to reduce school closures by hiring a second teacher in Seva Mandir's NFEs, and reduced absenteeism by only 15 percentage points (Banerjee, Jacob and Kremer, 2005), both because individual teacher absence rates remained high and because teachers coordinated to come on the same day.

The program effects on teacher attendance were pervasive—teacher attendance increased in both high- and low-attendance treatment schools. Figure 3A plots the observed density of absence rates in treatment and comparison schools for the 20 random checks conducted during the program, while Figure 3B graphs the estimated cumulative density function of the frequency of attendance assuming that the distribution of absence follows a beta-binomial distribution. The actual and estimated distributions are very similar, indicating that the assumption of a binomial distribution is quite accurate. Both figures clearly show that the incentive program shifted the entire distribution of absence for treatment teachers. Of the 20 days, not one of 60 teachers in the comparison schools was present on all days and only one was present for 19. Almost 15 percent of teachers were absent more than half the time. On the other hand, six of the 60 program teachers were present for all days, six were present for 19 of the 20 days, and

all teachers were present at least half the time. Therefore, the camera program was effective on two margins: it eliminated extremely delinquent behavior (less than 50 percent presence), and increased the number of teachers with perfect or very high attendance records.

The evidence suggests that there were very few instances of “gaming” of the system. The fact that treatment teachers had a lower absence rate at the random checks, which were conducted in the middle of the day, suggests that teachers did not attend class at the start and end of the day just to sit for the photographs, and then leave in the intervening period. A comparison of the random check data and the camera data provides direct proof of this. Table 4 shows that for the treatment schools, the camera data tends to match the random check data quite closely. Out of the 976 cases, 82 percent had perfectly matching random check and camera data, that is, the school was open and the photos were valid or the school was closed and the photos were not valid. In 15 percent of the cases, the school was found open at the random check, but the photos indicated that the day was not considered “valid” (which is of course not an instance of “gaming”). There are 43 cases (4 percent) where the school was closed and the photos were valid, but only 19 (1.7 percent) of these were due to teachers being absent in the middle of the day during the random check and shown as present both before and after. In the other 24 cases, the data did not match because the random check was completed after the school had closed for the day, or there was missing data on the time of the random check or photo (Table 4, Panel C). These instances declined over time: during the last five months of the program, there were no cases where teachers left in the middle of the day.

Of the 131 cases (15 percent) where the school was open but the photos were invalid, it was primarily because there was only one photo (37 percent of the cases) or because the school was open for less than the full five hours (29 percent).¹⁰ This suggests that for a small number of cases, the random check may have designated a comparison school as open for the day, even though it was open for only

¹⁰ As the program progressed, teachers fully grasped that their salary was fully determined by the photographs, and as a result, discrepancies between the camera and random check data decline over time. For example, in the first three months of the program, there were 58 cases where the school was open, but the photos were invalid. For the final three months, there were only 15 such cases.

part of the school day. Therefore, since the program may also have affected the length of each school day, the random check data may, if anything, underestimate the effect of the program on total teaching time a child received. Figure 4 provides some support for this hypothesis. It plots the difference in average teacher attendance for treatment and comparison schools at the time of the random check. The figure illustrates that the difference in the attendance rate increased throughout the day, suggesting that teachers in treatment schools not only attended more often, but also kept the schools open for more hours.

2. Teacher Behavior

Though the program increased teacher attendance and the length of the school day, the program could still be considered ineffective if the teachers compensated for increased presence by teaching less. We looked at activity data collected at the time of the random check to determine what the teachers were doing while present in the classroom. It is important to note that since we can only measure the impact of the program on teacher performance for schools that were open, the fact that treatment schools were open more may introduce selection bias. That is, if teachers who tended to be absent also tended to teach less when present, the treatment effect may be biased downward since more observations would be drawn from among such low-effort teachers in the treatment group than in the comparison group. Table 5 shows that there was no significant difference in the activities of teachers in program and comparison schools during the random visit. In the comparison schools, as in the treatment schools, teachers were as likely to be in the classroom, to have used the blackboard, and to be addressing students when the enumerator arrived. This does not appear to have changed during the duration of the program (Appendix Figure 1).

The fact that, as opposed to just showing up to class more, teachers did not reduce their effort in school suggests that the fears of multitasking and loss of intrinsic motivation were unfounded. Instead, our findings confirm that once teachers were forced to attend (and therefore to forgo the additional earnings they could get by doing something else, or their leisure time), the marginal cost of teaching may not have been that large.

The teachers' general acceptance of the incentive system may be an additional reason why multitasking was not a problem: Several months into the program, teachers were asked to fill out feedback forms, which gave us a qualitative impression of the program's perception among teachers. Seva Mandir also conducted a feedback session at their bi-annual sessions, which were attended by members of the research team. No teachers complained about the principle of the program, though many teachers had some specific complaints about the inflexibility of the rules. (For example, many did not like the fact that a day was not valid even if a teacher was present 4 hours and 55 minutes—the normal school day is six hours, but an hour's slack was given to the teachers—or the fact that getting eight children to assemble on time at the beginning of the day is difficult, or the fact that the program did not plan for any sick leave or leave for extenuating circumstances, such as a funeral.) However, many felt empowered by the fact that the bonus of performing better (and being better paid as a result) was actually in their hands: "Our payments have increased, so my interest in running the center has gone up." Others described how the payment system had made others in the community less likely to burden the teacher with other responsibilities (such as being the secretary of a town meeting) once town members knew that a teacher would be penalized if he did not attend school on a given day. This suggests that the program may actually have stronger effects in the long run, as it signals a change in the norms of what teachers are expected to do.

3. Child Attendance

On the feedback forms, many teachers said that the teachers had made children, as well as themselves, attend more regularly: "This program has instilled a sense of discipline among us as well as the students. Since we come on time, the students have to come on time as well." Unfortunately,

conditional on whether a school was open, the effect of the program on child attendance cannot be directly estimated without bias, because of selection of the observations where the school was open. If schools that were typically open also attracted more children, and the program induced the “worst” school (with fewer children attending regularly) to be open more often in the treatment schools than in the comparison schools, then the selection bias will tend to bias the effect of the program on child attendance downwards. In fact, selection bias is a realistic concern since, for the comparison schools, there is a positive correlation between the number of times a school is found open and the number of children found in school.

Even so, child attendance was actually higher in treatment schools, although the difference is insignificant. In Table 6, we present the participation rates of a child in an open school, by treatment status.¹¹ While an average child’s participation rate was slightly higher in treatment schools (51 percent) than in comparison schools (49 percent), this difference is not significant. Excluding children who left the NFE, child attendance is higher overall (64 percent for treatment and 61 percent for comparison schools), but the difference is also insignificant.

Treatment schools had more teaching days. Even if the program did not increase child attendance on a particular day, the increase in the number of days the school was open should result in more days of teaching per child. The impact of the program on child instruction time is reported in Rows 3 and 4 of Table 6. Taking into account days in which the schools were closed, a child in a treatment school received 10 percentage points (or 30 percent) more days of instruction than a child in a comparison school. Assuming 27 days of work in a month (schools are open six days a week), a child obtained 2.7 more days of instruction time a month at treatment schools. Since there are roughly 20 children per classroom, this figure translates into 54 more child-days of instruction per month in program schools than in comparison schools. This effect is larger than that of successful interventions that have used child

¹¹ “Participation” subsumes both attendance and enrollment. It is the correct concept to use in an environment when being enrolled does not necessarily indicate that the child actually attends school. The participation dummy is defined for every day a random check is conducted, and is equal to 1 if the child is present on that day and 0 otherwise.

participation as the principal lever to increase child participation, such as the PROGRESA program of conditional cash transfers, which increased enrollment by 3.4 percent in primary schools and had no impact on attendance (Schultz, 2004); de-worming, which increased participation by 7.5 percentage points (Miguel and Kremer, 2004); a child incentive program (Kremer, Miguel, and Thornton, 2004), which increased participation by 5 percentage points; and a child sponsorship program, which increased participation by 8 percentage points (Kremer et al., 2004). The effect is comparable to that of adding a second teacher in Seva Mandir NFEs (Banerjee, Jacob and Kremer, 2005), which increased the number of days of instruction per month by 3.1.

In summary, since children were as likely to attend class on a given day in treatment schools as in comparison schools, and because the school was open much more often, children received significantly more days of instruction in the treatment schools. This finding suggests that the high teacher absence we observed is not likely to be the efficient response to a lack of interest by the children: if it were the case that children came to school 55 percent of the time because they could afford to attend more than a certain number of days, then we would see a sharp reduction in children presence in treatment schools on days where the school was open. On the other hand, we do not see a sharp increase in the presence of children in treatment schools despite the increased presence of the teachers. This suggests that either teacher absence is not the main cause of the irregular child presence, or that the children have not yet had time to adjust to this new pattern.

4. Effects on Learning

Children in treatment schools, on average, received 30 percent more instruction time than children in comparison schools. Over the course of a year, this resulted in 34 more days of instruction per child. Did this result lead to an increase in test scores?

4.1 Attrition and Means of Mid- and Post-Test

Before comparing test scores in the treatment and comparison schools, we must first ensure that selective attrition does not invalidate the comparison. There are three possible sources of attrition. First, a few centers closed down immediately after the program started. These closures were unrelated to the program, and equally distributed among treatment and comparison schools. We made no attempt to track the children from these centers. Second, some children leave the NFEs, either because they drop out of school altogether or because they start attending regular primary schools. Finally, some children were absent on testing days. To minimize the impact of attrition on the study, we made considerable attempts to track down children who did not show for the last two tests (even if they had left the NFE) and administered the post-test to them. Consequently, attrition was fairly limited. Of the 2,230 students who took the pre-test, 1,893 also took the mid-test, and 1,760 also took the post-test. Table 7 shows the attrition in the treatment and comparison groups as well as the characteristics of the attriters. At the time of the mid-test, attrition was higher in the comparison group than in the treatment group. At the time of the post-test, attrition was similar in both groups, and children who dropped out of the treatment group were similar to those that dropped out of the comparison group.

Table 7 also provides some simple descriptive statistics, comparing the test scores of treatment and comparison children. The first row presents the percentage of children who were able to take the written exam, while subsequent rows provide the mean exam score (normalized by the mid-test comparison group). Relative to the pre-test and mid-test, many more children, in both the treatment and comparison schools, were able to write by the post-test. On the post-test, students did slightly worse in math relative to the mid-test comparison, but they performed much better in language.

Table 7 also shows the simple differences between treatment and comparison at the mid- and post-tests. On both tests, in both language and math, the treatment students did better than the comparison students (a 0.16 standard deviation increase and 0.11 standard deviations in language at the post-test

score), even though the differences are not significant at 95 percent. To obtain more statistical precision, we control for individual pre-test results.

4.2. Test Results

In Table 8, we report the impact of the program on the mid-test (conducted in April) and the post-test (conducted in October). We compare the average test scores of students in the treatment and comparison schools, conditional on a child's pre-program competency and preparedness level. In a regression framework, we model the effect of being in a school j that is being treated ($Treat_j$) on child i 's test score ($Score_{ikj}$) on test k (where k denotes either the mid- or post-test exam):

$$Score_{ikj} = \beta_1 + \beta_2 Treat_j + \beta_3 Pre_Writ_{ij} + \beta_4 Oral_Score_{ij} + \beta_5 Written_Score_{ij} + \varepsilon_{ijk}. \quad [1]$$

Because test scores are highly autocorrelated, controlling for a child's test scores before the program increases the precision of our estimate. However, the specific structure of the pre-test (i.e. the fact that children either took the written or the oral test) does not allow for a traditional difference-in-difference (DD) strategy. Instead of a DD approach, we include a variable containing the child's pre-test score for the oral test if he took the oral pre-test and 0 otherwise ($Oral_Score_{ij}$), the child's pre-test score on the written test if he took the written test and 0 otherwise ($Written_Score_{ij}$), and an indicator variable for whether he took the written test at the pre-test (Pre_Writ_{ij}).¹² Standard errors are clustered by school. Each cell in Table 8 represents the treatment effect (β_2) obtained in a separate regression. For ease of

¹² At the pre-test, children were given either the oral or the written score. At the mid- and post-test, every child took the oral part, and every child who could write took the written exam (all children were given a chance to try the written exam; if they could not read, they were given a zero for the written test).

interpretation, the mid-test results (Columns 1 to 4) and post-test results (Columns 5 to 8) are expressed in the standard deviation of the distribution of the mid-test score in the comparison schools.¹³

The tables reveal that the program had a significant impact on learning, even as early as the mid-test. Children in treatment schools gained 0.16 standard deviations of the test score distribution in language, 0.15 standard deviations in math, and 0.17 overall (Panel A). Children with higher initial test scores gained the most from the program: those able to write at the pre-test had mid-test test scores 0.25 standard deviations higher in treatment schools than in comparison schools (Panel C). Children whose scores were below the median scores on the pre-test show no significant gains in test scores (Panel D).

The differences between students in the treatment and comparison schools persisted in the post-test (Columns 5 to 8). Children in treatment schools gained 0.21 standard deviations in language, 0.16 in math, and 0.17 overall (Panel A). Similar to the mid-test, much of the gains came from children with higher initial learning levels. The treatment effect of 0.17 standard deviations compares favorably to other successful educational interventions, such as the Tennessee Star experiment in the United States (Kruger and Whitmore, 2001), the Balsakhi Remedial Education Program in India during its first year (Banerjee, et al., 2005), and a girls' incentive program in Kenya (Kremer, Miguel and Thornton, 2004).

We compare the impact of the program on girls versus boys in Table 9. As in Table 8, we continue to compare for the pre-test scores. The first two rows of Panel A list the individual treatment effects for girls and boys, respectively, while the third row reports the difference in their treatment effects. The data shows that girls gained as much from the program as boys. On the mid-test, 7 percentage-points more of girls in the treatment schools were able to write relative to the comparison schools, compared to only 2 percentage-points of boys (this 5-percentage point difference is significant). The post-test also suggests that girls gained slightly more from the program than the boys, but these differences are not significant.

¹³ Scores are normalized such that the mean and standard deviation of the comparison group at the time of the mid-test exam is zero and one, respectively. (Specifically, we subtract the mean of the comparison group in the pre-test, and divide by the standard deviation.) This allows for comparison across samples, as well as with results from other studies. We could not normalize with respect to the pre-test score distribution since not every child took the same test at the pre-test.

4.3. Leaving the NFE

NFEs prepare children, who might not otherwise attend school, to join government schools at the age-appropriate grade level. To join a government school, children must demonstrate proficiency for a grade, either by passing an exam or through vetting by a government teacher. The ability to join schools is therefore a strong signal of the success of an NFE in fulfilling its mission. The program increased the number of children graduating to the government schools. As shown in Table 10, 14 percent of students in the treatment schools graduated to the government schools, compared to only 10 percent in the comparison schools, a 40 percent increase.

In the final row of Table 10, we present the dropout rates for children who left school entirely (i.e. left the NFE and did not join a government school). The dropout rate is slightly lower for the treatment schools, but we cannot reject the hypothesis that the difference between treatment and comparison schools is zero.

4.4 Teacher Presence on Learning

The previous sections presented the reduced form analysis of the effect of the incentives program on child learning. Table 11 interprets what these estimates can tell us about the impact of teacher presence. Columns 1 to 3 report simple correlations between teacher presence and test scores. Specifically, they report the coefficient estimate of the number of times a school was found open ($Open_j$) on a regression of the mid-test or post-test scores:

$$Score_{ikj} = \beta_1 + \beta_2 Open_j + \beta_3 Pre_Writ_{ij} + \beta_4 Oral_Score_{ij} + \beta_5 Written_Score_{ij} + \varepsilon_{ijk}. \quad [2]$$

As in the previous tables, we continue to control for the child's pre-test score and to cluster standard errors by school.

Column 1 reports OLS estimation of Equation 2 for comparison schools in order to obtain the correlation between presence and child achievement levels. In this case, the random check data is used to estimate the number of times a school is found open. The coefficient is 0.20, indicating that the test scores of children in centers open 100 percent of the time would be 0.10 standard deviations higher than those of children in a center open 50 percent of the time. The coefficient is also insignificant.

This point estimate is similar to those reported in other studies (Chaudhury, et al., 2005a) and suggests that the effect of teacher attendance on learning is not that large. Chaudhury et al. (2005a) conjectures that the measurement of absence rates based on a few random visits per school have considerable error, and may thus bias the results downwards. Consistent with this theory, the effect on the post-test scores, where having more months of random check data allows us to better estimate the absence rate per school, becomes larger (0.58 standard deviations). Our study provides a much more direct test of this hypothesis, since, for treatment teachers, the photograph data gives us the actual attendance. We present the OLS estimate of the effect of presence for treatment teachers using the random check data (Column 2) and camera data (Column 3). Overall, the effect of teacher presence is larger in the treatment schools than the comparison schools (compare 0.39 in Column 2 to 0.20 in Column 1, both obtained with random check data). More interestingly, consistent with the measurement error hypothesis, the effect of teacher presence is larger and much more significant when using the more accurate measure of presence, especially for the mid-test scores (the estimate is 0.87 standard deviations in the Column 3, compared to 0.39 in Column 2). For the post-test, where we have a much more accurate measure of presence from the random check data, the results from the two methods are instead similar (0.98 in Column 3 versus 1.17 in Column 2).

Finally, in Column 4, we pool both samples and instrument $Open_j$ (as measured by the random check) by the treatment status of the school to obtain exogenous variation in the percentage of time the school was found open in the random check. Since we have shown that the program had a direct effect on the length of the school day, as well as whether or not the school opened at all, the 2SLS estimate captures the joint effect of outright absence and of a longer school day. The 2SLS estimates are higher

than the OLS results found in Column 1, and they are indistinguishable from the OLS results in Column 3, obtained with the precisely measured absence. This suggests that the relatively low correlation between teacher absence and test scores that was observed in previous studies is indeed likely to be due to measurement error in the teacher absence data, and that reducing absence would have the potential to greatly increase test scores. Even a 10 percentage point reduction in the absence rate would result in a 0.10 standard deviation increase in test scores.

4.5. Teacher and Child Characteristics

In Table 12, we examine whether the treatment effect varies based on teacher and student characteristics. Each cell in Table 12 reports the coefficient estimate (β_4) of the interaction of being in a treated school and a school's characteristic ($Char_j$) on a regression of the test score:

$$Score_{ijk} = \beta_1 + \beta_2 Treat_j + \beta_3 Char_j + \beta_4 Treat_j * Char_j + \beta_5 X_{ij} + \varepsilon_{ijk} \quad [3]$$

X_{ij} includes controls for pre-test scores and controls for the interaction of the pre-test scores with the school characteristic. In Columns 1 and 2, we interact the treatment effect with a teacher's academic abilities at the start of the program; the treatment effect is slightly larger for teachers with higher test scores and for teachers with more years of schooling, but this effect is small and not always significant. The treatment effect does not vary based on the infrastructure level of the school (Column 3), and does not vary much based on teacher pedagogy (Column 4) or student behavior (Column 5) at the time of the school observations in October 2003. This suggests that regardless of the level of school infrastructure or teaching competency, initiating the incentive program can result in positive gains to learning.

5. Costs-Benefit Analysis

The evaluation presented in this paper shows that a straightforward monitoring and incentive program can effectively reduce teacher truancy. The benefits (in terms of child learning) of running such a program, relative to costs, are high, and comparable to other successful education programs in developing countries (evaluated with randomized evaluations).

Table 13 presents an estimate of the administrative costs of the program for one year. For the treatment schools, the average teacher salary was nearly Rs 1,000. Since the flat salary paid to comparison teachers was also Rs1,000, the program did not increase expenditure on teacher salaries. Other program costs (administration, developing the pictures, and buying the cameras) amounted to Rs5,379 per center per year. This cost corresponds to 40 percent of a teacher's yearly salary, but to only Rs268 (\$6) per child per year (assuming about 20 children per teacher).¹⁴ Expressed in terms of cost per outcome, this program cost 11 cents for each additional instruction day per child, \$60 per additional school year, and \$3.58 for increasing test scores by 0.10 standard deviations.

The cost per standard deviation improvement in test scores is higher than that of the Balsakhi Remedial Education Program evaluated in Banerjee et al. (2005). In the Balsakhi program, a second teacher (often a woman) was hired to provide remedial tutoring to children who had been identified as lagging behind their peers. The Balsakhi program resulted in a 0.14 increase for Rs 107 during its first year (and larger increases in its second year), which makes it over 2.5 times more cost effective. However, the Balsakhi program was evaluated in an urban setting, in the cities of Mumbai and Vadodara, where the external monitoring of teachers is cheaper. In contrast, the second teacher program evaluated in Udaipur district by Banerjee, et al. (2005), while it reduced school closures by 15 percent and increased the number of child-days, did not result in any improvement in test scores. The cost-effectiveness of the Seva Mandir camera program is comparable to that of other successful education programs in rural Africa: the cost per 0.10 standard deviations of the camera program (\$3.58) is similar to that of a girl's

¹⁴ This estimate does take into account the opportunity cost for teachers and children. Note, however, that the effects are larger than they could be if the program was implemented on a large scale, and more cost-effective technology (such as digital cameras) could be used.

scholarship program (\$3.53) that was evaluated in Kremer, Miguel and Thornton (2004). The scholarship program is currently the only program that has been proven to *durably* improve test scores in Africa.¹⁵

Using the estimate in Table 6, we calculate that the cost per year of schooling is $6/0.10 = \$60$ per additional year of schooling due to the program. This is much higher than the cost of the de-worming program in Africa (evaluated to be only \$3.53 per additional year of schooling),¹⁶ but lower than that of any other programs evaluated there, such as the child incentive program (\$90 per extra year), or a child sponsorship program which delivered uniforms to children (\$99 per extra year).¹⁷ It is also just over half the cost of the two-teacher program, previously implemented in Seva Mandir, which, evaluated at the current teacher's salary, cost \$115 per extra year of schooling.¹⁸ Thus, the camera program, even in its pilot form (which used an expensive way to develop photographs) is a cost-effective program compared to many others, both in terms of increasing instruction time and in terms of increasing learning.

IV. Conclusion

Addressing the startlingly high rates of teacher absenteeism in developing countries is a critical step for increasing school quality. The failure of school systems to carry out their own rules regarding teacher presence has led some to believe that only community pressure can increase school quality. However, several recent studies have shown that, for a variety of reasons, community monitoring often delivers disappointing results.

¹⁵ The test-based teacher incentive program that was evaluated in Glewwe, Ilias and Kremer (2003) had a cost of \$3.41 per 0.10 standard deviations, its gains on test scores were considered to be temporary, and it reflects gaming rather than real learning improvement.

¹⁶ In making this comparison, it is worth noting that Kremer and Miguel (2004) use the cost of the de-worming program if implemented on a large scale, whereas we use the cost of the program as implemented in this small scale pilot. However, even the cost of the program they actually evaluated was only about three times larger than what they used for the cost-benefit evaluation, which still makes the de-worming program a more cost-effective way to improve instruction time.

¹⁷ The cost per year of the PROGRESA program in primary schools is substantially larger (\$5,902.45). However, the PROGRESA program is primarily a transfer program to families, and its cost effectiveness should probably not be based on its effect on school outcomes alone.

¹⁸ The cost-effectiveness figure reported by Banerjee, Jacob and Kremer (2005) is \$4.82 per extra month, or \$58 per extra year, but the teachers were then paid Rs400, which was, according to the authors, untenable even then, in the face of competition for teachers, and was subsequently increased to Rs100.

In this paper, we show that in contrast to community monitoring, external monitoring, coupled with high-powered incentives, can be a cost-effective method to improve school quality. In particular, we show that the direct monitoring of teachers, combined with simple and credible incentives based on teacher presence can lead to a large increase in teacher attendance, even if implemented in a difficult environment. The program cut teacher absence from an average of 42 percent in the comparison schools to 22 percent in the treatment schools. As a result, students in program schools benefited from about 30 percent more instruction time. The program had a statistically and economically significant impact on test scores. After a year, child test scores in program schools were 0.17 standard deviations higher than in comparison schools, and children were more likely to be admitted to regular primary schools. Despite being implemented on a small scale, the program is cost-effective.

Our findings show that external monitoring systems can succeed in reducing absenteeism in situations where internal systems have failed. Often times, monitoring systems have failed because individuals within institutions have chosen to ignore their own rules. For example, top-down monitoring systems have been shown to fail when the headmasters are in charge of implementing them (Kremer and Chen, 2001), because the headmasters have marked the teachers present even if they were absent. In contrast, mechanical systems, such as using cameras, have the advantage of not being subject to the discretion of any one individual: a commitment at a senior level would make its implementation viable.

These results suggest that extending Seva Mandir's incentive program to other non-formal schools has the potential to increase learning levels for India's most vulnerable children. However, the one question that remains is whether it is politically feasible to implement an effective monitoring system in government schools. Since teachers in government schools are much more politically powerful than NFE teachers, it may prove impossible to institute a system where they would be monitored daily using a camera or similar device such as a date-time stamp, and other methods may prove necessary (such as having more frequent inspections). However, these results do tell us that finding ways to monitor and reward presence can in principle be effective in tackling the absence problem and improving learning. Moreover, our findings suggest that whatever barriers currently prevent teachers from attending school

regularly (distance, other activities, lack of interest by children, etc.), they are not insurmountable. Given the political will, it is therefore likely that solutions to the absence problem could be found in government schools as well. These results thus tell us both that absenteeism can be addressed, and that doing so would be worthwhile.

Works Cited

- Banerjee, Abhijit, Rukmini Banerji, Esther Duflo, Rachel Glennerster, Stuti Khemani, Sendhil Mullainathan, and Marc Shotland (2005), "The Impact of Information, Awareness and Participation on Learning Outcomes" Mimeo, Poverty Action Lab, MIT.
- Banerjee, Abhijit, and Esther Duflo (2005), "Addressing Absence," forthcoming, *Journal of Economic Perspectives*.
- Banerjee, Abhijit, Angus Deaton and Esther Duflo (2004), "Wealth, Health and Health Services in Rural Rajasthan," *American Economic Review Papers and Proceedings*.
- Banerjee, Abhijit, Suraj Jacob and Michael Kremer, with Jenny Lanjouw and Peter Lanjouw (2005), "Moving to Universal Education! Costs and Trade offs" Mimeo, MIT.
- Banerjee, Abhijit, Shawn Cole, Esther Duflo and Leigh Linden (2005), "Remedying Education: Evidence from Two Randomized Experiments in India" Poverty Action Lab working paper.
- Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, F. Halsey Rogers (2005a), "Provider Absence in Schools and Health Clinics," forthcoming, *Journal of Economic Perspective*,
- Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, F. Halsey Rogers (2005b), "Teacher Absence in India: A Snapshot," forthcoming in *Journal of the European Economic Association*.
- Education for All Forum (2000), EFA Country Assessment Country Reports.
- Fehr, Ernst, and Schmidt (2004), "Fairness and Incentives in a Multi-task Principal-Agent Model" *Scandinavian Journal of Economics* 106(3), 453-474
- Fehr, Ernst and Lorenz Gotte (2002), "Do Workers Work More if Wages are High? Evidence from a Randomized Field Experiment," University of Zurich Working Paper 125.
- Figlio, David and Lawrence S. Getzler (2002) "Accountability, Ability and Disability: Gaming The System," NBER Working Paper 9307.
- Figlio, David and Josh Winicki (2002), "Food for Thought? The Effects of School Accountability Plans on School Nutrition," NBER Working Paper 9319.
- Glewwe, Paul, Nauman Ilias and Michael Kremer (2003), "Teacher Incentives," Mimeo, Harvard.
- Glewwe, Paul, Michael Kremer, and Sylvie Moulin. (1997). "Textbooks and Test scores: Evidence from a Prospective Evaluation in Kenya", unpublished working paper.
- Holmstrom, Bengt and P. Milgrom (1991), "Multi-Task Principal-Agent Problems: Incentive Contracts, Asset Ownership and Job Design," *Journal of Law, Economics and Organization*.

Jacob, Brian and Steve Levitt (2003), "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," *Quarterly Journal of Economics*.

Kremer, Michael and Daniel Chen (2001), "An Interim Report on a Teacher Attendance Incentive Program in Kenya," Mimeo, Harvard University

Kremer, Michael, Edward Miguel and Rebecca Thornton (2004) "Incentives to Learn" NBER Working Paper #10971, December.

Kremer, Michael and Christel Vermeersch (2005), "School Committee Empowerment: Preliminary Notes," Mimeo, Harvard University.

Kreps, David (1997), "Intrinsic Motivation and Extrinsic Incentives," *American Economic Review*.

Krueger, Alan and Diane M. Whitmore (2001), "The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR," *Economic Journal*.

Lavy, Victor (2004), "Paying for Performance: The Effect of Individual Financial Incentives on Teachers' Productivity and Students' Scholastic Outcomes," Mimeo, Hebrew University.

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

Olken, Ben (2004), "Monitoring Corruption: Evidence from a Field Experiment in Indonesia," Mimeo, Harvard.

Schultz, Paul (2004), "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program" *Journal of Development Economics*.

World Bank (2004), "Making Service Work for Poor People," World Development Report, Washington and Oxford: World Bank and Oxford University Press.

Figure 1: Photographs from Program



Table 1: Is School Quality Similar in Treatment and Control Groups Prior to Program?

	Treatment (1)	Control (2)	Difference (3)
<i>A. Teacher Attendance</i>			
Percent of Schools Open	0.66	0.63	0.02 (0.10)
	44	41	85
<i>B. Student Participation (Random Check)</i>			
Number of Students Present	17.72	15.54	2.19 (2.23)
	29	26	55
<i>C. Teacher Qualifications</i>			
Teacher Test Scores	34.99	33.62	1.37 (2.01)
	53	56	56
Teacher Highest Grade Completed	10.21	9.80	0.41 (0.46)
	57	54	111
<i>D. Teacher Performance Measures (Random Check)</i>			
Percentage of Children Sitting Within Classroom	0.85	0.84	0.01 (0.09)
	29	26	55
Percent of Teachers Interacting with Students	0.79	0.73	0.06 (0.12)
	29	26	55
Blackboards Utilized	0.86	0.85	0.01 (0.11)
	22	26	48
<i>E. School Infrastructure</i>			
Infrastructure Index	3.39	3.20	0.19 (0.30)
	57	55	112
Fstat(1,115)			1.32

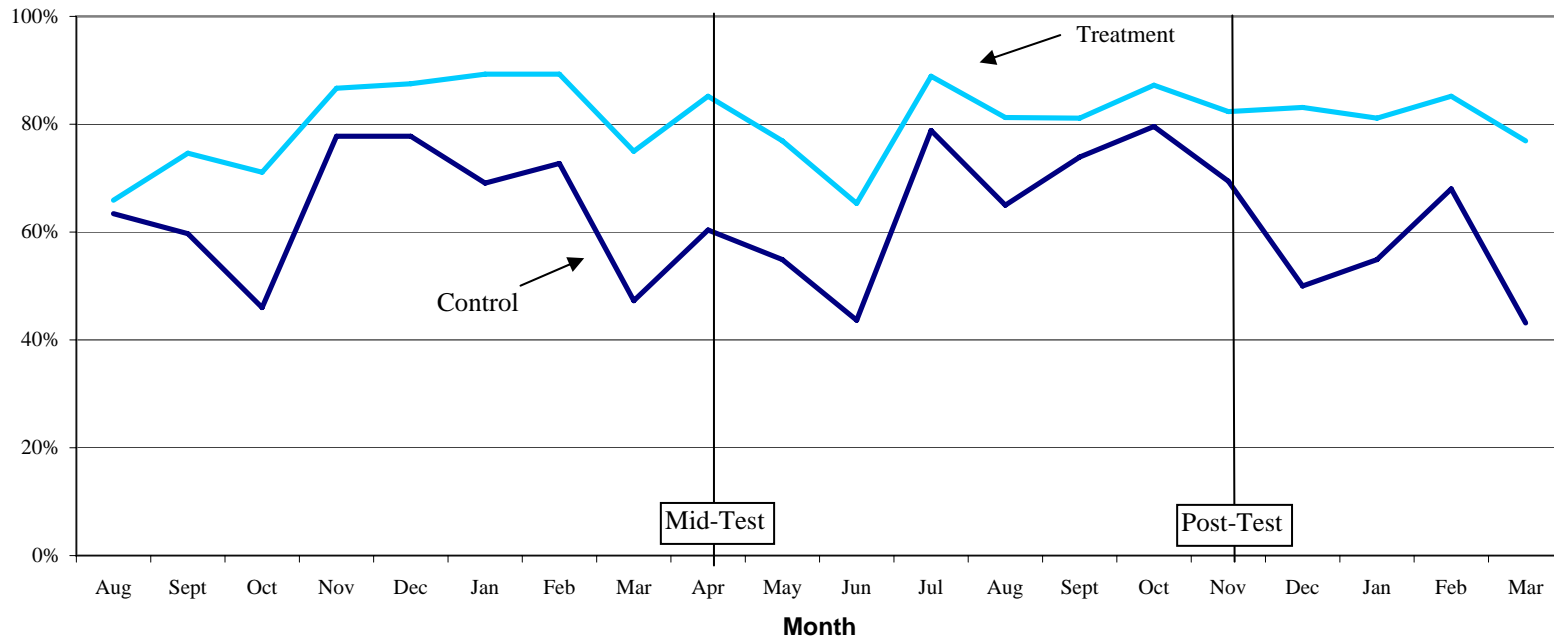
Notes: (1) Teacher Performance Measures from Random Checks only include schools that were open during the random check. (2) Infrastructure Index: 1-5 points, with one point given if the following school attribute is sufficient: Space for Children to Play, Physical Space for Children in Room, Lighting, Library, Floor Mats

Table 2: Are Students Similar Prior To Program?

	Levels			Normalized by Control		
	Treatment (1)	Control (2)	Difference (3)	Treatment (4)	Control (5)	Difference (6)
<i>A. Can the Child Write?</i>						
Took Written Exam	0.17	0.19	-0.02 (0.04)			
	1136	1094	2230			
<i>B. Oral Exam</i>						
Math Score on Oral Exam	7.82	8.12	-0.30 (0.27)	-0.10	0.00	-0.10 (0.09)
	940	888	1828	940	888	1828
Language Score on Oral Exam	3.63	3.74	-0.10 (0.30)	-0.03	0.00	-0.03 (0.08)
	940	888	1828	940	888	1828
Total Score on Oral Exam	11.44	11.95	-0.51 (0.48)	-0.08	0.00	-0.08 (0.07)
	940	888	1828	940	888	1828
<i>C. Written Exam</i>						
Math Score on Written Exam	8.62	7.98	0.64 (0.51)	0.23	0.00	0.23 (0.18)
	196	206	402	196	206	402
Language Score on Written Exam	3.62	3.44	0.18 (0.46)	0.08	0.00	0.08 (0.20)
	196	206	402	196	206	402
Total Score on Written Exam	12.17	11.41	0.76 (0.90)	0.16	0.00	0.16 (0.19)
	196	206	402	196	206	402

Notes: (1) Sample includes every student present at pre-test exam. (2) Children who could write were given a written exam. Children who could not write were given an oral exam. (3) Standard errors are clustered by school.

Figure 2: Percentage of Schools Open during Random Checks



Note: (1) The program began in Sept 2003. August only includes schools checked before announcement of program (August 25). September includes all random checks between August 26 through the end of September. (2) Child learning levels were assessed in a mid-test (April 2004) and a post-test (November 2004). After the post-test, the "official" evaluation period was ended. Random checks continued in both the treatment and control schools.

Table 3: Teacher Attendance

	Sept 2003-March 2005			Difference Between Treatment and Control Schools		
	Treatment (1)	Control (2)	Diff (3)	Until Mid-Test (4)	Mid to Post Test (5)	After Post Test (6)
Open	0.78	0.58	0.2 (0.04)	0.2 (0.04)	0.14 (0.04)	0.26 (0.05)
	1103	1086	2189			

Notes: (1) Child learning levels were assessed in a mid-test (April 2004) and a post-test (November 2004). After the post-test, the "official" evaluation period was ended. Random checks continued in both the treatment and control schools. (2) Standard errors are clustered by school.

**Figure 3A: Impact of the Cameras
(out of 20 visits)**

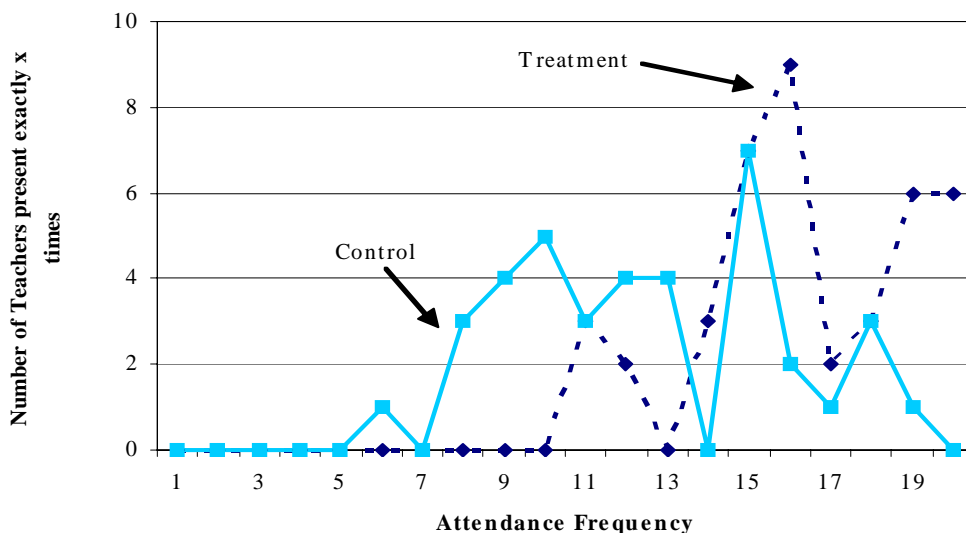
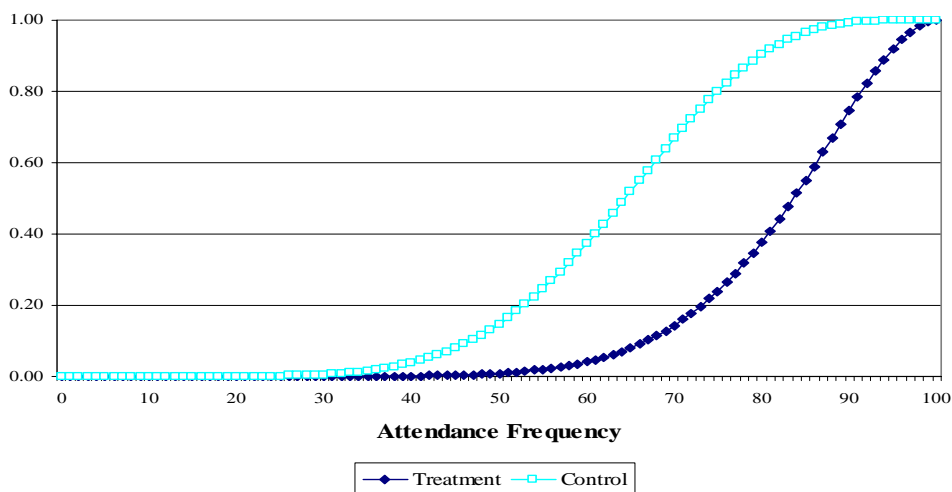


Figure 3B: Teacher Attendance



Note: Figure 3B is the estimated CDF of attendance, assuming that absence follows a beta-binomial distribution.

Table 4: Comparing Random Checks to Photo Data for Treatment Schools

Scenario	Number	Percent of Total
<i>A. Possible Scenarios</i>		
School Open and Valid Photos	673	69%
School Open and Invalid Photos	131	13%
School Closed and Valid Photos	43	4%
School Closed and Invalid Photos	129	13%
<i>B. Out of 131 where School is Open, the photos are invalid because....</i>		
School not open for full 5 hours	31	24%
Only one photo	44	34%
Not enough Children	28	21%
Instructor not in Photo	8	6%
No photograph	20	15%
<i>C. Out of 43 where School is Closed and the photos are valid.....</i>		
Random check completed after the school closed	4	9%
Teacher left in the middle of the day	19	44%
Random Check Time Missing	17	40%
Photo Data Missing	3	7%

Figure 4: Difference in the Percent of Open Schools Between Treatment and Control, By Hour

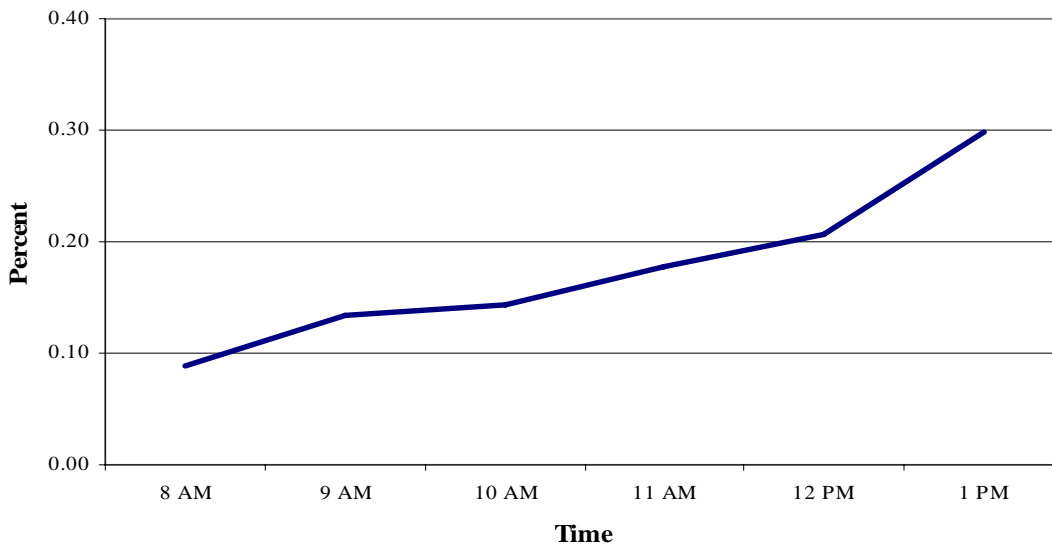


Table 5: Teacher Performance

	Sept 2003-March 2005			Difference Between Treatment and Control Schools		
	Treatment (1)	Control (2)	Diff (3)	Until Mid-Test (4)	Mid to Post Test (5)	After Post Test (6)
Percent of Children Sitting Within Classroom	0.89 865	0.88 633	0.01 (0.01) 1498	0.01 (0.02)	0.04 (0.03)	-0.04 (0.03)
Percent of Teachers Interacting with Students	0.68 865	0.69 633	-0.02 (0.03) 1498	-0.03 (0.02)	0.01 (0.02)	0.02 (0.03)
Blackboards Utilized	0.93 843	0.93 615	0.00 (0.01) 1458	-0.04 (0.04)	0.08 (0.05)	-0.07 (0.05)

Notes: (1) Teacher Performance Measures from Random Checks only include schools that were open during the random check. (2) Standard errors are clustered by school.

Table 6: Child Attendance

	Sept 03-March 05			Difference Between Treatment and Control Schools		
	Treatment (1)	Control (2)	Diff (3)	Until Mid-Test (4)	Mid to Post Test (5)	After Post Test (6)
Attendance of Students Present at Pre-Test Exam	0.51	0.49	0.02 (0.03)	0.02 (0.03)	0.03 (0.04)	0.02 (0.04)
	15536	11217	26753			
Attendance for Children who did not leave NFE	0.64	0.61	0.03 (0.03)	0.02 (0.03)	0.05 (0.03)	0.05 (0.04)
	10890	8193	19083			
Presence for Students Present at Pre-Test Exam	0.42	0.32	0.1 (0.03)	0.1 (0.03)	0.09 (0.04)	0.11 (0.04)
	18889	17194	36083			
Presence for Student who did not leave NFE	0.53	0.40	0.13 (0.03)	0.1 (0.03)	0.13 (0.04)	0.18 (0.05)
	13162	12524	25686			

Notes: (1) Standard errors are clustered at the level of the school. (2) Child attendance data collected during random check. (3) Pre-test exam determined child enrollment at the start of the program.

Table 7: Descriptive Statistics for Mid Test and Post Test

	Mid Test			Post Test		
	Treatment	Control	Difference	Treatment	Control	Difference
	<i>A. Attrition Process</i>					
Percent Attrition	0.11	0.22	-0.10 (0.05)	0.24	0.21	0.03 (0.04)
Difference in Percent Written of Pre-Test attriters-stayers	0.01	0.03	0.02 (0.06)	0.06	-0.03	0.10 (0.06)
Difference in Verbal Test of Pre-Test attriters-stayers	0.05	0.08	-0.03 (0.14)	0.02	0.12	-0.10 (0.14)
Difference in Written Test of Pre-Test attriters-stayers	-0.41	-0.23	-0.18 (0.34)	-0.19	-0.13	-0.06 (0.29)
	<i>B. Exam Score Means</i>					
Took Written	0.36	0.33	0.03 (0.04)	0.61	0.57	0.04 (0.05)
Math	0.14	0.00	0.14 (0.10)	-0.08	-0.24	0.16 (0.15)
Language	0.14	0.00	0.14 (0.10)	1.71	1.60	0.11 (0.11)
Total	0.14	0.00	0.14 (0.10)	0.35	0.24	0.12 (0.11)

Notes: (1) Test Scores in Panel B are normalized by the mean of the mid-test control. (2) Standard Errors are clustered by school.

Table 8: Estimation of Treatment Effects for the Mid- and Post-Test

Mid-Test				Post-Test			
Took Written (1)	Math (2)	Lang (3)	Total (4)	Took Written (5)	Math (6)	Lang (7)	Total (8)
<i>A. All Children</i>							
0.04 (0.03) 1893	0.15 (0.07) 1893	0.16 (0.06) 1893	0.17 (0.06) 1893	0.06 (0.04) 1760	0.21 (0.12) 1760	0.16 (0.08) 1760	0.17 (0.09) 1760
<i>B. Took Pre-Test Oral</i>							
	0.14 (0.08) 1550	0.13 (0.06) 1550	0.15 (0.07) 1550		0.2 (0.14) 1454	0.13 (0.09) 1454	0.16 (0.10) 1454
<i>C. Took Pre-Test Written</i>							
	0.19 (0.12) 343	0.28 (0.11) 343	0.25 (0.11) 343		0.28 (0.18) 306	0.28 (0.11) 306	0.25 (0.12) 306
<i>D. Below Median Rank on Pre-Test</i>							
	0.06 (0.07) 958	0.11 (0.06) 958	0.09 (0.06) 958		0.07 (0.16) 897	0.01 (0.11) 897	0.05 (0.11) 897
<i>E. Above Median Rank on Pre-Test</i>							
	0.23 (0.09) 935	0.21 (0.08) 935	0.23 (0.08) 935		0.32 (0.12) 863	0.32 (0.09) 863	0.28 (0.08) 863

Notes: (1) The table presents the coefficient estimate of being in a treated school on the sum of a child's score on the oral and written exams. All regressions include controls for the child's learning levels prior to the program. (2) The mid and post test scores normalized by mid test control group. (3) Standard errors are clustered by school.

Table 9: Treatment Effects for the Mid- and Post-Test, by Gender

	Mid-Test		Post-Test	
	Took Written (1)	Total Score (2)	Took Written (3)	Total Score (4)
Girls	0.07 (0.03)	0.2 (0.07)	0.07 (0.05)	0.18 (0.09)
	891	891	821	821
Boys	0.02 (0.04)	0.14 (0.07)	0.05 (0.04)	0.16 (0.10)
	988	988	929	929
Interaction of Female and Treat	0.06 (0.03)	0.07 (0.07)	0.04 (0.04)	0.03 (0.08)
	1879	1879	1750	1750

Notes: (1) The table presents the coefficient estimate of being in a treated school on the sum of a child's score on the oral and written exams. All regressions include controls for the child's learning levels prior to the program. (2) The mid and post test scores normalized by mid test control group. (3) Standard errors are clustered by school.

Table 10: Dropouts and Movement into Government Schools

	Treatment (1)	Control (2)	Diff (3)
Child Left NFE	0.30	0.28	0.02 (0.04)
	1136	1061	2197
Child Enrolled in Government School	0.14	0.10	0.04 (0.03)
	1136	1061	2197
Child Dropped Out of School	0.16	0.18	-0.02 (0.03)
	1136	1061	2197

Notes: (1) Standard errors are clustered at the level of the school. (2) Dropouts are defined as being absent for the last 5 random checks in which a school was found open.

Table 11: Does the Random Check Predict Test Scores?

Method:	OLS	OLS	OLS	2SLS
Sample:	Control Schools	Treatment Schools	Treatment Schools	All Schools
Data:	Random Check	Random Check	Photographs	Random Check
	(1)	(2)	(3)	(4)
<i>A. Mid-test (Sept 03-April 04)</i>				
Took Written	0.02 (0.10)	0.28 (0.08)	0.36 (0.11)	0.26 (0.19)
Total Score	0.20 (0.19)	0.39 (0.21)	0.87 (0.22)	1.07 (0.43)
N	878	1015	1015	1893
<i>B. Post-test (Sept 03 -Oct 04)</i>				
Took Written	0.24 (0.16)	0.51 (0.15)	0.59 (0.20)	0.33 (0.22)
Total Score	0.58 (0.35)	1.17 (0.36)	0.98 (0.53)	0.97 (0.47)
N	883	877	877	1760

Notes: (1) The table presents the coefficient estimate of the teacher's attendance on the sum of a child's score on the oral and written exams. All regressions include controls for the child's learning levels prior to the program. (2) The mid and post test scores normalized by mid test control group. (3) Standard errors are clustered by school.

Table 12: Interactions with Teacher Skills and Performance

	Teacher Skills				
	Test Scores	Highest Grade Completed	Infrastructure Index	Good Teacher Behavior	Good Student Behavior
	(1)	(2)	(3)	(4)	(5)
<i>A. Mid Test</i>					
Took Written	0.00 (0.00)	0.03 (0.01)	0.00 (0.02)	0.08 (0.06)	0.11 (0.06)
Total Score	0.01 (0.01)	0.06 (0.03)	0.01 (0.04)	0.07 (0.12)	0.17 (0.12)
<i>B. Post Test</i>					
Took Written	0.00 (0.00)	0.01 (0.02)	0.04 (0.03)	0.08 (0.08)	-0.08 (0.08)
Total Score	0.01 (0.01)	0.02 (0.04)	-0.05 (0.04)	0.14 (0.18)	-0.08 (0.18)

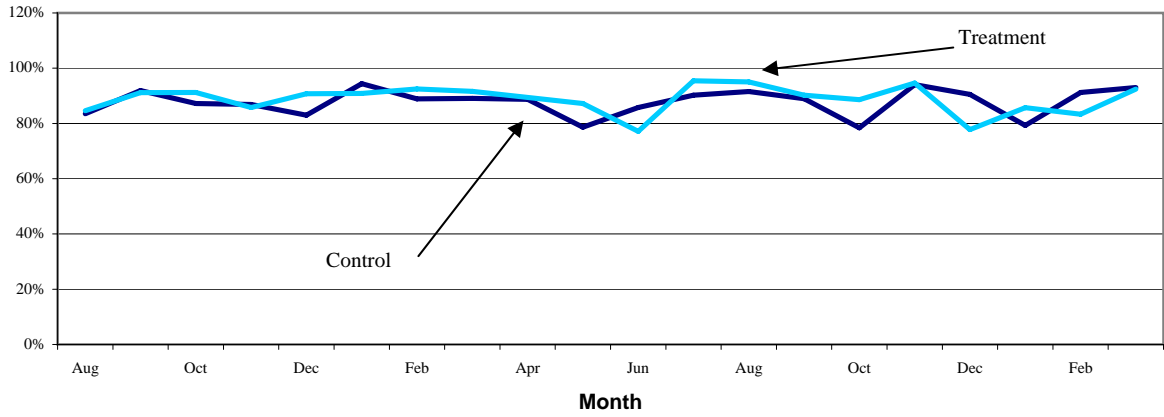
Notes: (1) Standard Errors are clustered by school. (2) Teacher observations were conducted in September thru October 2004. (3) Test Scores and Highest Grade Completed are in levels. The Infrastructure Index is the same as in Table 1. The Teacher and Student Behaviors are measured as being above the median in terms of each behavior.

Table 13: Cost of Program Per Center over 12 Month Period

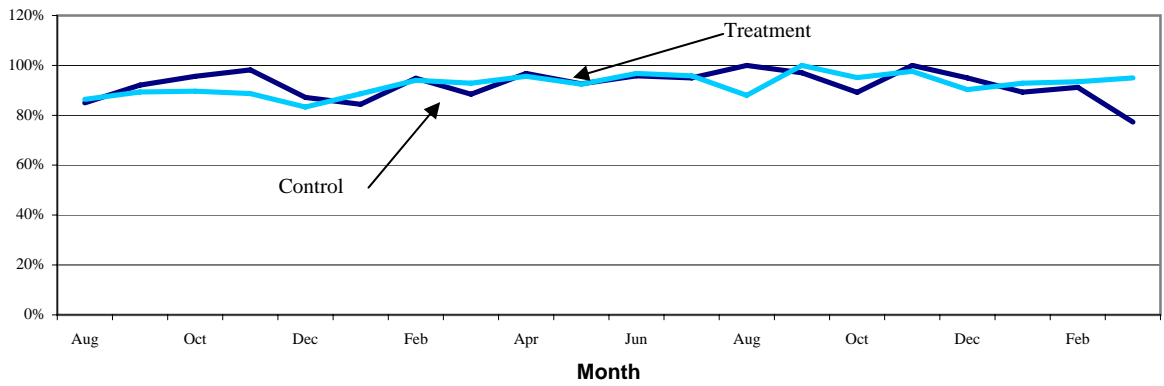
Item	Cost
<i>A. Camera Cost</i>	
Camera Cost ¹	1133
Film Cost	1392
Battery Cost	552
Photo Development and Printing:	1852
<i>B. Salaries</i>	
Teacher Salaries ²	0
Labor Cost to Run Program ³	450
Total Costs to Run Program	5379

Notes: (1) Assumes cameras last 3 years (2) Average Teacher Salary is Rs1000 under program. In the absence of the program, it would be Rs1000. (3) It takes approximately 50 man hours to process 115 schools per month. Assume a staff worker being paid Rs 10,000 per month and works a 40 hour week. Thus, it takes 1/2 hour of labor at Rs37.5 to complete one center per month.

Appendix 1A: Ratio of Students Inside the Classroom to Outside the Classroom



Appendix Figure 1B: Blackboards Used



Appendix Figure 1C: Teacher Interacting with Student

