Child Control in Education Decisions: An Evaluation of Targeted Incentives to Learn in India

James Berry*,†

Cornell University

Abstract

I report the results of a field experiment in Gurgaon, India, in which the recipient of an incentive to learn was randomly assigned either to the parent or to the child. While I find no evidence that the identity of the recipient matters in the aggregate, incentives to parents are more effective for initially high-performing children, while the reverse is true for low-performing children. To explore the mechanisms behind this result, I present a model of household education production and find additional empirical results consistent with the model.

*I am especially grateful to my advisers Esther Duflo, Abhijit Banerjee and Ben Olken as well as Nava Ashraf, Neil Bhutta, Camilo Dominguez, Greg Fischer, Raymond Guiteras, Kirabo Jackson, Cynthia Kinnan, Jeanne Lafortune, Leigh Linden, Karthik Muralidharan, Trang Nguyen, Jesse Shapiro, Christopher Smith, David Tabak, José Tessada, Tom Wilkening, and numerous seminar participants for helpful comments and suggestions. I would also like to thank Hansa Parmar for her tireless work coordinating the field activities in Gurgaon, Swati Gaur for research assistance, the staff at Pratham Gurgaon, especially Brij Kaul, Pawan Sharma, and Rajeev Ranjan, and the Gurgaon district officer for his support. All errors are my own. This project was supported by the Russell Sage Foundation Small Grants Program in Behavioral Economics and the George and Obie Schultz Fund.

†E-mail: *jimberry@cornell.edu*

1 Introduction

An increasingly popular intervention to encourage schooling behavior in developing countries offers cash rewards or other incentives to households when their children enroll in, attend, or achieve in school (UNESCO, 2010). While interventions that provide these incentives condition the transfer on different behaviors (enrollment, attendance, or achievement), those that target families in developing countries primarily provide the rewards to parents. Implicit in the design of these interventions is the idea that rewarding the parents, rather than the children, produces the best results. In contrast to a large literature suggesting that the recipient of a transfer matters between adults within a household (Duflo, 2003; Lundberg et al., 1997), there is very limited evidence on whether it is more effective to target parents or children.[1]

As with transfers received by different adult members of the household, transfers received by parents or children could yield different results because of frictions within the household's decision process. For example, if child effort is an important input into education production and if parents are unable to motivate their children to the full extent to which they are themselves motivated by incentives, then incentives provided to children may result in better outcomes than incentives provided to parents. As I explore in this paper, the transfers that parents provide their children may be dampened either by agency problems within the household or by parents' inability to commit to rewarding their children for positive outcomes.

In this paper I fill this gap in the literature by presenting the results of a field experiment conducted with primary school students in urban slums in Gurgaon, India. The experiment offered prizes to the parent or to the child if the child reached a literacy goal after two months. Each child was given a goal based on his pretest score and was tested again after two months to determine if the goal had been reached.[2] In order to isolate the effects of changing control over the rewards between the parent and child, program families were randomly offered either money given to the child's mother ("parent money"), money given to the child ("child money"), or a toy of equivalent value

[1]The one other study that I am aware of that directly examines this issue, Baird et al. (2011), is discussed below.
[2]Throughout the paper, I use masculine pronouns to refer to the child and feminine pronouns to refer to the parent.

given to the child as a reward. Because the parent could not easily capture the toy, and because baseline survey responses indicated that money given to the child would be simply transferred to the parent, I focus on the difference between the aggregated money treatments and the toy treatment in the analysis.

The experiment also tested whether parents want to reward their children for positive outcomes but cannot commit to doing so. For example, parents might be unable to credibly commit to reward their children for good performance on a test because they are unable to put the resources aside to purchase the reward.[3] To test this hypothesis, the experiment included an additional treatment that offered the parent a choice of money for herself or a toy for her child as the incentive. The parents therefore had the opportunity to commit to reward their children with toys if the goal was achieved.

All children in the program, regardless of treatment, were given the opportunity to attend free after-school classes to assist them in improving their reading skills. These classes were held to give the children a greater chance to achieve the goals set out by the program. In addition, attendance in these classes provides an objective measure of effort and serves as an intermediate outcome in the analysis.

The key finding of the analysis is that the identity of the recipient has a significant effect on educational outcomes. I first show that the incentives had a substantial impact on test scores across all incentive treatments.[4] While I find no significant differences in average attendance in the after-school classes or achievement of the literacy goal between the money and toy treatments, this result masks important heterogeneity in treatment effects by pretest score. Children with lower pretest scores perform better when provided a toy as an incentive relative to money, while the reverse is true for children with higher pretest scores. These results hold for attendance in the supplemental classes as well as for achievement of the literacy goal. Finally, I find no evidence that offering parents the opportunity to commit to rewarding their children with toys improves outcomes.

[3]This hypothesis was formed based on focus group discussions conducted during pilot activities.

[4]While institutional constraints prevented the inclusion of a randomized control group, I construct a control group using households included in the randomization but not reached for the program announcement. As I show in Section 3, this group had strikingly similar pretest scores to the program group.

To explore the mechanisms behind the heterogeneity in effects by pretest score, I hypothesize that the child's test score may reflect the parent's ability to teach her child and motivate her child to learn. When a parent is more able, a parent incentive will be more effective than a child incentive. To develop this hypothesis further, I present a model of education production in which both the parent and child exert costly effort. Households vary by parental ability, which is modeled as the productivity of the parent's input. The model implies that higher parental productivity will result in both higher initial test scores and more effective parent incentives relative to child incentives. As a result, the interaction of parent incentives with the component of test scores attributable to parental productivity will be stronger than than the interaction with overall test scores.

I then test these implications empirically. To isolate the variation in test scores attributable to parental productivity, I use the predicted values from a regression of initial test scores on parental characteristics. Consistent with the model, I find that the interaction of the toy treatment with the portion of test scores related to parental productivity is more negative than the interaction with overall test scores.

This paper offers several contributions to existing research. It is one of only two studies that I am aware of that directly compares the effectiveness of incentives given to parents with incentives given to children. Baird, et al. (2011) evaluate the effectiveness of a conditional cash transfer program for girls in Malawi that provided transfers to both parents and girls, varying the amounts given to each recipient. Conditional on the minimum level of transfers, the amount given to each recipient has no significant effect on enrollment or test scores.

My study complements the findings of Baird, et al. (2011) in a number of ways. First, in Baird, et al. (2011), the transfers are always in the form of money, and both the mother and child always receive some transfers. Because money may be pooled in the household, I vary both the *recipient* and the *type* of reward in my study. I also offer the entire reward to either the parent or child in order to create a stronger contrast between the types of incentives. Second, I explore the distributional implications of the different incentives and find that lower-performing children have better outcomes when the incentives are provided directly to them. Finally, Baird, et al. (2011)

study a conditional cash transfer program that rewarded enrollment and attendance in school, while my study examines an incentives-to-learn program. While conditional cash transfer programs are popular in the developing world, they may not always be appropriate when enrollments are already high and when learning is a key constraint to success in school. While net primary school enrollment in India improved from 79% in 2000 to 92% in 2008 (World Bank, 2013), performance in school still lags. A recent survey found that only 39% of children in grade 3 could read at a first-grade level (ASER Centre, 2013).

This paper also adds to a new but growing literature evaluating incentives-to-learn programs in developing countries. Consistent with my results, this prior research has found that incentives programs are effective in improving school performance. Kremer, Miguel and Thornton (2009) evaluate a cash incentive program for primary-aged girls in Kenya and find a significant impact on learning outcomes of girls in the program. Blimpo (2010) finds significant gains in exam scores as a result of individual and team incentives for performance on secondary-school certification exams in Benin.[5]

More generally, this paper contributes to the theoretical and empirical literature on how parents provide incentives to their children. Becker's Rotten Kid Theorem (1974) provides an early theoretical foundation for this line of research. The Rotten Kid Theorem shows that under certain assumptions, a parent can control her child's actions indirectly through transfers to her child. External incentives provided to the parent will therefore produce equivalent results to incentives provided to the child. However, the Rotten Kid theorem does not hold under moral hazard (Bergstrom, 1989), an important "real-world" feature that I incorporate into my model.[6] A recent study by

[5]Studies of developed-country incentives programs find mixed evidence on their effectiveness. Angrist and Lavy (2009) find a significant impact on girls, but not on boys, in an evaluation of rewards for performance on high-school matriculation exams in Israel. Angrist, Lang and Oreopoulos (2009) find that a program rewarding good grades in a Canadian university raised grades for women, but not for men. Jackson (2010) finds that a program combining teacher and student incentives for student performance on Advanced Placement exams in the United States had a significant impact on exam performance. Fryer (2010) evaluates incentives programs in schools in three U.S. cities and finds no evidence that the programs affected student achievement. Bettinger (2010) finds that an incentive program for primary school students in Ohio affected scores in math, but not in other subjects.

[6]Gatti (2005) explores the theoretical implications of a moral hazard model for bequests and intergenerational transfers between parents and children. Weinberg (2001) finds empirical evidence in favor of a moral hazard model by examining the relationship between household income and the use of corporal punishment.

Bursztyn and Coffman (2012) finds that parents in a Brazilian conditional cash transfer program prefer conditional transfers over unconditional transfers. The results suggest that information asymmetries between parents and children may lead parents to desire methods to help them control their children's schooling behavior.

The remainder of this paper is organized as follows. Section 2 describes the design of the intervention and the outcome measurement. Section 3 presents my findings on the effects of the treatments on class attendance and test scores. Section 4 lays out a model of education production that provides a framework to analyze the heterogeneity in impacts. Section 5 tests the implications of the model. Section 6 concludes.

2 Experimental Design

The intervention was conducted from July through September of 2007 in Gurgaon, a suburb of Delhi. The main intervention consisted of a pretest, announcement of the child's incentive scheme, and a post-test approximately two months later. Children were initially tested in schools to determine baseline learning levels. Each child was given a goal competency based on his pretest score. Children were then re-tested using a similar testing instrument after two months. If the child achieved the goal, he or his parent would receive a prize. The treatments were randomly assigned at the individual level, after stratifying by pretest score. The prize value was set at 100 rupees (about $2.50 at the prevailing exchange rate) for all treatments. At the time of the study, one hundred rupees was the approximate daily wage for an unskilled laborer in these areas.

Eight government-run primary schools were selected based on proximity to the city center and availability of public transportation nearby. In seven schools all 1st, 2nd and 3rd grade students participated in the program. In one school, 1st-grade children were excluded due to administrative difficulties in obtaining these students' addresses.[7]

[7] The sample also includes 38 students from two additional schools, all living in a slum community near the city center. These children were tested at home, rather than at school. The results are robust to the exclusion of these additional students.

2.1 Treatment groups

The experiment exogenously varied the type of reward along two dimensions: the direct recipient of the reward (either the parent or child) and the form of the reward (either money or a toy). The parent money treatment offered a reward of 100 rupees to the parent if the child achieved the goal. This represents the case in which the parent has complete control over the reward. The child money treatment offered a reward of 100 rupees to the child if the child achieved the goal. If the parent and child consider money given to the child as earmarked for child consumption and there is no compensating behavior by the parent, this treatment would represent more control over the reward by the child; however, if income from the parent and child is pooled in the household, this treatment would be equivalent to the parent money treatment.

Evidence from the baseline survey suggests that in this context money given to the child was indeed pooled within the household. When asked what they would do if given 100 rupees, over 80% of children in the sample reported that they would give it directly to their parents. In addition, while 51% of parents reported giving their children spending money within a day of the survey, the majority (74%) of the time the money was given specifically for food items. This suggests that money was rarely given to the child to be spent at his own discretion.

The next treatment offered a reward of a toy valued at 100 rupees to the child if the child achieved the goal.[8] Rewarding the child with a toy gave him control over the reward in two ways. First, it gave the child an item whose value could not be easily appropriated by the parent. While the parent still retained the right to take the toy away from the child, it would have been difficult to sell the toy and convert its value to other household consumption. The value of the toy was also high enough that the parents were unlikely to be able to adjust the child's consumption of these goods. Only 4% of parents reported having given their child a toy during the week before

[8]Households in the toy treatment was randomly split into 2 sub-treatments. In the "child toy" treatment, the child was given a choice between 5 different toys, each with a retail value of 100 rupees. In the "voucher" treatment, the child was given a voucher worth 100 rupees that was redeemable at a local toy store. The voucher treatment was included in case the toys offered in the child toy treatment were unpopular. In practice, however, the limited number of toys selected for the child toy treatment proved to be very popular, and the shopkeepers reported that those redeeming the voucher often chose toys that were available in the child toy treatment. Therefore, these treatments are combined for the purposes of the analysis.

the baseline survey, and anecdotal evidence suggests that the value of these toys was substantially less than that of the toys offered as part of the program. Second, as in the case of the child money treatment, the toy was given directly to the child.

In order to test whether parents want to reward their children for positive outcomes but cannot commit to doing so, I included two additional treatments that offered the parents a choice between money for themselves and a toy for their children. In *ex ante* choice treatment, the parent made her choice when the program was announced. In the *ex post* choice treatment, the parent made her choice after the child had reached the goal.

The *ex post* choice treatment was included to confirm that the salience or the convenience of the choice itself does not cause a positive impact of the *ex ante* choice. If the choice itself drives the results, one would expect a positive impact of either choice treatment on outcomes. On the other hand, if the results are driven by an actual desire to commit, only the *ex ante* choice treatment will positively affect outcomes. In addition, the *ex ante* and *ex post* choice treatments can serve to check for consistency in choices. If parents desire to commit *ex ante* because they know they will not reward the child *ex post*, they will be more likely to choose money *ex post*.

Regardless of treatment category, all children were invited to attend free after-school classes run as part of the program. The classes were led by teachers trained to assist the children in achieving their literacy goals. The profile and training of the teachers followed the para-teacher model of Pratham, a large India-wide NGO specializing in literacy and numeracy (see Banerjee et al., 2007). In each school, enough teachers were provided so that there was at least one teacher for every 20-30 students who attended the classes.[9] Classes ran for three hours every afternoon that school was in session. Children were free to attend on a drop-in basis, and teachers were given flexibility to customize lessons based on the reading levels of the children who attended. Tutorials held outside of school hours are common in India, and thus the extra classes provided a learning environment familiar to the households in the study.

There were two primary reasons for including the classes. First, government schools in India

[9] In one school, the principal did not allow the after-school class teachers access to the school premises to conduct the class, and no suitable alternative location was found.

are often poor platforms for learning, and the classes provided a greater chance for the children to reach the goals set by the program. Second, the classes present a unique opportunity to measure effort. Existing studies typically rely on attendance in school, either copied from the school's administrative records or collected through random, unannounced checks by outside surveyors. Administrative records are notoriously inaccurate in India, as schools may have incentives to inflate recorded attendance (Shastry and Linden, 2011). Random checks are usually unable to measure attendance on a daily basis, since they disrupt class and are difficult to take accurately. Teachers of the after-school classes in this study were monitored carefully to ensure accuracy.

2.2 Pretest

Children were initially tested for reading ability during school time. The test used an instrument developed by Pratham and used in national assessments of child reading ability (ASER Centre, 2013). Each child was evaluated on a five-point scale: 0) the child could not recognize letters, 1) the child could recognize letters, 2) the child could read simple words, 3) the child could read a simple paragraph, and 4) the child could read and understand a multi-paragraph story.

Based on each child's pretest score, the child was given a goal competency to be reached when he was re-tested after two months. If the child achieved the goal, he or his parent would receive a prize as per his treatment category. Children reading at levels 0 and 1 were each given a goal one level above their respective pretest scores, while children at levels 2 and 3 were given a goal of 4. Goals were selected such that approximately half of the children would reach the goal, based on a pilot study. Children at the highest reading level (4) at the pretest were excluded from the study and were instead given an unconditional prize at the end of the program.

In the analysis I use two measures of test scores. First, I use the simple categorical variable described above. Second, I construct a continuous measure of the child's test score, relative to the other children in his grade, by adding a fraction to the categorical level reflecting the number of correct or incorrect answers that the child gave in the next level.[10] The continuous measure is then normalized by subtracting the mean pretest score in the child's grade and dividing by the standard deviation.[11]

2.3 Baseline Survey and Program Announcement

After the pretest, each child was randomly assigned to one of the five treatment groups outlined in Section 2.1. The randomization was conducted at the individual level and was stratified by pretest score within each school, grade and classroom.

Approximately one week after the pretest, a baseline survey was conducted at the child's home, and the household's assigned incentive scheme was announced to the mother and child.[12] The survey and the program announcement were conducted with the child's mother rather than his father because pilot surveys indicated that the mother was usually more involved in the child's education. The survey collected demographic information as well as information on the transfers that the parent had made to her child over the past week.

After finishing the survey, the surveyor read a script announcing the incentives program to the mother and child. The script was individualized based on each child's treatment group and the child's literacy goal. The mother and child were informed that the child would be tested again in school after two months, and if the goal competency was reached, the mother or child would receive the specified prize. In addition to the announcement of the incentive scheme, the mother and child were informed that extra classes would be conducted after school in order to assist the child in reaching the goal.

[10] To reach letter and word levels, the testing procedure requires that a child must read at least 6 letters or words on the testing instrument. Therefore, for students who had fewer than 6 correct answers, each correct answer was given 1/6 of a point. For paragraph and story levels, a child must make less than 3 mistakes in the respective sections. Therefore, three mistakes were given 0.75 points, 4 mistakes were given 0.5 points, and 5 mistakes were given 0.25 points.

[11] In order for the grade-specific means to be representative of all children in the schools, I include the highest pretest scores in constructing this variable.

[12] Announcement and reminders of the incentive schemes were done at each child's home to minimize contamination across treatment groups. As described below, reminders were also done on an individual basis, and at the time of the reminder, an individualized card was given to each household reminding them of the program and the assigned prize.

Table 1 shows the sample composition by pretest score, grade and treatment category. Out of 1466 children who took the pretest, 331 were excluded from the study because they achieved the highest possible test score, and 49 others were excluded because they lived too far from the schools, making surveying impractical. 1086 children were thus available for the randomization. 85% of children out of the randomized group of 1086 were reached for the baseline survey and program announcement. The attrition between the randomization and program announcement was primarily due to difficulty in locating the children's homes and in reaching the parents at home.[13] Of the 925 children offered the program, 900 (97%) took the post-test after two months. Most of the 25 students who were not available for the post-test had moved away since the program announcement.[14]

Table 2 confirms the effectiveness of the randomization by examining correlations between treatment status and household characteristics, as well as the incentives that the parent had given the child at the time of the baseline survey.[15] Column 2 presents the p-values of a joint test of equality of means of each characteristic across the five treatment categories. All 14 p-values are above 0.1. The p-value is close to 0.1 for the amount of money given to the child over the past week, but this is to be expected given the large number of comparisons in the table. Column 3 presents p-values for the comparison of means of the toy treatment with the aggregated parent and child money treatments. Again, none of the differences is significant at the 10% level. While both mother's and father's education are almost significant at this level, these slight correlations are to be expected given the number of comparisons performed.

[13] Schools in Gurgaon do not keep detailed information on the addresses of their students. It was therefore necessary to have every child in the study show the surveyor his home at the time of the pretest so that the surveyor could note the child's address information. In some cases, the children were not available to show the surveyors their homes.

[14] Attrition between the program announcement and post-test is not significantly related to either assignment to the five treatment categories or to the broader money and toy categories described below. Given the small levels of attrition and the lack of correlation with treatment groups, it is reasonable to conclude that attrition does not bias the achievement results. The attendance results include all children offered the program regardless of whether they took the post-test.

[15] The mother and father education variables were mistakenly excluded from the baseline survey and had to be measured at the second follow-up. Since these are objective measures, however, it is unlikely that survey responses were biased by the treatments.

2.4 Reminder and Follow-up Survey

Surveyors visited the schools approximately one month after the program announcement to remind children individually of the program. Each child was given a card to take home specifying his goal, prize and the date of the post-test.

Approximately 10 days before the post-test, surveyors visited the homes of children in the money and toy treatments to conduct another survey that collected information on transfers between parents and children.[16] At the end of the survey, a short script was read reminding the parent and child of the program.

2.5 Post-test and Second Follow-up Survey

Approximately two months after the program announcement, a post-test, similar in form but not exact content to the pretest, was conducted in the schools. Prize distribution took place the day after the post-test either at school or at the child's home, according to the child's treatment category.[17] Approximately one week after the post-test, a second follow-up survey was conducted to again measure transfers given by the parent to her child. The purpose of this survey was to examine how parents reacted after the post-test had been conducted and rewards had been distributed.

3 Results

The following three subsections present the main results of the experiment. Section 3.1 shows that the incentives treatments had large and significant impacts on child test scores. Section 3.2 compares average outcomes across the incentives treatments and finds no significant differences between the money, toy and choice treatments. Section 3.3 presents evidence of heterogeneity in treatment effects by initial test score: lower-performing children had better outcomes in the toy

[16] The treatment effects on self-reported transfers are analyzed in Appendix A.

[17] Prizes for the child money treatment, toy treatment, and toys chosen in the in the *ex ante* treatment were distributed in school. Parents in the parent money treatment and those who chose money in the *ex ante* treatment were given the money at home. Parents in the *ex post* choice treatment were also visited at home the day after the post-test and were given the prizes upon making the choice.

treatment, while higher-performing children had better outcomes in the money treatments.

3.1 Overall Effect of the Incentives Program

The overall effects of incentives programs are important from a policy perspective, and it is useful to show that the incentives used in this study did influence test scores overall. In order to gain support of the local school committee, the study was designed to treat all eligible children and therefore did not include a pure control group. However, features of the implementation of the program allow me to construct a quasi-experimental control group to estimate the overall effects of the incentives schemes on test scores.

As shown in Table 1, a number of children and their mothers in the randomized sample were not reached at the time of the baseline survey. Some children were not in school when addresses were initially collected, but they were included in the randomization in case they could be found at a later date. In other households, the children and their mothers were not available during the surveyors' 2-3 daytime visits when the baseline was conducted. There were 161 students included in the randomization but who were not reached for the program announcement. Out of these 161 students, 152 were in school when the post-test was administered. This group (the "no program" group) will serve as a control group for the analysis of this section. The results in this section include both children whose addresses were not collected and those who were not available at the time of survey, but restricting the sample to either of these groups leaves the results largely unchanged.

Figure 1 presents the distributions of raw pretest and post-test scores of the program group and the no-program group. As shown in the top panel, the pretest scores are remarkably similar between the two groups. A Pearson χ^2 test fails to reject the equality of the two distributions (p-value = 0.88). The bottom panel presents the post-test scores of the two groups. The program group now has a much lower proportion of test scores of zero and higher proportions of test scores of 2 and 4. There is a lower proportion of scores of 3, but this result is not surprising given that this score category was not one of the goals given to students in the program group. A Pearson χ^2

test now strongly rejects equality of the two distributions (p-value < 0.01).

Table 3 presents the results of regressions of a dummy variable that indicates if the mother and child were reached at the baseline on the child's relative pretest score and the set of characteristics that was observable for both the program and no-program groups. Column 1 indicates that being reached at the baseline is not significantly related to the pretest score, grade or gender of the child. Columns 2 and 3 add controls for classroom and surveyor dummies.[18] Of the three specifications presented in this table, the only significant difference in test scores between the program and no program group is the specification with both sets of dummies, displayed in Column 3. The point estimate in this specification indicates that an increase in pretest score by one standard deviation is associated with a 2.6% lower likelihood of being reached at the baseline. Overall, however, the results of this table imply few differences between the program and no-program groups.

The results in Figure 1 and Table 3 suggest that although the no-program group was not randomly assigned, children in this group may serve as a plausible control group for the purposes of determining the overall impact of the program.[19] Table 4 presents the difference-in-differences estimates of the effects of participation in the program on test scores. The first and second rows display the estimates using raw and relative pretest scores, respectively. The estimated effects of the program are 0.48 points using raw test scores and 0.53 standard deviations using relative test scores. These results are significant at the 1% level.[20]

The remaining rows of Table 4 display the estimated program effects on the probability that students fall into each test score category. Column 3 shows that there are no significant differences in the fraction of students in any of the these categories at the pretest. Consistent with Figure 1, the difference-in-differences estimates imply that the program group decreased the proportion of children at the nothing level (significant at the 1 percent level) and increased the proportion of students at the word and story levels (significant at the 10 and 5 percent levels, respectively).

[18]The female dummy is not included in Columns 2 and 3 of Table 3 because the majority of classes are organized by gender.

[19]The improvement in test scores in the no-program group during the intervention is almost identical to an estimate of the counterfactual improvement for the program group based on cross-sectional differences in test scores between grades. See Appendix B for details.

[20]Appendix B compares these results with those found in Kremer, Miguel and Thornton (2009) and Blimpo (2010).

One caveat is in order with respect to interpretation of the results presented in this section. While the after-school classes were open to any child who wished to attend, children in the incentives program were notified individually when the program was announced. In practice, children in both the program and no-program groups were often reminded of the classes during school time by their teachers, but this was not controlled as part of the experiment. Therefore, the program effects estimated in this section are the combined effects of receiving an incentive treatment in addition to individual notification of the classes. However, several pieces of evidence suggest that individual notification of the classes does not drive the program effects found in this section. First, while 24% of children in the incentive treatments attended the classes, 13% of children in the no-program group attended. Based on this 11% difference in attendance, the classes would have to have been responsible for an improvement of over 1 standard deviation in order to account for just 25% of the estimated program effect. It is therefore unlikely that differential class attendance between the program and no-program groups could be driving a substantial fraction of the estimated program effect. Second, running the analysis separately for the 176 second- and third-grade students in the one school that did not have after-school classes, the estimated effects of the program equal 0.46 points using raw scores and 0.29 standard deviations using relative scores. Both estimates are significant at the 1% level. These estimates are almost identical to the effects on second- and third-graders over the entire sample, which equal 0.41 points and 0.30 standard deviations.

3.2 Comparison of Treatment Groups

3.2.1 Money and Toy Treatments

As noted earlier, survey responses indicated that most parents have discretion over money received by their children. This suggests that in the context of the experiment, the *form* of the reward, either money or a toy, may be a more relevant dimension of variation than the *recipient* of the reward. I therefore focus on the comparison of the pooled money and toys treatment in the analysis. Panel A of Table 5 presents OLS estimates comparing average outcomes across these groups. The first two columns use attendance in the after-school classes as the outcome. The dependent variable is a dummy indicating any attendance in the after-school classes over the two-month period in which they were held. Columns 3 and 4 present the results of similar regressions using achievement of the goal competency as the outcome variable. Comparing the toy and money treatments, there are no significant differences between categories, using either attendance or achievement as the outcome.

Panel B of Table 5 disaggregates the two money treatments and compares average outcomes in the child money and toy treatments with the parent money treatment. The differences between treatment groups are again small and insignificant, although less precisely estimated than those of Panel A.

3.2.2 Choice Treatments

As described above, the *ex ante* and *ex post* choice treatments test whether allowing parents to commit to rewarding children can improve experimental outcomes. The choices of the parents in these two treatments can provide preliminary evidence of a commitment problem on the part of parents. If parents cannot commit and are aware of this problem, they will be willing to reward their children with a toy *ex ante* but will decide to keep the money for themselves *ex post*. Therefore, if the choice of a toy is primarily driven by a desire for commitment, one would expect more parents to choose the toy in the *ex ante* treatment than in the *ex post* treatment.

However, the results show that substantially *more* parents in the *ex post* treatment chose to reward their children with toys. Table 8 tabulates the choices in both treatments. Thirty-three percent of parents in the *ex ante* treatment and 51% of parents in the *ex post* treatment chose toys. The percentage of parents in the *ex ante* treatment choosing the toy remains virtually unchanged (32%) when restricting the sample to those who achieved the goal. Because more parents chose the toy in the *ex* post *treatment,* these results provide suggestive evidence against a commitment problem on the part of parents. Instead, it is possible that parents initially chose money because they were uncertain of their needs for cash at the end of the program. Once the uncertainty was resolved, they were willing to choose the toy.

Panel B of Table 5 compares mean attendance and achievement outcomes of the choice treatments with the parent money treatment. In the attendance regressions, the coefficient on the *ex ante* choice treatment is positive, but it is significant in only one of the two specifications. However, the effect is negative and insignificant for achievement. The estimated effects for the *ex post* treatment are small and insignificant across both outcomes. On balance, there is weak evidence that the opportunity to commit did improve attendance in the after-school classes, but there is no evidence that it affected achievement.

3.3 Heterogeneity by Pretest Score

In this subsection I present the results of regressions interacting pretest score with the toy, *ex ante* and *ex post* treatment groups, using the money treatments as the omitted category.[21] Table 6 estimates these interaction effects on attendance in the after-school classes. Columns 1 and 2 use the categorical pretest score in the interactions, while Columns 3 and 4 use the continuous measure of the child's pretest score relative to his grade. In all specifications, the interaction of the test score and the toy treatment is negative and significant, implying that children with higher initial test scores have higher attendance in the money treatments, while those with lower test scores have higher attendance in the toys treatment. The magnitudes of the coefficients in Columns 1 and 2 indicate that a student who scores one point higher on the pretest is almost 14 percentage points more likely to attend in the money treatments relative to the toys treatment. Columns 3 and 4 show that a student who scores 1 standard deviation higher on the pretest is 10-11 percentage points more likely to attend in the money treatments relative to the toys treatment.

The interactions of the choice treatments and initial test score are also negative, but the magnitudes are smaller and none of the coefficients is significant. Thus, there is no evidence that the effect of the choice treatments relative to the money treatments differs by pretest score.

Table 7 repeats this analysis using achievement of the literacy goal as the outcome of interest. A similar pattern emerges: the interaction term of the toys treatment and the child's pretest score

[21] Appendix C presents an analysis of the interactions disaggregating the money treatments. Relative to the parent money treatment, there is no statistically significant interaction of pretest score and the child money treatment.

is again negative and significant, implying that children with higher pretest scores achieve the goal more often in the money treatments, and vice versa. In Columns 1 and 2, the interaction terms indicate that a student who scores one point higher on the pretest is 8-9% more likely to achieve the goal in the money treatments relative to the toys treatment. The interactions using relative pretest score show that students who score 1 standard deviation higher on the pretest are about 8 percentage points more likely to achieve the goal in the money treatments relative to the toys treatment. The interactions of the choice treatments and pretest scores are again small and insignificant.

These results imply that the toys treatment reduced inequality in outcomes relative to the money treatments: initially low-performing children improved more in the toys treatment, while initially high-performing children improved more in the money treatments. This suggests that incentives to children are more effective than incentives to parents for lower-performing children, while incentives to parents are more effective for higher-performing children. The next two sections explore the mechanisms behind these effects.

4 Theoretical Framework

One potential mechanism for the interaction of the toy treatment with pretest score is that test scores may serve as a proxy for the parent's ability to monitor her child's time and provide direct inputs into the education process. A parent with higher ability may have higher performing children and may be able to respond more strongly to parent incentives. This section formalizes this idea with a simple model of household education production.

In this model, the key friction within the household is moral hazard in the education production process. As noted in the introduction, previous work has utilized moral hazard in parent-child interactions to explain household behaviors (e..g, Weinberg, 2001; Gatti, 2005). In my model, education is produced by both parent and child effort, and the production process involves is two-sided moral hazard similar to the sharecropping model of Eswaran and Kotwal (1985). The parent and child cannot contract on inputs, but the parent can decide how much to give the child to induce effort. Based on the amount given to the child and the remaining value of education that accrues to the parent, the agents decide on how much effort to exert.

There are two periods in the model. The education output produced in each period represents achievement in an individual educational task, such as grade promotion or exam performance. The parent and child each provide inputs. Parent and child inputs reflect each agent's direct influence on education production. Parent input represents the parent's own effort in assisting the child with studies or in directly managing the child's activities, while child input represents the child's direct effort.[22] The productivity of each input in the production function varies across households. Productivity represents the parent's and child's ability to influence education production by providing inputs.

The model's two periods represent production before and during an external incentive program. In the first period, the parent and child decide how much effort to exert in the absence of external incentives. The parent decides how much to transfer to the child based on her own value of achievement in that period. In the second period, an experimenter augments the parent's value of achievement with additional incentives to learn.

4.1 Initial Period

The parent and child contribute inputs p_t and c_t in periods $t = \{1,2\}$. The costs of input to the parent and child are given by the quadratic functions $\frac{bp_t^2}{2}$ and $\frac{ac_t^2}{2}$ in each period.

The parent places a normalized value of 1 on success in each period. That is, she receives 1 if the task is completed successfully, and 0 if it is not.[23] The probability of success in each period is given by

$$f(p_t, c_t) = \beta p_t + \delta c_t$$

[22] This conceptualization of parent and child inputs in education production follows the seminal work of Leibowitz (1974).

[23] I assume that all parents place the same value on success in the task, and that parents only value transfers to their children for the motivational effect of these transfers. The implications of relaxing these two assumptions are discussed in Appendix E.

The parameters $\beta \geq 0$ and $\delta \geq 0$ capture the productivity of the parent's and child's inputs, respectively, in education production.[24]

Before first-period input choices are made, the parent announces a fraction γ_1 of the value of success that she will give to the child if the task is completed successfully. I assume that the parent must commit to this division of value until the first-period outcome has been realized. I further assume that the parent cannot make negative transfers to the child. In equilibrium, this limited liability constraint will bind.[25] If the child is unsuccessful, he will receive no transfer, and if he is successful, he will receive a fraction γ_1 of the value of success.

Based on the fraction γ_1, the parent and child maximize their respective shares of the value of success net of costs over their first-period inputs. The first-order conditions of these maximization problems form two incentive-compatibility constraints that the parent faces in deciding on γ_1.

I assume that both the parent and child are risk neutral and therefore make their decisions based on the expected value of success $f(p_1, c_1)$. The parent's and child's incentive-compatibility constraints are formed by the maximization of their share of the value of success net of costs over p_1 or c_1, taking the parent's choice of γ_1 as given:

$$p_1(\gamma_1) = \arg\max_{p_1} (1-\gamma_1)\{\beta p_1 + \delta c_1\} - \frac{bp_1^2}{2} \tag{1}$$

$$c_1(\gamma_1) = \arg\max_{c_1} \gamma_1\{\beta p_1 + \delta c_1\} - \frac{ac_1^2}{2} \tag{2}$$

Subject to the incentive-compatibility constraints (1) and (2), the parent maximizes her expected share of the value of success over the fraction γ_1 she gives to the child:

$$\max_{\gamma_1} (1-\gamma_1)\{\beta p_1 + \delta c_1\} - \frac{bp_1^2}{2} \tag{3}$$

This yields the sharing rule $\gamma_1^*(\beta, \delta, a, b)$ and the first-period probability of success, denoted by

[24] I assume that the cost parameters b and a are sufficiently large that the probability of success is always less than 1.

[25] In the absence of a limited liability constraint, the parent would be able to sufficiently punish the child so that the child received full incentives, thereby imposing the first-best level of effort on the part of the child.

$f_1 \equiv \beta p_1(\gamma_1^*) + \delta c_1(\gamma_1^*)$.

Proposition 1 *The first-period probability of success (i.e., $f_1 \equiv \beta p_1(\gamma_1^*) + \delta c_1(\gamma_1^*)$) is increasing in both parental and child productivity.*

Proof. *See Appendix D.* ■

4.2 Experiment

Now suppose that after the first-period input decisions are made, an experimenter observes the expected output, f_1, by administering a pretest. The experimenter then offers an external incentive of value π to either the parent or the child conditional on success in the second period. As noted above, the parent places a value of 1 on success in the second period in addition to this incentive.

If this external incentive accrues to the parent, the parent decides on an updated share γ_2 based on the external incentive plus the parent's own value of success $(\pi + 1)$.

This yields the new sharing rule $\gamma_2^*(\beta, \delta, a, b)$, and second-period probability of success of

$$f_{2,parent} \equiv \beta p_2(\gamma_2^*) + \delta c_2(\gamma_2^*)$$

Now suppose that the external incentive of π is given to the child and cannot be appropriated by the parent. This restriction imposes a share of at least $\frac{\pi}{\pi+1}$ to be given to the child. Denote this share by γ_{child}.

This restriction binds if the parent would have given the child less than π if she had received the incentive herself. That is, the restriction binds if

$$\gamma_2^* < \gamma_{child} \tag{4}$$

I assume throughout that π is large enough such that equation (4) holds. The probability of success in this case is given by

$$f_{2,child} \equiv \beta p_2(\gamma_{child}) + \delta c_2(\gamma_{child})$$

The difference in the second-period probability of success between incentives to the parent and incentives to the child is given by

$$f_{2,parent} - f_{2,child} = \beta p_2(\gamma_2^*) + \delta c_2(\gamma_2^*) - [\beta p_2(\gamma_{child}) + \delta c_2(\gamma_{child})] \tag{5}$$

The remainder of this subsection uses equation (5) to make predictions for the relative effectiveness of parent incentives versus child incentives when parent and child productivity (β and δ) vary across households.

Proposition 2 *At low values of parental productivity, incentives provided to the child will result in a higher probability of success than incentives provided to the parent. At high values of parental productivity, incentives provided to the parent will result in a higher probability of success than incentives provided to the child.*

Proof. *See Appendix D.* ■

The intuition behind this result is as follows. When the parent's input is relatively unproductive, the probability of success is higher when the child is rewarded for effort. However, when the parent is given the incentives she will distort the rewards towards herself instead. When the parent's input is relatively productive, the probability of success is higher when the parent is rewarded for effort. This occurs when the parent is given the incentives rather than the child.

Proposition 3 *At low values of child productivity, incentives provided to the parent will result in a higher probability of success than incentives provided to the child. At high values of child productivity, incentives provided to the parent will result in a lower probability of success than incentives provided to the child.*

Proof. *See Appendix D.* ■

The intuition behind this result is similar to that for Proposition 2. When the child's input is relatively unproductive, the parent will reward the more productive input (i.e., her own) when she is given the incentives. When the child's input is relatively productive, the parent is still inclined to distort the rewards towards herself when she is given the incentives. Therefore, it becomes more effective to give the incentives directly to the child.

This discussion shows how parental and child productivity are directly related to both the child's initial test score and the relative effect of child incentives. It also implies that test scores will have a negative interaction with child incentives to the extent that they reflect parental productivity, rather than child productivity. Thus, the variation in pretest scores attributable to parental productivity will yield a stronger interaction with child incentives than pretest scores overall.

5 Interactions Between Parental Productivity and Treatments

In this section I test the key implication of the theory. I first create an index of parental productivity from the predicted values of a regression of initial test scores on baseline survey responses that reflect the parent's ability teach her child and manage her child's time. I then estimate the effects of interactions between the resulting index and the toys treatment.

I include eight variables from the baseline survey as proxies for parental productivity.[26] These variables fall into four broad categories. First, I include two variables that reflect household composition. More children below age 15 in the household should take the parent's time away from the program child and therefore are expected to negatively affect the parent's ability to contribute to the child's education. On the other hand, the number of household members at or above age 15 are expected to positively affect productivity, since these members represent resources the child can use for help with his studies. Second, I include three variables that reflect the work and education status of the child's parents. Mother's employment status could affect her ability to contribute to her child's education because employed mothers will have less time to devote to their children.[27] Mother's and father's education are also included because more educated parents should be more

[26] See Haveman and Wolfe (1995) for a discussion of the empirical evidence on the effects of parental characteristics and behaviors on child attainment.

[27] Father's employment status is not included because only 4% of fathers were reported to be out of work. It is also not clear whether an out-of-work father represents additional parental resources, or whether the father is not working because he is sick or injured.

able to help the child with studies. Third, I include durables ownership, a measure of household wealth. Household wealth is expected to be positively related to parental productivity because parents in more wealthy households spend less time meeting basic needs and can therefore devote more time to their children. In addition, more wealthy households can contribute resources such as school supplies to facilitate their children's education. Finally, I include two measures of productive behavior: an indicator for whether anyone in the household has helped the child with his studies over the past day, and the amount of money that the parent spent on tutoring for the child over the past month.

To capture the extent to which these variables influence child test scores, I first regress the child's relative pretest score on these variables. The results of the first-stage regression of relative pretest score on the parental productivity variables are reported in Table 9.[28] In the specification of Column 1, all of the estimates except for father's education and durables ownership have the expected sign, and of the six with the expected sign three are significant at the 10% level. The results are similar when classroom dummies are added in Column 2.[29]

The predicted values of the the regressions in Columns 1 and 2 of Table 9 form a parental productivity index, as they represent variation in test scores attributable to parental productivity. Table 10 presents the effects of the index interacted with the toys treatment.[30,31] Columns 1 and 2 use attendance in the after-school classes as the outcome of interest. The estimated coefficients on the interaction terms are negative and significant at the 1% level in both specifications. The magnitudes of these coefficients are more than four times the estimates using the actual pretest scores, suggesting that the relative effectiveness of the toy treatment is more strongly related to the portion of pretest scores driven by parental characteristics than to pretest scores overall. The difference in

[28]Because households with children whose pretest scores were in the highest category were not surveyed, these scores are not included in this regression.

[29]It is somewhat surprising that durables ownership is significantly related to *lower* initial achievement. However, as discussed below, the inclusion of the durables measure does not drive the results presented in this section.

[30]The index used in Columns 2 and 4 of Table 10 includes only the effects of the productivity variables and not the classroom dummies.

[31]Because the parental productivity index is generated from a first-stage regression, standard errors for these regressions are bootstrapped based on 500 replications. Classes are drawn to form the bootstrap samples to account for clustering.

magnitudes is consistent with the theory, since the theory predicts that relative effectiveness of the toy treatment will be directly related to the *share* of test scores that reflect parental productivity.

Columns 3 and 4 repeat the analysis using achievement of the literacy goal as the outcome. As with the attendance results, the magnitudes of the estimated coefficients on the interaction terms increase considerably, although the estimate is significant at the 10% level in only one of the two specifications.

Following the theory, the parental productivity index is constructed to reflect the parent's ability to monitor her child and manage her child's time, and this mechanism drives the interaction effects between the productivity index and the toys treatment found in Table 10. Appendix E explores whether the parental productivity index could reflect other aspects of household preferences that in turn could drive these interaction effects. In particular, I evaluate in detail whether the productivity index could reflect altruism, preferences for toys, or the parent's value of education. I find no evidence in favor of any of these alternative explanations.

Appendix F provides additional analysis of the components of the productivity index and its interaction with the toys treatment. I show that the interaction results in Table 10 are robust to dropping any one variable from the index. While this indicates that no one variable is driving the interaction results, I find that the largest absolute decrease in the interaction terms occur when the variable indicating the number of children in the household is dropped. The influence of this variable suggests that the amount of time that a parent can spend with an individual child is an important constraint on the parent's ability to motivate that child.

The results of this section provide empirical support for the hypothesis that parental productivity is driving the interaction between pretest scores and the toys treatment. Consistent with the theory, the variation in initial test scores attributable to parental characteristics produces stronger interactions with the toys treatment than test scores overall. I find no evidence consistent with several competing explanations for these results. These results imply that incentives to parents are more effective when parents are more able to teach their children and motivate them to learn, while incentives to children are more effective when their parents are less able to teach and motivate them.

6 Conclusion

In this paper I present the results of a field experiment in Gurgaon, India, designed to test the effects of varying the recipient of incentives to learn between parents and children. The experiment offered incentives for families of first, second, and third graders in government primary schools to increase their children's reading ability. While I find no significant differences in attendance in after-school classes or achievement between the groups that were offered money or toys as an incentive, there is significant heterogeneity in the treatment effects. Children with higher initial test scores performed better in the money treatments, while children with lower test scores performed better in the toys treatments. This heterogeneity in treatment effects is consistent with a model of education production in which pretest scores reflect parental productivity, defined as parents' ability to monitor and motivate their children. When parents are more productive, it is more effective to reward the parents, and when parents are less productive, it is more effective to reward the children.

Thus far, research on targeting of cash transfer programs in developing countries has focused primarily on which adult within the household should receive the transfer rather than including the possibility of the child receiving the transfer directly. In the case of education, outcomes depend on child effort, and schemes that reward parents must depend on the parents both to exert their own effort *and* motivate their children. My results imply that when initial learning levels are low, providing incentives to learn directly to the child may be more effective than providing incentives to the parent. While this paper has focused on children early in the education process, future research should examine how the decision process changes as children grow older. If children's inputs become more important as they grow older, incentives targeted to children will be more effective for these older children. [32]

This paper also has implications for understanding how households respond to other external factors such as changes in returns to education. If parents' perceived returns to education change, households with more productive parents may respond more strongly because these households suffer from fewer agency problems between the parents and children. Since households with more productive parents tend to have higher-performing children to begin with, increases in returns may exacerbate inequality in education outcomes. This paper provides a starting point for future research examining these effects.

[32] Indeed, Attanasio and Kaufmann (2009), using evidence from Mexico, find that both parents' and youths' subjective expectations matter for high school attendance decisions, while only youths' expectations matter for college attendance decisions.

References

Joshua Angrist and Victor Lavy. The effects of high stakes high school achievement awards: Evidence from a group randomized trial. *American Economic Review*, 99:1384–1414, 2009.

Joshua Angrist, Daniel Lang, and Philip Oreopoulos. Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, 1: 136–163, 2009.

ASER Centre. *Annual Status of Education Report (Rural) 2012*. Pratham Resource Center: Mumbai, 2013.

Orazio Attanasio and Katja Kaufmann. Educational choices, subjective expectations and credit constraints. *NBER Working Paper 15087*, 2009.

Sarah Baird, Crag McIntosh, and Berk Ozler. Cash or condition? Evidence from a cash transfer experiment. *Quarterly Journal of Economics*, 126(4):1709–1753, 2011.

Abhijit Banerjee, Shawn Cole, Esther Duflo, and Leigh Linden. Remedying education: Evidence from two randomized experiments in India. *Quarterly Journal of Economics*, 122(3):1235–1264, 2007.

Gary Becker. A theory of social interactions. *Journal of Political Economy*, 82(6):1063–1093, 1974.

Theodore Bergstrom. A fresh look at the rotten kid theorem-and other household mysteries. *Journal of Political Economy*, 97(5):1138–1159, 1989.

Eric Bettinger. Paying to learn: The effect of financial incentives on elementary school test scores. *NBER Working Paper 1633*, 2010.

Moussa Blimpo. Team incentives for education in developing countries: A randomized field experiment in Benin. *Mimeo, Stanford University*, 2010.

Leonardo Bursztyn and Lucas Coffman. The schooling decision: family preferences, intergenerational conflict, and moral hazard in the Brazilian favelas. *Journal of Political Economy*, 3: 359–397, 2012.

Esther Duflo. Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in South Africa. *World Bank Economic Review*, 17(1):1–25, 2003.

Mukesh Eswaran and Ashok Kotwal. A theory of contractual structure in agriculture. *American Economic Review*, 75(3):352–367, 1985.

Roland Fryer. Financial incentives and student achievement: Evidence from randomized trials. *Quarterly Journal of Economics*, 126:1755–1798, 2011.

Roberta Gatti. Family altruism and incentives. *Scandinavian Journal of Economics*, 107(1):67–81, 2005.

Robert Haveman and Barbara Wolfe. The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature*, 33(4):1829–1878, 1995.

C. Kirabo Jackson. A little now for a lot later: a look at the Texas Advanced Placement Incentive Program. *Journal of Human Resources*, 45(3):591–639, 2010.

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

Arleen Leibowitz. Home investments in children. *The Journal of Political Economy*, 82(2):S111–S131, 1974.

Shelly Lundberg, Robert Pollak, and Terence Wales. Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit. *Journal of Human Resources*, 32(3):463–480, 1997.

Gauri Kartini Shastry and Leigh Linden. Grain inflation: Identifying agent discretion in response to a conditional school nutrition program. *Mimeo, Wellesley College*, 2011.

Bruce Weinberg. An incentive model of the effect of parental income on children. *Journal of Political Economy*, 109(2):266–280, 2001.

World Bank. *Edstats Database*. Accessed March 27, 2013.

Figure 1: Distributions of Raw Test Scores
Program and No-Program Groups

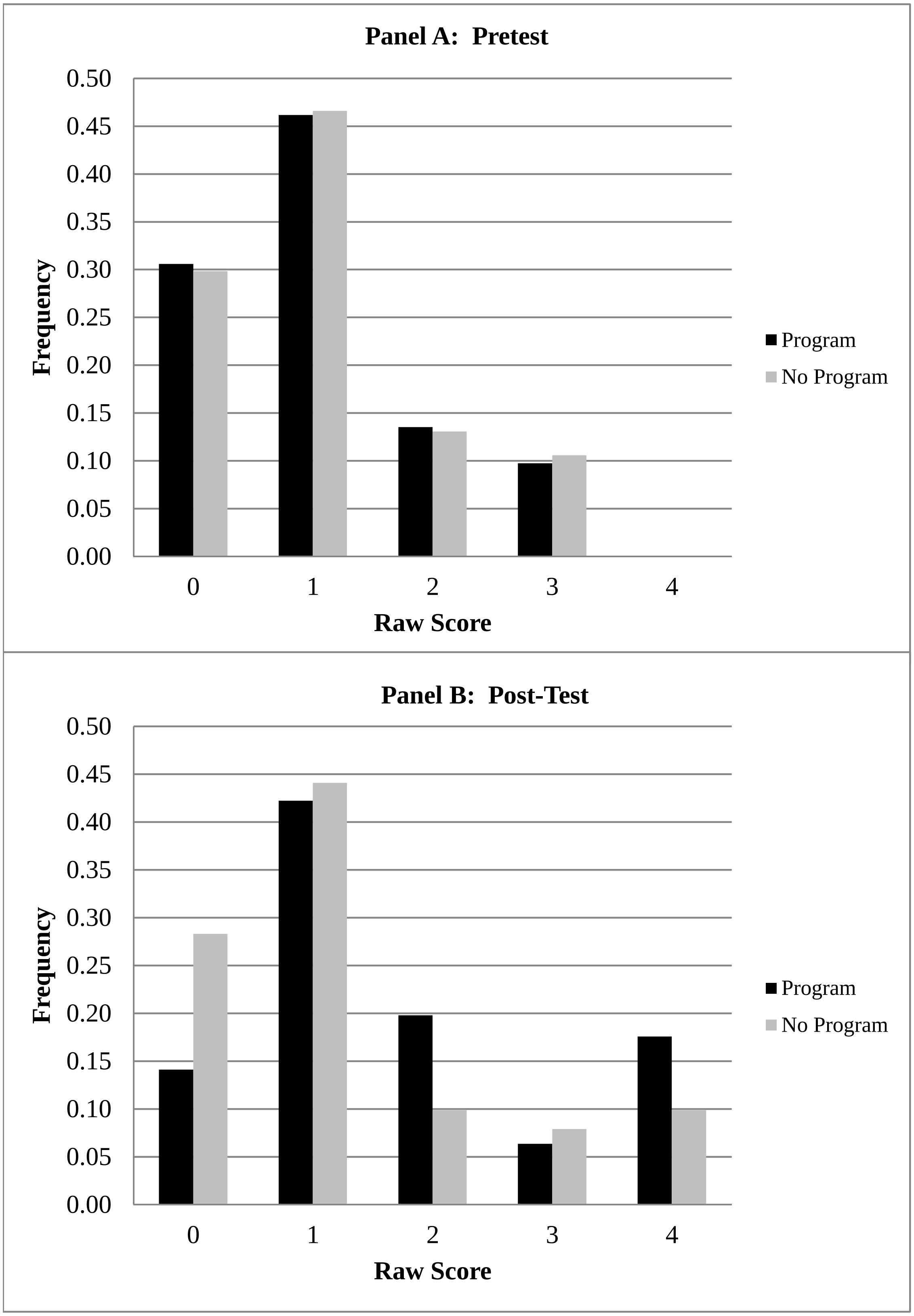

Table 1: Sample Composition

	Reached at Pretest	Randomized Sample	Reached for Program Announcement	Reached at Post-test
	(1)	(2)	(3)	(4)
Total	1466	1086	925	900
Raw Pretest Score				
0	349	331	283	276
1	528	502	427	414
2	151	146	125	124
3	107	107	90	86
4	331	0	0	0
Grade				
1	410	384	331	319
2	552	431	363	353
3	504	271	231	228
Treatment				
Parent Money		179	156	150
Child Money		181	156	152
Toy		362	305	296
Ex Ante Choice		183	153	151
Ex Post Choice		181	155	151

Table 2: Balance of Observables Across Treatment Groups

		P-Value of Differences	
	Mean	5 Treatment Categories	Money vs. Toy
Variable	(1)	(2)	(3)
Raw Pretest Score	1.024	0.570	0.449
Relative Pretest Score	-0.310	0.942	0.853
Female	0.571	0.639	0.318
Number of Children 0-14	2.915	0.796	0.787
Number of Adults 15+	2.442	0.215	0.118
Mother Employed	0.345	0.541	0.610
Mother Education	3.176	0.363	0.108
Father Education	6.421	0.512	0.101
Durables	1.236	0.369	0.467
Helped with Studies	0.358	0.126	0.973
Tutoring Fees Paid	26.585	0.589	0.955
Money Given	13.065	0.101	0.127
Gave Toys	0.037	0.237	0.699
Gave Other Item	0.088	0.896	0.582

Notes:

Column 2 presents the p-values of a test of equality of means across the parent money, child money, toy, *ex ante* choice and *ex post* choice treatment categories.

Column 3 presents the p-values of a test of equality of means between the aggregated money and toy categories.

"Raw Pretest Score" represents the child's integer pretest score.

"Relative Pretest Score" represents the difference between the continuous measure of child's score and the grade specific average, divided by the grade-specific standard deviation.

"Durables" is constructed as the first principal component of a set of dummies indicating household ownership of a bicycle, motorcycle, radio, dvd player, tv, refrigerator, gas stove, cooler, landline and mobile phone.

"Helped with Studies" is a dummy variable indicating whether anyone in the household helped the child with studies in the past day.

"Tutoring Fees Paid" represents the amount of money (in Rs.) paid for private tutoring over the past month.

"Money Given" represents the amount of money given to the child over the past week.

"Gave Toys" and "Gave Other Item" are dummy variables indicating whether the parent gave the child a toy or other item over the past week.

Standard errors are clustered at the classroom level.

Table 3: Determinants of Being Reached At Baseline

	Dependent Variable: Reached At Baseline (Dummy)		
	(1)	(2)	(3)
Relative Pretest Score	-0.001	-0.014	-0.026**
	(0.015)	(0.015)	(0.010)
Grade 2	-0.005		
	(0.042)		
Grade 3	0.001		
	(0.049)		
Female	0.029		
	(0.035)		
Classroom Dummies	NO	YES	YES
Surveyor Dummies	NO	NO	YES
Observations	1052	1052	1052
R-squared	0.002	0.098	0.409

Notes:

The sample includes all children in the randomized sample who took both the pretest and post-test.

In Column 3 a separate dummy is included for each surveyor in addition to a dummy indicating whether the household was surveyed by a team of two surveyors. An additional dummy indicates whether no address was collected (and no surveyor was assigned to the household).

Standard errors are clustered at the classroom level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4: Difference-in-Differences Estimates
Overall Effect of the Incentive Program

	Pretest Mean (No Program)	Pretest Mean (Program)	Difference	Post-test Mean (No Program)	Post-test Mean (Program)	Difference	Difference-in-Differences
	(1)	(2)	(3) (2) - (1)	(4)	(5)	(6) (5) - (4)	(7) (6) - (3)
Raw Test Score	1.059	1.022	-0.037 (0.104)	1.270	1.710	0.440*** (0.128)	0.477*** (0.082)
Relative Test Score	-0.301	-0.317	-0.016 (0.079)	-0.142	0.374	0.516*** (0.129)	0.532*** (0.104)
Could Read Nothing	0.289	0.307	0.017 (0.051)	0.283	0.141	-0.142*** (0.049)	-0.159*** (0.050)
Could Read Letters	0.474	0.460	-0.014 (0.046)	0.441	0.422	-0.019 (0.049)	-0.005 (0.064)
Could Read Words	0.125	0.138	0.013 (0.034)	0.099	0.198	0.099*** (0.031)	0.086* (0.045)
Could Read Paragraph	0.112	0.096	-0.016 (0.028)	0.079	0.064	0.016 (0.026)	0.001 (0.037)
Could Read Story	0	0	0	0.099	0.176	0.077** (0.032)	0.077** (0.032)

Notes:
"Raw Pretest Score" represents the child's integer pretest score.
"Relative Test Score" represents the difference between the continuous measure of child's score and the grade specific average pretest score, divided by the grade-specific standard deviation.
The rows indicating specific reading levels represent the fraction of children at those levels.
Standard errors are clustered at the classroom level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5: Overall Treatment Effects

	Dependent Variable: Attendance in After-School Classes		Dependent Variable: Achievement of Literacy Goal	
	(1)	(2)	(3)	(4)
Panel A: Money vs. Toy				
Toy	0.004	0.002	0.007	0.008
	(0.035)	(0.034)	(0.048)	(0.053)
Pretest Dummies	NO	YES	NO	YES
Classroom Dummies	NO	YES	NO	YES
Observations	502	502	598	598
R-squared	0.000	0.154	0.000	0.190
Mean of Dep. Var.	0.231	0.231	0.537	0.537
Panel B: 5 Main Treatment Groups				
Child Money	0.038	0.042	-0.04	-0.019
	(0.051)	(0.053)	(0.057)	(0.060)
Toy	0.023	0.022	-0.013	0.001
	(0.047)	(0.046)	(0.060)	(0.065)
Ex Ante Choice	0.093*	0.077	-0.05	-0.041
	(0.054)	(0.053)	(0.069)	(0.070)
Ex Post Choice	0	-0.004	0.016	0.024
	(0.044)	(0.039)	(0.057)	(0.062)
Pretest Dummies	NO	YES	NO	YES
Classroom Dummies	NO	YES	NO	YES
Observations	755	755	900	900
R-squared	0.005	0.133	0.002	0.172
Mean of Dep. Var.	0.240	0.240	0.537	0.537

Notes:

In Columns 1 and 2, the dependent variable is a dummy which equals 1 if the child attended the after-school classes on at least 1 day. In Columns 3 and 4, the dependent variable is a dummy which equals 1 if the child reached the literacy goal.
In Panel A the sample is restricted to the money and toy treatments.
The omitted treatment categories in Panel A are parent and child money.
The omitted treatment category in Panel B is parent money.
Standard errors are clustered at the classroom level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6: Interactions of Toy Treatment and Pretest Scores
Outcome: Attendance in After-School Classes

	Categorical Pretest Score		Relative Pretest Score	
	(1)	(2)	(3)	(4)
Toy	0.141***	0.135***	-0.023	-0.023
	(0.051)	(0.049)	(0.040)	(0.036)
Ex Ante Choice	0.125	0.105	0.066	0.046
	(0.082)	(0.075)	(0.048)	(0.046)
Ex Post Choice	0.049	0.040	-0.034	-0.031
	(0.060)	(0.059)	(0.037)	(0.033)
Toy * Pretest	-0.137***	-0.137***	-0.111**	-0.097**
	(0.043)	(0.043)	(0.041)	(0.041)
Ex Ante * Pretest	-0.050	-0.049	-0.035	-0.031
	(0.056)	(0.053)	(0.062)	(0.058)
Ex Post * Pretest	-0.069	-0.065	-0.049	-0.027
	(0.047)	(0.048)	(0.055)	(0.054)
Pretest Dummies	YES	YES	YES	YES
Classroom Dummies	NO	YES	NO	YES
Observations	755	755	755	755
R-squared	0.025	0.146	0.023	0.143

Notes:
The dependent variable is a dummy which equals 1 if the child attended the after-school classes on at least one day.
The omitted treatment categories are parent and child money.
Columns 3 and 4 control for relative pretest score in addition to categorical pretest score dummies.
Standard errors are clustered at the classroom level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 7: Interactions of Toy Treatment and Pretest Scores
Outcome: Achievement of Literacy Goal

	Categorical Pretest Score		Relative Pretest Score	
	(1)	(2)	(3)	(4)
Toy	0.088	0.099	-0.018	-0.014
	(0.070)	(0.072)	(0.055)	(0.056)
Ex Ante Choice	-0.008	0.004	-0.041	-0.044
	(0.067)	(0.068)	(0.052)	(0.050)
Ex Post Choice	0.040	0.053	0.029	0.033
	(0.067)	(0.070)	(0.049)	(0.053)
Toy * Pretest	-0.079*	-0.088**	-0.079*	-0.081*
	(0.041)	(0.042)	(0.045)	(0.045)
Ex Ante * Pretest	-0.027	-0.033	-0.022	-0.032
	(0.048)	(0.048)	(0.043)	(0.042)
Ex Post * Pretest	-0.006	-0.018	-0.015	-0.022
	(0.051)	(0.052)	(0.049)	(0.052)
Pretest Dummies	YES	YES	YES	YES
Classroom Dummies	NO	YES	NO	YES
Observations	900	900	900	900
R-squared	0.046	0.176	0.048	0.194

Notes:
The dependent variable is a dummy which equals 1 if the child reached the literacy goal.
The omitted treatment categories are parent and child money.
Columns 3 and 4 control for relative pretest score in addition to categorical pretest score dummies.
Standard errors are clustered at the classroom level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 8: Choice Between Toy and Money

	Sample		
	All	Achievers Only	
	Ex Ante Treatment	*Ex Ante* Treatment	*Ex Post* Treatment
	(1)	(2)	(3)
Chose Toy	0.327	0.316	0.506
Chose Money	0.673	0.684	0.494
Observations	153	76	79

Notes:
This table displays the choices of the parents in the *ex ante* and *ex post* choice treaments.
Columns 2 and 3 include only parents of children who reached the program goal.

Table 9: Relationship Between Relative Pretest Score and Parental Productivity Measures

	Dependent Variable: Relative Pretest Score	
	(1)	(2)
# Children under 15	-0.053**	-0.041*
	(0.023)	(0.022)
# Adults 15+	0.016	0.053
	(0.053)	(0.057)
Durables	-0.054***	-0.033**
	(0.015)	(0.014)
Mother employed	-0.076	-0.054
	(0.075)	(0.074)
Mother education	0.022*	0.016
	(0.011)	(0.011)
Father education	-0.004	-0.001
	(0.009)	(0.008)
Helped with studies	0.067	0.007
	(0.063)	(0.057)
Tutoring fees paid	0.179***	0.152***
	(0.044)	(0.048)
Classroom Dummies	NO	YES
Observations	925	925
R-squared	0.044	0.246

Notes:
See Table 2 notes for variable definitions.
Tutoring fees paid are measured in hundreds of rupees.
Dummy variables are included to account for missing values.
Standard errors are clustered at the classroom level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10: Interactions of Toy Treatment and Productivity Index

	Dependent Variable: Attendance in After-School Classes		Dependent Variable: Achievement of Literacy Goal	
	(1)	(2)	(3)	(4)
Toy	-0.146***	-0.170***	-0.068	-0.104
	(0.054)	(0.058)	(0.079)	(0.074)
Productivity Index	0.291***	0.289***	0.106	0.162
	(0.094)	(0.089)	(0.141)	(0.131)
Toy * Productivity	-0.494***	-0.572***	-0.240	-0.354*
	(0.141)	(0.146)	(0.199)	(0.183)
Pretest Dummies	YES	YES	YES	YES
Classroom Dummies	YES	YES	YES	YES
Predictions Control for Classroom	NO	YES	NO	YES
Observations	502	502	598	598
R-squared	0.165	0.165	0.192	0.193

Notes:
The sample used in this table includes the parent money, child money, and toy treatments.
The dependent variable is a dummy which equals 1 if the child achieved the literacy goal.
The omitted treatment categories are parent and child money.
In Columns 1 and 3, the productivity index represents the predicted values of the regression in Column 1 of Table 6.
In Columns 2 and 4, the productivity index represents the predicted values of the regression in Column 2 of Table 6, using the average effect of the classroom dummies.
Standard errors are constructed based on 500 bootstrap draws, sampling by classroom.
* significant at 10%; ** significant at 5%; *** significant at 1%

A Appendix: Self-reported Transfers from Parents to Children

Understanding the nature of transfers within the household can shed light on the extent to which the treatments provide binding constraints on parents. Further, the timing of transfers (either before or after the goal has been reached) indicates whether the parents provide incentives to their children based on effort or outcomes.

Appendix Table 1 presents estimates of the effects of the treatment groups on transfers from parents to children, controlling for the transfers reported at the baseline. The behaviors examined are the amount of money given to the child over the past week, a dummy for whether the parent gave a toy to the child over the past week and a dummy for whether the parent gave the child any other item over the past week. I also include an aggregate measure of the three types of transfers by averaging the z-scores of the measures. Panel A uses data from the first follow-up survey, taken just before the post-test. The estimates are small and insignificant, except for the difference in non-toy items, which is significant at the 5% level.

Panel B of Appendix Table 1 repeats the exercise using data from the second follow-up survey, taken just after the post-test. Using the aggregate measure of transfers in Column 1, the coefficient on the toys treatment is large and significant, indicating that parents in the money treatments did provide more transfers after the outcome of the test had been realized. This result is driven by the difference in transfers of non-toy items, which were 26% lower in the toy treatment than in the money treatments. These results suggest that parents who received money as a prize often used this money to purchase more practical items such as clothes and books for their children. While the money was often spent on items used by the children, it is likely that it was not spent in the manner most preferred by the children themselves. In addition, the size of the effect is about half of the fraction of children achieving the goal, indicating that not all of the prize money was spent on child-related goods. Thus, while parents did provide transfers to incentivize their children, they likely did not make the children the full residual claimants of the prize money.

While there were few differences in transfer behavior in the *ex ante* treatment relative to the money treatments, it is notable that the transfers in the *ex post* treatment follow a pattern similar to that of the toys treatment. One explanation for this pattern is that parents in the *ex post* treatment used the salient possibility of a toy as a motivator for their children. This explanation is speculative, however, because in this case one would expect the analysis of outcomes for the *ex post* treatment to mirror that of the toys treatment, and this analysis instead does not reject equivalence between the *ex post* treatment and the money treatment.

B Appendix: Additional Analysis of Overall Program Effects

B.1 Estimation of the Counterfactual Improvement using Cross-sectional Differences Across Grades

In order to verify that the no-program group represents a reasonable counterfactual for the program group, I compare the increase in test scores for the no-program group to the cross-sectional differences in test scores between grades 1, 2, and 3. On average, the no-program group increased by 0.21 points during the two months between the pretest and post-test. On average, a first grader is 1.04 points below a second grader at the pretest, and a second grader is 1.15 points below a third grader. Gurgaon schools are in session 11 months out of the year, and if a student improves by an equal amount each month, one would expect a first-grade student to improve 2/11*1.04 = 0.19 points on average, and a second-grade student to improve 2/11*1.15 = 0.21 points. These estimated improvements are very close to the 0.21-point increase observed among the no-program group.

B.2 Effect Size Comparison

This section compares estimates of the overall effectiveness of the incentives program with estimates from Kremer, Miguel and Thornton (2009) and Blimpo (2010). As shown in Table 4, the difference-in-differences estimate of the Gurgaon program's impact on normalized test scores equals 0.53 standard deviations. Kremer, Miguel and Thornton (2009) find an effect size of 0.19

standard deviations, while Blimpo (2010) finds an effect of 0.29 standard deviations for the individual incentives treatment.

While the three incentives programs had a number of structural differences, it is particularly important to note that the incentives in my study were based on performance on a basic reading test. In contrast, Kremer, Miguel and Thornton (2009) and Blimpo (2010) use broad-based tests covering multiple subjects. Thus, one plausible explanation for the larger effects that I find is that it is easier for a student to focus only on reading skills rather than the more broad-based skills incentivized in the other studies.

C Appendix: Interaction Analysis of Disaggregated Money Treatments

Appendix Table 2 repeats the interaction analysis of Tables 6 and 7, separating the child money treatment from the parent money treatment. In this table, the child money, toy, *ex ante* choice and *ex post* choice treatments are interacted with the child's initial test score. The omitted category is the parent money treatment.

I focus on the interaction terms of the child money and toys treatments. Relative to the parent money treatment, the interaction of the toy treatment and pretest score is negative and significant across all four specifications. The coefficients are of similar magnitude to the coefficients comparing the toys treatment with the aggregated money treatments in Tables 6 and 7. On the other hand, the interaction of the child money treatment and the child's pretest score are not significantly different from the parent money treatment in any of the specifications. The child money interaction is statistically different from the toy interaction term in one of the two attendance specifications, but in neither of the achievement specifications. Because child money is statistically indistinguishable from parent money, this supports the assertion that child and parent money are pooled in the household. It should be noted, however, that the coefficients are imprecisely estimated, and I cannot consistently rule out equivalence of the child money and toys interactions.

D Appendix: Proofs

D.1 Proposition 1: First-period Probability of Success is Increasing in β and δ

Maximizing equations (1) and (2) yields the incentive-compatibility constraints $p_1 = \frac{\beta(1-\gamma_1)}{b}$ and $c_1 = \frac{\delta\gamma_1}{a}$. Substituting these equations into the parent's optimization for γ_1 (equation (3)) yields

$$\max_{\gamma_1}(1-\gamma_1)\left\{\frac{\beta^2(1-\gamma_1)}{b}+\frac{\delta^2\gamma_1}{a}\right\}-\frac{\beta^2(1-\gamma_1)^2}{2b}$$

The first-order condition for γ_1 is given by

$$\frac{\delta^2(1-\gamma_1)}{a}-\frac{\delta^2\gamma_1}{a}-\frac{\beta^2(1-\gamma_1)}{b}=0 \tag{6}$$

Solving this equation for $\gamma_{1,}$, and noting that $0 \leq \gamma \leq 1$, yields

$$\gamma^* = \frac{b\delta^2 - a\beta^2}{2b\delta^2 - a\beta^2} \tag{7}$$

when $\beta^2 < \delta^2 \frac{b}{a}$, and 0 otherwise, where I omit the time subscript for ease of exposition.

To prove that the probability of success is increasing in β or δ, it is sufficient to show that $\frac{(1-\gamma^*)\delta^2}{a}$ is also increasing in β or δ.[33] Because $\frac{(1-\gamma^*)\delta^2}{a}$ is decreasing in γ^*, I can show that γ^* is decreasing in β to show that the probability of success is increasing in β. I do this by differentiating (7) with respect to β:

$$\frac{\partial\gamma^*}{\partial\beta} = -\frac{2a\beta}{2b\delta^2-a\beta^2}-\frac{2a\beta\left(b\delta^2-a\beta^2\right)}{\left(2b\delta^2-a\beta^2\right)^2}$$

This quantity will be negative when γ^* is at an interior solution, i.e., $\beta^2 < \delta^2\frac{b}{a}$.

[33]This follows because on the interior, the first-order condition (6) can be rewritten as $\frac{(1-\gamma^*)\delta^2}{a} = \beta p_1 + \delta c_1$.

To prove that $\frac{(1-\gamma^{*})\delta^{2}}{a}$ is increasing in δ, differentiate with respect to δ :

$$\frac{\partial}{\partial\delta}\frac{(1-\gamma^{*})\,\delta^{2}}{a}=-\frac{\partial\gamma^{*}}{\partial\delta}\frac{\delta^{2}}{a}+(1-\gamma^{*})\frac{2\delta}{a} \tag{8}$$

where

$$\frac{\partial\gamma^{*}}{\partial\delta}=\frac{2ab\delta\beta^{2}}{\left(2b\delta^{2}-a\beta^{2}\right)^{2}} \tag{9}$$

After substituting (7) and (9) into (8) and simplifying, I have

$$\frac{\partial}{\partial\delta}\frac{(1-\gamma^{*})\,\delta^{2}}{a}=\frac{4b\delta^{3}(b\delta^{2}-a\beta^{2})}{a\left(2b\delta^{2}-a\beta^{2}\right)^{2}}$$

This quantity will be positive when $\gamma^{*}$is on the interior.■

D.2 Proposition 2: Incentives to Parents Are Relatively More Effective when the Parent's Input Is More Productive

In this proof I show that when parental productivity is below a certain point, incentives to parents are more effective, and when parental productivity is above that point, incentives to children are more effective.

I first solve for the parent's choice of γ_2 if she is given the incentive directly. In period 2, the incentive-compatibility constraints and are given by $p_2=(\pi+1)\frac{\beta(1-\gamma_2)}{b}$ and $c_2=(\pi+1)\frac{\delta\gamma_2}{a}$. Performing the parent's maximization over γ subject to these incentive-compatibility constraints yields

$$\gamma_2^{*}=\frac{b\delta^{2}-a\beta^{2}}{2b\delta^{2}-a\beta^{2}} \tag{10}$$

when $\beta^{2}<\delta^{2}\frac{b}{a}$, and 0 otherwise.

Now I can use the solution for γ_2^{*} to examine the how equation (5) varies with δ. Substituting the incentive-compatibility constraints into (5), and noting that $\gamma_{child}=\frac{\pi}{\pi+1}$, equation (5) can be written as

$$\beta^2 \frac{(\pi+1)(1-\gamma_2^*)}{b} + \delta^2 \frac{(\pi+1)\gamma_2^*}{a} - \beta^2 \frac{1}{b} - \delta^2 \frac{\pi}{a} \tag{11}$$

Equation (11) will be negative when $\beta < \delta\sqrt{\frac{b}{a}}$ and positive when $\beta > \delta\sqrt{\frac{b}{a}}$. Equation (11) equals zero only when $\beta = \delta\sqrt{\frac{b}{a}}$, and at this point it has a positive partial derivative with respect to β, implying that it is negative at lower values of β and positive at higher values of β.■

D.3 Proposition 3: Incentives to Children Are Relatively More Effective when the Child's Input Is More Productive

The proof of this proposition follows the same logic as in the proof of Proposition 2. Equation (11) will be positive when $\delta < \beta\sqrt{\frac{a}{b}}$ and negative when $\delta > \beta\sqrt{\frac{a}{b}}$. Equation (11) equals zero only when $\delta = \beta\sqrt{\frac{a}{b}}$, and at this point it has a negative partial derivative with respect to δ, implying that it is positive at lower values of δ and negative at higher values of δ.■

E Appendix: Robustness Checks

This appendix evaluates several alternative explanations for the interaction between the productivity index (constructed as predicted values from the regressions in Table 9) and the toy treatment. The next three subsections consider heterogeneity in altruism, preferences for toys, and the parent's value of education. I find no evidence that the results from Section 5 are driven by any of these three effects.

E.1 Altruism

I first explore whether the productivity index could reflect different levels of altruism on the part of the parent. The theory developed in Section 4 can be extended to incorporate altruism, where the parent places a value of α (where $0 < \alpha \leq 1$) on the child's utility. Doing so increases the fraction γ given to the child. This occurs because the parent now values a contribution to the child both for its motivational effect as well as for its effect on the child's consumption. In case of perfect altruism, when $\alpha = 1$, the parent transfers all of the benefits to the child, that is, $\gamma = 1$.

When parent and child productivity are fixed, and altruism varies across households, altruism could be positively related to the relative effect of parent incentives. This occurs because altruism raises γ and thereby lowers the extent to which the toy incentive puts a constraint on the parent. Suppose that for low-altruism households the constraint (4) is binding and equation (5) is negative, so that toys are relatively more effective. For households in which altruism is high enough, (4) will no longer bind, and toys and money will be equally effective. Thus, if the productivity index reflects altruism, this altruism could result in a negative interaction of the index and the toy treatment.

Note that the results presented in Section 5 do not conform to this theory in that this theory predicts that toys will be more effective for low-altruism households and that toys and money will be equivalent for high-altruism households. The results from Table 10 imply that the toy treatment is more effective for households with low values of the productivity index and less effective for households with high values of the index.[34]

I also test whether the productivity index is related to altruistic behavior. The theory predicts that an altruistic parent will provide more transfers to her child in the absence of external incentives. Panel A of Appendix Table 3 regresses transfers from the parent to the child in the week before the baseline on the productivity index, as generated from Column 2 of Table 9.[35] The transfers examined are the same as those in Appendix Table 1. I find no significant relationship between the productivity index and any of these measures.

[34] The effect of the toy treatment at the mean value of the index is approximately 0 for both attendance and achievement.

[35] Repeating the regressions in Appendix Table 3 using the productivity index generated from Column 1 of Table 9 produces similar results.

E.2 Preferences

A second alternative explanation for the interaction between the productivity index and the toy treatment is that the index could reflect heterogeneous preferences for child rewards across households.

Suppose that households maximize a single utility function over child rewards (i.e., toys) and other goods, but households differ by the extent to which they value toys. A household with a strong preference for toys will value these goods at their cash equivalent, but a household that dislikes toys will value these toys at less than their cash equivalent. If a higher value of the productivity index reflects a preference against toys, then the interaction of the productivity index and the toy treatment will be negative. Note, however, that this model predicts that toys will *never* be preferred to money; rather, money will be preferred for some households. Contrary to this prediction, in Section 5 I find that toys are more effective for households with low values of the productivity index and less effective for households with high values of the index.

I also examine whether the productivity index could reflect preferences by correlating the productivity index with several proxies for preferences. Panel B of Appendix Table 3 presents regressions of these proxies on the productivity index, as generated from Column 2 of Table 9. The first three columns use the parent's choice of toy in the *ex ante* or *ex post* treatments to proxy for preferences. The coefficient on the productivity index is always insignificant and has an inconsistent sign. The last column proxies preferences with a dummy that equals 1 if the child indicated that he would buy a toy if he was given 100 rupees. Although there is a negative coefficient on the productivity index, this relationship is not at all significant. In sum, there is little evidence that preferences for the toy rewards are driving the negative interaction between the productivity index and the toy treatment.

E.3 Value of Education

The theory presented in Section 4 assumes that all parents place the same value on success in the task. This subsection explores whether heterogeneity in this value could be driving the interactions between the productivity index and the toy treatment.

Suppose that parents place different values on success in the task. As shown in Appendix D.2, the value of success does not affect the share of the rewards γ that goes to the child. It does, however, alter the *total* amount received by the parent and child. This, in turn, will influence whether the toy treatment puts a binding constraint on the distribution of rewards after the goal has been reached. If equation (5) is negative for households with low values of education (implying that child incentives are relatively more effective), the difference between the two types of incentives will shrink as the value of education increases. For a high enough value of success, the constraint (4) will not bind and the two types of incentives will produce equivalent results. Therefore, if the value of education is reflected in the productivity index, then heterogeneity in the value of education could result in the negative interaction between the productivity index and the relative effectiveness of the toy treatment. As with the case of heterogeneity in altruism, however, households with a low value of education would have better outcomes in the toy treatment compared with the money treatment, while the treatments would be equivalent for households with a high value of education. As stated above, the analysis in Section 5 shows that the money treatments are relatively *more* effective for high values of the productivity index.

While I do not have a direct measure of the parents' values of education, a model with heterogeneity in the value of education predicts that parents with higher values will transfer more to their children in the absence of external incentives. Thus, the lack of correlation between transfers and the productivity index, as shown in Panel A of Appendix Table 3, is not consistent with the hypothesis that the productivity index reflects the parent's value of education.

F Components of the Parental Productivity Index

In this section I explore which of the components of the parental productivity index are most influential in driving the interactions between the index and toys treatment in in Table 10.

I examine these components through two sets of regressions. First, I drop one variable at a time from the first-stage regressions in Column 2 of Table 9 and re-run the regressions in Columns 2 and 4 of Table 10. The results are presented in Appendix Table 4.[36] Overall, dropping one variable from the index does not change the estimated coefficients substantially. For both outcome measures, the largest decreases in the interaction coefficients occur after dropping the variable indicating the number of children in the household.

Second, I re-run the regressions in Columns 2 and 4 of Table 10, interacting the toy treatment with each component of the index. The results are presented in Appendix Table 5. In the attendance regression, the only variable that is significant when interacted with they toy treatment is the tutoring variable (significant at the 5% level). In the achievement regression, the only significant variable is the number of children in the household (significant at the 5% level).

This analysis suggests that while no single variable drives the interactions between the productivity index and the toy treatment, the number of children in the household and the baseline amount spent on tutoring are most influential in these interactions. The tutoring variable may be especially influential in the attendance regression because it reflects the parent's ability to send her children to classes outside of school time. It should be noted, however, that parents in the money treatments did not use the classes provided as part of the program as a substitute for prior tutoring classes. There is no significant difference in average spending on tutoring between the money and toy treatments in the month before the post-test (results not shown).

[36]Repeating the regressions in Appendix Table 4 using the productivity index generated from Column 1 of Table 9 produces similar results.

Appendix Table 1: Treatment Effects on Self-Reported Transfers

	Average Transfers	Money given	Gave Toys	Gave Other Item
	(1)	(2)	(3)	(4)
Panel A. First Follow-up vs. Baseline				
Toy	-0.042	0.054	0.018	-0.061**
	(0.061)	(0.634)	(0.024)	(0.029)
Baseline	0.068	0.028**	0.081	-0.023
	(0.048)	(0.012)	(0.072)	(0.071)
Pretest Dummies	YES	YES	YES	YES
Classroom Dummies	YES	YES	YES	YES
Observations	607	615	608	607
R-squared	0.136	0.219	0.12	0.088
Panel B. Posttest Follow-up vs. Baseline				
Toy	-0.398***	-0.297	-0.052	-0.260***
	(0.074)	(0.972)	(0.033)	(0.045)
Ex Ante Choice	-0.141	0.257	-0.019	-0.101*
	(0.112)	(1.502)	(0.044)	(0.059)
Ex Post Choice	-0.385***	-0.573	-0.097***	-0.177***
	(0.093)	(0.968)	(0.029)	(0.053)
Baseline	0.077**	0.026	-0.086**	0.140**
	(0.034)	(0.017)	(0.038)	(0.061)
Pretest Dummies	YES	YES	YES	YES
Classroom Dummies	YES	YES	YES	YES
Observations	905	920	908	907
R-squared	0.198	0.221	0.145	0.142

Notes:
In Panel A the sample is restricted to the money and toy treatments.
The omitted treatment categories are parent and child money.
"Average Transfers" is an average of the z-scores of the three individual transfer categories. The baseline means and standard deviations were used in computing the z-scores.
"Baseline" represents the baseline value of the transfer variable.
See Table 2 notes for variable definitions.
Standard errors are clustered at the classroom level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix Table 2: Interactions of Toy Treatment and Pretest Scores
Disaggregating Money Treatments

	Attendance		Achievement	
	Categorical Score	Relative Score	Categorical Score	Relative Score
	(1)	(2)	(3)	(4)
Child Money	0.037	0.028	-0.031	-0.021
	(0.070)	(0.057)	(0.079)	(0.068)
Toy	0.156**	-0.009	0.083	-0.024
	(0.065)	(0.050)	(0.089)	(0.071)
Ex Ante Choice	0.126	0.061	-0.012	-0.055
	(0.077)	(0.056)	(0.097)	(0.070)
Ex Post Choice	0.061	-0.015	0.037	0.023
	(0.066)	(0.046)	(0.091)	(0.070)
Child Money * Pretest	0.011	-0.076	0.014	-0.044
	-0.069	-0.058	-0.058	-0.071
Toy * Pretest	-0.133**	-0.136**	-0.080*	-0.103*
	(0.059)	(0.062)	(0.046)	(0.057)
Ex Ante * Pretest	-0.046	-0.070	-0.026	-0.054
	(0.060)	(0.073)	(0.064)	(0.058)
Ex Post * Pretest	-0.062	-0.065	-0.011	-0.043
	(0.059)	(0.067)	(0.064)	(0.074)
Pretest Dummies	YES	YES	YES	YES
Classroom Dummies	YES	YES	YES	YES
P-value: Toy vs. Child Money Inter.	0.007	0.12	0.101	0.308
Observations	755	755	900	900
R-squared	0.147	0.146	0.176	0.194

Notes:
The dependent variable is a dummy which equals 1 if the child attended the after-school classes on at least one day. In Columns 3 and 4, the dependent variable is a dummy which equals 1 if the child reached the literacy goal.
The omitted treatment category is parent money.
Columns 2 and 4 control for relative pretest score in addition to categorical pretest score dummies.
Standard errors are clustered at the classroom level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix Table 3: Correlates of the Productivity Index

Panel A: Productivity Index and Baseline Transfers

	Dependent Variable			
	Average Transfers	Money Given	Gave Toys	Gave Other Item
	(1)	(2)	(3)	(4)
Productivity Index	-0.027	3.692	0.021	-0.110
	(0.146)	(3.530)	(0.040)	(0.078)
Observations	922	925	923	922
R-squared	0.000	0.001	0.000	0.004

Panel B: Productivity Index and Choices

	Dependent Variable			
	Parent Chose Toy			Child Chose Toy
	(1)	(2)	(3)	(4)
Productivity Index	0.099	0.121	-0.161	-0.038
	(0.137)	(0.121)	(0.122)	(0.030)
Sample	Both Choice Treatments	*Ex Ante* Treatment	*Ex Post* Treatment	All Treatments
Observations	231	153	78	907
R-squared	0.001	0.002	0.002	0.001

Notes:

The productivity index is constructed as the predicted values of the regression in Column 2 of Table 9.

"Average Transfers" is an average of the z-scores of the three individual transfer categories. The baseline means and standard deviations were used in computing the z-scores.

See Table 2 notes for defnitions of the dependent variables in Panel A.

The dependent variable in Panel B, Columns 1-3 is a dummy which equals 1 if the parent chose the toy.

The dependent variable in Panel B, Column 4 is a dummy which equals 1 if the child indicated that he would buy a toy if he had 100 rupees.

Standard errors are constructed based on 500 bootstrap draws, sampling by classroom.

* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix Table 4: Importance of Components of the Productivity Index
Toy-Productivity Interactions, Dropping One Variable at a Time from Productivity Index

	Variable Dropped							
	Number of Children 0-14	Number of Adults 15+	Mother Employed	Mother Education	Father Education	Durables	Helped with Studies	Tutoring Fees Paid
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Outcome: Attendance								
Toy * Productivity	-0.444***	-0.535***	-0.591***	-0.638***	-0.662***	-0.666***	-0.566***	-0.515**
	(0.140)	(0.149)	(0.155)	(0.136)	(0.151)	(0.171)	(0.138)	(0.229)
Observations	502	502	502	502	502	502	502	502
R-squared	0.161	0.164	0.164	0.165	0.167	0.165	0.164	0.160
Panel B: Outcome: Achievement								
Toy * Productivity	-0.214	-0.344*	-0.447**	-0.447**	-0.234	-0.383**	-0.352*	-0.474*
	(0.199)	(0.191)	(0.199)	(0.215)	(0.229)	(0.175)	(0.193)	(0.281)
Observations	598	598	598	598	598	598	598	598
R-squared	0.192	0.192	0.194	0.194	0.191	0.193	0.193	0.194

Notes:
Within each panel, each column replicates the regressions from Columns 2 and 4 of Table 10, where one variable at a time is dropped from the productivity index.
* significant at 10%; ** significant at 5%; *** significant at 1%

Appendix Table 5: Importance of Components of the Productivity Index
Individual Interactions with Toy Treatment

	Interacted Variable							
	Number of Children 0-14	Number of Adults 15+	Mother Employed	Mother Education	Father Education	Durables	Helped with Studies	Tutoring Fees Paid
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Outcome: Attendance								
Toy * Variable	0.056	-0.050	0.010	-0.009	-0.006	-0.007	-0.080	-0.160**
	(0.037)	(0.043)	(0.079)	(0.011)	(0.011)	(0.015)	(0.082)	(0.072)
Observations	500	487	501	488	478	495	502	496
R-squared	0.170	0.165	0.163	0.152	0.155	0.158	0.156	0.163
Panel B: Outcome: Achievement								
Toy * Variable	0.063**	-0.061	-0.049	0.006	0.007	0.005	-0.030	-0.022
	(0.029)	(0.039)	(0.092)	(0.015)	(0.012)	(0.024)	(0.070)	(0.095)
Observations	594	577	597	590	579	590	598	591
R-squared	0.198	0.200	0.190	0.194	0.202	0.193	0.191	0.190

Notes:
Within each panel, each column replicates the regressions from Columns 2 and 4 of Table 10, where the productivity index is replaced by a single variable.
Standard errors are clustered at the classroom level.
* significant at 10%; ** significant at 5%; *** significant at 1%