Child Control in Education Decisions

An Evaluation of Targeted Incentives to Learn in India

James Berry

ABSTRACT

I report the results of a field experiment in Gurgaon, India that offered cash and noncash incentives to learn either to children or to their parents. While I find no evidence that the identity of the recipient or form of the reward mattered in the aggregate, noncash incentives targeted to children were more effective for initially low-performing children while cash incentives were more effective for high-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. 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). In contrast to a relatively large literature on the overall effectiveness of incentive programs in education, rigorous evidence directly comparing the effectiveness of different types of incentives is relatively rare.[1]

1. See Rawlings and Rubio (2005) for a review of the literature on conditional cash transfers that incentivize enrollment and attendance. Kremer, Miguel, and Thornton (2009) and Blimpo (2010) provide evaluations of incentives-to-learn programs.

*James Berry is an assistant professor of economics at Cornell University.**He is especially grateful to 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, three anonymous referees, and numerous seminar participants for helpful comments and suggestions. He 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 his own. This project was supported by the Russell Sage Foundation Small Grants Program in Behavioral Economics and the George and Obie Schultz Fund. The data used in this article can be obtained beginning March 2016 through April 2019 from James Berry, 486 Uris Hall, Ithaca, NY 14853, jimberry@cornell.edu.*
[Submitted April 2013; accepted August 2014]
ISSN 0022-166X E-ISSN 1548-8004

In particular, little is known about how the recipient within the household and form of incentives influences the effectiveness of these programs. Conditional cash transfer programs that provide incentives for enrollment and attendance typically target mothers rather than fathers because it is assumed that mothers will invest more in household public goods (Duflo 2003; Lundberg, Pollak, and Wales 1997). Similarly, frictions within the household decision-making process could result in differences in outcomes depending on whether incentives are targeted to children or to parents. For example, if child effort is an important input into education production and if parents are unable to motivate their children to provide effort, then incentives provided to children may improve educational outcomes more than incentives provided to parents.

In addition to the direct recipient of the reward, the form of the reward could also influence outcomes. Cash incentives received by children may be appropriated by parents, or parents may adjust child consumption so that the recipient of the transfer does not matter (Becker 1974). In-kind transfers, on the other hand, may be less appropriable by parents, or parents may not be able to adjust child consumption along other dimensions in order to offset the transfers.

In this paper I test the effects of targeted cash and in-kind incentives to learn using 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 pre-test 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 incentives of either money given to the child's mother ("parent money"), money given to the child ("child money"), a toy of equivalent value given to the child ("child toy"), or a voucher redeemable for a toy given to the child. The experiment therefore varied both the *recipient* of the incentive (parent or child) and the *form* of the incentive (money or toy).

All children in the program, regardless of treatment, were given the opportunity to attend free afterschool 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.

I find no evidence that the type of incentive or the identity of the recipient affected outcomes in the aggregate, but there is evidence that these impacts varied by the child's initial test score. There were no significant differences in average attendance in the afterschool classes or achievement of the literacy goal between the parent money, child money, voucher, or child toy treatment groups. Confidence intervals are able to rule out differences above about half of the estimated effect of any incentive. This result suggests that the average response of households did not depend on the recipient or form of the incentive. However, this result masks important heterogeneity in treatment effects by pre-test score. Children with lower pre-test scores performed better when provided a toy or voucher as an incentive relative to parent or child money while the reverse was true for children with higher pre-test scores. These results provide evidence that the form of the reward can be important in determining the effectiveness of incentives.

2. Throughout the paper, I use masculine pronouns to refer to the child and feminine pronouns to refer to the parent.

To explore the mechanisms behind the heterogeneity in effects by pre-test score, I adapt a model of education production to the context of the experiment. In the model, 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 the interaction with overall test scores.

I then test these implications empirically. To create an index of parental productivity, I use the predicted values from a regression of initial test scores on a set of parental characteristics that proxy for parental productivity. I find a negative interaction between the productivity index and the combined toy and voucher treatments. In addition, consistent with the model, this interaction is more strongly negative than the interaction using overall test scores.

This paper offers several contributions to existing research. It is one of only two evaluations that I am aware of that directly compares the effectiveness of incentives given to parents with incentives given to children. Baird, McIntosh, and Ozler (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. The authors find that the amount given to each recipient had no significant effect on enrollment or test scores.[3]

My study complements the findings of Baird, McIntosh, and Ozler (2011) in two key ways. First, in Baird, McIntosh, and Ozler (2011), the transfers always were in the form of cash. Because money may be pooled in the household, I varied both the recipient and the type of reward in my experiment. Second, I explore the distributional implications of the different incentives and find that lower-performing children had better outcomes when provided toy incentives.[4]

This paper also adds to the literature on the effects of in-kind transfers on the distribution of resources within the household. Under a standard unitary household model, in-kind transfers will yield different household consumption than cash transfers only if the in-kind transfers are not inframarginal — that is, if they are large enough to constrain household choices. However, it is possible that even inframarginal transfers can "stick" to their intended recipients without compensatory shifts in other consumption. For example, food transfers may be seen as women's income if women are the main purchasers of food (Haddad, Hoddinott, and Alderman 1997). Jacoby (2002) finds evidence for what he coins the "intrahousehold flypaper effect," where a school feeding program in the Philippines was not mitigated by shifts in consumption outside of school.

One other study that I know of compares cash with noncash incentives to learn. Levitt et al. (2012) compare cash with trophy incentives on test-taking effort among

3. Baird, de Hoop, and Ozler (2013) examine the effects of the same intervention on psychological distress of the children and find suggestive evidence that larger transfers to the household increased psychological distress while larger transfers to the girls decreased psychological distress.

4. Several other differences are worth noting. In the Baird, McIntosh, and Ozler (2011) study, the key variation is in the amount of transfers conditional on both the mother and child receiving a positive transfer. My study offered the entire reward to either the parent or child in order to create a stronger contrast between the types of incentives. In addition, Baird, McIntosh and Ozler (2011) study a conditional cash transfer program that rewarded enrollment and attendance in school, while my study examines an incentives-to-learn program.

schoolchildren in Chicago. The authors find evidence that nonfinancial rewards were as effective as more costly financial rewards. This study differs from mine in that the rewards were given based on test-taking effort rather than learning over time, and the trophies were designed primarily to remind the children of achievement rather than providing consumption-based utility.

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.[5] A recent study by Bursztyn and Coffman (2012) finds that parents in a Brazilian conditional cash transfer program preferred 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.

Finally, this paper also adds to the growing literature evaluating incentives-to-learn programs in developing countries. 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.[6]

The remainder of this paper is organized as follows. Section II describes the design of the intervention and the outcome measurement. Section III presents my findings on the effects of the treatments on class attendance and test scores. Section IV lays out a model of education production that explores the mechanisms behind the heterogeneity impacts. Section V tests the implications of the model. Section VI concludes.

II. Experimental Design

The intervention was conducted from July through September of 2007 in Gurgaon, a suburb of Delhi.[7] Eight government-run primary schools were selected based on proximity to the city center and availability of public transportation nearby.

5. 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.
6. Studies of developed-country incentives programs find mixed evidence on their effectiveness (Angrist and Lavy 2009; Angrist, Lang, and Oreopoulos 2009; Jackson 2010; Fryer 2011; Bettinger 2010).
7. Between December 2006 and March of 2007, a pilot was conducted with 138 children in three schools that were not chosen for the study. Households were randomized into money, child toy, and choice treatments. The pilot allowed the research team to practice the protocols, refine the survey instruments, and calibrate the goal levels so that approximately 50 percent of children would achieve these goals.

In seven schools, all first, second, and third grade students participated in the program. In one school, first grade children were excluded due to administrative difficulties in obtaining these students' addresses.[8] The experiment consisted of a pre-test, announcement of the child's incentive scheme, and a post-test approximately two months later.[9] Children were initially tested in schools to determine baseline learning levels. Each child scoring below the highest level on the test was given a goal competency based on his pre-test score and was administered one of six randomly assigned incentive schemes. The treatments were assigned at the individual level. In order to increase power to detect heterogeneity by pre-test score, the randomization was stratified by pre-test score within each school, grade, and classroom.

Award of the incentive depended on the performance of the children on a post-test conducted two months after the program announcement. 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, 100 rupees was the approximate daily wage for an unskilled laborer in these areas.

A. Treatment Groups

The experiment consisted of six treatment groups. Four treatment groups assigned the household a reward that varied 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 two remaining groups offered the parent a choice between money for herself and a toy, either upon program announcement or conditional on reaching the goal.

The first two treatments (hereafter referred to as the "money" treatments) offered money either to the parent or to the child. The parent money treatment offered a reward of 100 rupees to the parent if the child achieved the goal. The child money treatment offered a reward of 100 rupees to the child if the child achieved the goal.

Comparison of results of the money treatments tests whether the recipient of the money has an impact on the effectiveness of the incentives. 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, the child money treatment would represent more control over the reward by the child; however, if income from the parent and child is pooled in the household, the child money treatment would be equivalent to the parent money treatment.

The next two treatments (hereafter referred to as the "toy" treatments) offered a reward of a toy valued at 100 rupees to the child if the child achieved the goal, either in the form of a toy or a voucher redeemable at a local toy store. In the offer scripts, the cost of the toys in each treatment was emphasized to equalize subjective valuations of the toys to the extent possible. In the child toy treatment, the script identified the four

8. 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 main results are robust to the exclusion of these additional students.

9. The timeframe was chosen to roughly correspond to the Pratham NGO's "Learn to Read" program timeline. This program focuses on the same competencies tested in this evaluation. The objective of the program is to improve reading skills by one or two levels within one to two months (Pratham 2014).

toys that the child could choose from (cricket set, doll, car, or drawing set) and stated that each cost 100 rupees.[10] In the voucher treatment, the script stated that the voucher would be redeemable for a toy that cost up to 100 rupees.

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. Although 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.[11] 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 percent of parents reported having given their child a toy during the week before 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 or the voucher was given directly to the child.

Although the two toy treatments were both designed to ultimately reward the children with toys, they differed in several ways. First, while the toys chosen for the child toy treatment were selected to appeal to as many children as possible, it is possible that some children would not value them as much as others. The vouchers would allow the children to choose from a wider variety of toys.[12] Second, because the voucher had to be taken to a toy store, the parents had some control over whether and when the voucher could be redeemed. In that case, the child may have less control over the reward than in the toy treatment, where he chose and received the toy at school.

The experiment also included two treatment groups that tested whether parents want to reward their children for positive outcomes but cannot commit to doing so. To test this hypothesis, 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.[13]

Regardless of treatment category, all children were invited to attend free afterschool 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 parateacher 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 to 30 students who

10. In practice, stores often allow consumers to negotiate, as was the case in the toy store from which the toys were obtained. Thus, the toys selected were chosen so that the price a typical shopper would pay was 100 rupees.

11. While I do not have evidence that resale of toys occurred during the experiment, a small number of interviews conducted after the pilot indicated that households receiving the toys did not sell them.

12. In practice, 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.

13. The online Appendix 1 provides additional discussion of these treatment groups. Online appendices are available at http://jhr.uwpress.org/.

attended the classes.[14] 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 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 an 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 2012). Random checks are usually unable to measure attendance on a daily basis because they disrupt class and are difficult to take accurately. Teachers of the afterschool classes in this study were monitored carefully to ensure accuracy.

It should be noted that extra classes are not typically part of incentives-to-learn or conditional cash transfer programs, and as such this program does differ somewhat from what is commonly implemented. However, the primary purpose of the program was to understand household decision-making based on the different incentive schemes, and this departure from the standard intervention was chosen to provide a measure of effort and to provide a uniform opportunity for children to learn during the program. The influence of the classes on the overall effects of the incentives schemes is discussed in the online Appendix 3.

B. Pre-test

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 2014). 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 multiparagraph story.

Based on each child's pre-test score, the child was given a goal competency to be reached when he was retested. If the child achieved the goal, he or his parent would receive a prize as per his treatment category. Children reading at levels zero and one were each given a goal one level above their respective pre-test scores, while children at levels two and three were given a goal of four. 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 at the pre-test 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

14. In one school, the principal did not allow the afterschool class teachers access to the school premises to conduct the class, and no suitable alternative location was found.

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.[15] The continuous measure is then normalized by subtracting the mean pre-test score in the child's grade and dividing by the standard deviation.[16]

C. Baseline Survey and Program Announcement

After the pre-test, each child was randomly assigned to one of the six treatment groups outlined in Section IIA. As noted above, the randomization was conducted at the individual level and was stratified by pre-test score within each school, grade, and classroom.

Approximately one week after the pre-test, a baseline survey was conducted at the child's home, and the household's assigned incentive scheme was announced to the mother and child.[17] 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.[18]

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 and if the goal competency were reached the mother or child would receive the specified prize.[19] 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.

Figure 1 displays the progression of the sample into the treatment groups, and Table 1 shows the sample composition by pre-test score, grade, and treatment category. Out of 1,466 children who took the pre-test, 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. One thousand eighty-six children were thus available for the randomization. Eighty-five percent of children out of the randomized group of 1,086 were reached for the baseline survey and program

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

16. In order for the grade-specific means to be representative of all children in the schools, I include the highest pre-test scores in constructing this variable.

17. Announcement of the incentive schemes was 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. While I do not find statistically significant evidence of peer effects, point estimates suggest that having a friend in the toys treatments lowered attendance in the afterschool classes for those in the money treatments (results not shown).

18. Survey instruments and treatment scripts are available from the author upon request.

19. The script initially stated that the post-test would be conducted after one month, but the timeline was subsequently extended to two months. The exact date of the post-test was printed on the reminder cards described below.

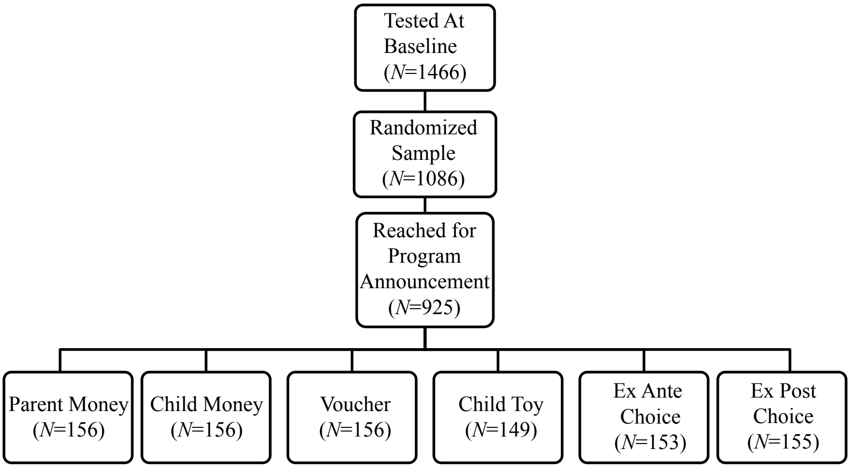

Figure 1
Sample Composition Across Treatment Groups

Table 1
Sample Composition

	Reached at Pre-Test 1	Randomized Sample 2	Reached for Program Announcement 3	Reached at Post-test 4
Total	1,466	1,086	925	900
Raw pre-test 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
Child toy		182	156	151
Voucher		180	149	145
Ex ante choice		183	153	151
Ex post choice		181	155	151

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.[20] Of the 925 children offered the program, 900 (97 percent) 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. The online Appendix 2 analyzes attrition patterns across treatment groups and baseline characteristics of attritted households. The appendix shows that attrition is not significantly related to treatment assignment. In addition, bounds using Lee's (2009) trimming method produce very small intervals around the treatment effects estimates and do not substantively change the interpretation of the achievement results. The attendance results include all children offered the program regardless of whether they took the post-test.

The final analysis sample contains approximately 150 children in each of the six treatment groups. A larger sample of 330 children per treatment group had initially been planned, but the sample was ultimately limited due to budget constraints. Precision of the estimated effect sizes is discussed in the results section. In addition to analyzing the separate treatment groups, the results section also aggregates similar groups to increase precision.

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.[21] Column 2 presents the *p*-values of a joint test of equality of means of each characteristic across the six treatment categories. All 14 *p*-values are at least 0.1. The *p*-value is almost exactly 0.1 for transfers of sweets to the child, 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 aggregated parent and child money treatments with the aggregated voucher child toy treatments. Again, none of the differences are significant at the 10 percent level. Although 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.

D. Reminder and Followup 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 ten 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. At the end of the survey, a short script was read reminding the parent and child of the program. The surveyor also asked to see the reminder cards that had been previously given to the households. Ninety percent of households were able to show the reminder cards, and there are no significant differ-

20. 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 pre-test 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.

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

Table 2
Balance of Observables Across Treatment Groups

Variable	Full Sample Mean 1	*P*-Value of Differences: Six Treatment Categories 2	*P*-Value of Differences: Money Versus Toy 3
Raw pre-test score	1.024	0.291	0.449
Relative pre-test score	–0.310	0.760	0.853
Female	0.571	0.768	0.318
Number of children 0–14	2.915	0.855	0.787
Number of adults 15+	2.442	0.131	0.118
Mother employed	0.345	0.446	0.610
Mother education	3.176	0.468	0.108
Father education	6.421	0.641	0.101
Durables	0.000	0.455	0.429
Helped with studies	0.358	0.181	0.973
Tutoring fees paid	26.585	0.521	0.955
Money given	13.065	0.119	0.127
Gave toys	0.037	0.320	0.699
Gave sweets	0.190	0.100	0.604
Gave clothes	0.015	0.745	0.780
Gave school supplies	0.064	0.709	0.700

Notes: Sample includes all households surveyed at baseline. Column 1 presents the mean of each variable over the entire sample. Column 2 presents the *p*-values of a test of equality of means across the parent money, child money, voucher, child 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 treatment groups. "Raw pre-test score" represents the child's integer pre-test score. "Relative pre-test 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 rupees) 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," "Gave sweets," "Gave clothes," and "Gave school supplies" are dummy variables indicating whether the parent gave the child the item indicated over the past week. Standard errors are clustered at the classroom level.

ences across treatment groups in retaining the cards (results not shown). To the extent that retaining the cards reflected trust that the prizes would be delivered, there is thus no evidence that trust differed across treatment groups.

E. Post-Test and Second Followup Survey

Approximately two months after the program announcement, a post-test, similar in form but not in exact content to the pre-test, 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.[22] Approximately one week after the post-test, a second followup survey was conducted to again measure transfers given by the parent to her child. The purpose of this survey was to examine transfers after the post-test had been conducted and rewards had been distributed.

III. Results

In this section, which presents the main results of the experiment, I restrict attention to the impacts of the four money and toy treatment groups; impacts of the choice treatments are presented in the online Appendix 1. Section IIIA compares average outcomes across the incentives treatments and finds no significant differences between the money and toy treatment groups. Section IIIB presents evidence of heterogeneity in treatment effects by initial test score: Lower-performing children had higher attendance and achievement in the toy treatments while higher-performing children had higher attendance and achievement in the money treatments.

The experiment did not contain a randomized control group, and therefore I cannot offer experimental evidence on the overall effectiveness of the incentives. I can, however, offer quasi-experimental evidence that the incentives substantially improved test scores. The group of children included in the randomization but not reached for the program announcement serves as a quasi-experimental control group. These children had remarkably similar pre-test scores compared with those reached for the announcement. At the post-test, children who were reached for the program announcement were 27 percent more likely to achieve the program goal and on average scored 0.61 standard deviations higher than children who were not reached. This analysis is presented in detail in the online Appendix 3.

Before turning to the impact estimates across treatment groups, it is useful to examine how reward money was spent in the parent money and child money treatments. Parents in the money treatments could have spent the money on toys, thereby negating the differences between the toy and money treatments. Table 3 tabulates self-reports on how the reward money was spent in the parent money and child money groups. Parents were asked how they or their children spent the money and could give multiple answers (although detailed data on amounts spent in each category were not collected). In both treatments, 27 percent of the time parents reported that at least some of the prize money was spent on toys. About half the time money was spent on clothes and another 21 percent of the time it was spent on school supplies. Seven percent of the time, the money was spent on household items or had not been spent at the time of the second followup survey. Thus, while 73 percent of households did not spend reward money on toys, 71 percent spent at least some of the money on clothing or school supplies. These items could have been used as rewards to motivate the children, or parents could have decided to buy those items with the reward money because the

22. Prizes were distributed in school for the child money, voucher, and child toy treatments, in addition to toys chosen in the ex ante choice treatment. Parents in the parent money treatment and those who chose money in the ex ante choice 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.

Table 3
Spending of Reward Money

	Treatment Group		
	Parent and Child Money 1	Parent Money 2	Child Money 3
Toys	0.27	0.25	0.28
Sweets	0.03	0.05	0.01
Clothes	0.54	0.53	0.55
School supplies	0.21	0.23	0.19
Household items / not spent	0.07	0.11	0.04
Observations	161	83	78

Notes: Each cell represents the proportion of parents reporting spending the reward money in the listed category. Categories are not mutually exclusive. Joint test that proportions in Columns 2 and 3 are equal: p-value = 0.40.

money was associated with a child learning program. I do not have detailed data on child preferences across toys, clothes, and school supplies, and the extent to which each would motivate the child to learn. Nonetheless, these results show that the majority of parents were not using the money to mimic the toy treatments.

It should be noted that reported spending of the prize money could represent inframarginal spending on children. The online Appendix 4 extends this analysis by examining self-reported transfers to children in each treatment group both before and after the experiment. Before the post-test, there is evidence of small differences in transfers between the money and toy treatment groups. Consistent with reported spending of the prize money, parents in the money treatment groups bought substantially more clothing and school supplies for their children after the post-test.

A. Comparison of Treatment Groups

Panel A of Table 4 presents OLS estimates of a regression of outcomes on dummy variables for each treatment group using the parent money group as the omitted category. In the first two columns, the outcome is a dummy variable that equals one if the child attended the afterschool classes. None of the differences is statistically significant. Columns 3 and 4 present the results of similar regressions using a dummy for achievement of the goal competency as the outcome variable. Again, the estimates are not statistically significant. Taken together, these results provide no evidence for differences in average attendance or achievement between treatment groups.

While these results are not statistically different from zero, it is useful to consider the effect sizes that can be ruled out based on the confidence intervals. To frame the effect sizes, 23 percent of students attended the afterschool classes and 54 percent achieved the literacy goal across all four treatments. The estimated effect of any incen-

Table 4
Overall Treatment Effects

	Dependent Variable: Attendance in Afterschool Classes		Dependent Variable: Achievement of Literacy Goal	
	1	2	3	4
Panel A: Four main treatment groups				
Child money	0.038	0.046	–0.040	–0.020
	(0.051)	(0.053)	(0.057)	(0.060)
Voucher	0.046	0.056	–0.037	–0.025
	(0.054)	(0.051)	(0.063)	(0.069)
Child toy	0.000	–0.005	0.012	0.022
	(0.056)	(0.058)	(0.062)	(0.068)
Pre-test dummies	NO	YES	NO	YES
Classroom dummies	NO	YES	NO	YES
Observations	502	502	598	598
R-squared	0.003	0.158	0.002	0.191
Mean of dependent variable	0.210	0.210	0.553	0.553
Panel B: Money versus toy				
Toy	0.004	0.002	0.007	0.008
	(0.035)	(0.034)	(0.048)	(0.053)
Pre-test 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 dependent variable	0.229	0.229	0.533	0.533

Notes: In Columns 1 and 2, the dependent variable is a dummy that equals one if the child attended the afterschool classes on at least one day. In Columns 3 and 4, the dependent variable is a dummy that equals one if the child reached the literacy goal. The omitted treatment category in Panel A is parent money. The omitted treatment categories in Panel B are parent and child money. "Toy" represents the aggregated child toy and voucher treatments. The mean of the dependent variable is calculated for the omitted category in each regression. Standard errors are clustered at the classroom level. * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

tive on achievement, as determined in the online Appendix 3, is 27 percent. Using the estimates from Column 1, the 95 percent confidence interval of the effect of the child toy treatment on attendance relative to the parent money treatment admits effects as large as 11 percent and as small as –11 percent. This interval rules out differences above half of the attendance rate in the sample. The corresponding interval for the effect of the child toy treatment on achievement ranges from –11 percent to 14 percent. This interval rules out differences above about half of the estimated effect of any incentive.

I next aggregate the treatment groups by the form of the reward: Households in the money treatments are compared with households in the toy treatments. Several pieces of evidence suggest 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. When asked what they would do if given 100 rupees, over 80 percent of children in the baseline reported that they would give it directly to their parents. In addition, while 51 percent of parents reported giving their children spending money within a day of the survey, the majority (76 percent) 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. Finally, as shown in Table 3, reported spending of the reward money was remarkably similar across the parent and child money treatments. A joint test of proportions fails to reject equivalence (p-value = 0.40).

Panel B of Table 4 presents OLS estimates comparing average outcomes between the aggregated money treatments with the aggregated toy treatments. There are no significant differences between categories using either attendance or achievement as the outcome. As expected, the estimates are more precise than those using the disaggregated treatments. For example, using the estimates on achievement from Column 3, the confidence interval rules out effect sizes below minus nine percentage points and above ten percentage points.

B. Heterogeneity by Pre-test Score

Table 5 analyzes heterogeneity in treatment effects by pre-test score. Panel A regresses either attendance in the afterschool classes or achievement of the literacy goal on interactions between pre-test scores and the child money, voucher, or child toy treatment groups. As before, the omitted category is parent money. Columns 1 and 2 display the results using attendance as the outcome of interest using either the categorical pre-test score or the continuous measure of pre-test score as the interacted variable. The interaction between test score and the child money treatment is small in magnitude, inconsistently signed across both specifications, and statistically insignificant. The interaction with the voucher treatment is negative and is significant in one of the two specifications. The child toy treatment interaction is large, negative, and statistically significant at least the 10 percent level in both specifications. The magnitude of the estimate in Column 2 indicates, for example, that a child scoring one standard deviation higher on the pre-test was 17 percentage points more likely to attend the afterschool classes in the parent money treatment compared with the child toy treatment.

Columns 3 and 4 of Panel A in Table 5 repeat this analysis using achievement of the literacy goal as the outcome of interest. A similar pattern emerges: The interaction effect of child money is small and inconsistently signed. The interaction of the voucher treatment is negative but statistically insignificant. The child toy treatment has a negative interaction with pre-test scores, and the estimate using relative pre-test score is statistically significant at the 5 percent level. The magnitude of the estimate implies that a student who scored one standard deviation higher on the pre-test was 11 percentage points more likely to achieve the goal in the parent money treatment relative to the child toy treatment.

Panel B of Table 5 repeats this analysis comparing the aggregated toy treatments with the aggregated money treatments. The results show consistent evidence of a negative interaction of pre-test scores and toys relative to money. Using attendance as the

Table 5
Interactions of Treatment Groups and Pre-test Scores

	Attendance		Achievement	
	Categorical Score 1	Relative Score 2	Categorical Score 3	Relative Score 4
	Panel A: Six main treatment groups			
Child money * pre-test	0.000	–0.076	0.011	–0.04
	(0.070)	(0.060)	(0.057)	(0.071)
Voucher * pre-test	–0.134**	–0.094	–0.085	–0.067
	(0.054)	(0.071)	(0.054)	(0.071)
Child toy * pre-test	–0.154*	–0.174**	–0.073	–0.111**
	(0.076)	(0.065)	(0.049)	(0.053)
Pre-test dummies	YES	YES	YES	YES
Classroom dummies	YES	YES	YES	YES
Observations	502	502	598	598
R-squared	0.180	0.182	0.197	0.220
	Panel B: Money versus toy			
Toy * pre-test	–0.144***	–0.099**	–0.084*	–0.069
	(0.044)	(0.042)	(0.042)	(0.046)
Pre-test dummies	YES	YES	YES	YES
Classroom dummies	YES	YES	YES	YES
Observations	502	502	598	598
R-squared	0.176	0.171	0.195	0.218

Notes: In Columns 1 and 2, the dependent variable is a dummy that equals one if the child attended the after-school classes on at least one day. In Columns 3 and 4, the dependent variable is a dummy that equals one if the child reached the literacy goal. The omitted treatment category in Panel A is parent money. The omitted treatment categories in Panel B are parent and child money. "Toy" represents the aggregated child toy and voucher treatments. All regressions include dummies for treatment categories. Columns 2 and 4 control for relative pre-test score in addition to categorical pre-test score dummies. Standard errors are clustered at the classroom level. * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

outcome, both coefficients are significant at the 5 percent level, while using achievement as the outcome, one of the two coefficients is significant at the 10 percent level. The magnitudes imply that a child who was one category higher at pre-test was 14 percentage points more likely to attend the classes and eight percentage points more likely to achieve the goal in the money treatments relative to the toy treatments. Alternatively, a child who scored one standard deviation higher on the pre-test was ten percentage points more likely to attend the classes and seven percentage points more likely to achieve the literacy goal in the money treatments relative to the toy treatments.

Figure 2 presents less parametric estimations of the treatment effects by initial learning level. Panel A graphs the estimated differences in attendance between the toy treatments and the money treatments by the students' raw pre-test score, controlling for classroom dummies. The treatment effects are decreasing across all four categories of test scores. In the lowest score category, children in the toy treatment were 16 percent more likely to attend the classes, a difference significant at the 1 percent level. In the highest two score categories, children in the money treatments were 18 percent and 24 percent more likely to attend the classes for baseline scores of two and three, respectively. The treatment effect for a baseline score of two is significant at the 5 percent level, while the effect for a baseline score of three is not significant. The F-test for the joint significance of all four effects definitively rejects the null hypothesis that all effects are zero (p-value < 0.01). Panel B repeats this exercise using achievement of the goal competency as the outcome of interest. As with the effects on attendance, there is a monotonically decreasing relationship between the relative effect of the toy treatments and the children's pre-test scores. Children with the lowest pre-test score were 7 percent more likely to achieve the goal in the toy treatments relative to the money treatments. In the highest test score category, children were 18 percent less likely to achieve the goal in the toy treatments than in the money treatments. In this case, however, none of the individual effects is significant, and the F-test does not reject the null hypothesis that all effects are zero (p-value = 0.36).

While these results suggest an interaction between the effectiveness of incentives and initial learning level, one alternative interpretation of this heterogeneity is that it might reflect differences in difficulty of reaching the different goal competencies. For example, it may be more difficult to move from a score of zero to one than from one to two. There are several pieces of evidence that suggest this is not the case. First, the attendance measure represents the same level of effort across all baseline test scores, and thus the attendance results are less likely to reflect differences in difficulty of reaching the different goals. Second, the achievement levels of children in the nonexperimental control group (as described in the online Appendix 3) suggest a nonmonotonic relationship between pre-test score and difficulty of achievement. Achievement in this group was 43 percent, 24 percent, 21 percent, and 41 percent for pre-test scores of zero, one, two, and three, respectively. These results suggest that the most difficult goals were for children with the middle two pre-test scores. Indeed, as shown in the online Appendix 5, there is no evidence that overall program effects varied by pre-test score.

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

IV. Theoretical Framework

To explore the mechanisms behind the heterogeneous treatment effects found in the previous section, I adapt a model of education production from the literature to the context of the experiment. In the model, the parent and child both contribute inputs toward child learning. By modeling the education production pro-

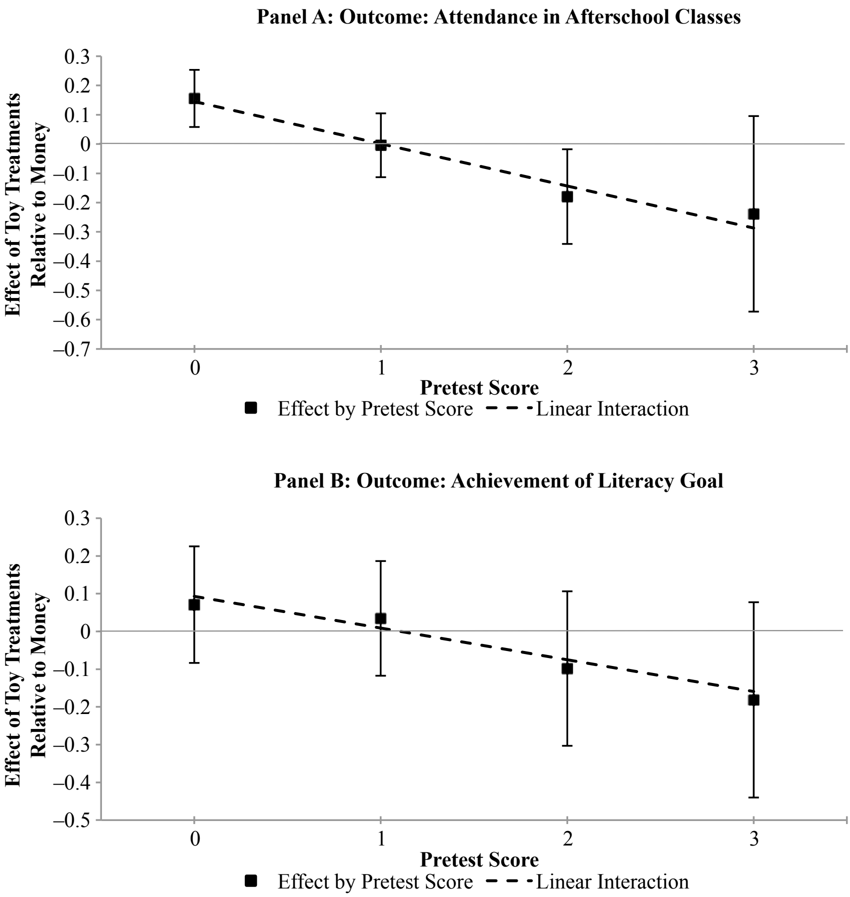

Figure 2
Relative Effects of Toy Treatments by Pre-test Score

Notes: Dots represent the coefficients of the regressions in Columns 1 and 3 of Panel B in Table 5, where the raw pre-test score is replaced by dummies for each pre-test category. Whiskers represent the 95 percent confidence interval on each coefficient. Dotted lines were constructed using the estimates in Columns 1 and 3 of Table 5, Panel B.

cess, I find that the key variation driving heterogeneity in impact is the productivity of the parent's input relative to the child's. Based on this result, I then derive testable implications regarding how the productivity of parental inputs influence test scores and in turn influence the relative effectiveness of parent incentives.

In the model, current learning levels are a function of prior learning levels plus contemporaneous inputs, similar to a "value-added" education production model (Todd and Wolpin 2007). To incorporate uncertainty, the child's learning level l is discrete, and the probability that the child moves up one level in period t is a function of parent and child inputs in that period:

(1) $P(l_t = l_{t-1} + 1) = g(p_t, c_t)$

where the parent and child contribute inputs p_t and c_t, respectively.

The probability that a child reaches the next level in period t is modeled as a linear function of the inputs:

(2) $g(p_t, c_t) = \beta_p p_t + \beta_c c_t$

where the parameters $\beta_p \geq 0$ and $\beta_c \geq 0$ capture the productivity of the parent's and child's inputs, respectively, in education production.

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.[23] 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.

To match the context of the experiment, I first model education production before the experiment, to generate the children's initial learning level, and then during the experiment, to generate the child's achievement within the experiment.

A. Preexperiment

In each preexperimental period, the parent and child decide how much input to contribute. I assume that $l_0 = 0$ and that production has taken place over m periods for a child of age m. Because the production function is additively separable across periods, each period can be analyzed separately. I suppress the time subscript from variables for ease of exposition.

In this model, the key friction within the household is a two-sided moral hazard similar to the sharecropping model of Eswaran and Kotwal (1985).[24] The parent values success in achieving an additional level of learning, but the child does not. The parent and child cannot contract on inputs, but the parent can decide how much to give the child to induce effort. The child's effort depends on the amount transferred by the parent, and the parent's effort depends on the amount that the parent keeps for herself.

The parent places a normalized value of one on success in each period. That is, she receives one if the child successfully moves up a level and zero if he does not.[25] The costs of input to the parent and child are given by the quadratic functions $k_p p^2 / 2$ and $k_c c^2 / 2$ in each period.[26]

Before input choices are made, the parent announces a fraction γ of the value of

23. This conceptualization of parent and child inputs in education production follows Leibowitz (1974).

24. As noted in the introduction, previous work has utilized moral hazard in parent-child interactions to explain household behaviors (see, for example, Weinberg 2001, Gatti 2005).

25. I assume that all parents place the same value on success 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 Section VC and in the online Appendix 7.

26. I assume that the cost parameters k_p and k_c are large enough to ensure that the probability of success is always less than one. A sufficient condition for this probability to be below one both before and during the experiment is that $k_p > \beta_p^2(1 + \pi)$ and $k_c > \beta_c^2(1 + \pi)$, where π is the value of the reward within the experiment. This follows because during the experiment, the probability of success equals $[(1 - \gamma)(1 + \pi)\beta_p^2 / k_p] + [\gamma(1 + \pi)\beta_c^2 / k_c]$, where $0 \leq \gamma \leq 1$.

success that she will give to the child if the child moves up a level. I assume that the parent must commit to this division of value until the 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.[27] If the child is unsuccessful, he will receive no transfer, and if he is successful, he will receive a fraction γ of the value of success.

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

I assume that both the parent and child are risk neutral and therefore make their decisions based on the expected value of success $g(p, c)$. 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 or c, taking the parent's choice of γ as given:

$$(3) \quad p(\gamma) = \arg\max_{p}(1-\gamma)\{\beta_p p + \beta_c c\} - \frac{k_p p^2}{2}$$

$$(4) \quad c(\gamma) = \arg\max_{c} \gamma\{\beta_p p + \beta_c c\} - \frac{k_c c^2}{2}.$$

Subject to the incentive-compatibility constraints of Equations 3 and 4, the parent maximizes her expected share of the value of success over the fraction γ she gives to the child:

$$(5) \quad \max_{\gamma}(1-\gamma)\{\beta_p p + \beta_c c\} - \frac{k_p p^2}{2}.$$

This yields the sharing rule $\gamma^*(\beta_p, \beta_c, k_p, k_c)$ and the probability of success, denoted by $g \equiv \beta_c p(\gamma^*) + \beta_c c(\gamma^*)$.

If the experiment begins after m periods, the child's learning level at the beginning of the experiment will have a binomial distribution:

$$(6) \quad P(l_m = l) = (m\,/\,l)g^l(1-g)^{m-1}.$$

The actual learning level can be written as the expected learning level based on Equation 6 plus an error term:

$$(7) \quad l_m = m(\beta_p p(\gamma^*) + \beta_c c(\gamma^*)) + \varepsilon.$$

PROPOSITION 1: *At the beginning of the experiment, the child's relative learning level is increasing in both parental and child productivity.*

PROOF. See the online Appendix 6.

27. 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.

B. Experiment

After m periods, the experimenter offers an external incentive of value π to either the parent or the child conditional on success in the experimental period. Assume the parent places a value of one on success in this period in addition to this incentive.

If this external incentive accrues to the parent, the parent decides on the share to transfer to the child γ_e based on the external incentive plus the parent's own value of success ($\pi + 1$). This yields a new sharing rule $\gamma^*_e(\beta_p, \beta_c, k_p, k_c)$, and a probability of success of

(8) $g_{e,parent} \equiv \beta_p p_e(\gamma^*_e) + \beta_c c_e(\gamma^*_e)$.

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 $\pi / (\pi + 1)$ 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

(9) $\gamma^*_e < \gamma_{child}$.

I assume throughout that π is large enough such that Equation 9 holds.[28] The probability of success in this case is given by

(10) $g_{e,child} \equiv \beta_p p_e(\gamma_{child}) + \beta_c c_e(\gamma_{child})$.

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

(11) $g_{e,parent} - g_{e,child} = \beta_p p_e(\gamma^*_e) + \beta_c c_e(\gamma^*_e) - [\beta_p p_e(\gamma_{child}) + \beta_c c_e(\gamma_{child})]$.

The remainder of this subsection uses Equation 11 to make predictions for the relative effectiveness of parent incentives versus child incentives when parent and child productivity (β_p and β_c) vary across households.

Proposition 2: *When* $\beta_p\sqrt{k_c} > \beta_c\sqrt{k_p}$, *incentives to the parent will result in a higher probability of success than incentives provided to the child. When* $\beta_p\sqrt{k_c} < \beta_c\sqrt{k_p}$, *incentives provided to the child will result in a higher probability of success than incentives provided to the parent.*
Proof. See the online Appendix 6.

This result shows that when the parent's input is relatively more productive, incentives to the parent are more effective than incentives to the child. When the child's input is more productive, incentives to the child are more effective than incentives to the parent.

To develop the intuition behind these results, consider the extreme case where the parent's input is completely unproductive, that is, $\beta_p = 0$. In this case, the probability of success will be maximized when the child is given the entire incentive. However, when the parent is given the incentive, she will divert some of the rewards toward herself instead. Thus, an intervention that allocates the incentive to the child will result in higher achievement than an intervention that allocates the incentive to the parent. At the other extreme, consider the case where the child's input is completely unproduc-

28. The online Appendix 4 shows that after the post-test children received different transfers across the money and toy treatments, suggesting that this restriction was indeed binding.

tive, that is, $\beta_c = 0$. In this case, any rewards or transfers to the child will not influence the likelihood of success. Thus, an intervention that rewards the child directly will not be effective while rewarding the parent will be effective. The formal proposition can therefore be seen as an extension of these cases to intermediate levels of parent and child productivity.

Propositions 1 and 2 show how parental and child productivity are directly related to both the child's initial test score and the effect of child incentives relative to parent incentives. To the extent that test scores reflect parental productivity rather than child productivity, children with higher test scores will perform better when their parents are given incentives, and children with lower test scores will perform better when they are given incentives directly. Thus, the variation in pre-test scores attributable to parental productivity will yield a stronger interaction with child incentives than pre-test scores overall.

Note that in this model, conditional on the current learning level, inputs in prior periods do not affect current-period learning. This assumption could be violated if, for example, a child who has attempted a level but failed finds it easier to achieve that level in the next period. Alternatively, a child who has just reached a level in the previous period may find it harder to learn the next level in the current period. In the context of the experiment, this would imply that the treatments could interact with previous inputs beyond the basic interaction with learning level. One way to test for this effect would be to examine whether children who were initially closer to their goals, as measured by their continuous pre-test scores, had different treatment effects from those who were initially farther from their goals. However, I find no evidence for these differential effects (results not shown).

V. 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 to teach her child and manage her child's time. I then estimate the effects of interactions between the resulting index and the toy treatments.

A. Results

To proxy for parental productivity, I use a broad set of eight variables from the baseline survey. These variables reflect a set of characteristics and behaviors that could potentially reflect the parent's ability to manage the child's time and motivate the child, following the literature on parental investments in children (see, for example, Liebowitz 1974, Haveman and Wolfe 1995).[29,30]

29. The online Appendix 7 provides additional analysis of the components of the productivity index and the robustness of the results to exclusion of various components.

30. Measures of direct transfers are excluded because they are endogenously determined in the model. Several other measures are excluded due to lack of variation.

These variables fall into four broad categories. First, I include three variables that reflect household composition and mother availability at home. More children younger than 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 older than age 15 are expected to positively affect productivity because these members represent resources the child can use for help with his studies. 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.[31] Second, I include two variables that reflect the education status of the child's parents. Mother's and father's education are also included because more-educated parents should be more 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.

These variables are chosen to proxy for parental productivity in the absence of direct measures. The variables could, to some extent, reflect other parameters in the model. For example, while these variables are meant to capture parental productivity, some may relate to costs of effort. In the case of the number of children in the household, a larger number of children may be considered a larger opportunity cost of providing attention to any one child rather than the parent's ability to provide attention conditional on cost. However, the key implication of the model cannot distinguish between a higher productivity parameter β_p and a lower cost parameter k_p because parental incentives will be more effective when $\beta_p / \sqrt{k_p} > \beta_c / \sqrt{k_c}$. Thus, while I am unable to disentangle the extent to which each measure represents lower costs rather than higher productivity, the theoretical implications are the same.

In addition, some of the variables chosen may reflect the actual input rather than the parent's ability to provide that input. For example, the amount of money spent on tutoring could reflect parental input rather than productivity. In the model, the amount of the parent's input depends positively on parental productivity. Therefore, although these variables may reflect actual input at baseline, they still can serve as proxies for parental productivity.

To generate the index of parental productivity, I first regress the child's relative pre-test score on these variables. The results of the first-stage regression of relative pre-test score on the parental productivity variables are reported in Table 6.[32] 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

31. Father's employment status is not included because only 4 percent 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.

32. Because households with children whose pre-test scores were in the highest category were not surveyed, these scores are not included in this regression.

Table 6
Relationship Between Relative Pre-Test Score and Parental Productivity Measures

	Dependent Variable: Relative Pre-test Score	
	1	2
Children under 15	–0.053**	–0.041*
	(0.023)	(0.022)
Adults 15+	0.016	0.051
	(0.053)	(0.057)
Durables	–0.063***	–0.034*
	(0.019)	(0.017)
Mother employed	–0.074	–0.052
	(0.075)	(0.075)
Mother education	0.021*	0.015
	(0.011)	(0.011)
Father education	–0.004	–0.001
	(0.009)	(0.008)
Helped with studies	0.059	0.002
	(0.062)	(0.057)
Tutoring fees paid	0.178***	0.150***
	(0.044)	(0.048)
Classroom dummies	NO	YES
Observations	925	925
R-squared	0.045	0.245

Notes: The sample used in this table includes all households surveyed at baseline. 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 percent; ** significant at 5 percent; *** significant at 1 percent.

are significant at the 10 percent level. The results are similar when classroom dummies are added in Column 2.[33]

The predicted values of the regressions in Columns 1 and 2 of Table 6 form a parental productivity index, as they represent the extent to which these variables contribute to test scores. Table 7 presents the effects of the index interacted with the toy treatments.[34,35] Columns 1 and 2 use attendance in the afterschool classes as the

33. It is somewhat surprising that durables ownership is significantly related to *lower* initial achievement. However, as discussed in the online Appendix 7, the inclusion of the durables measure does not drive the results presented in this section.

34. The index used in Columns 2 and 4 of Table 7 includes only the effects of the productivity variables and not the classroom dummies.

35. 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.

Table 7
Interactions of Toy Treatments and Productivity Index

	Dependent Variable: Attendance in Afterschool Classes		Dependent Variable: Achievement of Literacy Goal	
	1	2	3	4
Toy	–0.158***	–0.186***	–0.062	–0.105
	(0.054)	(0.058)	(0.077)	(0.073)
Productivity index	0.296***	0.298***	0.072	0.151
	(0.094)	(0.091)	(0.139)	(0.133)
Toy * productivity	–0.535***	–0.627***	–0.222	–0.358**
	(0.138)	(0.144)	(0.194)	(0.181)
Pre-test 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	0.167	0.166	0.191	0.193

Notes: In Columns 1 and 2, the dependent variable is a dummy that equals one if the child attended the afterschool classes on at least one day. In Columns 3 and 4, the dependent variable is a dummy that equals one if the child reached the literacy goal. The omitted treatment categories are parent and child money. "Toy" represents the aggregated child toy and voucher treatments. 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 percent; ** significant at 5 percent; *** significant at 1 percent.

outcome of interest. The estimated coefficients on the interaction terms are negative and significant at the 1 percent level in both specifications. The magnitudes of these coefficients are more than five times the estimates using the actual pre-test scores, suggesting that the relative effectiveness of the toy treatments is more strongly related to the portion of pre-test scores driven by these parental characteristics than to pre-test scores overall. The difference in magnitudes is consistent with the theory because the theory predicts that relative effectiveness of the toy treatments will be directly related to the *share* of test scores that reflect parental productivity. To the extent that test scores reflect both parental and child productivity, isolating a component of test scores related to parental productivity results in a stronger interaction.

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 5 percent level in only one of the two specifications.

Note that while the productivity index reflects the extent to which the measured parental characteristics and behaviors correlate with child test scores, it does not fully isolate the variation in test scores attributable to parental productivity. Once accounting for these characteristics, the remaining variation in test scores could reflect unobserved factors that contribute to either parent or child productivity. Indeed, after accounting for the interaction with the productivity index, the interactions between test scores and the toy treatments become smaller in magnitude but remain negative (see online Appendix 7).

Further, while the variables used in the parental productivity index are chosen to reflect parental productivity, they could, to some extent, reflect child productivity. The key implication is that the interaction between the toy treatments and the index will be negative *to the extent that* they reflect parental productivity rather than child productivity. If these variables were more correlated with child productivity than with parental productivity, the theory would predict an interaction in the opposite direction: At high levels of the index, the toy treatments would be more effective and vice versa. Instead, I find a large and negative interaction with the toy treatments. The fact that the interaction is larger in magnitude than the interaction with test scores provides further evidence that the index is related to parental productivity rather than child productivity.

In sum, the results of this section provide empirical evidence that parental productivity contributes to the interaction between pre-test scores and the toy treatments. Consistent with the theory, the variation in initial test scores attributable to proxies for parental productivity produces stronger interactions with the toy treatments than do test scores overall. These results imply that incentives to parents were more effective when parents were more able to teach their children and motivate them to learn while incentives to children were more effective when their parents were less able to teach and motivate them.

B. Robustness: Components of the Index

The online Appendix 7 provides additional analysis of the components of the productivity index and its interaction with the toy treatments. I show that the interaction results in Table 7 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 term occurs 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 could spend with an individual child was an important constraint on the parent's ability to motivate that child.

C. Robustness: Alternative Explanations for Productivity Interaction

This subsection evaluates several alternative explanations for the interaction between the productivity index and the toy treatments. In this subsection, I consider heterogeneity in preferences for toys and the parent's value of education. Heterogeneity in parental altruism is discussed in the online Appendix 7. I find no evidence that the results from Section V are driven by any of these three effects. I note, however, that the experiment was not optimally designed to fully test these explanations and as such I cannot definitively rule them out.

First, I consider whether the index could reflect heterogeneous preferences for child rewards across households.[36] Suppose that households maximize a single util-

36. Different preferences could be the result of different discounting behavior on the part of children. Children with higher discount rates may prefer toys while children with lower discount rates may prefer a reward to their parent, some of which may be reinvested in the child.

ity function over child rewards (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 money will be more effective at higher values of the index, as found in Section VA. Note, however, that heterogeneity in preferences implies that toys will never be more effective than money; rather, money will be more effective than toys for some households. Contrary to this prediction, the results in Section V show that toys were 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 for toys by correlating the productivity index with several proxies for these preferences. Panel A of Table 8 presents regressions of these proxies on the productivity index, as generated from Column 2 of Table 6.[37] 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 uses a dummy that equals one 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.

Second, I consider whether the interactions between the toy treatments and the productivity index could be driven by heterogeneity in the parent's value of education.[38] Parents with higher values of education may provide more transfers to their children, and therefore the toy reward would be less of a constraint for these households. If the productivity index reflects the value of education, then it is possible that for low values of the index the toy treatments were more effective because they presented more of a constraint on child consumption. However, this implies that the treatments would be equivalent for households with a high value of education. On the contrary, the analysis in Section V shows that the money treatments were 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. Panel B of Table 8 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 6. The transfers examined include money, toys, sweets, clothes, and school supplies. I find no significant relationship between the productivity index and any of these measures. Thus, this evidence is not consistent with the hypothesis that the productivity index reflects the parent's value of education.

37. Repeating the regressions in Table 8 using the productivity index generated from Column 1 of Table 6 produces similar results.

38. In the model, the future value of education is discounted to the present. Thus, the value of success could reflect differences in time discounting on the part of parents.

Table 8
Correlates of the Productivity Index

	Dependent Variable			
	Parent Chose Toy			Child Chose Toy
	1	2	3	4
	Panel A: Productivity Index and Choices			
Productivity	0.114	0.126	–0.115	–0.033
	(0.131)	(0.128)	(0.113)	(0.033)
Sample	Both choice treatments	Ex ante treatment	Ex post treatment	All treatments
Observations	231	153	78	907
R-squared	0.001	0.002	0.001	0.001
Mean dependent variable	0.390	0.327	0.513	0.032

	Dependent Variable					
	Average Transfers 1	Money Given 2	Gave Toys 3	Gave Sweets 4	Gave Clothes 5	Gave School Supplies 6
	Panel B: Productivity Index and Baseline Transfers					
Productivity	–0.135	4.363	0.021	–0.117	–0.034	–0.064
	(0.109)	(3.655)	(0.042)	(0.089)	(0.028)	(0.048)
Observations	912	925	923	923	913	913
R-squared	0.002	0.001	0.000	0.002	0.002	0.002
Mean dependent variable	–0.004	13.065	0.037	0.190	0.015	0.064

Notes: The productivity index is constructed as the predicted values of the regression in Column 2 of Table 6. The dependent variable in Panel A, Columns 1–3 is a dummy that equals one if the parent chose the toy in the treatment(s) indicated. The dependent variable in Panel A, Column 4 is a dummy that equals one if the child indicated that he would buy a toy if he had 100 rupees. In Panel B, the sample includes all households surveyed at baseline. “Average Transfers” is an average of the z-scores of the five individual transfer categories. See Table 2 notes for definitions of the dependent variables in Panel B. Standard errors are constructed based on 500 bootstrap draws, sampling by classroom. * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent.

VI. Conclusion

In this paper, I present the results of a field experiment in Gurgaon, India, designed to test the effects of varying the recipient and form of incentives to learn. The experiment offered incentives of either cash or toys to 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 afterschool 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 toy treatments. These results show that while the recipient of a cash incentive may not influence the effectiveness of incentives, in-kind incentives can have important distributional impacts on outcomes relative to cash. In particular, providing in-kind incentives that target children can reduce inequality in outcomes compared with cash incentives.

Through adapting the context to a model of education production, this paper shows that heterogeneity in treatment effects is, in part, driven by differences in 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; when parents are less productive, it is more effective to reward the children. Future research could build upon these results by experimentally increasing parental productivity and examining how this manipulation interacts with targeted incentives.

While this paper has focused on children early in the education process, the decision process may change 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. Future research examining how the effects of targeted transfers change as children progress through school would thereby inform how the education production process changes over time.

References

Angrist, Joshua, Daniel Lang, and Philip Oreopoulos. 2009. "Incentives and Services for College Achievement: Evidence from a Randomized Trial." *American Economic Journal: Applied Economics* 1(1):136–63.

Angrist, Joshua, and Victor Lavy. 2009. "The Effects of High Stakes High School Achievement Awards: Evidence from a Group Randomized Trial." *American Economic Review* 99:1384–414.

ASER Centre. 2014. *Annual Status of Education Report (Rural) 2013*. New Delhi: ASER Centre.

Baird, Sarah, Jacobus de Hoop, and Berk Ozler. 2013. "Income Shocks and Adolescent Mental Health." *Journal of Human Resources* 48(2):370–402.

Baird, Sarah, Crag McIntosh, and Berk Ozler. 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment." *Quarterly Journal of Economics* 126(4):1709–53.

Banerjee, Abhijit, Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *Quarterly Journal of Economics* 122(3):1235–64.

Becker, Gary. 1974. "A Theory of Social Interactions." *Journal of Political Economy* 82(6):1063– 93.

Bergstrom, Theodore. 1989. "A Fresh Look at the Rotten Kid Theorem – and Other Household Mysteries." *Journal of Political Economy* 97(5):1138–59.

Bettinger, Eric. 2010. "Paying to Learn: The Effect of Financial Incentives on Elementary School Test Scores." NBER Working Paper 1633.

Blimpo, Moussa. 2010. "Team Incentives for Education in Developing Countries: A Randomized Field Experiment in Benin." Stanford University. Unpublished.

Bursztyn, Leonardo, and Lucas Coffman. 2012. “The Schooling Decision: Family Preferences, Intergenerational Conflict, and Moral Hazard in the Brazilian Favelas.” *Journal of Political Economy* 3:359–97.

Duflo, Esther. 2003. “Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa.” *World Bank Economic Review* 17(1):1–25.

Eswaran, Mukesh, and Ashok Kotwal. 1985. “A Theory of Contractual Structure in Agriculture.” *American Economic Review* 75(3):352–67.

Fryer, Roland. 2011. “Financial Incentives and Student Achievement: Evidence from Randomized Trials.” *Quarterly Journal of Economics* 126:1755–98.

Gatti, Roberta. 2005. “Family Altruism and Incentives.” *Scandinavian Journal of Economics* 107(1):67–81.

Haddad, Lawrence, John Hoddinott, and Harold Alderman. 1997. “Policy Issues and Intrahousehold Resource Allocation: Conclusions.” In *Intrahousehold Resource Allocation in Developing Countries: Models, Methods and Policy*, ed. Lawrence Haddad, John Hoddinott, and Harold Alderman, 275–92. Baltimore: Johns Hopkins University Press.

Haveman, Robert, and Barbara Wolfe. 1995. “The Determinants of Children’s Attainments: A Review of Methods and Findings.” *Journal of Economic Literature* 33(4):1829–78.

Jackson, C. Kirabo. 2010. “A Little Now for a Lot Later: A Look at the Texas Advanced Placement Incentive Program.” *Journal of Human Resources* 45(3):591–639.

Jacoby, Hanan G. 2002. “Is There an Intrahousehold ‘Flypaper Effect’? Evidence from a School Feeding Programme.” *Economic Journal* 112(476):196–221.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. “Incentives to Learn.” *Review of Economics and Statistics* 91(3):437–56.

Lee, David. 2009. “Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects.” *The Review of Economic Studies* 76(3):1071–102.

Leibowitz, Arleen. 1974. “Home Investments in Children.” *Journal of Political Economy* 82(2):S111–31.

Levitt, Steven, John List, Susanne Neckermann, and Sally Sadoff. 2012. “The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance.” NBER Working Paper 18165.

Lundberg, Shelly, Robert Pollak, and Terence Wales. 1997. “Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit.” *Journal of Human Resources* 32(3):463–80.

Pratham. 2014. “Pratham Learning Support Classes.” http://www.pratham.org/S-10-3-Learning Support-Classes.aspx. Accessed May 1, 2014.

Rawlings, Laura, and Gloria Rubio. 2005. “Evaluating the Impact of Conditional Cash Transfer Programs.” *World Bank Research Observer* 20(1):29–55.

Shastry, Gauri Kartini, and Leigh Linden. 2012. “Grain Inflation: Identifying Agent Discretion in Response to a Conditional School Nutrition Program.” *Journal of Development Economics* 99(1):128–38.

Todd, Petra, and Kenneth Wolpin. 2007. “The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps.” *Journal of Human Capital* 1(1):91–136.

Weinberg, Bruce. 2001. “An Incentive Model of the Effect of Parental Income on Children.” *Journal of Political Economy* 109(2):266–80.

UNESCO. 2010. *Education for All Global Monitoring Report, 2010*. Oxford: Oxford University Press.