Information, Non-Financial Incentives, and Student Achievement: Evidence from a Text Messaging Experiment*

Roland G. Fryer, Jr.

Harvard University and NBER

October 2016

Abstract

This paper describes a field experiment in Oklahoma City Public Schools in which students were provided with free cellular phones and daily information about the link between human capital and future outcomes via text message in one treatment and minutes to talk and text as an incentive in a second treatment. Students' reported beliefs about the relationship between education and outcomes were influenced by the information treatment. However, there were no measurable changes in student effort, attendance, suspensions, or state test scores, though there is evidence that scores on college entrance exams four years later increased. The patterns in the data appear most consistent with a model in which students have present-bias or lack knowledge of the educational production function, though other explanations are possible.

* Special thanks to Karl Springer, superintendent of Oklahoma City Public Schools, for his support and leadership during this experiment. I am grateful to my colleagues Lawrence Katz, Andrei Shleifer, and Robert Jensen for helpful comments and suggestions. I would also like to thank the Editor and two anonymous referees for extraordinarily detailed feedback. Brad Allan, Matt Davis, Blake Heller, William Murdock, and Namrata Narain provided excellent research assistance and project management support. Financial and in-kind support from the Sandridge Foundation, Droga5, and TracFone Wireless Inc. is gratefully acknowledged. Correspondence can be addressed to the author by email at rfryer@fas.harvard.edu. The usual caveat applies.

In an effort to increase student achievement, a wide variety of innovative reforms have been put forth by school districts across the United States and by not-for-profits across the globe. One particularly cost-effective strategy, not yet tested in American public schools, is providing frequent information about the link between human capital and future outcomes.[1]

Theoretically, providing such information could have ambiguous effects. First, as Wilson (1987) argues, if students lack accurate information on the returns to education, and their expectations are lower than the true returns to human capital, then providing information could motivate students to increase effort and spur achievement.[2] Second, if students are more optimistic than historical returns suggest they should be – as Smith and Powell (1990), Avery and Kane (2004), and Rouse (2004) argue – providing information could lead to reduced effort and thereby decreases in achievement. Third, providing information will likely have no effect on effort or achievement if students do not know the education production function, heavily discount the future, have other self-control constraints, or already hold accurate beliefs about returns to schooling (Mickelson 1990, Fryer 2011b).

In addition, the effect of providing students with information about the link between human capital and future outcomes could interact with the provision of incentives to increase one's human capital. Fryer (2011b) finds no effect of providing financial incentives alone on average student achievement, but as Fryer, Devi, and Holden (2016) note, both providing incentives for multiple tasks for a given agent or for multiple agents for a given task substantially increase achievement. The incentives here could complement the provided information by reducing the perceived cost of investing in human capital in the present. Input incentives (e.g. rewards for reading books) rather than output incentives (e.g. rewards for increasing test scores)

[1] Some informational programs have been attempted in the United States to motivate students by providing accurate information on the returns to schooling or "rebranding" achievement. Since 1972, The United Negro College Fund has run a series of Public Service Announcements (PSAs) promoting education among low-income students with the "A Mind is a Terrible Thing to Waste" campaign. Since 2000, with the launch of "Operation Graduation," the U.S. Army has sponsored Ad Council media campaigns to encourage students to stay in school. Their most recent collaboration, "Boost Up", follows the lead of non-profit organizations like the Gates Foundation, using interactive web sites and online videos in addition to traditional visual and print media to engage youth and promote academic achievement among vulnerable populations. On the other hand, the Gates Foundation's "Get Schooled" campaign utilizes the influence of celebrities, partnering with MTV, DefJam, and others to generate excitement around school improvement and to implore students to stay in school to reach their potential. While government agencies and not-for-profit organizations continue to invest millions of dollars to engage youth through these informational campaigns and others, no rigorous evaluation of their effect on student learning or other educational outcomes has been attempted in the context of the US.

[2] Neal and Johnson (1996) argue that, if anything, the returns to test scores are higher for blacks than for whites.

may complement the provision of information particularly well if students are unaware of *how* to increase their human capital.

In the 2010-2011 school year, we conducted a randomized field experiment in Oklahoma City Public Schools (three treatment arms containing 490 students each and a control group of 437 students – a total of 1,907 students in sixth and seventh grades) that tested the effect of providing information, incentives, and a combination of the two on various outcome measures.[3] In partnership with the largest pre-paid mobile phone provider in the US and an internationally recognized advertising firm, we launched a campaign titled "The Million," designed to provide accurate information to students about the importance of education for future life outcomes such as unemployment, incarceration, and wages.[4] The key vehicle of the experiment was a cellular telephone, pictured in Appendix Figure 1.

Students in each of the three treatment groups were given cellular phones free of charge, which came pre-loaded with 300 credits that could be used to make calls or send text messages.[5] Students in our main treatment arm – (Treatment 1: Information Only) – received 300 credits per month to use as they wanted and received one text message per day delivered at approximately 6:00 p.m.[6] A second treatment arm – (Treatment 2: Information and Incentives) – provided the same information on the link between human capital and future outcomes, and also included non-financial incentives – credits to talk and text were earned by reading books outside of school. A third treatment – (Treatment 3: Incentives Only) – allowed students to earn credits by reading books and included no information. There is also a pure control group that received neither free cellular phones, information, nor incentives.[7]

[3] Throughout the text, I depart from custom by using the terms "we," "our," and so on. While this is a sole-authored work, it took a team of dedicated project and finance managers to implement the experiment. Using "I" seems disingenuous.
[4] Given the complexities involved in the field experiment, an operational pilot program was conducted in seven public schools in New York City in the Spring of 2008.
[5] One credit is equivalent to one minute of talking, one text message, or 20 seconds of time on the internet. Anecdotally, we heard many reports that the pre-loaded credit was used within days.
[6] When to send the text messages was an important experimental design question, for which theory provided little guidance. We chose 6 p.m. because it was likely after students' extracurricular activities, but before dinner and bed time. We chose not to send messages in the morning because the corresponding appropriate time window was less obvious.
[7] The inclusion of the second treatment, which included both information and incentives, was to understand whether there might be important complementarities between the two. Yet, if the interaction were positive, it would be impossible to tell if this was due to complementarities or the inclusion of incentives. The third treatment was designed to disentangle these effects. We opted for a "pure" control group because data from student focus groups in Oklahoma City revealed that forty percent of middle school students in Oklahoma City already used a cellular

A total of 183 text messages were sent over the duration of the experiment. No message was ever repeated.[8] Given the character limit of text messaging and the age group of our subjects, we did not distinguish between evidence that was drawn from correlations and evidence gleaned from quasi- or field experiments.

For the direct outcomes of students in the two informational treatments, we examine students' ability to answer specific questions about the relationship between human capital and outcomes, such as income and incarceration, whose answers were sent to the treatment students in text messages during the year. Treatment effects are generally positive. For example, using the raw data, students who were randomly assigned to the information and incentives treatment were 3.47 percentage points more likely to correctly answer the question on the income differences between college graduates and dropouts than students in the control. Similarly, they were also 16.82 percentage points more likely to know the correct answer for differences in incarceration rates between high school graduates and dropouts than the control group average.

As a robustness test, we included a "placebo" question on the unemployment rate of college graduates, about which students never received information. The difference in the probability of answering this question correctly between informational treatments and the control group was trivial and statistically insignificant. Moreover, 54 percent of control students believe that incarceration rates for high school graduates and dropouts are "no differen[t]" or "really close," suggesting that students in Oklahoma Public Schools do not have accurate knowledge of the returns to schooling.

For indirect outcomes, such as state test scores, attendance, and an index of self-reported effort, we estimate reasonably precise zeros. ITT estimates of the effect of treatment on an index of effort, composed from self-reported answers to our survey administered to all students in our experimental sample, are statistically insignificant across all treatment arms. More precisely, the treatment effect on the effort index is 0.009 (0.048) standard deviations (hereafter, σ) for the information only treatment; -0.016σ (0.046) for individuals who received both information and incentives; and -0.033σ (0.046) for those who received only incentives.

---

phone. At roughly $400 per phone, we did not believe the marginal benefits of simply providing a phone to the control group outweighed the marginal costs.

[8] Typical messages read: "High school dropouts are 3 to 4 times more likely to go to prison than high school graduates;" "When blacks and whites have the same education, they make the same money per hour"; and "If you need help, it's okay to ask your parents."

On administrative outcomes – measured at the end of the experimental year – math and English Language Arts (ELA) test scores, student attendance, and behavioral incidence – there is no evidence that any treatment had a statistically significant impact, though due to imprecise estimates we cannot rule out small to moderate effects that might have a positive return on investment.

Four years after the end of the experiment, we collected data on whether students in treatment and control took the ACT college entrance exam and, for those who did, their scores. The impact of the information only treatment on whether or not a student takes the ACT is -0.03 (0.030). All other treatments had a similar effect. None of these coefficients are statistically different from students in the control group or each other.

There are some interesting results using the scores of the ACT tests, conditional upon taking the exam in OKCPS. The impact of the information only treatment on a student's maximum ACT composite score – which is the average of the four subject test scores – is 0.13σ (0.064). Students in the incentives only treatment scored 0.10σ (0.060) higher and students who received both incentives and information scored 0.09σ (0.062) higher. The latter two coefficients are not statistically different from zero. Perhaps the most compelling part of these results is the fact that the composite scores are driven by relatively large gains on the English portion of the ACT. The impact of the information only treatment on English ACT scores is 0.18σ (0.072). The two other treatments yield almost identical results.

We demonstrate that our four findings – providing students with information on the returns to schooling changes their beliefs, does not alter measureable effort, has no impact on state test scores or attendance, but, four years after the experiment concluded, has statistically significant impact on college entrance exams – is robust to sample attrition and bounding, as well as to adjusting the standard errors on the treatment effects to account for the family-wise error rate.

The paper concludes with a simple model of human capital investment. In the model, exerting effort in school incurs costs, but yields long-term benefits that increase with the perceived returns to education. This yields simple equilibrium conditions from which we derive comparative statics. The magnitude of our identified treatment effects depends on two features of the model: the responsiveness of effort to the change in beliefs, and the shape of the production technology around the pre-treatment equilibrium. We use this setup to frame our empirical results and explain why, in our experimental sample, beliefs changed, but immediate effort and academic achievement did not, while there was evidence of improvement on long-run academic outcomes such as college entrance exams.

We provide speculative evidence that the data is most consistent with either a model in which students do not know the education production function, have high discount rates – as is evidenced in adolescents (see Mischel 2014 for a nice review) – or a combination of the two. A lack of knowledge of the particulars of the education production function is appealing because it may also reconcile our results from those gleaned in developing countries. In stark contrast to our results, Jensen (2010) and Nguyen (2013) report significant treatment effects on educational attainment and achievement from implementing informational experiments in the Dominican Republic and Madagascar, respectively. In our framework, higher costs of effort lead to higher marginal productivity in equilibrium, following directly from the first-order conditions. If the effort costs of investing in education are higher in less developed countries, then under certain conditions investment will be more sensitive to changes in the perceived return to education as compared to a developed country like the US, where most students attend school and the cost of schooling is generally low. The key idea is that there is more "low-hanging fruit" for students in developing countries.

Other explanations, such as complementarities in production, all seem to contradict the data in important ways, but our ability to test them is limited. We find no evidence of complementarities between information and teacher quality, neighborhood quality (measured by poverty rates), or residential segregation.

Our experiment has several important caveats. First, our results are only from one medium-sized school district in Oklahoma. Second, we did not assess individual student beliefs on the returns to human capital, across any dimension, before the experiment commenced. Third, our ability to pin down the precise mechanism generating the results is quite limited by data.

The next section provides a brief review of the literature on how much students know about returns to schooling. Section III describes details of our field experiment that aimed at providing accurate information regarding the link between education and future outcomes to students in the informational arms of the treatment. Section IV outlines our research design and details the data used in our analysis. The main statistical results are presented in Section V. Section VI attempts to reconcile our results and the data gleaned from similar experiments in developing countries with a range of potential theories. The final section concludes. There are two online appendices. Online Appendix A is an implementation supplement that provides details on the timing of our experimental roll-out and critical milestones reached. Online Appendix B is a data appendix that provides details on how we construct our covariates, our samples from the school district administrative files, and the survey data used in our analysis.

II. A Brief Review of Related Literature

A growing body of research examines student perceptions of the value of education in the US and abroad (Dominitz and Manski 1996, Avery and Kane 2004, Rouse 2004, Harris 2008, Kaufmann 2009, Attanasio and Kaufmann 2009), as well as the effects of informational treatments on educational outcomes (Beyer et al. 2015, Jensen 2010, Nguyen 2013, Wiswall and Zafar 2014). Below, we describe each of these literatures in turn.

Survey Data on Attitudes and Beliefs about Returns to Schooling

The anthropology and sociology literatures are divided on whether and the extent to which minority or low-income students know the link between educational achievement and future outcomes. Ogbu (1978) and Lieberson (1980) suggest that the historically discriminatory job ceiling has led educated members of the black community to provide negative feedback to their communities regarding returns to education. They hypothesize that this causes black students and their parents to lower their expectations about the returns to educational attainment and question its instrumental value. In contrast, using data from the 1990 National Education Longitudinal Study (NELS), Ainsworth-Darnell and Downey (1998) question Ogbu's (1978) oppositional culture explanation, reporting that black students are more likely than their white peers to report that education is important to getting a job later on.[9]

Economists have also documented similarities in the perceived expected costs and benefits of education across racial and income groups. Surveying a group of low-income, mainly

[9] To explain why blacks report more optimistic beliefs about returns to human capital investment than their white counterparts, Mickelson (1990) distinguishes between "abstract" and "concrete" attitudes toward education. "Abstract" attitudes are defined as a respondent's expressed beliefs about the general value of education in society. "Concrete" attitudes relate to a respondent's expressed beliefs about the value of education and barriers to enjoying its full value for themselves, personally. Consistent with Ainsworth-Darnell and Downey's (1998) analysis, Mickelson (1990) notes that in survey results, black respondents have "abstract" attitudes toward education that are similar to that of their white peers, but relatively less-positive "concrete" attitudes, which are rooted in life experience.

minority youth in Boston and a group of relatively affluent, white suburban students from a nearby suburb, Avery and Kane (2004) find striking similarities between the perceived costs and payoffs from attending college among members of these two groups. Similarly, Rouse (2004) finds little evidence of differential expected returns to education between racial or socioeconomic groups, but notes that high expectations in the low-income group are not as strongly correlated with actual college enrollment as in the higher-income group.

Field Experiments in Developing Countries

The papers most closely related to the current project are from field experiments conducted in the Dominican Republic (Jensen 2010) and Madagascar (Nguyen 2013).[10] Jensen (2010) considers the role that the perceived returns to education play in students' schooling choices. Jensen (2010) demonstrates that the eighth grade boys in his sample dramatically underestimate measured returns to education. While the mean earnings of Dominicans who finish secondary school are 40% higher than those who do not, the typical student perceives that his earnings will increase by only 9.2% if he completes secondary school. More importantly, a random subset of students who received information on the real returns to education enrolled only in an additional 0.20 – 0.35 years of high school, on average.[11]

Nguyen (2013) also shows that providing information about returns to education to parents and students can have a positive impact on academic outcomes, especially when parents underestimate the value of schooling. Teachers in 80 randomly selected treatment schools presented parents and students with information about the distribution of jobs and the expected earnings of 25 year-old males and females in Madagascar by educational attainment level. Nguyen (2013) finds that providing accurate statistics on the value of additional schooling to parents and students in Madagascar raised test scores by 0.202σ (0.106) and improved attendance

[10] See also Wiswall and Zafar (2014), who inform college students in the US of the true income distribution by college major/degree status. The authors find that this influences their beliefs about future earnings and intended major, but they do not observe whether students actually change their behavior (e.g. switching majors).
[11] It is unclear whether the treatment in Jensen (2010) or the current approach is "stronger." Treated students in Jensen's sample were read a single paragraph that cited the average salary earned by Dominican men with a primary education, a high school education, and a college education. Our treatment provided daily messages over the school year on a wider variety of returns (i.e. incarceration, unemployment, etc.) While it is possible that delivering the message in person results in a larger change in beliefs, Karlan et al. (2010) show that text messages can lead to measurable changes in behavior in a different setting. Similarly, York and Loeb (2014) find that using text messages to communicate with parents of preschoolers improves their children's early literacy development.

by 3.5 percentage points. Test scores increased by 0.365σ (0.156) among those who underestimated the returns to education during a baseline survey.

Our paper makes three contributions to the current literature. Perhaps most importantly, we conduct the first field experiment aimed at exploring the role of information on student achievement in the US – where the survey evidence is ambivalent as to whether minorities know the true returns to human capital. Second, our text message technology potentially improves upon the previous literature. While past efforts have relied upon pamphlets or one-time conferences to distribute information, mobile technology allowed us to provide a multi-faceted stream of information directly to students over the course of an entire school year.[12] Third, we inform students of a variety of outcomes that are correlated with educational attainment and achievement – unemployment, probability of incarceration, life expectancy – rather than concentrating solely on labor market returns. Conceptually, this may provide even more impetus to invest in human capital.

III. Field Experiment Details

Oklahoma City Public Schools (OKCPS) is a typical medium-sized urban school district serving 42,567 students in eighty-nine schools. Seventy-seven percent of OKCPS students are black, Hispanic, or Native American. Roughly 85 percent of all students are eligible for free or reduced-price lunch and twenty-eight percent of the students are English language learners. There is a large racial achievement gap in OKCPS by 6th grade; within the twenty-two experimental schools, black and Hispanic students' 2009-2010 test scores are 0.404σ (0.042) and 0.317σ (0.044) behind their white peers in reading and math respectively, controlling for socioeconomic status, free lunch eligibility, English Language Learner status, Special Education status, and gender. This is consistent with overall national trends (Fryer 2011a, Jencks and Phillips 1998).

A. Description of Treatment

Table 1 provides a bird's eye view of the experiment. First, we – together with local philanthropists, TracFone (the mobile device provider), and Droga5 (an internationally-

[12] Karlan et al. (2010) use text message reminders to promote and incentivize monthly savings among bank customers in Peru and Bolivia. They find that reminders coupled with incentives based upon account interest rates increased the amount saved and the likelihood of reaching a savings goal.

recognized advertising agency) – garnered support from the district superintendent. Following the superintendent's approval, we held an information session for the principals and instructional leaders of all twenty-two schools in the district with sixth and/or seventh grade students that were not designated "alternative education academies" to provide an overview of the proposed experiment. All twenty-two eligible schools signed up to participate. At the end of September 2010, information packets (containing a letter about the program to families and a parent consent form) were distributed to principals and library media specialists (LMS) from the twenty-two elementary and middle schools. The LMS had been jointly determined to act as school-based coordinators to help oversee implementation for a small stipend that was not tied to performance.

Sixth and seventh grade students attending the twenty-two elementary and middle schools in OKCPS who signed up for the program were eligible to participate.[13] Students received information packets on September 28, 2010 and were required to return a signed consent form by October 1, 2010 in order to be eligible for the lottery that determined participation. We received 1,907 student consent forms (out of a possible 4,810) and randomized students into one of four groups: (Treatment 1) 490 students received a cell phone (pre-loaded with 300 minutes) with daily informational text messages and a fixed allocation (non-performance-based) of 200 credits on a monthly schedule; (Treatment 2) 490 students received a cell phone (pre-loaded with 300 minutes) with daily informational text messages, and were also required to read books and complete quizzes to confirm their understanding of those books in order to receive additional credits; (Treatment 3) 490 students received a cell phone (pre-loaded with 300 minutes) and were required to read books and complete quizzes about those books in order to receive additional credits on a biweekly schedule, but did not receive any informational text messages; and (Control) 437 students did not receive a phone, informational messages, or non-financial incentives.[14] Sending one outgoing text message or talking on the phone for one minute or a fraction of a minute deducted one credit from the student's balance. Incoming text messages were free of charge.

[13] We chose sixth and seventh grades because they were old enough to have a cellular phone, but only 39% of students in OKC had them. This number is almost double in urban centers such as New York City (where we conducted the operational pilot), which makes OKC an ideal location on this dimension.
[14] Students completed the reading quizzes through an online portal and schools provided students time to do this at school (e.g. before school, after school, study hall, lunch, time in the library/computer lab, etc.). However, students could access the website and complete the quizzes from home if they wished.

Phones were distributed to each of the twenty-two schools on the morning of October 8, 2010. Students in treatments (2) and (3) were eligible to earn credits by reading books. Upon finishing a book, each student took an *Accelerated Reader* (AR) computer-based comprehension quiz, which provided evidence as to whether the student had read the book. Each book in the AR is assigned a point value based on length and difficulty. Students were allowed to select and read books of their choice and at their leisure, not as a classroom assignment. The books came from the existing stock available at their school (in the library or in the classroom), though additional copies of books that proved to be particularly popular were ordered during the year. This is almost identical to the reading incentive program described in Fryer (2011b).

For those students required to read books in order to receive additional credits, the incentive scheme was strictly linear: each point earned during each biweekly reward period translated to ten credits which could be used to talk or text. Because credits could only be distributed (i.e. uploaded electronically) in increments of 200, point earnings in excess of a multiple of 20 were banked and carried over to subsequent reward periods. Once a student reached or passed any 20-point interval, blocks of 200 credits were uploaded at the next scheduled "payday" according to the predetermined biweekly reward schedule.

Text messages were sent to students in the appropriate treatment groups on a daily basis, including weekends, at approximately 6:00 p.m. We worked closely with Droga5, an advertising firm based in New York City, to determine the messaging and branding components of the program. We met initially to discuss the types of text messages that would be written and sent to students on a daily basis. Writing text messages throughout the year was a collaborative and iterative process. Drawing upon advertising research suggesting that consumers respond to both informative and persuasive messages, (Nelson 1974, Mullainathan, Schwartzstein, and Schleifer 2008, Shapiro 2006) and recognizing our comparative advantages, Droga5 created the persuasive messages while we created the informative messages based on information from the Bureau of Labor Statistics, the National Center for Education Statistics, the Census Bureau, and other sources.[15] Project teams met monthly to finalize upcoming text messages. Approximately 25% of

[15] Examples of informational texts include "Each year, H.S. dropouts make $21,023. College graduates make $58,613. Do the math" (United States Census Bureau 2011) and "High school dropouts are more than three times as likely to be unemployed as college graduates" (Bureau of Labor Statistics 2011). Persuasive examples include "People don't look down on someone for being too educated" and "Graduates never regret staying in school, but dropouts often regret leaving it."

the sent messages were informational and 75% were designed to be persuasive. Approved messages were sent to TracFone for distribution to students in Treatments 1 and 2.

Implementation Monitoring

Implementation of experimental protocols was monitored along several dimensions. First, we visited each school and project managers reviewed the basics of the program with treatment students to reinforce their understanding of the program details. To diagnose specific misunderstandings of the reward algorithm or distribution system, brief quizzes were administered to check for student understanding, covering topics including the incentive structure, reward schedule, and how to report phone problems. After the first three months of implementation, students answered 79% of quiz questions correctly. Second, administrative access to the AR program enabled us to follow usage of the software on a daily basis for students in the incentive treatments and produce and deliver program-, school-, and student-level dashboards weekly. Third, every month, project managers conducted site visits to schools.

By the end of the experiment, 77 percent of students who received a phone and were required to earn AR points in order to receive credits had earned at least a fraction of a point.[16] Twelve of the twenty-two schools had a rate of participation of at least 90 percent. The largest and second largest schools (in terms of number of students with cell phones in incentivized treatment groups) had participation rates of 65 percent and 75 percent, respectively. The average number of AR points earned in the information and incentives treatment was 35.25 over the year. For the incentives only treatment, it was 36.17. Overall, ten percent of the sample earned more than 100 points, and the top one percent of the sample read earned more than 350 points. To get a sense of the magnitude, this is equivalent to reading seventy Dr. Seuss level books, three books such as “Harry Potter and the Philosopher’s Stone”, or one “Moby-Dick or The Whale.”

In total, incentive and hardware costs were $230,365 for the program with 1,470 subjects in treatment. Administrative costs were approximately $139,000, which includes the AR registration fees, software installation, and a district-based program manager. Total cost of implementation was approximately $369,365 – or $251.27 per student (this does not include potentially billable hours of the advertising firm.)

[16] This figure includes the approximately 11 percent of students who exited the experiment during the year for a variety of reasons: lost phone, moved out of district, etc.

IV. Data, Research Design, and Econometrics

A. Data

We collected administrative data from all schools in OKCPS and survey data from students in the experimental group in each school. We begin with an overview of the administrative data.

Administrative Data

The administrative data includes first and last name, birth date, race, gender, free lunch eligibility, record of behavioral incidents, daily attendance, matriculation with course grades, special education status (SES), English language learner (ELL) status, and Oklahoma Core Curriculum Criterion Referenced Test (CRT) assessment data for math and the English language arts (ELA). We use administrative data from 2008-09 and 2009-10 (pre-treatment years) to construct baseline controls, and for 2010-11 (post-treatment year) for outcome measures.

For ease of interpretation, test scores have been normalized to have a mean of zero and a standard deviation of one within grades and subjects (across all schools) for 2010-2011 scores, when they are used as outcomes in our analysis. We do not report testing results for the 7% of students who take the Oklahoma Modified Alternative Assessment Program. Pooling the results for the two tests together does not change our findings. Students took the CRT in mid-April of 2011, approximately seven months after the start of the treatment.

Individual attendance rates account for all presences and absences for each student, regardless of which school the student had enrolled in when the absence occurred, as long as the student was enrolled in OKCPS. The attendance rate is calculated by dividing the number of days present by the number of days a student was enrolled in the district during the 2010-2011 school year.[17] Attendance rates have also been normalized to have a mean of zero and a standard deviation of one within grades, across all of OKCPS.

[17] Oklahoma law requires that absences be recorded daily for both the morning and afternoon portions of the school day. If a student misses more than one hour of school in the morning, he incurs a half-day's absence. If he also misses more than one hour of school in the afternoon, he is marked as absent for the day.

Behavioral incidents are recorded individually by date of infraction. Our measure of behavior is the total number of suspensions each student incurs during the year, regardless of the length of the suspension or the nature of the infraction. Using the total number of recorded infractions yields identical results.

We use a parsimonious set of controls to aid in precision.[18] The most important controls are reading and math achievement scores from the previous two years, as well as their squares and cubes, which we include in all regressions. Previous years' test scores are available for most students who were in the district in the previous year (See Table 2 for exact percentages of experimental group students with missing test scores from the previous year). We also include a set of indicator variables that take a value of one if a student is missing a given test score from the previous year and zero otherwise.

Other individual-level controls include a mutually exclusive and collectively exhaustive set of race dummies extracted from each school's district administrative files, indicators for free lunch eligibility, special education status (SES), and whether a student is an English Language Learner (ELL).[19] Special education and ELL status are determined by the OKCPS Special Services office and the OKCPS Language and Cultural Services Office, respectively.

Survey Data

To supplement each district's administrative data, we administered a survey, post-treatment, to all students in the experimental group in each school. In total, 66 percent of student surveys were completed and returned in experimental schools; 61 percent of control students and 68 percent of treatment students completed and returned a survey.[20] We consider the possible implications of differential attrition for our results in Section VI.

---

[18] Excluding all controls does not alter the qualitative conclusions of the paper.

[19] A student is income-eligible for free lunch if her family income is below 130 percent of the federal poverty guidelines, or categorically eligible if (1) the student's household receives assistance under the Food Stamp Program, the Food Distribution Program on Indian Reservations (FDPIR), or the Temporary Assistance for Needy Families Program (TANF); (2) the student was enrolled in Head Start on the basis of meeting that program's low-income criteria; (3) the student is homeless; (4) the student is a migrant child; or (5) the student is identified by the local educational liaison as a runaway child receiving assistance from a program under the Runaway Youth and Home Youth Act.

[20] More specifically, 70 percent of students in the information only treatment, 69 percent in the information and non-financial incentives treatment, and 65 percent in the non-financial incentives only treatment completed and returned student surveys.

The data from the student survey includes questions about student motivations for entering the experiment, phone use, phone problems and troubleshooting, student perceptions of school-wide impact, and a series of questions designed to assess various dimensions of student effort. In addition, the survey included questions that quizzed students on specific facts about the importance of education that were delivered via text messages to students in the informational treatment arms during the year. We asked students “Are high school dropouts more likely to go to prison than high school graduates?”, which referenced the text messages “male high school dropouts go to prison four times more often than men who went to college” and “high school dropouts are 3-4 times more likely to go to prison than high school graduates.” The survey also asked “True or false: college graduates make 54% more money than college dropouts” – a statistic pulled directly from an earlier text message. A third question asked, “Your income as an adult increases by (.) for every year you spend in school,” which referred to the similarly worded text message “Your income goes up by 10% for every year you spend in school.” The last question asked for the unemployment rate among college graduates, a statistic that was not referenced in any text message, and is therefore a placebo question for which we expect zero effect.

Table 2 provides descriptive statistics of all the 6th and 7th grade students in OKCPS, divided (not mutually exclusively) into six columns: students in eligible schools who did not choose to participate in the experiment (column 1); students who opted into the experiment (column 2); students randomly selected into the information only treatment (column 4); students randomly selected into the information and incentives treatment (column 5); students randomly selected into the incentives only treatment (column 6), and a pure control group (column 7). Each column provides the mean and standard deviation for each variable used in our analysis (see Online Appendix B for details on how each variable was constructed).

As students could opt in to the randomization, there are some statistically significant differences between participants and non-participants. Participating students are 3.5 percentage points less likely to be male and also 3.7 percentage points less likely to be white. They are also poorer on average – 91.7% of participating students are eligible for free or reduced price lunch, relative to 85.7% of non-participants – and roughly 10 percentage points more likely to have non-missing baseline testing data.

Within the experimental group, the treatment and control groups are well-balanced, although the control group has more male students ($p = 0.022$). A joint significance test, testing the equality of means across all covariates, yields a p-value of 0.387, suggesting that the randomization is collectively balanced along the observable dimensions we can consider. Moreover, the treatment groups are also quite balanced if one conducts pairwise balance tests, although the joint p-value comparing control and the information only group is 0.047 (see Appendix Table 1).

B. Research Design

There is an active debate as to which randomization procedures have the best properties under different circumstances (e.g. Greevy, Lu, and Silber 2004, Bruhn and McKenzie 2009, Imai, King, and Nall 2009, Imbens 2011, Kasy 2016). In samples with more than 300 units, Bruhn and McKenzie (2009) provide evidence that there is little gain from different methods of randomization over a pure single draw. Consistent with this, we used a pure single random draw to sort the 1,907 students who turned in consent forms into one of the three treatments and control.

C. Econometric Model

To estimate the causal impact of each treatment, we estimate Intent-To-Treat (ITT) effects, i.e. differences between treatment and control group means for each treatment arm. Let Z_i be an indicator for assignment to a given treatment arm that takes a value of one if a student is in that treatment group and a value of zero if a student is in the control group. Let X_i be a vector of baseline covariates measured at the individual level; X_i and a school fixed effect γ_s comprise our set of controls. Given our research design, results with or without controls are virtually identical. Controls are included to aid in precision. All regressions without controls are available from the author by request.

The ITT effect, π, is estimated from the equation below:

$$outcome_{i,s} = \alpha + \beta X_i + \gamma_s + \pi Z_i + \epsilon_{i,s} .$$

Each ITT estimate is an average of the causal effects for students who were randomly selected into a given arm of treatment at the beginning of the year and students who signed up for treatment but were not chosen. In other words, ITT provides an estimate of the impact of being offered a chance to participate in a given arm of the experiment. All student mobility and disruptions in phone service due to theft, loss, or malfunction is ignored.[21] We only include students who were enrolled in OKCPS as of the date of randomization, October 4, 2010. In OKCPS, school began on August 19, 2010, and students in the incentive treatment were eligible to earn credits starting October 11, 2010.

V. Results

In this section, we describe the main results of our experiment across three domains. First, using survey data, we investigate the direct effect of receiving daily text messages about the link between human capital and life outcomes on students' knowledge of similar correlations. We also use this data to measure the heterogeneity of treatment effects for various predetermined subgroups. Second, we examine several short-run indirect outcomes: test scores, behavior, and attendance – collected from the district's administrative files – and survey measures of effort. Third, we investigate longer-term outcomes such as whether or not students take college entrance exams and, if so, their scores. We discuss all three in turn.

A. Direct Outcomes

Knowledge of the Link Between Human Capital and Future Outcomes

Recall, to assess whether students better understood the link between human capital and future outcomes, we asked them questions for which students in the informational treatments had received multiple text messages with the answers throughout the year. We also asked them a "placebo" question designed to test whether treatment students became generally more knowledgeable about returns to education or whether they only retained knowledge about the specific information they were provided. The three questions students in the information treatments were provided information about via text message were: (1) "True or false? College

[21] Roughly 27% of our sample either lost their phone or experienced technical problems that prevented them from receiving text messages for part of the year. Hence, there is some variation in the treatment dosage after random assignment. As a separate specification, we also estimate two-stage least squares models in which we use the treatment assignment to instrument for the percentage of the year in which a student had a working phone. We report only ITT estimates in the text and put 2SLS results in Appendix Table 2.

graduates make 54% more money than college dropouts", (2) "Your income as an adult increases by (.) for every year your spend in school", and (3) "Are high school dropouts more likely to go to prison than high school graduates?" The placebo question was "15.5% of high school dropouts are unemployed. What percentage of college graduates are unemployed?"

Table 3 presents treatment effects on students' ability to correctly identify links between human capital and life outcomes, which are mostly positive for the informational treatment arms. Students in the information only arm of the treatment were 5.4 (standard error = 3.2) percentage points more likely to correctly identify the wage gap between college graduates and college dropouts [control mean = 81.9 percent], 17.4 (4.5) percentage points more likely to correctly identify the relationship between schooling and incarceration [control mean = 45.9 percent], and 0.5 (4.4) percentage points less likely to know about the schooling and income relationship [control mean = 48.2]. They also got right, on average, 0.23 (0.073) more questions than students in the control group [control mean = 1.744]. Students in the information only treatment were no more likely to answer the placebo question correctly, suggesting that the informational treatments did not improve students' general understanding of the linkages between human capital and life outcomes covered. Moreover, as mentioned previously, 54.1 percent of students underestimated the relationship between educational attainment and incarceration, which implies that students in OKCPS do not have accurate information about the returns to schooling.[22]

B. Indirect Outcomes

Survey Outcomes

We gleaned five measures of effort from our survey. Students were asked questions about the impact of the program, such as "Since the Million program started, do you think you are more focused on or excited about doing well in school?"[23] The results on individual questions assessing effort are reported in Appendix Table 6.

[22] Jensen (2010) and Nguyen (2013) also demonstrate a low baseline level of knowledge about the returns to schooling among students and parents. While the measures of the effect of treatment on student knowledge are not directly comparable across studies, they seem similar enough that differences on this margin are unlikely to explain our different findings of the effect of informational treatment on educational achievement.

[23] Other questions were as follows: "How many books did you read during the school year?", "About how much of your math homework do you usually complete, either during school hours or outside of school?", "About how much of your reading homework do you usually complete, either during school hours or outside of school?", and "Which of these is closest to the amount of time you usually spend on homework outside of school each day (Monday-Thursday)? (Options: less than one hour, one-to-two hours, etc.)"

Students in the information only treatment were 16.6 (4.4) percentage points more likely to report feeling more focused or excited about doing well in school, while students in the information plus incentives treatment arm were 13.2(4.3) percentage points more likely to report feeling so compared to the control mean of 43.1 percent. Similarly, students in the incentives only treatment were 15.8 (4.3) percentage points more likely to report feeling more focused or excited about doing well in school.

These are the only positive impacts that the treatment had on measures of effort. All others were statistically zero or negative. For instance, students in the information only treatment group reported reading, on average, 0.722 fewer books than the control mean of 15.50 books. Similarly, students in the information plus incentives and the incentives only arms of the treatment read a statistically significant 1.555 (0.603) and 1.890 (0.622) fewer books than those in the control group, respectively.[24]

We attempt to summarize the impact of our information and incentive treatments using an index measure that combines all five individual effort measures. We standardize each individual measure to have a mean of zero and a standard deviation of one in the control group. We then take the (unweighted) average of each standardized z-score measure. We include all students with at least one non-missing outcome.[25] Using this approach, the impact of the information only treatment on effort is 0.01σ (0.05). These results are displayed in panel A of Table 4. The impact of the other two treatments is nearly identical. There is no evidence that any treatment had a measureable effect on the effort index.

Summarizing, only one of the five questions designed to measure effort – whether the student felt more "focused" or excited about school since the experiment began – has a positive treatment effect. It is unclear how to weigh this question relative to the others. For instance, if it is more correlated with long term outcomes of interest than the others then one might conclude that the experiment increased effort on dimensions that matter most. In the absence of such data that links our questions to outcomes, we prefer a more agnostic approach that simply gives all

[24] In previous work in which we provide incentives for reading books, data constraints made it impossible to estimate the effect of treatment on the number of books read. Note that the negative effects on the number of books read by students in the incentives groups reported in this paper could be the unintended consequence of giving students a cell phone, rather than the effect of the incentives themselves. Another possibility is that students in the treatment group were simply more cognizant of the number of books they read and less likely to over-report on the survey.

[25] Dobbie and Fryer (2015) use an identical approach to create indices which measure the impact of attending the Promise Academy Charter school on a series of outcomes.

the questions equal weight in the index. In this case, the treatment effects on effort are particularly small and statistically indistinguishable from zero.[26]

Administrative Data Outcomes

Panel B of Table 4 presents ITT estimates of the effect of each treatment on state math and ELA standardized test scores, attendance, and student suspensions. Test scores are normalized by grade level and subject to have a mean of zero and a standard deviation (σ) of one within the full OKCPS sample. Again, treatment effects are reported in σ units and standard errors are presented in parentheses below each estimate. Attendance is measured as the proportion of days present in OKCPS divided by days enrolled in the school district in 2010-11 and is then normalized to have a mean of zero and a standard deviation of one. Total suspensions are counted and summed for each student.

Across the three treatment arms, there are no statistically significant treatment effects on any administrative outcomes, though due to imprecise estimates we cannot rule out small to moderate effects which might have a positive return on investment (the experiment was designed to detect 0.15σ effects with eighty percent power). The effect on ELA achievement is 0.070σ (0.046) for the information only treatment and 0.027σ (0.050) for the incentive only treatment. The ITT effects on math achievement are 0.012σ (0.046) and -0.034σ (0.047) for the information only and incentives only treatments, respectively. The ELA and math effects for students in the information plus incentives arm of the treatment are 0.014σ (0.047) and -0.062σ (0.045) respectively. Similar results obtain for attendance and suspensions.

To assess heterogeneity in treatment effects across subgroups of students, Appendix Tables 7A, 7B, and 7C report treatment effects for the three treatment arms on a subset of direct and indirect outcomes for a number of predetermined subgroups. For ease of comparison, the first row of Appendix Tables 7A, 7B, and 7C shows the ITT estimate for the sample for whom we observe the demographic data used to create the subgroups. These estimates are nearly identical to the full-sample estimates in Tables 3 and 4. The final row in each panel reports a p-value on the null hypothesis of equal treatment effects within the panel.

[26] These self-reported measures of effort may have been influenced by mentions of the Million program in the survey instrument, potentially making interpretation more difficult.

There are few consistent patterns. Male students show a much larger increase in the probability of answering both quiz questions correctly in the information only and information plus incentives arms of the treatment. However, the treatment seems to reduce males' math scores across all treatment arms, particularly for the information plus incentives arm of the treatment, where scores reduce by 0.178σ (0.068). Information only students who are not eligible for special education accommodations answer 0.224 (0.075) more quiz questions correctly, while students who are eligible for the accommodations answer 0.22 (50.5) fewer quiz questions correctly. There is very little observable heterogeneity along measures of baseline ability for any of the three treatment groups. Unfortunately, without a measure of students' baseline knowledge about the value of education, we cannot test for heterogeneous effects on this margin.

C. Robustness Checks

Sample Attrition and Bounding

If students selectively exit the sample, then the treatment effects we reported above may be biased. A standard test for attrition bias is to check for differential response rates among treatment and control groups. In Table 5, we regress an indicator for having non-missing data for our main outcome measures on treatment dummies and our full set of controls. While we find no evidence of differential attrition on test score outcomes, students in all three treatments are more likely to provide valid survey data than the control group (to be precise, response rates for the treatment groups are higher by 3.8 (2.4), 7.9 (2.4), and 6.9 (2.4) percentage points in Treatments, 1, 2, and 3 respectively.)

Conceptually, the direction of the potential attrition-induced bias is unclear. If the students in the treatment who gleaned more valuable information are more likely to respond to our survey, then the estimates in Table 3 may be biased upward. If, on the other hand, these students naturally would have absorbed more information – independent of treatment – then our estimates would be too low.

In Table 6, we use calculate Lee (2009) bounds to explore the extent to which differential survey attrition between treatment and control can account for our set of results. Given that we have flat priors on the direction of the bias, we present both upper and lower bounds using the methods described in Lee (2009).

The bounds in Columns (2) and (4) are generated by trimming the sample to equalize response rates between the treatment and control groups. To estimate a lower bound, the sample is trimmed by dropping the fraction of treatment students who have the largest predicted residuals from a regression of the survey outcome of interest on baseline test scores and demographics. Samples for upper bounds are created analogously. We then re-estimate our main ITT specification on the resulting sample.

This exercise confirms the robustness of our results. In the information only treatment, the Lee lower bounds for one coefficient – knowing the wage gap – is no longer statistically significant. The other positive survey estimates all maintain p-values below 0.1. Throughout, none of the attrition-adjusted coefficients are statistically distinguishable from the main ITT results for all direct outcomes, suggesting that differential survey attrition is not an important factor for our results.

A final concern is that our single-comparison tests do not correct for biases introduced by testing multiple hypotheses. The p-values on our main outcomes with positive treatment effects – number of questions answered correctly and ACT scores – are small enough to survive even conservative methods to adjust for multiple comparison bias (i.e. the Holm-Bonferroni method).

D. Longer-Term Outcomes

Four years after the experiment ended, we collected another set of administrative data, including dropout rates, whether or not a student took the ACT in OKCPS, and, if so, their scores on the test. Unfortunately, in the most recent data we were able to obtain (2014-2015 school year), the experimental cohort just finished their tenth or eleventh grade year and measures of dropout behavior at this stage are far less than ideal. The way that OKC Public Schools keep data, the vast majority of kids (over 97%) are considered "still in school," though their historical graduation rate hovers around 72.5%.[27, 28]

We did, however, obtain data on whether or not students took the ACT – and for those who did – their performance. Table 7 reports treatment effects on whether students took the ACT, students' scores on the first time they took the test, and the max scores they ever obtained

[27] With these caveats in mind, there are no differences across any of the treatments and control in drop-out behavior.
[28] The first number is the author's calculation, while the second is the graduation rate for the year 2013-2014 for Oklahoma City Public Schools (http://oklahomawatch.org/2015/07/29/graduation-rates-drop-in-all-categories/.)

on the test. We report results for the total score (i.e. composite score) and for the various subtests.

Interestingly, there are no differences across any of the treatments and control in whether students took the ACT, but there is some evidence of treatment effects on composite ACT scores. The impact of the information only treatment on a student's maximum ACT composite score – which is the average of the four subject test scores – is 0.13σ (0.064). Students in the incentives only treatment scored 0.10σ (0.060) and students who received both incentives and information scored 0.09σ (0.062) higher. The latter two coefficients are not statistically different from zero. Moreover, the coefficients are all statistically identical to each other. The impact of the information only treatment on maximum English ACT scores attained by a student, on average, is 0.167σ (0.069). The two other treatments yield almost identical results.[29]

VI. Discussion and Speculation

The experimental results establish four facts. First, receiving information via text messages causes students to update their beliefs about the returns to education, and their updated beliefs are more "correct." Second, students *report* no discernible increases in effort, according to our index, although the survey question on effort related to the messaging program demonstrated positive treatment effects. Third, there was no measurable increase in educational attainment or achievement when measured at the end of the experiment. Fourth, four years after the experiment, students in the three treatment arms were no more likely to take the college entrance exam than the control group but scored significantly higher on their overall score and in particular on the English portion.

To better understand what mechanisms might lead to these conclusions, we propose a simple two-period model of human capital investment and consider the conditions that might generate these facts. This section is, by necessity, more speculative than our previous analysis.

[29] These results are robust to conservative bounding procedures presented in Appendix Table 8. Students who go on to take the ACT also experience short-term effects similar to the rest of the sample, although students in the information treatment arm who go on to take the ACT score higher on the OK state math test and attend more school, and students in the incentives only treatment arm who go on to take the ACT score higher on the state reading test in 2010-11 (Appendix Table 9).

Consider the problem of a representative student choosing the optimal level of effort *E* to invest in her studies.[30] The production function for academic achievement follows the production function *A*, such that $A=F(E,\boldsymbol{K})$, where $\boldsymbol{K}$ is an *n*-dimensional vector of school, neighborhood, and family "capital" levels that are fixed prior to the student's decision. We impose the following restrictions*:* (a) *F()* is twice continuously differentiable in all inputs, (b) production exhibits diminishing marginal returns to effort – i.e. $\frac{\partial F}{\partial E} > 0$ and $\frac{\partial^2 F}{\partial E^2} < 0$ – and (c) capital and effort are complements – i.e. $\frac{\partial^2 F}{\partial E \partial k_i} > 0$, where k_i is the i^{th} element of the vector $\boldsymbol{K}$.

Academic achievement yields long-term benefits in the forms of higher wages, increased employment opportunities, and other social opportunities. Let *V(A,r)* denote the long-run perceived benefits of achievement, where *r* is a parameter that measures the student's perceived return to achievement. We assume that $\frac{\partial V}{\partial A} > 0$ and $\frac{\partial^2 V}{\partial A^2} < 0$. We also assume that increases in *r* increase payoffs at all levels of *A*: $\frac{\partial V}{\partial r} > 0$.

The student's problem can then be summarized as:

$$\max_{E} \beta V(A,r) - C(E)$$

where *C(E)* is the cost of effort and β is a standard discount factor. Assume that $C'(0) = 0$ and $F'(0, K) > 0$ to ensure an interior solution. The equilibrium level of effort is then defined by the value E^* that solves:

$$\beta \frac{\partial V}{\partial E}(r) = \beta \frac{\partial V}{\partial A}(r) * \frac{\partial A}{\partial E} = C'(E^*).$$

In what follows, we use this simple model to frame a discussion of explanations for our set of facts. In this admittedly limited framework, there are three potential mechanisms to generate a change in beliefs without a change in achievement: discount rates, complementarities in production, and uncertainty about the production function.

A. High Discount Rates

[30] Here we do not differentiate between academic achievement and attainment. This is in part due to empirical necessity, at this time we do not know whether the intervention encouraged students to stay in school longer. As a theoretical matter, the intuition provided in this section still holds so long as students do not substitute academic effort for additional years in school.

The key challenge in interpreting our results is explaining why neither short-run academic achievement nor effort seemingly increased despite the change in perceived returns, while four years after the experiment there is evidence of an increase on test scores.

If the benefits of education occur primarily in the future, then excessive discounting could explain this paradox. In other words, even if the information treatment causes students to foresee additional rewards for investing in their education, the payoff arrives so far in the future that it is not worth expending effort in the current period.[31] In our framework, this is equivalent to having β small enough such that $\frac{\partial E^*}{\partial r}$ is roughly zero. Yet, as time progresses, given students know that academic achievement has important correlates, the incentive to exert effort increases as they get closer to benchmarks like high school graduation or college entry.

Our data cannot reject this hypothesis. High-discount rates are consistent with student achievement remaining flat even after an increase in r and a lack of measurable increase in effort. Yet, recall that treatment students reported being "more focused" across all arms of the treatment. Conversely, our effort index and other (administrative) proxies for effort – such as attendance – show no treatment effects.[32]

Further complicating this theory is the idea that discount rates could also change over time, particularly during adolescence. Steinberg et al. (2009) find that discount rates (measured in a laboratory setting) decrease with age, with a discreet difference between teenagers aged 12-13 and those aged 16-17, when they approach levels more common among adults. This is consistent with our findings given that most students take the ACT in their junior year of high school.

B. Lack of Knowledge of the Production Function

The standard economic model implicitly assumes that students know their production functions – that is, the precise relationship between the vector of inputs and the corresponding output. If students only have a vague idea of how to increase achievement, then there may be

[31] A slightly different interpretation is that students lack self-control – i.e. they recognize that effort will result in large benefits in the future, but cannot commit to studying, going to class, etc. The empirical predictions of this model are identical to the discount-rate explanation.

[32] Whether one believes that present-bias can explain our results depends on our equal weighting of the self-reported measures of effort in surveys. Put differently, it depends on any weighting that does not put too much emphasis on the sole effort question that exhibits positive treatment effects.

little reason for them to increase effort in response to new information, or their effort may not result in measureable output. In our framework, one might imagine that F represents the students' beliefs about the production function, though not necessarily the true relationship.

In this scenario, the treatments changed beliefs, students put in more effort, but the effort was not effective at producing test scores given their lack of knowledge of how to translate effort into output. This explanation may also reconcile our set of facts with those presented in Nguyen (2013) and Jensen (2010). Less than half of the parents in Nguyen's sample finished their primary education, and 45% of the eighth graders in Jensen's control group do not enroll in high school the following year. This suggests that these populations are investing extremely little in their education at baseline, leaving significant "low-hanging fruit" unclaimed.[33] In other words, simple changes in effort – like coming to school – can produce large gains in this context. In the US, there is not likely such low hanging fruit, since most kids come to school and basic management problems such as rampant teacher absenteeism are also much smaller.

This is not the first time that similar educational interventions have shown much larger effects in the developing world than the United States. For instance, series of experiments in India (Duflo, Hanna, and Ryan 2012, Muralidharan and Sundararaman 2011) and Kenya (Glewwe, Ilias, and Kremer 2010) have revealed important achievement gains after the introduction of teacher incentives. Comparable merit pay initiatives have been ineffective in the United States (Fryer 2013, Springer et al. 2010, Fryer et al. 2012). A frequent explanation for these differences is similar to our discussion of low-hanging fruit; in the absence of incentives, teachers do not pursue simple measures to improve student achievement (for instance, unannounced visits revealed 35% of the schools in Duflo, Hanna, and Ryan's (2012) sample from India were closed due to teacher absenteeism).

Intuitively, the mapping from effort to academic success ought to be clear at low levels of investment. The decision to attend school or drop out, for instance, has a clear relationship to

[33] Theoretically, systematic differences in discount rates between the populations could also explain why similar treatments are more successful in developing countries. Since we do not directly observe discounting behavior in any of these experiments, evaluating this claim is difficult. Wang, Rieger, and Hens (2011) analyze survey data from 45 countries and find that citizens of poorer countries do have higher discount rates. However, Lawrance (1991) shows that low-income Americans in the Panel Study of Income Dynamics exhibit higher-than-average discounting behavior, suggesting that the national average may not be a good proxy for our population. Given the paucity of clear evidence, we can neither confirm nor rule out that discount rates explain the divergent findings.

academic achievement. At higher levels of investment, however, the ways in which different kinds of effort produce achievement is less clear. Once students are in school, they have to choose not just how much to study, but which particular types of studying to invest in.

This theory is also consistent with our set of facts if one assumes that students become more aware of the production function as they age. Moreover, this theory is in no way inconsistent with high-discount rates. In other words, the "true" model may have both present-bias and lack of knowledge of the production function.

Lack of knowledge of the production function explains the lack of effort – that is, students, being aware of not knowing how to produce achievement, lack incentive to exert effort even when they are made more aware of the rewards – and why short-term outcomes such as test scores did not increase.

The difficulty, it seems, with the lack of knowledge of production function theory is how to explain why ACT scores increased. To do this, we need a bit more structure. If, as students age, their knowledge of the production function increases, then this may explain the results. In this thought experiment, our treatment provided useful knowledge that students acted upon when the incentive to do so was large enough.[34]

C. Complementary Inputs

A third interpretation that may explain our findings is that the educational production function has important complementarities that are out of the student's control. For instance, student effort may need to be coupled with effective teachers, an engaging curriculum, safe neighborhoods, involved parents, or other inputs in order to yield increased achievement. In the parlance of our model, if capital levels $\boldsymbol{K}$ are so low that there is a very small return to effort, then students have little reason to work hard. In symbols: for small enough k_i, $\frac{\partial A}{\partial E}|_{E=E*} \approx 0$.

[34] Finally, it is important to note that the longer-term effects of treatment on ACT scores is statistically identical across treatments. To explain this, it may be the case that present-bias is important for the informational treatments and lack of knowledge of the production function explains the incentive treatments. Or, whatever effects that our information treatments had were similar to providing incentives. We cannot distinguish between these and other explanations for our results.

For intuition, consider a special case that lends itself to graphical exposition. Let the production technology be Cobb-Douglas with a single capital input, such that $F(E,K) = aE^{\alpha}K^{1-\alpha}$, and assume that the long-run benefits are linear in units of achievement: $V(A) = rA$. This allows us to use units of academic achievement as the numeraire and represent benefits and achievement on the same axis.

Figure 1 considers how achievement A responds to changes in returns r for different levels of capital K. The gray lines show the marginal product of effort at low levels of capital, and thin black lines depict the high-capital scenario. For each capital level, the solid curve represents the base case, in which we normalize the return r to one. The dashed lines show marginal payoffs after an increase in r.

The graph clarifies the two channels through which missing complements reduce treatment effects. First, because labor and capital are complements, the marginal return to a unit of effort is lower in equilibrium when capital levels are lower. Second, an increase in r results in a larger increase in equilibrium effort at higher levels of marginal productivity.

There are several (admittedly weak) tests of elements of this model that are possible with our data.[35] If effective teachers or environmental factors are an important complementary input to student incentives in producing test scores, we should notice a correlation between these inputs and the impact of providing information on achievement.[36] To test this hypothesis, we partition our sample based on three measures of external "capital" that are plausible complements of academic effort: (1) Teacher Quality (measured by teacher value-added (TVA) estimates calculated for ELA or math teachers for roughly 85% of our sample), (2) Neighborhood Quality (measured by the zip-code level poverty rates recorded in the American Community Survey), Neighborhood Segregation (measured by zip codes' Black Dissimilarity Indices). See Online Appendix B for precise details on how we calculate each of these measures.

[35] We are not, however, able to test potentially important (but far subtler) complementarities such as perceived discrimination or other obstacles, as Edin and Kefalas (2011) do in the context of poor women who choose to have children before marriage. These authors note that even when poor women delay marriage and choose to have children out of wedlock, they continue to *believe* in the institution of marriage, but don't necessarily act upon it, sometimes fearing that marriage is unattainable to them.

[36] A limitation to this analysis is that our measures of external capital investment may be correlated with students' baseline levels of knowledge about the returns to education.

To create subgroups, we rank all students in the experimental group and split the sample at the median. Appendix Tables 11A, 11B, and 11C presents treatment effects for our information treatment within each of these groups on our four main outcome measures.[37]

If anything, the resulting estimates demonstrate the opposite of what one might expect if complementarities in production were a driving force. Students from more segregated neighborhoods show *larger* increases in both math and reading scores. Similarly, students in the information only treatment arm assigned to low-TVA teachers show treatment effects of 0.088σ (0.070) in reading, relative to a -0.030σ (0.068) effect in math. The results are similar for high-TVA classrooms. These differences are inconsistent with a model in which complementarities are the driving force behind the lack of statistically significant effects on direct outcomes – though, as we have stated throughout, we only test a small subset of the possible channels in which complementarities might exist.

VII. Conclusion

In an effort to increase achievement and narrow achievement gaps, school districts have become incubators of innovative reforms. One potentially cost effective and quickly scalable strategy, not yet tested in American public schools, is to provide information to students about the returns to human capital.

This paper reports estimates of the impact of providing this type of information from a field experiment in Oklahoma City Public Schools during the 2010-2011 school year. Four facts emerge: (1) students update their beliefs about the returns to education in response to the text messages, (2) students report that they are putting more effort into their work, (3) there are no detectable changes in short-run academic achievement, but (4) long-term academic outcomes show improvement. How to interpret these facts in a model of human capital acquisition is less clear. We argue that high discount rates among youth or lack of adequate knowledge of the education production, or both, may explain our set of findings.

[37] In Appendix Tables 12A, 12B, and 12C, we report covariate means and balance tests within each of these subgroups. In the low-dissimilarity group for the information only arm of the treatment, the p-value on a joint significance test is 0.713; the other two subgroups are all well-balanced.

References

Ainsworth-Darnell, James W., and Douglas B. Downey. 1998. "Assessing the Oppositional Culture Explanation for Racial/ Ethnic Differences in School Performance." *American Sociological Review,* 63: 536-553.

Angrist, Joshua D. and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton: Princeton University Press.

Attanasio, Orazio, and Katja Kaufmann. 2009. "Educational Choices, Subjective Expectations, and Credit Constraints." NBER Working Paper No. 15087.

Avery, Christopher, and Thomas J. Kane. 2004. "Student Perceptions of College Opportunities: The Boston COACH Program." In *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It*, Caroline M. Hoxby, ed., Chicago: University of Chicago Press.

Beyer, Harald, Justine Hastings, Christopher Neilson, and Seth Zimmerman. 2015. "Connecting Student Loans to Labor Market Outcomes: Policy Lessons from Chile." *American Economic Review: Papers and Proceedings,* 105(5): 508-513.

Bruhn, Miriam, and David McKenzie. 2009. "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics*, 1(4): 200-232.

Bureau of Labor Statistics. 2011. "Economic News Release: Employment Status of the Civilian Population 25 Years and Over by Educational Attainment." http://www.bls.gov/news.release/empsit.t04.htm, accessed January 2011.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review*, 104(9): 2593-2632.

Devi, Tanaya, Roland G. Fryer, Jr., and Richard T. Holden. "Multitasking, Learning, and Incentives: A Cautionary Tale." Working Paper. Harvard University.

Dobbie, Will and Roland G. Fryer, Jr. 2015. "The Medium-Term Impacts of High Achieving Charter Schools." *Journal of Political Economy*, 123(5): 985-1037.

Dominitz, Jeff, and Charles F. Manski. 1996. "Eliciting Student Expectations of the Returns to Schooling." *Journal of Human Resources*, 31: 1–26.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review*, 102(4): 1241-78.

Edin, Kathryn, and Maria Kefalas. 2011. *Promises I Can Keep: Why Poor Women Put Motherhood Before Marriage, With a New Preface.* University of California Press.

Fryer Jr., Roland G. 2011a. “Racial Inequality in the 21st Century: The Declining Significance of Discrimination.” In *Handbook of Labor Economics*, Orley Ashenfelter and David Card, eds. Vol. 4. Amsterdam: Elsevier Science/North-Holland.

Fryer Jr., Roland G. 2011b. “Financial Incentives and Student Achievement: Evidence from Randomized Trials.” *Quarterly Journal of Economics,* 126: 1755-1798.

Fryer Jr., Roland G. 2013. “Teacher Incentives and Student Achievement: Evidence from New York City Public Schools. *Journal of Labor Economics*, 31(2): 373-407.

Fryer Jr, Roland G., Steven D. Levitt, John List, and Sally Sadoff. 2012. “Enhancing the Efficacy of Teacher Incentives through Loss Aversion: A Field Experiment.” NBER Working Paper No. 18237.

Glewwe, Paul, Nauman Ilias, and Michael Kremer. 2010. “Teacher Incentives.” *American Economic Journal: Applied Economics*, 2(3): 205-227.

Greevy, Robert, Bo Lu, and Jeffrey H. Silber. 2004. “Optimal Multivariate Matching before Randomization." *Biostatistics*, 5: 263-275.

Harris, Angel L. 2008. “Optimism in the Face of Despair: Black-White Differences in Beliefs About School as a Means for Upward Social Mobility.” *Social Science Quarterly*, 89(3): 608–630.

Imai, Kosuke, Gary King, and Clayton Nall. 2009. “The Essential Role of Pair Matching in Cluster Randomized Experiments." *Statistical Science*, 24(1): 29-53.

Imbens, Guido. 2011. “Experimental Design for Unit and Cluster Randomized Trials.” Conference Paper, International Initiative for Impact Evaluation.

Jahn, Julius A., Calvin F. Schmid, and Clarence Schrag. 1947. “The Measurement of Ecological Segregation.” *American Sociological Review*, 12: 293-303.

Jencks, Christopher, and Meredith Phillips, eds. 1998. *The Black-White Test Score Gap*. Washington D.C.: Brookings Institution Press.

Jensen, Robert. 2010. “The (Perceived) Returns to Education and the Demand for Schooling,” *Quarterly Journal of Economics,* 125(2): 515-48.

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. 2010. “Getting to the Top of Mind: How Reminders Increase Saving.” NBER Working Paper No. 16205.

Kasy, Maximilian. 2016. "Why Experimenters Should Not Randomize, and What They Should do Instead." *Political Analysis*, (2016): 1-15.

Kaufmann, Katja M. 2014. "Understanding the Income Gradient in College Attendance in Mexico: The Role of Heterogeneity in Expected Returns to College." *Quantitative Economics, 5(3): 583-630.*

Lawrance, Emily C. 1991. "Poverty and the Rate of Time Preference: Evidence from Panel Data." *Journal of Political Economy,* 99(1): 54-77.

Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies*, 76(3): 1071-1102.

Lieberson, Stanley. 1980. *A Piece of the Pie: Blacks and White Immigrants Since 1880*. Berkeley: University of California Press.

Mickelson, Roslyn A. 1990. "The Attitude-Achievement Paradox among Black Adolescents." *Sociology of Education*, 63*:* 44-61.

Mischel, Walter. 2014. *The Marshmallow Test: Why Self-Control is the Engine of Success*. Little Brown.

Mullainathan, Sendhil, Joshua Schwartzstein, and Andrei Shleifer. 2008. "Coarse Thinking and Persuasion." *Quarterly Journal of Economics*, 123: 577–619.

Muralidharan, Karthik and Venkatesh Sundararaman. 2011. "Teacher Performance Pay: Experimental Evidence from India." *Journal of Political Economy*, 119 (1).

Neal, Derek, and William Johnson. 1996. "The Role of Premarket Factors in Black-White Wage Differentials." *Journal of Political Economy*, 104: 869–95.

Nelson, Phillip. 1974. "Advertising as Information." *Journal of Political Economy*, 81: 729-754.

Nguyen, Trang. 2013. "Information, Role Models and Perceived Returns to Education: Experimental Evidence from Madagascar," *enGender Impact: The World Bank's Gender Impact Evaluation Database.* Washington DC: World Bank.

Ogbu, John U. 1978. *Minority Education and Caste: The American System in Cross-Cultural Perspective.* New York: Academic Press.

Rouse, Cecelia E. 2004. "Low-Income Students and College Attendance: An Exploration of Income Expectations." *Social Science Quarterly*, 85: 1299–1317.

Shapiro, Jesse. 2006. "A 'Memory-Jamming' Theory of Advertising." Available at SSRN: http://ssrn.com/abstract=903474.

Smith, Herbert L., and Brian Powell. 1990. "Great Expectations: Variations in Income Expectations among College Seniors." *Sociology of Education*, 63: 194–207.

Springer, Matthew G., Dale Ballou, Laura S. Hamilton, Vi-Nhuan Le, J.R. Lockwood, Daniel F. McCaffrey, Matthew Pepper, and Brian M. Stecher. 2010. "Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching." Nashville, TN: National Center on Performance Incentives at Vanderbilt University.

Steinberg, Laurence, Sandra Graham, Lia O'Brien, Jennifer Woolard, Elizabeth Cauffman, and Marie Banich. 2009. "Age Differences in Future Orientation and Delay Discounting." *Child Development,* 80 (1): 28-44.

United States Census Bureau. 2011. "Current Population Survey Data on Educational Attainment." http://www.census.gov/hhes/socdemo/education/data/cps/index.html, accessed May 2011.

Wang, Mei, Marc O. Reiger, and Thorsten Hens. 2011. "How Time Preferences Differ: Evidence From 45 Countries." Department of Business and Management Science, Norwegian School of Economics, Discussion Paper No. 2011/18.

Wilson, William Julius. 1987. "The Truly Disadvantaged: The Inner City, the Underclass, and Social Policy." Chicago: University of Chicago Press.

Wiswall, Matthew and Basit Zafar. 2014. "Determinants of College Major Choice: Identification using an Information Experiment." Federal Reserve Bank of New York Staff Reports No. 500.

York, Benjamin N. and Susanna Loeb. 2014. "One Step at a Time: The Effects of an Early Literacy Text Messaging Program for Parents of Preschoolers." NBER Working Paper No. 20659.

Table 1: Summary of The Million Experiment

A. Overview	
Schools	All 22 non-alternative OKCPS schools serving grades 6 and 7 opted in to participate. All experimental schools were provided complete Accelerated Reading software, training, and implementation materials. All treatment students received a Samsung t401g mobile phone.
Treatment Group	1,470 6th and 7th grade students: 31.1% black, 44.3% Hispanic, 91.7% free lunch eligible
Control Group	437 6th and 7th grade students: 30.9% black, 43.5% Hispanic, 91.8% free lunch eligible
Outcomes of Interest	Student Knowledge of Returns to Education, Oklahoma Core Curriculum Criterion Referenced Test (CRT), Measures of Student Effort and Motivation, Attendance, Suspensions
Test Dates	CRT: April 11-26, 2011
Operations	$230,365 worth of hardware and incentives distributed to treatment students, 34.3% consent rate. 1 dedicated project managers.

B. Treatments	(1) Information Only	(2) Information & Incentives	(3) Non-Financial Incentives Only
Phone	Free Samsung t401g mobile phone	Free Samsung t401g mobile phone	Free Samsung t401g mobile phone
Basic Reward Structure	Fixed allotment of 300 minutes per month	Students earned 10 cell phone minutes per Accelerated Reader point earned, distributed in blocks of 200 minutes	Students earned 10 cell phone minutes per Accelerated Reader point earned, distributed in blocks of 200 minutes
Informational Campaign	Students received one informational or persuasive message per day	Students received one informational or persuasive message per day	None
Reward Frequency	Monthly, unconditional	Biweekly, contingent upon AR points earned	Biweekly, contingent upon AR points earned

Notes. In panel A, each row describes an aspect of treatment indicated in the first column. In panel B, each column represents a different arm of treatment. Entries are descriptions of the schools, students, outcomes of interest, testing dates, and basic operations of each phase of the incentive treatment. See Online Appendix A for more details. The number of treatment and control students given are for those students who have non-missing reading or math test scores.

Table 2: Student Baseline Characteristics

	Non Participating	Participating	*p-value* (1) = (2)	Information	Incentives & Information	Incentives	Control	*p-value* (4)=(5)=(6)=(7)
Student Characteristics	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Male	0.521	0.486	0.019	0.453	0.504	0.453	0.538	0.022
	(0.500)	(0.500)		(0.500)	(0.500)	(0.498)	(0.499)	
White	0.200	0.163	0.002	0.149	0.167	0.163	0.172	0.799
	(0.400)	(0.369)		(0.374)	(0.374)	(0.370)	(0.377)	
Black	0.290	0.311	0.125	0.294	0.327	0.314	0.309	0.740
	(0.454)	(0.463)		(0.469)	(0.469)	(0.465)	(0.463)	
Hispanic	0.435	0.443	0.634	0.469	0.424	0.441	0.435	0.534
	(0.496)	(0.497)		(0.495)	(0.495)	(0.497)	(0.496)	
Asian	0.025	0.026	0.824	0.018	0.031	0.018	0.037	0.199
	(0.155)	(0.158)		(0.172)	(0.172)	(0.134)	(0.188)	
Other Race	0.051	0.058	0.288	0.069	0.051	0.063	0.048	0.453
	(0.220)	(0.234)		(0.220)	(0.220)	(0.244)	(0.214)	
Special Education Services	0.149	0.139	0.326	0.131	0.141	0.147	0.137	0.903
	(0.356)	(0.346)		(0.348)	(0.348)	(0.354)	(0.345)	
English Language Learner	0.154	0.159	0.612	0.165	0.153	0.159	0.160	0.964
	(0.361)	(0.366)		(0.360)	(0.360)	(0.366)	(0.367)	
Free Lunch	0.857	0.917	0.000	0.922	0.920	0.908	0.918	0.857
	(0.351)	(0.276)		(0.271)	(0.271)	(0.289)	(0.275)	
Economically Disadvantaged	0.741	0.915	0.000	0.922	0.914	0.908	0.915	0.886
	(0.438)	(0.279)		(0.280)	(0.280)	(0.289)	(0.279)	
Baseline Math	0.010	0.030	0.565	-0.009	0.028	0.006	0.108	0.400
	(1.022)	(0.983)		(1.035)	(1.035)	(0.993)	(0.916)	
Baseline Reading	0.037	-0.021	0.098	-0.086	0.007	-0.053	0.062	0.204
	(1.015)	(1.010)		(1.049)	(1.049)	(0.989)	(0.905)	
Missing: Baseline Math	0.319	0.216	0.000	0.169	0.224	0.237	0.233	0.036
	(0.466)	(0.411)		(0.418)	(0.418)	(0.426)	(0.423)	
Missing: Baseline Reading	0.326	0.219	0.000	0.176	0.231	0.231	0.243	0.055
	(0.469)	(0.414)		(0.422)	(0.422)	(0.422)	(0.429)	
p-value from joint F-test			0.000					0.387
Observations	2903	1907	4810	490	490	490	437	1907

Notes: This table reports summary statistics for the field experiment. Columns (1), (2), (4), (5), (6) and (7) represent the sample means of the variable indicated in each row for the group indicated in each column The treatment groups are restricted to randomly selected 6th and 7th grade students in Oklahoma City Public Schools experimental schools who opted into the randomization for the field experiment. Columns (3) and (8) report the p-value from a test of equality across treatment indicators (or experimental group indicators) from a regression of the variable in each row on indicators for each treatment group and the control group (or experimental group status). The joint F-tests report the p-value from a test of the null hypothesis that there are no differences between the given groups across all reported variables in the table.

Table 3 - Mean Effect Sizes (Intent-to-Treat) on Direct Outcomes

	Control Mean	Information	Information & Incentives	Incentives	*p-value*
	(1)	(2)	(3)	(4)	(5)
A. Treatment Questions					
Knows Wage Gap btw BA and Dropouts	0.819	0.054*	0.042	0.014	0.676
		(0.032)	(0.031)	(0.033)	
		569	592	589	
Knows Schooling & Income Relationship	0.482	-0.005	-0.023	0.030	0.674
		(0.044)	(0.043)	(0.042)	
		563	577	581	
Knows Prison Rates	0.459	0.174***	0.172***	-0.043	0.000
		(0.045)	(0.043)	(0.043)	
		561	587	585	
Number of Questions Correct	1.774	0.228***	0.195***	-0.005	0.048
		(0.073)	(0.070)	(0.072)	
		544	563	564	
B. Placebo Question					
Knows Unemployment Rate of College Grads	0.327	0.036	-0.011	0.039	0.628
		(0.042)	(0.041)	(0.043)	
		573	590	590	

Notes: This table reports ITT estimates for the effect of being offered a chance to participate in the field experiment on students' ability to correctly answer questions about human capital development. Column 1 presents means for students that were randomly assigned to the the control group. Questions are coded as a 1 if the student answered the question correctly and a 0 otherwise. All regressions include school fixed effects and controls for student grade, gender, race, SES, special education status, and English language learner status, as well as 2009 state test scores, 2010 state test scores, and their squares and cubes. The sample is restricted to randomly selected 6th and 7th grade students in Oklahoma City Public Schools. Randomization was done at the student level. Treatment is defined as returning a signed consent form to participate and being lotteried into the specified treatment group. Heteroskedasticity-robust errors are reported in parentheses below each estimate. The number of observations in each regression is reported directly below the standard errors. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Table 4 - Mean Effect Sizes (Intent-to-Treat) on Indirect Outcomes

	Control Mean	Information	Information & Incentives	Incentives	*p-value*
	(1)	(2)	(3)	(4)	(5)
A. Survey Outcomes					
Effort Index	-0.009	0.009	-0.016	-0.033	0.814
		(0.048)	(0.046)	(0.046)	
		582	604	607	
B. Administrative Data Outcomes					
OK State Math Test Post-Treatment	0.126	0.012	-0.062	-0.034	0.504
		(0.046)	(0.045)	(0.047)	
		794	790	782	
OK State Reading Test Post-Treatment	0.029	0.068	0.014	0.027	0.688
		(0.046)	(0.047)	(0.049)	
		786	790	780	
OK State Math Test Post-Treatment, Nationally Normed	-0.275	0.009	-0.044	-0.025	0.505
		(0.033)	(0.032)	(0.033)	
		796	790	782	
OK State Reading Test Post-Treatment, Nationally Normed	-0.378	0.049	0.008	0.018	0.635
		(0.032)	(0.033)	(0.034)	
		789	790	780	
Attendance rate (std.)	0.110	0.010	0.020	0.052	0.871
		(0.059)	(0.061)	(0.059)	
		856	863	861	
Number of Suspensions	0.471	0.020	0.019	0.023	0.999
		(0.069)	(0.074)	(0.074)	
		927	927	927	

Notes: This table reports ITT estimates for the effect of being offered a chance to participate in the field experiment on survey and administrative data outcomes. Column 1 presents means for students that were randomly assigned to the the control group. Survey measures are coded as a 1 if the student answered a question indicating that he or she agreed with the statement in the corresponding row and a 0 otherwise. Test score and attendance variables are standardized to have mean zero and standard deviation one by grade in the full OKCPS 6th and 7th grade samples. Test score variables are also reported normed to the national distribution of scores on the National Assessment of Educational Progress (NAEP), for details see Appendix B. All regressions include school fixed effects and controls for student grade, gender, race, SES, special education status, and English language learner status, as well as 2009 state test scores, 2010 state test scores, and their squares and cubes. The sample is restricted to randomly selected 6th and 7th grade students in Oklahoma City Public Schools. Randomization was done at the student level. Treatment is defined as returning a signed consent form to participate and being lotteried into the specified treatment group. Heteroskedasticity-robust errors are reported in parentheses below each estimate. The number of observations in each regression is reported directly below the standard errors. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Table 5 - Mean Effect Size on Attrition

	Control Response Rate	Differential Follow-up		
		Information	Information & Incentives	Incentives
	(1)	(2)	(3)	(4)
Reading	0.931	-0.007	0.004	0.001
		(0.007)	(0.004)	(0.005)
Mathematics	0.931	0.001	-0.001	-0.001
		(0.005)	(0.005)	(0.005)
Survey	0.611	0.038	0.079***	0.069***
		(0.024)	(0.024)	(0.024)
Number of Observations		927	927	927

Notes: This table reports differential rates of attrition for individuals in the field experiment's experimental groups. Column (1) reports the share of control students with non-missing values for the post-treatment outcomes indicated in each row. Columns (2), (3) and (4) report coefficients from regressions of an indicator variable equal to one if the outcome in the same row is non-missing on an indicator for being randomly selected into the indicated treatment group. All regressions include the full set of covariates and fixed effects used in the preceding tables. Heteroskedasticity-robust errors are reported in parentheses below each estimate. The number of observations in each regression is reported in the final row. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Table 6 - Bounding

	Survey ITT	Lee Lower Bound	*p-value* (1)=(2)	Lee Upper Bound	*p-value* (1)=(4)
	(1)	(2)	(3)	(4)	(5)
A. Information Treatment versus Control					
Knows Wage Gap btw BA and Dropouts	0.054*	0.048	0.891	0.092***	0.390
	(0.032)	(0.032)		(0.030)	
	569	556		556	
Knows Schooling & Income Relationship	-0.005	-0.035	0.633	0.017	0.724
	(0.044)	(0.044)		(0.044)	
	563	550		550	
Knows Prison Rates	0.174***	0.158***	0.801	0.190***	0.795
	(0.045)	(0.045)		(0.044)	
	561	550		550	
Number of Questions Correct	0.228***	0.181**	0.654	0.256***	0.785
	(0.073)	(0.072)		(0.072)	
	544	533		533	
Effort Index	0.009	-0.026	0.597	0.049	0.561
	(0.048)	(0.047)		(0.048)	
	582	568		568	
B. Information & Incentives Treatment versus Control					
Knows Wage Gap btw BA and Dropouts	0.042	0.017	0.570	0.123***	0.047
	(0.031)	(0.031)		(0.027)	
	592	564		564	
Knows Schooling & Income Relationship	-0.023	-0.079*	0.355	0.015	0.533
	(0.043)	(0.043)		(0.043)	
	577	555		555	
Knows Prison Rates	0.172***	0.123***	0.417	0.227***	0.359
	(0.043)	(0.043)		(0.042)	
	587	560		560	
Number of Questions Correct	0.195***	0.115*	0.419	0.275***	0.410
	(0.070)	(0.069)		(0.067)	
	563	541		541	
Effort Index	-0.016	-0.099**	0.200	0.055	0.270
	(0.046)	(0.045)		(0.046)	
	604	575		575	
C. Incentives Treatment versus Control					
Knows Wage Gap btw BA and Dropouts	0.014	-0.007	0.658	0.079**	0.157
	(0.033)	(0.033)		(0.031)	
	589	569		569	
Knows Schooling & Income Relationship	0.030	-0.023	0.375	0.069*	0.506
	(0.042)	(0.042)		(0.042)	
	581	560		560	
Knows Prison Rates	-0.043	-0.100**	0.348	-0.014	0.637
	(0.043)	(0.043)		(0.043)	
	585	565		565	
Number of Questions Correct	-0.005	-0.080	0.454	0.059	0.522
	(0.072)	(0.071)		(0.070)	
	564	546		546	
Effort Index	-0.033	-0.099**	0.308	0.020	0.407
	(0.046)	(0.045)		(0.045)	
	607	582		582	

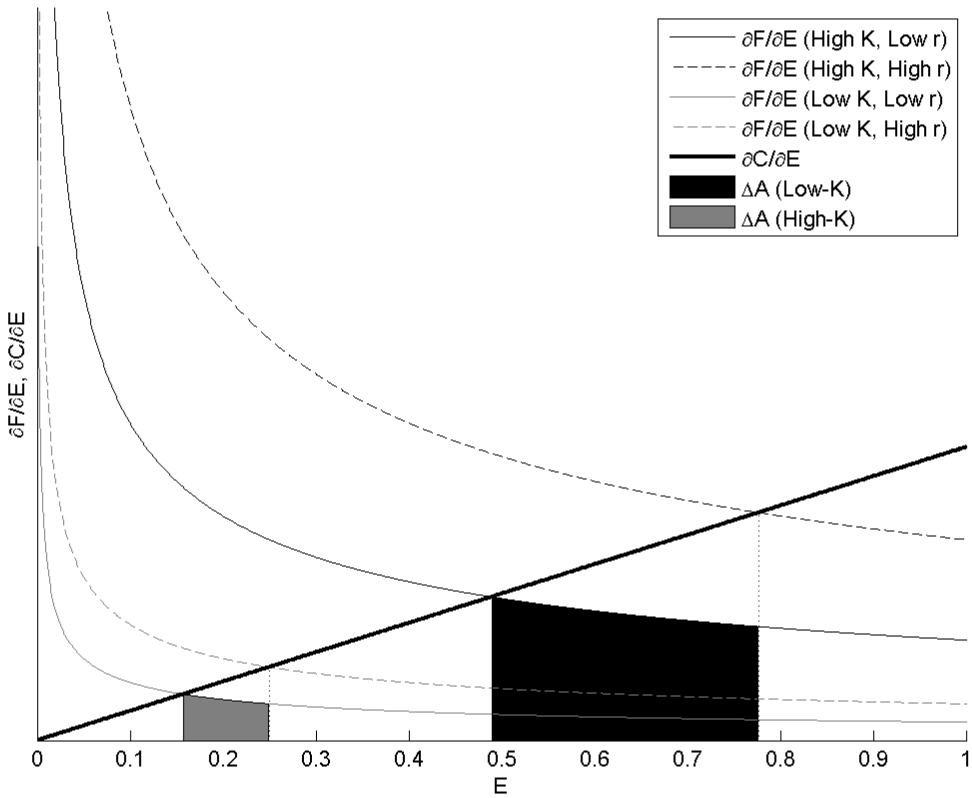
Notes: This table reports upper and lower Lee bounds to account for survey attrition. For ease of comparison, Column (1) reproduces the survey results from Table 3 for our direct survey outcomes of interest. Column (2) reports lower Lee Bounds. These bounds are generated by predicting the residuals from a regression of the survey outcome of interest on baseline test scores, demographics, and treatment-year test scores within the control group only. The treatment group is then sorted and individuals with the largest residuals from the regressions are removed from the regression to equate response rates between treatment and control. The resulting Lee lower bounds are from an OLS regression identical to our main specification after trimming the sample in this way. Column (4) reports upper Lee Bounds. These bounds are generated by the same process as lower Lee Bounds, except individuals with the smallest residuals are removed from the regression to equate response rates between treatment and control. Columns (3) and (5) report p-values on the null hypothesis that the treatment coefficients from the LEE bounds are equal to the treatment coefficient from the main ITT specification for the treatment group indicated in the panel title. Heteroskedasticity-robust errors are reported in parentheses below each estimate. The number of observations in each regression is reported directly below the standard errors. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Table 7 - Mean Effect Sizes (Intent-to-Treat) on ACT Scores

	Control Mean	Information	Information & Incentives	Incentives	*p-value*
	(1)	(2)	(3)	(4)	(5)
First ACT Comprehensive Score	-0.774	0.143**	0.069	0.091	0.685
	(0.787)	(0.063)	(0.060)	(0.058)	
	154	154	173	166	
Max ACT Comprehensive Score	-0.711	0.132**	0.090	0.101*	0.890
	(0.853)	(0.064)	(0.062)	(0.060)	
	154	154	173	166	
First ACT Math Score	-0.677	0.089	0.030	0.069	0.761
	(0.743)	(0.060)	(0.055)	(0.053)	
	154	154	173	166	
Max ACT Math Score	-0.641	0.082	0.048	0.065	0.918
	(0.777)	(0.059)	(0.058)	(0.055)	
	154	154	173	166	
First ACT English Score	-0.816	0.186***	0.100*	0.153**	0.625
	(0.819)	(0.067)	(0.060)	(0.067)	
	154	154	173	166	
Max ACT English Score	-0.765	0.167**	0.135**	0.158**	0.937
	(0.873)	(0.069)	(0.063)	(0.067)	
	154	154	173	166	
First ACT Reading Score	-0.709	0.168**	0.068	0.066	0.566
	(0.851)	(0.077)	(0.078)	(0.077)	
	154	154	173	166	
Max ACT Reading Score	-0.630	0.131	0.064	0.059	0.775
	(0.924)	(0.081)	(0.079)	(0.078)	
	154	154	173	166	
First ACT Science Score	-0.655	0.117	0.060	0.043	0.767
	(0.795)	(0.078)	(0.074)	(0.069)	
	154	154	173	166	
Max ACT Science Score	-0.595	0.137*	0.081	0.082	0.833
	(0.865)	(0.076)	(0.075)	(0.071)	
	154	154	173	166	
Took the ACT	0.352	-0.031	0.009	-0.004	0.629
	(0.478)	(0.030)	(0.030)	(0.030)	
	437	490	490	490	

Notes: This table reports ITT estimates for the effect of being offered a chance to participate in the field experiment on students' ACT scores, normalized to the national distribution of high school graduates in 2015-2016. We report results for both the first test a student took and the test where they scored their maximum comprehensive score. We report estimates for the comprehensive score and each subscore. All regressions include school fixed effects and controls for student grade, gender, race, SES, special education status, and English language learner status, as well as 2009 state test scores, 2010 state test scores, and their squares and cubes. The sample is restricted to randomly selected 6th and 7th grade students in Oklahoma City Public Schools. Randomization was done at the student level. Treatment is defined as returning a signed consent form to participate and being lotteried into the specified treatment group. Heteroskedasticity-robust errors are reported in parentheses below each estimate. The number of observations in each regression is reported directly below the standard errors. *** = significant at 1 percent level, ** = significant at 5 percent level, * = significant at 10 percent level.

Figure 1: Treatment Effects Under High and Low Capital Endowments



Notes: The figure depicts how achievement changes with an increase in perceived returns r in a low-capital and high-capital scenario. The model is described in Section VI of the text and is parameterized as follows: $a=1$, $\alpha=0.5$, $K_{high}=30$, $K_{low}=1$, and $C(E) = 4E^2$.