Achieving Academic Success After School:

A Randomized Evaluation of the Higher Achievement Program[1]

September 2011

Leigh L. Linden
(The University of Texas at Austin, NBER, J-PAL, IZA)

Carla Herrera
(Public/Private Ventures)

Jean Baldwin Grossman
(Public/Private Ventures, Princeton University)

Abstract:
We conduct a RCT of a high-quality, comprehensive, and time-intensive after-school and summer supplemental-education program targeted at students in middle school. We find that the program significantly increases students' problem solving and reading comprehension scores two years after baseline. The program also causes students to reassess their perceptions of their own academic abilities. We find no evidence that the program improves test scores after one year or that the program changes students perceptions of the support they receive from adults or their peers. We also do not find that the program improves students' test scores or other outcomes over the summer of 2010. However, their test scores also do not decrease over the summer as they do for many students. Finally, we find evidence that the program increases students' desire to attend competitive area high schools rather than their local public schools. Surprisingly, the program also increases the rate at which students report misbehaving.

[1] This study could not have been completed without the assistance of many individuals. We are, of course, grateful to the parents and youth who took the time to provide the data used in this manuscript. At Higher Achievement, Edsson Contreras, Maureen Holla, Lynsey Jeffries, Nicole Levine, Richard Tagle, and Gail Williams all provided significant help with recruitment and other aspects of the study. Annelies Raue provided exceptional assistance in conducting the presented analyses. Ama Baafra Abeberese, Evan Borkum, and Mariesa Herrmann also assisted with data analysis during earlier phases of the project. At P/PV, Jennifer McMaken co-directed the project throughout the first several years of its implementation. Survey Research Management provided all data collection services, and Vestal McIntyre edited the manuscript. The William T. Grant Foundation, The Atlantic Philanthropies, and The Wallace Foundation all provided the financial support necessary to complete this study. Please direct all correspondence regarding the manuscript to Leigh L. Linden at leigh.linden@austin.utexas.edu.

I. Introduction

The achievement gap has long been a major challenge of the American educational system. Students from historically disadvantaged backgrounds—minority and low social economic status—underachieve academically relative to their peers (Stern, 1989). These children enter school at a deficit and fall further behind as they progress through school (Campbell, Hombo and Mazzeo, 2000; Neal, 2006). While some of the differences may be driven by socio-cultural factors, the quality of schools typically attended by these youths may also play a significant role. Far too often, youths from disadvantaged backgrounds attend lower quality schools than those available to their more advantaged peers (Morgan and Sirageldin, 1968; Johnson and Stafford, 1973). Although options do exist for academically oriented students to attend more competitive schools that would improve their academic performance (Neal, 1997; Neal et al., 2000), disadvantaged families often undervalue or lack the knowledge of the administrative procedures necessary to take advantage of such opportunities (Research for Action, 2010)—barriers that may also prevent college enrollment (Hoxby, 2011).

Given the difficulty of improving the school environments available to disadvantaged youth (and their knowledge of and desire to attend higher-quality schools), out-of-school-time (OST) programs are seen as a strategy that could offset these educational disparities through supplemental educational programming. Academically focused OST programs, in particular, aim to improve students' academic attitudes, behaviors and performance by increasing youths' access to high-quality academic supports and opportunities. But these programs range widely in their focus and intensity—variations that likely affect their ability to make such improvements. For example, programs that provide youths with long-term (multi-year) academic support and that carefully integrate both school-year (i.e., after-school) *and* summer learning opportunities would

appear to be particularly promising, as they extend the amount of time youths devote to learning and other positive activities across the entire year. Those that target youths as they transition to middle school also might be especially helpful because they reach youths at a time when even strong students can experience academic slides (Eccles and Midgley, 1989; Seidman et al., 1994; Blyth et al., 1983). Finally, OST programs that explicitly encourage application and matriculation at competitive local high schools, may serve as a catalyst for moving disadvantaged children into higher quality educational tracks—the same ones experienced by their relatively advantaged peers.

Unfortunately, very few programs have all of these features. And, for those that do, there is little rigorous evidence supporting claims that they can significantly boost students' academic performance over the long term and even less evidence that they can alter high school selection. We evaluate the long-term effects of a mature, time-intensive, "Cadillac" version of an OST academic enrichment program that boasts all of these characteristics. If this comprehensive program fails to generate long-term improvements in academic outcomes, then OST programs in general may not be capable of doing so, but if it succeeds, then it will demonstrate that such programs can have lasting impacts on students' lives.

This study is based on a sample of youths that we recruited to participate in a two-year random assignment impact evaluation of the Higher Achievement (HA) program in Washington, DC. Recruitment for the study started in late spring of 2006. Over a three-year period, we recruited 951 students in three cohorts, randomly assigning them to either a treatment group that was offered a slot in the program or a control group that was not. We then conducted follow-up surveys one and two years after baseline to evaluate the direct effect of the overall program, as well as two follow-ups conducted before and after the summer of 2010 to assess the effects of

the program on the summer learning loss. A final ongoing component of the project will assess the effects of the program on students four years after baseline.[2]

We find that the program increases students' standardized test scores in both problem solving and reading comprehension and causes students to reassess their perceptions of their own academic abilities. At the first-year follow-up, we find the program had no effect on test scores, but by the end of the second year, we find statistically significant intent-to-treat effects of 0.12 standard deviations on students' problem solving skills and 0.09 standard deviations on students' reading comprehension skills. The estimated effects of attending an academically oriented OST program is 0.21 standard deviations on problem solving scores and 0.16 standard deviations on reading comprehension scores.

The program's effects on students' academic attitudes vary by whether the students enter the program the summer before grade five or grade six. Treatment students in both grades both experience declines in their self-perceptions, typical of students entering middle school. However, possibly because the control group fifth graders are finishing their last year of elementary school, their scores remain steady, while the treatment fifth graders experience a net negative treatment effect. At the same time, the sixth graders experience no net effect because their control peers (who are also transitioning to middle school) decline as well. By the second year, the treatment effect on the youths who started as rising sixth graders is positive while the effect for the fifth graders is still negative, but smaller in magnitude as the control group's attitudes drop significantly during this period as they enter middle school. We find no effects on students' perceptions of support from adults and peers, and we find a consistent negative effect

[2] This component of the study is being funded by the William T. Grant Foundation and the Spencer Foundation. To date, we have completed surveys for the first two cohorts and hope to complete survey administration for the final cohort in late spring of 2012 with an anticipated release of the results in mid-2013.

on students' self-reported misbehavior, suggesting that the program causes students to act out more than they otherwise would.

When examining the effects of HA over the summer of 2010, we find evidence that the program increases students' desire to attend a competitive area high school as well as their reports of having taken some of the steps necessary to apply, such as visiting a school, acquiring information about it, and having made a decision about where to apply. Unlike many disadvantaged students, we also find that the students who apply to HA but do not have access to the program (i.e., the control group) do not experience the summer learning loss over the summer of 2010, perhaps due to their access to other OST enrichment opportunities. HA students also experience no decline in test scores over the summer, but they do not experience an increase in scores over those gains experienced by the control group, yielding no relative impacts in test scores during summer 2010.

The observed impacts on test scores stand in contrast to much of the OST literature which provides mixed results on the effectiveness of OST programs. To date, there have only been two large-scale randomized controlled trials of OST programs. James-Burdumy, Dynarski and Deke (2007) find no academic effects of the 21st Century Community Learning Centers in a large-scale experimental and quasi-experimental evaluation. However, the evaluation has been heavily criticized regarding possible methodological problems and the fact that the evaluated programs were immature, had very low participation rates, and lacked significant, formal academic programming (Mahoney and Zigler, 2006). In a follow-up one-year study that compared such undeveloped programs to ones that employed a well-delivered, research-based math and reading curriculum, Black et al. (2008) find, also using a large-scale experimental design, that the

stronger programs increase math scores by only 0.06 standard deviations but had no effect on reading scores.[3,4]

The remainder of the manuscript is organized as follows: Section II provides an overview of the HA program. Section III describes the research design, and Section IV assesses the internal validity of the study. We present the results in Section V. Finally, we conclude in Section VI.

II. Intervention

Higher Achievement is an established OST program that strives to develop positive academic behaviors and attitudes in academically motivated but underserved fifth- through eighth-grade youth. Its long-term objective is to support participants' academic achievement and encourage their application and matriculation at competitive area high schools. Most participants are African American and come from low-income families. As our data show (Section V.B), students are typically about average in their academic performance when compared to the general population and include both high- and low-performing students.

HA participants ("scholars") typically begin the program the summer before they enter fifth grade or before they transition to middle school (sixth grade)—a period when most youths are still academically engaged. HA involves youths at this time because research has indicated

[3] There are a number of other RCT's in the literature. However, they suffer from various methodological problems. The vast majority simply have very small sample sizes—often under 100 students (Beckett et al., 2009). Chaplin and Capizzano (2006) estimate the effects of the Build Educated Leaders for Life (BELL) summer program on students' reading comprehension scores to be 0.08 standard deviations using a larger sample. However, while this study is built around a randomized controlled trial, the control group experienced more total days of school than the treatment group. The 0.08 standard deviations estimate is based on a comparison that adjusts for the days of school received by each student. The unadjusted intent-to-treat estimates show no effects on students' reading scores.

[4] Several meta-analyses have demonstrated larger positive aggregate effects on students' standardized test scores. While some of these estimates are comparable to the intent-to-treat estimates in our two-year study (Lauer et al., 2006, for example), these studies (a.) rely on non-experimental studies or experimental studies with very small samples for the larger treatment effect estimates in their sample and (b.) are subject to the significant analytical problems associated with aggregating results across studies, including the issue of publication bias.

that this period is critical in determining students' future academic success. They aim to keep students academically motivated throughout middle school and, in the process, substantially improve their skills so that they can transition into DC's best high schools where they can continue to receive high quality academic instruction after leaving HA.

The program is extremely time-intensive, offering approximately 650 hours a year of academic instruction, enrichment activities and mentoring during the after-school and summer hours. During the school year, scholars attend the "After-School Academy" three days a week. This 25-week program, which runs from 3:30 to 8:00 p.m., includes homework help, dinner, an elective, a 15-minute group meeting and two hours of academic instruction in small groups of about two to three scholars—one day a week in mathematics, one day in literature and one in technology. Volunteer mentors lead these groups. Scholars participate in monthly field trips and community service projects. During the summer, the six-week "Summer Academy" operates from 8:00 a.m. to 4:00 p.m., five days a week. Students attend four classes a day, taught by trained faculty, in mathematics, science, social studies and literature as well as two electives. Scholars take weekly field trips and participate in a three-day university trip during which they experience college life by attending classes, sleeping in dorms, going to lectures and eating in dining halls.

In both the After-School and Summer Academies, students receive high school placement services. These include general information about area schools as well as specific details on application processes and deadlines. Scholars and their families receive help (often one-on-one) with interviewing, selecting schools and completing application and financial aid materials. Scholars also visit local high schools to learn about them firsthand. As youths approach the eighth grade, this level of support increases. For example, in the fall of the eighth grade year, one

mentoring session each week is devoted to high school applications. More generally, HA encourages scholars to value and believe they are capable of academic success to lay the attitudinal foundation that will encourage students to apply to and matriculate at competitive high schools.

Both developmental theory as well as empirical research support HA's program structure. HA serves youths during a turbulent time in their development. Many students experience declines in academic motivation, confidence and achievement during the transition from elementary to middle school (Anderman and Maehr, 1994; Seidman et al., 1994; Wigfield et al., 1991; Eccles and Midgley, 1989), and minority youths in particular may struggle with an increasing awareness of racial stereotypes (Simmons and Blythe, 1987).

Eccles and Midgley (1989) argue that middle school fails to meet early-adolescent developmental needs, resulting in an environment that negatively affects youths (e.g., Seidman et al., 1994). For example, young adolescents have a strong need for close relationships with adults and desire more autonomy and control over their learning than they did in elementary school (Eccles and Wigfield, 2000; Eccles and Midgley, 1989). However, middle schools are typically less conducive to fulfilling these needs than are elementary schools. They are larger, rely on teacher-led instruction, and require students to rotate among teachers. As a result, students have fewer opportunities to develop close relationships with individual teachers.

A number of studies show a strong correlation between the quality of students' relationships with their instructors and academic attitudes and behaviors (e.g., Goodenow, 1993; Hamre and Pianta, 2001) and between active learning styles and engagement (Stipek, 2002). By providing scholars with opportunities to develop close relationships with mentors and to learn in small-group interactive settings, HA hopes to counteract the dips in academic attitudes that many

youths experience during this period (Fuligni, Eccles and Barber, 1995). By providing scholars with a challenging curriculum, it is also hoped that the environmental "fit" will be better than in the schools that students attend during the school day.

HA's high school placement services are designed to ensure that these experiences do not end when students graduate from HA at the end of middle school. As noted, several studies suggest that disadvantaged families lack the skills and knowledge necessary to navigate the admissions process and the financial aid applications required to take advantage of the competitive academic opportunities that are available to them. By informing youths about the high school application process, supporting them during this process and equipping them with the skills needed to succeed in competitive high schools, HA's high school placement program strives to "even the playing field" between HA scholars and their more advantaged peers. Ultimately, HA aims to place its scholars in high school environments that will continue to provide the same kind of academic challenge and support the students received in middle school through their program.

III. Research Methods

A. Experimental Design

Many students in HA perform quite well academically. However, HA's strict admission policies screen for motivated parents and what the program defines as "academically motivated" students. While academic performance is not a criteria for admission to the program, it is of course possible that such families may succeed academically even without access to the HA program. To identify the effects of attending HA above and beyond what these families would have achieved without the program, we built the study around an oversubscription randomized controlled trial.

Each year HA recruits youths for a fixed number of available slots in their centers. During recruitment for the study, HA recruited more eligible youths than were required to fill the available positions, allowing us to randomly assign the available slots among the applicants. Specifically, interested families contacted HA, completed applications, and met with HA staff for interviews. Staff then determined the families' eligibility for the program. Eligible youths were required to attend one of several survey sessions held near each HA center[5] where they completed both the baseline standardized test and the baseline survey. Families were only allowed to participate in the lottery once—controls were not allowed to reapply for admission.

All eligible students who completed the baseline test and survey were then randomly assigned to either the treatment or control group. We conducted a separate randomization for each HA center (or ward) stratifying by grade, gender, and the problem solving score on the baseline test[6] both to minimize variance for the study and to ensure that HA received an incoming cohort of students that was balanced by gender, grade level, and academic performance.[7] Intake of subjects lasted three years with the final sample comprising 951subjects in three cohorts of students recruited annually starting in 2006. Cohorts 1, 2, and 3 contain 276, 276, and 399 subjects respectively, and given the ratio of the sample to the number of available slots each year, the ratio of treatment to control students for the first cohort was 2:1 and for the latter two cohorts, 1:1.

[5] To ensure that treatment and control students had as similar an experience as possible during survey administration, we chose not to hold the survey sessions at the HA centers. Instead they were held in a school located near the centers that was accessible to all families.

[6] Given the time required to score the standardized tests, each test was graded by our survey firm, and we stratified by the number of correct answers.

[7] To avoid imposing the burden of having only one child admitted to HA, we also ensured that siblings who applied together either were both admitted or both assigned to the control group by treating them as a single unit and stratifying by whether or not a child applied with a sibling. Siblings of children already attending HA were automatically offered a slot and were not included in the study.

Table 1 presents the data collection schedule for the evaluation. For each cohort, the table provides the timing of each survey round, the grades of the students at each round, and the overall response rates on the surveys. In this study, we focus on the effects of HA after one and two years of participation (i.e., at the first- and second-year follow-up surveys). We also investigate the changes in students' outcomes over the summer of 2010 using the data collected in the late spring and early fall of that year. As described in Section I, the final component of the study will investigate the effects of the program four years after baseline.

B. Data Collection

Follow-up and baseline surveys were organized using the same process. Subjects were sent a reminder card three to four weeks in advance of the testing sessions inviting them to the first round of testing. Subjects that failed to attend the first session where then contacted by mail and phone to notify them of subsequent testing sessions. Typically, we held three to four testing sessions for each survey round from April to June, and a small number of families unable to attend any session were tested at home. For each round, all surveys (i.e., baseline, first-year follow-up, second-year follow-up) were administered simultaneously, during the same testing session, with youths divided in groups according to their age. As compensation for their efforts, all families received a payment of $120 for completing each follow-up survey session.

The data used in this study include four main components: the original application to HA, the standardized test, a parent survey, and a youth survey. The application was completed either online or in person at the HA interview. It included an informed consent form and asked parents to provide a range of demographic and socioeconomic information. This included the child's gender, grade, age, race, and whether the child qualified for free or reduced-price lunch at

school. Parents also indicated their household income, their educational attainment, household composition, and the languages spoken at home.

For testing, the subjects completed the abbreviated versions of the Reading Comprehension and Problem Solving sections of the Tenth Edition of the Stanford Achievement Test. The tests were scored by the publisher who provided us with standardized scores normalized relative to the national distribution. Unless otherwise stated, in subsequent analysis, we used a measure of the Normal Curve Equivalent, rescaled to have a mean of zero and a standard deviation of one. These are described in Panel A of Table 2.

The parent surveys were designed primarily to collect information about youth's participation in after-school and summer programs. Parents were asked whether or not their children participated in any out-of-school time programs (both academic and non-academic) and if so, to indicate their frequency of participation. On the baseline survey, parents were also asked a series of questions designed to predict their eventual participation in the program such as the distance of the program center from their home, their plans for picking up and delivering their child to the center, the proximity of public transportation, and whether they owned a car.

The youth survey included two types of questions. For all youths, we measured students' academic attitudes and beliefs using several psychometric scales, and we asked students questions about their behavior both in and out of school. These questions are described in more detail in Panels B and C of Table 2. As a service contrast measure, we also collected detailed information about whether children engaged in specific activities usually associated with academic OST programs. These included questions about whether children received homework help outside of school, visited a college campus, etc. We use this information to contrast the experiences of treatment children to those youths in the control group and to compare the

experience of children enrolled in HA to those enrolled in other academically oriented OST programs.

C. Sample Description

Table 3 provides a descriptive tabulation of the sample used in this study. Columns one through three provide information on each cohort while column four provides the characteristics of the full sample. As shown in Panel A, the first two cohorts are the same size while the third is 45 percent larger. The cohorts are similar in the demographic characteristics listed in Panel B, but they do differ by average academic performance measures provided in Panel C. The first cohort is the weakest academically, scoring 0.4 standard deviations lower than the other cohorts in problem solving. Youths are generally evenly distributed across the centers with the fraction from Alexandria being the largest in cohort 2, the first year the center opened, due to the larger number of available slots. The only other important difference to note is the difference in the ratio of treatment-to-control students in each cohort. This difference requires the inclusion of cohort fixed effects in the analyses that follow.

The first column of Table 4 provides a summary of key baseline and demographic characteristics for all of the control students. The sample of students closely matches the types of students that HA tries to serve. As shown in the first column of Table 3, for example, 42 percent of the students entered the study as rising fifth graders, with the remainder starting before their sixth grade year. Girls constitute 59 percent of the sample. Racially, most students are African American (76 percent) with 14 percent identifying as Latino, and 66 percent of students are eligible for free or reduced-price lunches. The distribution of participants' baseline test scores is remarkably close to the national distribution. The mean of the students' scores is only slightly

higher than the national average: a difference of 0.10 standard deviations.[8] Our sample distribution is slightly more compact than the national average (the standard deviation of our students' standardized scores for reading comprehension and problems solving is 82 and 92 percent respectively of that of the national distribution). Most families earn less than $50,000 a year, though some earn more than $75,000 (14 percent). Slightly less than half of the students live with both parents, and 17 percent of families report speaking a language other than English at home.

D. Statistical Models

We use two primary models to conduct the analyses that follow. First, we use a simple difference estimator to estimate the average differences between treatment and control groups without controlling for any covariates. To do so, we estimate the following linear model using ordinary least squares:

$$y_i = \beta + \tau Treat_i + \theta' Cohort_i + \varepsilon_i \quad (1)$$

In this model, y_i is the characteristic of interest (either a baseline variable as in Tables 4 and 5 or an outcome variable as in Tables 9 and 10). The variable $Treat_i$ is an indicator variable assigned a value of one if the child is assigned to the treatment group, and $Cohort_i$ is a vector of cohort fixed effects to correct for the varying treatment assignment ratios. The estimate of τ is then the estimated difference between the treatment and control groups.

As we demonstrate in the next section, the randomization succeeded in creating balanced treatment and control groups. Because of this, adding variables to control for baseline characteristics of the students will not change the estimated differences between the treatment

[8] The test scores described in this section are normalized relative to the national distribution and expressed in terms of standardized normal scores. Thus, the national distribution has a mean of zero and a standard deviation of one.

and control groups. However, adding controls will improve the precision of the estimated treatment effects by reducing the unexplained variance in outcomes. To use this information, we estimate the following variant of equation (1):

$$y_i = \beta + \tau Treat_i + \delta' X_i + \theta' Cohort_i + \varepsilon_i \quad (2)$$

This model is identical to equation (1) except that it includes a vector of baseline characteristics X_i.[9] While we vary the included controls, our preferred specification includes the following characteristics: baseline values for problem solving, reading comprehension, self-perceptions of academic abilities, creativity, industry and persistence, enjoyment of learning, and curiosity as well as indicator variables for grade at baseline, age, income,[10] race, primary parent's education, receipt of free or reduced-price lunch, household language, and family composition.

IV. Internal Validity

A randomized controlled trial establishes internal validity by ensuring that the treatment is allocated independently of all other characteristics of the subjects. If the treatment and control groups are similar in all characteristics except that the treatment group receives the treatment, then any differences that emerge at follow-up can be causally attributed to the treatment. In our study, this requires three conditions to hold. First, the treatment and control groups should have been similar at the time of the randomization. Second, the attrition patterns in the two groups should not differ sufficiently to generate meaningful differences in the research groups. And third, because we can only offer access to Higher Achievement, we must check that enough of the treatment group did, in fact, take up the opportunity that significantly more children in the

[9] Note that because the randomization was conducted at the child level, we do not account for possible correlation between individual students in outcomes (Duflo, Glennerster, and Kremer, 2007).
[10] Income was collected on the application in ranges of $5,000 rather than as a continuous measure.

treatment group participated in HA and academic OST programs in general than children in the control group. We assess each of these conditions in turn.

A. Baseline Comparison

To assess whether the randomization did, in fact, create two similar research groups, we compare students in the treatment group with those in the control group on youth characteristics collected in the applications and the students' baseline test scores. The results are presented in the second column of Table 4. All differences are estimated using equation (1).

None of the differences are statistically significant and all of them are small in magnitude. The differences in test scores are less than a hundredth of a standard deviation, for example. To assist in judging the magnitude of these differences, columns three through six provide estimates of the coefficients on most of the baseline variables from a regression of each test score for each follow-up period on the full set of controls. Consider age, for example. On average, older applicants perform worse than younger applicants on the first-year follow-up reading test. However, because the correlation is -0.064 standard deviations per year and the average difference in age is only 0.05 years, the small imbalance in the averages ages can only generate a difference of 0.0032 standard deviations on the reading exam at follow-up.

B. Attrition

Because this study tracks children over time, we will inevitably not be able to survey all of the children who entered the study at baseline. Overall the follow-up rate is high at 85.9 percent for the first-year and 81.6 percent for the second. If the rates differ significantly by research group or if different types of students drop out of the two groups, however, then significant differences

could emerge between the remaining research groups—in which case, we would not be able to attribute any differences in these groups at follow-up to the treatment group's access to HA. We assess this possibility in Table 5.

Panel A of Table 5 shows that the attrition rates are balanced between the two research groups. On average, 84 percent of baseline students were retained from the control group for the first-year follow-up with treatment subjects being three percentage points less likely to attrit. The difference for the second-year follow-up is smaller at one percentage point on a control average of 81 percent. Neither difference is statistically significant at conventional significance levels.

Despite these similarities in rates, however, it is possible that different types of students may have attrited from each group. We assess this possibility in the remaining panels by comparing non-attriting students in the two research groups. While the differences are slightly larger than those presented in Table 4, they are still practically quite small. For example in the first year, only two of the 17 comparisons are statistically significant—the 7-percentage-point difference in youths receiving free or reduced-price lunch and the 6-percentage-point difference in students whose households report income below $25,000—but both are practically small, given the coefficients presented in columns three through six of Table 4.

Table A1 of the Appendix investigates these patterns in more detail. Overall, there seem to be some modest differences between attriting and non-attriting students. Generally, the students form the highest socioeconomic groups seem to be slightly more likely to attrit. Attritors tend to be less likely to receive free or reduced-price lunches, more likely to report incomes over $75,000, and more likely to be living with both parents. Such families may have more demands on their time that interfere with the follow-up surveys, or they may not find the $120 incentive as attractive as less-well-off families. There are also a few differences in the types of attriting

students (relative to the non-attritors) in the treatment group compared with those in the control group, but these differences do not follow a consistent pattern. As described in the previous paragraph, they are also not large enough to generate differences in the non-attriting students given the follow-up rates. Finally, all differences are more pronounced in the first-year follow-up survey than in the second-year, although the patterns are similar.

C. Treatment Differential

We cannot force youths to attend the HA program—the best we can do is to offer them admission. However, for the experiment to provide a meaningful test of the effects of the program, students offered the treatment must enroll at a sufficiently high rate that treatment youths are more likely to attend than youths in the control group.[11] To assess whether this was the case, we compare the participation rates of children in the treatment and control groups in columns one through three of Table 6. Panel A presents the results for the first-year follow-up and Panel B presents the results for the second-year follow-up.

The results suggest that the experiment succeeded in generating a large treatment differential in students' HA participation. As expected, none of the control students participated in HA. Of the treatment students, however, 87 to 88 percent of students attend HA at some point prior to the follow-up surveys, and 75 percent and 70 percent of youths are currently participating in HA at the first- and second-year follow-ups respectively.[12,13] On average, by the

[11] Students assigned to the control group were not allowed to enroll in Higher Achievement and we had no instances of children being erroneously admitted to the program despite being assigned to the control group.

[12] This information is based on self-reported participation rates from the parent surveys. We were able to obtain administrative participation records from Higher Achievement for the summer of 2010 (records were not retained for other periods). These records matched the self-reported rates very closely. Of the 444 subjects who completed our survey in fall 2010, 97.3 percent reported their summer participation consistently with the administrative records. Seven claimed to have attended HA without HA having a record of their participation, and five claimed not to have attended HA but were recorded as having attended by the program.

first follow-up, treatment students experienced a total of 8.9 hours a week of HA during the academic year.

While control students cannot attend HA, they can attend other academically oriented OST programs. Columns four through six of Table 6 compare the treatment and control groups in their participation rates in other academic OST programs. Overall, their participation rates are fairly similar. There is a small difference in that fewer treatment children report ever having attended a non-HA academic OST program than controls, but for those who continue to attend the programs at follow-up, the rates are virtually the same, and the overall averages of participation differ by slightly more than a half an hour, a difference that is not statistically significant.

Columns seven through nine of Table 6 estimate the overall participation rates in any academically oriented OST program, and as one would expect from the previous results, students in the treatment group are much more likely to have attended an academically oriented OST program than youths in the control group. At each follow-up, 54 and 50 percentage points more children in the treatment group are attending an academic OST program, and students in the treatment group experience an average of 10.3 and 8.5 more hours a week of academic OST instruction than children in the control group.

The similarities in non-HA academic OST participation are important for the interpretation of our results. The fact that the rates do not differ between the two groups means that the difference in outcomes results from students attending HA who otherwise would not have attended any academic OST program. This rules out the possibility, for example, that

[13] It is important to note that while these are very high participation rates for the sample of students that we were able to survey, these are not the overall retention rates for the program. This is the rate for students who did not attrit from the study, and attriting students were much less likely to participate in HA than non-attriting students. In the summer of 2010, we estimate the two-year retention rate for cohort 3 to be 47.7 percent using the administrative data from HA, and the three-year retention rate for cohort 2 to be 44.1 percent.

admissions to HA simply changed which OST program a child attended rather than whether the child attended an academic OST program. If this had been the case, then we would observe children in the control group being much more likely than children in the treatment group to be attending a non-HA academic OST program. This also would have changed the interpretation of our results because the comparison would have been between attending different types of OST programs rather than between attending HA and attending no academic OST program.

The fact that some children attended other academic OST programs also provides us with an opportunity to compare youths' experiences in HA to their experience in other programs to better understand the characteristics that define the HA treatment. As shown in Table 7, one of the biggest differences is the intensity of the experience. For each survey round, we compare the average time spent in an academic OST program in the treatment group to that in the control group, and then, just using students who were actively participating in an academic OST program at follow-up, we compare the time spent in any academic OST program to time spent in HA. As we show in the previous table, treatment students spend an average of ten more hours a week in an academic OST during the academic year as of the first-year follow-up. During the summer, this difference increases to 20 hours a week. Both comparisons are statistically significant at the one-percent level. These differences are partly due to the treatment students spending 1.5 to 2.2 days more a week in such a program in the academic year and summer respectively. These differences drop slightly as of the second year, but they are of a similar magnitude and still statistically significant at the one-percent level.

Comparing time spent in HA relative to that spent in academic OST programs, the differences are smaller, but HA is still more intense. During the academic year, students spend fewer days in HA than in other academic programs, but they spend sufficiently more time in HA

each day that, on average, they spend 2.7 and 4.2 more hours a week in HA at the first- and second-year follow-ups respectively. During the summer, however, they spend more days a week and 16 more hours a week in HA than in other academic OST programs for both survey periods.

At each spring follow-up, we also asked students what types of academic and enrichment activities they engaged in during the previous year. These results are presented in Table 8 for the first-year follow-up. The results for the second year are similar and presented in Table A2 in the Appendix. With the exception of helping other kids with their school work (row six), treatment students report engaging in all of these activities more frequently than control students, and ten of the fifteen differences are statistically significant—eight at the one-percent level. Given the number of outcomes tested, we conducted a simultaneous test of the significance of all of the differences using a Seemingly Unrelated Regressions model and find the overall test to be statistically significant at the one-percent level.

Looking more closely at the experiences of students by the type of OST program in which they participated (presented in the last three columns of Table 8), these differences seem to be driven primarily by the experiences of children in Higher Achievement. Overall, the experiences of youths enrolled in other academic OSTs are strikingly similar to those of the students who are not attending an OST. The only difference is the fraction of students who report having visited a business or organization. However, compared with students who attend other academic OSTs, students in HA report experiencing more of these activities, and as with the overall treatment difference, the joint test of all of the differences is statistically significant at the one-percent level.

V. Impact Estimates

V.A. Test Scores

One of the main goals of the Higher Achievement program is to improve students' academic achievement. To test whether the program is reaching that goal, we estimate the effects of the program on students' test scores on the Problem Solving and Reading Comprehension sections of the Tenth Edition of the Stanford Achievement Test. We find that the program had little effect in the first year, but by the second year, it significantly improved students' scores on both tests.

The results are presented in Table 9. Panel A presents the results for the first-year follow-up survey and Panel B presents the results for the second year. In columns one through three, each cell presents the estimated treatment effect for the indicated test (Reading Comprehension in the first row and Problem Solving in the second) using equations (1) and (2) including the indicated control variables. The final row of each panel presents the results of a joint test of the significance of both treatment effects by simultaneously estimating both effects using Seemingly Unrelated Regressions.

Starting with column one in Panel A, it is clear that the program has little effect on students' test scores after a year of program involvement. The estimated differences for Reading Comprehension are 0.04 standard deviations and for Problem Solving are 0.05, neither of which is statistically significant at conventional levels. Emphasizing the comparability of the research groups, adding the additional control variables in columns two and three has little effect on the estimated effects. Including them only reduces the estimated effects, and in our preferred estimate in column three the effect on both tests is estimated to be only 0.01 standard deviations.

The effects in the second year, however, are positive and statistically significant. Turning to Panel B, the estimated effects are statistically significant for both tests at the ten-percent level

without including any control variables (column one). Adding the additional controls, improves the precision of the estimates. In our preferred specification, we find an effect of 0.09 standard deviations on students' reading scores (significant at the ten-percent level) and 0.12 standard deviations on their problem solving scores (significant at the five-percent level).

Because not all children in the treatment group eventually enrolled in HA at some point during the evaluation periods, we can estimate a local average treatment effect to determine the effect of having ever attended HA relative to the other options available to students. To do this, we estimate equation (2), including an indicator variable for whether the child ever participated in HA during the previous evaluation period in place of the treatment indicator. We then instrument for this participation measure with the treatment indicator using Two-Stage Least Squares. These results are presented in column four of Table 9. All of the estimates are very close to the intent-to-treat estimates presented in column three, presumably because of the very large proportion of treatment children who report having ever been enrolled in HA, as shown in Table 6.

Table 6 also demonstrated that a number of children in both the treatment and control groups reported attending academic OST programs other than Higher Achievement. These programs may have had an effect on children's test scores as well. Thus, in column five of Table 9 we estimated an average effect of academic OST participation using assignment to the treatment group as an instrument. We used the same methodology that we used in column four, but we include an indicator for whether the child has ever attended an academic OST program in place of the indicator of participation in HA. Because such a significant portion of the control group had been enrolled in another academic OST program, the estimated effects are larger than those in column four. We find that the average effect of an academic OST program is to change

students' reading comprehension scores by 0.16 standard deviations and their problem solving scores by 0.21 standard deviations, significant at the ten and five-percent level respectively.

In results not presented in this manuscript, we conducted two other sets of analyses to better understand the effects on test scores. First, we estimated treatment effects for a range of subgroups including divisions by baseline test scores, gender, entering grade level, ward, free or reduced-price lunch status, and HA center of application. We find no statistically significant differences in the impacts on any of these groups for either follow-up survey. Second, we attempted to use information collected on the baseline survey about how accessible the local HA center was for families[14] as a predictor (along with our other control variables) of later participation in the program. We find that the model predicts participation fairly well, but we do not find differences in the treatment effect that are correlated with projected participation.[15]

V.B. Academic Attitudes and Peer and Adult Support

In addition to test scores, we also measured students' attitudes toward school, their self-perceptions, their experiences of peer and adult support and other factors that might influence their success. The estimated treatment effects for these outcomes are presented in Table 10. We estimate the results using the same three sets of control variables that we used when conducting the test score analyses presented in Table 9. We present the results for the first-year follow-up survey in the first three columns followed by the estimates for the second-year follow-up in the last three columns. Estimated effects on academic attitudes are presented in Panel A, and the effects for peer and adult support are presented in Panel B. For each set of outcomes, we also

[14] This included questions such as the distance of the center to the child's home, whether the family had access to a car, who would be responsible for sending/taking a child to/from the center, etc.

[15] It is important to note that this is not a test of whether or not the intensity of participation affects test scores. There is no source of random variation in participation levels within this experiment. Rather, this is a test of heterogeneity in outcomes along the dimensions correlated with higher levels of performance.

present the results of a joint significance test for all outcomes estimated through Seemingly Unrelated Regressions.

Quite surprisingly we observe that, rather than improving students' perceptions of their academic abilities, the program seems to have instead caused setbacks in these variables in the first year. Again, the point estimates are consistent across the different specifications, but with the exception of ability to change the future through effort, we observe negative treatment effects of between -0.11 and -0.18 standard deviations, all of which are statistically significant at the ten-percent level or higher. The joint significance test is also statistically significant at the five-percent level. However, we find that this effect disappears in the second year; we observe no treatment effects on attitudes in the second-year follow-up survey.

This negative treatment effect in the first year seems to be solely driven by the students entering the lottery as rising fifth graders. Figure 1 shows the average for the six outcomes for each survey period by grade and research group. Because we are investigating changes over time, we normalize the follow-up outcomes relative to the baseline control group, rather than the contemporaneous control group distribution as in Table 10. The short dashed lines are the rising fifth graders and the long-dashed lines are the rising sixth graders. The lines with dots are the treatment group and those without dots are the control group.

From this figure, it is clear that the experiences of the treatment fifth graders are perfectly consistent with those of the sixth graders in the treatment group. Entering grade does not seem to be correlated with different trajectories of outcomes for those youths in the treatment group. The difference is the counter-factual experiences for each grade level. Control sixth graders experience a slight decline in outcomes during the sixth-grade year, consistent with the experiences of youth in general as they enter middle school (Eccles and Midgley, 1989). Fifth-

grade control students, however, are experiencing the last year of elementary school and their perceptions actually improve slightly at the first-year follow-up, but then decline sharply in their first year of middle school. By placing fifth-grade students in a more competitive environment, the HA program seems to have caused the middle-school decline a year before these students would otherwise experience it, causing them to experience a change in attitudes very similar to the older students in the study.

Table 11 presents numerical estimates of these patterns. To replicate the functional forms presented in Figure 1, we stack the dataset so that we can estimate changes in the overall average academic attitude over each period. Estimates are performed at the survey-round level and the standard errors are clustered by child to account for correlation in the performance of an individual child over time. Columns one and two provide estimates of the difference in scores relative to baseline by grade for the control and treatment groups respectively. As shown in Figure 1, fifth graders experience a slight increase in the first year and sixth graders experience a slight decrease (neither of the changes though is statistically significant at conventional levels), but both experience much larger declines in the second year. For the treatment group, fifth and sixth graders both experience declines of about a tenth of a standard deviation in both years, and the differences between the two grades are not statistically significant.

Columns three through five in Table 11 demonstrate how this difference generates the observed negative impact in the first year. Column three presents the estimate for the full sample showing the statistically significant negative effect in the first year (row six). Column four provides the estimates for only the fifth graders, resulting in a negative effect of -0.16 standard deviations (statistically significant at the one-percent level). And column five shows that sixth graders experience no difference in outcomes in either follow-up period.

Finally, Table 12 presents the estimates by grade for the overall average and individual academic attitudes, as well as test scores, using equation (2) and dividing the treatment effect by grade. As in Table 11, the outcomes are normalized relative to the baseline control distribution,[16] but the statistical analysis is run at the child level. Panel A presents the results for the first year, showing the negative effect for fifth graders and no effect for sixth graders. As shown in the last row, the treatment effect by grade for the overall average is statistically significant at the one-percent level, and the difference is statistically significant for three of the six individual outcomes at the ten percent level or better. In Panel B, we see that the magnitude of the effect for fifth graders diminishes for the second year with an overall difference of -0.07 standard deviations. Entering fifth graders still have weaker perceptions than their peers when compared to the difference for entering sixth graders, presumably because they have experienced a more competitive academic environment for a year longer than their entering fifth-grade peers in the control group. For entering sixth graders, we find that the program actually has positive effects on self-perceptions of academic abilities (0.30 standard deviation, significant at the one-percent level). The joint test over all six outcomes is statistically significant at the five-percent level as well, indicating that for students who started the program as rising sixth graders, the program eventually improves their self-perceptions above what they would have been without the experience of Higher Achievement.

Importantly, the accelerated decline in rising-fifth-graders' attitudes does not seem to be related to their academic achievement. The final two columns in Table 12 provide estimates of the effect of the program on test scores by entering grade level. In the first year, we find no effects on either group of students. At the second-year follow-up, students who entered the

[16] We have estimated the same models with the outcome normalized relative to the contemporaneous control group distribution. The results are virtually identical.

program as fifth graders experience gains on the math test while those who started as sixth graders experience gains on the reading test.[17]

Finally, we turn back to the adult and peer support outcomes in Table 10. We find no evidence of an effect on students' perceptions of the amount of support they receive from adults or peers despite the fact that HA offers the assistance of mentors and instructors and a cohort of academically inclined peers. In the first year, the effect on their perceptions of the influence of their friends is negative and of adult support is positive, but neither is statistically significant at conventional levels. The effects in the second year are both positive, but again statistically insignificant. The joint tests in both rounds are also statistically insignificant.

V.C. Behavior

In Table 13, we estimate the effects of HA on students' self-reported behavior. Panel A presents the results for whether a child reports ever engaging in general types of misconduct (e.g. stealing, hitting someone, etc.) within the last three months. Panel B then presents the measures of school-related misconduct. Joint tests of the significance of all treatment effects are presented in Panel C. Results for the first-year follow-up are presented in the first three columns followed by the results for the second-year follow-up in the last three columns.

Overall, the experience of the HA program increases the rate at which children report engaging in many of these negative behaviors in both follow-up periods. The probability of taking something that did not belong to the child increases by 14 percentage points in the first year and 19 in the second year, statistically significant at the ten- and five-percent levels

[17] In results not presented in this manuscript, we also conducted a mediation analysis (Baron and Kenny, 1986) in which we assessed whether the observed changes in academic attitudes mediated the eventual changes in test scores. Despite testing several different specifications for the mediating variables, we find that they do not explain the increases in test scores. They are positively correlated with test scores, but the treatment effect on test scores seems to occur independently of these changes in attitudes.

respectively. And the probability of being sent to the principal's office increases by 22 percentage points in both rounds. The joint tests are statistically significant at the five-percent level in the preferred specification for the first year and at the ten-percent level for the second year. This is clearly an important area for future investigation. Because these are self-reported measures, they could reflect an actual increase in negative behavior, but they also could reflect differences in reporting practices. These effects seem at odds with the observed changes in test scores, but given the consistency of the results, they are difficult to dismiss.

V.D. High School Preparation

To begin assessing the high school preparation component of the Higher Achievement program, the survey administered in fall 2010 to students in cohort 3 and those in cohort 2 who entered as rising fifth graders contained a series of questions asking students if they had engaged in several different activities over the previous summer related to applying to high school. The treatment effect estimates are presented in Table 14. Given the high level of motivation of this sample of students as evidenced by their engaging in academically oriented OST programs even without access to HA, it is perhaps not surprising that the control students reported engaging in a number of activities related to high school choice including talking with their parents (69 percent) or other adults (51 percent) about high schools, and deciding where to apply (46 percent). Yet, HA significantly increased the likelihood that students reported engaging in a number of these activities above and beyond those reports by the controls. Namely, they were more likely to visit a high school, get information about a high school, talk with non-parental adults and peers about high schools, and decide where to apply. The joint test on all of these activities is statistically significant at the one-percent level.

V.E. Changes in Outcomes, Summer 2010

As described in Section II, many students experience a decline in their academic performance during the summer. To evaluate the effects of HA over a summer period, we surveyed those students who were eligible for the HA summer program in the summer of 2010 (i.e., they had not yet begun the eighth grade). This included all students in cohort 3 and those in cohort 2 who applied as rising fifth graders. The results are presented in Table 15. We evaluate the effects of the program on students' test scores (Panel A), academic attitudes (Panel B), perceptions of external support (Panel C), and high school preferences (Panel D). We present treatment-control differences measured in the spring in columns one and two. Columns three and four present treatment-control differences measured in the fall, and columns five through seven present the relative changes in scores from spring to fall for the two groups. Column six presents the impact estimates from a differences-in-differences specification while column seven presents the results from a specification like equation (2) but including the spring scores as independent variables. It is important to note that these latter specifications are not strictly identified because treatment students enter the spring of 2010 having had two or three years of HA programming compared to the controls.

Starting with test scores in Panel A, we find that as of the spring survey, treatment students (who had had the opportunity to participate in HA for two to three years) were already experiencing statistically significant treatment effects of 0.11 standard deviations on the problem solving test and 0.16 standard deviations on the reading comprehension test. In the fall, we find that, rather than experiencing a decline in test scores, the control students' reading comprehension scores increased to 0.12 standard deviations while their problem solving scores remained roughly stable. The treatment group's reading scores also increased, but not as much as

those for the control group, resulting in a decline in the treatment effect to a statistically insignificant 0.03 standard deviations. The treatment effect for problem solving is slightly smaller and also no longer statistically significant. As a result, we see no relative differences in the changes in scores between treatment and control students presented in either of the specifications in columns six and seven. The results are very similar for academic attitudes and peer and adult support, except that the initial differences in the spring are not statistically significant for these outcomes.

Youths' high school preferences, however, do show a marked change. In the spring, nine percentage points more students in the treatment group express a desire to attend a competitive area high school compared to the control group. In the fall, this increases to 17 percentage points; and ten percentage points fewer students express a desire to attend their local public school. In the differences-in-differences specifications, we find that the relative changes are of similar magnitudes and statistically significant at the one-percent level. With the exception of the spring difference, the joint test on the treatment effects are all statistically significant at the one-percent level.

VI. Conclusion

The study presented in this manuscript documents the ability of a mature, comprehensive, and long-term academically oriented OST program to improve students' academic performance during their middle school years. HA increases students' problem solving and reading comprehension scores, and significantly alters their perceptions of their own academic abilities. We also find evidence that the program increases students' desire to attend a competitive area high school and increases the probability that students have taken the steps to do so. More

research is needed on the effects of such programs on students' behavior. Troublingly, we find that the program negatively affects students' self-reported misbehavior, although because our measures are self-reported this finding could simply reflect a change in the way that students respond to the questionnaire.

VII. References

Afterschool Alliance. (2011) "Senators Call for 'Significant Investment' in 21st CLCC," *Afterschool Advocate*, 12:6. http://www.afterschoolalliance.org/advocate.cfm?story_id=4001561 (Accessed July 26, 2011).

Anderman, E. M. and M. L. Maehr. (1994) "Motivation and Schooling in the Middle Grades," *Review of Educational Research*. 64: 287-309.

Banerjee, A., S. Cole, E. Duflo and L. L. Linden. (2007) "Remedying Education: Evidence from Two Randomized Experiments in India," *Quarterly Journal of Economics*. 122(3): 1235-1264.

Baron, R. B. and D. A. Kenny. (1986) "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*. 51: 1173-1182.

Beckett, M., G. Borman, J. Capizzano, D. Parsley, S. Ross, A. Schirm and J. Taylor. (2009) *Structuring Out-of-School Time to Improve Academic Achievement: A Practice Guide* (NCEE #2009-012). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides.

Black, A. R., F. Doolittle, P. Zhu, R. Unterman and J. B. Grossman. (2008) *The Evaluation of Enhanced Academic Instruction in After-School Programs: Findings After the First Year of Implementation* (NCEE 2008-4021). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Blyth, D. A., R. G. Simmons, and S. Carlton-Ford. (1983) "The Adjustment of Early Adolescents to School Transitions," *Journal of Early Adolescence.* 3 (1–2): 105–120.

Campbell, J. R., C. M. Hombo and J. Mazzeo. (2000) *NAEP 1999 Trends in Academic Progress: Three Decades of Student Performance*. National Center for Educational Statistics, Office of Educational Research and Improvement, U.S. Department of Education. NCES 2000-469.

Chaplin, D. and J. Capizzano. (2006) "Impacts of a Summer Learning Program: A Random Assignment Study of Building Educated Leaders for Life (BELL)," *Urban Institute Report*. http:// www.urban.org/url.cfm?ID=411350 (Accessed August 3, 2011).

Durlak, J. and R. Weissberg. (2007) *The Impact of After School Programs that Promote Personal and Social Skills.* Chicago, IL: Collaborative for Academic, Social, and Emotional Learning (CASEL).

Durlak, J., R. P. Weissberg and M. Pachan. (2010) "A Meta-Analysis of After-School Programs That Seek to Promote Personal and Social Skills in Children and Adolescents," *American Journal of Community Psychology*. 45: 294-309.

Eccles, J. S. and C. Midgley. (1989) "Stage-Environment Fit: Developmentally Appropriate Classrooms for Young Adolescents," in C. Ames and R. Ames (Eds.), *Research on Motivation in Education: Goals and Cognitions* (Vol. 3, pp. 139-186). New York: Academic Press.

Eccles, J., M. Vida and B. Barber. (2004) "The Relation of Early Adolescents' College Plans and Both Academic Ability and Task-Value Beliefs to Subsequent College Enrollment," *Journal of Early Adolescence*. 24: 63–77.

Eccles J., and A. Wigfield. (2000) "Schooling's Influences on Motivation and Achievement," in S. Danziger & J. Waldfogel (Eds.) *Securing The Future: Investing in Children from Birth to College.* New York: Russell Sage Foundation.

Eccles, J., A. Wigfield, R. Harold and P. Blumenfeld. (1993) "Age and Gender Differences in Children's Self- and Task Perceptions during Elementary School," *Child Development.* 64: 830-847.

Fuligni, A., J. Eccles, and B. Barber. (1995) "The Long-Term Effects of Seventh Grade Ability Grouping in Mathematics," *Journal of Early Adolescence*. 15: 58-89.

Gennetian, Lisa A., C. Miller and J. Smith. (2005) *Turning Welfare into a Work Support: Six Year Impacts on Parents and Children from the Minnesota Family Investment Program.* New York: Manpower Demonstration Research Corporation.

Duflo, E., R. Glennerster, and M. Kremer. (2007) "Using Randomization in Development: A Tookit," Center for Economic Policy Research, Working Paper No. 6059.

Goodenow, C. (1993) "Classroom Belonging Among Early Adolescent Students' Relationships to Motivation and Achievement," *Journal of Early Adolescence,* 13: 21-43.

Grossman, J. B. and J. E. Rhodes. (2002) "The Test of Time: Predictors and Effects of Duration in Youth Mentoring Relationships," *American Journal of Community Psychology.* 30(2): 199–219.

Gutman, L. M. and C. Midgley. (2000) "The Role of Protective Factors in Supporting the Academic Achievement of Poor African American Students During the Middle School Transition," *Journal of Youth and Adolescence.* 29: 223-248.

Hamre, B. K. and R. C. Pianta. (2001) "Early Teacher-Child Relationships and the Trajectory of Children's School Outcomes Through Eighth Grade," *Child Development*. 72:625-38.

Hoxby, C. M. (2011) "Economics of Education," *NBER Reporter*, 2011:(1): 1-5.

James-Burdumy, S., M. Dynarski and J. Deke. (2007) “When Elementary Schools Stay Open Late: Results from the National Evaluation of the 21st Century Community Learning Centers Program,” *Educational Evaluation and Policy Analysis.* 4: 296–318.

Johnson, G. E. and F. P. Stafford. (1973) “Social Returns to Quantity and Quality of Schooling,” *The Journal of Human Resources.* 8(2), 139-155.

Kling, J. and J. Liebman. (2007) "Experimental Analysis of Neighborhood Effects on Youth," *Econometrica*. 75(1): 83-119.

Lauer, P., M. Akiba, S. Wilerson, H. Apthorp and D. Snow. (2006) “Out-of-School Time: A Meta-Analysis of Effect for At-Risk Students,” *Review of Educational Research.* 76: 275–313.

Mahoney, J. L. and E. F. Zigler. (2006) “Translating Science to Policy Under the No Child Left Behind Act of 2001: Lessons from the National Evaluation of the 21st-Century Community Learning Centers,” *Journal of Applied Developmental Psychology*. 27(4) 282–294.

Morgan, J. and I. Sirageldin. (1968) “A Note on the Quality Dimension in Education,” *Journal of Political Economy. 76*(5): 1069-1077.

Neal, D. (1997) “The Effects of Catholic Secondary Schooling on Educational Achievement,” *Journal of Labor Economics. 15*(1): 98-123.

Neal, D. (2006) "Why Has Black-White Skill Convergence Stopped?" in E. Hanushek and F. Welch (Eds.) *Handbook of Economics of Education.* Elsevier.

Neal, D., J. Grogger, E. A. Hanushek and R. M. Schwab. (2000) “Further Evidence on the Effects of Catholic Secondary Schooling,” *Brookings-Wharton Papers on Urban Affairs.* 151-201.

Research for Action. (2010) *Uneven Playing Field: Demographic Differences and High School “Choice” in Philadelphia*. Policy Brief. March 2010.

Sattin-Bajaj, C. (2009) “Informing Immigrant Families about High School Choice in New York City: Challenges and Possibilities,” *New York University Prepared for School Choice and School Improvement: Research in State, District and Community Contexts*. Vanderbilt University, October 25–27, 2009.

Scott-Little, C., M. S. Hamann and S. G. Jurs. (2002) “Evaluations of After-School Programs: A Meta-Analysis of Methodologies and Narrative Synthesis of Findings,” *American Journal of Evaluation.* 23: 387–419.

Seidman, E., L. Allen, J. L. Aber, C. Mitchell and J. Feinman. (1994) "The Impact of School Transitions in Early Adolescence on the Self-System and Social Context of Poor Urban Youth," *Child Development*. 65: 507-522.

Simmons, R.G. and D. A. Blythe. (1987) *Moving into Adolescence: The Impact of Pubertal Change and School Context*. Hawthorne, NY: Aldine de Gruyter.

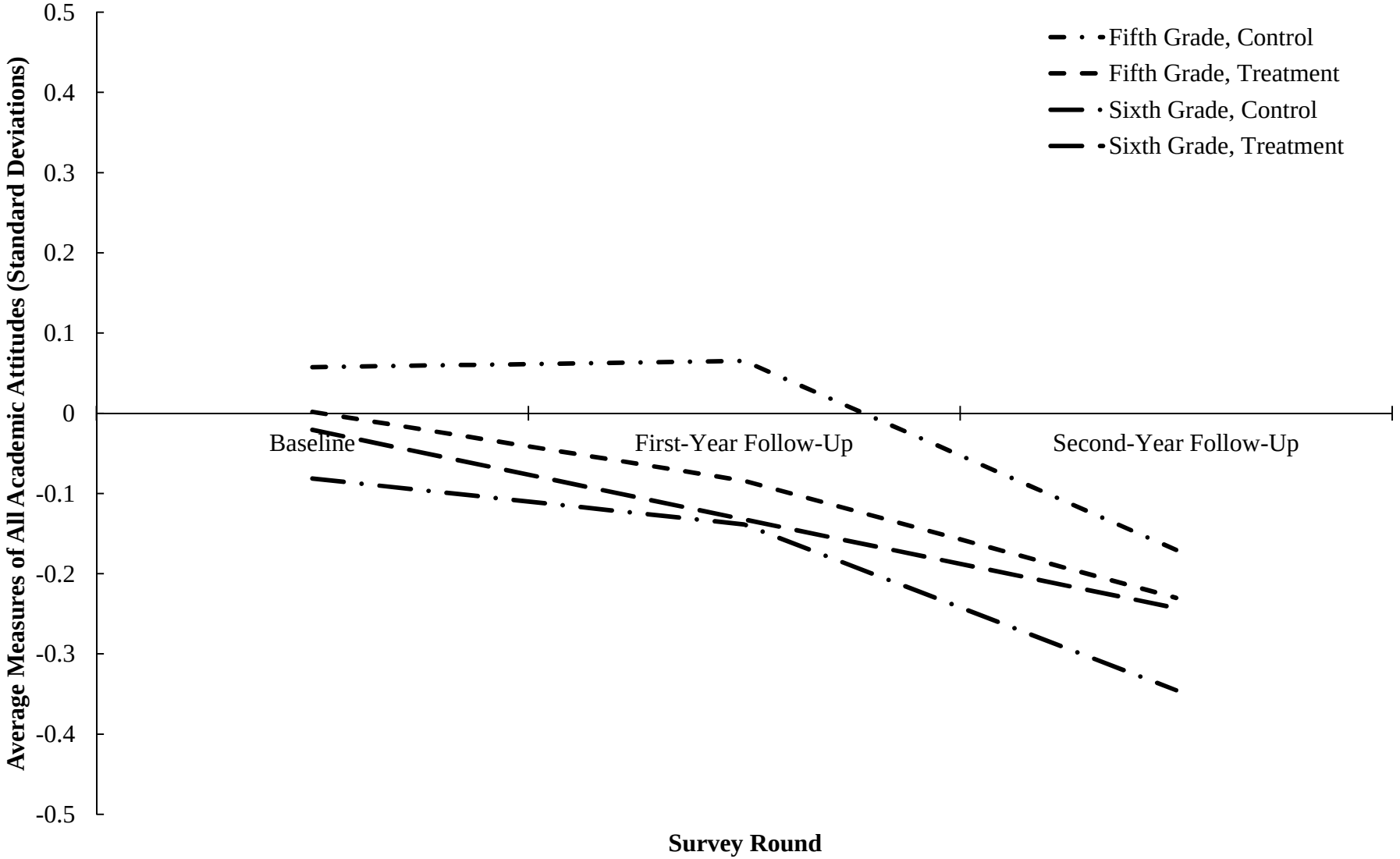
Stern, D. (1989) "Educational Cost Factors and Student Achievement in Grades 3 and 6: Some New Evidence," *Economic of Education Review. 8*(2): 149-158.

Stipek, D. (2002) *Motivation to Learn: Integrating Theory and Practice*. Boston: Allyn and Bacon.

U.S. Department of Education. (2011) "21st Century Community Learning Centers: Program Description" http://www2.ed.gov/programs/21stcclc/index.html (Accessed July 28, 2011).

Wigfield, A., J. Eccles, D. Mac Iver, D. Reuman, and C. Midgley. (1991) "Transitions at Early Adolescence: Changes in Children's Domain-Specific Self-Perceptions and General Self-Esteem across the Transition to Junior High School," *Developmental Psychology*. 27: 552-565.

Figure 1: Average Academic Attitudes by Survey Round



Note: This figure presents the average of all six measures of academic attitudes by entering grade and treatment assignment for each survey period.

Table 1: Schedule of Survey Activities

	2006 Spring	2007 Spring	2008 Spring	2009 Spring	2010 Spring	2010 Fall	2011 Spring	2012 Spring
Cohort 1 (N=276)								
Survey Round	Baseline	FU1	FU2		FU4			
Grade Entering	5th/6th	6th/7th	7th/8th		9th/10th			
Cohort 2 (N=276)								
Survey Round		Baseline	FU1	FU2	FUSp	FUFa	FU4	
Grade Entering		5th/6th	6th/7th	7th/8th	8th	8th	9th/10th	
Cohort 3 (N=399)								
Survey Round			Baseline	FU1	FU2/FUSp	FUFa		FU4
Grade Entering			5th/6th	6th/7th	7th/8th	7th/8th		9th/10th

Note: This table depicts the schedule of baseline and follow-up surveys. Students were recruited at the start of the summer before either their fifth or sixth grade years. FU1, FU2, and FU4 designate the first-, second-, and fourth-year follow-up surveys respectively. FUSp and FUFa reflect surveys conducted before and after (end of spring beginning of fall) the summer of 2010. The surveys FUSp and FUFa included only eighth grade students from Cohort 2 because the ninth grade students had aged-out of the HA program.

Table 2: Itemization of Outcome Variables

Youth Survey Outcomes	Title of Measure	Author(s)	Sample Items	Items	Alpha Follow-Up 2
Panel A: Academic Performance					
Reading Comprehension	The Stanford Achievement Test, 10th edition	Pearson Education Inc.	*Proprietary*	30	--
Problem Solving	The Stanford Achievement Test, 10th edition	Pearson Education Inc.	*Proprietary*	30	--
Panel B: Academic Attitudes					
Industry and Persistence	Industry subscale from Values in Action Youth Survey	Peterson and Seligman, 2004	*Proprietary*	9	0.77
Creativity	Creativity subscale from Values in Action Youth Survey	Peterson and Seligman, 2004	*Proprietary*	8	0.79
Enjoyment of Learning	Learning subscale from Values in Action Youth Survey	Peterson and Seligman, 2004	*Proprietary*	7[a]	0.76
Curiosity	Curiosity subscale from Values in Action Youth Survey	Peterson and Seligman, 2004	*Proprietary*	8	0.75
Ability to Change the Future through Effort	RAPS Manual	Institute for Research and Reform in Education, 1998	If I get bad grades, it's because I didn't try hard enough.	6	0.76
School Liking	Adapted from a scale tested with middle-school youth	Jacque Eccles	In general, I like school a lot.	3	--[b]
Self-Perceptions of Academic Abilities	Adapted from the Manual for the Self-Perception Profile for Children	Harter, 1985	I do very well at my classwork.	5[a]	0.62
Panel C: Peer and Adult Support					
Adult Support	Adult Support (subset of original items)	Gambone & Arbreton, 1997	How many adults who are not relatives…could you talk to about personal problems?	5	0.82
Academically Supportive Friends	PALS	Midgley et al., 2000	My friends try to get me to do my best in school.	5	0.83

Note: This table presents a summary of the variables collected for each category of outcome variables.

[a] One item was dropped from the original scale to improve reliability.

[b] School liking was only administered in the spring and fall 2010 survey administrations and thus was not included in all follow-up 2 administrations

Table 3: Tabulation of Subject Characteristics

Student Characteristic	Cohort 1	Cohort 2	Cohort 3	All
Panel A: Sample Composition				
Number of Students	276	276	399	951
Percentage of Students	0.29	0.29	0.42	1
Panel B: Demographic Characteristics				
Female	0.57	0.62	0.59	0.59
Grade	5.47	5.42	5.39	5.42
Age at Application	9.84	9.85	9.88	9.86
Panel C: Normalized Test Scores				
Reading Comprehension	-0.1	0.23	0.1	0.08
	(0.86)	(0.79)	(0.75)	(0.80)
Problem Solving	-0.21	0.19	0.19	0.07
	(0.92)	(0.90)	(0.90)	(0.92)
Panel D: Center				
Ward 1	0.26	0.12	0.14	0.17
Ward 4	0.2	0.15	0.19	0.18
Ward 6	0.35	0.2	0.2	0.24
Ward 7	0.19	0.22	0.2	0.2
Alexandria	0	0.32	0.26	0.2
Panel E: Treatment-to-Control Assignment Ratio				
Student Assigned to Treatment	0.67	0.5	0.5	0.55

Note: This table provides a tabulation of the students in our sample. Data are provided for each cohort in columns one through three and for the overall sample in column four. Panel A provides the breakdown of the sample across cohorts. Panel B provides the average demographic characteristics, and Panel C provides the mean and standard deviation of the baseline normalized test scores relative to the national reference sample. Panel D provides the breakdown by the HA center to which each child applied to attend, and Panel E provides the probability that a child was assigned to the treatment group.

Table 4: Baseline Comparison of Student Characteristics by Research Group

			Estimated Correlation for Controls:			
			Reading Comprehension		Problem Solving	
Student Characteristic	**Control Average**	**Treatment-Control**	**FU1**	**FU2**	**FU1**	**FU2**
Panel A: Demographic Characteristics						
Female	0.59	< 0.01	0.080*	0.06	-0.061	0.035
	(0.49)	(0.03)	(0.046)	(0.049)	(0.047)	(0.050)
Grade	5.42	< 0.01	-0.106*	-0.038	-0.136**	0.052
	(0.49)	(0.03)	(0.054)	(0.058)	(0.055)	(0.058)
Age	9.83	0.05	-0.064**	-0.077**	-0.051	-0.102***
	(0.78)	(0.06)	(0.031)	(0.033)	(0.032)	(0.033)
African-American	0.76	-0.02	-0.007	-0.016	0.024	0.089
	(0.43)	(0.03)	(0.068)	(0.073)	(0.069)	(0.074)
Latino	0.14	< 0.01	-0.025	0.075	-0.135	0.061
	(0.34)	(0.02)	(0.105)	(0.111)	(0.108)	(0.113)
Free or Reduced-Price Lunch	0.66	-0.05	-0.035	-0.061	0.054	0.043
	(0.47)	(0.03)	(0.052)	(0.056)	(0.053)	(0.057)
Panel B: Normalized Test Scores						
Reading	0.10	< 0.01	0.430***	0.492***	0.219***	0.175***
	(0.82)	(0.05)	(0.038)	(0.040)	(0.039)	(0.041)
Problem Solving	0.10	< 0.01	0.333***	0.263***	0.685***	0.578***
	(0.92)	(0.06)	(0.034)	(0.036)	(0.034)	(0.037)
Panel C: Household Characteristics						
Total Household Income:						
Below $25,000	0.32	-0.03	-0.056	-0.025	-0.033	-0.129*
	(0.47)	(0.03)	(0.069)	(0.074)	(0.070)	(0.074)
$26,000-$50,000	0.38	0.03	-0.008	-0.03	0.037	-0.045
	(0.49)	(0.04)	(0.063)	(0.068)	(0.064)	(0.069)
$51,000-$75,000	0.16	0.03	-0.036	-0.022	-0.066	-0.121
	(0.37)	(0.03)	(0.080)	(0.087)	(0.082)	(0.088)
Over $75,000	0.14	-0.04	-0.083	-0.043	-0.069	-0.136
	(0.35)	(0.02)	(0.097)	(0.104)	(0.099)	(0.105)
Parent Education:						
College Degree	0.33	-0.01	-0.008	0.071	0.039	0.112*
	(0.47)	(0.03)	(0.058)	(0.062)	(0.059)	(0.063)
Some College	0.31	< 0.01	0.042	0.136**	0.032	0.05
	(0.46)	(0.03)	(0.056)	(0.060)	(0.058)	(0.061)
Living with Both Parents	0.41	-0.04	0.038	-0.025	0.003	-0.03
	(0.49)	(0.04)	(0.055)	(0.058)	(0.056)	(0.058)
Non-English Language	0.17	< 0.01	-0.063	-0.078	0.182*	0.106
Spoken at Home	(0.37)	(0.03)	(0.102)	(0.106)	(0.104)	(0.107)
Constant			1.096***	0.789**	1.107***	0.640**
			(0.293)	(0.309)	(0.299)	(0.313)
Observations			800	759	800	759
R-squared			0.48	0.47	0.6	0.49

Note: This table provides a comparison of all students assigned to either the treatment or the control group using information collected either on the HA application or the baseline survey. Column one contains the average characteristics of the control group. Column two provides the estimated difference between the treatment and control group using equation (1). Columns three through six contain estimates from a linear regression of the respective follow-up test score on the listed baseline and demographic characteristics using onl the students assigned to the control group. Significance at the ten-, five-, and one-percent levels are indicated by *, **, and *** respectively.

Table 5: Comparison of Non-Attriting Students by Research Group

	Baseline to Follow-Up 1		Baseline to Follow-Up 2	
	Control Average	Treatment-Control	Control Average	Treatment-Control
Panel A: Attrition Rates				
Probability of Completing Survey	0.84	0.03	0.81	0.01
		(0.02)		(0.03)
Panel B: Demographic Characteristics				
Female	0.58	0.02	0.59	0.02
		(0.04)		(0.04)
Grade	5.41	0.01	5.41	0.01
		(0.04)		(0.04)
Age	9.82	0.08	9.84	0.06
		(0.06)		(0.06)
African-American	0.76	-0.02	0.77	-0.03
		(0.03)		(0.03)
Latino	0.14	-0.01	0.14	-0.01
		(0.03)		(0.03)
Reduced Lunch	0.68	-0.07*	0.7	-0.09**
		(0.03)		(0.04)
Panel C: Normalized Test Scores				
Reading	0.09	0.05	0.09	0.04
		(0.06)		(0.06)
Problem Solving	0.1	0.04	0.09	0.03
		(0.06)		(0.07)
Panel D: Household				
Total Household Income:				
Below $25,000	0.34	-0.06*	0.34	-0.05
		(0.04)		(0.04)
$26,000-$50,000	0.38	0.05	0.37	0.06
		(0.04)		(0.04)
$51,000-$75,000	0.16	0.03	0.16	0.02
		(0.03)		(0.03)
Over $75,000	0.13	-0.02	0.13	-0.03
		(0.03)		(0.03)
Parent Education:				
College Degree	0.33	0.01	0.34	0
		(0.04)		(0.04)
Some College	0.31	0.01	0.3	0.01
		(0.03)		(0.04)
Living with Both Parents	0.39	0	0.4	-0.02
		(0.04)		(0.04)
Non-English Language Spoken at Home	0.17	0	0.17	0
		(0.03)		(0.03)

Note: This table provides a comparison of students in the treatment and the control group who completed the respective follow-up survey using information collected either on the HA application or the baseline survey. Columns one and three contain the average characteristics of the control group for the first- and second-year follow-up surveys respectively. Columns two and four provide the estimated difference between the non-attriting treatment and control group using equation (1). Significance at the ten-, five-, and one-percent levels are indicated by *, **, and *** respectively

Table 6: Effects of Treatment Assignment on Out-of-School Time Program Participation

	Higher Achievement			Non-HA Academic			Any Academic		
	Ever Attended	Current Attending	Weekly Hours	Ever Attended	Currently Attending	Weekly Hours	Ever Attended	Currently Attending	Weekly Hours
Panel A: First-Year Follow-Up									
Treatment	0.868***	0.747***	8.919***	-0.095***	-0.036	-0.563	0.524***	0.539***	10.341***
	(0.018)	(0.023)	(0.412)	(0.033)	(0.031)	(0.508)	(0.027)	(0.029)	(0.673)
R-Squared	0.75	0.57	0.37	0.02	0.02	0	0.32	0.31	0.23
F-Statistic	2380.63	1047.54	468.85	8.01	1.34	1.23	368.54	337.98	236.15
Prob > F	0	0	0	0	0.25	0.27	0	0	0
Sample Size	818	818	818	818	818	818	818	818	818
Control Average	0	0	0	0.35	0.26	2.98	0.35	0.26	2.98
Panel B: Second-Year Follow-Up									
Treatment	0.883***	0.698***	7.325***	-0.067**	-0.037	0.346	0.562***	0.500***	8.517***
	(0.018)	(0.025)	(0.399)	(0.034)	(0.032)	(0.544)	(0.028)	(0.032)	(0.674)
R-Squared	0.75	0.5	0.31	0.01	0	0	0.34	0.24	0.18
F-Statistic	2305.9	773.44	337.07	3.87	1.39	0.4	391.21	242.93	159.83
Prob > F	0	0	0	0.05	0.24	0.52	0	0	0
Sample Size	774	774	774	774	774	774	774	774	774
Control Average	0	0	0	0.35	0.26	2.98	0.35	0.26	2.98

Note: This table provides estimates of the effects of treatment assignment on academic OST participation. Panel A and B contain the estimates for the first- and second-year follow-up surveys respectively. Columns one through three contain estimates for participation in the HA program. Columns four through six present estimates for participation in academic OST programs other than the HA program. Columns seven through nine present estimates for participation in any academic OST program. All treatment effects are estimated using equation (1). Significance at the ten-, five-, and one-percent levels by *, **, and *** respectively.

Table 7: Academic OST Intensity

	First-Year Follow-Up				Second-Year Follow-Up			
	All Students		Attending an OST		All Students		Attending an OST	
	Control Average	Treatment-Control	Non-HA	HA - Non-HA	Control Average	Treatment-Control	Non-HA	HA - Non-HA
Panel A: Academic Year								
Days per week	1.66	1.52***	3.27	-0.87***	1.32	1.39***	3.16	-0.49***
		(0.14)		(0.14)		(0.14)		(0.17)
Hours per week	3.94	10.34***	14.8	2.66**	2.98	8.52***	13.59	4.23***
		(0.67)		(1.07)		(0.67)		(1.05)
Panel B: Summer								
Days per week	0.83	2.20***	2.85	1.32***	0.98	2.28***	3.93	1.40***
		(0.13)		(0.18)		(0.16)		(0.23)
Hours per week	4.15	19.76***	23	15.72***	5.06	19.09***	29.53	15.88***
		(1.10)		(1.68)		(1.26)		(1.99)

Note: The table presents comparisons of the intensity of participation in academic OST programs. Panel A presents estimates for the academic year, while Panel B presents estimates for the summer. Columns one, two, five, and six present estimates comparing treatment and control students, with the first column presenting the control average, and the second the estimated treatment effect estimated using equation (1). Columns three, four, seven, and eight present estimates comparing the participation rates of youth who report participating either in HA or another academic OST program at the time of the follow-up survey. The first column presents the average participation rates for children attending a program other than HA (irrespective of treatment assignment), and the second column presents the difference in participation rates between those participating in HA and those attending other programs. The comparison is made by estimating an equation identical to equation (1) but including an indicator variable for HA participation instead of an indicator for treatment assignment. Significance at the ten-, five-, and one-percent levels are indicated by *, **, and *** respectively.

Table 8: OST Related Activities, First-Year Follow-Up

	By Research Group		By Type of OST Program		
	Control Average	Treat-Control	No OST	Other OST-No OST	HA - Other OST
Done some type of community service or volunteer work?	0.53	0.01	0.52	0.04	-0.02
		(0.04)		(0.05)	(0.05)
Talked with other kids about a math or science problem outside of school?	0.65	0.05	0.65	0	0.05
		(0.03)		(0.05)	(0.05)
Spoken to a group outside of school about your ideas or your work?	0.57	0.07*	0.57	0.03	0.05
		(0.04)		(0.05)	(0.05)
Visited a college campus to see what it would be like to be a college student?	0.44	0.28***	0.48	0.04	0.25***
		(0.03)		(0.05)	(0.05)
Gotten praise for your achievements from your peers?	0.82	0.02	0.82	-0.01	0.03
		(0.03)		(0.04)	(0.04)
Helped other kids with their school work?	0.88	-0.01	0.85	0.03	0
		(0.02)		(0.04)	(0.03)
Read books that are not for school?	0.71	0.03	0.71	-0.02	0.09**
		(0.03)		(0.05)	(0.04)
Written things (like poems, letters, or essays) not assigned at school?	0.66	0.07**	0.67	0.02	0.07
		(0.03)		(0.05)	(0.05)
Stood up in front of a group of children to present your ideas?	0.82	0.02	0.83	-0.04	0.04
		(0.03)		(0.04)	(0.04)
Spoken with an adult (not your parent) about what you need to do to get into a good high school?	0.61	0.11***	0.63	0.01	0.08*
		(0.03)		(0.05)	(0.05)
Spoken with an adult (not your parent) about going to college or college applications?	0.53	0.13***	0.57	-0.02	0.09*
		(0.03)		(0.05)	(0.05)
Spoken with an adult (not your parent) about what you need to do to get a good job?	0.65	0.09***	0.67	-0.02	0.08*
		(0.03)		(0.05)	(0.05)
Spoken with an adult (not your parent) about what job you might want to have in the future?	0.77	0.08***	0.79	-0.02	0.06
		(0.03)		(0.04)	(0.04)
Visited a business or organization to see what it would be like to work there?	0.52	0.04	0.47	0.12**	-0.01
		(0.04)		(0.05)	(0.05)
Gone to events outside of your neighborhood with your after-school program?	0.67	0.11***	0.66	0.06	0.09**
		(0.03)		(0.05)	(0.04)
Participated in academic contests at your after-school program?	0.55	0.13***	0.53	0.08	0.09*
		(0.03)		(0.05)	(0.05)
Average participation in all activities	0.65	0.08***	0.65	0.02	0.06***
		(0.02)		(0.02)	(0.02)
Joint Test of All Activities					
$Chi^2_{(2)}$		79.58***		10.59	44.04***
P-value		< 0.001		0.834	< 0.001

Note: This table presents comparisons of the activities in which youth report engaging at the second-year follow-up survey. Columns one and two divide the sample by research group. Column one contains the averages for the control group, and column two presents the treatment effects estimated using equation (1). Columns three and four compare the reports of youth who report participating either in HA or another academic OST program at the time of the follow-up survey. The first column presents the average participation rates for children attending a program other than HA (irrespective of treatment assignment), and the second column presents the difference between those participating in HA and those attending other programs. The comparison is made by estimating an equation identical to equation (1), but including an indicator variable for HA participation instead of an indicator for treatment assignment. Significance at the ten-, five-, and one-percent levels are indicated by *, **, and *** respectively.

Table 9: Effects of HA on Standardized Test Scores

	(1)	(2)	(3)	(4)	(5)
Panel A: First-Year Follow-Up					
Reading Comprehension	0.04	0.01	0.01	0.01	0.01
	(0.06)	(0.05)	(0.05)	(0.05)	(0.09)
Problem Solving	0.05	0.03	0.01	0.02	0.03
	(0.07)	(0.05)	(0.05)	(0.05)	(0.09)
Joint Test					
$Chi^2_{(2)}$	0.59	0.38	0.10	0.09	0.10
P-value	0.744	0.825	0.953	0.954	0.953
Panel B: Second-Year Follow-Up					
Reading Comprehension	0.11*	0.09*	0.09*	0.10*	0.16*
	(0.06)	(0.05)	(0.05)	(0.06)	(0.09)
Problem Solving	0.12*	0.13**	0.12**	0.14**	0.21**
	(0.07)	(0.05)	(0.05)	(0.06)	(0.09)
Joint Test					
$Chi^2_{(2)}$	3.74	7.81**	7.12**	7.55**	7.46**
P-value	0.154	0.020	0.029	0.023	0.024
Model	ITT	ITT	ITT	TOT: HA	TOT: OST
Cohort FE	Y	Y	Y	Y	Y
Baseline Child Characteristics	N	Y	Y	Y	Y
Household Characteristics	N	N	Y	Y	Y
Ward FE	N	N	Y	Y	Y

Note: This table presents the estimated treatment effects on test scores. Columns one, two, and three present the intent-to-treat estimates with column one esitmated using equation (1) and columns two and three estimated using equation (2) with the indicated sets of control variables. Column four contains the local average treatment effects for HA participation, and column five presents the estimated local average treatment effects for any academic OST participation. Joint tests of the significance of both treatment effects in the respective survey period are made by estimating all treatment effects simultaneously using Seemingly Unrelated Regressions. Significance at the ten-, five, and one-percent levels are indicated by *, **, and *** respectively.

Table 10: Effects on Psychometric Outcomes

	First-Year Follow-Up			Second-Year Follow-Up		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Academic Attitudes						
Industry and Persistence	-0.14**	-0.16**	-0.18***	-0.02	-0.06	-0.06
	(0.07)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)
Ability to be Creative	-0.13*	-0.14**	-0.16**	0.02	0.01	0.01
	(0.07)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)
Self-Perceptions of Academic Abilities	-0.1	-0.09	-0.11*	0.1	0.11*	0.08
	(0.07)	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)
Enjoyment of Learning	-0.14**	-0.14**	-0.16**	-0.05	-0.08	-0.11
	(0.07)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)
Curiosity	-0.11	-0.11*	-0.14**	-0.01	-0.01	-0.02
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Ability to Change the Future	0.06	0.04	0.04	0.02	-0.01	-0.01
through Effort	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)
Joint Test						
$Chi^2_{(6)}$	6.94	9.24	12.66**	5.85	9.76	9.11
P-value	0.326	0.160	0.049	0.440	0.135	0.167
Panel B: Peer and Adult Support						
Academically Supportive Friends	0	-0.01	-0.03	0.03	0.06	0.05
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Adult Support	0.07	0.09	0.06	0.08	0.1	0.08
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Joint Test						
$Chi^2_{(2)}$	1.15	1.69	1.16	1.28	2.30	1.54
P-value	0.563	0.430	0.559	0.529	0.316	0.463
Cohort FE	Y	Y	Y	Y	Y	Y
Baseline Child Characteristics	N	Y	Y	N	Y	Y
Household Characteristics	N	N	Y	N	N	Y
Ward FE	N	N	Y	N	N	Y

Note: This table presents estimates of the treatment effects on the indicated psychometric scores. Columns one through three present the results for the first-year follow-up survey, and the remaining columns contain the results for the second-year. Estimates in columns one and four are estimated using equation (1) while those in the other columns are estimated using equation (2) with the respective set of control variables. Joint tests of the significance of all treatment effects in the respective category of outcome variables are made by estimating all treatment effects simultaneously using Seemingly Unrelated Regressions. Significance at the ten-, five-, and one-percent levels are indicated by *, **, and *** respectively.

Table 11: Changes in Average Academic Attitudes by Grade

	Control	Treatment	All	Fifth Grade	Sixth Grade
	(1)	(2)	(3)	(4)	(5)
First-Year*Grade Five	0.023	-0.106***			
	(0.045)	(0.040)			
First-Year*Grade Six	-0.058	-0.106**			
	(0.059)	(0.052)			
Second-Year*Grade Five	-0.213***	-0.263***			
	(0.048)	(0.047)			
Second-Year*Grade Six	-0.261***	-0.219***			
	(0.062)	(0.055)			
First-Year			-0.007	0.037	-0.053
			(0.033)	(0.042)	(0.054)
First-Year*Treat			-0.103**	-0.158***	-0.053
			(0.042)	(0.053)	(0.067)
Second-Year			-0.227***	-0.193***	-0.262***
			(0.036)	(0.046)	(0.057)
Second-Year*Treat			-0.023	-0.085	0.042
			(0.047)	(0.061)	(0.071)
Observations	1144	1401	2545	1475	1070
R-squared	0.500	0.460	0.46	0.490	0.480
H0: F1*Grade Five = F1*Grade Six					
F-stat	1.18	<0.01			
P-value	0.278	0.990			
H0: F2*Grade Five = F2*Grade Six					
F-stat	0.37	0.38			
P-value	0.546	0.536			

Note: This table presents the treatment effects on the overall average of the academic attitude measures by entering grade and survey period. Estimates are performed at the child-survey-period level using a regression model similar to equation (2), including interactions between the follow-up period, grade, and treatment indicator. Significance at the ten-, five-, and one-percent levels are indicated by *, **, and *** respectively.

Table 12: Psychometric Treatment Effects by Grade

	Academic Attitudes							Standardized Test Scores	
	Avg Academic Attitudes	Industry and Persistence	Ability to be Creative	Perceptions of Acad Abilities	Enjoyment of Learning	Curiosity	Change Future through Effort	Reading Comp.	Problem Solving
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(9)	(10)
Panel A: First-Year Follow-Up									
Treat*Grade Five	-0.19***	-0.36***	-0.18**	-0.25***	-0.23***	-0.13	0.03	-0.02	0.05
	(0.06)	(0.09)	(0.08)	(0.09)	(0.08)	(0.09)	(0.09)	(0.06)	(0.06)
Treat*Grade Six	-0.03	0.04	-0.12	0.08	-0.07	-0.16	0.05	0.04	-0.04
	(0.07)	(0.10)	(0.09)	(0.10)	(0.10)	(0.11)	(0.11)	(0.07)	(0.07)
Grade Five	0.14*	0.29**	0.11	0.17	0.15	0.02	0.09	0.15*	0.08
	(0.07)	(0.12)	(0.10)	(0.11)	(0.11)	(0.12)	(0.12)	(0.08)	(0.08)
R-squared	0.348	0.301	0.325	0.315	0.302	0.272	0.100	0.520	0.631
Observations	818	817	816	816	818	817	817	817	817
H_0:Treat*Grade Five = Treat*Grade Six									
t-statistic	3.21	8.33	0.26	6.22	1.55	0.03	0.01	0.43	0.86
P-value	0.073	0.004	0.611	0.013	0.213	0.860	0.913	0.514	0.354
Panel B: Second-Year Follow-Up									
Treat*Grade Five	-0.07	-0.20**	-0.02	-0.06	-0.18*	0.02	0.03	0.05	0.19***
	(0.07)	(0.10)	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)	(0.07)	(0.07)
Treat*Grade Six	0.05	0.12	0.05	0.30***	-0.03	-0.09	-0.06	0.14*	0.03
	(0.08)	(0.11)	(0.11)	(0.10)	(0.12)	(0.12)	(0.11)	(0.08)	(0.08)
Grade Five	0.02	0.22*	-0.04	0.20*	0.05	-0.22*	-0.1	0.05	-0.16*
	(0.08)	(0.13)	(0.12)	(0.11)	(0.13)	(0.13)	(0.12)	(0.08)	(0.08)
R-squared	0.29	0.27	0.25	0.29	0.22	0.23	0.11	0.51	0.54
Observations	776	774	775	775	776	775	774	776	776
H_0:Treat*Grade Five = Treat*Grade Six									
t-statistic	1.44	4.72	0.22	7.54	0.98	0.49	0.33	0.77	2.70
P-value	0.231	0.030	0.640	0.006	0.323	0.484	0.566	0.380	0.101

Note: This table presents estimates of the treatment effects on academic attitudes and standardized test scores by entering grade. Panels A and B present the estimates for the first- and second-year follow-up surveys respectively. All estimates are performed using equation (2) in which the treatment indicator is interacted with an indicator variable for the grade at which a child entered the study and the specification includes the full set of control variables. Significance at the ten-, five-, and one-percent levels are indicated by *, **, and *** respectively.

Table 13: Effects on Self-Reported Behavior

	First-Year Follow-Up			Second-Year Follow-Up		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: General Misconduct						
In the last three months, have you…						
1. taken something on purpose that didn't belong to you?	0.11	0.15*	0.14*	0.12	0.18**	0.19**
	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
2. broken something on purpose?	0.15**	0.20***	0.20***	0.06	0.09	0.09
	(0.07)	(0.07)	(0.08)	(0.08)	(0.09)	(0.09)
3. taken something from a store without paying for it?	0	0.02	0.01	-0.03	-0.01	-0.01
	(0.06)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)
4. really hit someone because you didn't like what they said or did?	0.07	0.16	0.15	0.14	0.18	0.23**
	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)
Panel B: School-Related Misconduct						
In the last three months, have you…						
5. been sent to the principal's office because you had done something wrong?	0.17**	0.21***	0.22***	0.13	0.19**	0.22**
	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)	(0.09)
6. been tardy to class?	0.07	0.1	0.1	0.16	0.17	0.18
	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)
7. had to have your parents come to school about a problem?	0.05	0.08	0.06	0.07	0.06	0.08
	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)	(0.09)
8. skipped school without your parent/guardian's permission?	0.09**	0.10**	0.10**	0.06	0.09	0.08
	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)
Panel C: Joint Test of All Individual Items						
$Chi^2_{(8)}$	12.29	16.68**	18.99**	8.65	11.94	14.22*
P-value	0.139	0.034	0.015	0.373	0.154	0.076
Cohort FE	Y	Y	Y	Y	Y	Y
Baseline Characteristics	N	Y	Y	N	Y	Y
Household Characteristics	N	N	Y	N	N	Y
Ward FE	N	N	Y	N	N	Y

Note: This table presents estimates of the treatment effects on self-reported behavior. Columns one through three present the results for the first-year follow-up survey, and the remaining columns contain the results for the second-year. Estimates in columns one and four are estimated using equation (1) while those in the other columns are estimated using equation (2) with the respective set of control variables. Joint tests of the significance of all treatment effects in the respective category of outcome variables are made by estimating all treatment effects simultaneously using Seemingly Unrelated Regressions. Significance at the ten-, five-, and one-percent levels are indicated by *, **, and *** respectively.

Table 14: Effects on High School Preparation Activities, Summer 2010

Student Characteristic	Control Average	Treatment-Control
Over the summer break, have you…		
a. visited a high school to learn more about it?	0.19	0.14***
	(0.39)	(0.05)
b. talked with students who attend a specific high school to see what they think about it?	0.53	0.07
	(0.50)	(0.05)
c. gotten information about a specific high school?	0.48	0.15***
	(0.50)	(0.05)
d. learned how to get information about specific high schools that interest you?	0.47	0.09*
	(0.50)	(0.05)
e. talked with your parents about different high schools?	0.69	0.02
	(0.46)	(0.05)
f. talked with adults other than your parents about high school possibilities?	0.51	0.17***
	(0.50)	(0.05)
g. talked with kids your age about high school possibilities?	0.58	0.14***
	(0.49)	(0.05)
h. decided where you will apply to high school?	0.46	0.11**
	(0.50)	(0.05)
i. practiced writing essays for high school applications?	0.12	0.06
	(0.32)	(0.04)
j. practiced being interviewed for high school applications?	0.12	0.04
	(0.32)	(0.04)
k. practiced filling out applications for high school?	0.13	0.03
	(0.34)	(0.04)
l. gotten information about what it will be like when you go to high school?	0.52	0.06
	(0.50)	(0.05)
Joint Test		
$Chi^2_{(12)}$		31.07***
P-value		0.002

Note: This table presents estimates of the effects of treatment assignment on self-reported preparatory activities for the high school application process. The sample includes only those students surveyed in FUFa. The first column presents the average for the control group, while the second column presents the estimated treatment effect using equation (2) with a full set of control variables. The joint significance test is estimated by simultaneously estimating the treatment effects for each outcome using Seeming Unrelated Regressions. Significance at the ten-, five, and one-percent levels are indicated by *, **, and *** respectively.

Table 15: Changes in Outcomes, Summer 2010

	Spring Outcomes		Fall Outcomes		Summer Differences		
Student Characteristic	**Control Average**	**Treatment-Control**	**Control Average**	**Treatment-Control**	**Control Average**	**Treatment-Control (Differences)**	**Treatment-Control (Spr Controls)**
Panel A: Standardized Test Scores							
Reading Comprehension	-0.03	0.16**	0.12	0.03	0.14	-0.09	-0.03
	(0.92)	(0.06)	(0.94)	(0.07)	(0.61)	(0.07)	(0.06)
Problem Solving	0.16	0.11*	0.18	0.1	0.02	0.02	0.06
	(0.92)	(0.06)	(0.92)	(0.07)	(0.65)	(0.07)	(0.06)
Joint Test							
$Chi^2_{(2)}$		7.89**		2.50		2.10	1.81
P-value		0.019		0.287		0.350	0.405
Panel B: Academic Attitudes							
Industry and Persistence	0	-0.14	0	-0.04	-0.02	0.09	0.06
	(1.00)	(0.09)	(1.00)	(0.09)	(0.72)	(0.07)	(0.07)
Ability to be Creative	0	-0.01	0	-0.03	0	0	0
	(1.00)	(0.09)	(1.00)	(0.09)	(0.77)	(0.08)	(0.07)
Self-Perceptions of Academic Abilities	0	0.06	0	0	-0.02	-0.08	-0.04
	(1.00)	(0.09)	(1.00)	(0.09)	(0.80)	(0.09)	(0.08)
Enjoyment of Learning	0	-0.14	0	0.01	-0.01	0.18**	0.12*
	(1.00)	(0.09)	(1.00)	(0.09)	(0.69)	(0.08)	(0.07)
Curiosity	0	-0.09	0	0.06	0	0.1	0.06
	(1.00)	(0.09)	(1.00)	(0.09)	(0.85)	(0.09)	(0.08)
Ability to Change the Future through Effort	0	0.08	0	0.11	0	0.03	0.07
	(1.00)	(0.10)	(1.00)	(0.10)	(0.95)	(0.10)	(0.09)
Joint Test							
$Chi^2_{(6)}$		9.99		3.47		10.86*	4.65
P-value		0.125		0.748		0.093	0.589
Panel C: External Influences							
Academically Supportive Friends	0	-0.05	0	0.02	0.01	0.05	0.02
	(1.00)	(0.10)	(1.00)	(0.10)	(0.95)	(0.10)	(0.09)
Adult Support	0	0.14	0	0.08	-0.02	-0.02	0.01
	(1.00)	(0.09)	(1.00)	(0.10)	(0.94)	(0.10)	(0.09)
Panel D: High School Preferences							
Competitive HS	0.42	0.09*	0.46	0.17***	0	0.16***	0.19***
	(0.49)	(0.05)	(0.50)	(0.05)	(0.52)	(0.06)	(0.05)
Local Neighborhood HS	0.32	-0.06	0.33	-0.10**	0.06	-0.14***	-0.15***
	(0.47)	(0.04)	(0.47)	(0.04)	(0.48)	(0.05)	(0.04)
Joint Test							
$Chi^2_{(2)}$		4.31		15.89***		10.97***	19.86***
P-value		0.116		< 0.001		0.004	< 0.001

Note: This table estimates the differences in outcomes for treatment and control students before and after the summer of 2010. Columns one and two present estimates of the differences prior to the summer of 2010 while columns three and four present estimates for the period following the summer of 2010. Columns one and three present the control averages while columns two and four present the estimated treatment effects using equation (2) with the full set of control variables. Columns five through seven present esitmates of the difference in scores after the summer of 2010 while controlling for the differences prior to the summer. Column five presents the average change in test scores for the control students over the summer, and column six presents the differences-in-differences estimate using equation (2) with the difference in test scores as the dependent variable. Column seven presents estimates of the difference in test scores after the summer of 2010 using equation (2) but including the test score prior to the summer as a control variable. Significance at the ten-, five-, and one-percent levels are indicated by *, **, and *** respectively.

Table A1: Attrition Patterns by Research Group, Baseline to Second-Year Follow-Up

	Baseline to Follow-Up 1			Baseline to Follow-Up 2		
Student Characteristic	**Control Average Non-Attrit**	**Control Attrit - Non-Attrit**	**Difference in Differences**	**Control Average Non-Attrit**	**Control Attrit - Non-Attrit**	**Difference in Differences**
Panel A: Demographic Characteristics						
Female	0.58	0.06	-0.18*	0.58	0.03	-0.12
		(0.07)	(0.09)		(0.06)	(0.08)
Grade	4.41	0.06	-0.08	4.41	0.02	-0.05
		(0.07)	(0.09)		(0.06)	(0.08)
Age	9.82	0.11	-0.17	9.82	-0.03	-0.04
		(0.10)	(0.16)		(0.10)	(0.14)
African-American	0.76	0	0.03	0.76	-0.06	0.08
		(0.06)	(0.08)		(0.05)	(0.07)
Latino	0.14	-0.03	0.05	0.14	-0.03	0.06
		(0.05)	(0.07)		(0.04)	(0.06)
Reduced Lunch	0.68	-0.15**	0.13	0.68	-0.21***	0.21**
		(0.06)	(0.09)		(0.06)	(0.08)
Panel B: Normalized Test Scores						
Reading	0.09	0.11	-0.41***	0.09	0.08	-0.24*
		(0.11)	(0.15)		(0.10)	(0.13)
Problem Solving	0.1	0.02	-0.24	0.1	0.09	-0.18
		(0.12)	(0.17)		(0.11)	(0.15)
Panel C: Household						
Total Household Income:						
Below $25,000	0.34	-0.12*	0.21**	0.34	-0.11*	0.12
		(0.07)	(0.10)		(0.06)	(0.09)
$26,000-$50,000	0.38	0	-0.17*	0.38	0.04	-0.16*
		(0.07)	(0.10)		(0.07)	(0.09)
$51,000-$75,000	0.16	0.02	0.02	0.16	0	0.06
		(0.05)	(0.08)		(0.05)	(0.07)
Over $75,000	0.13	0.11**	-0.05	0.13	0.07	-0.02
		(0.05)	(0.07)		(0.05)	(0.06)
Parent Education:						
College Degree	0.33	0.01	-0.12	0.33	0	-0.03
		(0.06)	(0.09)		(0.06)	(0.08)
Some College	0.31	0.06	-0.04	0.31	0.05	-0.05
		(0.06)	(0.09)		(0.06)	(0.08)
Living with Both Parents	0.39	0.15**	-0.26**	0.39	0.06	-0.16
		(0.07)	(0.10)		(0.07)	(0.10)
Non-English Language Spoken at Home	0.17	0	0	0.17	0	0
		(0.06)	(0.08)		(0.05)	(0.07)

Note: This table compares youth who attrit from the study at the first- and second-year follow-up surveys to those who do not. The first three columns present the results for the first-year follow-up survey while the last three columns present the results for the second-year. The first column in each group presents the average characteristics for the non-attriting control youth. The second column then presents the difference in characteristics for youth who attrit from the sample relative to those who do not using a model similar to equation (1). Finally, the third column compares the difference in attriting and non-attriting youth from the treatment group to the same difference for the control group, also using a variant of equation (1). Significance at the ten-, five-, and one-percent levels are indicated by *, **, and *** respectively.

Table A2: OST Related Activities, Second-Year Follow-Up

	By Research Group		By Type of OST Program		
	Control Average	Treat-Control	No OST	Other OST-No OST	HA - Other OST
Done some type of community service or volunteer work?	0.53	0.01	0.55	0.05	0.01
		(0.04)		(0.05)	(0.05)
Talked with other kids about a math or science problem outside of school?	0.65	0.05	0.66	0.03	0.03
		(0.03)		(0.05)	(0.05)
Spoken to a group outside of school about your ideas or your work?	0.57	0.07*	0.56	0.03	0.05
		(0.04)		(0.05)	(0.05)
Visited a college campus to see what it would be like to be a college student?	0.44	0.28***	0.51	0.02	0.25***
		(0.03)		(0.05)	(0.05)
Gotten praise for your achievements from your peers?	0.82	0.02	0.85	-0.02	0.02
		(0.03)		(0.04)	(0.04)
Helped other kids with their school work?	0.88	-0.01	0.83	0.04	0.04
		(0.02)		(0.04)	(0.03)
Read books that are not for school?	0.71	0.03	0.74	0.04	0.01
		(0.03)		(0.05)	(0.04)
Written things (like poems, letters, or essays) not assigned at school?	0.66	0.07**	0.7	-0.03	0.13***
		(0.03)		(0.05)	(0.05)
Stood up in front of a group of children to present your ideas?	0.82	0.02	0.8	0.08*	-0.02
		(0.03)		(0.04)	(0.04)
Spoken with an adult (not your parent) about what you need to do to get into a good high school?	0.61	0.11***	0.73	0.01	0.08*
		(0.03)		(0.05)	(0.04)
Spoken with an adult (not your parent) about going to college or college applications?	0.53	0.13***	0.64	0	0.11**
		(0.03)		(0.05)	(0.05)
Spoken with an adult (not your parent) about what you need to do to get a good job?	0.65	0.09***	0.7	-0.03	0.12**
		(0.03)		(0.05)	(0.05)
Spoken with an adult (not your parent) about what job you might want to have in the future?	0.77	0.08***	0.82	-0.01	0.07*
		(0.03)		(0.04)	(0.04)
Visited a business or organization to see what it would be like to work there?	0.53	0.10***	0.54	0.06	0.04
		(0.04)		(0.05)	(0.05)
Gone to events outside of your neighborhood with your after-school program?	0.74	0.07**	0.73	0.07	0.03
		(0.03)		(0.05)	(0.04)
Participated in academic contests at your after-school program?	0.55	0.13***	0.55	0.04	0.15***
		(0.03)		(0.05)	(0.05)
Average participation in all activities	0.65	0.08***	0.68	0.02	0.07***
		(0.02)		(0.03)	(0.02)
Joint Test of All Activities					
$Chi^2_{(16)}$		79.58***		12.44	40.83***
P-value		< 0.001		0.71	< 0.001

Note: The table presents comparisons of the activities in which youth report engaging in the second-year follow-up survey. Columns one and two divide the sample by research group. Column one contains the averages for the control group, and column two presents the treatment effects estimated using equation (1). Columns three and four compare the reports of youth who report participating either in HA or another academic OST program at the time of the follow-up survey. The first column presents the average participation rates for children attending a program other than HA (irrespective of treatment assignment), and the second column presents the difference between those participating in HA and those attending other programs. The comparison is made by estimating an equation identical to equation (1), but including an indicator variable for HA participation instead of an indicator for treatment assignment. Significance at the ten-, five-, and one-percent levels are indicated by *, **, and ** respectively.