

U.S. Department of Education
Institute of Education Sciences
National Center for Education
Evaluation and Regional
Assistance

When Schools Stay Open Late: The National Evaluation of the 21st Century Community Learning Centers Program

New Findings

October 2004

When Schools Stay Open Late: The National Evaluation of the 21st Century Community Learning Centers Program

New Findings

October 2004

Mark Dynarski
Susanne James-Burdumy
Mary Moore
Linda Rosenberg
John Deke
Wendy Mansfield

Mathematica Policy Research, Inc.

Elizabeth Warner, Project Officer
Institute of Education Sciences

U.S. Department of Education / Institute of Education Sciences
National Center for Education Evaluation and Regional Assistance (NCEE)

U.S. Department of Education

Rod Paige

Secretary

Institute of Education Sciences

Grover J. Whitehurst

Director

National Center for Education Evaluation and Regional Assistance

Phoebe Cottingham

Commissioner

October 2004

This report was prepared for the Institute of Education Sciences under Contract No. ED-99-CO-0134. The project officer was Elizabeth Warner in the National Center for Education Evaluation and Regional Assistance. The Charles Stewart Mott Foundation also contributed to the evaluation under Grant No. 20-205. The views expressed herein are those of the contractor.

This publication is in the public domain. Authorization to reproduce it in whole or in part for educational purposes is granted.

Suggested Citation

Dynarski, Mark, Susanne James-Burdumy, Mary Moore, Linda Rosenberg, John Deke, and Wendy Mansfield. *When Schools Stay Open Late: The National Evaluation of the 21st Century Community Learning Centers Program: New Findings*. U.S. Department of Education, National Center for Education Evaluation and Regional Assistance. Washington, DC: U.S. Government Printing Office, 2004.

For ordering information on this report, write:

U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794-1398

Call toll free 1-877-4ED-Pubs; or order online at <http://www.edpubs.org>

This report is also available on the Department's Web site at:

<http://www.ed.gov/ies/ncee>

On request, this publication is available in alternate formats, such as Braille, large print, audiotape, or computer diskette. For more information, please contact the Department's Alternate Format Center at (202) 260-9895 or (202) 205-8113.

Acknowledgments

This report resulted from the combined efforts of researchers, data collection experts, and school staff who are too numerous to all be thanked by name. The authors want to recognize Elizabeth Warner at the Institute of Education Sciences for her encouragement and support and for her incisive reviews as the report went from draft to final, and Ricky Takai and Phoebe Cottingham at the Institute of Education Sciences for their helpful comments. We thank Alan Krueger and two anonymous referees for reviewing and commenting on drafts. We thank the staff of the 21st-Century program office at the U.S. Department of Education for their assistance and support in launching the evaluation and the Charles Stewart Mott Foundation for financial support that enabled us to broaden the evaluation.

Keith Appleby, Heather Berry, Jan Fertig, Jessica Wilkins, Valerie Williams, and Claire Wilson assisted in managing the data collection. Richard Heman-Ackah, Larry Snell, and Marianne Stevenson, and their interviewing and data collection staff at phone centers in Columbia, Maryland, and Princeton, N.J., put in many hours of effort to collect the data. We also thank Angela Richardson, Peter Crosta, Anne Bloomenthal, Josh Hart, Leonard Hart, Ece Kalay, Barbara Kolln, Sedhou Ranganathan, for information systems and programming support, and Angela Richardson and Peter Crosta for their steady managerial assistance. We would also like to thank Mike Sinclair and Frank Potter for creating nonresponse weights for the study.

Many school district and after-school program staff assisted the evaluation by providing data and by assisting in site visits. We appreciate their willingness to respond to many requests and to contribute their perspectives and insights during interviews.

The report was produced by Jill Miller with assistance from Marjorie Mitchell.

Contents

Chapter	Page
Executive Summary	xv
I Introduction	1
A. Features of the Evaluation Design	3
B. Key Findings From First Report	5
C. Report Organization	7
II Implementation and Impacts at Elementary School Centers	9
A. Features of Elementary School 21st Century Centers	9
1. Center Goals and Structure	10
2. Activities and Services	11
3. Characteristics of Center Staff	13
B. Attendance at Centers	13
C. Impacts of Centers	15
1. Centers Affected Who Students Were With and Where Students Were After School	18
2. Centers Increased How Many Mothers Worked or Looked for Work	21
3. Centers Did Not Increase Working on or Completing Homework	22
4. Centers Did Not Improve Academic Outcomes	26
5. Centers Increased Feelings of Safety	27
6. Centers Increased Some Types of Parent Involvement	27
7. Few Improvements in Social and Interpersonal Outcomes	30
8. Few Impacts for Subgroups	30

Contents *(continued)*

Chapter	Page
III Implementation and Outcome Differences at Middle School Centers.....	37
A. Middle School Centers in the 2001-2002 School Year	38
1. Center Goals and Structure	38
2. Centers Reported Placed Greater Emphasis on Academics.....	40
3. Centers Experienced Heavy Staff Turnover	41
B. Student Attendance Was Low in the Second Year	42
C. Middle School Center Outcome Differences.....	47
1. Some Differences in Supervision, Location, and After-School Activities	51
2. Few Differences in Academic Outcomes.....	54
3. Higher Educational Aspirations for Treatment Students	59
4. No Differences in Social and Emotional Outcomes	59
5. No Differences in Parent Outcomes	59
6. No Differences in Feelings of Safety	61
7. Mixed Evidence on Negative Behaviors.....	61
D. Few Differences for Subgroups	63
E. Comparing Estimates	64
References	71
Appendix A: Response Rates and Data Quality	75
Appendix B: Study Design and Methods for Estimating Impacts	87
Appendix C: Sensitivity Tests and Results for Alternative Specifications.....	97
Appendix D: Subgroup Tables.....	139

Tables

Table	Page
II.1 21 st Century Elementary School Center Attendance, Year 1	14
II.2 Characteristics of Treatment And Control Group Students At Baseline, Elementary School Centers	17
II.3 Impacts on Maternal Employment and Students’ Location, Supervision, and Activities After School, Elementary School Centers, Year 1	19
II.4 Impacts on Academic and Other In-School Outcomes, Elementary School Centers, Year 1	23
II.5 Impacts on Type of Homework Assistance, Elementary School Centers, Year 1	25
II.6 Impacts on Other Outcomes, Elementary School Centers, Year 1	28
II.7a Impacts on Location and Care After School, Student Effort, Maternal Employment, and Student Discipline Outcomes for White, Black, and Hispanic Subgroups, Elementary School Centers, Year 1	31
II.7b Impacts on Location and Care After School, Student Effort, Maternal Employment, and Student Discipline Outcomes for Two-Parent and One-Parent Subgroups, Elementary School Centers, Year 1	32
II.8a Impacts on Student Attendance, Academic Achievement, and Other Outcomes for White, Black, and Hispanic Subgroups, Elementary School Centers, Year 1	33
II.8b Impacts on Student Attendance, Academic Achievement, and Other Outcomes for Two-Parent and One-Parent Subgroups, Elementary School Centers, Year 1	34
III.1 21 st Century Middle School Center Attendance, Year 2	44
III.2 Characteristics of Center Participants and Comparison-Group Students: Middle School Centers	48
III.3 Outcome Differences in Maternal Employment and Students’ Location, Supervision, and Activities After School, Middle School Centers, Year 2	52
III.4 Outcome Differences in Homework Completion and on Behavior and Level of Effort in the Classroom, Middle School Centers, Year 2.....	55

Tables *(continued)*

Table	Page
III.5 Outcome Differences in Teacher-Reported Achievement and Grades, Middle School Centers, Year 2	57
III.6 Outcome Differences in Types of Homework Assistance, Middle School Centers, Year 2	58
III.7 Outcome Differences in Social Engagement, Educational Expectations, and Parental Involvement, Middle School Centers, Year 2	60
III.8 Outcome Differences in Student Safety, Negative Behavior, and Victimization, Middle School Centers, Year 2	62
III.9a Outcome Differences in Maternal Employment and Students' Location, Care, and Activities Outcomes After School For White, Black, and Hispanic Subgroups, Middle School Centers, Year 2	65
III.9b Outcome Differences in Maternal Employment and Students' Location, Care, and Activities After School for Low and High Baseline Grade Subgroups, Middle School Centers, Year 2	66
III.10a Outcome Differences in Other Student and Parent Outcomes for White, Black, and Hispanic Subgroups, Middle School Centers, Year 2	67
III.10b Outcome Differences in Other Student and Parent Outcomes for Low and High Baseline Grades Subgroups, Middle School centers, Year 2	68
III.11 Effect Sizes For Selected Outcomes From First Report And Current Report	69
A.1 Data Sources by Data-Collection Wave	78
A.2 Sample Sizes and Response Rates for the Baseline and First Followup Elementary School Sites	79
A.3 Distribution Of Response Rates For Elementary School Sites	79
A.4 Sample Sizes And Response Rates For Second Followup Middle School Sites	82
A.5 Distribution Of Response Rates, By Site, For Middle School Second Followup	83
A.6 Sample Sizes and Response Rates: Data Collected From School and After-School Center Staff Members In 2001-2002	84
C.1 Sensitivity of Impact Estimates To Alternative Specifications, Elementary School Centers, Year 1	101

Tables *(continued)*

Table	Page
C.2 Number of Sites With Positive Or Negative Impacts On Other Outcomes, Elementary School Centers, Year 1	102
C.3 Sensitivity of Various Self-Care Impact Estimates To Alternative Specifications, Elementary School Centers, Year 1	104
C.4 Examining the Effect of Regressors on Baseline Differences Between Treatment and Comparison Groups, Middle School Centers	106
C.5 Sensitivity of Outcome Differences to Alternative Specifications, Middle School Centers, Year 2	108
C.6 Number of Sites With Positive or Negative Outcome Differences on Student Safety, Negative Behavior, And Victimization, Middle School Centers, Year 2	109
C.7 Sensitivity Of Alternative Self-Care Outcome Differences to Alternative Specifications, Middle School Centers, Year 2	111
C.8 Characteristics of Center Participants and Comparison-Group Students: Middle School Centers	113
C.9 Outcome Differences in Maternal Employment and Students' Location, Supervision, and Activities After School, Middle School Centers, Year 2	115
C.10 Outcome Differences in Homework Completion and Level of Effort and Behavior in the Classroom, Middle School Centers, Year 2	118
C.11 Outcome Differences in Teacher-Reported Achievement and Grades, Middle School Centers, Year 2	120
C.12 Outcome Differences in Quality of Homework Assistance, Middle School Centers, Year 2	121
C.13 Outcome Differences in Social Engagement, Educational Expectations, and Parental Involvement, Middle School Centers, Year 2	122
C.14 Outcome Differences in Student Safety, Negative Behavior, and Victimization, Middle School Centers, Year 2	124
C.15 Differences in Maternal Employment and Students' Location, Supervision, and Activities After School By Attendance, Middle School Centers, Year 2	127

Tables *(continued)*

Table	Page
C.16 Differences in Homework Completion, Level of Effort, and Behavior In the Classroom by Attendance, Middle School Centers, Year 2.....	129
C.17 Differences in Teacher-Reported Achievement and Grades By Attendance, Middle School Centers, Year 2	130
C.18 Differences in Social Engagement, Educational Expectations, and Parental Involvement by Attendance, Middle School Centers, Year 2.....	131
C.19 Differences in Student Safety, Negative Behavior, and Victimization By Attendance, Middle School Centers, Year 2	132
C.20 Sensitivity of Attendance Estimates to Specification for Teacher-Reported Achievement and Grades, Middle School Centers, Year 2.....	135
C.21 Sensitivity of Attendance Estimates to Estimation Technique for Maternal Employment and Students’ Location, Supervision, and Activities After School, Middle School Centers, Year 2.....	137
C.22 Sensitivity of Attendance Estimates to Estimation Technique for Student Safety, Negative Behavior, and Victimization, Middle School Centers, Year 2	138
D.1a Impacts on Maternal Employment and Students’ Location, Care, and Activities After School by Subgroup, Elementary School Centers, Year 1	141
D.1b Impacts on Maternal Employment and Students’ Location, Care, and Activities After School by Subgroup, Elementary School Centers, Year 1	142
D.1c Impacts on Maternal Employment and Students’ Location, Care, and Activities After School by Subgroup, Elementary School Centers, Year 1	143
D.2a Impacts on Homework Completion, Level of Effort, and Classroom Behavior by Subgroup, Elementary School Centers, Year 1	144
D.2b Impacts on Homework Completion, Level of Effort, and Classroom Behavior by Subgroup, Elementary School Centers, Year 1	145
D.2c Impacts on Homework Completion, Level of Effort, and Classroom Behavior by Subgroup, Elementary School Centers, Year 1	146
D.3a Impacts on Student Attendance and Academic Achievement by Subgroup, Elementary School Centers, Year 1	147

Tables (continued)

Table	Page
D.3b Impacts on Student Attendance and Academic Achievement by Subgroup, Elementary School Centers, Year 1	148
D.3c Impacts on Student Attendance and Academic Achievement by Subgroup, Elementary School Centers, Year 1	149
D.4a Impacts on Other Student and Parent Outcomes by Subgroup, Elementary School Centers, Year 1	150
D.4b Impacts on Other Student and Parent Outcomes by Subgroup, Elementary School Centers, Year 1	151
D.4c Impacts on Other Student and Parent Outcomes by Subgroup, Elementary School Centers, Year 1	152
D.5a Outcome Differences on Maternal Employment and Students' Location, Care, and Activities After School by Subgroup, Middle School Centers, Year 2.....	153
D.5b Outcome Differences on Maternal Employment and Students' Location, Care, and Activities After School By Subgroup, Middle School Centers, Year 2	154
D.6a Outcome Differences on Homework Completion, Level of Effort, and Classroom Behavior by Subgroup, Middle School Centers, Year 2.....	155
D.6b Outcome Differences on Homework Completion, Level of Effort, and Classroom Behavior by Subgroup, Middle School Centers, Year 2.....	156
D.7a Outcome Differences on Other Student and Parent Outcomes by Subgroup, Middle School Centers, Year 2.....	157
D.7b Outcome Differences on Other Student and Parent Outcomes by Subgroup, Middle School Centers, Year 2.....	158

Figures

Figures	Page
II.1 Days Attended Per Week, Year 1.....	15
III.1 Attendance in Second Year.....	43
III.2 Average Days Attended Each Week, Second Year and First Year.....	46
III.3 Average Days Attended Each Week, Second-Year Participants and First-Year Participants.....	46
B.1 Effect of 10 More Days Attended on Achievement.....	94

When Schools Stay Open Late: New Findings From the National Evaluation

Executive Summary

After-school programs have grown rapidly in recent years, spurred by rising employment rates of mothers, pressure to increase academic achievement, and concerns about risks to children who are unsupervised during after-school hours. The percentage of public schools offering “extended day” programs (which include before- and after-school programs) more than tripled from 1987 to 1999, from about 13 percent to 47 percent.

The federal government’s investment in after-school programs has grown rapidly as well. Funding for the 21st Century Community Learning Centers program, created in 1994, rose from \$40 million in 1998 to \$1 billion in 2002. The program now provides funding to 2,250 school districts to support school-based programs in 7,000 public schools.

Some studies of after-school programs have found that these programs increase academic achievement and student safety, as well as reduce negative behaviors such as drug and alcohol use. However, other studies have found that after-school programs have no effect on—and even worsen—certain outcomes, leading to debate over whether the evidence supports increased investment in after-school programs.

In 1999, the U.S. Department of Education contracted with Mathematica Policy Research, Inc., and Decision Information Resources, Inc., to evaluate the 21st Century Community Learning Centers program. The evaluation team collected student outcome data in five areas: after-school supervision, location, and activities; academic performance and achievement; behavior; personal and social development; and safety. Because the purpose of the 21st Century Community Learning Centers program by law is safe and drug-free learning environments for students that support academic achievement, this evaluation focused on student and school outcomes. It did not explore the full range of parental needs and satisfaction that might be affected by the availability of after-school programs. It did collect parent outcome data on involvement in school activities and employment status.

In its first year of data collection, the team gathered data for roughly 1,000 elementary school students in 18 schools in 7 school districts, and 4,300 middle school students in 61 schools in 32 school districts. The elementary study was based on random assignment, in which outcomes of students assigned to the program were compared with outcomes of students not assigned to the program. The middle school evaluation was based on a matched-comparison design, in which outcomes of students who participated in programs were compared with outcomes of similar students who did not. Findings from these data were presented in the study’s first report (hereafter referred to as the “first report”), which was released in February 2003.

For the second year of data collection, researchers gathered additional data in two ways. First, they added more elementary school programs and students. Second, they followed middle school students for a second year, which enabled the evaluation to explore whether there were

outcome differences after two years. The results are summarized in this new report, which contains findings from this second year of data collection. A third report will analyze impacts for elementary students after two years.

Key Findings From the Second Year

The findings from the second year of the study are generally consistent with those from the first year. Specifically, the study found

- ***Supervision by Other Adults Increased.*** Students in programs were more likely to be with adults who were not their parents after school and less likely to be with their parents or older siblings.
- ***Self-Care Was Unaffected.*** Participation in programs had no effect on whether students were in self-care (so-called latch-key children) after school. Multiple definitions of self-care were analyzed with similar results.
- ***Few Impacts on Academic Achievement.*** Programs did not affect reading test scores or grades for elementary students. Grades for middle school students in programs were higher in social studies relative to the comparison group but not in English, mathematics, and science. Programs did not increase whether elementary or middle school students completed their homework. Middle school students in programs missed fewer days of school and were more likely to aspire to attend college.
- ***Elementary Students Felt Safer.*** Elementary students in after-school programs reported feeling safer during after-school hours. Middle school students did not report feeling safer.
- ***Mixed Evidence on Negative Behavior for Middle School Students.*** Some estimates pointed to higher levels of negative behaviors for middle school students, while others indicated no differences between treatment and comparison groups.
- ***Some Impacts on Parent Outcomes.*** Parents of participating elementary school students were more likely to report that they attended school events. Other measures of parent involvement did not increase. There was some evidence that programs increased whether mothers of elementary students worked or looked for work. Involvement of middle school parents did not differ between the treatment and comparison groups. No employment difference was observed for mothers of middle school students.
- ***Few Impacts on Developmental Outcomes.*** Elementary students were more likely to report helping other students after school. They were no more likely to report being able to work with others on a team, believe the best of other people, or set goals and work to achieve them. Middle school students showed no differences in these outcomes.

- ***Low Middle School Attendance in Second Year.*** Two attendance patterns emerged in the study's second year. First, many students who had access to programs in the second year (53 percent) did not attend. Second, among those who did attend, average attendance was low (30 days) and similar to attendance during the first year (33 days).
- ***Moderate Elementary School Attendance.*** The first report noted that elementary school students attended programs an average of 58 days in the school year. With five additional sites and a larger student sample, average attendance was 63 days.
- ***Stable Program Leadership, But High Staff Turnover Between the First and Second Years.*** Eighty-two percent of project directors who worked in programs during the first year still worked for the programs in the second year. However, two-thirds of the line staff and one-third of center coordinators who worked in programs during the first year of the study were no longer working for the programs in the second year.

Study Methodology

The national evaluation of the 21st Century Community Learning Centers program includes an elementary school study and a middle school study.

The elementary school study uses random assignment of students to treatment and control groups. The study involved 12 school districts and 26 centers, which were included in the evaluation because the centers had more students interested in attending than the centers could serve, a precondition for random assignment. The findings are based on baseline and follow-up data collected from students, parents, teachers, principals, program staff members, and school records. The baseline and follow-up data were collected for 589 treatment group students and 384 control group students in 7 school districts in the 2000-2001 school year, and for 693 treatment group students and 666 control group students in 5 school districts in the 2001-2002 school year. The total elementary school sample was 2,308 students.

The middle school study is based on a nationally representative sample of 21st Century programs serving middle school participants and a matched comparison group of students who are similar to participants. Similar students were identified in host schools or in other schools in the participating districts. Student data were collected from 32 school districts and 61 centers in those districts. The sample includes 1,782 participants who were matched to 2,482 comparison students.

The U.S. Department of Education has funded seven cohorts of grantees. The middle school study includes grantees from the first three cohorts of grants, and the elementary school study includes grantees from the first five. When the study began, all grantees were in their second or third year of a 3-year grant. In 2001, the No Child Left Behind (NCLB) law changed the program to state administration; this study does not include 21st Century programs from the state-administered program.

The implementation analysis was based on site visits that were conducted to all grantees, with visits lasting between two and four days. Each center was visited twice, once during each of the two years of the study.

Characteristics and Impacts of Elementary School Programs

The two most common objectives of administrators of elementary school programs were to offer students a safe place after school and to help students improve academically. These goals mirror those of parents, who said they enrolled their children in the programs to help them do better in school (79 percent of parents) or to provide “a safe place for my child after school” (63 percent of parents).

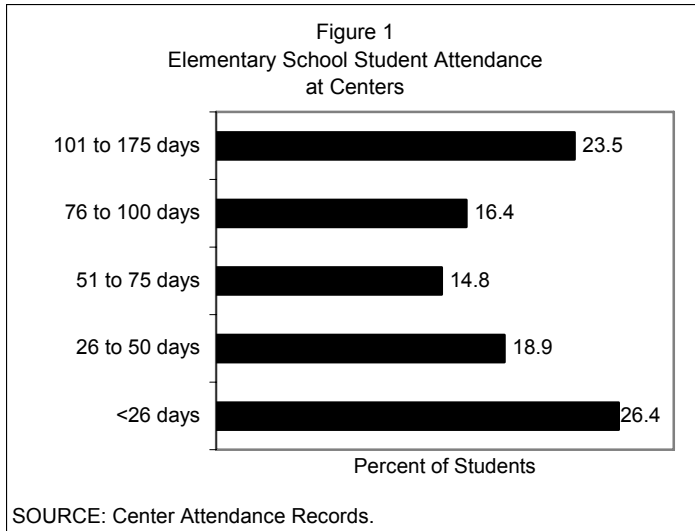
A Typical Elementary School Center

The center is open five days a week for three hours a day, serving students in kindergarten through grade six. About 120 students a day come to the center. The first 75 minutes is snack time followed by homework. Certified teachers and aides supervise the homework sessions. The next two sessions are 40 minutes each and include academic and enrichment activities. For the first session, students alternately work on computers to enhance their reading or math skills or meet with a certified teacher for lessons that complement what students worked on with computers. For the next session, students are grouped with other students in their grade and rotate through enrichment activities such as arts and crafts, karate, and fitness and dance. A mix of teachers, instructional aides, and outside organizations lead the enrichment activities. On Fridays, students have free choice for one 40-minute block and use the time to play board games or basketball.

Generally programs were open for three hours after school four or five days a week. A typical day included one hour for homework and a snack, one hour for another academic activity such as a computer lab, and one hour for recreational or cultural activities.

Eighty-five percent of the centers offered homework assistance, mostly by setting aside time for students to do their homework. Eighty-five percent also provided academic activities, such as teaching or tutoring, in addition to, or instead of, homework help.

Moreover, programs provided recreational, cultural, and interpersonal activities. Nearly all centers—92 percent—offered recreational opportunities, ranging from unstructured free time to organized sports. Programs also offered dance, drama, and music, and workshops on developmental topics, such as building leadership skills and resolving conflicts with peers.

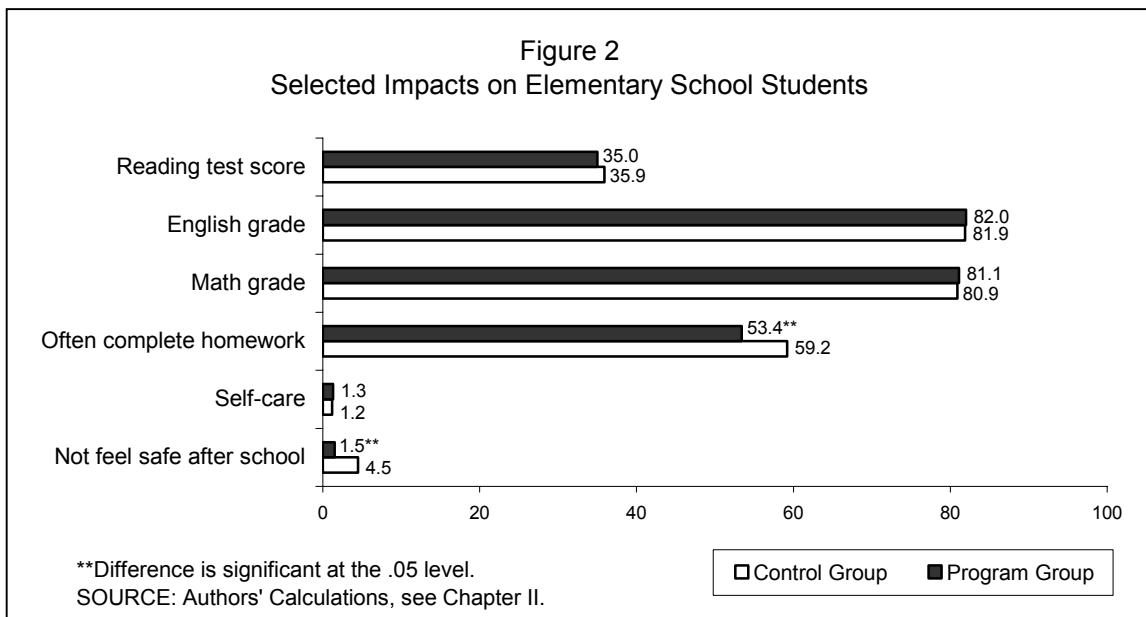


Attendance at programs was moderate (Figure 1). Students attended, on average, 63 days a year, or about two days a week. Almost one-fourth of the students attended more than 100 days a year and one-fourth attended fewer than 26 days.

Supervision After School.

Students who attended after-school programs were more likely to be with adults who were not their parents, and less likely to be with their parents after school. Students also were less likely to be with an older sibling after school. Programs

did not affect the frequency with which students reported “self-care,” or the number of days when they were at home after school without a parent, another adult, or an older sibling. Just over one percent of both groups of students said they were in self-care three or more days in a typical week (Figure 2).



Academic Achievement. Students attending after-school programs scored no better on reading tests than their peers who did not participate; nor did their grades in English, mathematics, science, and social studies increase (Figure 2). In addition, there were no statistically significant differences between the two groups of students in time spent on homework, student effort in class, preparation for class, and absenteeism; and, according to teachers, program students were *less* likely to complete homework often.

Safety After School. Programs improved students’ reported feelings of safety after school; 1.5 percent of participants, compared with 4.5 percent of nonparticipants, reported feeling “not at all safe” after school (Figure 2).

Developmental Outcomes. Programs had few impacts on developmental outcomes. For example, treatment group students were no more likely to report getting along with others their age, to rate themselves highly on working with others on a team, or to be able to set goals and work to achieve them than nonparticipants. Students in programs were more likely to report helping other students after school.

Negative Behaviors. Students in programs were no less likely than students in the control group to be suspended, to receive detention, or to be sent to the office for misbehaving. Students in programs were as likely as control students to report negative behaviors, such as breaking things, arguing with parents, or giving teachers a hard time.

Parent Outcomes. Parents of students in programs were more likely to attend after-school events in schools. There was no effect on parents attending parent-teacher organization meetings or school open houses, or volunteering at school. There was some evidence that programs increased whether mothers of elementary students worked or looked for work. Mothers of students in programs were more likely to be in the labor force (working full time, part time, or looking for work) than mothers of control students.

Subgroup Impacts. The study looked at subgroup impacts for elementary students but found few groups with significant impacts. Students from two-parent households had larger impacts on some outcomes than students in single-parent households, but after controlling for membership in other subgroups, many of these impacts were no longer significant.

Characteristics and Outcome Differences of Middle School Programs

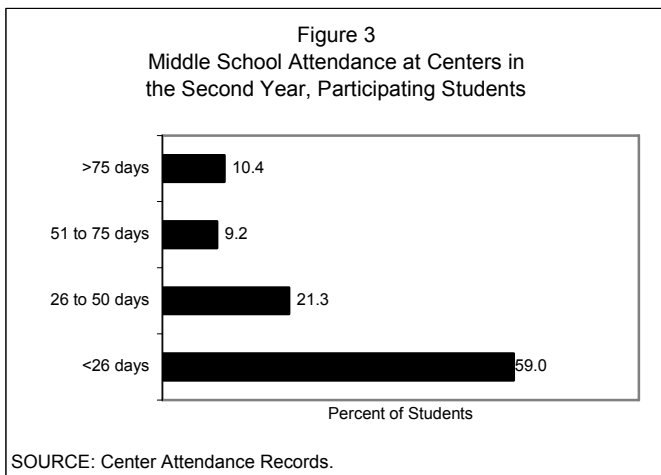
During the second year of the study’s data collection, program administrators indicated that their major objectives for programs serving middle school students were to help students improve academically and to provide a safe place for them after school. About 80 percent of centers offered homework sessions and 60 percent offered other types of academic assistance, such as additional help in language arts or mathematics. The emphasis on academics increased from the first to second year, according to site visitors, principals, center coordinators, and project directors. While our site visit data cannot confirm this shift, there clearly was a perception that centers were focusing more on academic activities.

A Typical 21st Century Middle School Center

About 45 students participate on an average day. After the school day ends at 1:30 p.m., students gather in the school cafeteria to get a snack followed by homework. After homework time ends, students choose from a variety of activities, such as free time in the gym, board games, table tennis, computer lab, and arts and crafts. A mix of certified teachers and paraprofessionals supervise the homework session and other activities.

Programs experienced considerable staff turnover during the 2 years of the study. Two-thirds of the staff did not return in the second year; almost one-third of the schools where centers were located had a new principal, and one-third had a new center coordinator. Only about 20 percent of programs had a new project director. Staff most commonly cited the demands on time that after-school work posed rather than pay as the reason for not returning.

Program attendance was much lower in the second year, averaging just 8.8 days. This was in large part because many students—59 percent of the program group—transferred to high



schools or other middle schools that had no 21st Century programs. Among the 41 percent of the program group who had access to the program in the second year, 47 percent attended at least 1 day; for the year, their attendance averaged 30 days. This is similar to the average number of days attended in the first year (33 days). Ten percent of participating students attended more than 75 days and 59 percent of participating students attended fewer than 26 days (Figure 3). Week-to-week attendance patterns also were similar to first-year patterns.

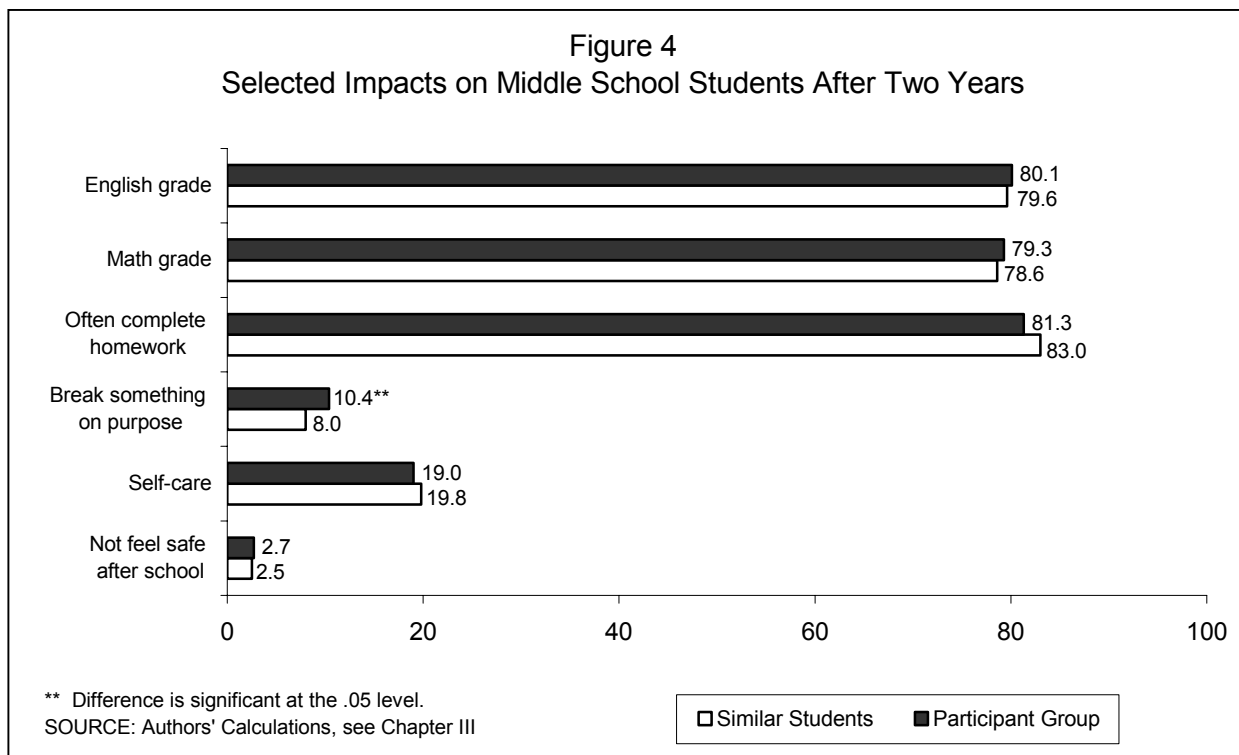
Supervision After School. The program group was less likely to be with siblings than the comparison group. There were no differences in self-care, with roughly 19 percent of participants and nonparticipants indicating that they were not with an adult or older sibling three or more days a week after school.

Academic Achievement. There were few differences between the program and comparison groups on academic outcomes (Figure 4). The program group had higher grades in social studies. Other outcomes—including grades in mathematics, science, and English, as well as teacher reports of achievement—did not differ. The level of homework completion also did not differ.

Safety After School. There were no differences between the program group and comparison group in feelings of safety after school.

Developmental Outcomes. The program group was more likely than the comparison group—82 percent versus 79 percent—to expect to graduate from college. No differences were observed in other developmental areas.

Negative Behaviors. Findings on one of several drug-use questions indicated that the program group had a higher incidence of drug use (use for both groups was low). There were no differences on the other measures of drug use. There were mixed findings on other measures of behavior. Treatment students were more likely than comparison students to report breaking things on purpose and had higher values on an index of negative behaviors, but there were no



differences on other outcomes such as punching someone, stealing, selling drugs, or getting arrested.

Parent Outcomes. No differences were found in parent involvement.

Subgroup Impacts. The study examined six subgroups: (1) grade level, (2) whether students had low or high reading test scores at baseline, (3) whether students had low or high discipline problems at baseline, (4) student race and ethnicity, (5) student gender, and (6) whether students lived in two-parent or one-parent households. None showed distinct patterns of difference, with one exception: students with low grades (at baseline) had more positive impacts than did students with high grades. Reasons for the difference were not clear.

Comparison of Findings of the First and Second Reports

The comparison below is presented separately for elementary and middle school students because the basis for differences in findings differs for the two groups. For elementary school students, differences in findings between the first and second reports are due to the addition of new sites to the study; for middle school students, differences in findings relate to an additional follow-up year.

Elementary School Students

Supervision and Location After School. Both reports found that elementary school students attending programs were less likely to be supervised by parents and siblings and more

likely to be supervised by other adults. They also were more likely to be at school and less likely to be at home during after-school hours.

Academic Achievement. Both reports found that programs generally did not improve academic outcomes such as grades or test scores. In the first report, elementary school students had higher grades in social studies but not in English, mathematics, or science. In the second report, grades were not higher in any of the four subjects. Both reports found no difference in reading test scores. Both reports found homework completion was lower; the second report's finding was statistically significant.

Safety After School. Both reports found that students reported feeling safer after school; only the second report's finding (based on a larger sample size) was statistically significant.

Social, Emotional, and Developmental Outcomes. Both reports found that students were more likely to help other students after school. There were no differences in other outcomes, such as the extent to which students reported getting along with others or setting goals and working toward them.

Negative Behaviors. Students were equally likely to be disciplined for bad behavior, be suspended, or receive detention.

Parent Outcomes. Both reports found that parents were more likely to attend after-school events, to help their children with homework, and to ask their children about class.

Subgroup Outcomes. Neither report found noteworthy patterns of subgroup outcomes. In the second report, students from two-parent households had larger impacts on some outcomes than students from single-parent households, but these differences were no longer significant after controlling for students' membership in other subgroups. This subgroup was not examined in the first report.

Middle School Students

Supervision and Location After School. The first report found that program students were more likely than comparison-group students to be supervised by other adults and less likely to be supervised by parents or siblings. Students also were more likely to be at school and less likely to be at home during the after-school hours. In the second report, the only significant findings were a reduction in being supervised by siblings and an increase in being at school during the after-school hours.

Academic Achievement. Both reports found few differences in academic outcomes. In the first report, students had higher grades in math but not in English, science, or social studies. In the second report, students had higher grades in social studies but not in English, math, or science. Both reports found no differences in homework completion. School absences were lower for treatment students relative to comparison students in both reports.

Safety After School. Both reports found no differences in feelings of safety after school.

Social, Emotional, and Developmental Outcomes. Both reports found an increase in students who expected to go to college.

Negative Behaviors. Both reports found mixed evidence on negative behaviors. Some estimates indicated that program students were more likely to engage in negative behaviors and others showed no difference.

Parent Outcomes. The first report indicated that parents were more likely to attend open houses, parent/teacher organization meetings, and after-school events, and more likely to volunteer at school. The second report found no differences in parent involvement.

Subgroup Outcomes. The first report found some increases in academic outcomes for black and Hispanic middle school students. The second report did not find such increases.

I. Introduction

The number of after-school programs has grown quickly in recent years, spurred by rising employment rates of mothers, pressures on districts and schools to increase academic achievement, concerns about risks to children who are unsupervised in after-school hours, and the expansion of federal funding for after-school programs. The percentage of public schools offering extended-day programs tripled from 1987 to 1999 (National Center for Education Statistics 2002) and estimates from the National Household Education Survey indicate that the number of children in kindergarten through 2nd grade in after-school programs grew from 1.6 million in 1995 to 2.5 million in 2001 (Brimhall and Reaney 1999; Kleiner et al. 2004).

Federal funding for after-school programs through the federal 21st Century Community Learning Centers program rose from \$40 million in 1998 to \$1 billion in 2002. In addition, federal funding from other sources, such as Temporary Aid for Needy Families or the Child Care and Development Fund, now supports after-school programs.

Some research studies have reported that after-school programs increase academic achievement, enhance safety, and reduce negative behaviors such as drug and alcohol use (Brooks et al. 1995; Hamilton and Klein 1998; Tierney et al. 1995; Welsh et al. 2002). However, most studies report negative or neutral findings for some outcomes and positive findings for others, patterns that have been noted by observers and researchers reviewing the literature (Fashola 1998; Hollister 2003; National Research Council and Institute of Medicine 2002; Roth et al. 1998). Some researchers and policymakers have argued that the research base supports increased investments in after-school programs (Afterschool Alliance 2003; Fight Crime: Invest in Kids 2003; Schwarzenegger 2003), while others have argued the opposite (Olsen 2000).

In 1999, the U.S. Department of Education contracted with Mathematica Policy Research, Inc., and Decision Information Resources, Inc., to evaluate the 21st Century Community Learning Centers program. The program was authorized in the Improving America's Schools Act of 1994 and began awarding grants to school districts in 1998, primarily to support after-school programs.

The evaluation's elementary school study was based on random assignment, in which outcomes of students assigned to the program were compared with outcomes of students who were not assigned to the program for lack of space. The evaluation's middle school study was based on a comparison design in which outcomes of students who participated in programs were compared with outcomes of similar students who did not.

In the first year of data collection, school year 2000-2001, researchers collected data for roughly 1,000 elementary school students in 18 schools in 7 school districts. In the second year of data collection, researchers collected data for five additional grantees, which brought the sample up to 2,308 students in 26 schools in 12 districts. Adding five grantees to the study enhanced its statistical power for detecting program impacts. This report combines the two grantee cohorts and reanalyzes the program's impacts on elementary school students after 1 year. The study is collecting a second year of data for elementary school students, and a future report will present findings based on these data.

The middle school study comprised 4,300 middle school students in 61 schools in 32 school districts. Unlike the elementary school study, the middle school study did not add new grantees or schools in its second year. In its second year, it gathered more data from students, teachers, and parents, which allowed the study to examine outcome differences after two school years.

A. Features of the Evaluation Design

The key features of the evaluation's design are noted below. Additional information about the study design can be found in chapter I of the first report and in the evaluation's design report (Dynarski et al. 2001).

Elementary Schools. The evaluation identified 21st Century programs that had waiting lists or were turning students away for lack of space and implemented experimental designs. In fall 2000, roughly 1,000 students from 18 schools in seven grantees applied to 21st Century programs and were randomly assigned (findings in the first report were based on this sample). In fall 2001, an additional 1,300 students from eight schools in five school districts applied to 21st Century programs and were randomly assigned. This new report presents results for the full sample of 2,300 students (in 26 elementary schools) after one school year.

Middle Schools. The evaluation used a comparison design for a nationally representative sample of grantees that operated 21st Century programs in middle schools. Thirty-four grantees were sampled randomly and agreed to participate in the study, and baseline data were collected for 32 grantees (for two grantees, delays in starting data collection were too long to include in the report). Students who had attended the program at least three days in a one-month window in fall 2000 (according to program records) constituted the study's "treatment" group. Also in fall 2000, the study administered questionnaires to more than 21,000 students who were and were not participating in 21st Century programs. Using data from the questionnaires, propensity score methods matched program students to similar students who were not attending.

Propensity score methods involved selecting as comparison students those students whose characteristics most closely resemble the students in the treatment group on a range of demographic and academic characteristics (Rosenbaum and Rubin 1983). Propensity score models were estimated separately for each grantee; comparison students were matched only to

treatment students in the same school district. In about half of the grantees, comparison students were drawn from the same schools attended by participants. In the other half, comparison students were drawn from similar schools in the district.

The matching followed three main steps. First, for each district, the study team estimated a logistic regression model in which the dependent variable was participation status and the independent variables were student demographic characteristics, indicators of student social development, measures of academic performance, and measures of student behavior. (See Table B.1 in the first report for a listing of matching variables). In most districts, the matching was based on 38 student characteristics.¹ For participants and potential comparison group students, propensity scores were generated based on the estimated models. Second, for each participant, we identified the potential comparison group student whose propensity score was numerically closest to the participant's propensity score. To allow for possible attrition in case parent consent was not received, we also identified potential comparison group students whose propensity scores ranked them as the second- or third-best match. Once we identified matching students for each participant, we conducted statistical tests of the equality of the set of characteristics for participants and the samples of first-best, second-best, and third-best matches. Third, we created an algorithm to generate 2,000 model specifications (created by drawing random combinations of characteristics and interactions of characteristics) and carried out steps one through four to find the most equivalent comparison groups.

¹In some districts, characteristics had to be dropped from the matching models because they did not vary enough. For example, the characteristic for whether students were Hispanic was dropped in some districts that had few or no students who were Hispanic.

The study used propensity scoring because of its appealing theoretical properties and its feasibility. However, in practice, comparison designs generally have lower validity than random assignment, a caution that applies to the findings here as well.²

Ultimately, about 4,400 students were included in the evaluation's middle school sample. The evaluation collected follow-up data in spring 2001 that formed the basis of the findings presented in the first report. It collected another round of follow-up data in spring 2002 that forms the basis for the findings presented in this new report.

Outcomes. The study collected data on a wide array of outcomes that were described in the design report and the first report. Outcomes spanned five domains: supervision and location after school, academic performance, social and emotional indicators, behavior, and safety. Specific outcomes included location and supervision after school, grades, test scores, teachers' perceptions of classroom behavior and effort, school absences, parental involvement, victimization, incidents of delinquent behavior, and feelings of safety after school. The wide range of outcomes reflects the many objectives embraced by after-school programs.

Data Sources. The evaluation's data sources also were described in the first report. They include questionnaires completed by students, parents, teachers, principals, and program staff members, as well as reading tests, school records, program attendance records, and site visits.

B. Key Findings From First Report

The findings from the first year of data collection, which appeared in the February 2003 report, provide a useful context for results presented in this report. For elementary schools, key results were:

²Rosenbaum and Rubin (1983) show that the propensity scoring technique can be equivalent to random assignment under specific conditions. Whether the conditions are met, however, can be verified only in rare circumstances.

- Students attended programs about 2 days per week, an average of 58 days.
- After-school program students (who had applied and been assigned to programs) were more likely than control-group students to be with an adult who was not their parent after school and less likely to be with a parent or sibling. They also were more likely to be at school or another place for activities and less likely to be at home after school. The incidence of self-care was the same for both groups.
- Program-group students were as likely as control-group students to report feeling safer after school.
- Program-group students did not improve their reading test scores or their grades in math, English, or science relative to the control group. They did improve their grades in social studies.
- Program-group students were as likely as control-group students to complete their homework.
- No relationships were found between impacts and program and student characteristics.

For middle school students, key first-year findings were:

- Students attended programs about 1 day per week, an average of 33 days.
- Students attending programs were more likely than control-group students to be with an adult who was not their parent after school and less likely to be with a parent or sibling. They also were more likely to be at school or another place for activities after school and less likely to be in their own homes or the home of someone else.
- Students attending programs were no more likely to feel safe after school.
- Students attending programs were not more likely to complete their homework and did not improve their grades in English, science, or history, relative to comparison students. They did improve their grades in math.
- Parents of students attending programs were more likely to attend school open houses, after-school events, and parent-teacher organization meetings.
- Students attending programs had higher levels of some negative behaviors, and were more likely to be victimized, such as having things taken from them.

Other recent studies of after-school programs have yielded similar results. For example, infrequent attendance has been found for such programs (Grossman et al. 2002; Walker and Arbretton 2004), as well as inadequate help with homework (Reisner et al. 2001; Walker and

Arbreton 2004), a lack of improvement in grades and test scores (Welsh et al. 2002; Walker and Arbreton 2004), and the possibility of more negative behaviors (Sherman et al. 1998).

However, the findings fall well short of consensus. For example, a recent review of research on “out-of-school time” programs (which included summer-school, after-school, and Saturday programs) reported that they increased reading and math achievement (Lauer et al. 2003).³

C. Report Organization

The report describes the implementation and impacts of elementary school programs after one year of the study, followed by the implementation and outcome differences of middle school programs after two years of the study. Because the first report described program implementation in detail, this new report focuses on describing key features of the programs’ implementation and on identifying areas in which evidence from our implementation study may help inform the study’s impact findings. (A future report will assess implementation and impacts for elementary school programs after two years of operation.) The appendixes present detailed information about response rates and data quality, methods for estimating impacts, and additional findings not presented in the main text.

Two types of additional analyses are presented in the appendix. First, because some middle school students graduated or transferred to other schools and did not have access to a 21st Century center in the study’s second year, we estimated outcome differences for students who had access to centers in the second year. Second, we examined the relationship between center attendance and outcomes. Both of these analyses address questions of interest, however, neither of the analyses rely on the original matched treatment and comparison groups, therefore, they

³The review includes findings for intensive programs (such as programs whose purpose was to tutor students in reading or math) that differ from those more commonly delivered during out-of-school hours by schools or community organizations.

provide estimates of lower validity than the estimates presented in the body of the report. For this reason, they are presented in the appendix.

II. Implementation and Impacts at Elementary School Centers

The addition of five grantees and more than 1,300 students to the study's sample allows us to reexamine the 21st Century program's impacts on elementary school students after one school year with enhanced statistical power. This chapter provides an overview of the features of elementary school centers in the study, focusing on features that may be linked to impacts. It then examines student attendance at centers, impacts for the full student sample, and impacts for different types of students.

The estimates show that students attended centers about two days a week on average and that the students were more likely to be at school and with adults who were not their parents during after-school hours. Students in the control group were more likely to be at home and with a parent after school. Centers did not improve student academic achievement as measured by homework completion, grades, and reading test scores. These findings generally are consistent with findings presented in the first report, which were based on data for seven sites (Dynarski et al. 2003). One impact that differs from the first year is that students who attended centers reported feeling safer after school.⁴

A. Features of Elementary School 21st Century Centers

Three features of the 26 elementary school centers in the study are especially useful for understanding implementation and impacts: (1) goals and structures of centers, (2) activities and service offerings, and (3) characteristics of center staff members.⁵ Understanding program goals

⁴The first report had similar impact estimates, but the impacts were statistically insignificant.

⁵Throughout the report, a "center" refers to after-school services provided in one school, and a "site" refers to the group of centers in a school district. A "grantee" refers to a school district that received a 21st Century grant to operate centers. A grantee differs from a site because not all centers operated by some grantees were in the study.

can inform the impact analysis by highlighting the outcomes that may be affected by centers. If an important objective of centers was to improve student academic achievement, assessing whether centers improved academic outcomes is appropriate. Similarly, the activities and services offered by centers and the types of staff members who work with students are important for understanding how centers could affect students.

1. Center Goals and Structure

Providing students with a safe place after school and helping them improve academically were the two most frequently cited objectives for centers, based on responses to questionnaires completed by center administrators. These objectives mirrored the reasons parents most frequently gave for having their child attend a 21st Century center—that the center would “help my child do better in school” (79 percent of parents) and “it is a safe place for my child after school” (63 percent of parents). Improving relationships between schools and parents also was cited as a major objective by center administrators.

Percentage of Project Directors Indicating Item as One of Three Most Important Objectives	
Provide a Safe, Supervised After-School Environment	66%
Provide Tutoring/Other Activities to Enhance Students' Ability to Meet Specific Academic Goals	50%
Provide Academic Enrichment	33%
Enrich Relationship Between Parents and Schools	25%
Create a Positive Relationship Between Students and Their Schools	16%
Provide Cultural Opportunities not Available at Home or in the Community	16%
Improve Homework Completion	8%
Enhance Social Development	8%
SOURCE: Project Director Survey. Sample size is 12.	
NOTE: Percentages do not sum to 100 because project directors could indicate up to three “most important” objectives.	

(continued)

For example, some grantees operated centers in both elementary and middle schools, and the study may have included only the elementary school centers or only the middle school centers.

Centers typically were open during after-school hours for four or five school days a week (half were not open on Fridays) and for two to three hours a day. Centers often divided the after-school time into roughly hour-long sessions. The first session typically was used for students to eat a snack and do their homework. The second session might be for another academic activity, such as students' working on computers or with teachers on their basic skills. The third session often would be for development or recreational activities, such as arts and crafts, interpersonal skill building, or sports. In three-quarters of the centers, students were required to attend academic activities, but typically could choose their activity for the last session of the day or on Friday (for centers that were open on Fridays).⁶

2. Activities and Services

Eighty-five percent of centers offered homework assistance. About half of the centers used certified teachers for homework sessions and half used paraprofessionals (one site had certified teachers circulate among homework sessions monitored by paraprofessionals), with students working individually or in small groups.

Homework help sessions generally were unstructured, with students not required to work on or complete their homework. Homework help was more structured in one-quarter of the centers. For example, one center required

Examples of Academic Activities in 21st Century Centers

- Hands-on lessons, such as making exact change, solving pre-algebra problems
- Educational technology packages to reinforce basic skills or supplement classroom instruction
- Practice drills in addition, subtraction, multiplication, phonics
- Preparation for standardized tests, such as taking and reviewing practice tests, completing worksheets related to standardized tests

students to complete their homework before they could participate in other activities, and three

⁶At one site, students attended only when they were accompanied by a parent or grandparent, and could choose all their activities while the adult participated in technology-oriented instruction. Because the structure of this grantee's program differed from others in the study, we also estimated impacts excluding the grantee and found that the main findings were not affected.

sites tracked students' homework assignments. Another center that served children from several elementary schools used a homework log that students completed and their classroom teachers signed. If students said they did not have homework and their logs were unsigned (which occurred frequently), the center faxed their names to their schools to confirm they had no homework. Another site also asked teachers to initial a homework log. If students came to the center without an initialed homework log, staff members checked the classrooms to see if homework assignments had been posted. However, the procedures depended on cooperation between after-school center staff members and regular teachers.

Eighty-five percent of centers also had academic activities, such as teaching or tutoring, in addition to, or instead of, homework help. Some centers combined academic activities with homework sessions, while other centers set up distinct sessions. Most centers designed their academic activities, but some used commercially available packages. Certified teachers usually led the activities, sometimes with the help of an aide. Most centers provided activities designed to help students improve their reading and math skills. For example, in one academic activity observed by a site visitor, the teacher asked 2nd grade students to identify compound words in a storybook. In another center, 5th graders used manipulatives to solve pre-algebra problems. In a third center, 3rd grade students separated into three groups; while one group worked on reading with the aid of a tutor, another group worked independently on math worksheets, and the third group worked with a teacher to identify geometric shapes. A few centers helped students prepare for standardized tests by giving them practice tests or by working on skills covered by the tests.

Almost all centers (92 percent) offered recreational activities, which sometimes were unstructured—for example, free time, board games, or access to the computer lab. Most centers gave students the opportunity to use computers to improve their academic skills or access the

Internet to work on school projects. Some recreational activities were more structured, such as karate, basketball, and other organized sports that had coaches or instructors.

Most centers (69 percent) also offered activities to develop interpersonal skills. Activities included workshops or discussions on building leadership skills, resolving conflicts, or resisting drugs and alcohol. Paraprofessionals or community members typically led these activities. Cultural activities, such as arts and music, also were common.

3. Characteristics of Center Staff

A mix of certified teachers, paraprofessionals, and community members staffed elementary school centers. Centers had an average of 16 paid staff members on their rosters. Center coordinators worked an average of four-and-a-half days a week for four hours a day and earned just over \$19 per hour. Other staff members worked an average of about four days per week for three hours a day and earned \$15 per hour. The average student-staff ratio across the centers was about 7 to 1, ranging from about 4 to 1 to as high as 13 to 1.⁷ For most of the staff, the after-school job was a second job (71 percent of coordinators and 78 percent of other staff members reported that they had another job) and teaching was most often cited as the first job.

B. Attendance at Centers

Table II.1 indicates that students attended 63 days a year, or about 2 days a week (centers were open for 30 weeks on average). About one-fourth of students attended centers fewer than 25 days, half of students attended 26 to 100 days, and one-fourth of students attended more than 100 days. Almost 60 percent of participants attended less than half the days that centers were open.

⁷These estimates are based on the total number of students enrolled in a center and the total number of paid staff working with students; more precise estimates are difficult because of variations in the number of students and staff at a center on a given day.

Table II.1

21st Century Elementary School Center Attendance, Year 1

Average Days Attended in School Year ^a	62.5
Number of Days Attended	Percent of Students
1 to 25 Days	26.4
26 to 50 Days	18.9
51 to 75 Days	14.8
76 to 100 Days	16.4
101 to 125 Days	23.5
Attendance Rate ^b	Percent of Participants
10 Percent or less	19.3
11 to 25 Percent	14.5
26 to 50 Percent	23.1
51 to 70 Percent	19.5
71 to 85 Percent	15.4
86 to 100 Percent	8.3

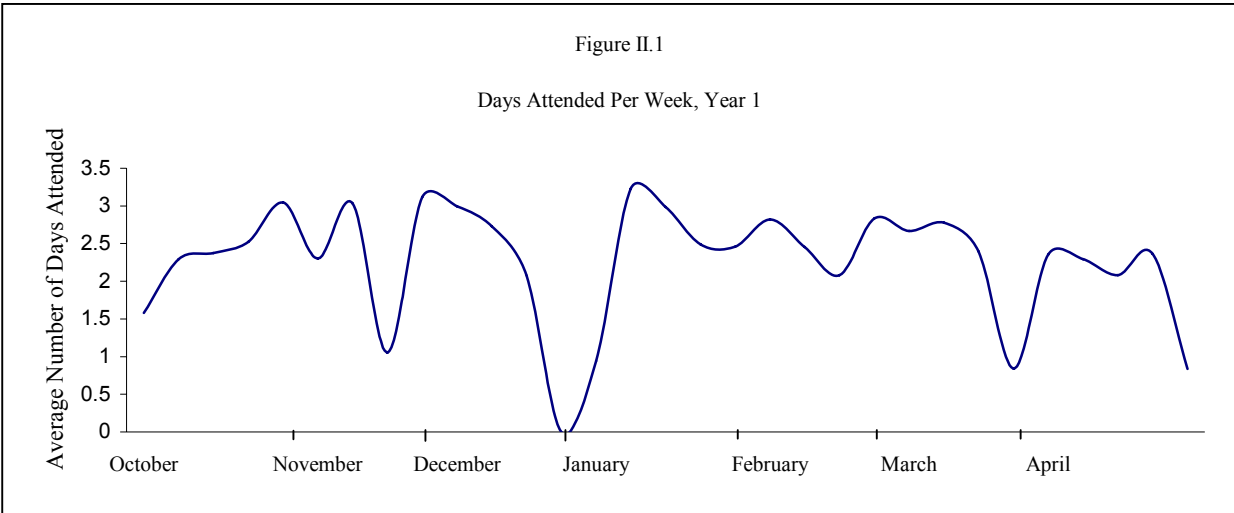
SOURCE: Center Attendance Records. Sample size is 980 students.

NOTE: Students who did not attend centers at least one day (19.5 percent of the treatment group) are excluded from the table.

^a Average number of days is calculated for center participants who attended the center at least one day after being randomly assigned to the center. Students who did not participate are not included in these calculations.

^b The attendance rate is the number of days participants attended as a proportion of the number of days centers were open, according to grantee annual performance reports.

Figure II.1 plots average days attended each week during the school year. The pattern of attendance is relatively flat, with sharp dips around major holidays.



SOURCE: Center Attendance Records.

NOTE: Students who transferred during the school year are not included in the figure.

Additional analysis found large differences in average student attendance across grantees. For example, one grantee had average student attendance of 44 days a year, whereas another had average student attendance of 78 days a year. Variations in average attendance across grantees explained much of the variation in student attendance.⁸

Few student characteristics were related to attendance at centers. We investigated 15 characteristics, only 3 of which were statistically significant. Students in younger grades (grades K through 2), students who were not on public assistance, and students who had not moved frequently in the past attended more often.

C. Impacts of Centers

Before turning to the impact estimates, it is useful to describe the treatment and control groups that are the basis for the estimates. Table II.2 shows that the treatment and control groups were similar on a range of characteristics, such as gender, grade level, mother's age, absences,

⁸Models of attendance explained 19 percent of its variation, with 17 percent of the variation in attendance related to grantees, and 2 percent related to student characteristics.

suspensions, and reading test scores, which is typical when random assignment is used. One characteristic differed at the 95 percent confidence level; students in the treatment group were more likely to report doing their homework. Considering the large number of variables reported in the table, some differences naturally would arise by chance, but the differences do not suggest that the treatment group was more or less advantaged than the control group. This same caution applies to the impact estimates reported in this chapter; because a large number of statistical tests were conducted, some findings may be significant by chance.

Follow-up response rates were high for student surveys (around 90 percent) but lower for parent surveys (75 percent), teacher surveys (79 percent), and student tests (82 percent). The study used nonresponse weights to adjust for possible differences in the characteristics of follow-up respondents in the treatment and control groups. (See Appendix A for a detailed description of the weights.)

The study also looked at attendance lists to detect whether control-group students attended centers. In principle, none of the control-group students would have attended centers. However, about nine percent of control-group students were found in the attendance records and, overall, the control group averaged 4 days of center attendance (compared to 63 days for the treatment group). During the enrollment period, the study detected some crossing over and corrected it. However, some crossing over was detected only after time had passed, and the study did not try to prevent students who had begun attending from continuing to attend centers.

Impacts were estimated using regression models to adjust for baseline differences and to improve the precision of the estimates. The study also used impact-estimation procedures to adjust for crossing over and for treatment group students not attending centers after going through random assignment. Appendix B provides details on the procedures used to estimate impacts, including the methods used to adjust for crossover by control students and

Table II.2

Characteristics of Treatment and Control Group Students at Baseline,
Elementary School Centers

Characteristic	Treatment Group	Control Group	p-value ^a
Demographics			
Gender			
Male	48.0	49.6	0.48
Female	52.0	50.4	0.48
Race/Ethnicity			
White (non-Hispanic)	6.6	4.8	0.06
Black (non-Hispanic)	54.2	55.0	0.06
Hispanic	35.3	36.2	0.06
Other	1.0	2.2	0.06
Mixed	2.9	1.7	0.06
Grade Level (percentages)			
Kindergarten	10.5	10.3	0.95
1st grade	17.9	18.2	0.95
2nd grade	17.7	19.1	0.95
3rd grade	14.9	13.7	0.95
4th grade	16.6	17.1	0.95
5th grade	16.3	16.0	0.95
6th grade	6.2	5.7	0.95
Mother's Age (Years)	34.7	34.3	0.28
Academic and Other Outcomes at Baseline			
SAT-9 Reading Score (Percentile)	32.6	30.4	0.18
Number of Absences from School	6.4	6.5	0.87
Percentage of Students Who Were Suspended At Least Once in Previous School Year	2.9	3.0	0.93
Percentage of Students Who Report Feeling the Following Level of Safety after School Up Until Dinnertime:			
Very safe	73.4	74.9	0.06
Somewhat safe	25.1	21.3	0.06
Not at all safe	1.9	3.8	0.06
Percentage of Students Who Report Doing the Homework Teachers Assign	42.9	38.8	0.04**
Sample Size^b	1,247	1,041	

SOURCE: Student Survey, Parent Survey, School Records.

^aThe p-value is the smallest level of significance at which the null hypothesis that the difference in means between program participants and control group members equals zero can be rejected. If the p-value is less than .01, the difference is significant at the 1 percent level. If the p-value is less than .05, the difference is significant at the 5 percent level, and so on. Chi-squared tests were conducted for categorical variables; for other variables, t-tests were conducted.

^bSample sizes differ depending on the data source. Sample sizes for demographic variables range from 746 to 1,041 for treatments and 936 to 1,247 for controls. Sample sizes on academic and other outcomes at baseline range from 501 to 721 for treatments and 567 to 847 for controls.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

nonparticipation by treatment students. Appendix C presents impacts based on simple treatment-control differences, which generally are similar in magnitude and statistical significance to the impacts presented in the text.

1. Centers Affected Who Students Were With and Where Students Were After School

Centers shifted who students were with during after-school hours.⁹ Centers reduced the likelihood that students were with their parents and older siblings after school, and increased the likelihood that students were with other adults (Table II.3).¹⁰ For example, 65 percent of treatment students were with parents after school at least 3 days in a typical week, compared with 75 percent of control students (effect size of .23). Forty-six percent of treatment students were with other adults after school at least 3 days in a typical week, compared with 35 percent of control students (effect size of .23).¹¹

Centers did not reduce the frequency of self-care reported by students' parents. For the self-care estimate reported in the table, students were defined as being in self-care if, for at least three days in a typical week, their parents said they were not with a parent, another adult, or an older sibling. Using this definition, just over 1 percent of treatment (and control) students were in self-care in a typical week. Defining self-care in other ways, such as whether students were home

⁹The first report used a hierarchy for the supervision categories in which the categories summed to 100 percent (a student who indicated that he was with a parent and a sibling after school was coded as being supervised by a parent). This report does not use a hierarchy and supervision categories therefore do not sum to 100 percent because students can be with different people after school (parents, other adults, siblings, and so on). This supervision construct more accurately captures the different types of supervision that students experience, but the overlap between categories makes interpreting impacts more difficult because changes in supervision in one category are not necessarily offset by changes in another category.

¹⁰We also examined impacts on sibling care using the hierarchical definition used in the first report, and found that treatment students were less likely to be supervised by siblings after school.

¹¹Tables and text indicate significance at 1 and 5 percent levels. The tables show two types of impact estimates. The first estimates are "intent to treat" estimates, which use the full treatment and control groups. The second estimates, which are shown in the column labeled "Estimated Impact on Participants," are the impacts after adjusting for the percentage of treatments who did not attend centers ("no-shows") and the percentage of controls who attended centers ("crossovers").

Table II.3

Impacts on Maternal Employment and Students' Location, Supervision, and Activities After School, Elementary School Centers, Year 1

Outcome	Treatment Group	Control Group	Estimated Impact	Estimated Impact on Participants
Percentage of Students with the Following Individuals at Least Three Days After School in a Typical Week, According to Parent Reports:				
Self-care ^a	1.3	1.2	0.1	-0.1
Parent	64.9	75.3	-10.4***	-13.0***
Nonparent adult	45.5	34.5	11.0***	14.4***
Sibling	20.8	26.3	-5.5**	-6.3
Mixed (not in any one category for at least three days)	2.2	1.5	0.7	1.0
Percentage of Students in the Following Locations After School at Least Three Days in a Typical Week, According to Parent Reports:				
Own home	61.2	79.5	-18.3***	-23.8***
Someone else's home	13.5	16.0	-2.4	-2.7
School or other place for activities	52.3	30.5	21.8***	29.3***
Somewhere to "hang out"	3.9	4.1	-0.3	-0.2
Mixed location (not in one location for at least three days)	1.3	0.9	0.4	0.7
Employment of Mother:				
Fulltime	55.4	52.2	3.2	4.8
Parttime	15.4	15.7	-0.3	0.0
Looking for work	13.9	11.3	2.6	2.8
Not in labor force	15.3	20.8	-5.5**	-7.6**
Mean Number of Days Stayed After School for Activities in a Typical Week, According to Parent Reports	2.1	0.8	1.3***	1.7***
Percentage of Students in the Following Activities after School at Least One Day in the Prior Week, According to Parent Reports:				
Homework	84.2	89.3	-5.1**	-7.6***
Tutoring	27.2	16.8	10.4***	12.5***
Non-homework reading, writing, or science activities	57.9	61.9	-4.0	-5.5
School activities (band, drama, etc.)	21.2	17.2	4.1	6.0
Lessons (music, art, dance, etc.)	22.8	19.2	3.7	5.8
Organized sports	23.0	25.9	-2.9	-2.9
Clubs (Boy and Girl Scouts, Boys and Girls Club, etc.)	16.0	15.9	0.2	-0.1
Activities at church, temple, or mosque	30.2	28.7	1.5	3.0
Watched TV or videos	78.5	82.8	-4.3	-5.0
Surfed the Internet or did other things on the computer	38.9	36.0	2.9	4.2
"Hung out" with friends	44.8	45.7	-0.9	-0.7
Did chores around the house	74.0	78.6	-4.6	-5.5
Took care of a brother or sister	16.9	21.8	-4.9**	-4.9
Mean Time Students Reported Watching Television in the Past Day (Hours)	2.1	1.9	0.2	0.3
Mean Time Students Reported Reading for Fun in the Past Day (Hours)	0.3	0.3	0.0	1.2
Sample Size^b	953	766		

SOURCE: Parent Survey, Student Survey.

NOTE: The tables show two types of impact estimates: (1) "intent to treat" estimates (in the "Estimated Impact" column) use the full treatment and control groups and (2) impacts on participants (in the "Estimated Impact on Participants" column) are the impacts after adjusting for the percentage of treatments who did not attend centers ("no-shows") and the percentage of controls who attended centers ("cross-overs"). The percentages and mean values of outcomes for treatment and control students have been regression-adjusted for baseline differences between the groups. The control variables in the regression included students' demographic characteristics, students' baseline test scores, and school attendance. Weights are used to adjust impact estimates for nonresponse. Impacts on participants are estimated using an instrumental variables method, and the significance levels may differ from significance levels of the intent-to-treat estimates. Appendix B describes methods used to estimate impacts. Percentages may not sum to 100 because of rounding.

^aStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least three days in a typical week.

Table II.3 (continued)

^bSample sizes differ for some outcomes due to nonresponse. Sample sizes for student-reported outcomes are 589 for the treatment group and 465 for the control group. Only students in third grade and above completed a student survey.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

alone at all during the week, or were home alone three or more days during the week, changed the levels of self-care (estimates range from 1 to 5 percent) but did not affect the impact estimates. (Appendix C reports impact estimates for alternative measures of self-care.)¹² However, the study's various measures of self-care are not sensitive to whether centers affected the total time students may have been in self-care, which would have required detailed information about the time that students spent with various people after school, which the study did not collect.¹³

Treatment students were more likely to be at school or another place outside the home during after-school hours (Table II.3). Fifty-two percent of treatment-group students were at school or another place outside the home at least three days in a typical week, compared with 31 percent of control-group students (effect size of .48). Treatment students were less likely to be at home during the after-school hours, with 80 percent of control students in their own homes after school at least 3 days in a typical week, compared with 61 percent of treatment students (effect size of .43).

2. Centers Increased How Many Mothers Worked or Looked for Work

Mothers of students in the treatment group were more likely than mothers of students in the control group to be “in the labor force,” which includes working full time, working part time, or looking for work (effect size of .15).¹⁴ No single measure of labor force status increased by a

¹²The incidence of self-care reported here is roughly consistent with national data. For example, in the National Household Education Survey, 2 percent of students in grades K to 2 and 8 percent of students in grades 3 to 5 were reported to be in self-care (Kleiner et al. 2004).

¹³A regression model to identify factors related to self-care found that older students (5th and 6th graders), students in one-parent households, and students in high-income households were more likely to be in self-care. Data from the National Survey of American Families and National Household Education Survey reported similar patterns (Vandivere et al. 2003; Kleiner et al. 2004).

¹⁴Whether mothers worked was related to after-school supervision. Students whose mothers worked full-time were less likely to be with a parent after school, and students whose mothers were looking for work or were not in the labor force were more likely to be with a parent after school. The National Household Education Survey found similar patterns (Kleiner et al. 2004).

statistically significant margin, but the shift from being out of the labor force to being in the labor force was statistically significant.¹⁵

3. Centers Did Not Increase Working on or Completing Homework

Treatment students were not more likely to work on or complete their homework, and some estimates suggested that attending centers may have reduced completing homework. For example, parents of treatment-group students reported that their child was less likely to work on homework after school (Table II.3).¹⁶

A similar mixed pattern was evident for whether students completed their homework. Teachers reported (Table II.4) that treatment-group students were less likely than control-group students to “often” complete their homework (53 percent of treatment students compared to 59 percent of control students, effect size is .12).¹⁷ As with working on homework, however, treatment students were as likely as control students to report that they had completed their homework, but the student sample (which excludes students in grades K to 2) is smaller than the teacher sample, and its statistical precision is lower.

The study gathered other data from student questionnaires that provide some basis for understanding the lack of differences in whether students worked on or completed their homework. In particular, Table II.5 shows that the availability of homework help in the after-school programs did not create differences in whether students had their homework checked, were asked whether it had been completed, or had parts of it explained to them by a parent or

¹⁵The study did not gather information about parental employment at baseline.

¹⁶Parents of treatment-group students reported that their child was more likely to be tutored after school. Because some parents may have considered receiving help from an adult on homework to be a form of tutoring, we also examined impacts on whether parents reported that students worked on homework *or* received tutoring in the after-school program. The results showed that parents of treatment students were less likely to report their child participating in either homework or tutoring.

¹⁷The first report had similar impact estimates, but the impacts were statistically insignificant.

Table II.4

Impacts on Academic and Other In-School Outcomes, Elementary School Centers, Year 1

Outcome	Treatment Group	Control Group	Estimated Impact	Estimated Impact on Participants
Mean Number of Days Student Was:				
Absent	7.9	8.0	0.0	-0.1
Late	4.5	4.2	0.3	0.5
Percentage of Students Whose Teachers Reported That they Are “Often” Late for Class	9.3	6.6	2.6	3.1
Percentage of Students Who Reported That They “Often” or “Always” Complete the Homework Teachers Assign ^a	81.1	80.3	0.8	0.9
Percentage of Students Whose Teachers Reported That They “Often” Complete Their Homework	53.4	59.2	-5.8**	-8.3**
Mean Amount of Time Students Spent Doing Homework the Last Time They Had Homework (hours) ^a	0.7	0.8	-0.1	-0.1
Percentage of Students Whose Teachers Reported the Following:				
“Agree” or “Strongly Agree” That Student Completes Assignments to the Teacher’s Satisfaction	53.1	56.5	-3.3	-5.4
Student Achieves at “Above Average” or “Very High” Level	24.8	28.4	-3.6	-4.7
“Agree” or “Strongly Agree” That Student Comes to School Prepared and Ready to Learn	56.5	59.8	-3.3	-5.3
Student “Usually Tries Hard” in Reading or English	52.1	49.0	3.1	3.7
Student “Often” Performs at or Above Their Ability	39.6	40.1	-0.5	-0.9
Percentage of Students Whose Parents “Agree” or “Strongly Agree” That Their Child Works Hard at School	81.1	84.1	-2.9	-5.3
Level of Effort Composite ^b (Mean)	3.5	3.6	-0.1	-0.1
Percentage of Students Whose Teachers Reported Doing the Following “Two or More Times”:				
Disciplining the child for misbehavior	50.4	45.4	5.0	7.1
Sending child to the office for misbehavior	12.8	11.1	1.6	2.2
Giving child detention	22.3	19.0	3.3	5.1
Calling parents about child’s behavior	27.5	24.5	3.1	3.8
Percentage of Students Who Were Suspended During Most Recent School Year	6.2	4.4	1.8	1.7
Mean Grade:				
Math	81.1	80.9	0.3	0.5
English/language arts	82.0	81.9	0.1	0.2
Science	82.2	82.0	0.2	0.3
Social studies/history	81.5	81.2	0.4	0.6
Mean Reading Test Score	35.0	35.9	-0.9	-0.7
Reading Confidence Composite ^c (Mean)	3.1	3.0	0.0	0.1
Sample Size^d	968	812		

SOURCE: Student Survey, Parent Survey, School Records, Teacher Survey.

NOTE: The tables show two types of impact estimates: (1) “intent to treat” estimates (in the “Estimated Impact” column) use the full treatment and control groups and (2) impacts on participants (in the “Estimated Impact on Participants” column) are the impacts after adjusting for the percentage of treatments who did not attend centers (“no-shows”) and the percentage of controls who attended centers (“cross-overs”). The percentages and mean values of outcomes for treatment and control students have been regression-adjusted for baseline differences between the groups. The control variables in the regression included students’ demographic characteristics, students’ baseline test scores, and school attendance. Weights are used to adjust impact estimates for nonresponse. Impacts on participants are estimated using an instrumental variables method, and the significance levels may differ from significance levels of the intent-to-treat estimates. Appendix B describes methods used to estimate impacts.

^aThe original set of seven sites was not asked these questions in the first year of the study.

^bThe level of effort composite is based on five teacher-reported items regarding student: (1) effort, (2) performance at ability level, (3) attentiveness, (4) participation, and (5) volunteering. Values on these items range from 1 to 5; a value of 1 on the composite indicates a low level, and a value of 5 indicates a high level.

^cThe reading confidence composite is based on student reports on three items: (1) reading is hard to learn, (2) they are a good reader, and (3) they would read better if they had more help. Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level, and a value of 4 indicates a high level.

^dSample sizes differ for some outcomes. For teacher-reported outcomes, the sample sizes are 968 treatment-group members and 812 control group members; for student-reported outcomes, the sample sizes are 578 treatment-group members and 462 control-group members; for records outcomes, the sample sizes range from 632 to 1038 for treatment-group members and from 504 to 866 for control-group members; for homework questions administered only to new sites, the sample sizes are 325 treatment-group members and 320 control-group members; for test scores, sample sizes are 1,039 for treatments and 848 for controls.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table II.5

Impacts on Types of Homework Assistance, Elementary School Centers, Year 1

Outcome	Treatment Group	Control Group	Estimated Impact	Estimated Impact on Participants
Percentage of Students Who Reported That Their Parent “Often” or “Always”: ^a				
Asks if homework is complete	85.4	82.3	3.1	3.1
Looks at homework to see if it is complete	76.3	76.8	-0.5	-0.3
Looks at homework to see if it is correct	72.3	76.3	-4.0	-5.2
Explains homework in a way that is easy to understand	71.6	74.0	-2.4	-3.8
Percentage of Students Who Reported That an Adult Who Is Not Their Parent “Often” or “Always”: ^a				
Asks if homework is complete	55.5	57.9	-2.4	-3.2
Looks at homework to see if it is complete	51.7	53.8	-2.1	-3.2
Looks at homework to see if it is correct	48.8	54.1	-5.2	-6.6
Explains homework in a way that is easy to understand	54.8	52.9	2.0	2.3
Percentage of Students Who Reported That Their Parent or an Adult Who Is Not Their Parent “Often” or “Always”: ^a				
Asks if homework is complete	90.7	86.6	4.1	4.7
Looks at homework to see if it is complete	82.2	82.3	-0.1	-0.1
Looks at homework to see if it is correct	79.7	81.9	-2.2	-3.2
Explains homework in a way that is easy to understand	79.5	81.7	-2.2	-3.1
Percentage of Students Who Were Asked to Correct Parts of Homework by: ^a				
Parent	90.8	89.5	1.3	1.5
An adult who is not their parent	75.0	76.0	-1.0	-1.2
A parent or an adult who is not their parent	91.3	93.8	-2.6	-3.2
Sample Size^b	325	320		

SOURCE: Student Survey.

NOTE: The tables show two types of impact estimates: (1) “intent to treat” estimates (in the “Estimated Impact” column) use the full treatment and control groups and (2) impacts on participants (in the “Estimated Impact on Participants” column) are the impacts after adjusting for the percentage of treatments who did not attend centers (“no-shows”) and the percentage of controls who attended centers (“cross-overs”). The percentages and mean values of outcomes for treatment and control students have been regression-adjusted for baseline differences between the groups. The control variables in the regression included students’ demographic characteristics, students’ baseline test scores, and school attendance. Weights are used to adjust impact estimates for nonresponse. Impacts on participants are estimated using an instrumental variables method, and the significance levels may differ from significance levels of the intent-to-treat estimates. Appendix B describes methods used to estimate impacts.

^aStudents in the original set of seven sites were not asked these questions in the first year of the study.

^bSample sizes differ for some outcomes due to nonresponse. Sample sizes in this table are smaller than the other elementary-school impact tables because all outcomes in the table are from the student survey, which was not administered to students in grades K-2.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

other adult. Evidently, although help with homework was a common activity in centers, parents continued to play an active role in helping with homework. Also, other adults' helping with homework was as common for control-group students as for treatment-group students. The lack of differences in who helped students with homework and in the kind of help they provided may explain the lack of differences in whether students worked on homework or completed it.

Site visitors also observed features of centers that may have contributed to their limited effect on homework completion. Some centers set aside only about 20 minutes for homework, which may not have been adequate for students to complete it. Whether program staff members knew what homework had been assigned also varied widely. Most sites had no procedures for monitoring homework assignments. Two of the three sites that had monitoring procedures relied on teachers to communicate assignments to center staffers. Indications were that teachers did not consistently do so. In some centers, site visitors observed homework sessions in which students were not engaged, program staff members made little effort to maintain order, and students received help only if they asked questions or made an effort to seek out help from staffers.¹⁸

4. Centers Did Not Improve Academic Outcomes

Centers did not improve reading test scores and grades in math, English, science, or social studies (Table II.4).¹⁹ Treatment students had an average reading score of 35.0 (in percentiles) on the Stanford Achievement Test—Version 9 (SAT-9) and control students had an average

¹⁸The After School Corporation's evaluation noted that the quality of homework help was low in many programs it studied, which it attributed to a lack of skills and knowledge among young and often part-time staffers who provided the help (Reisner et al. 2001). The Forum for Youth Investment characterized homework help in after-school programs as "spotty at best" Forum for Youth Investment (2003).

¹⁹The sample sizes enabled the study to have reasonable power to detect an effect size of 0.10 for reading test scores, which is equivalent to a change in reading scores of 3.2 percentile points.

reading score of 35.9.²⁰ No impacts were evident for student effort and preparedness for class (as reported by teachers) or for absenteeism (as reported in student records).

5. Centers Increased Feelings of Safety

Centers reduced whether students felt unsafe after school (Table II.6). Two percent of treatment students reported feeling “not at all safe” after school, compared to five percent of controls (effect size of .15).²¹ The increase in feelings of safety indicates that centers were meeting an objective that many indicated was a priority for them (see Section A.1).

6. Centers Increased Some Types of Parent Involvement

Parents of treatment students were more likely to help their child with homework, to ask about their child’s class work, and to attend after-school events (Table II.6). The increase in parents’ helping with homework seems counterintuitive for programs that provided homework help, but nonetheless is a form of involvement that may reflect greater parental engagement in their child’s education.²² Forty-six percent of parents of treatment students attended at least three after-school events in the past year, compared to 36 percent of parents of control students (effect size of .20). Centers did not improve attendance at school open houses or parent-teacher organization meetings, or the extent to which parents volunteered at school.

²⁰Baseline reading scores were imputed by calculating the mean baseline reading score among students with a baseline score and assigning the mean score to students who were missing the baseline score. Handling missing baseline reading scores in other ways, such as estimating impacts only for students with baseline reading scores and excluding the baseline reading score from the list of regressors, did not change the findings.

²¹The first report had similar impact estimates, but the impacts were statistically insignificant.

²²The results on parents’ helping with homework in Table II.6, which are based on data reported by parents, appear somewhat at odds with the results in Table II.5, which are based on data reported by the smaller sample of students. However, the items in Table II.6 focus more on the frequency of asking about homework and checking it three or more times a week, which can differ from whether parents asked about or checked homework at all.

Table II.6

Impacts on Other Outcomes, Elementary School Centers, Year 1

Outcome	Treatment Group	Control Group	Estimated Impact	Estimated Impact on Participants
Percentage of Students Who Reported Feeling the Following Levels of Safety After School up Until 6 p.m.:				
Very safe	76.8	75.3	1.5	0.6
Somewhat safe	21.7	20.3	1.4	3.2
Not at all safe	1.5	4.5	-3.0**	-3.8
Percentage of Students Who Reported the Following Are "Somewhat True" or "Very True":				
They get along with others their age	81.0	86.0	-5.1	-8.5**
They feel left out of things	32.5	32.4	0.1	-0.5
Percentage of Students Who Reported Doing the Following "Some" or "A Lot":				
Help another student in school	75.2	79.4	-4.2	-6.1
Help another student after school	60.1	52.2	8.0**	10.5
Percentage of Students Who Rated Themselves as "Good" or "Excellent" on the Following:				
Working with others on a team or in a group	78.8	81.6	-2.8	-2.0
Feeling bad for other people who are having difficulties	70.4	74.2	-3.9	-6.2
Believing the best about other people	79.4	79.5	-0.1	0.2
Percentage of Students Who Rated Themselves as "Excellent" on the Following:				
Using a computer to look up information	48.2	46.6	1.6	3.4
Setting a goal and working to achieve it	57.2	59.1	-2.0	-2.5
Percentage of Students Who Rated Themselves as "Excellent" on Sticking to What They Believe In, Even if Their Friends Don't Agree				
	56.1	56.8	-0.7	-0.8
Negative Behavior Composite ^a	1.6	1.7	0.0	0.0
Percentage of Students Whose Parents:				
Helped their child with homework at least three times last week	69.1	60.7	8.4***	8.9**
Checked on their child's homework completion at least three times last week	92.4	90.3	2.1	2.2
Asked their child about things they were doing in class at least seven times last month	70.4	64.1	6.3**	8.1
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:				
Attended an open house at the school	42.3	42.3	0.0	0.3
Attended parent-teacher organization meetings	50.1	47.6	2.6	3.1
Attended an after-school event	45.6	36.3	9.2***	12.4***
Volunteered to help out at school	29.9	33.9	-4.1	-4.7
Sample Size^b	862	677		

SOURCE: Student Survey, Parent Survey.

NOTE: The tables show two types of impact estimates: (1) "intent to treat" estimates (in the "Estimated Impact" column) use the full treatment and control groups and (2) impacts on participants (in the "Estimated Impact on Participants" column) are the impacts after adjusting for the percentage of treatments who did not attend centers ("no-shows") and the percentage of controls who attended centers ("cross-overs"). The percentages and mean values of outcomes for treatment and control students have been regression-adjusted for baseline differences between the groups. The control variables in the regression included students' demographic characteristics, students' baseline test scores, and school attendance. Weights are used to adjust impact estimates for nonresponse. Impacts on participants are estimated using an instrumental variables method, and the significance levels may differ from significance levels of the intent-to-treat estimates. Appendix B describes methods used to estimate impacts.

^aThe negative behavior composite is based on student responses to five questions regarding how often they: (1) break something on purpose, (2) punch or hit someone, (3) argue with their parents, (4) lie to their parents, and (5) give a teacher a “hard time.” Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level, while a value of 4 indicates a high level. Percentages may not sum to 100 because of rounding.

^bSample sizes differ for outcomes depending on the source. For some parent-reported outcomes, the sample sizes are 841 treatment group members and 663 control-group members; for student-reported outcomes, the sample sizes are 583 treatment-group members and 468 control group members.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

7. Few Improvements in Social and Interpersonal Outcomes

Centers did not improve most of the social and interpersonal skills that the study assessed. Treatment students were no more likely than control students to report getting along with others their age, to rate themselves highly on working with others on a team or in a group, or to rate themselves highly on sticking to their beliefs even if their friends do not agree (Table II.6). Treatment students were more likely to report helping other students after school (effect size of .16), which may reflect interactions students had with each other after school.

8. Few Impacts for Subgroups

It is possible that programs might be more effective with some types of students; if this is the case, programs could target their services to those students. For example, centers may be interested in targeting reading instruction to younger students, or academic help to students with low test scores. The study team estimated impacts for six subgroups defined by student or parent characteristics: (1) grade level, (2) whether students were above or below the median reading test score at baseline, (3) whether students were above or below the median value of the discipline composite variable at baseline (4) student race and ethnicity, (5) student gender, and (6) whether students lived in two-parent or one-parent households (see Tables II.7a through II.8b).²³ The full set of subgroup impacts estimated is presented in Appendix C; a smaller set of tables, which focus on key outcomes and subgroups, is presented here.

Tables II.7b and II.8b show that program impacts differed for students from two-parent households compared to students from single-parent families for some outcomes. The number of

²³We also estimated impacts for a subgroup defined by whether students participated in after-school programs (the 21st Century program or another after-school program) in the spring prior to the start of the study, to assess whether previous after-school program participation was associated with impacts. Subgroup impacts were not significantly different for students who had and students who had not participated in after-school programs.

Table II.7A

Impacts on Location and Care After School, Student Effort, Maternal Employment, and Student Discipline Outcomes for White, Black, and Hispanic Subgroups, Elementary School Centers, Year 1

Outcome	White (Non-Hispanic)			Black (Non-Hispanic)			Hispanic		
	Treatment Mean	Control Mean	Estimated Impact	Treatment Mean	Control Mean	Estimated Impact	Treatment Mean	Control Mean	Estimated Impact
Percentage of Students in the Following Locations After School at Least 3 Days in Typical Week (According to Parents):									
Own home	62.4	91.9	-29.5***	57.8	75.6	-17.8***	72.9	88.2	-15.3***
Someone else's home	5.8	8.3	-2.5	17.3	15.8	1.5	10.2	16.2	-6.0
School or other place for activities	46.2	20.8	25.5***	55.3	29.8	25.6***	43.5	17.9	25.6***
Somewhere to "hang out"	4.5	7.6	-3.1	4.3	1.3	3.1**	1.1	0.6	0.4
Mixed (no one location for at least 3 days)	3.1	0.3	2.8	0.5	1.2	-0.6	1.8	0.3	1.5
Percentage of Students with the Following Individuals After School at Least 3 Days in Typical Week (According to Parents):									
Self-care ^a	0.0	0.0	n.a. ^b	2.0	1.6	0.4	0.7	0.3	0.4
Parent	77.3	87.6	-10.3	60.1	71.9	-11.8***	78.7	78.8	0.0
Non-parent adult	38.5	25.6	12.9	46.4	32.1	14.3***	38.0	32.7	5.3
Sibling	5.9	12.3	-6.5	21.1	21.8	-0.7	25.7	37.1	-11.4**
Mixed (no one category for at least 3 days)	2.6	1.3	1.3	1.9	1.9	0.0	0.6	2.0	-1.4
Employment of Mother:									
Full-time	54.4	42.4	12.0	54.1	54.6	-0.5	51.9	44.8	7.1
Part-time	12.8	20.5	-7.7	18.6	14.9	3.7	10.6	18.3	-7.6
Looking for work	13.1	9.8	3.3	13.3	12.2	1.2	19.6	11.4	8.2
Not in labor force	19.8	27.3	-7.6	14.0	18.3	-4.4	17.9	25.5	-7.6
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" that the Student Completes Assignments to Their Satisfaction									
	64.2	58.5	5.7	52.0	54.9	-2.9	54.5	62.7	-8.2
Percentage of Students Whose Teachers Reported that the Child "Often" Performs at or Above His/Her Ability									
	57.7	50.2	7.5	38.5	37.0	1.5	46.3	50.0	-3.7
Teacher-Reported Level of Effort Composite (Mean)									
	3.9	3.7	0.2	3.5	3.6	-0.1	3.8	3.8	0.0
Student-Reported Disciplinary Problems Composite (Mean)									
	1.3	1.3	0.0	1.7	1.6	0.1	1.3	1.6	-0.2
Percentage of Students Who Were Suspended									
	7.7	0.0	7.7	6.8	3.5	3.3	0.3	4.6	-4.2**
Number of Observations:									
Student-reported outcomes		58			474			273	
Teacher-reported outcomes		95			739			464	
School records outcomes (suspensions)		50			531			227	
Parent-reported outcomes		88			843			474	

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Percentages may not sum to 100 because of rounding.

^aStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least three days in a typical week.

^bNo white students were reported to be in self-care.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table II.7B

Impacts on Location and Care After School, Student Effort, Maternal Employment, and Student Discipline Outcomes for Two-Parent and One-Parent Subgroups, Elementary School Centers, Year 1

Outcome	Two-Parent Household Structure ^a			One-Parent Household Structure ^a		
	Treatment Mean	Control Mean	Estimated Impact	Treatment Mean	Control Mean	Estimated Impact
Percentage of Students in the Following Locations After School at Least 3 Days in Typical Week (According to Parents):						
Own home	62.6	81.0	-18.5***	59.9	76.8	-16.9***
Someone else's home	11.8	13.0	-1.3	15.8	17.1	-1.3
School or other place for activities	50.0	33.0	17.0***	52.2	27.6	24.6***
Somewhere to "hang out"	2.2	5.4	-3.2**	5.9	3.0	2.9
Mixed (no one location for at least 3 days)	1.9	0.3	1.5	1.2	1.4	-0.2
Percentage of Students with the Following Individuals After School at Least 3 Days in Typical Week (According to Parents):						
Self-care^b	0.5	2.4	-1.9	1.9	0.9	1.0
Parent	68.0	75.9	-7.9	63.5	72.1	-8.6**
Non-parent adult	42.6	33.5	9.1	45.2	33.6	11.6***
Sibling	24.4	32.5	-8.1	19.9	17.8	2.1
Mixed (no one category for at least 3 days)	2.1	1.5	0.5	2.4	1.4	1.0
Employment of Mother:						
Full-time	60.4	56.6	3.8	49.6	52.9	-3.4
Part-time	15.4	14.1	1.4	15.8	14.9	1.0
Looking for work	13.5	9.6	3.9	16.0	13.3	2.7
Not in labor force	10.7	19.8	-9.1**	18.6	18.9	-0.2
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" that the Student Completes Assignments to Their Satisfaction						
	56.4	60.2	-3.8	52.2	53.4	-1.2
Percentage of Students Whose Teachers Reported That They "Often" Perform at or Above Their Ability						
	40.0	49.1	-9.0	42.7	35.5	7.2
Teacher-Reported Level of Effort Composite (Mean)						
	3.6	3.7	-0.2	3.6	3.6	0.1
Student-Reported Disciplinary Problems Composite (Mean)						
	1.5	1.5	0.1	1.6	1.7	-0.1
Percentage of Students Who Were Suspended						
	5.6	4.5	1.2	6.0	5.9	0.2
Number of Observations:						
Student-reported outcomes		396			437	
Teacher-reported outcomes		647			750	
School records outcomes (suspensions)		431			477	
Parent-reported outcomes		797			900	

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Percentages may not sum to 100 because of rounding.

^aStudents are in the "two-parent" subgroup if they live with a mother, stepmother, foster mother, or female guardian and a father, stepfather, foster father, or male guardian. If they do not live with both a male and a female parent or guardian, students are in the "one-parent" subgroup.

^bStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least three days in a typical week.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table II.8A

Impacts on Student Attendance, Academic Achievement, and Other Outcomes for White, Black, and Hispanic Subgroups, Elementary School Centers, Year 1

Outcome	White (Non-Hispanic)			Black (Non-Hispanic)			Hispanic		
	Treatment Mean	Control Mean	Estimated Impact	Treatment Mean	Control Mean	Estimated Impact	Treatment Mean	Control Mean	Estimated Impact
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6 p.m.:									
Very safe	73.0	76.3	-3.3	77.0	76.1	0.9	69.9	70.5	-0.5
Somewhat safe	27.0	23.7	3.3	20.8	22.6	-1.9	29.4	25.5	3.9
Not at all safe	0.0	0.0	0.0	2.2	1.2	1.0	0.6	4.0	-3.4
Mean Number of Days School Records Indicate Student Was:									
Absent	7.0	6.6	0.4	7.7	7.3	0.4	5.7	7.5	-1.8**
Late	4.5	3.5	1.0	4.9	3.9	1.0	1.6	4.0	-2.4***
Mean Class Grade:									
Math	87.8	85.2	2.6	80.3	79.5	0.8	83.0	84.2	-1.2
English	87.4	83.6	3.8	81.7	81.7	-0.1	83.5	83.2	0.3
Science	88.4	83.6	4.8	81.6	81.3	0.4	85.1	84.7	0.4
Social Studies	89.4	85.4	4.0	81.9	80.3	1.7	81.3	82.4	-1.2
Mean Reading Test Score	51.7	51.3	0.4	34.3	34.9	-0.5	34.3	37.2	-2.8
Percentage of Students Who Reported Helping Another Student After School									
	70.2	55.8	14.4	65.4	50.1	15.4***	49.7	49.2	0.5
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:									
Attended an open house at school	32.3	35.7	-3.4	46.3	39.2	7.1	51.3	54.7	-3.4
Attended a PTO meeting	48.9	31.3	17.7	49.3	48.4	0.9	54.2	47.7	6.5
Attended an after-school event	39.2	33.4	5.7	46.3	32.9	13.4***	38.8	32.4	6.4
Volunteered to help out at school	21.5	24.5	-3.0	30.8	34.6	-3.8	23.0	23.5	-0.6
Number of Observations:									
Student-reported outcomes		58			473			271	
Parent-reported outcomes		86			838			465	
School records outcomes (attendance)		86			786			468	
School records outcomes (grades)		49			679			420	
School records outcomes (reading scores)		99			785			474	

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Percentages may not sum to 100 because of rounding.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table II.8B
Impacts on Student Attendance, Academic Achievement, and Other Outcomes for Two-Parent and One-Parent Subgroups, Elementary School Centers, Year 1

Outcome	Two-Parent Household Structure ^a			One-Parent Household Structure ^a		
	Treatment Mean	Control Mean	Estimated Impact	Treatment Mean	Control Mean	Estimated Impact
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6 p.m.:						
Very safe						
Somewhat safe	74.6	67.2	7.5	70.2	81.5	-11.3
Not at all safe	24.7	30.6	-5.9	27.8	15.5	12.3**
	0.7	2.2	-1.5	2.0	3.1	-1.1
Mean Number of Days School Records Indicate Student Was:						
Absent	6.6	6.9	-0.3	7.8	7.5	0.2
Late	3.8	3.6	0.2	4.4	3.5	0.9
Mean Class Grade:						
Math	83.4	82.8	0.5	80.3	80.6	-0.2
English	83.9	83.3	0.6	81.5	81.9	-0.4
Science	84.4	81.3	3.0**	81.5	81.8	-0.2
Social Studies	83.4	82.1	1.2	81.1	80.6	0.5
Mean Reading Test Score	38.6	41.9	-3.3	33.8	31.6	2.2
Percentage of Students Who Report Helping Another Student After School	56.4	44.4	12.0	62.0	54.4	7.5
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:						
Attended an open house at school	40.0	39.5	0.4	42.4	45.8	-3.4
Attended a PTO meeting	53.2	47.9	5.3	47.0	47.1	-0.1
Attended an after-school event	47.4	33.6	13.8***	42.3	38.6	3.7
Volunteered to help out at school	38.5	30.6	8.0	24.0	36.6	-12.6***
Number of Observations:						
Student-reported outcomes		393			438	
Parent-reported outcomes		710			804	
School records outcomes (attendance)		693			764	
School records outcomes (grades)		568			646	
School records outcomes (reading scores)		693			770	

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Percentages may not sum to 100 because of rounding.

^aStudents are in the "two-parent" subgroup if they live with a mother, stepmother, foster mother, or female guardian and a father, stepfather, foster father, or male guardian. If they do not live with both a male and female parent or guardian, students are in the "one-parent" subgroup.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

impacts that differed significantly between students in two-parent and one-parent households was large enough to be unlikely to occur by chance alone. However, being in a two-parent household was also correlated with other factors, such as being Hispanic and having low discipline problems at baseline.

The evaluation investigated whether controlling for membership in other subgroups modified the two-parent findings, and found that the effects on self-care, safety, and parental volunteering were no longer significant after controlling for other subgroups. This suggests that being in a two-parent household *per se* may not be the factor that is giving rise to the differing impacts, but that factors related to being in a two-parent household may be.²⁴

Hispanic students in the treatment group were less likely to be suspended and late for school less often (Tables II.7b and II.8b). However, only two sites had enough Hispanic students to be included in the Hispanic subgroup impact analysis, and only one site had significant impacts.²⁵ Because of these limits, it is not clear whether the results should be attributed to the Hispanic subgroup or to the particular site in which the results were observed.

²⁴The data indicate that students in the two-parent household subgroup were more likely to be in other subgroups, such as the Hispanic subgroup and low baseline discipline problems subgroup, and membership in another subgroup may be responsible for a significant estimate for the two-parent subgroup. To assess how membership in other subgroups might affect the two-parent impacts reported in the table, we first estimated models that controlled only for membership in the two-parent and one-parent subgroups. We then estimated models that controlled for membership in other subgroups.

²⁵Because impacts are estimated as an average of site impacts, the estimation procedure used a minimum threshold of five treatment students and five control students for each site. Sites that did not meet the threshold were excluded from the estimation for that subgroup. For example, a site with seven control students and four treatment students for a particular subgroup would not be included in the estimation for that subgroup.

III. Implementation and Outcome Differences at Middle School Centers

The second year of follow-up data for middle school students enables us to examine longer-term outcome differences between the treatment and comparison groups and assess whether outcomes are affected after two school years. The first report examined academic, social, emotional, and other outcomes after one school year.

This chapter begins with a discussion of program implementation at middle school centers, focusing on key features in the study's second year of data collection. It then analyzes patterns of student attendance at centers during the second year and presents outcome differences between the treatment and comparison groups and for different types of students.

Generally, the second-year findings are consistent with the first-year findings. As in the first year, students attended centers infrequently. Students who attended in the second year averaged about 30 days of attendance, and many did not attend at all in the second year. As in the first year, treatment students spent more time at school during after-school hours than comparison students, had a higher grade in one of the four subjects for which the evaluation collected grades (in this case, social studies), and had higher educational expectations. There were no differences between treatment and comparison students on other academic outcomes or on feelings of safety after school. There was mixed evidence on negative behavior outcomes; treatment students had higher levels of negative behavior than comparison students on some outcomes, but there was no difference between the two groups on other outcomes.

A. Middle School Centers in the 2001-2002 School Year

This section sketches key program features in the 2001-2002 school year and discusses two changes that were evident between the first and second school years: (1) center staff indicated that they shifted activities and services toward more academic programming and (2) centers experienced high rates of staff turnover.

1. Center Goals and Structure

In the 2001-2002 school year, administrators of the 21st Century centers indicated that their major objectives were to help students improve academically and to provide a safe place for them after school. The first year report noted the same objectives.

Centers designed services and activities to appeal to students who had other options for their after-school time. A common approach was to let students choose most of their activities; more than half of the centers let students choose all their activities. When the centers restricted choice, they typically

Percent of Project Directors Indicating Item as One of Three Most Important Objectives	
Provide a Safe, Supervised After-School Environment	75%
Provide Tutoring/Other Activities to Enhance Students' Ability to Meet Specific Academic Goals	66%
Create a Positive Relationship Between Students and Their Schools	37%
Improve Homework Completion	33%
Enhance Social Development	33%
Provide Academic Enrichment	12%
Provide Cultural Opportunities Not Available at Home or in the Community	8%
Provide Sports/Recreation Opportunities	8%
SOURCE: Project Director Survey. Sample size is 24 programs operating with 21st Century funds.	
NOTE: Percentages do not sum to 100, because project directors could indicate up to three "most important" objectives.	

required students to participate first in an academic activity (usually a homework session) before taking part in recreational or cultural activities.

Homework help was the most prevalent academic activity, with about 80 percent of centers offering it. Generally, centers grouped students in common areas or in classrooms. Students worked on their own and could ask session monitors—typically teachers or paraprofessionals—

for help. Most centers did not have procedures in place to monitor homework assignments or to ensure that students completed the assigned homework. As in the visits during the first year, site visitors noted that homework help was an opportunity for students to do homework, but not one that students always took advantage of. Section C.2 provides more discussion of homework issues.

Site visitors reported that 60 percent of centers offered other types of academic activities, usually focused on help in reading, writing, or math. The format typically was

Examples of Academic Activities in 21st Century Centers

- Teacher instruction
- Educational technology packages to reinforce basic skills or supplement classroom instruction
- Practice drills and games to improve reading, writing, long division, multiplication
- Preparation for standardized tests, such as taking and reviewing practice tests

small-group instruction, with a teacher working with a group of students from the same grade on particular subject matter or skill development exercises. Some centers helped students prepare for state assessment tests, such as by administering practice tests and identifying areas in which students needed more help.

Most centers provided recreational, cultural, and developmental activities to students. Of these other activities, recreation was the most common. Site visitors reported that 74 percent of centers provided recreation at least

weekly, which often involved learning a particular game or skill (tennis or martial arts, for example). Centers also offered unstructured recreation, such as basketball or board games. Almost two-thirds of

Examples of Other Activities in 21st Century Centers

- Recreation: basketball, martial arts, cheerleading, board games, table tennis, swimming, free time in playground or gym
- Cultural enrichment: art and music classes, choir practice, dance and drama classes, cooking classes, trips to museums and theater, classes promoting awareness of different cultures
- Interpersonal development: team-building activities, leadership training activities, peer mediation and conflict resolution activities, teen discussion groups

centers had regular activities devoted to music, art, or other forms of cultural appreciation.

Interpersonal development activities—focused on students’ behavior and their relationships with others—were the least common; 42 percent of centers offered such activities weekly or daily.

2. Centers Reported Placing Greater Emphasis on Academics

Between the first and second year, centers reported shifting activities toward academics. Site visitors noted that about 75 percent of centers reported increasing their academic activities; almost 80 percent of principals indicated doing so. Data from center coordinator and project director surveys also showed reported increases in such activities. While our site visits cannot provide data to verify this shift occurred, there clearly was a perception of a shift in focus toward academic activities.

Among the reasons for this increased emphasis on academics were growing concerns about student academic performance in general, and test scores in particular. Site visitors noted that relatively few centers in the previous year (7 percent) had said that helping students on assessment tests was an objective, but in the second year about 20 percent said so. Center administrators, noting in interviews that centers were a way to improve the achievement of low-performing students, targeted academic services to these students or increased the academic content of centers in other ways.

Making room for this shift meant reducing nonacademic activities. For example, one center dropped some of its enrichment offerings to make room for literacy activities and tutoring sessions in math. Another center eliminated most of its enrichment activities to focus on providing extra academic help to students.

3. Centers Experienced Heavy Staff Turnover

Between the study's first and second year, centers experienced changes in staff at all levels.²⁶ Staff members were the most likely to turn over; two-thirds were not working for the centers in the study one year later. Almost one-third of schools operating a 21st Century center had a new principal, and almost one-third of the centers had a new coordinator. Project directors experienced the lowest turnover, with less than a fifth of grantees having a new project director.

Staff Turnover at Middle School Centers	
Percent Turnover In Two Years	
Principals	30.4
Project Directors	17.9
Center Coordinators	31.7
Line Staff	65.1
SOURCE: Questionnaires and Program Records. The sample size in the 2000-2001 school year was 46 principals, 28 project directors, 41 center coordinators, and 555 staff members.	

The high staff turnover is similar to turnover found in a national study of after-school programs (Seppanen et al. 1993) and higher than turnover found in child-care settings (Whitebook et al. 1998). Some turnover resulted from grantee efforts to substitute school-district staffers for staffers of outside organizations. Four grantees said they wanted to rely less on outside organizations so they could more effectively monitor services; two grantees no longer had the financial resources to purchase services from outside organizations; and one grantee made a policy decision to rely more on teachers as staff members.²⁷ Some turnover also may have resulted from administrators' efforts to scale back staffing. For example, as grant funds declined, administrators of one program reduced the amount of activities offered, which reduced

²⁶The study estimated staff turnover by comparing the names of staff members who were listed for the programs or schools in the two years. Grantees that no longer operated 21st Century programs as of fall 2001 were not included in the estimates.

²⁷Survey data confirm the hiring shift toward school district employees. The percentage of coordinators employed directly by the 21st Century program rather than employed by community or nonprofit organizations rose from 80 to 91 percent, and the percentage of other staff members employed by 21st Century programs rather than by community or nonprofit organizations rose from 82 to 92 percent.

the program's staffing needs. Both factors would contribute to the turnover numbers, but would not represent decisions by staff members to leave their center jobs. Even after accounting for these factors, however, staff turnover was substantial.

In surveys, administrators rated staff turnover as a minor issue for their centers, but indicated to site visitors that hiring new staff took more time than they had expected. Turnover of more senior administrators had mixed effects on centers. For example, site visitors observed that new principals at some host schools supported the program more, while new principals at other host schools supported it less. A new principal at one school, for example, moved the center's office from a portable classroom far from the main school building into the building itself. In another school, the new principal moved the center's office from next to the principal's office into the basement. In a third school, the relationship between the school and the center was unaffected when an assistant principal who shared similar views about the center became the new principal.

Whether turnover can be reduced is unclear. On a survey, most staffers who said they did not expect to return the following year cited personal commitments as the reason. During the site visits, center staffers noted that teachers who worked for the program usually left because of other commitments and because of burnout from teaching both during and after school.

B. Student Attendance Was Low in the Second Year

Two key attendance patterns for middle school students emerge in the second year. First, many students did not return to programs after having attended in the previous year. Second, among returning students who attended in the second year, attendance was low, comparable to the low attendance levels observed in the first year.

About 59 percent of students who attended in the first year no longer had access to centers in the second year, because they had gone on to high school or had transferred to other middle

schools that did not operate centers. Among the remaining 41 percent who had access to centers, about half (47 percent) attended. The remaining 53 percent did not attend (Figure III.1 shows the breakdown). For the total sample of students who had participated in centers in the first year, average attendance in the second year was nine days (Table III.1).

Among students who participated in the second year, attendance levels were similar to those in the first year. Average attendance in the second year was 30 days, compared with 33 days in the first year (Table III.1). Eighty percent of students attended 50 days or less, and 59 percent

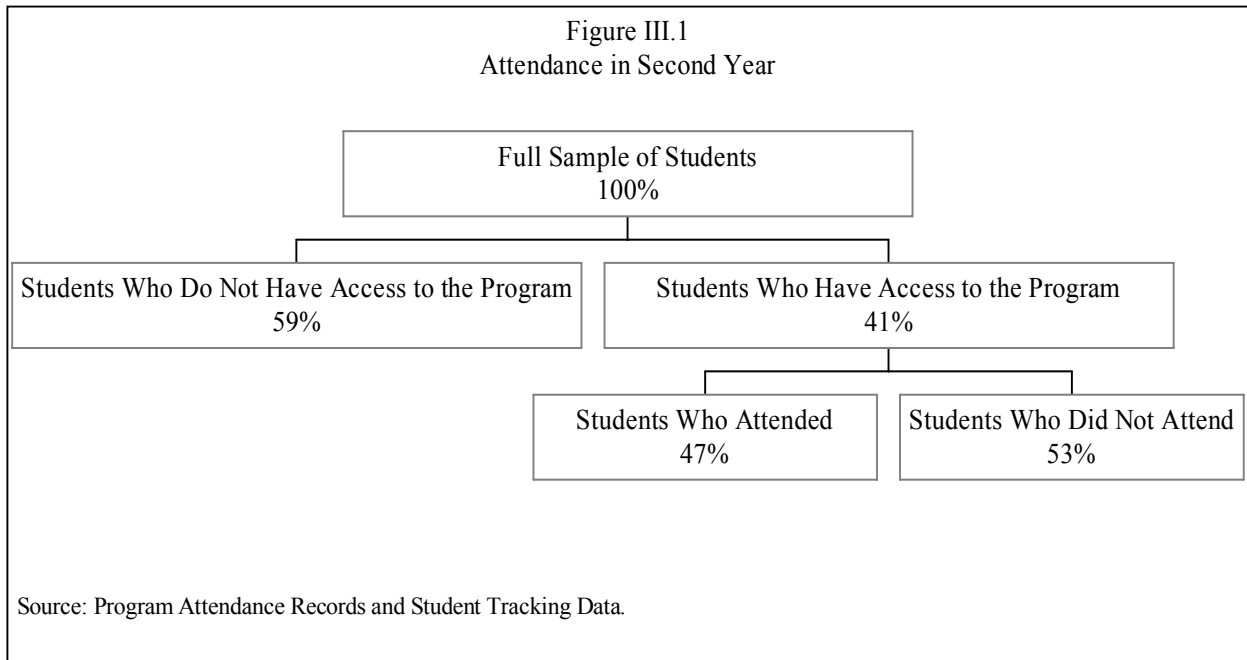


Table III.1

21st Century Middle School Center Attendance, Year 2

	All Treatment Students	Participating Treatment Students
Percentage of Students Who Attended the Program in the 2001-2002 School Year	29.5	100.0
Average Days Attended in 2001-2002	8.8	29.5
Number of Days Attended (Percentage of Students)		
0	70.0	0.0
1 to 25	17.7	59.0
26 to 50	6.4	21.3
51 to 75	2.8	9.2
76 to 150	3.1	10.4
Attendance Rate ^a (Percentage of Students)		
10 or less	80.7	35.2
11 to 25	6.9	23.2
26 to 50	6.3	21.3
51 to 70	2.4	8.1
71 to 85	2.1	7.0
86 to 100	1.5	5.2

SOURCE: Center Attendance Records. The sample size for all treatment students is 1,629. The sample size for participating treatment students—students who attended the program at least one day in the 2001-2002 school year—is 488.

^aThe attendance rate is the number of days students attended as a proportion of the number of days centers were open, which centers provided in their annual performance reports. Totals may not add to 100 percent because of rounding.

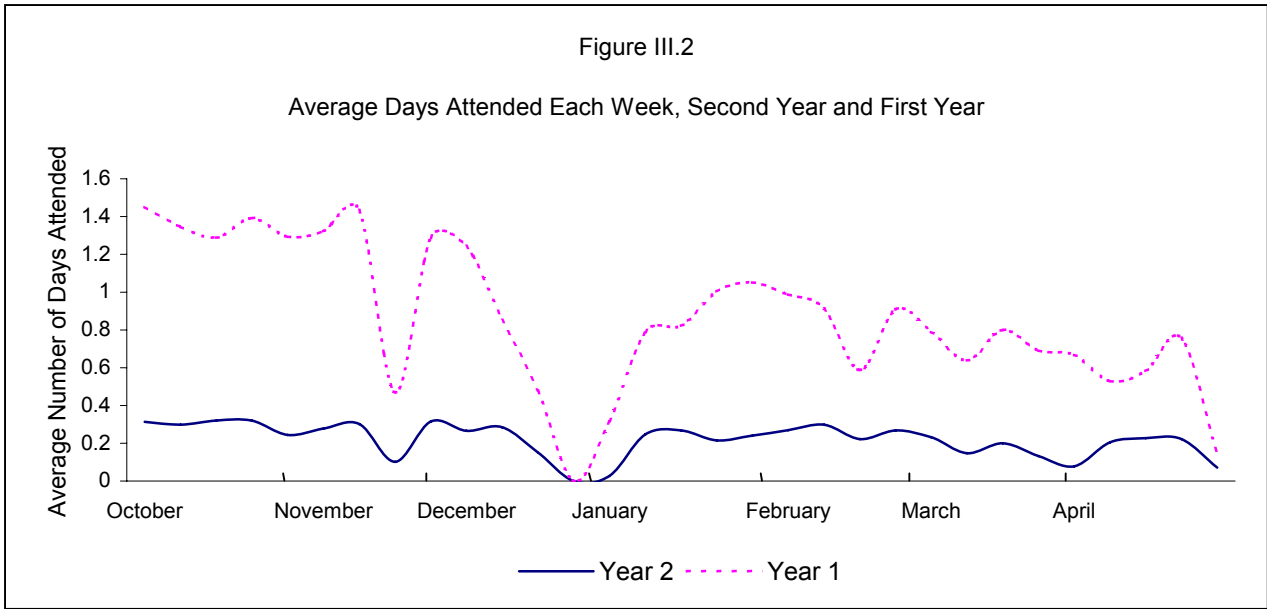
attended for 25 days or less. Sixty percent attended less than one-quarter of the days that centers were open (which averaged about 96 days). Students who participated in the second year were younger (in 6th or 7th grade rather than in 8th), as would be expected since many older students moved on to high school, but they were also more likely to be white, speak English at home, and had mothers who were more highly educated.

The week-to-week pattern of attendance in the second year was similar to the pattern in the first year (Figure III.2). The average number of days a week students attended fell throughout the school year and declined sharply around major holidays. The figure shows that the average attendance for the second year was well under the average attendance in the first year, because of the large number of students not attending during the second year. Restricting the sample to only students who attended (at least one day) shows that the attendance pattern in the second year was similar to that in the first year, with the same average frequency and about the same seasonal pattern (Figure III.3).

Additional analysis found large differences in average student attendance across grantees. For example, one grantee had average student attendance of 17 days a year, whereas another had average student attendance of 39 days a year. Variations in average attendance across grantees explained much of the overall variation in student attendance.²⁸

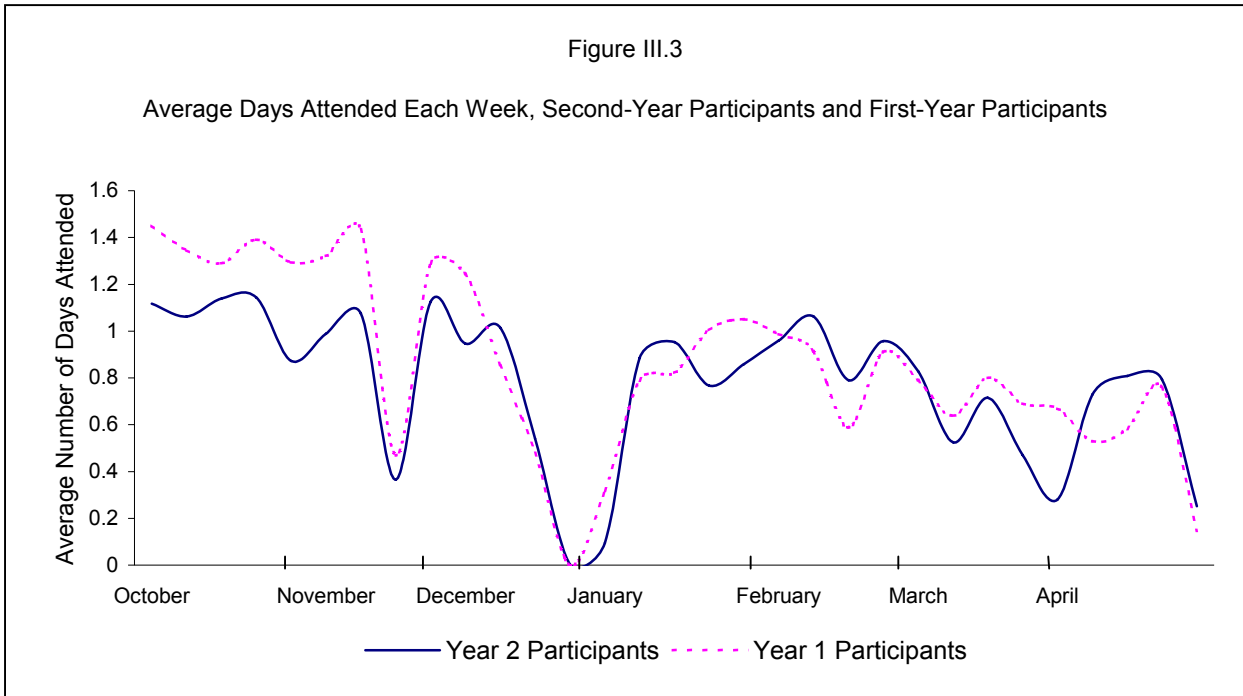
Few student characteristics were related to frequency of attendance at centers. We investigated almost 30 characteristics and only a few were statistically significant. Students who attended more frequently during the second year were younger, were more likely to be over-age for their grade, had fewer school-day absences, and rated their school more highly.

²⁸Models of student attendance explained 36 percent of its variation; 33 percent of the variation was explained by grantee variables, and 3 percent was explained by student characteristics.



Source: Center Attendance Records.

Note: Figure includes all participating students except students who transferred during the year.



Source: Center Attendance Records.

Note: Only students who attended in the second year are included, except for students who transferred during the year.

C. Middle School Center Outcome Differences

Table III.2 shows that the propensity score matching technique used to create the study's comparison group yielded groups that were similar along many dimensions. For example, the two groups both had slightly more females than males and similar racial and ethnic composition.

Some characteristics of the groups differed. Treatment group students had lower average grades, less-regular homework habits, and more discipline problems than comparison group students.²⁹ Regression models were used to adjust for these differences; Appendix Table C.4 presents results of tests of the effectiveness of the regression adjustments.³⁰ The tests show that the adjustments typically reduced differences to be statistically insignificant or close to insignificant.

How differences that remain after regression adjustment might affect the estimates is unclear. If students attending programs are more disadvantaged than comparison group students, for example, and if centers generally have larger impacts on disadvantaged students, the estimates here will overstate true effects of the programs. The study can only speculate on the direction because it has not been established by previous research. The method also may overestimate effects for some outcomes and underestimate effects for others. The study views the measured outcome differences as indications of how centers affected various outcomes, but recognizes that a more rigorous experimental design may have yielded findings that differ from those reported here.

²⁹A large number of variables are reported in Table III.2 and some differences will arise by chance. This caution also applies to the impact estimates reported in this chapter; because a large number of statistical tests were conducted, some findings may be significant by chance alone.

³⁰Other tests of the results were run, including tests of the sensitivity of impact results to the use of nonresponse weights. These tests are presented in Appendix C.

Table III.2

Characteristics of Center Participants and Comparison-Group Students:
Middle School Centers

Characteristic	Percentage of Program Participants	Percentage of Comparison-Group Members	p-value ^a
Demographics			
Gender			
Male	47.3	46.5	0.62
Female	52.7	53.5	0.62
Race/Ethnicity			
White (non-Hispanic)	38.2	40.6	0.33
Black (non-Hispanic)	27.7	24.7	0.33
Hispanic	12.3	12.0	0.33
Other	15.5	15.9	0.33
Mixed race	6.3	6.9	0.33
Grade Level			
6	20.7	21.6	0.19
7	37.8	38.2	0.19
8	33.7	34.1	0.19
Other or ungraded	7.8	6.2	0.19
Primary language in the home is not English	17.8	18.9	0.39
Academic and Other Outcomes at Baseline			
Student-Reported Baseline Grades			
Mostly A's	30.4	34.1	0.00***
Mostly B's	35.8	36.5	0.00***
Mostly C's	23.2	21.3	0.00***
Mostly D's or below	8.8	7.5	0.00***
Not graded	1.8	0.7	0.00***
Average Grades	83.1	84.0	0.01***
Homework			
Mother or father helps student with homework	63.1	63.2	0.93
Mean of homework habits index ^b	2.80	2.85	0.02**
Mean of Index of Positive Behavior ^c	3.0	3.0	0.52
Student-Based Discipline Problem Composite ^d (Mean)	1.39	1.33	0.00***
Mean of Parental Discipline Index ^e	2.9	2.9	0.46
Negative Behavior Composite ^f (Mean)	1.55	1.52	0.07
Mean of Index of Empathy ^g	3.1	3.1	0.94
Mean of Index of Controlling Destiny ^h	3.0	3.0	0.81
Student-Reported Tobacco, Alcohol, and Drug Use Composite (Mean) ⁱ	1.12	1.11	0.10
Mean of Safety Index ^j	3.33	3.37	0.03**
Sample Size^k	1,727	2,385	

SOURCE: Student Survey, Parent Survey.

NOTE: Percentages may not sum to 100 because of rounding.

Table III.2 (continued)

^a The p-value is the smallest level of significance at which the null hypothesis that the difference in means between program participants and comparison group members equals zero can be rejected. If the p-value is less than .05, the difference is significant at the 5 percent level, and if the p-value is less than .01, the difference is significant at the 1 percent level.

^b The homework habits index is based on student responses to how often they: (1) did the homework the teachers assign, (2) do homework in the same place each day, (3) do homework at the same time each day, and (4) write down homework assignments. The index is equal to the mean of the four variables. A value of 1 on the index indicates poor homework habits, whereas a value of 4 indicates good homework habits.

^c The positive behavior index is based on how often the student: (1) helps another kid in school, (2) helps her parents, and (3) goes to church, temple, or mosque. A value of 1 on the index indicates never doing the aforementioned, while a value of 4 indicates doing them often.

^d The student-based discipline problem composite is based on four items: the extent to which students report (1) skipping school or class, (2) getting sent to the office for doing something wrong, (3) getting detention, and (4) having their parents called to school about a problem they are having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems.

^e The parental discipline index is based on student responses to how often parents: (1) check on whether homework is completed, (2) limit the amount of time available to watch TV, (3) decide which TV shows their kids are allowed to watch, and (4) tell their children not to drink alcohol or use drugs. A value of 1 on the composite indicates parents who engage in less discipline, while a value of 4 indicates parents who engage in more discipline.

^f The negative behavior composite is based on student responses to eight questions regarding how frequently they: (1) break something on purpose, (2) punch or hit someone, (3) argue with their parents, (4) lie to their parents, (5) steal from a store, (6) give a teacher a hard time, (7) sell illegal drugs, and (8) get arrested or detained by police. Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level of negative behavior, while a value of 4 indicates a high level of negative behavior.

^g The empathy index is based on a student ratings of ability to: (1) work with others on a team or on a group project, (2) feel bad for other people when they are having a hard time, and (3) believe the best about other people. A value of 1 on the index indicates poor ability, while a value of 4 indicates excellent ability.

^h The controlling destiny index is based on a student ratings of ability to: (1) set goals and work to achieve them, (2) plan for things needed in the future, (3) work out conflicts or disagreements with others, (4) stick to beliefs even if friends disagree. A value of 1 on the index indicates poor ability, while a value of 4 indicates excellent ability.

ⁱ The tobacco, alcohol, and drug use composite is based on seven items: the extent to which students (1) smoke cigarettes, (2) use smokeless tobacco, (3) have at least one drink of alcohol, (4) have five or more drinks of alcohol in a row, (5) smoke marijuana, (6) use inhalants, and (7) use any other illegal drug. Values on these items range from 1 to 4; a value of 1 on the composite indicates no substance abuse, while a value of 4 indicates frequent substance abuse.

^j The safety index is based on how often the student feels safe: (1) walking in her neighborhood, (2) being at home alone, (3) on the ground outside school, (4) going to the bathroom at school, and (5) in the hallways at school. A value of 1 indicates feeling less safe and a value of 4 indicates feeling more safe.

^k Sample sizes may differ due to missing values.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

The study estimated two types of outcome differences. First, it estimated “intent-to-treat” differences by comparing average outcomes at the time of the second follow-up for the full treatment and comparison groups. The intent-to-treat estimates provide a general sense of the outcomes experienced by the treatment and comparison groups after two school years. These estimates are useful because they are based on the treatment and comparison groups as they were constructed at the start of the study.³¹ Looking at the full group of students (even students who are no longer attending in the second year) provides an opportunity to observe outcome differences in the second year that might occur due to attending the program in the first year and possibly in the second year. All results presented in this chapter are based on intent-to-treat estimators. Models were estimated using SUDAAN[®] to account for the complex sampling design.

Because some students graduated to high school or transferred to other schools and did not have access to a 21st Century center in the study’s second year, the study team also estimated outcome differences for students who had access to centers during the second year.³² These students could have attended centers had they wanted to, which is conceptually similar to students who were the basis for the first-year report except for the additional year.³³ The method used to conduct this analysis does not rely on the original treatment and comparison groups identified through propensity score matching. Instead, it relies on a sample that has been

³¹“Treatment” group refers to students who were participating in programs in fall 2000. Whether these students received program services—a treatment—in the second year depended on whether they attended.

³²For both types of estimates, when baseline outcomes were available, the study estimated outcome differences by comparing the change in outcomes from baseline to second followup for the treatment and comparison groups (the “differences of differences” method).

³³We investigated an “instrumental-variables” approach for estimating outcome differences for participants, which requires a variable that predicts participation in the program and that is not correlated with outcomes. However, we were unable to identify a variable that met both conditions.

rebalanced to increase the comparability of the treatment group (students from the original treatment group who had access to programs in the second year) and its comparison group. The analysis and its results are described in Appendix C. The study faced a similar situation when examining the relationship between attendance and outcomes. The analysis is based on regression models—not on the comparison design used to estimate outcome differences—and the results are presented in Appendix C.

The findings presented in Appendix C are generally consistent with the findings presented below. Estimating outcome differences for students with program access and examining the relationship between attendance and outcomes does not substantially alter the main findings for middle school students.

1. Some Differences in Supervision, Location, and After-School Activities

Several differences in location and supervision were evident between treatment and comparison students (Table III.3). Treatment students were more likely than comparison students to spend at least three days each week at school or another place for activities (28 percent versus 23 percent, effect size of 0.10). Most activities after school did not differ between the two groups (Table III.3), although treatment students were more likely than comparison students to have participated in lessons and clubs (effect sizes of 0.08 and 0.10, respectively).

Treatment students were less likely than comparison students to be with siblings after school (18 percent versus 21 percent, effect size of 0.09).³⁴ Levels of self-care were similar among treatment and comparison students. Multiple definitions of self-care were examined, but the

³⁴The reduction in “being with siblings” does not necessarily imply a reduction in being *only* with siblings, because students also can be with parents and other adults when they are with siblings. About 66 percent of middle school students who reported being with a sibling after school also reported being with a parent or other adult. We examined sibling supervision using the hierarchical definition from the first report (explained in more detail in chapter II) and found no differences.

Table III.3

Outcome Differences in Maternal Employment and Students' Location, Supervision, and Activities After School, Middle School Centers, Year 2

Outcome	Center Participants	Comparison Group	Difference
Percentage of Students with the Following Individuals at Least Three Days After School in a Typical Week:			
Self-care ^a	19.0	19.8	-0.8
Parent	50.9	53.0	-2.1
Nonparent adult	33.9	28.6	5.3
Sibling	17.5	21.2	-3.7**
Mixed (Not in any one category for at least three days)	4.0	5.4	-1.4
Percentage of Students in the Following Locations After School at Least Three Days in a Typical Week:			
Own home	69.2	71.5	-2.3
Someone else's home	12.6	11.8	0.8
School or other place for activities	27.5	23.2	4.4**
Somewhere to "hang out"	12.9	10.5	2.4
Mixed location (Not in one location for at least three days)	8.2	7.8	0.4
Employment of Mother (Parent-Reported):			
Full-time	59.9	62.6	-2.7
Part-time	15.7	13.4	2.2
Looking for Work	8.7	9.1	-0.4
Not in the labor force	15.7	14.9	0.9
Mean Number of Days Stayed After School for Activities in Typical Week	1.0	0.8	0.2**
Percentage of Students Who Participated in Following Activities After School:			
Homework	84.6	86.7	-2.2
Tutoring	18.1	15.1	3.0
Non-homework reading, writing, or science activities	43.9	41.9	2.0
School activities (band, drama, etc.)	32.1	29.3	2.7
Lessons (Music, art, dance, etc.)	23.8	20.7	3.2**
Organized sports	41.5	40.1	1.5
Clubs (Boy and Girl Scouts, Boys and Girls Club, etc.)	15.7	12.2	3.5**
Activities at church, temple, mosque	30.5	29.6	1.0
Watched TV or videos	89.1	87.7	1.5
Surfed the Internet or did other things on a computer	64.9	64.8	0.2
"Hung out" with friends	82.1	78.1	4.1***
Volunteered or did community service	17.8	15.4	2.4
Worked at a job	20.5	19.0	1.6
Did chores around the house	77.8	79.0	-1.3
Took care of a brother or sister	50.3	49.7	0.7
Mean Time Students Reported Watching Television in the Past Day (Hours)	2.0	2.0	0.02
Mean Time Students Reported Reading for Fun in the Past Day (Hours)	0.3	0.3	0.02
Sample Size^b	1,605	2,203	

SOURCE: Student Survey, Parent Survey.

NOTE: The percentages and mean values of outcomes for participants and comparison group members have been regression-adjusted for baseline differences between the groups. The control variables in the regressions include student characteristics such as indicators of students' demographic characteristics, students' baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated outcome differences shown in the table do not always equal the difference between center participants and the comparison group. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Appendix A describes how weights were constructed, and Appendix B describes methods used to estimate outcome differences. Percentages may not sum to 100 because of rounding.

^aStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least three days in a typical week.

Table III.3 (continued)

^bSample sizes differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

definitions did not change the findings. Appendix C shows estimates for alternative definitions of self-care.

2. Few Differences in Academic Outcomes

Most academic outcomes did not differ significantly between the treatment and comparison groups, with a few exceptions. Treatment students had better school attendance than comparison students, being absent 9 days on average versus 10 days for the comparison group (effect size of 0.09; see Table III.4). Subject grades differed for one subject, with treatment students having an average social studies grade of 82 and comparison students averaging an 80 (effect size is 0.14). There were no indications from site visits to suggest why social studies grades would differ between the two groups. Grades in math, science, and English—and student and teacher reports of achievement—did not differ (Table III.5).

According to teachers, student effort in class did not differ between treatment and comparison groups. Treatment students were less likely than comparison students to report paying attention to teachers in class (83 percent of treatment students reported paying attention relative to 87 percent of comparison students, an effect size of 0.10). Measures of school discipline problems showed no significant differences between treatment and comparison groups (Table III.4).

Homework completion and time spent doing homework did not differ significantly between the treatment and comparison groups (Table III.4). Table III.6 indicates potential reasons for the lack of differences. First, treatment students were no more likely than comparison students to receive help with homework (Table III.6). When parents and other adults are combined, about 80 percent of the treatment and comparison groups reported that they were asked “often” or “always” whether they had completed their homework. About 52 percent reported that their

Table III.4

Outcome Differences in Homework Completion and on Behavior and Level of Effort in the Classroom,
Middle School Centers, Year 2

Outcome	Center Participants	Comparison Group	Difference
Percentage of Students Who Reported That They “Often” or “Always” Complete the Homework Teachers Assign	81.3	83.0	-1.7
Percentage of Students Whose Teachers Reported That They “Often” Complete Their Homework	49.8	50.5	-0.8
Mean Amount of Time Students Spent Doing Homework the Last Time They Had Homework (Hours)	0.9	1.0	-0.1
Percentage of Students Whose Teachers Reported the Following:			
“Agree” or “strongly agree” that student completes assignments to the teacher’s satisfaction	53.4	55.2	-1.8
Student “usually tries hard” in English class	49.3	48.4	1.0
Student “often” performs at or above ability level	41.5	43.8	-2.3
Teacher-Reported Level of Effort Composite ^a (Mean)	3.5	3.5	0.0
Percentage of Students Who Reported that They Pay Attention to their Teachers in School	83.4	87.1	-3.7**
Percentage of Students Whose Parents “Agree” or “Strongly Agree” That Their Child Works Hard at School	78.5	76.0	2.5
Student-Based Discipline Problem Composite ^b (Mean)	1.4	1.4	0.0
Teacher-Based Discipline Problem Composite ^c (Mean)	1.4	1.4	0.0
Percentage of Students Who Were Suspended During 2001-2002 School Year	21.9	21.7	0.2
Mean Number of Days Student Was:			
Absent	9.0	10.0	-1.0**
Late	6.2	5.4	0.8
Sample Size^d	1,633	2,198	

SOURCE: Student Survey, Teacher Survey, Parent Survey, School Records.

NOTE: The percentages and mean values of outcomes for participants and comparison-group members have been regression-adjusted for baseline differences between the groups. The control variables in the regressions include student characteristics such as indicators of students’ demographic characteristics, students’ baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated outcome differences shown in the table do not always equal the difference between center participants and the comparison group. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Appendix A describes how weights were constructed and Appendix B describes methods used to estimate outcome differences.

^aThe level of effort composite is based on five items reported by teachers: whether the student (1) usually tried hard, (2) often performs at or above his or her ability level, (3) is attentive in class, (4) participates in class, and (5) volunteers in class. The composite is equal to the mean of the five variables. Values on these items range from 1 to 5; a value of 1 on the composite indicates a low level of effort, and a value of 5 indicates a high level of effort.

^bThe student-based discipline problem composite is based on four items: the extent to which students reported (1) skipping school or class, (2) getting sent to the office for doing something wrong, (3) getting detention, and (4) having their parents called to

Table III.4 (continued)

school about a problem they were having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems.

^cThe teacher-based discipline problem composite is based on four items: the extent to which the teacher reported that the student was (1) skipping school or class, (2) getting sent to the office for doing something wrong, (3) getting detention, and (4) having his or her parents called to school about a problem they were having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems.

^dSample sizes may differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table III.5

Outcome Differences in Teacher-Reported Achievement and Grades,
Middle School Centers, Year 2

Outcome	Center Participants	Comparison-Group Members	Difference
Percentage of Students Whose Teachers Reported That They Achieve at an “Above-Average” or “Very High” Level	31.3	33.8	-2.5
Percentage of Students Whose Teachers “Agree” or “Strongly Agree” That They Get Good Grades on Tests	50.8	51.8	-1.0
Teacher-Reported Achievement Composite (Mean) ^a	3.3	3.3	0.0
Mean Grade:			
Math	79.3	78.6	0.7
English	80.1	79.6	0.5
Science	79.6	79.0	0.6
Social studies/history	81.6	79.8	1.7***
Sample Size^b	1,533	2,126	

SOURCE: Teacher Survey, School Records.

NOTE: The percentages and mean values of outcomes for participants and comparison-group members have been regression-adjusted for baseline differences between the groups. The control variables in the regressions include student characteristics such as indicators of students’ demographic characteristics, students’ baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated outcome differences shown in the table do not always equal the difference between center participants and the comparison group. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Appendix A describes how weights were constructed and Appendix B describes methods used to estimate outcome differences.

^aThe teacher-reported achievement composite is based on teacher responses to five questions: (1) At what level is this student performing in reading? (2) Does this student get good grades on tests? (3) Does this student complete assignments to my satisfaction? (4) Does this student have good communication skills? (5) Is this student a proficient reader? Values on these items range from 1 to 5; a value of 1 on the composite indicates low achievement, and a value of 5 indicates high achievement.

^bSample sizes may differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table III.6

Outcome Differences in Types of Homework Assistance, Middle School Centers, Year 2

Outcome	Center Participants	Comparison Group	Difference
Percentage of Students Who Reported That Their Parent “Often” or “Always”:			
Asks if homework is complete	76.1	76.1	0.0
Looks at homework to see if it is complete	42.5	45.1	-2.7
Looks at homework to see if it is correct	38.5	41.8	-3.3
Explains homework in a way that is easy to understand	45.3	49.4	-4.1
Percentage of Students Who Reported That an Adult Who Is Not Their Parent “Often” or “Always”:			
Asks if homework is complete	38.8	35.3	3.5
Looks at homework to see if it is complete	29.1	28.4	0.8
Looks at homework to see if it is correct	29.4	25.8	3.6
Explains homework in a way that is easy to understand	35.3	33.7	1.6
Percentage of Students Who Reported That Their Parent or an Adult Who Is Not Their Parent “Often” or “Always”:			
Asks if homework is complete	80.5	80.4	0.1
Looks at homework to see if it is complete	52.0	52.6	-0.6
Looks at homework to see if it is correct	49.2	49.1	0.1
Explains homework in a way that is easy to understand	56.6	58.5	-1.9
Percentage of Students Who Had the Following Individual Ask the Child To Correct Parts of Homework:			
Parent	75.0	76.3	-1.3
An adult who is not their parent	57.1	54.6	2.5
A parent or an adult who is not their parent	83.3	83.1	0.1
Sample Size^a	1,633	2,198	

SOURCE: Student Survey.

NOTE: The percentages and mean values of outcomes for participants and comparison-group members have been regression-adjusted for baseline differences between the groups. The control variables in the regressions include student characteristics such as indicators of students’ demographic characteristics, students’ baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated outcome differences shown in the table do not always equal the difference between center participants and the comparison group. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Appendix A describes how weights were constructed and Appendix B describes methods used to estimate outcome differences.

^aSample sizes may differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

homework was looked at to see if it was complete; about 49 percent reported that their homework was checked to see if it was correct. With comparison students as likely as treatment students to have their homework checked, and more likely to have a parent check their homework, centers evidently had the effect of substituting other adults for parents in the role of helping with homework. Relatedly, Table III.6 also shows that parents reported having a relatively high level of involvement with their children's homework, regardless of participation in the program.

3. Higher Educational Aspirations for Treatment Students

Eighty-two percent of treatment students and 80 percent of comparison students said they expected to graduate from college (effect size of 0.06, Table III.7).

4. No Differences in Social and Emotional Outcomes

Social, emotional, and other developmental outcomes did not differ between the groups (Table III.7). For example, treatment students were no more likely than comparison students to rate themselves highly on working out conflicts with others, to report feeling more socially engaged, or to report better interactions with and empathy for others.

5. No Differences in Parent Outcomes

Parental involvement was about the same for both groups. For example, roughly 19 percent of parents from both groups attended at least three open houses at the school during the past year, and roughly 25 percent attended at least three parent-teacher organization meetings. In the first report, all four parent-involvement measures were statistically significant; in this report, none of the impacts was statistically significant.

Table III.7

Outcome Differences in Social Engagement, Educational Expectations, and Parental Involvement,
Middle School Centers, Year 2

Outcome	Center Participants	Comparison Group Members	Difference
Social Engagement Composite ^a (Mean)	3.54	3.56	-0.02
Peer Interaction/Empathy Composite ^b (Mean)	3.01	3.03	-0.02
Percentage of Students Who Rated Themselves as “Good” or “Excellent” at Working Out Conflicts with Others	57.4	60.7	-3.3
Percentage of Students Who Rated Themselves as “Good” or “Excellent” on Using a Computer to Look Up Information	36.9	36.6	0.3
Percentage of Students Who Think They Will:			
Graduate from college	82.1	79.6	2.5**
Graduate from high school but not college	16.5	18.5	-2.0
Attend high school but not graduate	1.4	1.9	-0.6
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:			
Attended an open house at the school	19.5	18.8	0.7
Attended parent-teacher organization meetings	26.8	25.4	1.4
Attended an after-school event	38.8	37.0	1.8
Volunteered to help out at school	16.1	14.2	1.9
Sample Size^c	1,601	2,208	

SOURCE: Student Survey, Parent Survey.

NOTE: The percentages and mean values of outcomes for participants and comparison-group members have been regression-adjusted for baseline differences between the groups. The control variables in the regressions include student characteristics such as indicators of students’ demographic characteristics, students’ baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated outcome differences shown in the table do not always equal the difference between center participants and the comparison group. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Appendix A describes how weights were constructed and Appendix B describes methods used to estimate outcome differences. Percentages may not sum to 100 because of rounding.

^aThe social engagement composite is based on five items: the extent to which students report that they (1) have friends to “hang out with,” (2) are never lonely, (3) get along with others their age, (4) find it easy to make new friends, and (5) never feel left out of things. The composite is equal to the mean of the five variables. Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level of social engagement, and a value of 4 indicates a high level of engagement.

^bThe peer interaction/empathy composite is based on three items: students’ rating of their ability to (1) work with others in a team or group, (2) feel bad for other people who are having difficulties, and (3) believe the best about other people. Values on these items range from 1 to 4; a value of 1 on the composite indicates poor peer interactions, while a value of 4 indicates excellent peer interactions.

^cSample sizes may differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

6. No Differences in Feelings of Safety

There were no differences between the treatment and comparison groups in their feelings of safety during the after-school hours. About two-thirds of students reported that they felt very safe after school (Table III.8).

7. Mixed Evidence on Negative Behaviors

There was mixed evidence on negative behavior outcomes. For example, more treatment students than comparison students reported breaking things on purpose (10 percent compared to 8 percent, effect size of 0.08). Treatment students had higher values of a negative behavior composite, which measured the frequency of eight behaviors including breaking things on purpose, punching someone, selling drugs, and being detained or arrested by the police (effect size of 0.08).³⁵ On other outcomes—such as stealing, selling illegal drugs, and getting arrested—there were no significant differences. Also, there were no significant differences between the treatment and comparison groups on four measures of victimization, including being threatened or hurt with a weapon and having property damaged.³⁶

There was mixed evidence on drug use (Table III.8). Treatment students were more likely than comparison students to report using cocaine, ecstasy, or LSD. The percentage of students reporting this type of drug use was small—less than 1 percent of students—and the effect size of 0.08 was small as well. There were no differences between the treatment and comparison groups on the extent to which students smoked cigarettes, drank alcohol, or smoked marijuana.

³⁵Other studies also have found evidence of increased negative behavior among program participants. For example, Mahoney et al. (2001) found evidence of increased criminal offenses among Swedish students attending youth recreation centers, and Weisman et al. (2002) found increased delinquent behavior among after-school program participants.

³⁶The study asked generally about whether students engaged in negative behaviors and not whether the behaviors occurred in the after-school program or elsewhere.

Table III.8

Outcome Differences in Student Safety, Negative Behavior, and Victimization, Middle School Centers, Year 2

Outcome	Center Participants	Comparison Group	Difference
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6:00 P.M.:			
Very Safe	64.6	66.9	-2.4
Somewhat safe	32.7	30.6	2.1
Not at all safe	2.7	2.5	0.3
Percentage of Students Who Reported That They Do the Following "Some" or "A Lot":			
Break something on purpose	10.4	8.0	2.4**
Punch or hit someone	22.4	19.7	2.7
Steal from a store	4.9	4.0	0.9
Sell illegal drugs	1.4	1.8	-0.3
Get arrested or detained by police	3.3	3.1	0.2
Negative Behavior Composite ^a (Mean)	1.56	1.53	0.03**
Percentage of Students Who Reported the Following Happened to Them "Some" or "A Lot":			
Been offered, sold, or given an illegal drug	18.1	19.1	-1.0
Been "picked on" after school	27.7	24.7	3.0
Been threatened or hurt with a weapon	6.8	5.9	1.0
Been threatened by a gang or gang member	7.2	7.0	0.2
Had property damaged on purpose	13.5	11.1	2.4
Percentage of Students Who Report That They Did the Following "Some" or "A Lot":			
Smoke cigarettes	4.7	4.1	0.6
Have at least one alcoholic drink	9.8	9.0	0.8
Smoke marijuana	4.8	4.3	0.5
Took illegal drugs such as cocaine, ecstasy, or LSD	0.8	0.2	0.6***
Tobacco, Alcohol, and Drug Use Composite ^b (Mean)	1.14	1.12	0.02
Sample Size^c	1,609	2,209	

SOURCE: Student Survey.

NOTE: The percentages and mean values of outcomes for participants and comparison-group members have been regression-adjusted for baseline differences between the groups. The control variables in the regressions include student characteristics such as indicators of students' demographic characteristics, students' baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated outcome differences shown in the table do not always equal the difference between center participants and the comparison group. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Appendix A describes how weights were constructed and Appendix B describes methods used to estimate outcome differences. Percentages may not sum to 100 because of rounding.

^aThe tobacco, alcohol, and drug use composite is based on seven items: the extent to which students report that they (1) smoke cigarettes, (2) use smokeless tobacco, (3) have at least one drink of alcohol, (4) have five or more drinks of alcohol in a row, (5) smoke marijuana, (6) use inhalants, and (7) use any other illegal drug. Values on these items range from 1 to 4; a value of 1 on the composite indicates no substance abuse, while a value of 4 indicates frequent substance abuse.

^bSample sizes may differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

*** Significantly different from zero at the .01 significance level, two-tailed test.

Increases in negative behavior may seem counterintuitive for programs that provided enrichment and recreational opportunities and whose goals typically included positive youth development. However, some researchers have noted that grouping youth together, particularly high-risk youth in unstructured settings, can increase negative behavior (Dishion et al. 1996; Dishion et al. 1999; Sherman et al. 1998). In addition, site visitors observed aspects of center operations that may have contributed to negative behaviors or at least contributed to not reducing them. Site visitors observed instances where students were disobeying staff members and arguing with staff members or other students, bothering other students, or talking when they were supposed to be doing homework or another activity, and instances where staffers were not maintaining control of students.

D. Few Differences for Subgroups

Results also were estimated for six subgroups based on student or parent characteristics: (1) grade level, (2) “high” versus “low” baseline grades, (3) “high” versus “low” baseline discipline problems, (4) race/ethnicity, (5) gender, and (6) a two-parent versus single-parent household structure.³⁷ Some of these characteristics might be used by programs to target program services, whereas others may be helpful in understanding estimates. The full set of subgroup estimates is presented in Appendix D; estimates for key outcomes and subgroups are presented here.³⁸

³⁷Students are defined as having low baseline grades if they reported average grades at baseline of C, D, or F; students are defined as having high baseline grades if they reported average grades at baseline of A or B. Students are defined as having low (high) discipline problems if their discipline problem composite variable is below (above) the median level of the discipline composite for the study sample. Results for a subgroup defined by participation in the program in the spring prior to the start of the study were estimated to investigate whether previous participation in an after-school program was associated with impacts. Most of the impacts were insignificant, except that students who had attended the program in the previous spring (before the study began) were more likely to say they expected to graduate from college.

³⁸Subgroup estimates should be interpreted with caution; because many statistical tests were done for the subgroup analysis, some findings will appear positive by chance alone.

The findings indicate few subgroup outcome differences (Tables III.9a through III.10b). An initial look at the subgroup findings indicated that outcome differences for one subgroup—students with high grades at baseline—were significantly different from the outcome differences for students with lower baseline grades. However, the differences were not significant after controlling for whether students were in other subgroups.³⁹ This suggests that having high grades may not be the factor that is giving rise to the outcome differences, but that factors related to having high grades may be.

The first report noted that black students had significantly larger outcome differences on some outcomes than white students—classroom effort, lateness to school, and math grades. Black students did not experience a similar pattern of larger outcome differences than white students in the second year.

E. Comparing Estimates

It will be helpful to recap the key findings for the two reports, focusing on the intent-to-treat estimates, which are based on the full sample of treatment and comparison students, and on several key outcomes: supervision after school, grades, absences, classroom effort, safety, and negative behaviors and victimization.

The two reports provide evidence that treatment students were less likely than comparison students to be with parents after school and more likely to be with other adults (Table III.11). These effects were more pronounced in the first than the second year, most likely due to the lower program participation observed in the study's second year.

³⁹To assess how impacts on other subgroups might affect the high baseline grade subgroup impacts listed in the report, we first ran models that regressed outcomes on a variable that interacted treatment and high grades or low grades. We then ran models that included other treatment-subgroup interaction terms. This procedure enabled us to assess whether subgroup impacts were evident in the presence of additional treatment and subgroup interactions.

Table III.9A

Outcome Differences in Maternal Employment and Students' Location, Care, and Activities After School for White, Black, and Hispanic Subgroups, Middle School Centers, Year 2

Outcome	White (Non-Hispanic)			Black (Non-Hispanic)			Hispanic		
	Treatment Mean	Comparison Mean	Difference	Treatment Mean	Comparison Mean	Difference	Treatment Mean	Comparison Mean	Difference
Percentage of Students Who Reported Being in the Following Locations After School at Least 3 Days in Typical Week:									
Own home	65.1	72.1	-6.9***	75.5	74.9	0.6	68.5	67.3	1.3
Someone else's home	10.9	11.4	-0.6	12.2	11.1	1.0	10.5	12.2	-1.7
School or other place for activities	30.8	30.4	0.4	29.1	24.0	5.0	27.9	22.4	5.5
Somewhere to "hang out"	11.5	10.5	1.0	14.8	8.6	6.2***	14.2	13.2	1.0
Mixed (no one location for at least 3 days)	10.2	4.6	5.6***	5.5	6.0	-0.5	7.3	9.4	-2.1
Percentage of Students Who Reported Being with the Following Individuals After School at Least 3 Days in Typical Week:									
Self care ^a	17.5	14.6	3.0	17.3	18.0	-0.8	21.7	22.6	-0.9
Parent	50.5	57.0	-6.5**	47.3	52.5	-5.3	50.5	46.5	4.0
Non-parent adult	35.6	32.9	2.6	30.8	28.2	2.6	29.7	25.2	4.4
Sibling	18.8	22.8	-4.0	20.9	19.8	1.1	13.2	21.8	-8.6***
Mixed (no one category for at least 3 days)	4.9	4.7	0.2	5.1	6.9	-1.8	5.3	7.2	-1.9
Employment of Mother (parent-reported):									
Full-time	60.7	66.1	-5.5	64.8	62.4	2.4	53.6	56.5	-2.9
Part-time	19.8	18.0	1.8	12.4	9.2	3.2	14.4	12.7	1.6
Looking for Work	4.9	3.3	1.6	10.3	11.6	-1.3	11.0	12.7	-1.7
Not in the labor force	14.6	12.5	2.1	12.5	16.8	-4.3	21.0	18.1	2.9
Mean Number of Days School Records Indicate Student Was:									
Absent	11.9	13.0	-1.1**	13.1	14.7	-1.6	11.4	12.1	-0.7
Late	5.8	5.3	0.5	7.6	5.8	1.8	8.0	7.4	0.6
Mean Class Grade:									
Math	81.6	80.7	1.0	74.7	74.0	0.7	76.4	75.7	0.7
English	79.5	79.2	0.3	74.6	74.0	0.6	75.6	74.9	0.7
Science	79.8	79.5	0.3	74.2	73.5	0.7	74.7	74.0	0.6
Social Studies	80.0	78.6	1.5	75.6	72.4	3.2***	74.5	73.9	0.7
Number of Observations									
Student-reported outcomes		1,334			909			1,020	
School records outcomes (attendance)		1,324			899			1,016	
School records outcomes (grades)		1,300			863			971	
Parent-reported outcomes		1,216			770			933	

SOURCE: Parent Survey, Student Survey, School Records.

NOTE: Subgroup estimates reported in bold indicate that the estimated outcome difference for one subgroup differed significantly from the estimated outcome difference for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Percentages may not sum to 100 because of rounding.

^aStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least three days in a typical week.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table III.9B

Outcome Differences in Maternal Employment and Students' Location, Care, and Activities After School
for Low and High Baseline Grade Subgroups, Middle School Centers, Year 2

Outcome	Low Baseline Grades ^a			High Baseline Grades ^a		
	Treatment Mean	Comparison Mean	Difference	Treatment Mean	Comparison Mean	Difference
Percentage of Students Who Reported Being in Following Locations After School at Least 3 Days in Typical Week:						
Own home	68.4	71.3	-3.0	69.5	71.5	-2.0
Someone else's home	12.5	13.6	-1.1	11.0	9.3	1.7
School or other place for activities	27.4	20.7	6.6***	29.8	27.0	2.7
Somewhere to "hang out"	16.5	13.0	3.5	10.7	8.8	1.9
Mixed (no one location for at least 3 days)	6.5	6.3	0.2	8.2	7.5	0.6
Percentage of Students Who Reported Being with the Following Individuals After School at Least 3 Days in Typical Week:						
Self ^b	20.3	23.8	-3.5	18.3	17.1	1.2
Parent	48.6	46.1	2.5	50.6	54.4	-3.8
Non-parent adult	31.1	25.9	5.3	34.6	29.9	4.7
Sibling	16.8	18.5	-1.7	17.4	23.1	-5.8***
Mixed (no one category for at least 3 days)	5.2	6.3	-1.1	4.1	5.9	-1.8
Employment of Mother (parent-reported):						
Full-time	59.4	63.8	-4.4	59.7	62.6	-2.9
Part-time	17.1	11.0	6.1**	15.1	14.2	0.9
Looking for Work	9.9	9.3	0.6	8.3	8.8	-0.5
Not in the labor force	13.6	15.9	-2.3	16.9	14.3	2.6
Mean Number of Days School Records Indicate Student Was:						
Absent	14.0	14.9	-0.93	11.2	12.2	-1.04**
Late	9.2	7.2	2.06**	6.0	5.8	0.21
Mean Class Grade:						
Math	71.6	70.7	0.94	79.6	79.2	0.46
English	71.4	69.8	1.56**	78.2	78.3	-0.13
Science	71.0	70.1	0.94	79.3	79.0	0.28
Social Studies	71.7	68.6	3.20***	78.7	77.7	0.92
Number of Observations						
Student-reported outcomes		1,130			2,593	
School records outcomes (attendance)		1,117			2,587	
School records outcomes (grades)		1,067			2,507	
Parent-reported outcomes		994			2,328	

SOURCE: Parent Survey, Student Survey, School Records.

NOTE: Subgroup estimates reported in bold indicate that the estimated outcome difference for one subgroup differed significantly from the estimated outcome difference for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Percentages may not sum to 100 because of rounding.

^aStudents are defined as having low baseline grades if they reported average grades of C, D, or F; students are defined as having high baseline grades if they reported average grades of A or B.

^bStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least three days in a typical week.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table III.10A

Outcome Differences in Other Student and Parent Outcomes for White, Black, and Hispanic Subgroups, Middle School Centers, Year 2

Outcome	White (Non-Hispanic)			Black (Non-Hispanic)			Hispanic		
	Treatment Mean	Comparison Mean	Difference	Treatment Mean	Comparison Mean	Difference	Treatment Mean	Comparison Mean	Difference
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6:00 P.M.:									
Very Safe	71.9	77.2	-5.3**	60.5	58.7	1.8	59.5	61.5	-1.9
Somewhat safe	25.4	21.5	3.9	35.3	38.0	-2.7	38.1	35.4	2.7
Not at all safe	2.6	1.2	1.4**	4.1	3.3	0.8	2.4	3.1	-0.7
Students' Educational Expectations (percentages):									
Graduate from college	82.8	84.8	-2.0	82.9	81.6	1.3	76.1	72.6	3.5
Graduate from high school	15.9	14.3	1.6	16.0	16.6	-0.5	21.0	24.1	-3.0
Drop out of high school	1.3	0.9	0.4	1.1	1.9	-0.7	2.9	3.4	-0.5
Percentage of Students Who Reported the Following Happened to Them "Some" or "A lot":									
Been offered, sold, or given an illegal drug	17.7	16.4	1.3	13.6	16.5	-2.9	21.7	24.9	-3.2
Been picked on after school	32.0	28.6	3.4	27.2	21.8	5.4	23.0	21.1	1.8
Been threatened or hurt with a weapon	7.1	5.7	1.3	6.5	5.4	1.2	7.0	6.2	0.8
Been threatened by a gang member	6.1	5.6	0.5	7.7	6.8	0.9	7.8	9.0	-1.1
Had your property damaged on purpose	12.0	12.1	-0.1	18.8	10.1	8.7***	9.4	11.2	-1.8
Percentage of Students Who Reported That They Do the Following "Some" or "A Lot":									
Smoke cigarettes	6.5	5.2	1.4	3.5	2.0	1.4	3.9	4.9	-1.0
Smoke marijuana	4.3	4.0	0.3	4.4	2.9	1.5	5.5	5.7	-0.2
Drink alcohol	10.7	9.4	1.3	8.2	5.1	3.1	12.3	12.6	-0.2
Student-Reported Tobacco, Alcohol, Drug Use Composite (mean)	1.1	1.1	0.0	1.1	1.1	0.0**	1.2	1.2	0.0
Number of Observations									
Student-reported outcomes	1,341			910			1,019		

SOURCE: Student Survey.

NOTE: Subgroup estimates reported in bold indicate that the estimated outcome difference for one subgroup differed significantly from the estimated outcome difference for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Percentages may not sum to 100 because of rounding.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table III.10B

Outcome Differences in Other Student and Parent Outcomes for Low and High Baseline Grades Subgroups, Middle School Centers, Year 2

Outcome	Low Baseline Grades ^a			High Baseline Grades ^a		
	Treatment Mean	Comparison Mean	Difference	Treatment Mean	Comparison Mean	Difference
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6:00 P.M.:						
Very safe	62.2	63.6	-1.4	64.4	68.5	-4.1
Somewhat safe	34.9	34.0	1.0	32.7	29.2	3.5
Not safe at all	2.9	2.5	0.4	2.8	2.2	0.6
Students' Educational Expectations (Percentages):						
Graduate from college	73.3	69.1	4.2	86.7	85.2	1.5
Graduate from high school	24.2	28.4	-4.2	12.0	13.1	-1.1
Drop out of high school	2.5	2.5	0.0	1.3	1.7	-0.4
Percentage of Students Who Reported the Following Happened to Them "Some" or "A lot":						
Been offered, sold, or given an illegal drug	18.0	24.1	-6.2***	18.1	17.0	1.1
Been picked on after school	30.6	25.7	4.9	26.1	24.1	2.0
Been threatened or hurt with a weapon	10.5	7.6	2.9	5.3	4.9	0.3
Been threatened by a gang member	8.6	8.3	0.3	6.3	6.4	-0.1
Had your property damaged on purpose	16.0	13.5	2.4	11.8	10.1	1.6
Percentage of Students Who Reported That They Do the Following "Some" or "A Lot":						
Smoke cigarettes	6.9	5.5	1.4**	3.7	3.5	0.2
Smoke marijuana	5.3	7.4	-2.0	4.7	3.0	1.7***
Drink alcohol	10.0	12.7	-2.6	10.0	7.5	2.5
Student-Reported Tobacco, Alcohol, and Drug Use Composite (mean)	1.2	1.2	-0.02	1.1	1.1	0.03**
Number of Observations						
Student-reported outcomes	1,133			2,600		

SOURCE: Student Survey.

NOTE: Subgroup estimates reported in bold indicate that the estimated outcome difference for one subgroup differed significantly from the estimated outcome difference for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Percentages may not sum to 100 because of rounding.

^aStudents are defined as having low baseline grades if they reported average grades of C, D, or F; students are defined as having high baseline grades if they reported average grades of A or B.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table III.11

Effect Sizes for Selected Outcomes From First Report and Current Report

Outcome	Year One Full Sample Outcome Differences	Year Two Full Sample Outcome Differences
In Self-Care After School	0.00	-0.02
With Parent After School	-0.12***	-0.04
With Other Adult After School	0.24***	0.11
With Sibling After School	-0.11***	-0.09**
In Mixed Care After School	0.00	-0.06
Grade in Math	0.06	0.06
Grade in English	0.01	0.04
Grade in Science	0.01	0.05
Grade in Social Studies/History	0.03	0.14***
Number of School Absences	-0.11***	-0.09**
Teacher-Reported Effort in Class	0.10***	0.01
Feel Very Safe After School	-0.03	-0.05
Feel Somewhat Safe After School	0.03	0.04
Feel Unsafe After School	0.00	0.02
Negative Behavior Composite	0.09***	0.08**
Drug Use Composite	0.01	0.05
Been Picked on After School	0.04	0.07
Had Property Damaged	0.08**	0.07

**Significantly different from zero at the .05 level, two-tailed test.

***Significantly different from zero at the .01 level, two-tailed test.

There were no differences between the treatment and comparison groups in terms of self-care after school. Self-care was not significantly lower for the treatment group relative to the comparison group in either the first or second year. In both years, however, sibling care was lower among the treatment group relative to the comparison group.

The two reports provide consistent evidence that treatment students do not have higher grades than comparison students in most subjects. Grades in English and science were similar for the two groups. In the first report, math grades were higher (at the 10 percent significance level); in the second report, social studies grades were higher.

The two reports provide evidence that treatment students had fewer absences and greater classroom effort than comparison students. Both reports showed significantly fewer absences and the first-year estimate showed greater classroom effort (the classroom effort results were corroborated by additional estimates presented in Appendix C).

The two reports also provide consistent evidence that centers did not improve perceptions of safety. Neither of the safety estimates was significantly different.

The evidence for the two years shows mixed findings on negative behaviors. In the first and second reports, the negative-behavior composite variable was significant, as were some of the individual negative behaviors that were part of the composite. On other behavior outcomes, however, there were no differences between the treatment and comparison groups, suggesting that the program had a mixed effect overall on negative behaviors.

Evidence from the two years is inconsistent in terms of the program's effects on victimization. There was evidence in the first report of higher levels of victimization among the treatment group, although this was not found in the second report.

References

- Afterschool Alliance. "New Data: Federal Funding Shortfall Leaves 75 Percent of Afterschool Needs Unmet." [www.afterschoolalliance.org/press_archives/budget_chart_nat_2003_nr.doc]. Accessed March 13, 2003.
- Brooks, Pauline, Cynthia Mojica, and Robert Land. "Final Evaluation Report: Longitudinal Study of LA's BEST After School Education and Enrichment Program, 1992-94." Los Angeles: University of California, Los Angeles, Center for the Study of Education, spring 1995.
- Chamberlain, G. "Analysis of Covariance with Qualitative Data." *Review of Economic Studies*, vol. 47, 1980, pp. 225-238.
- DeAngelis, Karen, and Robert Rossi. "Schools Serving Family Needs: Extended-Day Programs in Public and Private Schools." National Center for Education Statistics Issue Brief (NCES 97-590). Washington, DC: National Center for Education Statistics, February 1997.
- Dishion, Thomas J., Kathleen M. Spracklen, David A. Andrews, and Gerald R. Patterson. "Deviancy Training in Male Adolescent Friendships." *Behavior Therapy*, vol. 27, 1996, pp. 373-390.
- Dishion, Thomas J., Joan McCord, and Francois Poulon. "When Interventions Harm: Peer Groups and Problem Behavior." *American Psychologist*, vol. 54, no. 9, 1999, pp. 755-764.
- Dynarski, Mark, Mary Moore, John Mullens, Phil Gleason, Susanne James-Burdumy, Linda Rosenberg, Carol Pistorino, Tim Silva, John Deke, Wendy Mansfield, Sheila Heaviside, and Daniel Levy. *When Schools Stay Open Late: The National Evaluation of the 21st Century Community Learning Centers Program, First-Year Findings*. Report submitted to the U.S. Department of Education. Princeton, NJ: Mathematica Policy Research, Inc, January 2003.
- Dynarski, Mark, Susanne James-Burdumy, Daniel Mayer, Mary Moore, John Mullens, Tim Silva, Carol Pistorino, and Douglas Hermond. "A Broader View: The National Evaluation of the 21st Century Community Learning Centers Program: Design Report." Report submitted to the U.S. Department of Education. Princeton, NJ: Mathematica Policy Research, Inc, March 2001.
- Fashola, Olatokunbo. "Review of Extended Day and After School Programs and their Effectiveness." [www.csos.jhu.edu/crespar/techReports/Report24.pdf]. Accessed May 17, 1999.
- Fight Crime: Invest in Kids. "America's After-School Choice: Juvenile Crime or Safe Learning Time." [www.fightcrime.org/reports/asTwoPager.pdf]. Accessed December 1, 2003.

- Forum for Youth Investment. "Out-of-School Time Policy Commentary #4: After-School For All? Exploring Access and Equity in After-School Programs." Washington, D.C.: The Forum for Youth Investment, July 2003.
- Grossman, Jean Baldwin, Marilyn Price, Veronica Fellerath, Linda Jucovy, Lauren Kotloff, Rebecca Raley, and Karen Walker. "Multiple Choices After School: Findings from the Extended-Service Schools Initiative." Philadelphia, PA: Public/Private Ventures, June 2002.
- Hamilton, Laura, and Stephen Klein. "Achievement Test Score Gains Among Participants in the Foundations School Age Enrichment Program." Report prepared for Foundations, Inc. Santa Monica, CA: RAND Corporation, September 1998.
- Hollister, Rob. "The Growth in After-School Programs and Their Impact." Report prepared for The Brookings Institution. Washington, DC: Brookings Institution, February 2003.
- Kleiner, Brian, Mary Jo Nolin, and Chris Chapman. "Before- and After-School Care, Programs, and Activities of Children in Kindergarten through Eighth Grade: 2001." National Center for Education Statistics Statistical Analysis Report (NCES 2004-008). Washington, DC: National Center for Education Statistics, April 2004.
- Lauer, Patricia, Motoko Akiba, Stephanie Wilkerson, Helen Aphthorp, David Snow, and Mya Martin-Glenn. "The Effectiveness of Out-of-School-Time Strategies in Assisting Low-Achieving Students in Reading and Mathematics: A Research Synthesis." Report prepared for the Institute of Education Sciences, Department of Education. Aurora, CO: Mid-continent Research for Education and Learning, October 2003.
- Mahoney, Joseph, Hakan Stattin, and David Magnusson. "Youth Recreation Centre Participation and Criminal Offending: A 20-Year Longitudinal Study of Swedish Boys." *International Journal of Behavioral Development*, vol. 25, no. 6, 2001, pp. 509-520.
- National Research Council and Institute of Medicine. *Community Programs to Promote Youth Development*. Committee on Community-Level Programs for Youth. Jacquelynne Eccles and Jennifer Appleton Gootman, eds. Board on Children, Youth, and Families, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press, 2002.
- Olsen, Darcy. "12-Hour School Days? Why Government Should Leave After-School Arrangements to Parents." *Policy Analysis*, no. 372, 2000, pp. 1-19.
- Reisner, Elizabeth, Richard White, Jennifer Birmingham, and Megan Welsh. "Building Quality and Supporting Expansion of After-School Projects, Evaluation Results from the TASC After-School Program's Second Year." Report submitted to The After-School Corporation. Washington, DC: Policy Studies Associates, Inc., February 2001.
- Rosenbaum, Paul, and Donald Rubin. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika*, vol. 70, 1983, pp. 41-45.

- Roth, Jodie, Jeanne Brooks-Gunn, Lawrence Murray, and William Foster. "Promoting Healthy Adolescents: Synthesis of Youth Development Program Evaluations." *Journal of Research on Adolescence*, vol. 8, no. 4, 1998, pp. 423-459.
- Schwarzenegger, Arnold. "Labor HHS Subcommittee Hearing Statement." Testimony before the Subcommittee on Labor, Health and Human Services, Education. U.S. Senate, Committee on Appropriations, Washington, DC, May 13, 2003.
- Seppanen, Patricia S., John M. Love, Dianne Kaplan deVries, Lawrence Bernstein, Michelle Seligson, Fern Marx, and Ellen Eliason Kisker. "National Study of Before- and After-School Programs." Portsmouth, NH: RMC Research Corporation, 1993.
- Sherman, L.; D. Gottfredson; D. MacKenzie; J. Eck; P. Reuter; and S. Bushway. "Preventing Crime: What Works, What Doesn't, What's Promising." A Report to the U.S. Congress, prepared for the National Institute of Justice. 1998.
- Tierney, Joseph, Jean Baldwin Grossman, and Nancy Resch. "Making a Difference: An Impact Study of Big Brothers/Big Sisters." Philadelphia, PA: Public/Private Ventures, November 1995.
- Vandivere, Sharon, Kathryn Tout, Jeffrey Capizzano, and Martha Zaslow. "Left Unsupervised: A Look at the Most Vulnerable Children." Child Trends Research Brief 2003-05. Washington, DC: Child Trends, 2003.
- Walker, Karen, and Amy Arbretton. "After-School Pursuits: An Examination of Outcomes in the San Francisco Beacon Initiative." Philadelphia, PA: Public/Private Ventures, forthcoming 2004.
- Weisman, Stephanie, David Soule, Denise Gottfredson, Shaoli Lu, Melissa Kellstrom, Shannon Womer, and Sean Bryner. "After-School Programs, Antisocial Behavior, and Positive Youth Development: Program Implementation and Youth Outcomes." Draft Report. College Park: University of Maryland, June 2002.
- Welsh, Megan, Christina Russell, Imeh Williams, Elizabeth Reisner, and Richard White. "Promoting Learning and School Attendance Through After-School Programming: Student-Level Changes in Educational Performance Across TASC's First Three Years." Draft report. Washington, DC: Policy Studies Associates, Inc., September 2002.
- Whitebook, M., C. Howes, and D. Phillips. *Worthy Work, Unlivable Wages: The National Child Care Staffing Study, 1988-1997*. Washington, DC: Center for the Child Care Work Force.

Appendix A

Response Rates and Data Quality

This appendix describes the results of the baseline and first follow-up wave of data collection conducted in the 12 elementary-school sites, and the second follow-up wave of data collection for the middle-school sites in the 2001-2002 academic year. The baseline and first follow-up experience for middle-school sites and for the first cohort of elementary-school sites were presented in our first report. A future report (the third and final report of the evaluation) will describe the second follow-up wave of data collection in the elementary-school sites.

The first follow-up for elementary-school sites was administered in spring 2001 for cohort 1 sites and in spring 2002 for cohort 2 sites. The second follow-up for middle-school sites was administered in spring 2002. As part of the enhanced study supported by the grant from the C. S. Mott Foundation, students participating in the 21st Century program at six middle-school sites completed an additional module of questions on their after-school activities.

The study collected data from a variety of respondents at 46 sites (34 middle-school sites and seven elementary-school sites in cohort 1 and five elementary-school sites in cohort 2). At the elementary-school sites, we conducted baseline surveys with students and parents and administered standardized reading tests to the students in fall 2000 for cohort 1 sites and fall 2001 for cohort 2 sites. At the middle-school sites, we conducted baseline surveys with students in fall 2000.

At all sites, we administered surveys to students, parents, teachers, school principals, and after-school program staff members (including directors). In each follow-up wave, we also collected students' school records and program-attendance records, and for elementary-school students we administered reading tests or collected past reading test scores (Table A.1).

Table A.1

Data Sources by Data-Collection Wave

Data Source	Data-Collection Wave		
	Baseline	1st Follow-up	2nd Follow-up
Elementary-School Student Questionnaire	T	T	T
Elementary-School Student Test	T	T	T
Elementary-School Parent Questionnaire	T	T	T
Middle-School Student Questionnaire	T	T	T
Middle-School Parent Questionnaire		T	T
Teacher Questionnaire		T	T
Principal Questionnaire		T	T
School Record ^a		T	T
After-School Program Attendance Record		T	T
After-School Program Project Director Questionnaire		T	T
After-School Program Center Coordinator Questionnaire		T	T
After-School Program Staff Member Questionnaire		T	T

^aBaseline records data were collected at the time of the first follow-up records collection.

A. Data Collection Procedures for Elementary-School Sites

1. Baseline

Baseline data collection consisted of an elementary-school student survey, a reading test, and a parent survey. The data were collected in the 2000-2001 school year for the first cohort of elementary-school sites and in the 2001-2002 school year for the second cohort. Questionnaires were given to all 3rd- to 6th-grade elementary-school students whose parents signed a consent form for their children to participate in the study. Questionnaires were generally self-administered during the school day (in some instances, teachers read the questions aloud to their class). We surveyed 90 percent of the 1,233 3rd- to 6th-grade elementary school students at baseline (Table A.2). Response rates ranged from 81 to 100 percent (Table A.3). Students at two sites were all in kindergarten through 2nd grade and were not surveyed.

Table A.2

Sample Sizes and Response Rates for the Baseline and First Followup
Elementary-School Sites

Instrument	Sample Size					Response Rate						
	Total		Treatment		Control		Total		Treatment		Control	
	N		N	%	N	%	N	%	N	%	N	%
Baseline												
Student Survey ^a	1,233		688	56	545	44	1,110	90	625	91	485	89
Student Test	2,308		1,258	55	1,050	45	1,568	68	847	67	721	69
Parent Survey	2,308		1,258	55	1,050	45	2,126	92	1,161	92	965	92
First Follow-Up												
Student Survey ^a	1,233		688	56	545	44	1,106	90	618	90	488	90
Student Test	2,308		1,258	55	1,050	45	1,902	82	1,044	83	858	82
Parent Survey	2,308		1,258	55	1,050	45	1,732	75	961	76	771	73
Teacher Survey ^b	2,308		1,258	55	1,050	45	1,831	79	995	79	836	80
School Record	2,308		1,258	55	1,050	45	2,016	87	1,110	88	906	86

^aSample includes only grades 3 to 6.

^bSample size and response rates are based on number of students; 88 percent of the 759 teachers in the sample completed surveys.

Table A.3

Distribution of Response Rates For Elementary-School Sites

Instrument	Total	Number of Sites					
		Percentage					
		90 to 100	80 to 89	70 to 79	60 to 69	50 to 59	Less than 50
Baseline							
Student Survey ^a	10	5	5	0	0	0	0
Student Test	12	2	1	2	1	3	3
Parent Survey	12	6	5	1	0	0	0
Follow-Up							
Student Survey ^a	10	4	2	4	0	0	0
Student Test	12	4	3	3	1	1	0
Parent Survey	12	0	2	4	5	1	0
Teacher Survey ^b	12	2	2	4	2	1	1
School Record	12	7	3	1	1	0	0

^aSurveys were administered only to 3rd- to 6th-grade students; one elementary-school site in each cohort had no sample in those grades at baseline and follow-up 1.

^bResponse rates are based on number of students, not teachers.

We obtained reading test scores for the Stanford Achievement Test 9 (SAT-9) for elementary school students in one of two ways: (1) We collected scores from sites that administered tests, or (2) field staff members administered the test at sites that did not use the test on their own. Field staff members administered the tests to most students during the school day, and did make-ups

with a few students in their homes. We obtained test scores for 68 percent of students at baseline (Table A.2). Most of the students who were not tested had transferred outside their district. Some students did not answer enough test questions for their test to be scored and some were not tested because of language barriers. Response rates across sites ranged from 43 to 98 percent, excluding one site that provided SAT-9 test scores for students in grades 2 to 5 but did not allow kindergartners and 1st graders to be tested.

We also asked elementary school parents to complete a baseline questionnaire, which 92 percent did (Table A.2). Response rates across sites ranged from 78 to 100 percent (Table A.3). Slightly more than two-thirds (69 percent) returned questionnaires by mail, and the rest (32 percent) completed them by telephone.

2. First Follow-up

a. Student Survey and Test Data

About six weeks before the end of the school year, field staff members administered follow-up questionnaires. Ninety percent of students in grades 3 to 6 completed the follow-up questionnaire (Table A.2). Response rates across sites ranged from 72 to 97 percent (Table A.3). About 9 of every 10 students who completed the questionnaire did so in school. The others (mostly students who had transferred to other schools) completed the questionnaire by telephone or mail.

We administered the reading component of the SAT-9 in school to students in kindergarten through 6th grade who had not taken a district-administered version of the SAT-9 that spring. We obtained test scores for 82 percent of students (Table A.2). Response rates by site ranged from 57 to 93 percent (Table A.3). Most of the students who were not tested had transferred outside their district, and a small percentage of students failed to answer enough questions for scoring or were not tested because of language barriers or impairment.

b. Data Collected from Parents, Teachers, and Records

Three-fourths of elementary-school parents completed the follow-up questionnaire (Table A.2). Slightly more than half (51 percent) responded to a mail survey and we interviewed the rest by telephone (49 percent). Response rates, by site, ranged from 51 to 89 percent (Table A.3).

About 88 percent of teachers completed questionnaires that provided data on 79 percent of students (Table A.2). Most teachers responded by mail (69 percent) or telephone (31 percent). Response rates across sites ranged from 46 to 100 percent (Table A.3).

We obtained school records for 87 percent of students (Table A.2). We collected more than 80 percent of records at all but two sites, with response rates ranging from 61 percent to 100 percent (Table A.3). Generally, students for whom we were unable to collect records had transferred to other schools outside the district.

B. Data Collection for Middle-School Sites for the Second Follow-Up Wave

1. Student Survey

About six weeks before the end of the 2001-2002 school year, field staff members administered the second follow-up questionnaire. Ninety percent of the 4,264 students in the study completed the questionnaire (Table A.4) and response rates were greater than 80 percent at all but one site (Table A.5).⁴⁰ Nearly all students completed the survey in school (84 percent).

⁴⁰At two of the 34 middle-school sites, the 2000-2001 baseline administration of student surveys was delayed because of the time needed to reach agreement about participating in the study and for obtaining parental consent. When reporting the data collected on individual students (from students, parents, teachers, and school records), we excluded those two sites from the baseline and follow-up analyses; when reporting data collected from center staff members, we included those two sites.

Table A.4

Sample Sizes and Response Rates for Second Followup
Middle-School Sites

Instrument	Sample Size		Response Rate			
	Total N	Treatment N %	Comparison N %	Total N %	Treatment N %	Comparison N %
Student Survey	4,264	1,782 42	2,482 58	3,856 90	1,620 91	2,236 90
Parent Survey	4,264	1,782 42	2,482 58	3,480 82	1,444 81	2,036 82
Teacher Survey ^a	4,264	1,782 42	2,482 58	3,686 86	1,540 86	2,146 86
School Record	4,264	1,782 42	2,482 58	3,905 92	1,641 92	2,264 91

^aSample size and response rates are based on number of students, not teachers; 89.1 percent of the 1,188 teachers in the sample completed surveys.

The others (16 percent), primarily transfer students, completed the questionnaire with computer-assisted telephone interviewers.

2. Data Collected from Parents, Teachers, and Records

Beginning in late spring of 2000-2001, we collected follow-up data from parents, English teachers, and school records, using the same instruments as in the first followup. Eighty-two percent of parents completed the follow-up questionnaire (Table A.4). A little more than half (54 percent) responded to a mail survey and we interviewed the rest by telephone (46 percent). Response rates by site ranged from 70 to 93 percent (Table A.5). Eighty-nine percent of teachers completed questionnaires, which provided data on 86 percent of students (Table A.4). Most teachers responded by mail (71 percent), with the remainder completing the questionnaire by telephone (29 percent). Response rates across sites ranged from 10 to 100 percent (Table A.5).

We obtained school records for 92 percent of students (Table A.4). With the exception of one site (which lost its 21st Century grant and did not cooperate with the study from that point forward), response rates ranged from 79 percent to 100 percent (Table A.5). Generally, students for whom we were unable to collect school records had transferred to other schools.

Table A.5

Distribution of Response Rates, By Site, for Middle-School Second Followup

Instrument	Number of Sites						
	Total	Percentage					
		90 to 100	80 to 89	70 to 79	60 to 69	50 to 59	Less than 50
Student Survey	32	19	12	1	0	0	0
Parent Survey	32	3	18	11	0	0	0
Teacher Survey	32	16	12	2	0	0	2
School Record	32	23	7	1	0	0	1

C. Data Collected From Center and School Staff Members

As we did for the 2000-2001 academic year, we collected data on schools and centers from principals and program staff members for the 2001-2002 school year. Ninety-six percent of principals, 98 percent of project directors, and 92 percent of project staff members completed a questionnaire (Table A.6). Most responded by mail.

We collected program attendance records from all centers that had active 21st Century programs (Table A.6). The centers provided copies of their records in whatever form they typically maintained attendance, such as by day or by activities offered each day. In a few cases, centers provided the total number of days students attended, rather than the daily attendance records. In principle, the elementary-school study design precluded attendance by students in the control group. However, records showed that 9 percent of control-group students attended the

Table A.6

Sample Sizes and Response Rates: Data Collected from School and After-School Center Staff Members In 2001-2002

Instrument	Sample Size	Response Rate	
		N	%
Principal Survey ^a	82	79	96
Project Director Survey ^a	44	43	98
Center Coordinator Survey ^{a,b}	90	78	87
Staff Survey ^{a,c}	323	296	92
Program Attendance Record ^d	74	74	100

^aIncludes 44 sites (the 21st Century program at two middle-school sites had closed in the 2001-2002 year; those sites were not included).

^bTen after-school programs had two center coordinators; both coordinators returned surveys at nine after-school programs.

^cWe drew a random sample of staff members from all after-school programs.

^dIncludes attendance collected from all programs that had a 21st Century after-school program during the 2001-2002 academic year, from three sites that were funding their after-school program with non-21st Century sources, and from two sites that previously did not have a 21st Century program. Two sites had no 21st Century program that year; two sites had a program in the fall semester only; and another site had a program in the spring semester only. At least partial attendance was obtained from each site, and complete attendance was obtained from 91 percent of the sites.

program for at least one day. Of the control-group students who attended the program, average attendance was 43 days. For the control group as a whole, average attendance was four days.⁴¹

The middle-school study design did not restrict comparison-group students from attending the program. Eleven percent of comparison-group students attended the program at least once. Most attended from one to 25 days (71 percent) and average attendance was 20 days. For the comparison group as a whole, average attendance was two days.

⁴¹Reasons control-group members attended the program were related mostly to program-staff changes and miscommunication. New staff members were not always aware that some students had been assigned not to attend the program.

D. Procedures for Constructing Nonresponse Weights

Preliminary analyses of missing data found evidence of systematic patterns that were related to treatment status. For example, in the elementary-school study, treatment-group students who were Hispanic were more likely to lack data for some outcomes. Nonresponse weights were used to adjust for the missing data.

Nonresponse weights were calculated by identifying how nonrespondents differed from respondents in terms of baseline characteristics. Respondents who were most similar to nonrespondents were given a greater weight, which enabled them to “represent” nonrespondents.

Nonresponse weights were constructed using a propensity-score approach. The probability of responding to the follow-up survey was modeled as a logistic function of student baseline characteristics similar to those used as control variables in estimating impacts. For each respondent, the predicted probability of response was calculated using the estimated model. Respondents who were most similar to nonrespondents generally were those with the lowest predicted probabilities of response. The nonresponse weight is the inverse of this predicted probability. For example, a respondent that had a predicted probability of responding to the follow-up survey of 0.25 was given a nonresponse weight of 4, whereas a respondent with a predicted probability of 0.90 was given a nonresponse weight of 1.1. Weights were then normalized so they summed to the original sample size.

The exact procedure used to estimate nonresponse weights followed three steps:

1. Estimate a logistic regression in which the dependent variable is a binary indicator of survey response and the independent variables are site indicators and baseline characteristics of students, and drop insignificant baseline characteristics (those with a p-value greater than 0.3). Retain site indicators regardless of significance. Use a stepwise procedure to identify significant interactions that might improve the model's explanatory power.
2. If any weights from the model in the first step are large (greater than 3), investigate trimming. (On inspection, none of the weights required trimming.)
3. Multiply the initial sample weight by the nonresponse weight and normalize the resulting weight so that it matches the sum of the original sampling weight. Use this final weight in impact regressions.

We used this procedure to construct nonresponse weights for the parent, teacher, student, and records surveys for both elementary and middle school, and for elementary-school reading tests, resulting in nine sets of weights. The goodness-of-fit of the propensity score models was high, with the models able to correctly predict 71 to 87 percent of responses (depending on the data source).

Appendix B

Study Design and Methods for Estimating Impacts

The impacts and outcome differences reported in the text are based on two study designs, one for elementary school students that used random assignment, the other for middle-school students that used matched-comparison groups. The first report described the two designs and presented evidence about how well the designs created treatment groups and control or comparison groups that were similar at baseline. Below, we briefly sketch aspects of the designs and methods used to estimate program impacts and outcome differences. We also describe the method used to estimate “attendance” impacts.

A. Study Designs

The design for measuring impacts at the elementary school sites was based on random assignment of students to treatment or control groups. Students and their parents applied to the program by completing a brief information form and consent form. Their applications were sent to Mathematica Policy Research, Inc. (MPR) and, after checking the information for completeness, MPR researchers conducted random assignment and sent the results to the staff members. For seven sites, random assignment took place at the beginning of the 2000-2001 school year; for the remaining five sites, random assignment took place at the beginning of the 2001-2002 school year.

The design for measuring outcome differences at the middle school sites was based on a matched-comparison group design. Thirty-five sites first were selected randomly to represent sites serving middle school students. One site declined to participate and two sites did not carry out timely baseline data collection, which left 32 sites as the basis for the estimates. Site weights were adjusted to maintain the representative nature of the sampling design.

At each site, nonparticipating students were matched to participating students using propensity-score matching techniques. Matching was based on 38 characteristics derived from

the student baseline questionnaire. Generally, matching resulted in similar groups on the 38 characteristics used to match students, but some characteristics differed after matching. All characteristics were included in regression models to adjust for remaining differences and to improve the precision of the estimates.

1. Methods for Estimating Impacts for the Elementary School Study

For elementary schools, the impact estimation approach used regression models that included outcomes at the first follow-up as dependent variables and variables created by interacting treatment status with the 12 site indicators and student baseline characteristics as independent variables. The models yielded 12 impact estimates, one for each site, and the overall impacts were then calculated as the simple mean of the 12 site-specific impacts. The variance of the estimator was derived from the variance-covariance matrix of the 12 site impact estimates.

A two-stage procedure was used to estimate impacts on elementary-school participants. In the first stage, an indicator for whether students participated in the program was regressed on treatment status and baseline characteristics. In the second stage, outcomes at the first follow-up were regressed on predicted participation from the first-stage and the baseline characteristics.⁴²

2. Methods for Estimating Outcome Differences for the Middle School Study

For middle schools, the estimation approach used regression models with outcomes at the second follow-up as dependent variables and treatment status and baseline characteristics as independent variables. Because sites were sampled with unequal probabilities and had

⁴²It is common in program evaluation for some treatment-group members not to participate in the program after random assignment occurs. A simple estimator of program impact on participants is to divide the overall impact estimate by the participation rate. The two-stage adjustment used in this study is the regression analog of that technique, which also adjusts for control-group students who cross over into the program.

associated sampling weights, the regression models were estimated using SUDAAN[®] so that variances of the estimates included design effects from sampling.⁴³

We tested the use of a two-stage procedure to estimate outcome differences for participating students. In the first stage, regression models were estimated with second-year participation status as the dependent variable, and treatment status and baseline characteristics as independent variables.⁴⁴ In the second stage, regression models were estimated with outcomes from the second follow-up as dependent variables, and the predicted level of participation from the first stage and the baseline characteristics as independent variables. We experimented with using access to centers as a variable in the first stage. Access to centers (whether students attended a school that operated a center) was correlated with whether students participated. However, access in the second year proved to be negatively correlated with a range of academic outcomes at baseline, which did not satisfy a condition for being an instrumental variable for the participation model. Using treatment status as an instrumental variable required assuming that all outcome differences observed in the second year were experienced only by students who participated in the second year, which is untenable.

B. Measuring the Impacts of Attendance

Policymakers often want to know whether greater participation in a program is related to larger effects.⁴⁵ This is especially important for after-school programs because attendance is voluntary and frequency of attendance is highly variable.

⁴³The stratified sampling design selected about the same number of sites from strata that included different numbers of grantees, which resulted in unequal selection probabilities.

⁴⁴Nonparticipation in the first year is not possible because first-year treatment status and participation are synonymous for middle school students. To be selected for the study's treatment group, students had to participate.

⁴⁵The relationship between attendance and outcomes sometimes is referred to as a "dosage" effect. However, because other aspects of program services and activities also can be viewed as related to dosage, such as how often

Since students and parents choose how often to attend, attendance decisions may be related to a host of factors not observed by the study, and estimates of the impact of more attendance that do not address the unobserved factors could be misleading. For example, the motivation to succeed in school is difficult to observe, and it may well increase program attendance and academic outcomes at the same time. Comparing academic outcomes for students with high and low attendance could reveal that students with high attendance had better academic outcomes; but the difference could be due more to motivation than to program attendance.

To explore this issue and understand how various methods of estimating impacts may be biased by unobserved factors, we ran a set of simulations to compare how various estimation methods are affected by unobserved factors. The simulations show that common approaches for estimating attendance impacts can be highly misleading. Results indicated that one method, “fixed effects,” yielded estimates that are least affected by bias, and we used this approach to estimate attendance effects.

The method does not eliminate the possibility that attendance impacts are mismeasured. If a circumstance arises that increases (or decreases) attendance and outcomes between the first and second year, and the circumstance is not related to program services and activities, the fixed-effect method nonetheless will attribute the outcome difference to program attendance. Compelling examples of these circumstances are not obvious, but they can be constructed. For example, suppose in the second year, an after-school program experienced many illness-related absences because of flu or some other contagion, while at the same time the absences reduced outcomes because students missed regular school. In this case, the fixed-effect approach will

(continued)

students participate in academic activities during the after-school program or the intensity of the academic activities, we focus the discussion here on attendance.

estimate a positive impact from program attendance, even if the program did not improve outcomes, because outcomes declined when program attendance declined. Note that if illness equally affected students in the program and students not in the program, the fixed-effect method would not yield a biased estimate, because the treatment and control or comparison groups would experience similar declines in outcomes from the first to the second year. The circumstance needs to affect only one group to create the potential for bias.

1. Implementing the Fixed-Effect Approach

Two aspects of the fixed-effect approach needed additional consideration for use in estimating attendance impacts. First, we wanted the models to allow for attendance to have different impacts at different attendance levels. For example, attending the program more often (say, 10 or 20 additional days) is likely to have different impacts on outcomes for a student who attended 30 days, compared to a student who attended 100 days. The models we estimated allowed for the nonlinear relationship by including a squared attendance variable (see Equation 1 below).

$$(1) \quad y_{i,t} = \beta_0 + \beta_1 X_{i,t} + \beta_2 d_{i,t} + \beta_3 d_{i,t}^2 + \beta_4 X_{i,t} u_i + \beta_5 u_i + \varepsilon_{i,t}^y$$

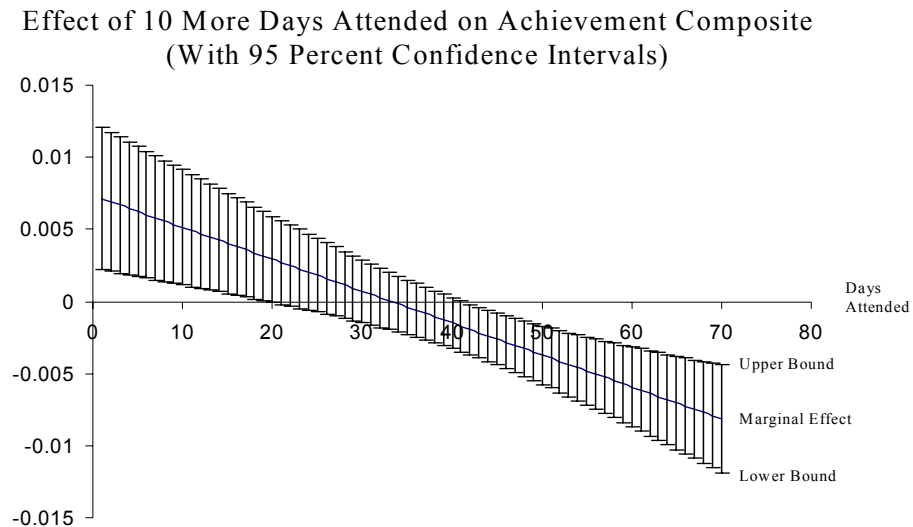
In Equation 1, there are two time periods, $t = 1, 2$. Time period 1 corresponds to the first follow-up, and time period 2 to the second follow-up. Variable d is a measure of program attendance, y is an outcome, observable characteristics used as regressors are represented by X . Finally, each student's "fixed effect" is designated by u , and the error term is represented by ε . For continuous outcomes, the marginal impact of attendance from Equation 1 is simply $\hat{\beta}_2 + 2\hat{\beta}_3 d$, where $\hat{\beta}_2$ and $\hat{\beta}_3$ are estimates of β_2 and β_3 .

The variance of the marginal impact is a function of the estimated parameters and depends on the assumed attendance level.

$$\text{Variance of Marginal Impact} = \text{Var}(\hat{\beta}_2) + 4d^2\text{Var}(\hat{\beta}_3) + 4d \cdot \text{Cov}(\hat{\beta}_2, \hat{\beta}_3)$$

Figure B.1 shows how the marginal impact for attendance varies with attendance, while also using the variance formula to create confidence intervals. The outcome is a composite measure of student achievement that was based on a set of items reported by teachers (the findings from Chapter III are used in the figure). The figure shows that the impact of additional attendance is larger at low levels of attendance and diminishes as attendance increases. The confidence interval also shows that the variance of the marginal impact increases as attendance moves away from average attendance.

Figure B.1



The second technical consideration is that we estimate a fixed-effect model for binary outcomes using logistic regression, rather than ordinary least squares, using the approach developed by Chamberlain (1980). A known feature of the fixed-effect logit model is that it uses only cases for which the binary dependent variable changes value between time periods (for a two-period model, the variable goes from zero to one or one to zero). The model requires coding students as a “1” if the dependent variable goes from 0 to 1, or as a “0” if the variable goes from 1 to 0. For example, a student who indicated that he or she always did homework in both time periods provides no information for the model and is not used in the estimation. However, if the student indicated that he or she did not always do homework in the first time period but always did homework in the second time period, the student receives a value of “1” for the dependent variable and is used in the estimation. The estimates for the model, therefore, are based on a sample size that varies by outcome.

For the estimates reported in the text, the sample of more than 4,000 students at baseline typically included 500 to 800 students who changed status over the 2 years and thus were included in the fixed-effect logit estimation. If we assume that the attendance impact is the same for all students, the dropping of cases will not bias impact estimates. However, if we assume some students have larger attendance impacts than others (some students respond more than others to attending the program), the fixed-effect logit estimator will estimate the attendance impact for students who have the larger values, which may overstate the attendance impact. In contrast, ordinary least squares models include all students but treat the binary outcome as if it were continuous, thereby introducing other possible specification errors. In Appendix C, we investigate this specification issue by comparing estimates from both approaches.

We calculated the variance of the marginal effect of attendance in the fixed-effect logit using a bootstrapping approach. We first calculated the predicted probability of the outcome at

the lower attendance level and the predicted probability of the outcome at the higher attendance level, and computed the difference between the two predicted probabilities. The difference is an estimate of the marginal impact of the difference in attendance. We then conducted 500 replications of the procedure (500 bootstrap replications) and estimated the variance of the marginal impact as the variance of the impact in the 500 replications.

Appendix C

Sensitivity Tests and Results for Alternative Specifications

This appendix presents results for alternative specifications and sensitivity tests that were conducted to assess the robustness of the findings. For both the elementary and middle school designs, we assessed the effects of using nonresponse weights and regression adjustment methods, the possibility that findings could be attributed to outlier sites, and the effects of using alternative definitions of self-care.

For the middle school design, we analyzed the same issues and took three additional steps: (1) analyzed the efficacy of using regression-adjustment models to reduce baseline differences in the treatment and comparison groups, (2) estimated outcome differences for students who had access to the program in the second year, and (3) analyzed the relationship between center attendance and outcomes. Step 3 included comparing attendance-outcome findings estimated via fixed-effects logit models with those estimated via ordinary least squares (OLS) models. This appendix presents the findings separately for the two designs.

A. Elementary Schools

As described in Appendix A, nonresponse weights were used because exploratory analyses found that missing data were correlated with treatment status and student characteristics. Using nonresponse weights improves the representativeness of the estimated impacts for the full sample.

Regression adjustment was used to increase the efficiency of the impact estimates. With random assignment, the variables used in the regression models are not correlated with treatment status by construction, but we expect greater precision in the regression-adjusted estimates.

1. Sensitivity of Estimates to Weights and Regression-Adjustment

We compared four sets of impacts: (1) those that use nonresponse weights and regression adjustment, (2) those that use the weights but not regression adjustment, (3) those that do not use

the weights but use regression adjustment, and (4) those that do not use the weights or regression adjustment. (The first set is used in the main text; the other three are included in appendix tables and figures.) Table C.1 presents the results. Comparing the first two columns provides a sense of how regression adjustment may have modified the impacts. The estimates are similar in the two columns and two of the 24 outcomes had a higher level of significance when regression adjustment was used. (Standard errors also were smaller with regression adjustment, although the reduction usually was not large enough to change the level of significance from .05 to .01.)

Comparing the first and third columns provides a sense of how nonresponse weighting may have modified the impacts. The last column presents impacts estimated as simple treatment-control differences. The point estimates are similar to the estimates in the first column. Three impacts that were significant in the first column were not significant in the fourth, which may reflect the lower precision of the simple estimator. Overall, the results appear to be robust to weights and regression adjustment.

2. Consistency of Impacts Across Sites

A measured impact could be attributable to an outlier site or set of sites, which would reduce confidence in the generalizability of the findings. For example, a positive impact that, on closer inspection, resulted from a large impact in one of 12 sites and no impact in 11 sites, might suggest an unusual experience in the one site.

To investigate this issue, we first compared the impact findings with the number of sites that had positive or negative impacts (regardless of statistical significance). We did the comparison for all main impacts, but here we show one table to illustrate how we did the analysis. Table C.2 shows insignificant impacts and significant impacts. For one outcome, whether students report

Table C.1

Sensitivity of Impact Estimates to Alternative Specifications, Elementary School Centers, Year 1

Outcome	With Nonresponse Weights and Regressors	With Nonresponse Weights, No Regressors	No Nonresponse Weights, with Regressors	No Nonresponse Weights, No Regressors
Percentage of Students Under the Following Types of Supervision at Least Three Days After School in a Typical Week, According to Parent Reports:				
Self-care ^a	0.1	0.0	0.0	-0.1
Parent care	-10.4***	-9.8***	-10.1***	-9.4***
Non-parent adult care	11.0***	10.5***	10.6***	10.0***
Sibling care	-5.5**	-5.5**	-5.1	-5.2
Mixed care (not in any one category for at least three days)	0.7	0.7	0.7	0.7
Percentage of Students in the Following Locations After School at Least Three Days in a Typical Week, According to Parent Reports:				
Own home	-18.3***	-17.8***	-17.9***	-17.4***
Someone else's home	-2.4	-2.7	-3.2	-3.5
School or other place for activities	21.8***	21.3***	21.2***	20.6***
Somewhere to "hang out"	-0.3	-0.3	-0.5	-0.5
Mixed location (not in one location for at least three days)	0.4	0.4	0.3	0.3
Percentage of Students Who Reported That They "Often" or "Always" Complete the Homework Teachers Assign ^b	0.8	0.6	0.4	0.8
Percentage of Students Whose Teachers Reported That They Often Complete the Homework Teachers Assign	-5.8**	-5.5	-4.8	-4.0
Mean Grade:				
Math	0.3	-0.1	0.2	-0.1
English/language arts	0.1	-0.2	0.0	-0.3
Science	0.2	-0.2	0.3	-0.2
Social studies/history	0.4	-0.1	0.3	-0.2
Mean Reading Test Score	-0.9	-0.5	0.0	0.2
Percentage of Students Who Reported Feeling the Following Levels of Safety After School up Until 6 p.m.:				
Very safe	1.5	0.3	0.0	-0.7
Somewhat safe	1.4	2.3	2.6	3.0
Not at all safe	-3.0**	-2.7**	-2.6	-2.3
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:				
Attended an open house at the school	0.0	-1.9	0.4	-1.4
Attended parent-teacher organization meetings	2.6	0.4	3.9	1.9
Attended an after-school event	9.2***	7.9**	9.6***	8.4***
Volunteered to help out at school	-4.1	-6.1**	-3.8	-5.7
Sample Size^c	1,719	1,719	1,719	1,719

SOURCE: Parent Survey, Student Survey.

^a Students are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least three days in a typical week.^b The original set of seven sites was not asked these questions in the first year of the study.^c Sample sizes differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table C.2

Number of Sites With Positive or Negative Impacts on Other Outcomes,
Elementary School Centers, Year 1

Outcome	Estimated Impact	Number of Positive Site Impacts	Number of Negative Site Impacts	Joint Significance of Site Impacts (p-Value) ^a
Percentage of Students Who Reported Feeling the Following Levels of Safety After School up Until 6 p.m.:				
Very safe	1.5	5	4	0.19
Somewhat safe	1.4	5	4	0.46
Not at all safe	-3.0**	3	6	0.38
Percentage of Students Who Reported the Following Are “Somewhat True” or “Very True”:				
They get along with others their age	-5.1	3	6	0.00***
They feel left out of things	0.1	5	4	0.44
Percentage of Students Who Reported Doing the Following “Some” or “A Lot”:				
Help another student in school	-4.2	4	5	0.07
Help another student after school	8.0**	8	1	0.96
Percentage of Students Who Rated Themselves as “Good” or “Excellent” on the Following:				
Working with others on a team or group	-2.8	5	4	0.02**
Feeling bad for other people who are having difficulties	-3.9	4	5	0.50
Believing the best about other people	-0.1	3	6	0.91
Percentage of Students Who Rated Themselves as “Excellent” on the Following:				
Using a computer to look up information	1.6	6	3	0.63
Setting a goal and working to achieve it	-2.0	5	4	0.06
Percentage of Students Who Rated Themselves as “Excellent” on Sticking to What They Believe In, Even if Their Friends Don’t Agree				
	-0.7	4	5	0.90
Negative Behavior Composite ^b	0.0	4	5	0.70
Percentage of Students Whose Parents Reported Doing the Following:				
Helped their child with homework at least three times last week	8.4***	8	4	0.00***
Checked on their child’s homework completion at least three times last week	2.1	7	5	0.02**
Asked their child about things they were doing in class at least seven times last month	6.3**	9	3	0.03**
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:				
Attended an open house at the school	0.0	6	6	0.09
Attended parent-teacher organization meetings	2.6	9	3	0.33
Attended an after-school event	9.2***	11	1	0.17
Volunteered to help out at school	-4.1	4	8	0.00***
Sample Size^c	1,539			

SOURCE: Student Survey, Parent Survey.

^aTo examine the joint significance of the site impacts, we tested whether they were jointly significantly equal to the mean of the site impacts.

^bThe negative behavior composite is based on student responses to five questions regarding how often they (1) break something on purpose, (2) punch or hit someone, (3) argue with their parents, (4) lie to their parents, and (5) give a teacher a “hard time.” Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level, while a value of 4 indicates a high level.

^cSample sizes differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

feeling unsafe after school, the impact was a reduction of three percentage points (an increase in students' feeling safe). Across the sites, six had an impact estimate with a negative sign and three sites had an impact with a positive sign. Similarly, we found a large positive impact for whether parents attended after-school events—and, on closer inspection, 11 of 12 sites also had a positive impact. We found no impact for whether parents attended a school open house, and sites were divided evenly between positive (six sites) and negative impacts (six sites).

We also conducted tests to determine whether site impacts differed from the average impact. Six of the 21 tests indicated that site impacts differed from the average impact (Table C.2). The likely reason for these differences is that there were one or two large site impacts in the opposite direction of the overall impact. This suggests that, for at least some outcomes, site-specific factors were related to impacts.

3. Self-Care Alternative Definitions

Changing the definition of self-care altered its levels but did not change the impact findings. We investigated four self-care definitions to assess the robustness of the estimated impact on self-care: a student was defined to be in self-care if (1) the student did not spend at least 3 days with a parent, a nonparent adult, or an older sibling in a typical week; (2) the student did not spend at least 1 day with a parent, a non-parent adult, or an older sibling in a typical week; (3) the student was alone at least 3 days in a typical week; and (4) the student was alone at least 1 day in a typical week. Overall, results from these additional analyses suggest that the definition of self-care used does not affect the findings (Table C.3). In addition, using nonresponse weights and regressors did not change the self-care findings. In all cases, the impact of the program on self-care is insignificant.

Table C.3

Sensitivity of Various Self-Care Impact Estimates to Alternative Specifications, Elementary School Centers, Year 1

Outcome	With Nonresponse Weights and Regressors	With Nonresponse Weights, No Regressors	No Nonresponse Weights, with Regressors	No Nonresponse Weights, No Regressors
Percentage of Students in Self-Care at Least Three Days After School in a Typical Week, According to Parent Reports (Self-Care Defined as Not Being in Parent, Non-Parent Adult, or Older Sibling Care)	0.1	0.0	0.0	-0.1
Any Self-Care After School in a Typical Week, According to Parent Reports (Self-Care Defined as Not Being in Parent, Non-Parent Adult, or Older Sibling Care)	-0.1	-1.7	-0.3	-0.4
Percentage of Students in Self-Care at Least Three Days After School in a Typical Week, According to Parent Reports (Self-Care Defined as Being Alone After School)	0.0	0.6	0.0	-0.1
Any Self-Care After School in a Typical Week, According to Parent Reports (Self-Care Defined as Being Alone After School)	1.1	1.6	0.8	0.8
Sample Size	1,719	1,719	1,719	1,719

SOURCE: Parent Survey, Student Survey.

B. Middle Schools

1. Regression Adjustment for Baseline Differences

Middle school treatment and comparison groups differed on several baseline characteristics after propensity score matching (see Chapter III, Table III.2). To increase the validity of the outcome difference estimates, we used regression models to adjust for baseline differences.

We tested the efficacy of regression adjustment by estimating regression models in which the baseline outcome is the dependent variable. If regression adjustment was successful, there should be no impact of being in the treatment group on the baseline outcome, because at that point students had not yet been “treated.” However, the regression models we used to estimate outcome differences included the baseline value of the outcome, which generally is the most powerful predictor of the follow-up value of the outcome. Testing the efficacy of regression adjustment for the same model would have required a pre-baseline value of the outcome (the model would have the baseline outcome as a dependent variable and have the pre-baseline value as a predictor variable along with other predictor variables). Since we have only the baseline value of the outcome, we can only investigate how regression adjustment for other variables reduces any baseline differences, which is a weaker test.

Table C.4 shows that regression adjustment substantially reduced baseline differences. The first column presents raw difference at baseline between the treatment and comparison groups for six variables. The third column shows the differences after adjusting for other baseline variables except the outcome itself. For average grades, for example, the raw difference of -0.94 (the treatment group’s average grades were 0.94 points lower than the comparison group’s on a 100-point scale) is statistically significant. The adjusted difference is very small, -0.04 , and not significant. The tests show that regression adjustment did not remove all differences, however, as is shown for the homework habits outcome. Instead, the tests suggest that the use of regression

Table C.4

Examining the Effect of Regressors on Baseline Differences between Treatment and Comparison Groups,
Middle School Centers

Outcome	Unadjusted Treatment, Comparison Baseline Difference	p- value ^a	Regression- Adjusted Treatment, Comparison Baseline Difference	p- value ^a
Average Grades	-0.94***	0.01	-0.04	0.26
Mean of Homework Habits Index ^b	-0.05**	0.02	-0.05**	0.04
Student-Based Discipline Problem Composite ^c (Mean)	0.06***	0.00	0.03	0.07
Negative Behavior Composite ^d (Mean)	0.03	0.07	0.01	0.64
Student-Reported Tobacco, Alcohol, and Drug Use Composite ^e (Mean)	0.02	0.10	0.00	0.83
Mean of Safety Index ^f	-0.04**	0.03	-0.01	0.50
Sample Size	4,128			

SOURCE: Student Survey.

^aThe p-value is the smallest level of significance at which the null hypothesis that the difference in means between program participants and comparison group members equals zero can be rejected. If the p-value is less than .01, the difference is significant at the 1 percent level. If the p-value is less than .05, the difference is significant at the 5 percent level, and so on.

^bThe homework habits index is based on student responses to how often they (1) do the homework the teachers assign, (2) do homework in the same place each day, (3) do homework at the same time each day, and (4) write down homework assignments. The index is equal to the mean of the four variables. A value of 1 on the index indicates poor homework habits, whereas a value of 4 indicates good homework habits.

^cThe student-based discipline problem composite is based on four items: the extent to which students report (1) skipping school or class, (2) getting sent to the office for doing something wrong, (3) getting detention, and (4) having their parents called to school about a problem they are having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems.

^dThe negative behavior composite is based on student responses to eight questions regarding how frequently they (1) break something on purpose, (2) punch or hit someone, (3) argue with their parents, and (4) lie to their parents, (5) steal from a store, (6) give a teacher a hard time, (7) sell illegal drugs, and (8) get arrested or detained by police. Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level of negative behavior, while a value of 4 indicates a high level of negative behavior.

^eThe tobacco, alcohol, and drug use composite is based on seven items: the extent to which students (1) smoke cigarettes, (2) use smokeless tobacco, (3) have at least one drink of alcohol, and (4) have five or more drinks of alcohol in a row, (5) smoke marijuana, (6) use inhalants, and (7) use any other illegal drug. Values on these items range from 1 to 4; a value of 1 on the composite indicates no substance abuse, while a value of 4 indicates frequent substance abuse.

^fThe safety index mean is based on how often the student (1) feels safe walking in his or her neighborhood, (2) feels safe being at home alone, (3) feels safe on the ground outside school, (4) feels safe going to the bathroom at school, and (5) feels safe in the hallways at school. A value of 1 indicates feeling less safe and a value of 4 indicates feeling more safe.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

adjustment models did reduce differences between treatment and comparison groups, and the use of baseline values for the outcome in the models that were used to estimate outcome differences reported in the text is likely to have reduced differences further.

2. Use of Nonresponse Weights and Regression Adjustment Models

Table C.5 presents the results for the same analysis previously shown for the elementary school design of the effects of using nonresponse weights and regression adjustment. Because of the role regression adjustment plays in reducing baseline differences, we expect outcome differences to differ when regression adjustment is used, and comparing the first two columns indicates that they do. For example, whether students are in their own home after school was not statistically significant when regression adjustment was used, and is more negative and significant at the five percent significance level when regression adjustment is not used. Using nonresponse weights also modified estimated outcome differences. Comparing the first and third columns, a number of outcome differences are numerically different and have different levels of statistical significance when weights are used. When regression adjustment is not used, weights have almost no effect on outcome differences, which can be seen by comparing the results in the second and fourth columns.

3. Consistency of Outcome Differences Across Sites

As with the elementary school design, we examined site-level outcome differences for the middle school design to assess whether outcome differences were associated with outlying sites. Table C.6 presents an example of the analysis. The table shows that statistically significant outcome differences generally are evident when a majority of sites have an outcome difference with the same sign. For example, we found a statistically significant increase in whether

Table C.5

Sensitivity of Outcome Differences to Alternative Specifications, Middle School Centers, Year 2

Outcome	With Nonresponse Weights, with Regressors	With Nonresponse Weights, No Regressors	No Nonresponse Weights, with Regressors	No Nonresponse Weights, No Regressors
Percentage of Students in the Following Types of Supervision at Least Three Days After School in a Typical Week:				
Self-care ^a	-0.8	-0.1	-1.4	-0.2
Parent care	-2.1	-5.6***	-4.1**	-5.8***
Nonparent adult care	5.3	6.8***	7.4***	6.6***
Sibling care	-3.7**	-4.8***	-5.1***	-4.6***
Mixed care (not in any one category for at least three days)	-1.4	-0.7	0.0	-0.6
Percentage of Students in the Following Locations After School at Least Three Days in a Typical Week:				
Own home	-2.3	-4.5**	-4.3**	-4.7***
Someone else's home	0.8	0.8	0.5	0.8
School or other place for activities	4.4**	5.9***	6.8***	5.6***
Somewhere to "hang out"	2.4	1.1	0.8	1.2
Mixed location (not in one location for at least three days)	0.4	-0.3	0.0	-0.2
Employment of Mother (parent-reported):				
Full-time	-2.7	-3.1	-2.4	-1.3
Part-time	2.2	1.0	1.7	1.0
Looking for work	-0.4	0.2	-0.5	0.3
Not in the labor force	0.9	1.9	1.2	2.1
Percentage of Students Whose Teachers Reported That They "Often" Complete Their Homework				
	-0.8	-5.3***	-0.8	-5.3***
Mean Grade:				
Math	0.7	-0.7	0.7	-0.6
English	0.5	-0.9	0.5	-0.9**
Science	0.6	-0.7	0.7	-0.6
Social studies/history	1.7***	0.6	1.5***	0.4
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:				
Attended an open house at the school	0.7	1.8	1.2	2.2
Attended parent-teacher organization meetings	1.4	2.5	2.0	3.0
Attended an after-school event	1.8	1.6	1.3	0.8
Volunteered to help out at school	1.9	1.5	1.6	1.1
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6:00 P.M.:				
Very safe	-2.4	-3.4	-3.0	-3.7**
Somewhat safe	2.1	2.8	2.7	3.1
Not at all safe	0.3	0.6	0.3	0.6
Sample Size^b	3,808	3,808	3,808	3,808

SOURCE: Student Survey, Parent Survey.

^aStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least three days in a typical week.^bSample sizes differ for some outcomes due to non-response.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table C.6

Number of Sites with Positive or Negative Outcome Differences on Student Safety, Negative Behavior, and Victimization, Middle School Centers, Year 2

Outcome	Difference	Number of Positive Site Differences	Number of Negative Site Differences	Joint Significance of Site Differences (p-Value) ^a
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6:00 P.M.:				
Very safe	-2.4	11	21	0.01***
Somewhat safe	2.1	19	13	0.01**
Not at all safe	0.3	19	13	0.99
Percentage of Students Who Reported That They Do the Following "Some" or "A Lot":				
Break something on purpose	2.4**	22	10	0.41
Punch or hit someone	2.7	17	15	0.41
Steal from a store	0.9	15	17	0.20
Sell illegal drugs	-0.3	11	21	0.88
Get arrested or detained by police	0.2	15	17	0.10
Negative Behavior Composite ^b (Mean)	0.03**	18	14	0.85
Percentage of Students Who Reported the Following Happened to Them "Some" or "A Lot":				
Been offered, sold, or given an illegal drug	-1.0	13	19	0.95
Been "picked on" after school	3.0	23	9	0.00***
Been threatened or hurt with a weapon	1.0	16	16	0.30
Been threatened by a gang or gang member	0.2	15	17	0.39
Had property damaged on purpose	2.4	20	12	0.00***
Percentage of Students Who Reported That They Do the Following "Some" or "A Lot":				
Smoke cigarettes	0.6	17	15	0.91
Have at least one alcoholic drink	0.8	17	15	0.09
Smoke marijuana	0.5	18	14	0.30
Take illegal drugs such as cocaine, ecstasy, or LSD	0.6***	19	13	0.96
Student-Reported Tobacco, Alcohol, and Drug Use Composite ^c (Mean)	0.02	16	16	0.27
Sample Size^d	3,818			

SOURCE: Student Survey.

^aTo examine the joint significance of the site estimates, we tested whether they were jointly significantly equal to the mean of the site estimates.

^bThe negative behavior composite is based on student responses to eight questions regarding how frequently they (1) break something on purpose, (2) punch or hit someone, (3) argue with their parents, (4) lie to their parents, (5) steal from a store, (6) give a teacher a hard time, (7) sell illegal drugs, and (8) get arrested or detained by police. Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level of negative behavior, while a value of 4 indicates a high level of negative behavior.

^cThe tobacco, alcohol, and drug use composite is based on seven items: the extent to which students (1) smoke cigarettes, (2) use smokeless tobacco, (3) have at least one drink of alcohol, (4) have five or more drinks of alcohol in a row, (5) smoke marijuana, (6) use inhalants, and (7) use any other illegal drug. Values on these items range from 1 to 4; a value of 1 on the composite indicates no substance abuse, while a value of 4 indicates frequent substance abuse.

^dSample sizes may differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

*** Significantly different from zero at the .01 significance level, two-tailed test

students reported that they “broke things on purpose,” and 22 of the 32 sites also had an increase in this behavior.

4. Alternative Definitions of Self-Care

As we did with the elementary school impacts, we investigated four self-care definitions to assess the robustness of the estimated outcome difference on self-care: a student was defined to be in self-care if (1) the student did not spend at least 3 days with a parent, a nonparent adult, or an older sibling in a typical week; (2) the student did not spend at least 1 day with a parent, a nonparent adult, or an older sibling in a typical week; (3) the student was alone at least 3 days in a typical week; and (4) the student was alone at least one day in a typical week. Overall, results from these additional analyses suggest that the definition of self-care used does not affect the findings (Table C.7).

Table C.7 also shows how estimates change with the inclusion of weights and regressors. In general, their inclusion does not affect estimates. However, estimates from one definition of self-care do change, depending on whether regressors and weights are included. When regressors are not included and nonresponse weights are included, the estimate of the outcome difference on self-care defined as being alone at least 3 days in a week is statistically significant. When regressors are included to control for baseline differences between students, which was shown to be necessary because of some baseline differences between treatment and comparison students—the estimates become insignificant.

5. Findings for Students With Program Access

The study estimated outcome differences for students who had access to centers during the second year of the study, because many students had graduated to high school or transferred to

Table C.7

Sensitivity of Alternative Self-Care Outcome Differences to Alternative Specifications, Middle School Centers, Year 2

Outcome	With Nonresponse Weights, with Regressors	With Nonresponse Weights, No Regressors	No Nonresponse Weights, with Regressors	No Nonresponse Weights, No Regressors
Percentage of Students Who Report Being in Self-Care at Least Three Days After School in a Typical Week (Self-Care Defined as Not Being in Parent, Nonparent Adult, or Older Sibling Care)	-0.8	-0.1	-1.4	-0.2
Any Self-Care After School in a Typical Week, According to Student Reports (Self-Care Defined as Not Being in Parent, Nonparent Adult, or Older Sibling Care)	-1.3	0.5	-1.0	0.1
Percentage of Students Who Report Being in Self-Care at Least Three Days After School in a Typical Week (Self-Care Defined as Being Alone After School)	-1.1	-1.7**	-1.1	-1.6
Any Self-Care after School in a Typical Week, According to Student Reports (Self-Care Defined as Being Alone After School)	-1.0	-1.3	-0.9	-1.3
Sample Size	3,808	3,808	3,808	3,808

SOURCE: Student Survey.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

other schools and, therefore, did not have access to a 21st Century center in the study's second year.

Including only students who had access required modifying the comparison group. We first identified students in the full comparison group who had been matched at baseline to treatment-group students who had access. This comparison group was older, on average, than the group of treatment students with program access, an artifact of the initial matching process. The age difference arises because treatment students with program access mostly were 6th and 7th graders (nearly all 8th graders had moved on to high school and no longer had access to centers). However, the initial comparison group still included 8th graders because in the original matching process some 6th- and 7th-grade treatment students were matched to 8th-grade comparison students. The matching method used to construct the treatment and comparison groups did not force students who were matched to be in the same grade.

To create more balance in the comparison group, we dropped from the treatment and comparison groups any students who did not have a matching student at the same grade level (6th graders to 6th graders and 7th graders to 7th graders). The result was that 356 treatment-group members who were matched at baseline with comparison-group members at other grade levels were dropped.

Baseline characteristics of students with access to the program were similar for the rebalanced groups (Table C.8). Only student feelings of safety differed between the two groups, a difference that also was found for the full sample of students.

Location, Supervision, and Activities After School. Students with access to centers were less likely to be in parent care and less likely to be in their own homes after school, and more likely to be in the care of other adults (Table C.9; effect sizes of 0.07, 0.14, and 0.18, respectively). Whether self-care was significantly reduced for students with program access is

Table C.8

Characteristics of Center Participants and Comparison-Group Students:
Middle School Centers

Characteristic	Full Sample			Students with Access to the Program		
	Percentage of Program Participants	Percentage of Comparison Group Members	p-value ^a	Percentage of Program Participants	Percentage of Comparison Group Members	p-value ^a
Demographics						
Gender						
Male	47.3	46.5	0.62	46.5	47.4	0.77
Female	52.7	53.5	0.62	53.5	52.7	0.77
Race/Ethnicity						
White (non-Hispanic)	38.2	40.6	0.33	41.9	41.2	0.19
Black (non-Hispanic)	27.7	24.7	0.33	26.1	23.3	0.19
Hispanic	12.3	12.0	0.33	12.6	10.8	0.19
Other	15.5	15.9	0.33	14.7	17.3	0.19
Mixed race	6.3	6.9	0.33	4.6	7.3	0.19
Grade Level						
6	20.7	21.6	0.19	29.5	29.8	0.47
7	37.8	38.2	0.19	55.7	56.4	0.47
8	33.7	34.1	0.19	4.2	5.5	0.47
Other or ungraded	7.8	6.2	0.19	10.5	8.4	0.47
Primary Language in the Home is Not English	17.8	18.9	0.39	14.3	17.3	0.19
Academic and Other Outcomes at Baseline						
Student-Reported Baseline Grades						
Mostly A's	30.4	34.1	0.00***	32.1	35.0	0.57
Mostly B's	35.8	36.5	0.00***	37.7	38.2	0.57
Mostly C's	23.2	21.3	0.00***	20.7	18.6	0.57
Mostly D's or below	8.8	7.5	0.00***	7.9	7.4	0.57
Not graded	1.8	0.7	0.00***	1.6	0.8	0.57
Average Grades	83.1	84.0	0.01***	84.1	84.5	0.41
Homework						
Mother or father helps student with homework	63.1	63.2	0.93	67.1	63.6	0.19
Mean of homework habits index ^b	2.80	2.85	0.02**	2.86	2.85	0.95
Mean of Index of Positive Behavior ^c	3.02	3.01	0.52	3.06	3.05	0.80
Student-Based Discipline Problem Composite ^d (Mean)	1.39	1.33	0.00***	1.36	1.34	0.47
Mean of Parental Discipline Index ^e	2.92	2.94	0.46	3.00	2.99	0.92
Negative Behavior Composite ^f (Mean)	1.55	1.52	0.07	1.51	1.50	0.98
Mean of Index of Empathy ^g	3.10	3.10	0.94	3.16	3.14	0.73
Mean of Index of Controlling Destiny ^h	3.00	3.00	0.81	3.05	3.03	0.67

Table C.8 (continued)

Characteristic	Full Sample			Students with Access to the Program		
	Percentage of Program Participants	Percentage of Comparison Group Members	p-value ^a	Percentage of Program Participants	Percentage of Comparison Group Members	p-value ^a
Student-Reported Tobacco, Alcohol, and Drug Use Composite (Mean) ⁱ	1.12	1.11	0.10	1.10	1.09	0.33
Mean of Safety Index ^j	3.33	3.37	0.03**	3.28	3.35	0.04**
Sample Size^k	1,727	2,385		664	604	

SOURCE: Student Survey, Parent Survey.

^aThe p-value is the smallest level of significance at which the null hypothesis that the difference in means between program participants and comparison group members equals zero can be rejected. If the p-value is less than .05, the difference is significant at the 5 percent level, and if the p-value is less than .01, the difference is significant at the 1 percent level.

^bThe homework habits index is based on student responses to how often they (1) do the homework the teachers assign, (2) do homework in the same place each day, (3) do homework at the same time each day, and (4) write down homework assignments. The index is equal to the mean of the four variables. A value of 1 on the index indicates poor homework habits, whereas a value of 4 indicates good homework habits.

^cThe positive behavior index is based on how often the student (1) helps another kid in school, (2) helps her parents, and (3) goes to church, temple, or mosque. A value of 1 on the index indicates never doing the aforementioned, while a value of 4 indicates doing them often.

^dThe student-based discipline problem composite is based on four items: the extent to which students report (1) skipping school or class, (2) getting sent to the office for doing something wrong, (3) getting detention, and (4) having their parents called to school about a problem they are having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems.

^eThe parental discipline index is based on student responses to how often parents (1) check on whether homework is completed, (2) limit the amount of time available to watch TV, (3) decide which TV shows their kids are allowed to watch, and (4) tell their children not to drink alcohol or use drugs. A value of 1 on the composite indicates parents who engage in less discipline, while a value of 4 indicates parents who engage in more discipline.

^fThe negative behavior composite is based on student responses to eight questions regarding how frequently they (1) break something on purpose, (2) punch or hit someone, (3) argue with their parents, (4) lie to their parents, (5) steal from a store, (6) give a teacher a hard time, (7) sell illegal drugs, and (8) get arrested or detained by police. Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level of negative behavior, while a value of 4 indicates a high level of negative behavior.

^gThe empathy index is based on student ratings of ability (1) work with others on a team or on a group project, (2) feel bad for other people when they are having a hard time, and (3) believe the best about other people. A value of 1 on the index indicates poor ability, while a value of 4 indicates excellent ability.

^hThe controlling destiny index is based on student ratings of ability (1) set goals and work to achieve them, (2) plan for things needed in the future, (3) work out conflicts or disagreements with others, and (4) stick to beliefs even if friends disagree. A value of 1 on the index indicates poor ability, while a value of 4 indicates excellent ability.

ⁱThe tobacco, alcohol, and drug use composite is based on seven items: the extent to which students (1) smoke cigarettes, (2) use smokeless tobacco, (3) have at least one drink of alcohol, (4) have five or more drinks of alcohol in a row, (5) smoke marijuana, (6) use inhalants, and (7) use any other illegal drug. Values on these items range from 1 to 4; a value of 1 on the composite indicates no substance abuse, while a value of 4 indicates frequent substance abuse.

^jThe safety index is based on how often the student (1) feels safe walking in his or her neighborhood, (2) feels safe being at home alone, (3) feels safe on the ground outside school, (4) feels safe going to the bathroom at school, and (5) feels safe in the hallways at school. A value of 1 indicates feeling less safe and a value of 4 indicates feeling more safe.

^kSample sizes may differ due to missing values.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table C.9

Outcome Differences in Maternal Employment and Students' Location, Supervision, and Activities After School, Middle School Centers, Year 2

Outcome	Center Participants	Comparison Group	Difference for Full Sample	Difference for Students with Program Access
Percentage of Students With the Following Individuals at Least Three Days After School in a Typical Week:				
Self-care ^a	19.0	19.8	-0.8	-4.7***
Parent	50.9	53.0	-2.1	-3.5***
Nonparent adult	33.9	28.6	5.3	8.4***
Sibling	17.5	21.2	-3.7**	-1.4
Mixed (Not in any one category for at least three days)	4.0	5.4	-1.4	-2.3
Percentage of Students in the Following Locations After School at Least Three Days in a Typical Week:				
Own home	69.2	71.5	-2.3	-6.6***
Someone else's home	12.6	11.8	0.8	1.2
School or other place for activities	27.5	23.2	4.4**	8.3**
Somewhere to "hang out"	12.9	10.5	2.4	1.0
Mixed location (Not in one location for at least three days)	8.2	7.8	0.4	1.8
Employment of Mother (Parent-reported):				
Full-time	59.9	62.6	-2.7	0.8
Part-time	15.7	13.4	2.2	0.6
Looking for work	8.7	9.1	-0.4	0.3
Not in the labor force	15.7	14.9	0.9	0.3
Mean Number of Days Stayed After School for Activities in Typical Week	1.0	0.8	0.2**	0.4***
Percentage of Students Who Participated in the Following Activities After School:				
Homework	84.6	86.7	-2.2	-0.6
Tutoring	18.1	15.1	3.0	5.1
Non-homework reading, writing, or science activities	43.9	41.9	2.0	7.1***
School activities (Band, drama, etc.)	32.1	29.3	2.7	6.6***
Lessons (Music, art, dance, etc.)	23.8	20.7	3.2**	3.5
Organized sports	41.5	40.1	1.5	2.7
Clubs (Boy and Girl Scouts, Boys and Girls Club, etc.)	15.7	12.2	3.5**	2.7
Activities at church, temple, mosque	30.5	29.6	1.0	0.6
Watched TV or videos	89.1	87.7	1.5	0.9
Surfed the Internet or did other things on a computer	64.9	64.8	0.2	3.9
"Hung out" with friends	82.1	78.1	4.1***	7.1***
Volunteered or did community service	17.8	15.4	2.4	4.2***
Worked at a job	20.5	19.0	1.6	4.4
Did chores around the house	77.8	79.0	-1.3	-2.3
Took care of a brother or sister	50.3	49.7	0.7	1.1
Mean Time Students Reported Watching Television in the Past Day (Hours)	2.0	2.0	0.02	-0.01
Mean Time Students Reported Reading for Fun in the Past Day (Hours)	0.3	0.3	0.02	0.03**
Sample Size^b	1,605	2,203		1,176

SOURCE: Student Survey, Parent Survey.

Table C.9 (continued)

NOTE: The percentages and mean values of outcomes for participants and comparison-group members have been regression-adjusted for baseline differences between the groups. The control variables in the regressions include student characteristics such as indicators of students' demographic characteristics, students' baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated outcome differences shown in the table do not always equal the difference between center participants and the comparison group. Weights are used to adjust estimates for non-response. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Appendix A describes how weights were constructed, and Appendix B describes methods used to estimate outcome differences.

^aStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least three days in a typical week.

^bSample sizes differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

unclear. We looked at four different definitions of self-care, and one of the four differed significantly between the treatment and comparison groups.

Students who had access to centers were more likely to participate in non-homework reading, writing, or science activities; school activities; and volunteering (effect sizes of 0.14, 0.14, and 0.12, respectively).

Academic Outcomes. Students with access to centers did not differ from comparison students on homework completion, time spent working on homework, suspensions, absenteeism, lateness, or math, English, science, or history grades (Tables C.10 and C.11). The outcome difference for history grades differed between the full sample of students and students with program access. The outcome difference for history grades was significant for the full sample of students, but was insignificant for students with program access.

Classroom effort is another area in which there are differences between the estimates based on the full sample of students and those based on students with program access. Students with access to centers had significantly higher levels of effort according to teachers (effect size of 0.10). This finding is consistent with the findings in the first report, which also showed increased classroom effort.

Homework Assistance. Students with access to centers were more likely to have their homework checked by other adults (Table C.12). In particular, they were more likely to have an adult ask if their homework was complete, look at their homework to see if it was correct, and explain homework in an understandable way. There was no difference in whether students received homework assistance either from parents or other adults.

Educational Aspirations. Here there were no differences between treatment students with program access and comparison students (Table C.13). This estimate differs from the full sample estimate, in which treatment students were more likely to aspire to graduate from college.

Table C.10

Outcome Differences in Homework Completion and on Behavior and Level of Effort in the Classroom,
Middle School Centers, Year 2

Outcome	Center Participants	Comparison Group	Difference for Full Sample	Difference for Students with Program Access
Percentage of Students Who Reported That They “Often” or “Always” Complete the Homework Teachers Assign	81.3	83.0	-1.7	-6.4
Percentage of Students Whose Teachers Reported That They “Often” Complete Their Homework	49.8	50.5	-0.8	1.6
Mean Amount of Time Students Spent Doing Homework the Last Time They Had Homework (Hours)	0.9	1.0	-0.1	-0.06
Percentage of Students Whose Teachers Reported the Following:				
“Agree” or “strongly agree” that student completes assignments to the teacher’s satisfaction	53.4	55.2	-1.8	3.3
Student “usually tries hard” in English class	49.3	48.4	1.0	2.9
Student “often” performs at or above ability level	41.5	43.8	-2.3	-1.3
Teacher-Reported Level of Effort Composite ^a (Mean)	3.5	3.5	0.0	0.10***
Percentage of Students Who Reported that They Pay Attention to their Teachers in School	83.4	87.1	-3.7**	-2.4
Percentage of Students Whose Parents “Agree” or “Strongly Agree” That Their Child Works Hard at School	78.5	76.0	2.5	-2.7
Student-Based Discipline Problem Composite ^b (Mean)	1.4	1.4	0.0	0.07***
Teacher-Based Discipline Problem Composite ^c (Mean)	1.4	1.4	0.0	-0.03**
Percentage of Students Who Were Suspended During 2001-2002 School Year	21.9	21.7	0.2	1.6
Mean Number of Days Student Was:				
Absent	9.0	10.0	-1.0**	-0.3
Late	6.2	5.4	0.8	0.8
Sample Size^d	1,633	2,198		1,150

SOURCE: Student Survey, Teacher Survey, Parent Survey, School Records.

Table C.10 (continued)

NOTE: The percentages and mean values of outcomes for participants and comparison-group members have been regression-adjusted for baseline differences between the groups. The control variables in the regressions include student characteristics such as indicators of students' demographic characteristics, students' baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated outcome differences shown in the table do not always equal the difference between center participants and the comparison group. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Appendix A describes how weights were constructed and Appendix B describes methods used to estimate outcome differences.

^aThe level of effort composite is based on five items reported by teachers: whether the student (1) usually tries hard, (2) often performs at or above his or her ability level, (3) is attentive in class, (4) participates in class, and (5) volunteers in class. The composite is equal to the mean of the five variables. Values on these items range from 1 to 5; a value of 1 on the composite indicates a low level of effort, and a value of 5 indicates a high level of effort.

^bThe student-based discipline problem composite is based on four items: the extent to which students report (1) skipping school or class, (2) getting sent to the office for doing something wrong, (3) getting detention, and (4) having their parents called to school about a problem they are having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems.

^cThe teacher-based discipline problem composite is based on four items: the extent to which the teacher reports that student are (1) skipping school or class, (2) getting sent to the office for doing something wrong, (3) getting detention, and (4) having their parents called to school about a problem they are having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems.

^dSample sizes may differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table C.11

Outcome Differences in Teacher-Reported Achievement and Grades,
Middle School Centers, Year 2

Outcome	Center Participants	Comparison- Group Members	Difference for Full Sample	Difference for Students with Program Access
Percentage of Students Whose Teachers Reported That They Achieve at an “Above-Average” or “Very High” Level	31.3	33.8	-2.5	-0.1
Percentage of Students Whose Teachers “Agree” or “Strongly Agree” That They Get Good Grades on Tests	50.8	51.8	-1.0	1.2
Teacher-Reported Achievement Composite (Mean) ^a	3.3	3.3	0.0	0.04
Mean Grade				
Math	79.3	78.6	0.7	1.1
English	80.1	79.6	0.5	0.6
Science	79.6	79.0	0.6	0.6
Social studies/history	81.6	79.8	1.7***	1.7
Sample Size^b	1,533	2,126		1,150

SOURCE: Teacher Survey, School Records.

NOTE: The percentages and mean values of outcomes for participants and comparison-group members have been regression-adjusted for baseline differences between the groups. The control variables in the regressions include student characteristics such as indicators of students’ demographic characteristics, students’ baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated outcome differences shown in the table do not always equal the difference between center participants and the comparison group. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Appendix A describes how weights were constructed and Appendix B describes methods used to estimate outcome differences.

^aThe teacher-reported achievement composite is based on teacher responses to five questions: (1) At what level is this student performing in reading? (2) Does this student get good grades on tests? (3) Does this student complete assignments to my satisfaction? (4) Does this student have good communication skills? (5) Is this student a proficient reader? Values on these items range from 1 to 5; a value of 1 on the composite indicates low achievement, and a value of 5 indicates high achievement.

^bSample sizes may differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table C.12

Outcome Differences in Quality of Homework Assistance, Middle School Centers, Year 2

Outcome	Center Participants	Comparison Group	Difference for Full Sample	Difference for Students with Program Access
Percentage of Students Who Reported That Their Parent “Often” or “Always”:				
Asks if homework is complete	76.1	76.1	0.0	0.8
Looks at homework to see if it is complete	42.5	45.1	-2.7	-3.3
Looks at homework to see if it is correct	38.5	41.8	-3.3	-2.0
Explains homework in a way that is easy to understand	45.3	49.4	-4.1	-3.5
Percentage of Students Who Reported That an Adult Who is Not Their Parent “Often” or “Always”:				
Asks if homework is complete	38.8	35.3	3.5	8.3***
Looks at homework to see if it is complete	29.1	28.4	0.8	1.7
Looks at homework to see if it is correct	29.4	25.8	3.6	5.9**
Explains homework in a way that is easy to understand	35.3	33.7	1.6	5.8**
Percentage of Students Who Reported That Their Parent or an Adult Who is Not Their Parent “Often” or “Always”:				
Asks if homework is complete	80.5	80.4	0.1	1.5
Looks at homework to see if it is complete	52.0	52.6	-0.6	-1.8
Looks at homework to see if it is correct	49.2	49.1	0.1	0.5
Explains homework in a way that is easy to understand	56.6	58.5	-1.9	0.1
Percentage of Students Who Had the Following Individual Ask the Child To Correct Parts of Homework:				
Parent	75.0	76.3	-1.3	-0.1
An adult who is not their parent	57.1	54.6	2.5	2.9
A parent or an adult who is not their parent	83.3	83.1	0.1	0.5
Sample Size^a	1,633	2,198		1,062

SOURCE: Student Survey.

NOTE: The percentages and mean values of outcomes for participants and comparison-group members have been regression-adjusted for baseline differences between the groups. The control variables in the regressions include student characteristics such as indicators of students’ demographic characteristics, students’ baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated outcome differences shown in the table do not always equal the difference between center participants and the comparison group. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Appendix A describes how weights were constructed and Appendix B describes methods used to estimate outcome differences.

^aSample sizes may differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table C.13

Outcome Differences in Social Engagement, Educational Expectations, and Parental Involvement,
Middle School Centers, Year 2

Outcome	Center Participants	Comparison-Group Members	Difference for Full Sample	Difference for Students with Program Access
Social Engagement Composite ^a (Mean)	3.54	3.56	-0.02	-0.01
Peer Interaction/Empathy Composite ^b (Mean)	3.01	3.03	-0.02	0.02
Percentage of Students Who Rated Themselves as “Good” or “Excellent” at Working Out Conflicts with Others	57.4	60.7	-3.3	-6.0***
Percentage of Students Who Rated Themselves as “Good” or “Excellent” on Using a Computer to Look Up Information	36.9	36.6	0.3	0.6
Percentage of Students Who Think They Will:				
Graduate from college	82.1	79.6	2.5**	2.8
Graduate from high school but not college	16.5	18.5	-2.0	-3.4
Attend high school but not graduate	1.4	1.9	-0.6	0.5
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:				
Attended an open house at the school	19.5	18.8	0.7	0.0
Attended parent-teacher organization meetings	26.8	25.4	1.4	3.9
Attended an after-school event	38.8	37.0	1.8	0.5
Volunteered to help out at school	16.1	14.2	1.9	3.2**
Sample Size^c	1,601	2,208		1,168

SOURCE: Student Survey, Parent Survey.

NOTE: The percentages and mean values of outcomes for participants and comparison-group members have been regression-adjusted for baseline differences between the groups. The control variables in the regressions include student characteristics such as indicators of students’ demographic characteristics, students’ baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated outcome differences shown in the table do not always equal the difference between center participants and the comparison group. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Appendix A describes how weights were constructed and Appendix B describes methods used to estimate outcome differences. Percentages may not sum to 100 because of rounding.

^aThe social engagement composite is based on five items: the extent to which students report that they (1) have friends to “hang out with,” (2) are never lonely, (3) get along with others their age, (4) find it easy to make new friends, and (5) never feel left out of things. The composite is equal to the mean of the five variables. Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level of social engagement, and a value of 4 indicates a high level of engagement.

^bThe peer interaction/empathy composite is based on three items: students’ rating of their ability to (1) work with others on a team or in a group, (2) feel bad for other people who are having difficulties, and (3) believe the best about other people. Values on these items range from 1 to 4; a value of 1 on the composite indicates poor peer interactions, while a value of 4 indicates excellent peer interactions.

^cSample sizes may differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Social, Emotional, and Developmental Outcomes. There were few significant differences in these outcomes among students with program access (Table C.13).

Social engagement, empathy for others, and the ability to use a computer to look up information did not differ between the two groups. Students with program access were less likely to rate themselves highly at working out conflicts with others.

Parental Involvement. For students who had access to centers, parents were more likely to volunteer at school (Table C.13). This finding is consistent with findings from the first-year report. It is worth noting, however, that only one of the four parent-involvement estimates is statistically significant for students with program access, while all four estimates were significant in the first report.

Feelings of Safety. There were no differences between treatment students with program access and comparison students in feelings of safety after school (Table C.14).

Negative Behaviors. There was mixed evidence on negative behaviors for students with program access (Table C.14). Among students who had access to centers, there was an increase in punching or hitting someone. However, there were no differences between the two groups on other measures of negative behavior, such as breaking something on purpose or stealing from a store.

Victimization. There was mixed evidence on victimization for students with access to the program in the study's second year (Table C.14). For students who had access to the program, there was an increase in being picked on after school. At the same time, there were no differences between the two groups on other outcomes such as being threatened with a weapon or by a gang.

Drug and Alcohol Use. Among students with program access, there was mixed evidence on the use of drugs and alcohol (Table C.14). Treatment students were more likely than

Table C.14

Outcome Differences in Student Safety, Negative Behavior, and Victimization, Middle School Centers, Year 2

Outcome	Center Participants	Comparison Group	Difference for Full Sample	Difference for Students with Program Access
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6:00 P.M.:				
Very Safe	64.6	66.9	-2.4	1.2
Somewhat safe	32.7	30.6	2.1	-1.9
Not at all safe	2.7	2.5	0.3	0.7
Percentage of Students Who Reported That They Do the Following "Some" or "A Lot":				
Break something on purpose	10.4	8.0	2.4**	1.8
Punch or hit someone	22.4	19.7	2.7	2.6**
Steal from a store	4.9	4.0	0.9	1.2
Sell illegal drugs	1.4	1.8	-0.3	-0.6
Get arrested or detained by police	3.3	3.1	0.2	0.9
Negative Behavior Composite ^a (Mean)	1.56	1.53	0.03**	0.05
Percentage of Students Who Reported the Following Happened to Them "Some" or "A Lot":				
Been offered, sold, or given an illegal drug	18.1	19.1	-1.0	-1.9
Been "picked on" after school	27.7	24.7	3.0	4.5**
Been threatened or hurt with a weapon	6.8	5.9	1.0	0.5
Been threatened by a gang or gang member	7.2	7.0	0.2	-0.1
Had property damaged on purpose	13.5	11.1	2.4	3.1
Percentage of Students Who Reported That They Do the Following "Some" or "A Lot":				
Smoke cigarettes	4.7	4.1	0.6	-0.2
Have at least one alcoholic drink	9.8	9.0	0.8	1.1
Smoke marijuana	4.8	4.3	0.5	-0.1
Take illegal drugs such as cocaine, ecstasy, or LSD	0.8	0.2	0.6***	0.5**
Tobacco, Alcohol, and Drug Use Composite ^b (Mean)	1.14	1.12	0.02	1.7
Sample Size^c	1,609	2,209		1,174

SOURCE: Student Survey.

NOTE: The percentages and mean values of outcomes for participants and comparison-group members have been regression-adjusted for baseline differences between the groups. The control variables in the regressions include student characteristics such as indicators of students' demographic characteristics, baseline test scores, attendance, disciplinary problems, and self-reported grades. Due to rounding, estimated outcome differences shown in the table do not always equal the difference between center participants and the comparison group. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN[®] to account for the statistical sampling design. Appendix A describes how weights were constructed and Appendix B describes methods used to estimate outcome differences. Percentages may not sum to 100 because of rounding.

^aThe negative behavior composite is based on student responses to eight questions regarding how frequently they: (1) break something on purpose, (2) punch or hit someone, (3) argue with their parents, (4) lie to their parents, (5) steal from a store, (6) give a teacher a hard time, (7) sell illegal drugs, and (8) get arrested or detained by police. Values on these items range from

Table C.14 (continued)

1 to 4; a value of 1 on the composite indicates a low level of negative behavior, while a value of 4 indicates a high level of negative behavior.

^bThe tobacco, alcohol, and drug use composite is based on seven items: the extent to which students report that they (1) smoke cigarettes, (2) use smokeless tobacco, (3) have at least one drink of alcohol, (4) have five or more drinks of alcohol in a row, (5) smoke marijuana, (6) use inhalants, and (7) use any other illegal drug. Values on these items range from 1 to 4; a value of 1 on the composite indicates no substance abuse, while a value of 4 indicates frequent substance abuse.

^cSample sizes may differ for some outcomes due to nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

*** Significantly different from zero at the .01 significance level, two-tailed test.

comparison students to report using cocaine, ecstasy, or LSD. However, there were no differences between the two groups on the extent to which students smoked cigarettes, drank alcohol, or smoked marijuana.

6. The Relationship Between Center Attendance and Outcomes

Having two years of attendance and outcome data allows us to explore the relationship between attendance and outcomes that could not be explored in the first report because only one year of data was available. Students could attend more or less often in the two years, and the differences in attendance could affect outcomes. Because we can observe the same students in two different time periods, the influence of unobservable factors that may vary across students, and that may affect both attendance and outcomes, can be reduced.

The analysis of the relationship between center attendance and outcomes found that some outcomes improved when students attended centers more often. The incidence of self-care was lower, students exerted greater effort in class, and parents were more involved (Tables C.15 through C.19). For example, if students attended centers for 40 days compared to 30 days, the likelihood that they were in self-care fell by 2.2 percentage points, and the likelihood that they were supervised by other adults increased by a roughly offsetting amount, 2.5 percentage points (Table C.15).⁴⁶ Attending centers more often also increased student participation in tutoring and school activities and the number of days students stayed after school for activities. (Tables C.15). However, attending centers more often did not improve academic outcomes such as course

⁴⁶We estimated models that allowed for impacts of attendance to differ, depending on the initial level of attendance. For example, additional attendance could have different impacts when the initial attendance level is low than when it is high. Statistical significance also could differ, depending on the estimated impact. Some impacts were statistically significant at lower levels of initial attendance but not at higher levels, and vice versa. For example, attending centers more often had a statistically significant impact on student achievement in the classroom as reported by teachers when the initial attendance level was 10 days, but attending centers more often did not have a statistically significant impact when the initial attendance level was 30 days (Table C.17).

Table C.15

Differences in Maternal Employment and Students' Location, Supervision,
and Activities After School by Attendance, Middle School Centers, Year 2

Outcome	Marginal Effect of Attending the After-School Program 10 More Days	
	Effect of 10 More Days For Those Attending 10 Days	Effect of 10 More Days For Those Attending 30 Days
Percentage of Students with the Following Individuals at Least Three Days after School in a Typical Week:		
Self-care ^b	-2.84	-2.23**
Parent	0.79	-0.07
Nonparent adult	1.29	2.50**
Sibling	0.74	1.02
Mixed (not in any one category for at least three days)	0.61	-1.94
Percentage of Students in the Following Locations After School at Least Three Days in a Typical Week:		
Own home	-1.38	-2.53**
Someone else's home	-4.78**	-3.13***
School or other place for activities	2.38	3.34***
Somewhere to "hang out"	-2.57	-0.84
Mixed location (not in one location for at least three days)	1.95	1.92
Employment of Mother:		
Full-time	0.88	-0.75
Part-time	-1.42	1.72
Looking for work	1.54	1.16
Not in the labor force	-1.42	-1.68
Mean Number of Days Stayed after School for Activities in Typical Week	0.18	0.16***
Percentage of Students Participating in the Following Activities After School:		
Homework	1.85	0.70
Tutoring	5.37***	3.98***
Non-homework reading, writing, or science activities	3.35	1.44
School activities (band, drama, etc.)	4.26***	2.94***
Lessons (music, art, dance, etc.)	0.86	1.98
Organized sports	2.41	0.66
Clubs (Boy and Girl Scouts, Boys and Girls Club, etc.)	2.61	1.58
Activities at church, temple, mosque	1.01	
Watched TV or videos	-2.13	0.81
Surfed the Internet or did other things on a computer	-0.22	-1.81
"Hung out" with friends	-0.41	0.83
Volunteered or did community service	-0.09	-0.09
Worked at a job	1.90	1.75
Worked at a job	1.59	1.50
Did chores around the house	0.18	-0.75
Took care of a brother or sister	0.56	0.24
Mean Time Students Reported Watching Television in the Past Day (Hours)	0.04	-0.02
Mean Time Students Reported Reading for Fun in the Past Day (Hours)	0.00	0.00
Sample Size^c	813	

SOURCE: Student Survey, Parent Survey.

Table C.15 (continued)

NOTE: All regressions include both linear and squared attendance terms to capture any diminishing returns to attendance. For binary outcomes, we use the *logit* command in STATA® to estimate fixed-effects logit models, and standard errors of the marginal effect are estimated by bootstrapping. Because bootstrapping does not account for the stratified sampling design, standard errors may be underestimated. For continuous outcomes, we estimate OLS fixed effects models by regressing the change in the outcome on the change in attendance and we use SUDAAN® to take into account the stratified sampling design when calculating standard errors. Weights are used to adjust estimates for nonresponse.

^aThe p-value is the smallest level of significance at which the null hypothesis that the difference in means between program participants and comparison group members equals zero can be rejected. If the p-value is less than .01, the difference is significant at the 1 percent level. If the p-value is less than .05, the difference is significant at the 5 percent level, and so on.

^bStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least three days in a typical week.

^cSample sizes can differ substantially both due to nonresponse and because the conditional logit drops all observations where the outcome does not change across time. The smallest sample size in this table is 223 for outcome “Mother not in labor force.” The largest sample size is 3,277 for outcome “Mean Time Students Reported Watching Television in the Past Day (Hours)”. The sample size reported in the table is the median sample size for the outcomes in this table.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table C.16

Differences in Homework Completion, Level of Effort, and Behavior in the Classroom
by Attendance, Middle School Centers, Year 2

Outcome	Marginal Effect of Attending the After-School Program 10 More Days	
	Effect of 10 More Days For Those Attending 10 Days	Effect of 10 More Days For Those Attending 30 Days
Percentage of Students Whose Teachers Reported That: Student "Often" Completes their Homework	1.62	
They "Agree" or "strongly agree" that student completes assignments to the teacher's satisfaction	3.87	-0.12
Student "usually tries hard" in English class	2.54	1.23
Student "often" performs at or above ability level	3.84	0.52
Teacher-Reported Level of Effort Composite ^a (Mean)	0.08***	2.16
Percentage of Students Who Report that They Pay Attention to Their Teachers in School	1.05	0.03**
Percentage of Students Whose Parents "Agree" or "Strongly Agree" That Their Child Works Hard at School	-1.34	0.41
Student-Based Discipline Problem Composite ^b (Mean)	-0.01	-1.17
Teacher-Based Discipline Problem Composite ^c (Mean)	-0.05	-0.01
Percentage of Students Who Were Suspended During 2001-2002 School Year	-0.05	-0.03**
Mean Number of Days Student Was: Absent	-5.42	-1.31
Late	-0.16	0.01
Sample Size^d	900	

SOURCE: Student Survey, Teacher Survey, Parent Survey, School Records.

NOTE: All regressions include both linear and squared attendance terms to capture any diminishing returns to attendance. For binary outcomes, we use the *logit* command in STATA[®] to estimate fixed-effects logit models, and standard errors of the marginal effect are estimated by bootstrapping. Because bootstrapping does not account for the stratified sampling design, standard errors may be underestimated. For continuous outcomes, we estimate OLS fixed effects models by regressing the change in the outcome on the change in attendance and we use SUDAAN[®] to take into account the stratified sampling design when calculating standard errors. Weights are used to adjust estimates for nonresponse.

^aThe level of effort composite is based on five items reported by teachers: whether the student (1) usually tries hard, (2) often performs at or above his or her ability level, (3) is attentive in class, (4) participates in class, and (5) volunteers in class. The composite is equal to the mean of the five variables. Values on these items range from 1 to 5; a value of 1 on the composite indicates a low level of effort, and a value of 5 indicates a high level of effort.

^bThe student-based discipline problem composite is based on four items: the extent to which students report (1) skipping school or class, (2) getting sent to the office for doing something wrong, (3) getting detention, and (4) having their parents called to school about a problem they are having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems.

^cThe teacher-based discipline problem composite is based on four items: the extent to which the teacher reports that the students: (1) skip school or class, (2) get sent to the office for doing something wrong, (3) get detention, and (4) have their parents called to school about a problem they are having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems.

^dSample sizes can differ substantially both due to nonresponse and because the conditional logit drops all observations where the outcome does not change across time. The smallest sample size in this table is 512 for outcome "Student 'Often' Completes His or Her Homework." The largest sample size is 3,267 for outcome "Student-Based Discipline Problem Composite (Mean)". The sample size reported in the table is the median sample size for the outcomes in this table.

Table C.15 (*continued*)

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table C.17

Differences in Teacher-Reported Achievement and Grades by Attendance,
Middle School Centers, Year 2

Outcome	Marginal Effect of Attending the After-School Program 10 More Days	
	Effect of 10 More Days For Those Attending 10 Days	Effect of 10 More Days For Those Attending 30 Days
Percentage of Students Whose Teachers Reported That They Achieve at an "Above- Average" or "Very High" Level	3.71	-0.07
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That - They Get Good Grades on Tests	2.20	1.55
Teacher-Reported Achievement Composite (Mean) ^a	0.05**	0.01
Mean Grade:		
Math	0.58	0.25
English	0.27	0.17
Science	-0.08	0.03
Social studies/history	0.05	-0.03
Sample Size^b	2,588	

SOURCE: Teacher Survey, School Records.

NOTE: All regressions include both linear and squared attendance terms to capture any diminishing returns to attendance. For binary outcomes, we use the *logit* command in STATA[®] to estimate fixed-effects logit models, and standard errors of the marginal effect are estimated by bootstrapping. Because bootstrapping does not account for the stratified sampling design, standard errors may be underestimated. For continuous outcomes, we estimate OLS fixed effects models by regressing the change in the outcome on the change in attendance and we use SUDAAN[®] to take into account the stratified sampling design when calculating standard errors. Weights are used to adjust estimates for nonresponse.

^aThe teacher-reported achievement composite is based on teacher responses to five questions: (1) At what level is this student performing in reading? (2) Does this student get good grades on tests? (3) Does this student complete assignments to my satisfaction? (4) Does this student have good communication skills? (5) Is this student a proficient reader? Values on these items range from 1 to 5; a value of 1 on the composite indicates low achievement, and a value of 5 indicates high achievement.

^bSample sizes can differ substantially both due to nonresponse and because the conditional logit drops all observations where the outcome does not change across time. The smallest sample size in this table is 459 for outcome "Percentage of Students Whose Teachers Report That They Achieve at an 'Above-Average' or 'Very High' Level." The largest sample size is 2,890 for outcome "English Grade." The sample size reported in the table is the median sample size for the outcomes in this table.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table C.18

Differences in Social Engagement, Educational Expectations, and Parental Involvement by Attendance,
Middle School Centers, Year 2

Outcome	Marginal Effect of Attending the After-School Program 10 More Days	
	Effect of 10 More Days For Those Attending 10 Days	Effect of 10 More Days For Those Attending 30 Days
Social Engagement Composite ^a (Mean)	0.00	0.00
Peer Interaction/Empathy Composite ^b (Mean)	-0.01	-0.01
Percentage of Students Who Rated Themselves as “Good” or “Excellent” at Working Out Conflicts with Others	-0.25	-0.74
Percentage of Students Who Rated Themselves as “Good” or “Excellent” on Using a Computer to Look Up Information	0.14	0.26
Percentage of Students Who Think They Will:		
Graduate from college	0.35	0.43
Graduate from high school but not college	-1.61	-0.77
Attend high school but not graduate	7.88**	1.52
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:		
Attended an open house at the school	3.39	2.92**
Attended parent-teacher organization meetings	4.21**	2.69***
Attended an after-school event	3.62	2.59**
Volunteered to help out at school	2.30	0.16
Sample Size^c	759	

SOURCE: Student Survey, Parent Survey.

NOTE: All regressions include both linear and squared attendance terms to capture any diminishing returns to attendance. For binary outcomes, we use the *logit* command in STATA[®] to estimate fixed-effects logit models, and standard errors of the marginal effect are estimated by bootstrapping. Because bootstrapping does not account for the stratified sampling design, standard errors may be underestimated. For continuous outcomes, we estimate OLS fixed effects models by regressing the change in the outcome on the change in attendance and we use SUDAAN[®] to take into account the stratified sampling design when calculating standard errors. Weights are used to adjust estimates for nonresponse.

^aThe social engagement composite is based on five items: the extent to which students report that they (1) have friends to “hang out with,” (2) are never lonely, (3) get along with others their age, (4) find it easy to make new friends, and (5) never feel left out of things. The composite is equal to the mean of the five variables. Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level of social engagement, and a value of 4 indicates a high level of engagement.

^bThe peer interaction/empathy composite is based on three items: students’ rating of their ability to (1) work with others on a team or in a group, (2) feel bad for other people who are having difficulties, and (3) believe the best about other people. Values on these items range from 1 to 4; a value of 1 on the composite indicates poor peer interactions, while a value of 4 indicates excellent peer interactions.

^cSample sizes can differ substantially both due to nonresponse and because the conditional logit drops all observations where the outcome does not change across time. The smallest sample size in this table is 108 for outcome “Attend high school but not graduate.” The largest sample size is 3,282 for outcome “Social Engagement Composite (Mean).” The sample size reported in the table is the median sample size for the outcomes in this table.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table C.19

Differences in Student Safety, Negative Behavior, and Victimization by Attendance, Middle School Centers, Year 2

Outcome	Marginal Effect of Attending the After-School Program 10 More Days	
	Effect of 10 More Days For Those Attending 10 Days	Effect of 10 More Days For Those Attending 30 Days
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6:00 P.M.:		
Very Safe	-2.20	-1.45
Somewhat safe	1.36	1.07
Not at all safe	5.27	2.25
Percentage of Students Who Reported That They Do the Following "Some" or "A Lot":		
Break something on purpose	-1.48	-0.24
Punch or hit someone	1.17	-0.42
Steal from a store	0.72	0.02
Sell illegal drugs	1.96	2.55
Get arrested or detained by police	1.34	0.48
Negative Behavior Composite ^b (Mean)	0.01	0.00
Percentage of Students Who Reported the Following Happened to Them "Some" or "A Lot":		
Been offered, sold, or given an illegal drug	-2.81	-1.25
Been "picked on" after school	3.11	1.10
Been threatened or hurt with a weapon	2.34	0.29
Been threatened by a gang or gang member	4.76	1.39
Had property damaged on purpose	2.42	-0.39
Percentage of Students Who Reported That They Do the Following "Some" or "A Lot":		
Smoke cigarettes	-4.33	-4.91
Have at least one alcoholic drink	-4.08	-2.89
Smoke marijuana	-3.51	-3.58
Student-Reported Tobacco, Alcohol, and Drug Use Composite ^c (Mean)	0.00	0.00
Sample Size^d	391	

SOURCE: Student Survey.

NOTE: All regressions include both linear and squared attendance terms to capture any diminishing returns to attendance. For binary outcomes, we use the *logit* command in STATA[®] to estimate fixed-effects logit models, and standard errors of the marginal effect are estimated by bootstrapping. Because bootstrapping does not account for the stratified sampling design, standard errors may be underestimated. For continuous outcomes, we estimate OLS fixed effects models by regressing the change in the outcome on the change in attendance and we use SUDAAN[®] to take into account the stratified sampling design when calculating standard errors. Weights are used to adjust estimates for nonresponse.

^aThe p-value is the smallest level of significance at which the null hypothesis that the difference in means between program participants and comparison group members equals zero can be rejected. If the p-value is less than .01, the difference is significant at the 1 percent level. If the p-value is less than .05, the difference is significant at the 5 percent level, and so on.

^bThe negative behavior composite is based on student responses to eight questions regarding how frequently they: (1) break something on purpose, (2) punch or hit someone, (3) argue with their parents, (4) lie to their parents, (5) steal from a store, (6) give a teacher a hard time, (7) sell illegal drugs, and (8) get arrested or detained by police. Values on these items range from 1 to 4; a value of 1 on the composite indicates a low level of negative behavior, while a value of 4 indicates a high level of negative behavior.

Table C.19 (continued)

^cThe tobacco, alcohol, and drug use composite is based on seven items: the extent to which students (1) smoke cigarettes, (2) use smokeless tobacco, (3) have at least one drink of alcohol, (4) have five or more drinks of alcohol in a row, (5) smoke marijuana, (6) use inhalants, and (7) use any other illegal drug. Values on these items range from 1 to 4; a value of 1 on the composite indicates no substance abuse, while a value of 4 indicates frequent substance abuse.

^dSample sizes can differ substantially both due to nonresponse and because the conditional logit drops all observations where the outcome does not change across time. The smallest sample size in this table is 117 for outcome “Sell illegal drugs.” The largest sample size is 3,271 for outcome “Negative Behavior Composite (Mean).” The sample size reported in the table is the median sample size for the outcomes in this table.

**Significantly different from zero at the .05 significance level, two-tailed test.

*** Significantly different from zero at the .05 significance level, two-tailed test.

grades (Table C.17), developmental outcomes such as working out conflicts and the extent of social engagement and empathy for others (Table C.18), or feelings of safety and negative behaviors (Table C.19).

The program access estimates presented in Tables C.9 to C.14 are based on whether students had access to the program in the second year, whereas the attendance estimates just discussed are based on how often students attended. We examined the direction and statistical significance of the two types of estimates to assess their consistency. Seventy percent of estimates were in a consistent direction, and almost 60 percent were consistent in both their direction and statistical significance. Both methods found increases in supervision by other adults, decreases in being at home after school, increases in being at school for activities, increases in students' reporting that they participated in school activities, and increased classroom effort. Also, both methods showed no significant estimates for most academic achievement measures, and student feelings of safety. The estimates were inconsistent for negative behaviors and drug and alcohol use.

7. Alternative Estimates of Attendance Outcome Differences

This section investigates how the estimates of attendance outcome differences just presented are affected by whether the estimate varies with attendance and by whether outcome differences are estimated using fixed-effects logit or ordinary least squares.

Nonlinear Effects of Attendance. Allowing for a nonlinear attendance effect generally did not modify the findings. For some outcomes, however, the nonlinear model indicated that the benefits of additional attendance depended on how frequently students already attended. Table C.20 presents the results for a selected set of outcomes that illustrate this point. The linear model indicates that the effect of attendance on the teacher-reported achievement variable was insignificant. The nonlinear model indicates that the effect of attendance on the variable was positive and significant at low attendance levels but insignificant at higher attendance levels

Table C.20

Sensitivity of Attendance Estimates to Specification For Teacher-Reported Achievement and Grades,
Middle School Centers, Year 2

Outcome	Attendance Estimates		
	Linear Model	Quadratic Model, Effect at 10 Days	Quadratic Model, Effect at 30 Days
Percentage of Students Whose Teachers Reported That They Achieve at an "Above-Average" or "Very High" Level	0.75	3.71	-0.07
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" That They Get Good Grades on Tests	1.66	2.20	1.55
Teacher-Reported Achievement Composite (Mean) ^a	-0.01	0.05**	0.01
Mean Grade:			
Math	0.06	0.58	0.25
English	0.11	0.27	0.17
Science	0.09	-0.08	0.03
Social studies/history	-0.09	0.05	-0.03
Sample Size^b	2,588		

SOURCE: Teacher Survey, School Records.

NOTE: The first column includes a linear attendance term; the last two columns also include a squared attendance term. All effects are scaled to represent the effect of an additional 10 days in the program. For continuous outcomes, we estimate OLS fixed effects models by regressing the change in the outcome on the change in attendance and other factors and we use SUDAAN® to take into account the stratified sampling design when calculating standard errors. For binary outcomes, we use the *logit* command in STATA® to estimate fixed-effects logit models, and standard errors of the marginal effect are estimated using a bootstrap method. Because bootstrapping does not account for the stratified sampling design, standard errors may be underestimated. Weights are used to adjust estimates for nonresponse.

^aThe teacher-reported achievement composite is based on teacher responses to five questions: (1) At what level is this student performing in reading? (2) Does this student get good grades on tests? (3) Does this student complete assignments to my satisfaction? (4) Does this student have good communication skills? (5) Is this student a proficient reader? Values on these items range from 1 to 5; a value of 1 on the composite indicates low achievement, and a value of 5 indicates high achievement.

^bSample sizes differ for some outcomes due to nonresponse and sample size restrictions imposed by the conditional logit. The number reported in the table is the median. The outcome with the smallest sample size is "Percentage of Students Whose Teachers Report That They Achieve at an 'Above-Average' or 'Very High' Level" with 459 and the outcome with the largest sample size is "English Grade" with 2,890.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

(Table C.20). A similar pattern is shown for the teacher-reported level of effort variable (not shown).

Fixed-Effects Logit and Ordinary Linear Squares. This report presents attendance outcome differences for discrete outcomes that were estimated using Chamberlain's (1980) fixed-effects logit model. Although the logit model is more appropriate for discrete outcomes, the fixed-effects logit excludes sample members when the outcome does not change across time periods. If the model is misspecified, the sample exclusions could result in biased estimates. In particular, if attendance estimates differ across students, the fixed-effects logit model may be excluding students for whom the effect is small and including students for whom the effect is larger.

An alternative estimation approach is to use a linear probability model and estimate fixed effects as if the outcome were continuous. We estimated outcome differences using both approaches and Tables C.21 and C.22 are examples of these estimates. The tables show the marginal impacts of attendance rather than the estimated coefficients so the estimates are in common units. The two methods often yield similar point estimates, but more of the OLS (ordinary least squares) estimates were statistically significant (possibly because of the larger sample size that OLS uses). For example, in Table C.21, 6 of the 29 fixed-effects logit estimates are statistically significant at the 5 percent level or better, whereas 12 of the 29 OLS estimates are significant. However, for Table C.22, none of the 16 fixed-effects logit estimates are significant at the 5 percent level, and only 1 of the OLS estimates is significant, a minor difference.

Table C.21

Sensitivity of Attendance Estimates to Estimation Technique for Maternal Employment and Students' Location, Supervision, and Activities After School, Middle School Centers, Year 2

Outcomes	Fixed Effects Logit	Fixed Effects OLS
Percentage of Students in the Following Types of Supervision at Least Three Days after School in a Typical Week:		
Self-care ^a	-2.22	-1.00**
Parent care	-0.20	-0.98**
Nonparent adult care	2.54**	2.62***
Sibling care	1.02	0.17
Mixed care (not in any one category for at least three days)	-1.06	-0.12
Percentage of Students in the Following Locations After School at Least Three Days in a Typical Week:		
Own home	-2.62***	-3.10***
Someone else's home	-3.26**	-1.27***
School or other place for activities	3.37***	3.35***
Somewhere to "hang out"	-0.82	-0.06
Mixed location (not in one location for at least three days)	1.94	0.65**
Employment of Mother (Parent-Reported):		
Full-time	-0.70	-0.23
Part-time	1.35	0.39
Looking for work	-1.72	0.21
Not in the labor force	-1.72	-0.38
Percentage of Students Who Reported Participating in the Following Activities After School:		
Homework	0.66	0.28
Tutoring	4.15***	2.56***
Non-homework reading, writing, or science activities	1.24	0.14
School activities (band, drama, etc.)	2.90***	2.14***
Lessons (music, art, dance, etc.)	2.15	0.87**
Organized sports	0.64	0.09
Clubs (Boy and Girl Scouts, Boys and Girls Club, etc.)	1.47	0.23
Activities at church, temple, mosque	0.77	0.52
Watched TV or videos	-1.78	-0.49
Surfed the Internet or did other things on a computer	0.98	1.49***
"Hung out" with friends	-0.07	0.61
Volunteered or did community service	1.72	0.68
Worked at a job	1.49	0.86**
Did chores around the house	-0.87	-0.38
Took care of a brother or sister	0.15	0.33
Sample Size^b	777	3,256

SOURCE: Student Survey, Parent Survey.

NOTE: All effects are scaled to represent the effect of an additional 10 days in the program. In each model, the regressor is attendance (fixed effects logit cannot include time invariant regressors). To estimate the fixed effects logit models, we use the *logit* command in STATA® and calculate the standard error of the marginal effect by bootstrapping. Because bootstrapping does not account for the stratified sampling design, standard errors may be underestimated. We use SUDAAN® to adjust for the stratified sampling design when calculating standard errors in the OLS regression models. Weights are used to adjust estimates for nonresponse.

^aStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least three days in a typical week.

^bSample sizes differ for some outcomes due to nonresponse. The sample sizes reported in the table are the medians for each column. In columns 1, the outcome with the smallest sample size is "Looking for work" with 223 and the outcome with the largest sample size is "Non-homework reading, writing, or science activities" with 1,271. In column 2, the outcome with the smallest sample size is "Not in the labor force" with 2,680 and the outcomes with the largest sample size are the location outcomes (Percentage of Students in the Following Locations After School at Least Three Days in a Typical Week) with 3,278.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table C.22

Sensitivity of Attendance Estimates to Estimation Technique For Student Safety, Negative Behavior, and Victimization, Middle School Centers, Year 2

Outcome	Fixed Effects Logit	Fixed Effects OLS
Percentage of Students Who Reported Feeling the Following Levels of Safety After School up Until 6:00 P.M.:		
Very safe	-1.39	-0.43
Somewhat safe	1.04	0.33
Not at all safe	2.97	0.10
Percentage of Students Who Reported That They Do the Following “Some” or “A Lot”:		
Break something on purpose	0.03	0.14
Punch or hit someone	-0.59	-0.09
Steal from a store	-0.09	0.05
Sell illegal drugs	2.52	0.10
Get arrested or detained by police	0.55	0.03
Percentage of Students Who Reported the Following Happened to Them “Some” or “A Lot”:		
Been offered, sold, or given an illegal drug	-1.14	0.40
Been “picked on” after school	1.07	-0.32
Been threatened or hurt with a weapon	0.19	-0.03
Been threatened by a gang or gang member	1.47	0.31
Had property damaged on purpose	-0.40	-0.52
Percentage of Students Who Reported That They Do the Following “Some” or “A Lot”:		
Smoke cigarettes	-4.79	-0.03
Have at least one alcoholic drink	-3.08	-0.40***
Smoke marijuana	-3.65	-0.12
Sample Size^a	345	3,251

SOURCE: Student Survey.

NOTE: All effects are scaled to represent the effect of an additional 10 days in the program. In each model, the regressor is attendance (fixed effects logit cannot include time invariant regressors). To estimate the fixed effects logit models, we use the *logit* command in STATA® and we calculate the standard error of the marginal effect by bootstrapping. Because bootstrapping does not account for the stratified sampling design, standard errors may be underestimated. We use SUDAAN® to adjust for the stratified sampling design when calculating standard errors in the OLS regression models. Weights are used to adjust estimates for nonresponse.

^aSample sizes may differ for some outcomes due to nonresponse. The sample sizes reported in the table are the medians for each column. In column 1, the outcome with the smallest sample size is “Sell illegal drugs” with 117 and the outcome with the largest sample size is “Very Safe” with 1,067. In column 2, the outcome with the smallest sample size is “Been threatened or hurt with a weapon” with 3,108 and the outcomes with the largest sample size are the Safety Outcomes (Percentage of Students Who Report Feeling the Following Levels of Safety After School Until 6:00 P.M.) with 3,273.

**Significantly different from zero at the .05 significance level, two-tailed test.

*** Significantly different from zero at the .01 significance level, two-tailed test.

APPENDIX D
SUBGROUP TABLES

Table D.1a

Impacts on Maternal Employment and Students' Location, Care, and Activities After School by Subgroup, Elementary School Centers, Year 1

Outcome	Estimated Impact						
	Grade Level			Baseline Test Scores ^a		Baseline Disciplinary Problems Composite ^b	
	K – 2	3 – 4	5 – 6	Low	High	Low	High
Percentage of Students in the Following Locations After School at Least 3 Days in a Typical Week (According to Parents):							
Own home	-19.7***	-21.1***	-1.8	-11.3**	-21.5***	-19.5***	-12.8*
Someone else's home	0.5	-5.3	6.0	-4.0	-3.2	-5.5	5.5
School or other place for activities	22.8***	20.0***	13.8	25.1***	19.6***	24.6***	15.8**
Somewhere to "hang out"	0.4	-1.6	-2.3	-1.6	-3.4**	-1.8	0.5
Mixed (no one location for at least 3 days)	0.7	0.4	-2.3	-0.7	-0.6	0.0	-1.2
Percentage of Students in the Following Types of Supervision After School at Least 3 Days in a Typical Week (According to Parents):							
Self-care ^c	-0.3	1.9	-3.1	2.6	-1.2	0.4	-0.9
Parent care	-10.0**	-12.8**	6.5	-1.6	-17.7***	-15.5***	7.1
Nonparent adult care	17.2***	13.3**	1.1	9.1	12.1**	14.3**	8.4
Sibling care	-4.7	-16.7***	9.3	-3.1	-9.5	-9.6	-5.6
Mixed care (no one type of care for at least 3 days)	0.8	-0.9	-1.4	-0.3	1.9	0.5	-4.6
Employment of Mother:							
Full time	2.3	-4.0	10.9	-1.9	4.0	6.7	8.1
Part time	0.1	4.8	0.5	0.1	0.2	-2.7	-1.0
Looking for work	6.5	0.3	-8.8	8.0**	-1.3	-0.2	-11.9**
Not in labor force	-9.0**	-1.2	-2.6	-6.3	-2.9	-3.8	4.8
Percentage of Students Who Participated in Each Activity at Least Once After School in the Past Week (According to Parents):							
Homework	-4.0	-6.8	-0.3	-4.7	-5.5	3.0	-9.1
Tutoring	19.0***	0.2	0.4	0.2	20.2***	-0.7	-2.5
Non-homework reading, writing, or science	-1.2	-0.4	-10.2	0.2	-5.1	-10.5	-8.7
Watched TV or videos	-5.3	-8.7	-1.0	-6.9	-1.5	-8.7	4.7
Surfed the Internet or did other things on a computer	8.0	-3.6	6.1	-7.3	6.0	-7.7	2.4
Hung out with friends	-6.3	2.8	0.0	-10.3	-1.1	-0.7	8.5
Mean Number of Hours Spent Watching TV in the Past Day (According to Students)							
	n.a. ^d	0.0	0.1	0.4**	-0.3	0.0	0.2
Mean Number of Hours Spent Reading for Fun in the Past Day (According to Students)							
	n.a. ^d	0.0	0.0	0.0	0.0	0.0	-0.1
Number of Observations:							
Student-reported outcomes	n.a. ^d	625	456	543	333	673	333
Parent-reported outcomes	704	481	310	660	519	536	254

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse.

^aStudents are defined as having low (high) scores if they scored below (above) the median reading test score for the study sample.^bThe baseline student discipline composite was based on students' responses to how frequently the following happened to them: (1) were sent to the office for doing something wrong, (2) have to miss recess or sit in the hall, and (3) parents had to come to school about a problem they are having. Students are defined as having low (high) levels of discipline problems if the composite falls below (above) the median of the composite for the study sample.^cStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least 3 days in a typical week.^dStudents in grades K-2 were not administered the student survey because of their age.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.1b

Impacts on Maternal Employment and Students' Location, Care, and Activities
After School by Subgroup, Elementary School Centers, Year 1

Outcome	Estimated Impact				
	Race/Ethnicity			Gender	
	White (Non-Hispanic)	Black (Non-Hispanic)	Hispanic	Male	Female
Percentage of Students in the Following Locations After School at Least 3 Days in a Typical Week (According to Parents):					
Own home	-29.5***	-17.8***	-15.3***	-18.4***	-16.4***
Someone else's home	-2.5	1.5	-6.0	-3.4	-3.1
School or other place for activities	25.5***	25.6***	25.6***	21.0***	19.1***
Somewhere to "hang out"	-3.1	3.1**	0.4	0.8	-1.3
Mixed (no one location for at least 3 days)	2.8	-0.6	1.5	1.1	-0.5
Percentage of Students Under the Following Types of Supervision After School at Least 3 Days in a Typical Week (According to Parents):					
Self-care ^a	n.a. ^b	0.4	0.4	1.1	-0.8
Parent care	-10.3	-11.8***	0.0	-7.7	-12.2***
Nonparent adult care	12.9	14.3***	5.3	9.7**	7.7
Sibling care	-6.5	-0.7	-11.4**	-4.9	-6.0
Mixed care (no one type of care for at least 3 days)	1.3	0.0	-1.4	0.6	1.3
Employment of Mother:					
Full time	12.0	-0.5	7.1	1.8	2.7
Part time	-7.7	3.7	-7.6	-0.9	-1.1
Looking for work	3.3	1.2	8.2	4.8	3.7
Not in labor force	-7.6	-4.4	-7.6	-5.6	-5.2
Percentage of Students Who Participated in Each Activity at Least Once After School in the Past Week (According to Parents):					
Homework	-2.9	-7.4**	1.0	-4.9	-3.8
Tutoring	-5.6	12.6***	6.7	8.7**	12.5***
Non-homework reading, writing, or science	-9.6	-5.3	6.2	0.1	-3.4
Watched TV or videos	-8.4	-6.1	-6.1	-7.2	-2.3
Surfed the Internet or did other things on a computer	30.1***	-2.0	-1.3	2.9	1.6
Hung out with friends	12.2	1.0	2.5	1.1	-0.6
Mean Number of Hours Spent Watching TV in the Past Day (According to Students)					
	-1.2***	0.2	0.2	0.2	0.0
Mean Number of Hours Spent Reading for Fun in the Past Day (According to Students)					
	-0.2***	0.0	-0.1	0.0	0.0
Number of Observations:					
Student-reported outcomes	58	474	273	464	548
Parent-reported outcomes	88	843	474	697	718

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse.

^aStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least 3 days in a typical week.

^bNo white students were reported to be in self-care.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.1c

Impacts on Maternal Employment and Students' Location, Care, and Activities
After School by Subgroup, Elementary School Centers, Year 1

Outcome	Household Structure ^a	
	Two Parent	One Parent
Percentage of Students in the Following Locations After School at Least 3 Days in a Typical Week (According to Parents):		
Own home	-18.5***	-16.9***
Someone else's home	-1.3	-1.3
School or other place for activities	17.0***	24.6***
Somewhere to "hang out"	-3.2**	2.9
Mixed (no one location for at least 3 days)	1.5	-0.2
Percentage of Students in the Following Types of Supervision After School at Least 3 Days in a Typical Week (According to Parents):		
Self-care ^b	-1.9	1.0
Parent care	-7.9	-8.6**
Nonparent adult care	9.1	11.6***
Sibling care	-8.1	2.1
Mixed care (no one type of care for at least 3 days)	0.5	1.0
Employment of Mother:		
Full time	3.8	-3.4
Part time	1.4	1.0
Looking for work	3.9	2.7
Not in labor force	-9.1**	-0.2
Percentage of Students Who Participated in Each Activity at Least Once After School in the Past Week (According to Parents):		
Homework	-1.3	-6.0**
Tutoring	9.6***	11.6***
Non-homework reading, writing, or science	-1.5	-5.3
Watched TV or videos	1.7	-5.1
Surfed the Internet or did other things on a computer	7.8	-0.7
Hung out with friends	6.2	-1.6
Mean Number of Hours Spent Watching TV in the Past Day (According to Students)	0.3	0.2
Mean Number of Hours Spent Reading for Fun in the Past Day (According to Students)	0.0	0.0
Number of Observations:		
Student-reported outcomes	396	437
Parent-reported outcomes	797	900

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse.

^aStudents are in the "two parent" subgroup if they live with a mother, stepmother, foster mother, or female guardian and a father, stepfather, foster father, or male guardian. If they do not live with both a male and female parent or guardian, students are in the "one parent" subgroup.

^bStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least 3 days in a typical week.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.2a

Impacts on Homework Completion, Level of Effort, and Classroom Behavior
by Subgroup, Elementary School Centers, Year 1

Outcome	Estimated Impact						
	Grade Level			Baseline Test Scores ^a		Baseline Disciplinary Problems Composite ^b	
	K – 2	3 – 4	5 – 6	Low	High	Low	High
Percentage of Students Whose Teachers Reported that They Often Complete Homework	-7.8	0.5	-12.2	0.9	-8.6	-5.5	-7.1
Percentage of Students Whose Teachers “Agree” or “Strongly Agree” that: Student completes assignments to my satisfaction	-5.6	2.5	-14.3	5.8	-4.8	-13.4**	6.2
Student comes prepared and ready to learn	-5.3	0.8	0.1	-2.8	-1.4	-6.6	3.3
Percentage of Students Whose Teachers Reported that They “Usually Try Hard” in Reading or English	4.6	6.8	-10.6	-2.4	5.1	3.4	0.8
Percentage of Students Whose Teachers Reported that They “Often” Perform at or above Their Ability	-5.0	6.4	1.0	-3.8	-0.3	2.9	3.9
Teacher-Reported Level of Effort Composite (Mean)	-0.1	0.1	-0.2	-0.1	0.0	0.0	0.0
Percentage of Students Whose Parents “Agree” or “Strongly Agree” that Child Works Hard at School	0.8	-5.9	-6.1	3.1	-2.4	-5.7	-17.0**
Percentage of Students Whose Teachers Reported Disciplining for Misbehaving “Two or More Times”	12.1**	-5.2	11.0	8.8	8.6	4.3	1.3
Percentage of Students Who Were Suspended	1.6	2.2	-0.4	4.1	1.3	-6.9**	4.9
Number of Observations:							
Parent-reported outcomes	791	525	333	652	516	529	252
Teacher-reported outcomes	862	541	346	641	595	564	277
School records outcomes (suspensions)	565	407	225	439	435	385	216

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse.

^aStudents are defined as having low (high) scores if they scored below (above) the median reading test score for the study sample.

^bThe baseline student discipline composite was based on students' responses to how frequently the following three things happened to them: (1) sent to the office for doing something wrong, (2) have to miss recess or sit in the hall, and (3) parents have to come to school about a problem they are having. Students are defined as having low (high) levels of discipline problems if the composite falls below (above) the median of the composite for the study sample.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.2b

Impacts on Homework Completion, Level of Effort, and Classroom Behavior
by Subgroup, Elementary School Centers, Year 1

Outcome	Estimated Impact				
	Race/Ethnicity			Gender	
	White (Non-Hispanic)	Black (Non-Hispanic)	Hispanic	Male	Female
Percentage of Students Whose Teachers Reported that They Often Complete Homework	4.4	-7.5	-1.3	-4.5	-7.6
Percentage of Students Whose Teachers “Agree” or “Strongly Agree” that:					
Student completes assignments to my satisfaction	5.7	-2.9	-8.2	-5.9	-1.6
Student comes prepared and ready to learn	15.1	-3.4	-8.2	-5.4	-2.5
Percentage of Students Whose Teachers Reported that They “Usually Try Hard” in Reading or English	8.2	-0.5	-2.0	3.6	2.2
Percentage of Students Whose Teachers Reported that They “Often” Perform at or above Their Ability	7.5	1.5	-3.7	-0.6	-1.2
Teacher-Reported Level of Effort Composite (Mean)	0.2	-0.1	0.0	-0.1	0.0
Percentage of Students Whose Parents “Agree” or “Strongly Agree” that Child Works Hard at School	-0.9	-1.3	-6.2	-2.0	-2.2
Percentage of Students Whose Teachers Reported Disciplining for Misbehaving “Two or More Times”	-6.7	8.8	4.9	3.2	6.3
Percentage of Students Who Were Suspended	7.7	3.3	-4.2**	0.4	3.5
Number of Observations:					
Parent-reported outcomes	125	902	548	765	803
Teacher-reported outcomes	95	739	464	820	846
School records outcomes (suspensions)	50	531	227	495	624

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.2c

Impacts on Homework Completion, Level of Effort, and Classroom Behavior
by Subgroup, Elementary School Centers, Year 1

Outcome	Household Structure ^a	
	Two Parent	One Parent
Percentage of Students Whose Teachers Reported that They Often Complete Homework	-8.6	-1.4
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" that: Student completes assignments to my satisfaction	-3.8	-1.2
Student comes prepared and ready to learn	-8.0	2.3
Percentage of Students Whose Teachers Reported that They "Usually Try Hard" in Reading or English	0.6	3.3
Percentage of Students Whose Teachers Reported that They "Often" Perform at or above Their Ability	-9.0	7.2
Teacher-Reported Level of Effort Composite (Mean)	-0.2	0.1
Percentage of Students Whose Parents "Agree" or "Strongly Agree" that Child Works Hard at School	-3.3	-0.4
Percentage of Students Whose Teachers Reported Disciplining for Misbehaving "Two or More Times"	9.5	4.1
Percentage of Students Who Were Suspended	1.2	0.2
Number of Observations:		
Parent-reported outcomes	794	891
Teacher-reported outcomes	647	750
School records outcomes (suspensions)	431	477

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse.

^aStudents are in the "two parent" subgroup if they live with a mother, stepmother, foster mother, or female guardian and a father, stepfather, foster father, or male guardian. If they do not live with both a male and female parent or guardian, students are in the "one parent" subgroup.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.3a

Impacts on Student Attendance and Academic Achievement by Subgroup,
Elementary School Centers, Year 1

Outcome	Estimated Impact						
	Grade Level			Baseline Test Scores ^a		Baseline Disciplinary Problems Composite ^b	
	K – 2	3 – 4	5 – 6	Low	High	Low	High
Mean Number of Days School Records Indicate Student Was:							
Absent	-0.5	-0.8	1.3	0.8	0.2	-0.3	-0.8
Late	0.2	0.2	0.3	0.3	-0.4	1.0	-1.3
Mean Student-Reported Reading Confidence Composite	n.a. ^c	0.1	-0.1	-0.1	0.0	0.0	0.0
Percentage of Students Whose Teachers Reported that They Achieve at an “Above Average” or “Very High” Level	-1.6	-5.8	-7.1	-4.5	-7.3	0.9	-7.4
Mean Class Grade:							
Math	0.0	1.7	-1.7	0.4	0.2	0.0	0.6
English	0.3	0.2	-1.4	0.9	-1.5	-0.6	-0.2
Science	0.1	1.4	-1.7	0.8	-0.6	0.1	0.9
Social Studies	-0.5	2.3**	-0.6	0.7	-1.2	0.9	-1.1
Mean Reading Test Score	-3.6	-3.3	-0.7	0.9	-2.1	-5.4	-0.8
Number of Observations:							
Student-reported outcomes	n.a. ^c	612	448	535	325	669	321
Teacher-reported outcomes	862	541	346	641	595	564	277
School records outcomes (attendance)	864	625	394	758	649	624	316
School records outcomes (grades)	666	490	367	647	553	501	261
School records outcomes (reading scores)	873	567	392	738	632	624	296

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse.

^aStudents are defined as having low (high) scores if they scored below (above) the median reading test score for the study sample.

^bThe baseline student discipline composite was based on students' responses to how frequently the following three things happened to them: (1) sent to the office for doing something wrong, (2) have to miss recess or sit in the hall, and (3) parents have to come to school about a problem they are having. Students are defined as having low (high) levels of discipline problems if the composite falls below (above) the median of the composite for the study sample.

^cStudents in grades K-2 were not administered the student survey because of their age.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.3b

Impacts on Student Attendance and Academic Achievement by Subgroup,
Elementary School Centers, Year 1

Outcome	Estimated Impact				
	Race/Ethnicity			Gender	
	White (Non-Hispanic)	Black (Non-Hispanic)	Hispanic	Male	Female
Mean Number of Days School Records Indicate Student Was:					
Absent	0.4	0.4	-1.8**	0.2	-0.3
Late	1.0	1.0	-2.4***	-1.3	1.5
Mean Student-Reported Reading Confidence Composite	-0.5**	0.0	0.1	0.0	0.1
Percentage of Students Whose Teachers Reported that They Achieve at an "Above Average" or "Very High" Level	-8.3	-1.4	7.2	-4.6	-3.6
Mean Class Grade:					
Math	2.6	0.8	-1.2	-0.1	0.1
English	3.8	-0.1	0.3	-0.2	0.0
Science	4.8	0.4	0.4	-0.5	0.4
Social Studies	4.0	1.7	-1.2	-0.5	1.1
Mean Reading Test Score	0.4	-0.5	-2.8	-2.1	2.9
Number of Observations:					
Student-reported outcomes	57	466	268	453	538
Teacher-reported outcomes	95	739	464	820	846
School records outcomes (attendance)	86	786	468	937	966
School records outcomes (grades)	49	679	420	736	771
School records outcomes (reading scores)	99	785	474	871	911

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.3c

Impacts on Student Attendance and Academic Achievement by Subgroup,
Elementary School Centers, Year 1

Outcome	Household Structure ^a	
	Two Parent	One Parent
Mean Number of Days School Records Indicate Student Was:		
Absent	-0.3	0.2
Late	0.2	0.9
Mean Student-Reported Reading Confidence Composite	-0.1	0.1
Percentage of Students Whose Teachers Report that They Achieve at an "Above Average" or "Very High" Level	-5.6	1.4
Mean Class Grade		
Math	0.5	-0.2
English	0.6	-0.4
Science	3.0**	-0.2
Social Studies	1.2	0.5
Mean Reading Test Score	-3.3	2.2
Number of Observations		
Student-reported outcomes	386	432
Teacher-reported outcomes	647	750
School records outcomes (attendance)	693	764
School records outcomes (grades)	568	646
School records outcomes (reading scores)	693	770

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse.

^aStudents are in the "two parent" subgroup if they live with a mother, stepmother, foster mother, or female guardian and a father, stepfather, foster father, or male guardian. If they do not live with both a male and female parent or guardian, students are in the "one parent" subgroup.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.4a

Impacts on Other Student and Parent Outcomes by Subgroup,
Elementary School Centers, Year 1

Outcome	Estimated Impact						
	Grade Level			Baseline Test Scores ^a		Baseline Disciplinary Problems Composite ^b	
	K – 2	3 – 4	5 – 6	Low	High	Low	High
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6 p.m.:							
Very safe	n.a. ^c	1.1	-5.4	2.9	-0.2	-1.6	-10.5
Somewhat safe	n.a. ^c	-0.6	7.9	0.9	-0.2	3.4	9.5
Not at all safe	n.a. ^c	-0.4	-2.5	-3.8	0.4	-1.8	1.0
Percentage of Students Who Reported Helping Another Student After School	n.a. ^c	10.9**	8.5	13.4**	-1.7	7.4	5.7
Student-Reported Disciplinary Problems Composite (Mean)	n.a. ^c	-0.1	0.2	0.0	0.0	0.0	0.1
Percentage of Students Whose Parents Reported that They Often Ask Student Things He or She Did in Class	7.5	4.3	9.4	7.4	0.6	-2.4	11.2
Percentage of Students Whose Parents Report Helping Them with Homework at Least Three Times Last Week	9.5**	3.4	-4.2	-4.0	-2.8	9.7	-6.7
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:							
Attended an open house at school	-0.9	-1.0	4.2	5.8	-2.8	-2.0	0.2
Attended a PTO meeting	4.3	6.9	7.6	12.7**	7.0	-1.4	3.2
Attended an after-school event	11.1**	10.7	16.7**	11.1**	12.7**	6.6	12.1
Volunteered to help out at school	-7.0	1.9	-7.2	-3.2	-3.5	-8.3	2.9
Number of Observations:							
Student-reported outcomes	n.a. ^c	625	454	538	332	674	330
Parent-reported outcomes	698	476	309	585	457	481	235

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse.

^aStudents are defined as having low (high) scores if they scored below (above) the median reading test score for the study sample.

^bThe baseline student discipline composite was based on students' responses to how frequently the following three things happened to them: (1) sent to the office for doing something wrong, (2) have to miss recess or sit in the hall, and (3) parents have to come to school about a problem they are having. Students are defined as having low (high) levels of discipline problems if the composite falls below (above) the median of the composite for the study sample.

^cStudents in grades K-2 were not administered the student survey because of their age.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.4b

Impacts on Other Student and Parent Outcomes by Subgroup,
Elementary School Centers, Year 1

Outcome	Estimated Impact				
	Race/Ethnicity			Gender	
	White (Non-Hispanic)	Black (Non-Hispanic)	Hispanic	Male	Female
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6 p.m.:					
Very safe	-3.3	0.9	-0.5	6.8	-9.9
Somewhat safe	3.3	-1.9	3.9	-3.9	11.4**
Not at all safe	0.0	1.0	-3.4	-2.9	-1.5
Percentage of Students Who Reported Helping Another Student After School	14.4	15.4***	0.5	16.2**	-3.4
Student-Reported Disciplinary Problems Composite (Mean)	0.0	0.1	-0.2	0.1	-0.1
Percentage of Students Whose Parents Reported that They Often Ask Student Things He or She Did in Class	-10.0	7.3	-6.5	9.8**	6.3
Percentage of Students Whose Parents Reported Helping Them with Homework at Least Three Times Last Week	13.4	2.2	8.0	11.7**	5.2
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:					
Attended an open house at school	-3.4	7.1	-3.4	5.3	-4.8
Attended a PTO meeting	17.7	0.9	6.5	1.6	6.6
Attended an after-school event	5.7	13.4***	6.4	12.1***	9.6**
Volunteered to help out at school	-3.0	-3.8	-0.6	-1.8	-2.4
Number of Observations:					
Student-reported outcomes	58	473	271	463	545
Parent-reported outcomes	86	838	465	690	711

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.4c

Impacts on Other Student and Parent Outcomes by Subgroup,
Elementary School Centers, Year 1

Outcome	Household Structure ^a	
	Two Parent	One Parent
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6 p.m.:		
Very safe	7.5	-11.3
Somewhat safe	-5.9	12.3**
Not at all safe	-1.5	-1.1
Percentage of Students Who Reported Helping Another Student After School	12.0	7.5
Student-Reported Disciplinary Problems Composite (Mean)	0.1	-0.1
Percentage of Students Whose Parents Reported that They Often Ask Student Things He or She Did in Class	1.9	10.9**
Percentage of Students Whose Parents Reported Helping Them with Homework at Least Three Times Last Week	9.7	6.6
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:		
Attended an open house at school	0.4	-3.4
Attended a PTO meeting	5.3	-0.1
Attended an after-school event	13.8***	3.7
Volunteered to help out at school	8.0	-12.6***
Number of Observations:		
Student-reported outcomes	393	438
Parent-reported outcomes	710	804

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup impacts reported in bold indicate that the estimated impact for one subgroup differed significantly from the estimated subgroup impact for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse.

^aStudents are in the "two parent" subgroup if they live with a mother, stepmother, foster mother, or female guardian and a father, stepfather, foster father, or male guardian. If they do not live with both a male and female parent or guardian, students are in the "one parent" subgroup.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.5a

Outcome Differences on Maternal Employment and Students' Location, Care, and Activities After School by Subgroup, Middle School Centers, Year 2

Outcome	Difference						
	Grade Level		Race/Ethnicity			Gender	
	5 – 6	7 – 8	White (Non-Hispanic)	Black (Non-Hispanic)	Hispanic	Female	Male
Percentage of Students Who Reported Being in the Following Locations After School at Least 3 Days in a Typical Week:							
Own home	-0.1	-3.0	-6.9***	0.6	1.3	-3.3	-1.1
Someone else's home	3.4	-0.2	-0.6	1.0	-1.7	2.8	-1.9
School or other place for activities	5.2	4.3***	0.4	5.0	5.5	4.2**	4.9
Somewhere to "hang out"	0.4	3.5	1.0	6.2***	1.0	3.4**	2.1
Mixed (no one location for at least 3 days)	0.4	0.4	5.6***	-0.5	-2.1	1.3	-0.6
Percentage of Students Who Reported Being in the Following Types of Supervision After School at Least 3 Days in a Typical Week:							
Self-care ^a	-1.0	-0.7	3.0	-0.8	-0.9	-1.3	-0.5
Parent care	3.3	-3.5	-6.5**	-5.3	4.0	-2.3	-1.4
Nonparent adult care	1.6	6.2**	2.6	2.6	4.4	4.5	5.9
Sibling care	-6.5***	-2.9	-4.0	1.1	-8.6***	-3.3	-4.2**
Mixed care (no one category for at least 3 days)	-1.6	-1.3	0.2	-1.8	-1.9	-0.6	-2.2
Employment of Mother (Parent-Reported):							
Full time	-0.8	-3.4	-5.5	2.4	-2.9	-1.3	-4.6**
Part time	2.9	2.1	1.8	3.2	1.6	2.3	2.4
Looking for work	-1.1	-0.3	1.6	-1.3	-1.7	-1.9	1.2
Not in the labor force	-1.0	1.6	2.1	-4.3	2.9	0.9	1.0
Percentage of Students Who Reported Participating in the Following Activities After School:							
Homework	-3.0	-1.8	0.6	-6.3	-2.7	-2.4	-1.7
Tutoring	1.9	3.9	2.8	13.1***	-3.0	4.2	2.9
Non-homework reading, writing, or science	6.4	1.2	1.6	7.6**	3.5	4.4	0.3
Watched TV or videos	-2.2	2.5	3.8	4.9	-2.6	1.4	1.5
Surfed the Internet or did other things on a computer	3.4	-1.2	-3.9	5.7	1.4	0.0	-0.5
Hung out with friends	6.0***	3.5***	3.6	7.7***	2.3	5.6***	2.3
Mean Number of Hours Spent Watching TV in the Past Day (According to Students)	-0.05	0.03	0.0	0.0	0.0	0.02	-0.01
Mean Number of Hours Spent Reading for Fun in the Past Day (According to Students)	-0.01	0.02	0.0	0.1	0.0	0.03	0.00
Number of Observations:							
Student-reported outcomes	1,080	2,725	1,334	909	1,020	2,041	1,763
Parent-reported outcomes	963	2,425	1,216	770	933	1,824	1,563

SOURCE: Parent Survey, Student Survey.

NOTE: Subgroup estimates reported in bold indicate that the estimated outcome difference for one subgroup differed significantly from the estimated outcome difference for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN© to account for the statistical sampling design.

^aStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least 3 days in a typical week.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.5b

Outcome Differences on Maternal Employment and Students' Location, Care, and Activities
After School by Subgroup, Middle School Centers, Year 2

Outcome	Difference					
	Baseline Grades ^a		Baseline Disciplinary Problems Composite ^b		Household Structure ^c	
	Low	High	Low	High	Two Parent	One Parent
Percentage of Students Who Reported Being in the Following Locations After School at Least 3 Days in a Typical Week:						
Own home	-3.0	-2.0	-3.6	-0.2	-1.5	-5.8***
Someone else's home	-1.1	1.7	1.3	-0.7	0.9	2.2
School or other place for activities	6.6***	2.7	4.9	4.0	4.5**	4.5
Somewhere to "hang out"	3.5	1.9	3.0	2.5	2.5	5.4**
Mixed (no one location for at least 3 days)	0.2	0.6	0.3	1.0	-0.8	4.2***
Percentage of Students Who Reported Being in the Following Types of Supervision After School at Least 3 Days in a Typical Week:						
Self-care ^d	-3.5	1.2	-1.4	0.5	0.2	1.3
Parent care	2.5	-3.8	-1.4	-3.0	0.9	-7.5***
Nonparent adult care	5.3	4.7	4.8	5.2	1.7	5.4
Sibling care	-1.7	-5.8***	-5.0**	-1.4	-5.8**	-2.4
Mixed care (no one category for at least 3 days)	-1.1	-1.8	-1.0	-1.9	-0.8	-1.3
Employment of Mother (Parent-Reported):						
Full time	-4.4	-2.9	-1.4	-7.5***	-2.2	-1.4
Part time	6.1**	0.9	2.2	2.5	1.6	0.3
Looking for work	0.6	-0.5	-2.0	3.7	-0.9	-0.2
Not in the labor force	-2.3	2.6	1.2	1.3	1.5	1.3
Percentage of Students Who Reported Participating in the Following Activities After School:						
Homework	-3.8	-2.0	-2.9**	-0.2	-2.2	-3.5
Tutoring	1.0	4.1**	2.1	5.3	2.2	5.3
Non-homework reading, writing, or science	2.5	1.1	3.0	2.1	1.4	6.4
Watched TV or videos	0.8	1.7	2.8	-1.2	1.7	2.8
Surfed the Internet or did other things on a computer	0.6	-1.1	-0.2	0.0	2.3	-2.2
Hung out with friends	3.4	3.9***	6.6***	-1.4	4.1***	5.7***
Mean Number of Hours Spent Watching TV in the Past Day (According to Students)	-0.05	0.05	0.07	-0.12	-0.04	0.07
Mean Number of Hours Spent Reading for Fun in the Past Day (According to Students)	0.02	0.02	0.02	0.01	0.00	0.06
Number of Observations:						
Student-reported outcomes	1,130	2,593	2,497	1,212	2,149	1,328
Parent-reported outcomes	994	2,328	2,217	1,079	2,007	1,197

SOURCE: Parent Survey, Student Survey.

NOTE: Subgroup estimates reported in bold indicate that the estimated outcome difference for one subgroup differed significantly from the estimated outcome difference for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN© to account for the statistical sampling design.

^aStudents are defined as having low baseline grades if they reported average grades of C, D, or F; students are defined as having high baseline grades if they reported average grades of A or B.

^bThe student-based discipline problem composite is based on four items: the extent to which students report (1) skipping school or class, (2) getting sent to the office for doing something wrong, (3) getting detention, and (4) having their parents called to school about a problem they are having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems. Students are defined as having high (low) levels of discipline problems if the composite falls above (below) the median of the composite for the study sample.

^cStudents are in the "two parent" subgroup if they live with a mother, stepmother, foster mother, or female guardian and a father, stepfather, foster father, or male guardian. If they do not live with both a male and female parent or guardian, students are in the "one parent" subgroup.

^dStudents are defined as being in self-care if they were not with a parent, a nonparent adult, or an older sibling at least 3 days in a typical week.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.6a

Outcome Differences on Homework Completion, Level of Effort, and Classroom Behavior by Subgroup, Middle School Centers, Year 2

Outcome	Difference						
	Grade Level		Race/Ethnicity			Gender	
	5 – 6	7 – 8	White (Non-Hispanic)	Black (Non-Hispanic)	Hispanic	Female	Male
Percentage of Students Whose Teachers “Agree” or “Strongly Agree” that: Student completes assignments to my satisfaction	3.4	-4.0	-5.1**	-5.2	2.0	-0.6	-4.1
Student comes prepared and ready to learn	-0.5	0.2	-4.0	1.3	2.6	-1.2	1.3
Percentage of Students Whose Teachers “Agree” or “Strongly Agree” that: The student is attentive in class	2.5	-2.1	-2.4	0.3	-2.3	-1.3	-0.6
The student participates in class	3.2	-1.0	-3.3	1.8	4.8	0.0	0.3
Percentage of Students Whose Teachers Reported that They “Usually Try Hard” in Class	1.7	0.4	0.5	3.3	-1.8	1.0	0.3
Percentage of Students Whose Teachers Reported that They “Often” Perform at or above Their Ability Level	-1.2	-2.8	-7.9**	0.6	-1.1	-1.8	-3.4
Teacher-Reported Level of Effort Composite (Mean)	0.04	0.00	-0.1	0.0	0.1	0.00	0.02
Teacher-Reported Disciplinary Problems Composite (Mean)	-0.02	0.02	0.0	0.1	0.0	0.02	-0.01
Mean Number of Days School Records Indicate Student Was: Absent	-1.11	-0.99**	-1.1**	-1.6	-0.7	-1.09**	-0.96**
Late	0.64	0.91	0.5	1.8	0.6	0.76	0.92
Percentage of Students Whose Teachers Reported They Achieve at an “Above Average” or “Very High” Level	-1.6	-3.2	-6.0	0.5	-2.2	-1.2	-4.7**
Mean Class Grade: Math	1.1	0.5	1.0	0.7	0.7	1.1	0.2
English	0.6	0.4	0.3	0.6	0.7	0.9	-0.1
Science	1.0	0.4	0.3	0.7	0.6	1.0	0.1
Social Studies	0.6	2.0***	1.5	3.2***	0.7	1.5**	1.9**
Number of Observations: Teacher-reported outcomes	1,082	2,560	1,288	881	941	1,947	1,693
School records outcomes (attendance)	1,060	2,728	1,324	899	1,016	2,016	1,771
School records outcomes (grades)	1,043	2,600	1,300	863	971	1,936	1,696

SOURCE: Teacher Survey, School Records.

NOTE: Subgroup estimates reported in bold indicate that the estimated outcome difference for one subgroup differed significantly from the estimated outcome difference for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN© to account for the statistical sampling design.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.6b

Outcome Differences on Homework Completion, Level of Effort, and Classroom Behavior by Subgroup, Middle School Centers, Year 2

Outcome	Difference					
	Baseline Grades ^a		Baseline Disciplinary Problems Composite ^b		Household Structure ^c	
	Low	High	Low	High	Two Parent	One Parent
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" that: Student completes assignments to my satisfaction	-1.8	-2.2	-0.5	-3.6	0.8	-5.0**
Student comes prepared and ready to learn	1.6	-0.7	1.5	-1.3	0.7	0.7
Percentage of Students Whose Teachers "Agree" or "Strongly Agree" that: Student is attentive in class	-0.8	-0.5	1.9	-5.7	0.2	-3.0
Student participates in class	0.8	-0.1	1.8	-2.1	1.5	-2.3
Percentage of Students Whose Teachers Reported that They "Usually Try Hard" in Class	2.2	0.7	1.8	-1.5	1.4	0.0
Percentage of Students Whose Teachers Reported that They "Often" Perform at or above Their Ability Level	0.6	-4.1	-3.2	-0.5	-1.3	-3.1
Teacher-Reported Level of Effort Composite (Mean)	0.06	-0.01	0.04	-0.04	0.03	-0.04
Teacher-Reported Disciplinary Problems Composite (Mean)	0.05	-0.02	-0.01	0.04	-0.04	0.03
Mean Number of Days School Records Indicate Student Was:						
Absent	-0.93	-1.04**	-0.93***	-1.13	-1.23***	0.02
Late	2.06**	0.21	0.85	0.79	0.47	1.20
Percentage of Students Whose Teachers Reported They Achieve at an "Above Average" or "Very High" Level	1.7	-4.6	-2.3	-4.4	-2.9	-2.7
Mean Class Grade:						
Math	0.9	0.5	1.5***	-0.8	1.0	0.7
English	1.6**	-0.1	1.0	-0.8	1.3***	-0.5
Science	0.9	0.3	0.9	-0.1	0.9	-0.2
Social Studies	3.2***	0.9	2.2***	0.8	1.8***	1.4
Number of Observations:						
Teacher-reported outcomes	1,090	2,470	2,374	1,167	1,917	1,188
School records outcomes (attendance)	1,117	2,587	2,468	1,227	2,019	1,220
School records outcomes (grades)	1,067	2,507	2,403	1,156	1,959	1,164

SOURCE: Teacher Survey, School Records.

NOTE: Subgroup estimates reported in bold indicate that the estimated outcome difference for one subgroup differed significantly from the estimated outcome difference for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN© to account for the statistical sampling design.

^aStudents are defined as having low baseline grades if they reported average grades of C, D, or F; students are defined as having high baseline grades if they reported average grades of A or B.

^bThe student-based discipline problem composite is based on four items: the extent to which students report (1) skipping school or class, (2) getting sent to the office for doing something wrong, (3) getting detention, and (4) having their parents called to school about a problem they are having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems. Students are defined as having high (low) levels of discipline problems if the composite falls above (below) the median of the composite for the study sample.

^cStudents are in the "two parent" subgroup if they live with a mother, stepmother, foster mother, or female guardian and a father, stepfather, foster father, or male guardian. If they do not live with both a male and female parent or guardian, students are in the "one parent" subgroup.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.7a

Outcome Differences on Other Student and Parent Outcomes
by Subgroup, Middle School Centers, Year 2

Outcome	Difference						
	Grade Level		Race/Ethnicity			Gender	
	5 – 6	7 – 8	White (Non-Hispanic)	Black (Non-Hispanic)	Hispanic	Female	Male
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6:00 P.M.:							
Very safe	1.5	-3.5	-5.3**	1.8	-1.9	-3.1	-1.5
Somewhat safe	-0.4	2.7	3.9	-2.7	2.7	3.2	0.7
Not at all safe	-1.1	0.8**	1.4**	0.8	-0.7	-0.1	0.8
Student-Reported Delinquent Behavior Composite (Mean)	0.05	0.01	0.0	0.1	0.0	0.03	0.01
Percentage of Students Who Rated Themselves as Skilled in Working out Conflicts with Others	-4.3	-2.5	-2.0	-8.0**	1.4	-3.8	-2.2
Student Educational Expectations (Percentages):							
Graduate from college	1.9	2.7**	-2.0	1.3	3.5	3.7**	1.3
Graduate from high school	-1.1	-2.5	1.6	-0.5	-3.0	-3.6	-0.6
Drop out of high school	-0.8	-0.3	0.4	-0.7	-0.5	-0.1	-0.7
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:							
Attended an open house at school	-5.0	2.3	0.4	1.9	-1.0	-0.9	2.2
Attended a PTO meeting	4.1	0.4	0.0	4.4	-1.3	0.0	2.7
Attended an after-school event	1.3	1.6	-1.0	3.6	2.0	0.3	3.3
Volunteered to help out at school	5.1**	1.1	1.3	0.6	5.5	2.2	2.1
Percentage of Students Who Reported the Following Happened to Them "Some" or "A lot"							
Been offered, sold, or given an illegal drug	1.1	-1.6	1.3	-2.9	-3.2	-3.1**	1.3
Been picked on after school	-2.0	4.6**	3.4	5.4	1.8	1.7	4.3
Been threatened or hurt with a weapon	1.0	1.0	1.3	1.2	0.8	0.8	1.4
Been threatened by a gang member	-2.8	1.0	0.5	0.9	-1.1	0.2	0.1
Had your property damaged on purpose	2.2	2.3	-0.1	8.7***	-1.8	1.9	2.7
Percentage of Students Who Reported that They Do the Following "Some" or "A lot"							
Break something on purpose	2.7	2.1	2.8	2.8	2.1	2.5	1.9
Punch/hit someone	2.4	3.0	4.0**	5.9	-1.0	3.7	1.8
Sell illegal drugs	-1.5**	0.0	-0.1	-1.0	-0.2	0.0	-0.7
Get arrested	0.7	-0.1	0.4	-1.1	0.3	-0.4	0.7
Percentage of Students Who Reported that They Do the Following "Some" or "A lot":							
Smoke cigarettes	0.6	0.7	1.4	1.4	-1.0	0.5	0.8
Smoke marijuana	0.4	0.4	0.3	1.5	-0.2	0.9	-0.1
Drink alcohol	2.8	0.1	1.3	3.1	-0.2	2.4	-1.1
Student-Reported Tobacco, Alcohol, Drug Use Composite (Mean)	0.02	0.01	0.0	0.0**	0.0	0.02**	0.01
Number of Observations:							
Student-reported outcomes	1,087	2,728	1,341	910	1,019	2,044	1,770
Parent-reported outcomes	967	2,436	1,226	783	927	1,832	1,570

SOURCE: Parent Survey, Student Survey.

NOTE: Subgroup estimates reported in bold indicate that the estimated outcome difference for one subgroup differed significantly from the estimated outcome difference for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN© to account for the statistical sampling design.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.

Table D.7b

Outcome Differences on Other Student and Parent Outcomes
by Subgroup, Middle School Centers, Year 2

Outcome	Difference					
	Baseline Grades ^a		Baseline Disciplinary Problems Composite ^b		Household Structure ^c	
	Low	High	Low	High	Two Parent	One Parent
Percentage of Students Who Reported Feeling the Following Levels of Safety After School Until 6:00 P.M.:						
Very safe	-1.4	-4.1	-2.8	-1.3	-4.1	-2.7
Somewhat safe	1.0	3.5	2.5	0.4	3.4	1.7
Not safe at all	0.4	0.6	0.3	0.9	0.7	1.0
Student-Reported Delinquent Behavior Composite (Mean)	0.02	0.02	0.01	0.03	0.00	0.04**
Percentage of Students Who Rated Themselves as Skilled in Working out Conflicts with Others	0.2	-4.8	-2.5	-2.7	-2.8	-2.9
Student Educational Expectations (Percentages):						
Graduate from college	4.2	1.5	1.7	5.0	2.7	2.8
Graduate from high school	-4.2	-1.1	-1.0	-4.8	-2.6	-2.5
Drop out of high school	0.0	-0.4	-0.6	-0.2	-0.1	-0.2
Percentage of Students Whose Parents Did the Following at Least Three Times Last Year:						
Attended an open house at school	-1.9	1.3	2.3	-3.8	2.8	-0.1
Attended a PTO meeting	1.2	2.0	1.6	-1.0	1.3	4.5
Attended an after-school event	3.0	1.0	1.0	3.4	2.3	-0.2
Volunteered to help out at school	2.3	2.2	3.3	0.0	3.3	-1.1
Percentage of Students Who Reported the Following Happened to Them "Some" or "A lot":						
Been offered, sold, or given an illegal drug	-6.2***	1.1	-1.1	0.2	-0.6	-2.0
Been picked on after school	4.9	2.0	0.9	6.3	2.0	8.0***
Been threatened or hurt with a weapon	2.9	0.3	0.9	1.4	-0.2	2.6
Been threatened by a gang member	0.3	-0.1	-0.3	0.9	0.4	1.8
Had your property damaged on purpose	2.4	1.6	0.3	6.2**	0.5	6.4***
Percentage of Students Who Reported that They Do the Following "Some" or "A lot":						
Break something on purpose	2.9**	2.1	1.5	3.3**	1.1	2.0
Punch/hit someone	3.1	3.4**	3.9***	0.8	3.6**	1.8
Sell illegal drugs	-0.6	-0.2	-0.2	-0.5	-1.1***	0.2
Get arrested	0.4	0.1	0.0	0.5	0.4	1.3
Percentage of Students Who Reported that They Do the Following "Some" or "A lot":						
Smoke cigarettes	1.4**	0.2	0.6	0.6	0.1	0.3
Smoke marijuana	-2.0	1.7***	0.2	1.0	0.6	0.4
Drink alcohol	-2.6	2.5	1.9	-0.7	0.1	-0.6
Student-Reported Tobacco, Alcohol, Drug Use Composite (Mean)	-0.02	0.03**	0.02	0.01	0.01	0.01
Number of Observations:						
Student-reported outcomes	1,133	2,600	2,499	1,220	2,154	1,331
Parent-reported outcomes	1,005	2,330	2,226	1,085	2,000	1,213

Table D.7b (continued)

SOURCE: Parent Survey, Student Follow-up Survey.

NOTE: Subgroup estimates reported in bold indicate that the estimated outcome difference for one subgroup differed significantly from the estimated outcome difference for the other related subgroup(s) at the .05 level or higher. Weights are used to adjust estimates for nonresponse. Variances are estimated using SUDAAN© to account for the statistical sampling design.

^aStudents are defined as having low baseline grades if they reported average grades of C, D, or F; students are defined as having high baseline grades if they reported average grades of A or B.

^bThe student-based discipline problem composite is based on four items: the extent to which students report (1) skipping school or class, (2) getting sent to the office for doing something wrong, (3) getting detention, and (4) having their parents called to school about a problem they are having. The composite is equal to the mean of the four variables. A value of 1 on the composite indicates infrequent discipline problems, while a value of 4 indicates frequent discipline problems. Students are defined as having high (low) levels of discipline problems if the composite falls above (below) the median of the composite for the study sample.

^cStudents are in the "two parent" subgroup if they live with a mother, stepmother, foster mother, or female guardian and a father, stepfather, foster father, or male guardian. If they do not live with both a male and female parent or guardian, students are in the "one parent" subgroup.

**Significantly different from zero at the .05 significance level, two-tailed test.

***Significantly different from zero at the .01 significance level, two-tailed test.



NCEE 2004-3001