

# **Boston Turnaround Using Increased Learning Time: Year 3 Student Outcomes Analysis**

**March 2016**

Martyna Citkowicz, Ph.D.

Erin Haynes, Ph.D.

**STATE & LOCAL  
EVALUATION** Center

at American Institutes for Research ■

20 North Wacker Drive, Suite 1231  
Chicago, IL 60606-2901  
312-288-7600  
[www.air.org](http://www.air.org)

Copyright © 2016 American Institutes for Research. All rights reserved.



[www.air.org](http://www.air.org)

5368\_03/16

# Contents

	<b>Page</b>
Summary.....	1
Methods.....	3
Results.....	5
Student Achievement.....	5
Attendance.....	8
Perception of School Environment.....	9
Student Behavior—Suspensions.....	12
Conclusion.....	14
References.....	15
Appendix A. CITS Technical Approach.....	16
Appendix B. Overall CITS Outcomes.....	24
Appendix C. CITS Outcomes by Grade.....	32
Appendix D. CITS Outcomes for Special Populations.....	40

## Summary

From 2012 to 2015, Boston Public Schools implemented an expanded learning time program called Turnaround Using Increased Learning Time (TILT) in two of its low-performing middle schools. TILT aimed to improve educational outcomes for students by increasing the amount of time students had for learning and teachers had for teaching. The primary components of TILT included the following: (1) additional time for core academic subjects, (2) additional time for enrichment activities (e.g., physical education and service learning), and (3) additional time for teachers to collaborate, plan, and engage in professional development.

Serving as the program's independent evaluator, American Institutes for Research (AIR) used a comparative interrupted time series (CITS) analysis<sup>1</sup> to determine the impact of the TILT program on student outcomes. The analysis compared the two middle schools that implemented TILT with two middle schools that were similar on various student- and school-level characteristics but did not implement the program. Analyzed impact variables included students' achievement, attendance, behavior, and perception of their school environment and learning engagement.

AIR also conducted CITS analyses of subgroup populations to examine whether the program had an impact on three at-risk groups of students. The three subgroups included students with English language learner (ELL) status, with free or reduced-price lunch eligibility, and with special education (SPED) status.

How effective was the TILT program? Overall, the program had mixed results that appeared to improve somewhat during the three years of implementation, driven by Grade 8 achievement. In the first year, impacts were minimal but primarily negative, with negative effects on mathematics achievement for the general population and other negative or nonsignificant effects for subpopulations. Year 2 saw a slight improvement, with significant positive effects for both mathematics and English language arts (ELA) achievement, as well as student engagement. In Year 3, the positive effect for ELA achievement continued, although no effect was found for mathematics achievement and a negative effect was found for school environment. In addition, no effect was found for science achievement or attendance during all three years of program implementation. Summaries for each year follow:

- **Year 1:** After the first year of the program, AIR found a significant but small negative effect of TILT on the mathematics state achievement test. TILT also was associated with an increase in the achievement gap on the ELA state test between students who received free or reduced-price lunch and those who did not receive these services in the Boston TILT program, in contrast to no change in the achievement gap between these students in the comparison schools. There also was an increase in the gap on student ratings of positive school environment between students receiving special education services and those who did not, relative to the change in the gap between these students in the comparison schools.

---

<sup>1</sup> CITS is an analytical approach that compares outcomes between two groups, presumably similar with regard to characteristics related to the outcome, over time, considering performance before and after the introduction of a new program.

- **Year 2:** After the second year of implementation, AIR found overall higher student ratings on engaged learning in Boston TILT schools than in comparison schools. AIR also found positive effects of TILT on both the ELA and mathematics state achievement tests. This finding is consistent with qualitative findings that there was improvement at both schools in Year 2 on how Academic League (i.e., the intervention) instruction is targeted, the quality of extended-day instruction and classroom management, and the structure of teachers' collaborative time. For some subgroups, however, negative effects were found after the second year of program implementation. The program was associated with a greater increase in the achievement gap between students receiving special education services and students who did not on the mathematics state achievement test than the achievement gap between these students in the comparison schools. An increase in the gap between student ratings of positive school environment between ELL students and non-ELL students also was found, relative to the change in the gap between these students in the comparison schools.
- **Year 3:** After the third year of implementation, AIR found a significant positive effect of TILT on the ELA state achievement test. AIR also found an overall greater decrease in student ratings on positive school environment in Boston TILT schools than in comparison schools. No subgroup effects were found in Year 3.

Finally, AIR found an increase in the number of suspension days in the Boston TILT schools, decreasing the gap in suspension rates between Boston TILT schools and comparison schools after the first, second, and third years of program implementation. No evidence was found, however, of any lower attendance figures in the Boston TILT schools than in comparison schools, which in an extended day program means that TILT students attended more hours of school. No evidence was found of any difference in science scores between the Boston TILT schools and comparison schools as a result of program implementation.

Findings from Years 1 and 2 of the study are detailed in Tucker, Hallberg, and Therriault (2014) and Citkowicz, Haynes, Hallberg, and Therriault (2014). This report discusses the Year 3 findings from the CITS analysis in more detail.

Qualitative analyses across all three years of the study (Haynes, Kistner, Therriault, & Hallberg, 2015) found that the TILT program offered students from a range of backgrounds access to academic instruction and extracurricular activities in a protected, supervised environment. These qualitative findings, coupled with the quantitative results, paint a cautiously optimistic picture for the success of the TILT program.

## Methods

AIR used a CITS design to measure impacts of the Boston TILT program on student outcomes. CITS is one of the strongest quasi-experimental designs that can be used when a comparison or control series can be constructed (Shadish, Cook, & Campbell, 2001). This method compares the outcomes of a treatment group and comparison group after a treatment occurs relative to their baseline trends prior to program implementation to determine program impact. In doing so, this method measures program impact by comparing changes in the treatment group outcomes after program implementation with the changes in outcomes one would expect to find in the same time period in the absence of the program. The CITS analysis for the Boston TILT program compares the changes in the outcomes of students in the Boston TILT schools after the first (2013), second (2014), and third (2015) years of program implementation in comparison with changes in outcomes of students in comparison schools.

To strengthen the internal validity of causal findings and to attribute any positive effects to the intervention with more confidence, comparison schools were selected prior to program implementation on the basis of their similarities to control schools in location, student demographics, and achievement.

Following is a brief description of all variables included in the models. See Appendix A for a detailed description of the CITS model.

The analysis included four outcome measures: (1) student achievement, (2) student attendance, (3) perception of school climate, and (4) behavior as measured by frequency of suspensions. See Tables A1–A6 in Appendix A for the distribution of the outcome measures.

- *Student achievement.* Student raw scores on the Massachusetts Comprehensive Assessment System (MCAS) in ELA, mathematics, and science were the primary measures of student achievement in the 2009–10, 2010–11, 2011–12, 2012–13, and 2013–14 school years. In the 2014–15 school year, BPS adopted the Partnership for Assessment of Readiness for College and Careers (PARCC) tests for the ELA and mathematics portions of the state achievement examinations. The MCAS raw scores and PARCC scaled scores were standardized within grade, year, and subject.<sup>2</sup> Science scores were available only for Grade 8.
- *Student attendance.* Student attendance rate was calculated as the total number of days a student attended school out of the total number of days the student was enrolled at the school. Students who attended 30 days or fewer were removed from the entire analysis.
- *Perception of school climate.* In the 2010–11, 2011–12, 2012–13, 2013–14, and 2014–15 school years, students were asked to complete school climate surveys (surveys were not administered during the first year of baseline data). Scale scores were calculated for two constructs from the survey: (1) positive school environment and (2) student engagement in learning.

---

<sup>2</sup> Because standardizing the data ensures that all the data is on the same norm-referenced metric, we believe that assessing the MCAS raw scores alongside PARCC scaled scores is a reasonable comparison.

- *Student behavior.* The student behavior measure included the total number of days a student was suspended during the school year.

In addition to the overall model, AIR conducted subgroup analyses of three populations to test whether the Boston TILT program had an impact on students most vulnerable for academic success. These populations were identified as ELLs, those eligible for free or reduced-price lunch, and those enrolled in SPED. Each subgroup model analyzed whether there was a difference in the outcome trends in the differences between students in that subgroup and those not in that subgroup (e.g., an increase or decrease in the achievement gaps) for students in the Boston TILT program in comparison with students in the comparison schools. AIR was unable to conduct analyses for the special populations for the suspension outcome because of the low frequency of suspensions for students within each of these subgroups.

Finally, the ELA and mathematics achievement models included a student's examination score in the prior year,<sup>3</sup> and most models included an indicator for a student's race, grade, and special student population classification (i.e., ELL, free or reduced-price lunch, SPED). The decision to include students' achievement scores from the prior year in the model was the subject of considerable debate, for the prior score represents a posttreatment variable for Grade 8 students in the second cohort. In the end, we found the inclusion of prior year's scores necessary in order to account for imbalance in the Boston TILT schools and comparison schools. As a result, each year's estimated program effect is not cumulative and, instead, represents solely that year's Boston TILT program effect.

See Tables A1–A8 in Appendix A for the distribution of variables across time and between the Boston TILT and comparison schools.

---

<sup>3</sup> We did not include a student's prior examination score in the science achievement models because the increase in effect size precision would be negligible over the three-year gap in which students took the examinations (Grades 5 and 8) and not worth the number of students we would lose from the sample from student mobility and attrition.

## Results

This section describes the overall and subgroup analyses for each student outcome.

### Student Achievement

Tables B1 and B2 in Appendix B display the overall results of the Boston TILT program on student achievement in ELA, mathematics, and science (Grade 8 only). Tables C1 and C2 in Appendix C display the results of the Boston TILT program on student achievement in ELA and mathematics for the three grades separately—Grades 6, 7, and 8. The results are presented in effect sizes because the students' scores were standardized in order to make the effects comparable across grade, year, and subject.<sup>4</sup> There is no absolute scale for measuring whether an effect is “big” or “small,” but an effect size can be compared with average impacts of other interventions. For example, the average effect size calculated for rigorous studies in middle schools is about 0.15, and the average effect size shown by whole-school reforms is 0.11. Effect sizes also may be judged against average gains that students are expected to make each year. Average expected gains from Grades 5 to 6, 6 to 7, and 7 to 8 are 0.32, 0.23, and 0.26 effect sizes, respectively, for reading and 0.41, 0.30, and 0.32 effect sizes, respectively, for mathematics (Lipsey et al., 2012).

The results suggest that students in the Boston TILT program performed significantly worse on the mathematics achievement test than students in comparison schools after the first year of program implementation ( $\beta_5 = -0.08$ ,  $p < 0.01$ ). Analyses disaggregated by grade suggest this finding is driven by an increase in the achievement gap between students in the Boston TILT schools and comparison schools in Grade 8 ( $\beta_5 = -0.13$ ,  $p < 0.01$ ). Further investigation, however, shows that there was a negative trend in mathematics achievement in Grade 8 prior to Boston TILT program implementation (see Figure 1f for the decrease in mathematics achievement from 2010 to 2012). Accounting for this trend by utilizing a linear trend model shows that students in the Boston TILT program actually displayed a significant increase on the mathematics test in Grade 8 after the first year of program implementation when compared with students in comparison schools ( $\beta_5 = 0.19$ ,  $p < 0.05$ ).

After the second year of the Boston TILT program's implementation, students in the Boston TILT schools performed significantly better on both the ELA ( $\beta_6 = 0.26$ ,  $p < 0.01$ ) and mathematics ( $\beta_6 = 0.07$ ,  $p < 0.05$ ) tests than students in comparison schools, although the effect size was relatively small for mathematics. The driving force appears to be students' achievement scores in Grade 8 ( $\beta_6 = 0.63$ ,  $p < 0.01$  for ELA;  $\beta_6 = 0.21$ ,  $p < 0.05$  for mathematics;  $\beta_6 = 0.68$ ,  $p < 0.01$  for mathematics, linear trend model), where the achievement gap between students in the Boston TILT schools and comparison schools decreased from 2013 to 2014 on both the ELA and the mathematics tests (see Figures 1e and 1f, respectively).

---

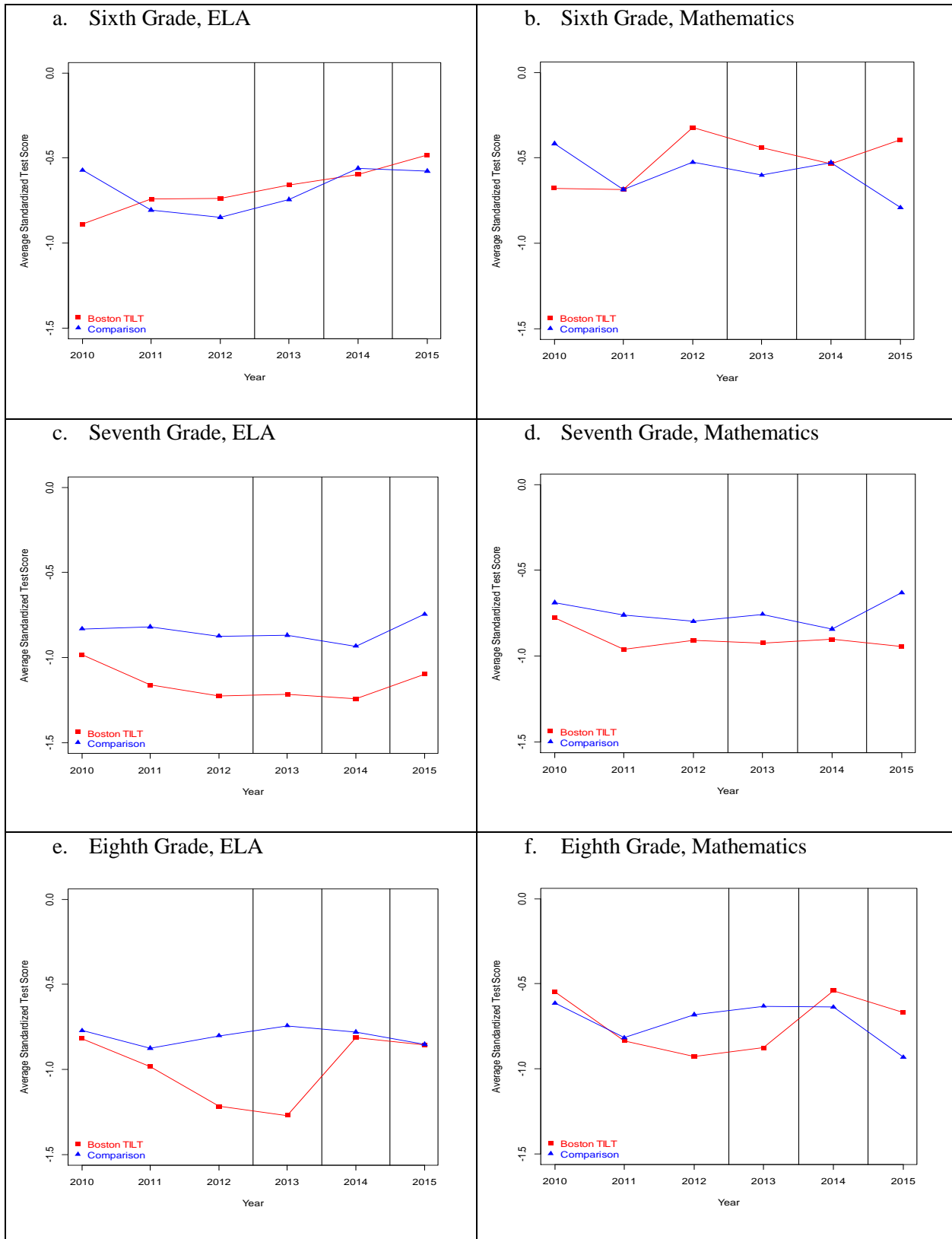
<sup>4</sup> Effect sizes were computed by subtracting the statewide average score from each student's score for that student's grade in each year and subject and dividing by the statewide standard deviation. Student raw scores on the MCAS in ELA, mathematics, and science were used in the analyses for the 2009–10, 2010–11, 2011–12, 2012–13, and 2013–14 school years, and scaled PARCC scores in ELA and mathematics were used for the 2014–15 school year.

After the third year of program implementation, students in the Boston TILT schools again performed significantly better on the ELA achievement test ( $\beta_7 = 0.40, p < 0.01$ ). This is because the achievement gap on the ELA test between students in the Boston TILT schools and comparison schools decreased even further from 2014 to 2015 in both Grade 7 ( $\beta_7 = 0.43, p < 0.05$ ; see Figure 1c) and Grade 8 ( $\beta_7 = 0.83, p < 0.01$ ; see Figure 1e). There was no significant effect for mathematics.

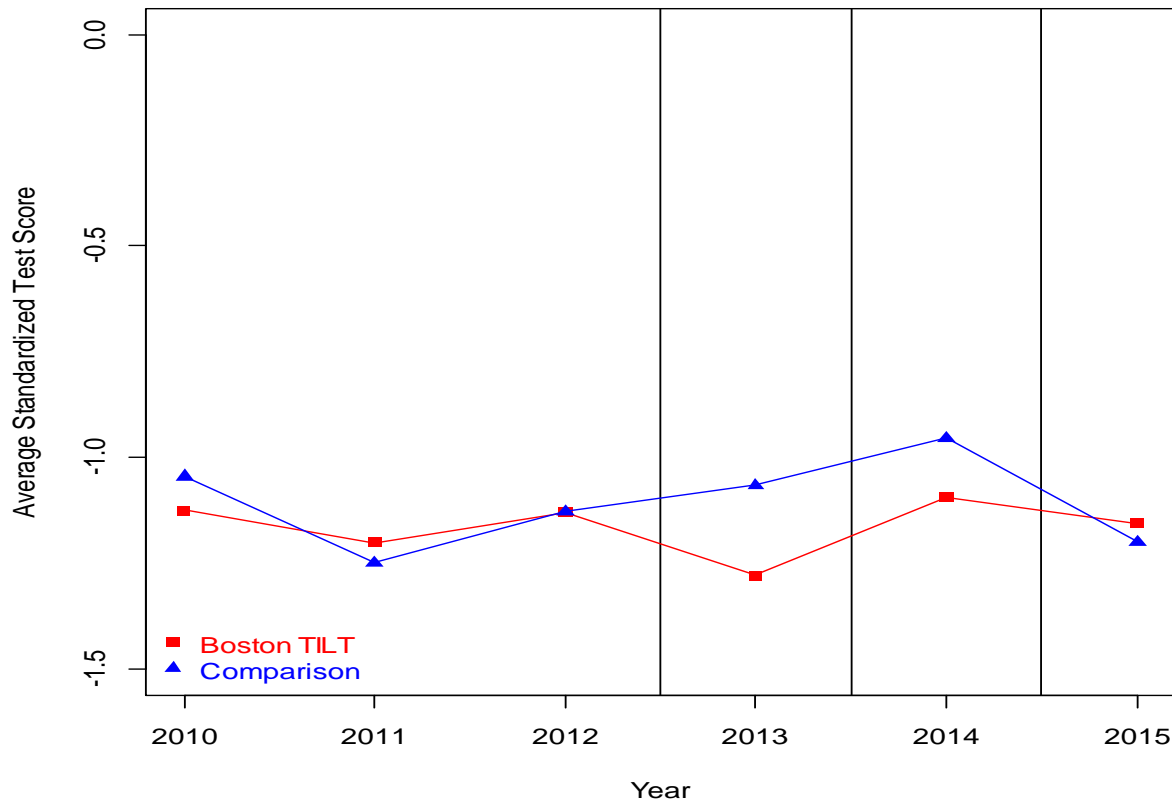
No changes in science achievement scores in Grade 8 were detected between students in the Boston TILT schools and comparison schools after the introduction of the Boston TILT program (Appendix B, Table B2). This finding is not surprising given the minimal change in average standardized test scores over the years between the Boston TILT schools and comparison schools (see Figure 2).



**Figures 1a–f. Average Standardized ELA and Mathematics Scores by Grade, School Year, and School Designation**



**Figure 2. Average Standardized Science Score by School Year and School Designation for Eighth Grade**



The subgroup analyses identified two statistical differences in the changes in the achievement gaps between the Boston TILT schools and comparison schools (Appendix D, Table D1). After the first year of program implementation, the achievement gap between students eligible for a free or reduced-price lunch and students not eligible increased on the ELA achievement test relative to the achievement gap between these students in the comparison schools ( $\beta_{12} = -0.21$ ,  $p < 0.05$ ). After the second year of program implementation, the achievement gap between students receiving special education services and those who do not also increased on the mathematics achievement test relative to the achievement gap between these students in the comparison schools ( $\beta_{13} = -0.17$ ,  $p < 0.05$ ).

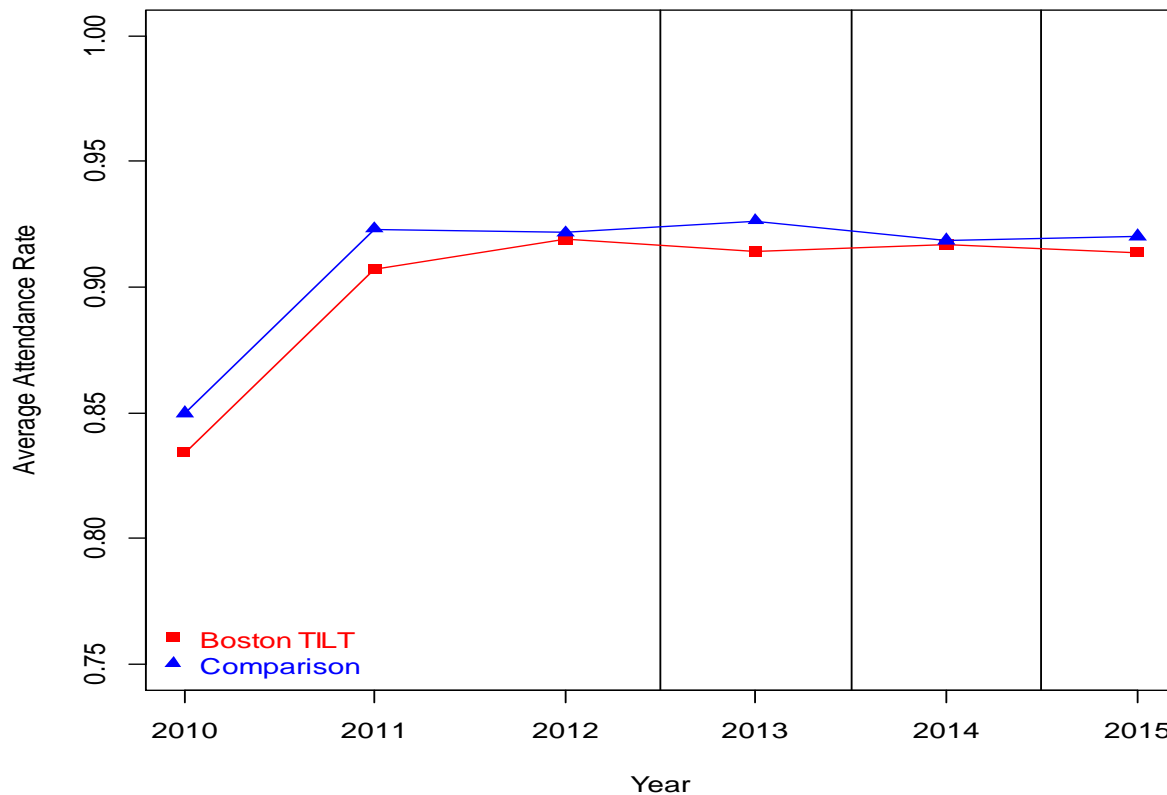
No changes were detected in the subgroups on the ELA and mathematics tests after the third year of the Boston TILT program’s implementation, nor were any changes detected on the science test in the subgroups after the first, second, or third year of program implementation (Appendix D, Table D2).

## Attendance

As shown in Figure 3, the attendance rate of students in both Boston TILT schools and comparison schools improved from 2010 to 2011. After the introduction of the Boston TILT program, there were no differences in the changes in student attendance rates between the students in the Boston TILT schools and those in comparison schools relative to prior attendance

rates overall<sup>5</sup> (Appendix B, Table B3) or by grade (Appendix C, Table C3). Because Boston TILT school students attend school for more hours than comparison school students each week, however, no change in attendance effectively indicates an increase in school participation.

**Figure 3. Average Student Attendance Rate by School Year and School Designation**



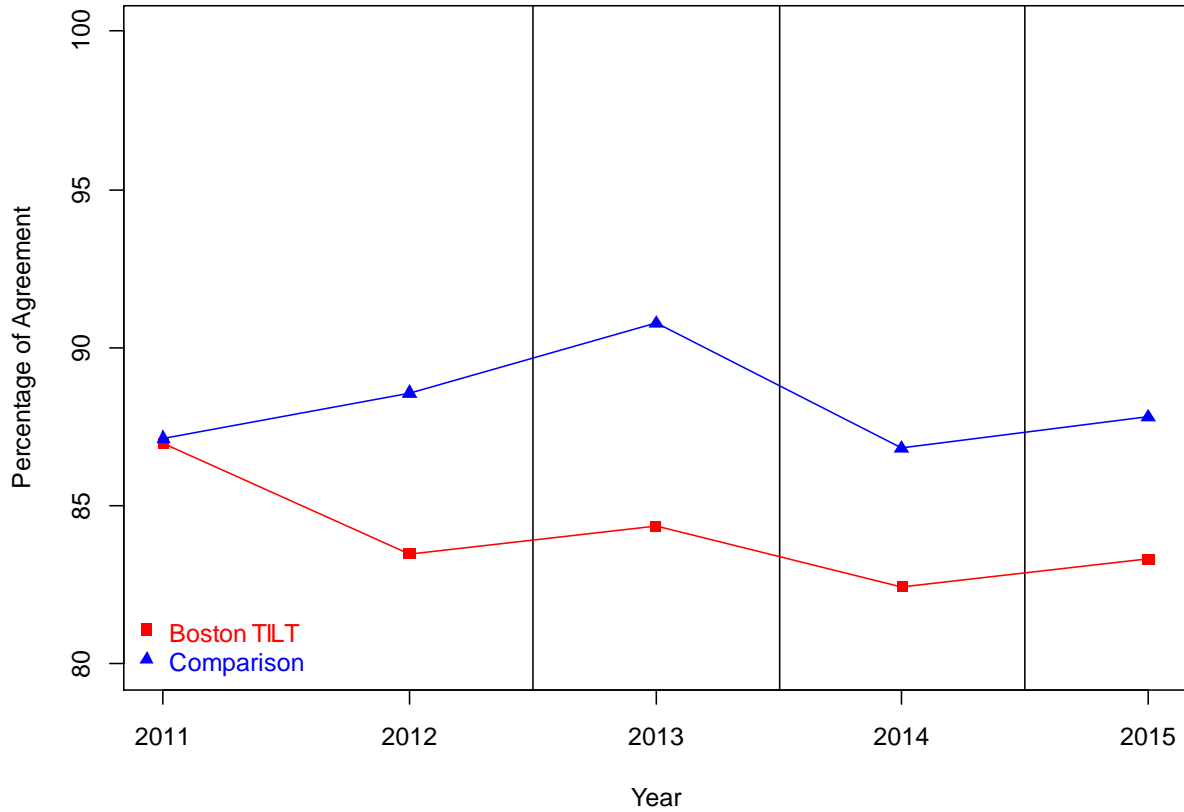
The subgroup analyses identified no statistical differences in the attendance rate gaps between students within a subgroup and other students in the Boston TILT schools versus comparison schools after the first, second, or third year of program implementation (Appendix D, Table D3).

## Perception of School Environment

As can be seen in Figure 4, students in the Boston TILT schools and comparison schools generally agreed or strongly agreed that their school had a positive environment. These ratings decreased slightly, however, over the years in the Boston TILT schools while increasing over the years in comparison schools, significantly increasing the gap in positive school environment ratings between students in the Boston TILT schools and comparison schools by 0.40 points ( $p < 0.01$ ) in the third year of program implementation (Appendix B,  $\beta_7$  in Table B4). This finding appears to be driven by Grade 6 students' perceptions, where the gap in positive school environment ratings between students in the Boston TILT schools and students in comparison schools increased by 0.83 ( $p < 0.01$ ) points from 2014 to 2015 (Appendix C,  $\beta_7$  in Table C5).

<sup>5</sup> The analyses for attendance were run in the arcsine metric to account for the fact that most attendance rates were close to one. Because the results remained unchanged, we chose to report the results without the arcsine transformation (i.e., in the original metric) for ease of interpretation.

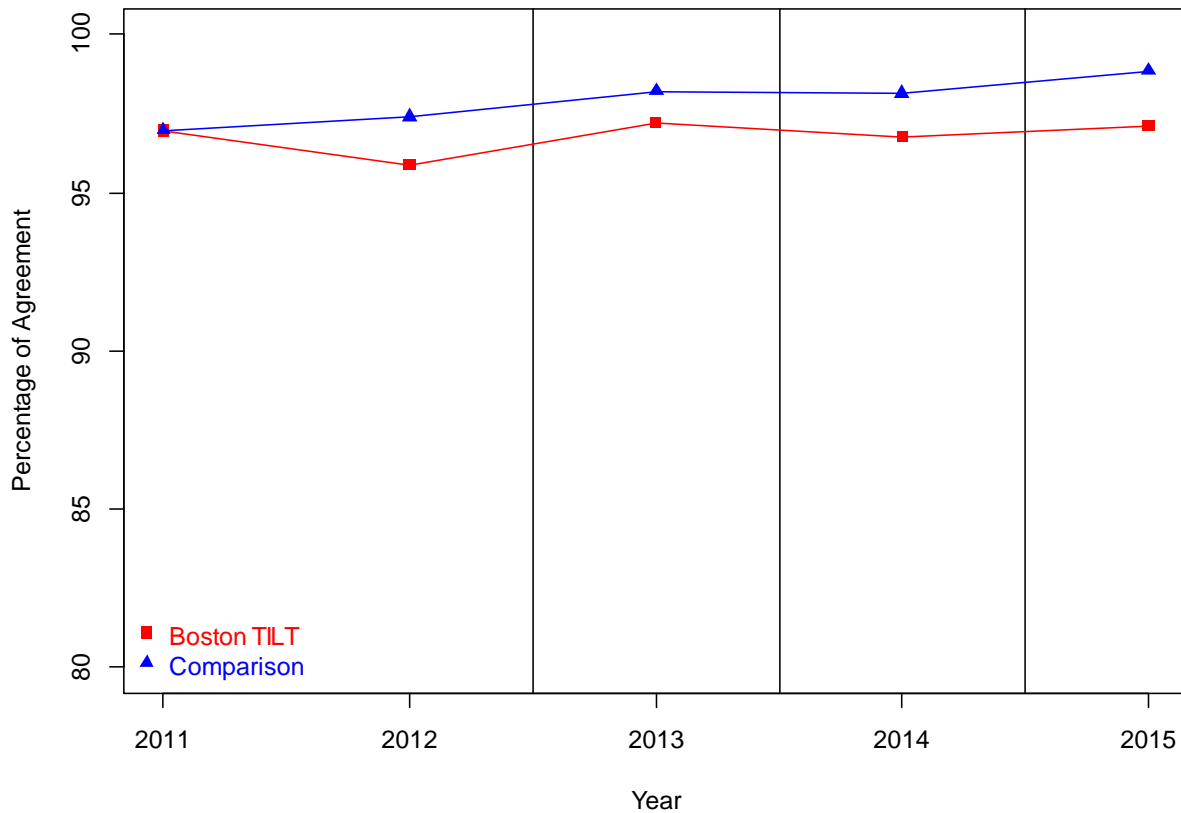
**Figure 4. Percentage of Agreement on the Positive School Environment Student Survey**



*Note.* The categories “Agree” and “Strongly Agree” were combined. The combining of categories is for presentation purposes only; the categories remained distinct in the analyses.

Similarly, Figure 5 shows that students in the Boston TILT schools and comparison schools generally agreed or strongly agreed that they were engaged in their learning. Although the trend remained relatively consistent across the years for both sets of schools, the gap in student engagement ratings between students in the Boston TILT schools and those in comparison schools decreased by 0.37 points ( $p < 0.05$ ) after the second year of program implementation (Appendix B, Table B4). This decrease in the gap from 2013 to 2014 can be found in both Grade 7 ( $\beta_6 = 0.66$ ,  $p < 0.01$ ) and Grade 8 ( $\beta_6 = 0.51$ ,  $p < 0.05$ ) but not Grade 6 (Appendix C, Table C4).

**Figure 5. Percentage of Agreement on the Engaged Learning Student Survey**



*Note.* The categories “Agree” and “Strongly Agree” were combined. The combining of categories is for presentation purposes only; the categories remained distinct in the analyses.

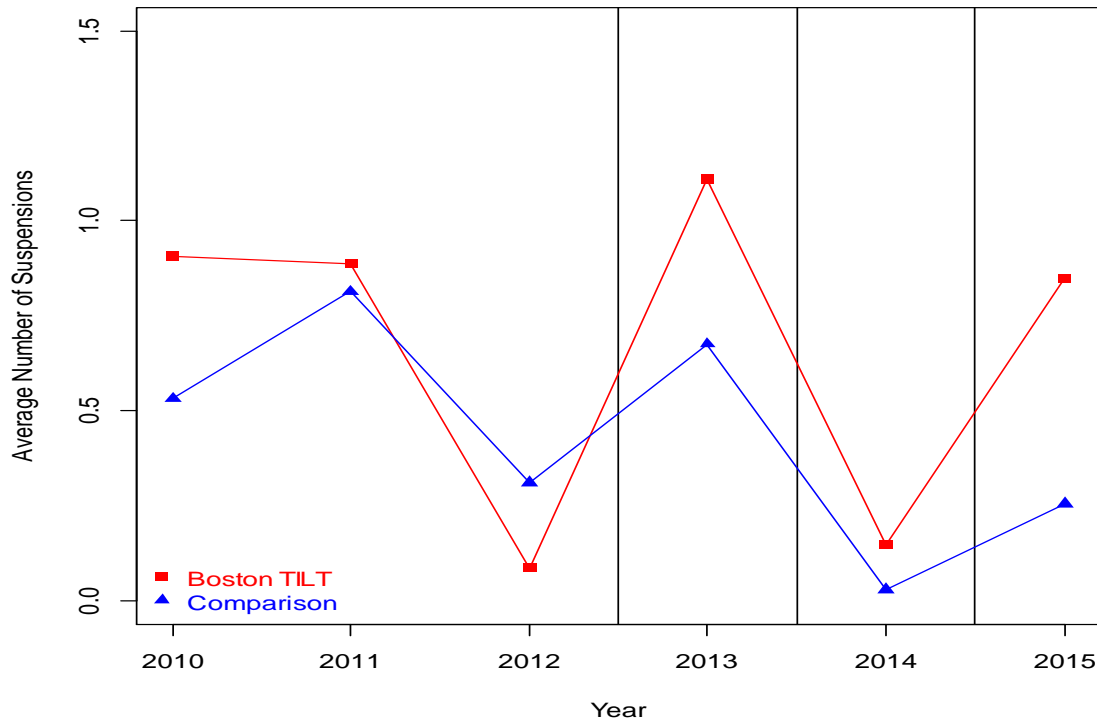
The subgroup analyses identified two statistical differences in student reports of a positive school environment after program implementation (Appendix D, Table D4). The first year after program implementation was associated with an increase in the gap on student ratings of positive school environment between students receiving special education services and those who do not ( $\beta_{12} = -0.51, p < 0.05$ ). For students receiving special education services, prior to program implementation students reported the positive school environment as higher than students who did not receive services. After program implementation, students in the Boston TILT schools who received special education services rated their school environment less highly than those who did not receive services, though the preimplementation pattern remained consistent in the comparison schools. The subgroup analyses also identified a greater gap on students’ ratings of positive school environment between ELL and non-ELL students in Boston TILT schools than in the comparison schools in the second year of program implementation ( $\beta_{13} = -0.51, p < 0.01$ ).

No changes were detected in the subgroups on students’ positive school environment ratings after the third year of the Boston TILT program’s implementation, nor were any changes detected on student’s engagement ratings in the subgroups after the first, second, or third year of program implementation.

## Student Behavior—Suspensions

Most students were not suspended for any days in a given year (Table 1), creating a rate of less than one suspension day a year on average (Figure 6). Nonetheless, the gap in the number of suspension days per student between students in the Boston TILT schools and students in comparison schools decreased by 0.16 ( $p < 0.01$ ) after the first year of program implementation, by 1.23 ( $p < 0.01$ ) after the second year of program implementation, and by 0.84 ( $p < 0.01$ ) after the third year of program implementation<sup>6</sup> (Appendix B, Table B5). This finding is due to slight increases in the number of suspension days in the Boston TILT schools and occasional decreases in the number of suspension days in the comparison schools. As Table C6 in Appendix C shows, the decrease in the gap between students in the Boston TILT schools and comparison schools can be found across most postimplementation years and grades (first postimplementation year:  $\beta_5 = 0.42$ ,  $p < 0.01$ , for Grade 6 and  $\beta_5 = 0.50$ ,  $p < 0.01$ , for Grade 8; second postimplementation year:  $\beta_6 = 1.80$ ,  $p < 0.01$ , for Grade 6 and  $\beta_6 = 1.51$ ,  $p < 0.01$ , for Grade 8; third postimplementation year:  $\beta_7 = 1.95$ ,  $p < 0.01$ , for Grade 6,  $\beta_7 = 0.27$ ,  $p < 0.05$ , for Grade 7, and  $\beta_7 = 0.84$ ,  $p < 0.01$ , for Grade 8). In Grade 7, however, there appears to be an increase in the gap in the number of suspension days per student between students in the Boston TILT schools and comparison schools after the first year of program implementation ( $\beta_5 = -0.47$ ,  $p < 0.01$ ).

**Figure 6. Average Number of Suspension Days per Student by School Year and School Designation**



<sup>6</sup> Because of the low frequency of student suspensions, the models were not able to account for grade and other student-level demographics. In addition, in 2012–13, the suspension days for two treatment students and one comparison student were excluded from these models because of the excessive number of days of suspension—65 days or more. In 2014–15, the suspension days for one treatment student was excluded.

**Table 1. Percentage of Students Without Suspensions by School Year and School Designation**

<b>School Year</b>	<b>Boston TILT</b>	<b>Comparison</b>
2009–10	79.7%	89.9%
2010–11	78.8%	78.9%
2011–12	96.0%	91.9%
2012–13	76.9%	86.7%
2013–14	97.5%	99.0%
2014–15	78.7%	90.5%

Special population analyses were not run because of the low frequency of suspensions within these groups.

## Conclusion

Overall, the CITS results suggest that the second (2013–14) and third (2014–15) years of the Boston TILT program’s implementation appear to have a more positive effect on students’ achievement in ELA and mathematics than the first year of program implementation (2012–13). Specifically, the Boston TILT program has a positive significant impact on students’ scores on the ELA achievement test after both the second and third years of program implementation and a positive significant impact on students’ scores on the mathematics achievement test after the second year of program implementation. This is not surprising because administrators, teachers, and students were adjusting to a new program in the first year, and qualitative data analyses revealed improvements in program implementation in Years 2 and 3, including stronger partnerships with outside providers and teacher-created structures to target students’ academic needs (see Haynes, Kistner, Therriault, & Hallberg, 2015). The lack of change detected in Boston TILT students’ eighth grade science scores is likely due to uneven implementation of science instruction in the two TILT middle schools.

School environment effects for student subgroups were mixed. Students reported higher ratings for engaged learning after the second year of program implementation, although they also reported lower positive school environment ratings after the third year of program implementation. There also was some evidence that the program has adverse effects for free or reduced-price-lunch-eligible and SPED students on the ELA and mathematics achievement tests, respectively. And the gap in students’ ratings of school environment increased for ELL and non-ELL students as well as SPED and non-SPED students in Boston TILT schools more than in the comparison schools.

The effects on students’ nonacademic outcomes also were mixed. There also was a positive significant difference in the change in suspension rates between the schools after all three years of program implementation, suggesting a reduction in the gap in suspension rates between Boston TILT schools and comparison schools. The percentage of students receiving suspensions in general was nevertheless very low. Last, no statistical differences attributable to the program were found between Boston TILT schools and comparison schools in attendance rates. With Boston TILT school students attending more hours (in general), no drop in attendance means that the students in Boston TILT schools are receiving more overall instruction.

Findings from a concurrent qualitative study over the three years of the TILT program (Haynes, Kistner, Therriault, & Hallberg, 2015) indicated that although implementation was a taxing experience for teachers, they were generally excited about its potential for students. One positive aspect of the program identified by numerous stakeholders at both schools was the availability of high-quality, supervised extracurricular programming for all students, regardless of background. At the same time, several challenges also were identified, including a lack of continuity in programming over the school day, inconsistent supports for SPED and ELL students, and overscheduled students and teachers. These implementation findings might account for the findings of mixed student outcomes. Taken together, the findings indicate an overall positive outcome of the TILT program, with room for improvement in implementation if continued.



## References

- Citkowicz, M., Haynes, E., Hallberg, K., & Therriault, S. B. (2014). *Boston turnaround using increased learning time: Student outcomes analysis*. Washington, DC: American Institutes for Research.
- Haynes, E., Kistner, A., Therriault, S., & Hallberg, K. (2015). *Implementation of the turnaround with increased learning time (TILT) model: Final evaluation memo*. Washington, DC: American Institutes for Research.
- Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., et al. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms* (NCSE 2013-3000). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research. Retrieved from <http://ies.ed.gov/ncser/pubs/20133000/pdf/20133000.pdf>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference* (2nd ed.). Mason, OH: Cengage Learning.
- Tucker, N., Hallberg, K., & Therriault, S. (2014). *Boston turnaround using increased learning time: Student outcomes analysis*. Washington, DC: American Institutes for Research.

## Appendix A. CITS Technical Approach

AIR used the following equation for the CITS baseline mean model to determine whether the Boston TILT program had an impact on student outcomes (including student achievement, attendance rate, and perception of school environment) after program implementation:

$$Y_{ij} = \beta_0 + \beta_1 TILT_j + \beta_2 Post_t + \beta_3 Post2_t + \beta_4 Post3_t + \beta_5 TILT_j Post_t + \beta_6 TILT_j Post2_t + \beta_7 TILT_j Post3_t + X_{ij} + v_j + u_{ij} + e_{ij}$$

In this model,  $Y_{ij}$  is the outcome measure (i.e., standardized achievement score, attendance rate, or survey scale score) for a student  $i$  at time  $t$  in school  $j$ ;  $TILT_j$  is an indicator for a school  $j$  that participated in the Boston TILT program;  $Post_t$ ,  $Post2_t$ , and  $Post3_t$  are indicators for whether a student was part of the program after its implementation in the first (2013), second (2014), and third (2015) postyears, respectively; and  $TILT_j Post_t$ ,  $TILT_j Post2_t$ , and  $TILT_j Post3_t$  are indicators of a student who was in a Boston TILT school *and* in the school after program implementation ( $TILT \times Post$ ) in the first, second, and third postyears, respectively. In this model, each indicator for a student is coded as 1 if it applies to a student and 0 otherwise. For example, a student in a Boston TILT school in 2013 (Post) would be coded 1 for Post, 1 for Boston TILT, and 1 for the  $TILT \times Post$  indicator. The model also included a set of student-level characteristics  $X_{ij}$  (i.e., grade, race, ELL status, free or reduced-price lunch status, special education status, and the prior year's student achievement standardized score for ELA and mathematics outcomes) that also may account for differences in the student outcomes. Random effects were included to account for school, time, and student effects by adding a random error term for each school ( $v_j$ ), cohort ( $u_{ij}$ ), and student ( $e_{ij}$ ).

Accounting for all student outcomes across time, the beta coefficients in the model represent the difference in differences between the Boston TILT schools and comparison schools prior to and after the start of the Boston TILT program. The  $\beta_1$  coefficient compares the mean outcome score (i.e., standardized achievement score, attendance rate, or student survey scale score) between students in Boston TILT schools and comparison schools in the pretreatment period. The  $\beta_2$ , and  $\beta_3$ , and  $\beta_4$  coefficients compare the mean outcome scores of students from all schools before and after program implementation in the first (2013), second (2014), and third (2015) postprogram implementation years, respectively. The  $\beta_5$ ,  $\beta_6$ , and  $\beta_7$  coefficients are the estimated effects of the Boston TILT program after the first, second, and third years of implementation, respectively.

*Model adjustment for the ELA achievement outcome.* In order to determine whether the results are robust to model selection, we also estimated linear trend models that include time indicators that allow for the modeling of slope trends across the years. The linear trend model was a more appropriate fit for the ELA achievement outcome:

$$Y_{ij} = \beta_0 + \beta_1 TILT_j + \beta_2 Post_t + \beta_3 Post2_t + \beta_4 Post3_t + \beta_5 TILT_j Post_t + \beta_6 TILT_j Post2_t + \beta_7 TILT_j Post3_t + \beta_8 Time_t + \beta_9 TILT_j Time_t + X_{ij} + v_j + u_{ij} + e_{ij}$$

In this model,  $Time_t$  is the outcome trend across time (years 2010–2015 are coded –3 through 2, respectively) and  $TILT_jTime_t$  is the interaction between  $TILT_j$  and  $Time_t$ , that allows for different pretreatment trends for TILT and non-TILT schools.

*Model adjustment for the suspension outcome.* Because of the low occurrence of student suspensions, the suspension model used a multilevel Poisson function to calculate the likelihood of a suspension:

$$Y_{ij} = \beta_0 + \beta_1 TILT_j + \beta_2 Post_t + \beta_3 Post2_t + \beta_4 Post3_t + \beta_5 TILT_j Post_t + \beta_6 TILT_j Post2_t + \beta_7 TILT_j Post3_t + v_j + u_{ij} + e_{ij}$$

In this model,  $Y_{ij}$  is the number of days a student  $i$  at time  $t$  in school  $j$  was suspended. The student-level characteristics were removed from the model because of the low occurrence of students who were within a subgroup and also were suspended.

*Subgroup analysis.* For the subgroup analysis, additional terms were added to the model to determine whether the Boston TILT program had an impact on student outcomes after Years 1 and 2 of program implementation:

$$Y_{ij} = \beta_0 + \beta_1 TILT_j + \beta_2 Post_t + \beta_3 Post2_t + \beta_4 Post3_t + \beta_5 TILT_j Post_t + \beta_6 TILT_j Post2_t + \beta_7 TILT_j Post3_t + \beta_8 TILT_j Subgroup_{ij} + \beta_9 Post_t Subgroup_{ij} + \beta_{10} Post2_t Subgroup_{ij} + \beta_{11} Post3_t Subgroup_{ij} + \beta_{12} TILT_j Post_t Subgroup_{ij} + \beta_{13} TILT_j Post2_t Subgroup_{ij} + \beta_{14} TILT_j Post3_t Subgroup_{ij} + X_{ij} + v_j + u_{ij} + e_{ij}$$

In this model, each subgroup analysis—ELL, free or reduced-price lunch status, and SPED—was run individually. The beta coefficients on the  $TILT_j Post_t Subgroup_{ij}$ ,  $TILT_j Post2_t Subgroup_{ij}$ , and  $TILT_j Post3_t Subgroup_{ij}$  variables were added to indicate whether there was a difference in the differences of a student in a subgroup versus not in a subgroup in a Boston TILT school versus a comparison school after the first, second, and third year of program implementation, respectively. Therefore, the beta coefficients represent the change in the gap between subgroup students and others in the Boston TILT schools versus changes in similar gaps in the comparison schools after program implementation in postyears 1 (2013), 2 (2014), and 3 (2015). In addition, four interaction terms— $Post_t Subgroup_{ij}$ ,  $Post2_t Subgroup_{ij}$ ,  $Post3_t Subgroup_{ij}$ , and  $TILT_j Subgroup_{ij}$ —were added to the model. Time indicators ( $Time_t$  and  $TILT_j Time_t$ ) were included in the ELA achievement models. Subgroup analyses were not run for the number of suspensions outcome.

Tables A1–A8 show the distribution of the variables included in the models.

## Distribution of Variables Over Time<sup>7</sup>

**Table A1. Mean Standardized ELA and Mathematics Scores**

Grade	School Year	ELA				Mathematics			
		N		Standardized Scores		N		Standardized Scores	
		Boston TILT	Comparison	Boston TILT	Comparison	Boston TILT	Comparison	Boston TILT	Comparison
Sixth	2009–10	420	335	-0.90 (1.27)	-0.56 (1.07)	419	339	-0.67 (1.13)	-0.49 (1.11)
	2010–11	399	344	-0.74 (1.20)	-0.81 (1.09)	405	345	-0.68 (1.18)	-0.73 (1.09)
	2011–12	409	314	-0.79 (1.20)	-0.83 (1.03)	414	315	-0.31 (1.14)	-0.56 (1.14)
	2012–13	408	298	-0.70 (1.23)	-0.74 (1.11)	413	299	-0.46 (1.23)	-0.62 (1.14)
	2013–14	388	186	-0.64 (1.27)	-0.56 (1.13)	392	186	-0.54 (1.22)	-0.53 (1.13)
	2014–15	274	152	-0.47 (1.17)	-0.53 (1.18)	282	164	-0.39 (1.13)	-0.77 (1.13)
Seventh	2009–10	336	252	-0.99 (1.12)	-0.79 (1.03)	335	253	-0.77 (1.04)	-0.68 (0.97)
	2010–11	339	281	-1.16 (1.14)	-0.78 (0.92)	344	283	-0.95 (1.00)	-0.82 (0.91)
	2011–12	337	303	-1.23 (1.10)	-0.83 (0.99)	341	304	-0.91 (0.99)	-0.81 (0.88)
	2012–13	339	270	-1.24 (1.06)	-0.81 (0.94)	340	273	-0.92 (0.92)	-0.80 (0.91)

<sup>7</sup> Percentages in tables may not sum to 100 as a result of rounding.

Grade	School Year	ELA				Mathematics			
		N		Standardized Scores		N		Standardized Scores	
		Boston TILT	Comparison	Boston TILT	Comparison	Boston TILT	Comparison	Boston TILT	Comparison
	2013–14	316	264	-1.27 (1.15)	-0.87 (1.01)	316	272	-0.91 (0.98)	-0.84 (0.92)
	2014–15	271	146	-1.03 (1.02)	-0.76 (0.98)	263	146	-0.89 (0.87)	-0.64 (0.91)
Eighth	2009–10	377	281	-0.82 (1.15)	-0.74 (1.05)	377	281	-0.54 (0.97)	-0.61 (1.00)
	2010–11	345	265	-0.98 (1.09)	-0.89 (1.02)	350	266	-0.84 (0.94)	-0.84 (0.95)
	2011–12	352	267	-1.22 (1.19)	-0.76 (1.07)	351	269	-0.92 (0.88)	-0.73 (0.90)
	2012–13	357	292	-1.27 (1.10)	-0.74 (0.97)	359	298	-0.88 (0.95)	-0.68 (0.94)
	2013–14	344	281	-0.83 (1.07)	-0.76 (0.93)	349	288	-0.52 (1.01)	-0.70 (0.92)
	2014–15	296	237	-0.85 (0.91)	-0.87 (0.89)	299	197	-0.64 (0.92)	-0.98 (0.70)

*Note.* Standard deviations are presented in parentheses.

**Table A2. Mean Standardized Science Scores for Eighth Grade**

School Year	<i>N</i>		Standardized Scores	
	Boston TILT	Comparison	Boston TILT	Comparison
2009–10	375	283	-1.12 (0.76)	-1.02 (0.80)
2010–11	347	264	-1.21 (0.90)	-1.22 (0.74)
2011–12	352	268	-1.13 (0.85)	-1.12 (0.80)
2012–13	357	295	-1.27 (0.91)	-1.10 (0.81)
2013–14	341	286	-1.10 (0.89)	-0.94 (0.83)
2014–15	316	260	-1.15 (0.92)	-1.16 (0.87)

*Note.* Standard deviations are presented in parentheses.

**Table A3. Mean Attendance Rate by School Type**

School Year	N		Mean Attendance Rate	
	Boston TILT	Comparison	Boston TILT	Comparison
2009–10	1,333	1,009	0.83 (0.27)	0.84 (0.28)
2010–11	1,156	950	0.91 (0.11)	0.92 (0.10)
2011–12	1,166	934	0.92 (0.09)	0.92 (0.09)
2012–13	1,159	916	0.91 (0.09)	0.93 (0.09)
2013–14	1,107	802	0.92 (0.09)	0.92 (0.09)
2014–15	940	640	0.91 (0.10)	0.92 (0.09)

Note. Standard deviations are presented in parentheses.

**Table A4. Mean Number and Range of Suspension Days per School Year by School Type**

School Year	N		Mean		Range	
	Boston TILT	Comparison	Boston TILT	Comparison	Boston TILT	Comparison
2009–10	1,333	1,009	0.88 (2.27)	0.39 (1.49)	0–18	0–14
2010–11	1,156	950	0.87 (2.31)	0.74 (1.84)	0–18	0–14
2011–12	1,166	934	0.08 (0.46)	0.30 (1.36)	0–5	0–21
2012–13	1,159	916	1.10 (4.66)	0.67 (2.95)	0–119	0–65
2013–14	1,107	802	0.14 (1.23)	0.03 (0.39)	0–29	0–9
2014–15	940	640	0.89 (4.60)	0.25 (1.01)	0–129	0–10

Note. Standard deviations are presented in parentheses.

**Table A5. Positive School Environment Student Survey Response Rates (Construct: Positive School Environment)**

School Year	School	N	Strongly Disagree	Disagree	Agree	Strongly Agree
2010–11	Boston TILT	867	4.2%	8.9%	79.0%	8.0%
	Comparison	700	4.1%	8.7%	80.9%	6.3%
2011–12	Boston TILT	915	5.6%	10.9%	76.0%	7.5%
	Comparison	778	4.9%	6.6%	79.3%	9.3%
2012–13	Boston TILT	843	6.5%	9.1%	74.3%	10.1%
	Comparison	682	3.2%	6.0%	81.7%	9.1%
2013–14	Boston TILT	826	8.4%	9.2%	70.0%	12.5%
	Comparison	554	5.4%	7.8%	75.6%	11.2%
2014–15	Boston TILT	671	7.9%	8.8%	70.6%	12.7%
	Comparison	361	3.6%	8.6%	69.0%	18.8%

**Table A6. Engaged Learning Student Survey Response Rates (Construct: Engaged Learning)**

School Year	Group	N	Strongly Disagree	Disagree	Agree	Strongly Agree
2010–11	Boston TILT	856	0.8%	2.2%	65.1%	31.9%
	Comparison	695	0.6%	2.4%	67.9%	29.1%
2011–12	Boston TILT	901	1.1%	3.0%	61.8%	34.1%
	Comparison	773	0.5%	2.1%	61.4%	36.0%
2012–13	Boston TILT	826	0.6%	2.2%	63.0%	34.3%
	Comparison	672	0.6%	1.2%	65.0%	33.2%
2013–14	Boston TILT	805	0.9%	2.4%	55.3%	41.5%
	Comparison	538	0.6%	1.3%	65.8%	32.3%
2014–15	Boston TILT	660	0.8%	2.1%	60.5%	36.7%
	Comparison	348	0.3%	0.9%	57.2%	41.7%



**Table A7. Student Demographics by School Type and Year**

School Year	N		SPED		Free or Reduced-Price Lunch Program		ELL	
	Boston TILT	Comparison	Boston TILT	Comparison	Boston TILT	Comparison	Boston TILT	Comparison
2009–10	1,333	1,009	25.7%	28.5%	82.9%	81.3%	40.5%	31.2%
2010–11	1,156	950	25.9%	27.7%	86.5%	86.1%	44.8%	35.7%
2011–12	1,166	934	27.0%	25.5%	77.0%	87.3%	46.7%	39.8%
2012–13	1,159	916	26.3%	25.0%	80.5%	81.4%	46.6%	40.6%
2013–14	1,107	802	26.8%	23.1%	88.1%	85.3%	47.2%	43.9%
2014–15	940	640	29.1%	26.1%	80.2%	72.2%	46.4%	48.0%

**Table A8. Student Demographics by School Type and Year**

School Year	N		Hispanic		Black		White		Asian		Other or More Than One Race	
	Boston TILT	Comparison	Boston TILT	Comparison	Boston TILT	Comparison	Boston TILT	Comparison	Boston TILT	Comparison	Boston TILT	Comparison
2009–10	1,333	1,009	42.0%	35.2%	37.5%	46.7%	10.4%	6.1%	8.0%	10.1%	2.1%	1.9%
2010–11	1,156	950	44.7%	33.4%	36.7%	49.1%	9.9%	6.8%	6.6%	8.8%	2.1%	1.9%
2011–12	1,166	934	46.2%	33.9%	34.7%	49.8%	10.2%	6.5%	6.5%	7.7%	2.3%	2.0%
2012–13	1,159	916	45.0%	29.7%	37.2%	51.6%	9.5%	8.0%	5.5%	7.8%	2.8%	2.9%
2013–14	1,107	802	42.5%	32.9%	38.1%	49.1%	9.1%	7.2%	7.3%	7.9%	3.0%	2.9%
2014–15	940	640	44.5%	38.0%	37.0%	42.2%	7.2%	8.4%	8.5%	8.6%	2.8%	2.8%

## Appendix B. Overall CITS Outcomes

Tables B1–B5 present the CITS results for each of the four outcomes, including student achievement, attendance rate, student perspective of school climate, and frequency of suspensions. For each model, the coefficients on Boston TILT  $\times$  Postyear 1 ( $\beta_5$ ), Boston TILT  $\times$  Postyear 2 ( $\beta_6$ ), and Boston TILT  $\times$  Postyear 3 ( $\beta_7$ ) represent the overall effects of the Boston TILT program after the first, second, and third year of program implementation, respectively. These effects represent changes in the outcomes of students in the Boston TILT schools after program implementation in comparison with changes in outcomes of students in comparison schools while controlling for various student-level demographics (i.e., grade, race, ELL, free or reduced-price lunch status, and SPED), as well as prior year’s achievement scores and time trends (where applicable). Student achievement outcomes are standardized, and thus effect sizes should be interpreted as standard deviation changes. Attendance outcomes are presented as rates, and thus effect sizes should be interpreted as percentage point changes. Suspension outcomes are presented as frequencies, and thus effect sizes should be interpreted as the number of suspension changes.

**Table B1. CITS Outcomes: Student ELA and Mathematics Achievement**

	ELA	Mathematics
Boston TILT ( $\beta_1$ )	-0.24** (0.05)	-0.01 (0.06)
Postyear 1 ( $\beta_2$ )	0.06 (0.04)	0.04 (0.09)
Postyear 2 ( $\beta_3$ )	0.04 (0.06)	0.02 (0.09)
Postyear 3 ( $\beta_4$ )	0.08 (0.08)	0.04 (0.09)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	0.06 (0.06)	-0.08** (0.03)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	0.26** (0.08)	0.07* (0.03)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	0.40** (0.10)	0.02 (0.04)
Time ( $\beta_8$ )	0.00 (0.02)	—
Boston TILT $\times$ Time ( $\beta_9$ )	-0.06** (0.02)	—

	<b>ELA</b>	<b>Mathematics</b>
Grade 7	0.02 (0.02)	0.04** (0.01)
Grade 8	0.01 (0.02)	0.03* (0.01)
Prior year's achievement score	0.68** (0.01)	0.74** (0.01)
SPED	-0.37** (0.02)	-0.25** (0.01)
Free or reduced-price lunch program	-0.09** (0.02)	-0.05** (0.02)
ELL	-0.01 (0.02)	0.01 (0.01)
Hispanic	-0.20** (0.03)	-0.18** (0.02)
Black	-0.19** (0.03)	-0.18** (0.02)
Asian	0.06 (0.04)	0.13** (0.03)
Other	-0.12* (0.05)	-0.13** (0.04)
Constant ( $\beta_0$ )	0.14** (0.05)	0.09 (0.06)
Time random-effects parameters	0.00 (0.00)	0.01 (0.07)
School random-effects parameters	0.00 (0.01)	0.00 (0.05)

*Note.* Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$

**Table B2. CITS Outcomes: Student Science Achievement—Eighth Grade**

	Science
Boston TILT ( $\beta_1$ )	-0.03 (0.05)
Postyear 1 ( $\beta_2$ )	-0.01 (0.06)
Postyear 2 ( $\beta_3$ )	0.15* (0.06)
Postyear 3 ( $\beta_4$ )	-0.05 (0.07)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	-0.08 (0.07)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	-0.09 (0.07)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	0.08 (0.08)
SPED	-0.60** (0.03)
Free or reduced-price lunch program	-0.10** (0.04)
ELL	-0.23** (0.03)
Hispanic	-0.25** (0.05)
Black	-0.29** (0.05)
Asian	0.17* (0.07)
Other	-0.11 (0.10)
Constant ( $\beta_0$ )	-0.55** (0.07)
Time random-effects parameters	0.00 (0.03)
School random-effects parameters	0.00 (0.04)

*Note.* Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$

**Table B3. CITS Outcomes: Attendance Rate**

	<b>Attendance Rate</b>
Boston TILT ( $\beta_1$ )	-0.009 (0.005)
Postyear 1 ( $\beta_2$ )	0.027 (0.032)
Postyear 2 ( $\beta_3$ )	0.017 (0.032)
Postyear 3 ( $\beta_4$ )	0.024 (0.032)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	0.001 (0.007)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	0.010 (0.008)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	0.003 (0.008)
Grade 7	-0.002 (0.003)
Grade 8	-0.005 (0.003)
SPED	-0.019** (0.003)
Free or reduced-price lunch program	0.031** (0.004)
ELL	0.031** (0.003)
Hispanic	-0.028** (0.005)
Black	0.002 (0.005)
Asian	0.026** (0.007)
Other	-0.005 (0.010)
Constant ( $\beta_0$ )	0.873** (0.017)

	<b>Attendance Rate</b>
Time random-effects parameters	0.001 (0.027)
School random-effects parameters	0.000 (0.003)

*Note.* Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$

**Table B4. CITS Outcomes: Perception of School Environment (Student Survey)**

	<b>Engaged Learning</b>	<b>Positive School Environment</b>
Boston TILT ( $\beta_1$ )	-0.08 (0.14)	-0.11 (0.23)
Postyear 1 ( $\beta_2$ )	0.05 (0.13)	0.09 (0.07)
Postyear 2 ( $\beta_3$ )	0.04 (0.14)	0.01 (0.07)
Postyear 3 ( $\beta_4$ )	0.48** (0.16)	0.45** (0.09)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	0.09 (0.14)	-0.07 (0.09)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	0.37* (0.15)	-0.01 (0.10)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	-0.29 (0.17)	-0.40** (0.11)
Grade 7	-0.56** (0.07)	-0.16** (0.04)
Grade 8	-0.65** (0.07)	-0.22** (0.04)
SPED	-0.04 (0.06)	0.16** (0.04)
Free or reduced-price lunch program	-0.03 (0.07)	0.07 (0.05)
ELL	0.40** (0.06)	0.27** (0.04)
Hispanic	-0.43** (0.11)	0.20** (0.07)
Black	-0.09 (0.10)	0.26** (0.07)
Asian	0.21 (0.14)	0.08 (0.09)
Other	-0.20 (0.19)	0.08 (0.12)

	<b>Engaged Learning</b>	<b>Positive School Environment</b>
Constant ( $\beta_0$ )	3.03** (0.16)	0.27 (0.18)
Time random-effects parameters	0.00 (0.06)	0.00 (0.00)
School random-effects parameters	0.01 (0.12)	0.05 (0.22)

*Note.* Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$



**Table B5. CITS Outcomes: Student Behavior—Suspension**

	Frequency of Suspensions
Boston TILT ( $\beta_1$ )	0.20 (0.25)
Postyear 1 ( $\beta_2$ )	0.41 (0.56)
Postyear 2 ( $\beta_3$ )	-2.56** (0.58)
Postyear 3 ( $\beta_4$ )	-0.49 (0.58)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	0.16** (0.06)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	1.23** (0.22)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	0.84** (0.09)
Constant ( $\beta_0$ )	-0.89** (0.33)
Time random-effects parameters	0.24 (0.49)
School random-effects parameters	0.06 (0.25)

*Note.* Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$

## Appendix C. CITS Outcomes by Grade

Tables C1–C6 present the CITS results for each of the four outcomes: student achievement, attendance rate, student perspective of school climate, and frequency of suspensions for the three grades separately—Grades, 6, 7, and 8. For each model, the coefficients on Boston TILT  $\times$  Postyear 1 ( $\beta_5$ ), Boston TILT  $\times$  Postyear 2 ( $\beta_6$ ), and Boston TILT  $\times$  Postyear 3 ( $\beta_7$ ) represent the overall effects of the Boston TILT program on each grade after the first, second, and third years of program implementation, respectively. Within each grade, these effects represent changes in the outcomes of students in the Boston TILT schools after program implementation in comparison with changes in outcomes of students in comparison schools while controlling for various student-level demographics (i.e., race, ELL, free or reduced-price lunch status, and SPED), as well as prior year’s achievement scores and time trends (where applicable).

**Table C1. CITS Outcomes by Grad: Student ELA Achievement**

	Sixth Grade	Seventh Grade	Eighth Grade
Boston TILT ( $\beta_1$ )	-0.05 (0.11)	-0.36** (0.10)	-0.33** (0.09)
Postyear 1 ( $\beta_2$ )	0.26** (0.07)	-0.09 (0.07)	-0.01 (0.08)
Postyear 2 ( $\beta_3$ )	0.32** (0.10)	-0.19 (0.10)	-0.04 (0.10)
Postyear 3 ( $\beta_4$ )	0.84** (0.14)	-0.18 (0.13)	-0.18 (0.13)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	-0.07 (0.09)	0.11 (0.10)	0.08 (0.10)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	0.01 (0.13)	0.13 (0.13)	0.63** (0.13)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	0.14 (0.41)	0.43* (0.18)	0.83** (0.17)
Time ( $\beta_8$ )	-0.09** (0.03)	0.07* (0.03)	0.04 (0.03)
Boston TILT $\times$ Time ( $\beta_9$ )	0.04 (0.04)	-0.08* (0.04)	-0.14** (0.04)
Prior year’s achievement score	0.68** (0.01)	0.68** (0.01)	0.66** (0.01)
SPED	-0.36** (0.03)	-0.33** (0.03)	-0.41** (0.03)
Free or reduced-price lunch program	-0.17** (0.03)	-0.01 (0.04)	-0.04 (0.03)

	Sixth Grade	Seventh Grade	Eighth Grade
ELL	-0.02 (0.03)	0.01 (0.03)	-0.01 (0.03)
Hispanic	-0.29** (0.04)	-0.13** (0.05)	-0.09 (0.05)
Black	-0.30** (0.04)	-0.07 (0.05)	-0.11* (0.05)
Asian	0.10* (0.05)	0.03 (0.07)	0.05 (0.06)
Other	-0.23** (0.08)	-0.07 (0.09)	0.02 (0.09)
Constant ( $\beta_0$ )	0.07 (0.09)	0.16 (0.09)	0.05 (0.09)
Time random-effects parameters	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
School random-effects parameters	0.01 (0.08)	0.00 (0.05)	0.00 (0.04)

*Note.* Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$

**Table C2. CITS Outcomes by Grade: Student Mathematics Achievement**

	Sixth Grade	Seventh Grade	Eighth Grade	Eighth Grade— Linear Trend Model
Boston TILT ( $\beta_1$ )	-0.02 (0.05)	-0.01 (0.05)	0.01 (0.10)	-0.31** (0.11)
Postyear 1 ( $\beta_2$ )	0.06 (0.12)	-0.07 (0.08)	0.13 (0.09)	0.06 (0.13)
Postyear 2 ( $\beta_3$ )	-0.02 (0.12)	-0.03 (0.08)	0.11 (0.09)	0.00 (0.18)
Postyear 3 ( $\beta_4$ )	0.06 (0.13)	0.13 (0.09)	0.03 (0.09)	-0.12 (0.23)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	-0.01 (0.06)	-0.16** (0.05)	-0.13** (0.05)	0.19* (0.08)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	0.04 (0.06)	-0.05 (0.05)	0.21** (0.05)	0.68** (0.10)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	-0.09 (0.32)	-0.18** (0.07)	0.11 (0.06)	0.73** (0.14)
Time ( $\beta_8$ )	–	–	–	0.04 (0.05)
Boston TILT $\times$ Time ( $\beta_9$ )	–	–	–	-0.16** (0.03)
Prior year’s achievement score	0.75** (0.01)	0.72** (0.01)	0.74** (0.01)	0.74** (0.01)
SPED	-0.33** (0.03)	-0.15** (0.02)	-0.26** (0.02)	-0.25** (0.02)
Free or reduced-price lunch program	-0.12** (0.03)	-0.04 (0.03)	0.04 (0.03)	0.03 (0.03)
ELL	-0.02 (0.02)	0.06** (0.02)	0.03 (0.02)	0.03 (0.02)
Hispanic	-0.28** (0.04)	-0.14** (0.04)	-0.03 (0.04)	-0.03 (0.04)
Black	-0.27** (0.04)	-0.10** (0.04)	-0.06 (0.04)	-0.06 (0.04)
Asian	0.08 (0.05)	0.18** (0.05)	0.20** (0.05)	0.21** (0.05)

	Sixth Grade	Seventh Grade	Eighth Grade	<b>Eighth Grade— Linear Trend Model</b>
Other	-0.17* (0.07)	-0.14 (0.07)	-0.01 (0.07)	-0.01 (0.07)
Constant ( $\beta_0$ )	0.27** (0.07)	0.03 (0.07)	-0.12 (0.09)	-0.05 (0.14)
Time random-effects parameters	0.01 (0.09)	0.00 (0.06)	0.01 (0.07)	0.00 (0.05)
School random-effects parameters	0.00 (0.04)	0.00 (0.04)	0.01 (0.09)	0.02 (0.15)

*Note.* Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$

**Table C3. CITS Outcomes by Grade: Attendance Rate**

	Sixth Grade	Seventh Grade	Eighth Grade
Boston TILT ( $\beta_1$ )	-0.005 (0.006)	-0.022** (0.006)	-0.001 (0.009)
Postyear 1 ( $\beta_2$ )	0.038 (0.05)	0.011 (0.025)	0.025 (0.017)
Postyear 2 ( $\beta_3$ )	0.028 (0.05)	0.009 (0.025)	0.010 (0.017)
Postyear 3 ( $\beta_4$ )	0.052 (0.05)	0.006 (0.026)	0.013 (0.017)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	0.003 (0.013)	0.018 (0.012)	-0.011 (0.012)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	0.000 (0.015)	0.009 (0.012)	0.024 (0.013)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	-0.012 (0.016)	0.014 (0.013)	0.009 (0.013)
SPED	-0.016** (0.006)	-0.017** (0.005)	-0.021** (0.005)
Free or reduced-price lunch program	0.071** (0.006)	0.021** (0.006)	0.001 (0.006)
ELL	0.027** (0.005)	0.033** (0.005)	0.033** (0.005)
Hispanic	-0.056** (0.009)	-0.014 (0.008)	-0.009 (0.009)
Black	-0.038** (0.008)	0.017* (0.008)	0.034** (0.009)
Asian	0.023* (0.011)	0.027* (0.012)	0.029* (0.012)
Other	-0.031 (0.017)	0.024 (0.015)	0.005 (0.017)
Constant ( $\beta_0$ )	0.864** (0.026)	0.877** (0.016)	0.869** (0.014)
Time random-effects parameters	0.002 (0.042)	0.000 (0.020)	0.000 (0.013)
School random-effects parameters	0.000 (0.000)	0.000 (0.003)	0.000 (0.006)

Note. Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$

**Table C4. CITS Outcomes by Grade: Perception of School Environment (Engaged Learning Student Survey)**

	Sixth Grade	Seventh Grade	Eighth Grade
Boston TILT ( $\beta_1$ )	0.08 (0.21)	-0.24 (0.16)	-0.17 (0.15)
Postyear 1 ( $\beta_2$ )	-0.16 (0.19)	0.26 (0.19)	0.03 (0.18)
Postyear 2 ( $\beta_3$ )	0.16 (0.22)	-0.04 (0.2)	-0.01 (0.19)
Postyear 3 ( $\beta_4$ )	0.51* (0.26)	0.76** (0.25)	0.23 (0.21)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	0.06 (0.24)	-0.01 (0.27)	0.37 (0.25)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	0.06 (0.27)	0.66** (0.27)	0.51* (0.26)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	-0.27 (0.31)	-0.42 (0.32)	-0.16 (0.28)
SPED	-0.17 (0.11)	-0.06 (0.11)	0.09 (0.11)
Free or reduced-price lunch program	0.06 (0.12)	-0.11 (0.14)	-0.09 (0.13)
ELL	0.26** (0.10)	0.49** (0.11)	0.48** (0.11)
Hispanic	-0.35* (0.16)	-0.60** (0.21)	-0.33 (0.20)
Black	-0.17 (0.16)	-0.18 (0.21)	0.11 (0.19)
Asian	0.15 (0.20)	0.27 (0.29)	0.32 (0.25)
Other	0.13 (0.29)	-0.57 (0.37)	-0.32 (0.36)
Constant ( $\beta_0$ )	3.00** (0.21)	2.64** (0.25)	2.30** (0.23)
Time random-effects parameters	0.00 (0.03)	0.00 (0.00)	0.00 (0.00)
School random-effects parameters	0.03 (0.16)	0.00 (0.07)	0.00 (0.00)

Note. Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$

**Table C5. CITS Outcomes by Grade: Perception of School Environment (Positive School Environment Student Survey)**

	Sixth Grade	Seventh Grade	Eighth Grade
Boston TILT ( $\beta_1$ )	0.01 (0.29)	-0.16 (0.23)	-0.21 (0.18)
Postyear 1 ( $\beta_2$ )	0.07 (0.12)	0.09 (0.12)	0.07 (0.11)
Postyear 2 ( $\beta_3$ )	0.03 (0.14)	-0.03 (0.13)	-0.01 (0.12)
Postyear 3 ( $\beta_4$ )	0.82** (0.17)	0.58** (0.16)	0.09 (0.13)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	-0.11 (0.16)	-0.21 (0.17)	0.19 (0.16)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	0.05 (0.18)	0.11 (0.17)	-0.12 (0.16)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	-0.83** (0.20)	-0.28 (0.20)	-0.17 (0.17)
SPED	0.19** (0.07)	0.09 (0.07)	0.18** (0.07)
Free or reduced-price lunch program	0.10 (0.08)	0.06 (0.09)	0.00 (0.08)
ELL	0.20** (0.07)	0.26** (0.07)	0.34** (0.07)
Hispanic	0.30** (0.11)	0.28* (0.14)	-0.05 (0.12)
Black	0.33** (0.11)	0.30* (0.13)	0.10 (0.12)
Asian	0.10 (0.13)	0.09 (0.18)	0.00 (0.16)
Other	0.24 (0.20)	-0.12 (0.23)	0.07 (0.22)
Constant ( $\beta_0$ )	0.11 (0.23)	0.11 (0.22)	0.34 (0.18)
Time random-effects parameters	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
School random-effects parameters	0.08 (0.28)	0.05 (0.21)	0.03 (0.16)

*Note.* Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$



**Table C6. CITS Outcomes by Grade: Student Behavior—Suspension**

	Sixth Grade	Seventh Grade	Eighth Grade
Boston TILT ( $\beta_1$ )	-0.15 (0.22)	0.36 (0.33)	0.34 (0.28)
Postyear 1 ( $\beta_2$ )	0.40 (0.59)	0.54 (0.56)	0.29 (0.59)
Postyear 2 ( $\beta_3$ )	-2.26** (0.69)	-2.28** (0.62)	-3.43** (0.77)
Postyear 3 ( $\beta_4$ )	-1.71** (0.65)	0.01 (0.56)	-0.47 (0.59)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	0.42** (0.11)	-0.47** (0.11)	0.50** (0.11)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	1.80** (0.38)	0.35 (0.33)	1.51** (0.53)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	1.95** (0.30)	0.27* (0.14)	0.84** (0.15)
Constant ( $\beta_0$ )	-0.98** (0.33)	-0.78* (0.36)	-0.92** (0.35)
Time random-effects parameters	0.25 (0.50)	0.22 (0.47)	0.25 (0.50)
School random-effects parameters	0.05 (0.21)	0.10 (0.32)	0.07 (0.27)

*Note.* Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$

## Appendix D. CITS Outcomes for Special Populations

Tables D1–D4 present the CITS results for three outcomes, including student achievement, attendance rate, and student perspective of school climate for three special populations of students—ELL, free or reduced-price lunch, and SPED status. For each model, the coefficients on Boston TILT  $\times$  Postyear 1  $\times$  Special Population ( $\beta_{12}$ ), Boston TILT  $\times$  Postyear 2  $\times$  Special Population ( $\beta_{13}$ ), and Boston TILT  $\times$  Postyear 3  $\times$  Special Population ( $\beta_{14}$ ) represent the overall effects of the Boston TILT program for students within this special population after the first, second, and third years of program implementation, respectively. These effects represent changes in the outcomes of students in the Boston TILT schools versus changes in outcomes of students in comparison schools after program implementation and relative to each special population’s reference group within their school while controlling for various student-level demographics (i.e., grade, race, ELL, free or reduced-price lunch, and SPED), as well as prior year’s achievement scores and time trends (where applicable). So, for example, the effects may represent the pre- to posttreatment change for ELL students in comparison with non-ELL students in Boston TILT schools versus the same change in comparison schools.

**Table D1. Regression Outcomes for Special Populations Within the Boston TILT Schools and Comparison Schools: Student ELA and Mathematics Achievement**

	ELL		Free or Reduced-Price Lunch		SPED	
	ELA	Mathematics	ELA	Mathematics	ELA	Mathematics
Boston TILT ( $\beta_1$ )	-0.25** (0.05)	0.00 (0.06)	-0.20** (0.06)	0.07 (0.07)	-0.20** (0.05)	0.00 (0.06)
Postyear 1 ( $\beta_2$ )	0.05 (0.05)	0.01 (0.09)	-0.02 (0.08)	0.07 (0.10)	0.05 (0.05)	0.04 (0.09)
Postyear 2 ( $\beta_3$ )	0.04 (0.06)	0.00 (0.09)	-0.05 (0.11)	0.06 (0.12)	0.06 (0.06)	-0.03 (0.09)
Postyear 3 ( $\beta_4$ )	0.06 (0.08)	0.01 (0.10)	0.08 (0.10)	-0.01 (0.11)	0.09 (0.08)	0.02 (0.09)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	0.04 (0.07)	-0.10* (0.04)	0.23* (0.10)	-0.14 (0.08)	0.03 (0.06)	-0.12** (0.04)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	0.27** (0.08)	0.09* (0.04)	0.37** (0.14)	0.05 (0.10)	0.25** (0.08)	0.11** (0.04)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	0.43** (0.11)	0.03 (0.06)	0.27 (0.15)	-0.04 (0.1)	0.37** (0.11)	0.00 (0.05)
Boston TILT $\times$ Subgroup ( $\beta_8$ )	0.01 (0.04)	-0.01 (0.03)	-0.05 (0.05)	-0.09* (0.04)	-0.16** (0.04)	-0.04 (0.03)
Subgroup $\times$ Postyear 1 ( $\beta_9$ )	0.00 (0.06)	0.06 (0.05)	0.10 (0.08)	-0.04 (0.06)	0.01 (0.06)	-0.03 (0.05)
Subgroup $\times$ Postyear 2 ( $\beta_{10}$ )	0.01 (0.06)	0.07 (0.05)	0.10 (0.11)	-0.04 (0.09)	-0.08 (0.07)	0.20** (0.06)
Subgroup $\times$ Postyear 3 ( $\beta_{11}$ )	0.03 (0.07)	0.07 (0.07)	0.00 (0.09)	0.07 (0.08)	-0.05 (0.08)	0.09 (0.07)

	ELL		Free or Reduced-Price Lunch		SPED	
	ELA	Mathematics	ELA	Mathematics	ELA	Mathematics
Boston TILT × Postyear 1 × Special Population ( $\beta_{12}$ )	0.03 (0.07)	0.02 (0.06)	-0.21* (0.10)	0.06 (0.08)	0.09 (0.08)	0.12 (0.07)
Boston TILT × Postyear 2 × Special Population ( $\beta_{13}$ )	-0.02 (0.08)	-0.05 (0.07)	-0.11 (0.13)	0.02 (0.11)	0.06 (0.09)	-0.17* (0.07)
Boston TILT × Postyear 3 × Special Population ( $\beta_{14}$ )	-0.06 (0.10)	-0.03 (0.09)	0.16 (0.13)	0.05 (0.11)	0.13 (0.11)	0.04 (0.09)
Time ( $\beta_{15}$ )	0.00 (0.02)	–	0.00 (0.02)	–	0.00 (0.02)	–
Boston TILT × Time ( $\beta_{16}$ )	-0.06** (0.02)	–	-0.06** (0.02)	–	-0.06** (0.02)	–
Grade 7	0.02 (0.02)	0.04** (0.01)	0.02 (0.02)	0.04** (0.01)	0.02 (0.02)	0.04** (0.01)
Grade 8	0.01 (0.02)	0.03* (0.01)	0.02 (0.02)	0.03* (0.01)	0.01 (0.02)	0.03* (0.01)
Prior year’s achievement score	0.68** (0.01)	0.74** (0.01)	0.68** (0.01)	0.74** (0.01)	0.68** (0.01)	0.74** (0.01)
SPED	-0.37** (0.02)	-0.25** (0.01)	-0.37** (0.02)	-0.25** (0.01)	-0.29** (0.03)	-0.26** (0.03)
FRPL	-0.09** (0.02)	-0.05** (0.02)	-0.07 (0.04)	-0.01 (0.03)	-0.09** (0.02)	-0.05** (0.02)
ELL	-0.02 (0.03)	0.00 (0.02)	-0.01 (0.02)	0.02 (0.01)	-0.01 (0.02)	0.02 (0.01)

	ELL		Free or Reduced-Price Lunch		SPED	
	ELA	Mathematics	ELA	Mathematics	ELA	Mathematics
Hispanic	-0.20** (0.03)	-0.18** (0.02)	-0.20** (0.03)	-0.18** (0.02)	-0.20** (0.03)	-0.18** (0.02)
Black	-0.19** (0.03)	-0.18** (0.02)	-0.19** (0.03)	-0.18** (0.02)	-0.19** (0.03)	-0.18** (0.02)
Asian	0.06 (0.04)	0.13** (0.03)	0.06 (0.04)	0.13** (0.03)	0.06 (0.04)	0.13** (0.03)
Other	-0.12* (0.05)	-0.13** (0.04)	-0.11* (0.05)	-0.13** (0.04)	-0.12* (0.05)	-0.13** (0.04)
Constant ( $\beta_0$ )	0.14** (0.05)	0.09 (0.06)	0.11* (0.06)	0.04 (0.07)	0.11* (0.05)	0.09 (0.06)
Time random-effects parameters	0.00 (0.00)	0.01 (0.07)	0.00 (0.00)	0.01 (0.07)	0.00 (0.01)	0.01 (0.07)
School random-effects parameters	0.00 (0.01)	0.00 (0.05)	0.00 (0.01)	0.00 (0.05)	0.00 (0.01)	0.00 (0.05)

*Note.* Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$

**Table D2. Regression Outcomes for Special Populations Within the Boston TILT Schools and Comparison Schools: Student Science Achievement, Eighth Grade**

	<b>ELL</b>	<b>Free or Reduced-Price Lunch</b>	<b>SPED</b>
Boston TILT ( $\beta_1$ )	-0.08 (0.06)	-0.02 (0.10)	0.02 (0.06)
Postyear 1 ( $\beta_2$ )	-0.10 (0.08)	-0.10 (0.13)	0.04 (0.07)
Postyear 2 ( $\beta_3$ )	0.13 (0.08)	-0.03 (0.14)	0.23** (0.07)
Postyear 3 ( $\beta_4$ )	0.02 (0.08)	-0.19 (0.12)	0.03 (0.07)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	0.02 (0.09)	-0.10 (0.18)	-0.12 (0.08)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	-0.02 (0.10)	-0.09 (0.21)	-0.11 (0.08)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	0.07 (0.10)	0.22 (0.17)	0.03 (0.09)
Boston TILT $\times$ Subgroup ( $\beta_8$ )	0.11 (0.07)	-0.02 (0.1)	-0.19* (0.08)
Subgroup $\times$ Postyear 1 ( $\beta_9$ )	0.25* (0.11)	0.11 (0.14)	-0.16 (0.12)
Subgroup $\times$ Postyear 2 ( $\beta_{10}$ )	0.05 (0.11)	0.21 (0.15)	-0.30* (0.13)
Subgroup $\times$ Postyear 3 ( $\beta_{11}$ )	-0.13 (0.11)	0.18 (0.13)	-0.26* (0.12)
Boston TILT $\times$ Postyear 1 $\times$ Special Population ( $\beta_{12}$ )	-0.26 (0.15)	0.03 (0.19)	0.13 (0.16)
Boston TILT $\times$ Postyear 2 $\times$ Special Population ( $\beta_{13}$ )	-0.16 (0.15)	-0.01 (0.22)	0.08 (0.17)
Boston TILT $\times$ Postyear 3 $\times$ Special Population ( $\beta_{14}$ )	0.00 (0.15)	-0.18 (0.19)	0.17 (0.17)
SPED	-0.60** (0.03)	-0.61** (0.03)	-0.42** (0.06)
Free or reduced-price lunch	-0.11** (0.04)	-0.17* (0.08)	-0.10** (0.04)
ELL	-0.29** (0.06)	-0.23** (0.03)	-0.24** (0.03)

	<b>ELL</b>	<b>Free or Reduced-Price Lunch</b>	<b>SPED</b>
Hispanic	-0.25** (0.05)	-0.25** (0.05)	-0.26** (0.05)
Black	-0.29** (0.05)	-0.29** (0.05)	-0.29** (0.05)
Asian	0.17** (0.07)	0.17* (0.07)	0.17** (0.07)
Other	-0.11 (0.10)	-0.11 (0.10)	-0.10 (0.10)
Constant ( $\beta_0$ )	-0.53** (0.07)	-0.50** (0.09)	-0.61** (0.07)
Time random-effects parameters	0.00 (0.03)	0.00 (0.03)	0.00 (0.03)
School random-effects parameters	0.00 (0.04)	0.00 (0.04)	0.00 (0.04)

*Note.* Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$

**Table D3. Regression Outcomes for Special Populations Within the Boston TILT Schools and Comparison Schools: Attendance Rate**

	<b>ELL</b>	<b>Free or Reduced-Price Lunch</b>	<b>SPED</b>
Boston TILT ( $\beta_1$ )	-0.009 (0.006)	0.024* (0.010)	-0.012* (0.005)
Postyear 1 ( $\beta_2$ )	0.030 (0.032)	0.106** (0.033)	0.028 (0.032)
Postyear 2 ( $\beta_3$ )	0.015 (0.032)	0.119** (0.034)	0.020 (0.032)
Postyear 3 ( $\beta_4$ )	0.026 (0.032)	0.114** (0.033)	0.033 (0.032)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	0.002 (0.010)	-0.021 (0.017)	0.006 (0.008)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	0.019 (0.010)	-0.003 (0.020)	0.012 (0.009)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	0.005 (0.011)	-0.016 (0.017)	0.004 (0.01)
Boston TILT $\times$ Subgroup ( $\beta_8$ )	0.000 (0.007)	-0.038** (0.010)	0.011 (0.008)
Subgroup $\times$ Postyear 1 ( $\beta_9$ )	-0.008 (0.011)	-0.095** (0.014)	-0.004 (0.013)
Subgroup $\times$ Postyear 2 ( $\beta_{10}$ )	0.004 (0.012)	-0.120** (0.016)	-0.012 (0.013)
Subgroup $\times$ Postyear 3 ( $\beta_{11}$ )	-0.005 (0.013)	-0.114** (0.015)	-0.035* (0.014)
Boston TILT $\times$ Postyear 1 $\times$ Special Population ( $\beta_{12}$ )	-0.001 (0.015)	0.023 (0.019)	-0.017 (0.017)
Boston TILT $\times$ Postyear 2 $\times$ Special Population ( $\beta_{13}$ )	-0.018 (0.015)	0.018 (0.022)	-0.005 (0.018)
Boston TILT $\times$ Postyear 3 $\times$ Special Population ( $\beta_{14}$ )	-0.004 (0.017)	0.024 (0.020)	0.000 (0.018)
Grade 7	-0.002 (0.003)	-0.002 (0.003)	-0.002 (0.003)
Grade 8	-0.005 (0.003)	-0.004 (0.003)	-0.005 (0.003)



	<b>ELL</b>	<b>Free or Reduced-Price Lunch</b>	<b>SPED</b>
SPED	-0.019** (0.003)	-0.017** (0.003)	-0.016** (0.006)
Free or reduced-price lunch	0.031** (0.004)	0.098** (0.007)	0.032** (0.004)
ELL	0.035** (0.006)	0.030** (0.003)	0.031** (0.003)
Hispanic	-0.028** (0.005)	-0.025** (0.005)	-0.028** (0.005)
Black	0.002 (0.005)	0.005 (0.005)	0.002 (0.005)
Asian	0.026** (0.007)	0.026** (0.007)	0.026** (0.007)
Other	-0.005 (0.010)	-0.003 (0.009)	-0.004 (0.010)
Constant ( $\beta_0$ )	0.872** (0.017)	0.814** (0.018)	0.872** (0.017)
Time random-effects parameters	0.001 (0.027)	0.001 (0.027)	0.001 (0.027)
School random-effects parameters	0.000 (0.004)	0.000 (0.005)	0.000 (0.003)

Note. Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$

**Table D4. Regression Outcomes for Special Populations Within the Boston TILT Schools and Comparison Schools: Perception of School Environment (Student Survey)**

	ELL		Free or Reduced-Price Lunch		SPED	
	EL <sup>a</sup>	PSE	EL	PSE	EL	PSE
Boston TILT ( $\beta_1$ )	-0.03 (0.16)	-0.06 (0.24)	-0.04 (0.23)	0.00 (0.26)	-0.11 (0.15)	-0.15 (0.23)
Postyear 1 ( $\beta_2$ )	0.11 (0.16)	-0.02 (0.09)	-0.31 (0.27)	0.00 (0.17)	0.01 (0.14)	0.06 (0.08)
Postyear 2 ( $\beta_3$ )	0.01 (0.17)	-0.22* (0.1)	-0.22 (0.30)	0.10 (0.19)	0.05 (0.15)	0.05 (0.09)
Postyear 3 ( $\beta_4$ )	0.47* (0.2)	0.40** (0.12)	0.33 (0.29)	0.23 (0.19)	0.54** (0.18)	0.50** (0.10)
Boston TILT $\times$ Postyear 1 ( $\beta_5$ )	0.21 (0.19)	0.05 (0.12)	0.37 (0.34)	0.00 (0.22)	0.20 (0.17)	0.06 (0.11)
Boston TILT $\times$ Postyear 2 ( $\beta_6$ )	0.36 (0.21)	0.23 (0.13)	0.47 (0.40)	-0.33 (0.25)	0.40* (0.17)	-0.02 (0.11)
Boston TILT $\times$ Postyear 3 ( $\beta_7$ )	-0.37 (0.24)	-0.41** (0.15)	-0.02 (0.37)	-0.22 (0.24)	-0.22 (0.20)	-0.48** (0.13)
Boston TILT $\times$ Subgroup ( $\beta_8$ )	-0.11 (0.17)	-0.08 (0.11)	-0.05 (0.22)	-0.13 (0.14)	0.13 (0.19)	0.19 (0.12)
Subgroup $\times$ Postyear 1 ( $\beta_9$ )	-0.16 (0.22)	0.25 (0.14)	0.44 (0.28)	0.11 (0.18)	0.17 (0.25)	0.11 (0.16)
Subgroup $\times$ Postyear 2 ( $\beta_{10}$ )	0.03 (0.23)	0.49** (0.15)	0.31 (0.32)	-0.10 (0.21)	-0.05 (0.27)	-0.16 (0.17)
Subgroup $\times$ Postyear 3 ( $\beta_{11}$ )	0.00 (0.27)	0.11 (0.18)	0.20 (0.33)	0.30 (0.21)	-0.23 (0.31)	-0.17 (0.2)

	ELL		Free or Reduced-Price Lunch		SPED	
	EL <sup>a</sup>	PSE	EL	PSE	EL	PSE
Boston TILT × Postyear 1 × Special Population ( $\beta_{12}$ )	-0.21 (0.29)	-0.29 (0.19)	-0.33 (0.37)	-0.09 (0.24)	-0.41 (0.33)	-0.51* (0.22)
Boston TILT × Postyear 2 × Special Population ( $\beta_{13}$ )	0.04 (0.30)	-0.51** (0.20)	-0.12 (0.43)	0.37 (0.27)	-0.12 (0.35)	0.06 (0.23)
Boston TILT × Postyear 3 × Special Population ( $\beta_{14}$ )	0.18 (0.35)	0.01 (0.22)	-0.34 (0.42)	-0.25 (0.27)	-0.27 (0.39)	0.30 (0.25)
Grade 7	-0.56** (0.07)	-0.16** (0.04)	-0.56** (0.07)	-0.16** (0.04)	-0.56** (0.07)	-0.16** (0.04)
Grade 8	-0.65** (0.07)	-0.22** (0.04)	-0.64** (0.07)	-0.22** (0.04)	-0.65** (0.07)	-0.22** (0.04)
SPED	-0.04 (0.06)	0.16** (0.04)	-0.05 (0.06)	0.16** (0.04)	-0.02 (0.14)	0.10 (0.09)
FRPL	-0.03 (0.07)	0.07 (0.05)	-0.10 (0.17)	0.09 (0.11)	-0.03 (0.07)	0.07 (0.05)
ELL	0.50** (0.13)	0.25** (0.08)	0.40** (0.06)	0.27** (0.04)	0.40** (0.06)	0.27** (0.04)
Hispanic	-0.42** (0.11)	0.21** (0.07)	-0.42** (0.11)	0.20** (0.07)	-0.43** (0.11)	0.20** (0.07)
Black	-0.08 (0.10)	0.28** (0.07)	-0.08 (0.11)	0.27** (0.07)	-0.08 (0.10)	0.26** (0.07)
Asian	0.21 (0.14)	0.08 (0.09)	0.22 (0.14)	0.08 (0.09)	0.21 (0.14)	0.08 (0.09)
Other	-0.18 (0.19)	0.10 (0.12)	-0.19 (0.19)	0.08 (0.12)	-0.19 (0.19)	0.08 (0.12)

	ELL		Free or Reduced-Price Lunch		SPED	
	EL <sup>a</sup>	PSE	EL	PSE	EL	PSE
Constant ( $\beta_0$ )	2.98** (0.16)	0.26 (0.18)	3.07** (0.21)	0.25 (0.20)	3.02** (0.16)	0.28 (0.18)
Time random-effects parameters	0.00 (0.06)	0.00 (0.00)	0.00 (0.06)	0.00 (0.00)	0.00 (0.06)	0.00 (0.00)
School random-effects parameters	0.01 (0.12)	0.05 (0.23)	0.01 (0.12)	0.05 (0.22)	0.01 (0.12)	0.05 (0.22)

<sup>a</sup> PSE, positive school environment; EL, engaged learning

*Note.* Standard errors are presented in parentheses (standard deviations for the random-effects parameters).

\* $p \leq .05$ ; \*\* $p \leq .01$