Boston Public Schools Expanded Learning Time Research Collaborative

Year 2 Findings Report

Award: R305H150013

September 2017

Stephanie Marek
Sarah Faude
Donna Muncey
Boston Public Schools

Alexandra Kistner
Dionisio García Píriz
Ryan Williams
Susan Therriault
American Institutes for Research

2465_09/17

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305H150013 to the American Institutes for Research. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Table of Contents

Introduction

Since 2006, Boston Public Schools (BPS) has implemented expanded learning time (ELT) with the goal of improving student outcomes and closing opportunity and achievement gaps. ELT, which has been supported by the U.S. Department of Education as part of its Title I, 1003(g) School Improvement Grants (SIG) program, generally includes some combination of additional instructional time for students, additional planning and collaboration time for teachers, and "enrichments" (Center on Education Policy, 2012).[1] Currently, schools in BPS can add time through a variety of funding channels, with differing restrictions and requirements. However, schools generally have substantial latitude for deciding how much time is added and how additional time is used.

In school year 2015–16, 46 Boston public schools implemented ELT. These schools added at least an additional 30 minutes per day, or 150 minutes per week, beyond the district standard.[2] These schools, which comprised roughly 36% of the district schools, adopted a longer day through multiple initiatives including autonomous or turnaround status funding, expanded day grants, or inclusion in a mayoral initiative[3], called Schedule A, to expand the day by 40 minutes in all traditional elementary and middle schools. By 2018 the mayoral initiative will increase the school day by 40 minutes in 60 schools, roughly half of the district's schools, including all traditional BPS schools serving Grades K–8. Given this rapid growth in schools adopting ELT, understanding how ELT schools use their extra time, and the impact of this time on student outcomes, are matters of considerable interest to the district and to the field of education overall.

Over the past two years, BPS has collaborated with American Institutes for Research (AIR) to study its ELT program types, with the goal of understanding the impact of ELT on student, staff, and parent outcomes. In Year 1 (2015–16) of the project, BPS identified 46 schools implementing ELT, and AIR and BPS staff interviewed school administrators and staff to begin to understand how schools were choosing to use expanded time, as well as strengths and challenges of implementing ELT.

Overall findings from Year 1 revealed the following:

- The district lacks centralized information on schools' time usage, a potential barrier to understanding ELT implementation.
- The amount of time added varies by ELT policy and funding type, with no single districtwide system. Allocation of time also varies, with some relationship to instructional focus.

[1] Center on Education Policy. (2012). *Increased learning time under stimulus-funded school improvement grants: High hopes, varied implementation.* Washington, DC: Author.
[2] A standard school day is defined as 6 hours for elementary schools; 6 hours, 10 minutes for middle schools; and 6 hours, 20 minutes for high schools. Grades K–8 schools were treated as elementary schools, and Grades 7–12 schools were treated as middle schools.
[3] The mayoral initiative began in 2015 as an agreement between the mayor of Boston and the Boston Teachers Union.

- The greatest reported strength of ELT—teacher satisfaction about its potential for student outcomes—is at odds with the greatest reported challenges—lack of teacher buy-in and teacher burnout.

One question from Year 1 interviews focused on understanding how school-based leadership and staff defined successful ELT. Responses to this question were central to the project's third advisory panel meeting, held August 30, 2016. This meeting, attended by BPS parents and teachers, directors of local educational nonprofits, representatives from the Boston Teachers Union, and Massachusetts Department of Elementary and Secondary Education staff, reviewed Year 1 findings and discussed how these findings should inform Year 2 quantitative analyses. The panel strongly agreed that how school respondents defined success should guide decisions about the outcomes selected and examined for Year 2 analyses.

How Year 1 interview respondents defined successful ELT fell into four major categories:

1. Improved student academic achievement
2. Teacher satisfaction and investment
3. Family satisfaction
4. Student social–emotional growth

With these success metrics framing Year 2 analyses, members of the BPS Office of Data and Accountability informed the panel of available BPS data that might serve as a measure or proxy of these four categories. For example, the panel discussed which measure of academic achievement would be most appropriate for the analysis. Although multiple measures were considered, including student grade point average, state test data were chosen due to their consistency and completeness over time. Additionally, data sources were considered based on understandability by multiple audiences in order to make findings accessible to readers representing the multitude of key BPS stakeholders.

Ultimately, the following data sources were chosen as measures for each of the four success metrics:

- Student composite performance index (CPI) math, English language arts (ELA), and science scores (Metric 1. Academic achievement)
- Teacher climate surveys (Metric 2. teacher satisfaction)
- Parent climate surveys (Metric 3. Family satisfaction)
- School choice data: Grades Pre-K–8 (Metric 3. Family satisfaction)
- Student attendance (Metric1. Academic achievement/ Metric 4. Student social-emotional growth)
- Student climate data (Metric 4. Student social emotional growth)

This report summarizes methodologies, findings, implications, and suggestions for future research for examining these metrics in the following sections: academic achievement, attendance, student and teacher climate, and parent engagement.

Academic Achievement

In order to produce robust, context-rich findings, several academic achievement analyses were conducted. More rigorous analyses looked at whether additional time alone improves student outcomes, while descriptive analyses delved deeper into how schools' use of time impacted student outcomes. Sample size was a major determinant in the level of rigor of the analyses. As expanded learning time schools were broken into smaller subgroups based on time usage, more rigorous analysis produced results that were less reliable due to small group or sample sizes.

Impact Analyses

Methodology

Looking at whether additional time alone affects student outcomes, researchers performed comparative interrupted time series (CITS) analyses to examine impacts on student performance in the first two years after ELT implementation. The CITS analyses compared outcome scores between ELT and non-ELT schools before and after ELT implementation, accounting for trends before ELT and differences in school composition. AIR conducted a confirmatory analysis with all schools with available data across all district schools, constituting this as the main achievement analysis. Thirty-one ELT schools met the data requirements to be included in the analysis, for which outcome scores for all of the students were compared with respective outcomes in 63 non-ELT schools.[4] School-level CPI score[5] data from mathematics and ELA state assessments[6] for Grades 3–8 and Grade 10, and science assessments scores for Grades 5, 8, and 10 were collected from school years 2005–06 to 2015–16 and used as outcomes. AIR also conducted a sensitivity analysis using a propensity score matched comparison sample, whereby ELT schools were paired with non-ELT schools based on prior achievement and school composition.

Additional subgroup analyses were conducted by student demographic groups (e.g., gender, racial groups, English language learners (ELLs), students with disabilities, and economically disadvantaged students[7]), as well as an exploratory analysis by school type (e.g., turnaround, pilot, innovation) using the same CITS analysis framework.

---

[4] To be included in the CITS analysis, ELT schools needed to have at least three years of data prior to implementing the ELT program. All non-ELT schools that were still active in 2016 with data available were included.

[5] The CPI is a 100-point index that assigns 100, 75, 50, 25, or 0 points to each student participating in the state assessment Massachusetts Comprehensive Assessment System [MCAS] or Partnership for Assessment of Readiness for College and Careers [PARCC]) based on their achievement. The CPI is a measure of the extent to which all students are progressing toward proficiency. CPIs are generated separately for ELA, mathematics, and science, and at all levels: state, district, school, and subgroup. See Massachusetts Department of Elementary and Secondary Education, *School Leader's Guide to the 2017 Accountability Determinations* (http://www.mass.gov/edu/docs/ese/accountability/annual-reports/school-leaders-guide.pdf).

[6] The state assessment was the MCAS for all grades in years 2006 through 2014. In 2015 and 2016 schools had the option of using the PARCC assessment or MCAS, so results from either assessment that each BPS school chose to take were used in these school years.

[7] Due to changing variables used for data collection and reporting over the period of data analyzed, the CPI scores used as an outcome for economically disadvantaged students or the information on percentage of economically disadvantaged students in a school used as a covariate, where combined data from the "low income" variable for years 2006–14, and the "economic disadvantage" variable for 2015 and 2016. "Low income" identified students who met any one of the following definitions of low income: is eligible for free or reduced-price lunch, receives Transitional Aid to Families benefits, or is eligible for food stamps. The inclusion of a student in the "economic

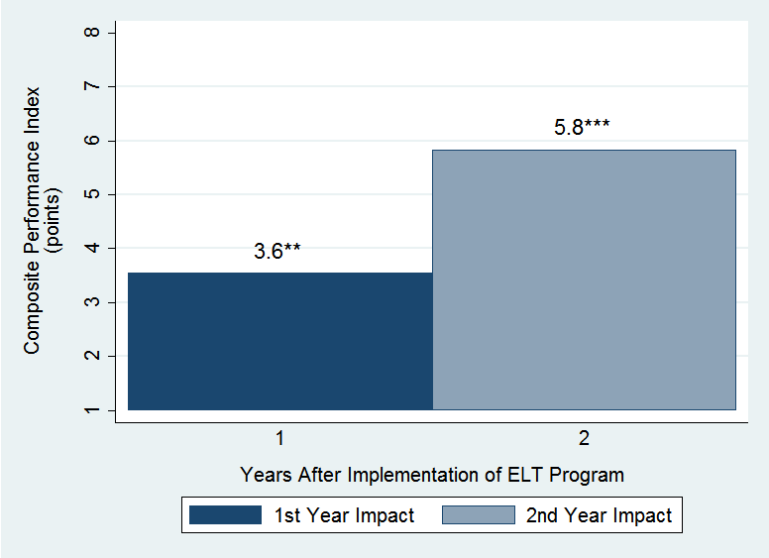
For more details on the analysis and methodology, please see Appendix A.

Findings

The main analysis was conducted for all students; a statistically positive impact was found on ELA and mathematics achievement in the first and second year following ELT implementation. The impact was especially strong in mathematics (both in magnitude and statistical significance). These results suggest that by extending their school days, ELT schools in BPS improved student achievement in these subjects. No impact was found for science achievement.[8]

- **Mathematics:** After the first year of ELT implementation, students in ELT schools scored 3.6 CPI points, on average, higher than students in non-ELT schools relative to their previous performance trajectories. In other words, students in ELT schools scored higher, on average, than students in non-ELT schools that had similar prior-achievement and student populations. Students in ELT schools also scored 5.8 CPI points higher, on average, than those in non-ELT comparison schools in the second year of ELT implementation. These estimates are large both in terms of magnitude and statistical significance with effect sizes of around 0.2 standard deviations in Year 1 and 0.4 in Year 2 (see Figure 1).

Figure 1. Difference in All Students Mathematics CPI Scores in ELT Schools Compared to Peers in Non-ELT Schools



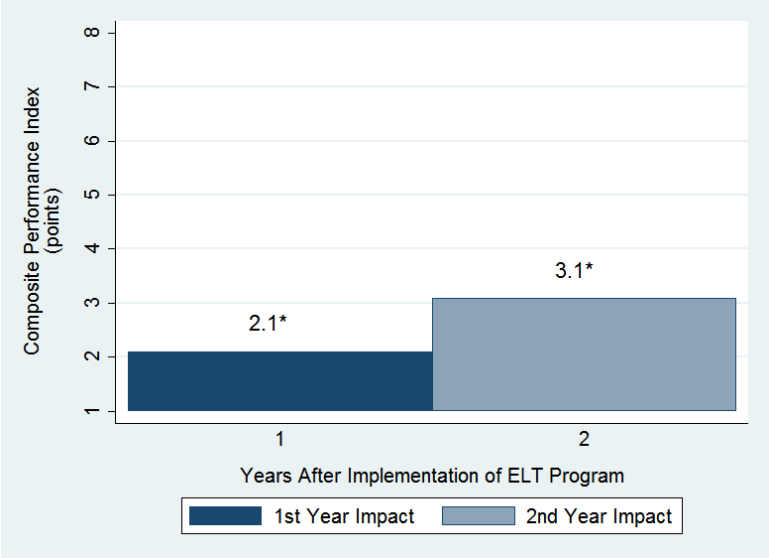
Note. ** $p < 5\%$; *** $p < 1\%$.

disadvantage" group is based on the student's participation in one or more of the following state-administered programs: Supplemental Nutrition Assistance Program (SNAP), Transitional Assistance for Families with Dependent Children (TAFDC), Department of Children and Families' (DCF) foster care program, and MassHealth (Medicaid).

[8] Results obtained in the sensitivity analysis using the matched comparison group were similar in magnitude and statistical significance. Please see Table B1 in Appendix for more detail.

- **ELA:** After the first year of ELT implementation, students in ELT schools scored 2.1 CPI points higher in ELA, on average, than students in non-ELT schools relative to their previous performance trajectories. In the second year of ELT implementation, the impact estimate gains were 3.1 CPI points, on average. These represent effect sizes of around 0.15 standard deviations in Year 1, and 0.25 in Year 2, although the statistical significance is more limited (see Figure 2).

Figure 2. Difference in All Students ELA CPI Scores in ELT Schools Compared to Peers in Non-ELT Schools



Note. * $p < 10\%$.

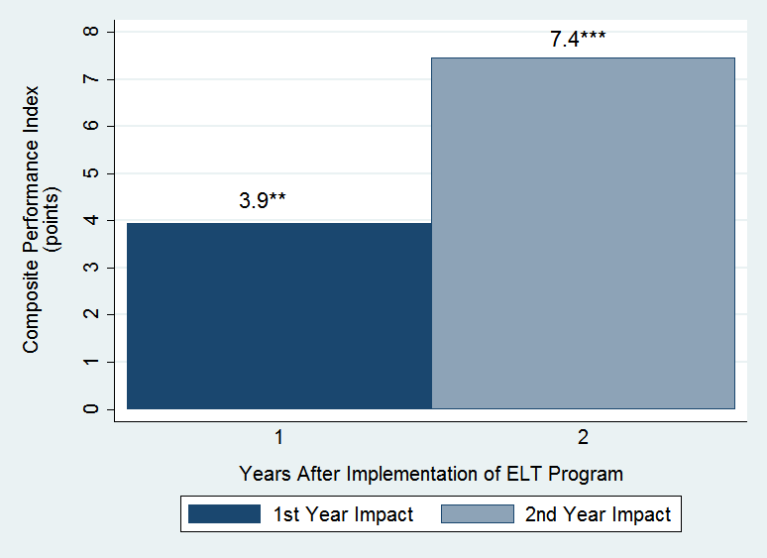
Findings by Student Demographic Groups

Analyses of student demographic subgroups showed evidence of impact gains for Black and Hispanic students in mathematics and ELA during both Year 1 and Year 2 ELT implementation. Some gains were also observed for economically disadvantaged students and ELLs, primarily in Year 2 of ELT implementation. No impact was found for White or Asian students or students with disabilities.[9]

- **Mathematics:** The subgroup analysis shows that ELT had a positive impact on mathematics performance across most student demographic groups, with particularly strong results for Black and Hispanic students.
    - After Year 1 of ELT implementation, **Black** students in ELT schools scored 3.9 CPI points higher, on average, than their peers in non-ELT schools with similar past trajectories. After Year 2 of implementation, Black students in ELT schools scored 7.4 CPI points higher, on average, compared to their peers in non-ELT schools (see Figure 3).

[9] Asian and White subgroups have low representation in schools that are part of the analysis. As a result, the models for these subgroups had significantly fewer observations given that many of these schools did not have a sufficiently large (or, in some cases, any) student subgroup population to have an achievement record.

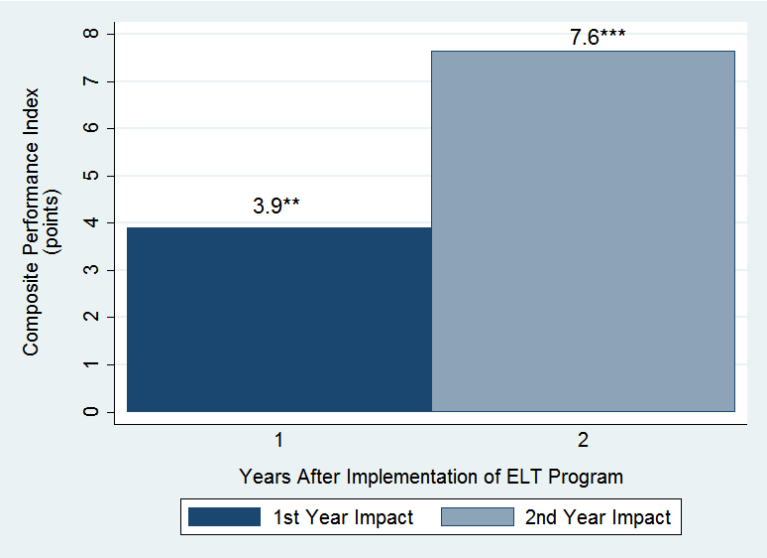
Figure 3. Difference in Mathematics CPI Scores for Black Students in ELT Schools Compared to Peers in Non-ELT Schools



Note. ** $p < 5\%$; *** $p < 1\%$.

- Similarly, **Hispanic** students in ELT schools scored 3.9 and 7.6 CPI points higher, on average, than their peers in Year 1 and Year 2, respectively (see Figure 4).

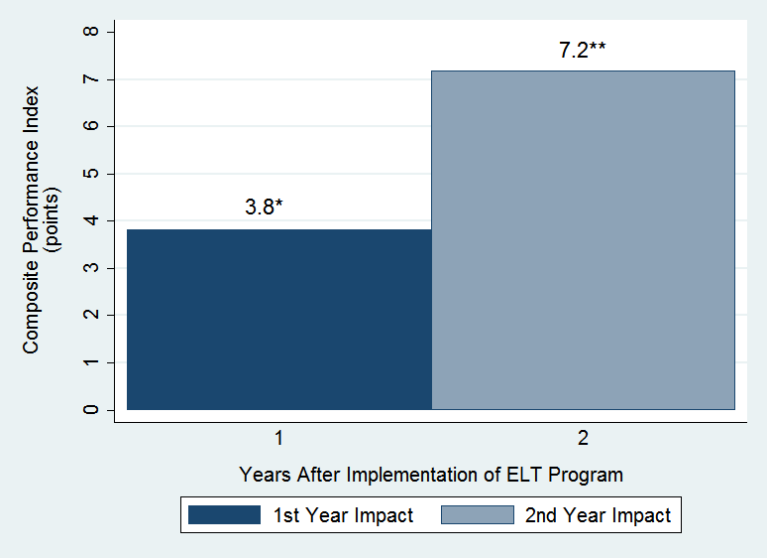
Figure 4. Difference in Mathematics CPI Scores for Hispanic Students in ELT Schools Compared to Peers in Non-ELT Schools



Note. ** $p < 5\%$; *** $p < 1\%$.

- Furthermore, **ELLs** in ELT schools scored 3.8 and 7.2 CPI points, on average, higher than their peers in non-ELT schools in the first and second years of implementation, respectively (see Figure 5).

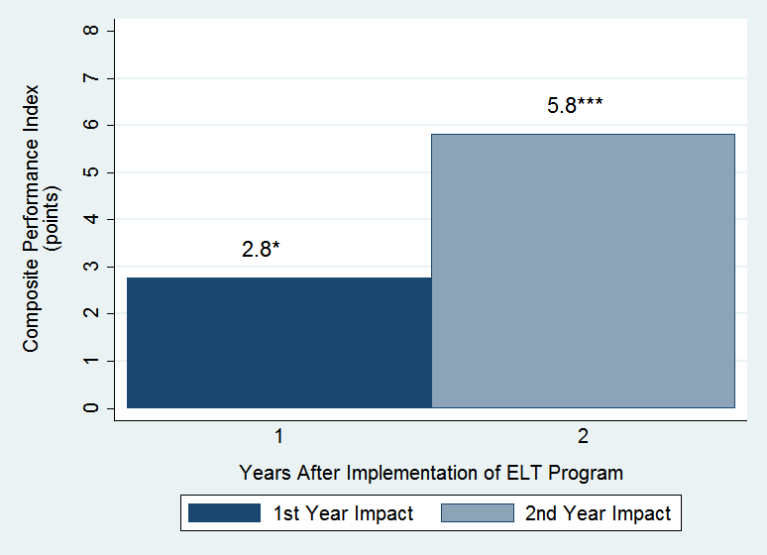
Figure 5. Difference in Mathematics CPI Scores for English Language Learners in ELT Schools Compared to Peers in Non-ELT Schools



Note. * $p < 10\%$; ** $p < 5\%$.

- **Economically disadvantaged** students scored 2.8 and 5.8 CPI points higher than their peers in non-ELT schools in the first and second years of implementation, respectively (see Figure 6).

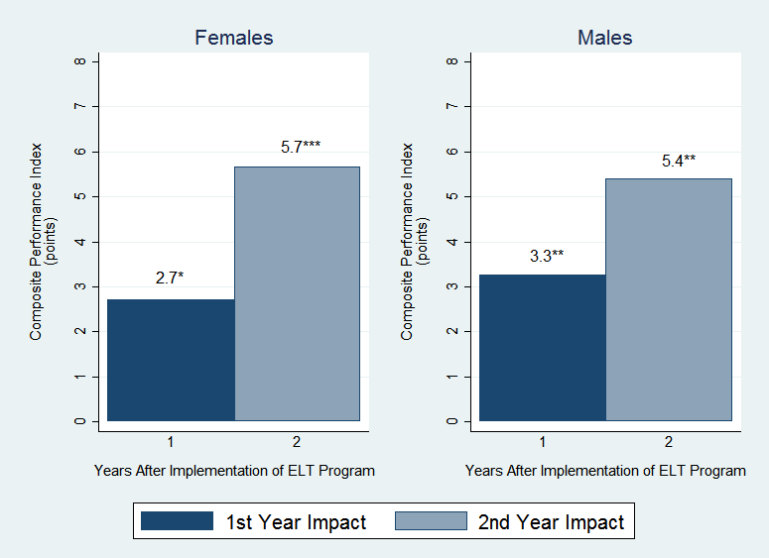
Figure 6. Difference in Mathematics CPI Scores for Economically Disadvantaged Students in ELT Schools Compared to Peers in Non-ELT Schools



Note. * $p < 10\%$; *** $p < 1\%$.

- In Year 1 of ELT implementation, **female** students scored 2.7 CPI points higher and **male** students scored 3.3 CPI points higher, on average than their peers in non-ELT schools. In Year 2, female and male students scored 5.7 and 5.4 CPI points higher, on average, respectively (see Figure 7).

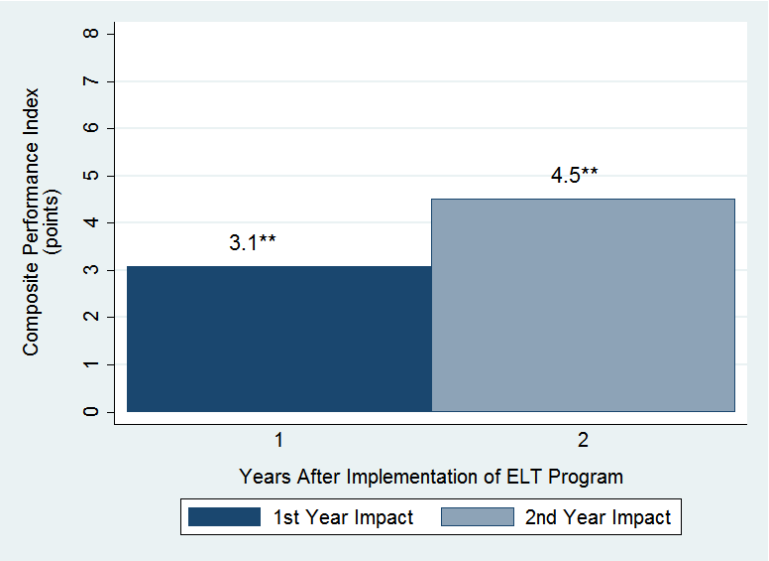
Figure 7. Difference in Mathematics CPI Scores for Female and Male Students in ELT Schools Compared to Peers in Non-ELT Schools



Note. * $p < 10\%$; ** $p < 5\%$; *** $p < 1\%$.

- **ELA:** The subgroup analysis found ELT had a positive impact on ELA performance across some student demographic groups, particularly in Year 2 of implementation.
  - The subgroup analysis found **Black** students in ELT schools scored 3.1 CPI points higher, on average, than their peers in non-ELT schools in Year 1 relative to their past trajectories than Black students in non-ELT schools and 4.5 CPI points higher in Year 2 (see Figure 8).

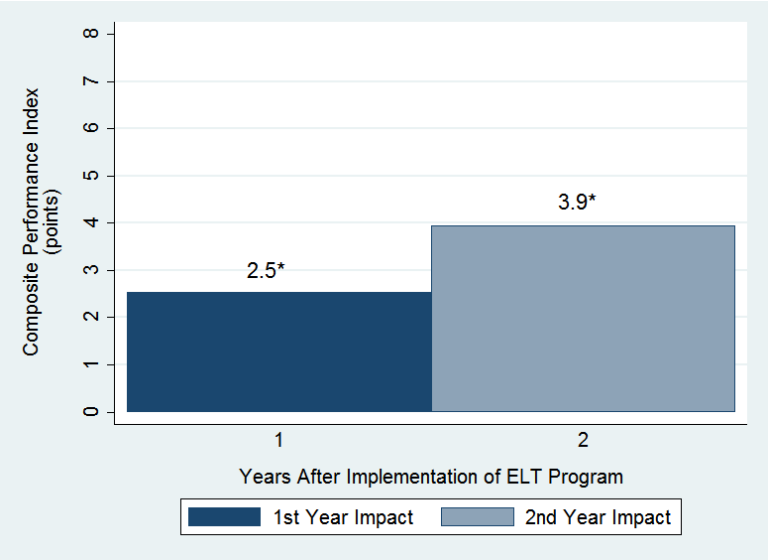
Figure 8. Difference in ELA CPI Scores for Black Students in ELT Schools Compared to Peers in Non-ELT Schools



Note. ** $p < 5\%$.

- Positive impacts were also found for **Hispanic** students compared to their peers in non-ELT schools, although at a lower degree: 2.5 and 4 CPI points in Year 1 and Year 2, on average, respectively (see Figure 9).

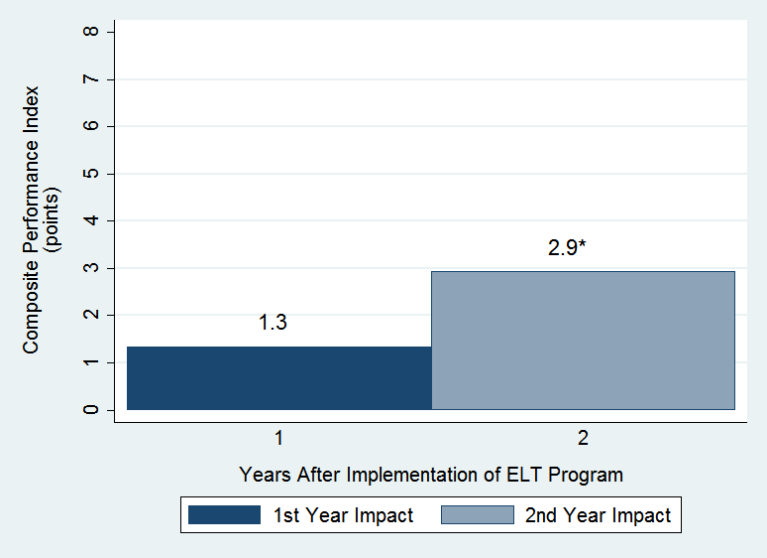
Figure 9. Difference in ELA CPI Scores for Hispanic Students in ELT Schools Compared to Peers in Non-ELT Schools



Note: * $p < 10\%$.

- Analyses found positive impacts of 2.9 CPI points, on average, for **economically disadvantaged** students in ELT schools compared to their peers in non-ELT schools in the second year of ELT (see Figure 10).

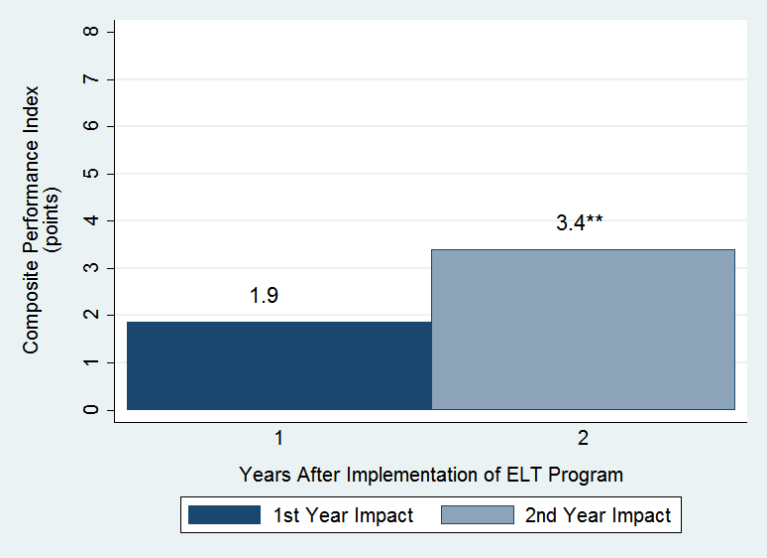
Figure 10. Difference in ELA CPI Scores for Economically Disadvantaged Students in ELT Schools Compared to Peers in Non-ELT Schools



Note. * $p < 10\%$.

- Similarly, positive impacts of 3.4 CPI points were found for **female** students in ELT schools compared to their peers in non-ELT schools in the second year of ELT, on average (see Figure 11).

Figure 11. Difference in ELA CPI Scores for Female Students in ELT Schools Compared to Peers in Non-ELT Schools



Note. ** $p < 5\%$.

Findings by School Type

The 31 ELT schools in the CITS analyses represent five types of BPS schools: innovation, pilot, Schedule A or mayoral initiative, traditional, and turnaround. For each school type, analyses estimated the impact for a given school type (e.g., innovation schools) and the impact for all other ELT schools (e.g., all ELT schools minus innovation schools) in the same model. Overall, because the number of schools in each school type category is small with the exception of Schedule A schools (see Table 1), the estimates obtained are highly imprecise.

Table 1. Number of Schools by School Type

School Type	Number of Schools
Innovation	5
Pilot	3
Schedule A	15
Traditional	2
Turnaround	6
Total	**31**

- **Mathematics:** Results suggest that pilot and turnaround schools may have had larger impacts on students' mathematics achievement than other ELT schools; however, this finding cannot be confirmed due to the small number of schools in each of these groups.
    - Pilot and turnaround schools had large impact estimates in Year 1 and Year 2 (9.2 CPI points in Year 1, and 12.5 points in Year 2 for pilot schools; 7.7 points in Year 1, and 9.7 points in Year 2 for turnaround schools). The statistical significance was small for pilot schools though (both estimates were only significant at the 10% level, mainly driven by large standard errors) as compared with turnaround schools whose estimates were significant at the 5% level.
    - These results suggest that pilot and turnaround schools *might* be obtaining greater achievement gains in math than other ELT schools. However, the model could not statistically distinguish between the gains from pilot schools versus all other ELT schools (including turnaround schools), or from turnaround schools versus all other ELT schools. In other words, the results could suggest that students in turnaround and pilot schools might be obtaining larger gains, but the analyses cannot say for certain because the confidence interval of these estimates overlap (mostly because of low precision of our estimates due to the small sample size).
    - For instance, students in turnaround schools scored 7.7 CPI points higher in math than students in non-ELT schools. The same model estimated that all other ELT schools scored 2.7 CPI points higher than non-ELT schools. However, due to the level of error in the estimates, the analyses could not determine whether these two values are actually different.
- **ELA:** Results suggest that turnaround schools may have had larger impacts on students' ELA achievement than other ELT schools; however, this finding cannot be confirmed due to the small number of schools in this group.
    - Turnaround schools had a large and significant impact estimate in Year 1, scoring 6 CPI points higher than non-ELT schools. This result could suggest that turnaround schools obtained larger gains in ELA than other ELT schools; however, analyses could not distinguish it from the estimate obtained for all other ELT schools.

The exploratory analysis shows some evidence that students in pilot and turnaround schools implementing ELT programs may have benefited more in math than students in other schools implementing ELT. Students in turnaround schools may have benefited more in ELA than students in other types of ELT schools. The estimates, however, are imprecise due to the small sample size and, as a result, the differences observed could be nonexistent. Further analysis is needed to confirm these exploratory results.

Descriptive Analyses

Methodology

During the first year of the research–practitioner partnership, AIR and BPS researchers interviewed teachers and administrators at BPS ELT schools, gathering qualitative data on how schools used time, as well as perceived strengths and challenges of implementing ELT. In order

to further clarify how schools with longer days in BPS used their time, BPS staff members contacted ELT schools in October 2016 to provide total minutes per week spent on the following activities[10] during the previous school year (2015–16):

- ELA (core time)
- Mathematics (core time)
- Science
- Social Studies
- Enrichment
- ELA intervention
- Mathematics intervention
- Teacher individual planning time
- Teacher collaborative planning time
- Teacher professional development time

Using these data, ELT schools within each time-use category were grouped by percentile based on the raw minutes per week spent on each category. For example, schools in the bottom 25% spent between 0 minutes and 30 minutes per week on ELA intervention, while schools in the top 75% spent between 163 minutes and 275 minutes per week on ELA intervention. Schools were also grouped by the proportion of time spent on each activity as a part of that school's total minutes per week (e.g., schools in the bottom 25% spent between 0% and 2% of the weekly instructional time on ELA intervention, whereas schools in the top 75% spent between 8% and 13% of the weekly instructional time on ELA intervention). AIR then conducted descriptive analyses looking at CPI scores in 2015–16 by group percentiles, as well as correlations. To account for school's prior achievement levels, multivariate regressions were run using 2014–15 CPI scores.

Use of Time Findings

In general, ELT activities were more likely to be positively correlated with higher CPI scores in mathematics than in ELA, even when the activities the school implemented were focused on ELA.

- Schools that spent more time in **math interventions** earned higher CPI scores in both ELA and mathematics. Controlling for prior achievement, there remains a positive relationship (significant at the 10% level) between time in math interventions and mathematics CPI scores (both looking at the proportion of time and raw minutes per week).
    - Accounting for prior achievement, there remains a positive relationship between mean CPI scores in mathematics and schools that spent a higher proportion of time on math when considering both time in core and math interventions.

[10] Analyses run during this grant were focused only on data related to ELA and mathematics (both core and intervention time) and enrichment due to time constraints.

- Schools that spent more time in **ELA interventions** also earned higher mathematics CPI scores (significant at the 5% level) for both proportion and minutes spent.
    - Accounting for prior achievement, there remains a positive relationship between mean CPI scores in mathematics and schools that spent a higher proportion of time or a higher number of minutes on ELA when considering both time in core and ELA interventions.
    - This could indicate that improvements in reading comprehension contributes to improvements in understanding mathematics problems that improve mathematics achievement.

Schools that spent more time in **enrichment** activities scored higher in ELA and mathematics; however, these differences disappear after accounting for prior achievement in a regression.

Implications of Academic Achievement Findings

These academic achievement findings indicate that ELT is having a positive impact on student achievement in the district. The results are especially positive for Black and Hispanic students, who make up a large percentage of the student population in many Boston schools. Overall, results have been most promising for mathematics. This was also seen in the descriptive results, in which ELT schools that integrated intervention time into their day see more growth in their students' mathematics performance than schools that have not used ELT to add intervention time. This is true both for schools with intervention time focused on ELA and/or mathematics: results appear to be most promising for mathematics achievement.

Future Research on Academic Achievement and ELT

BPS and AIR are exploring multiple avenues for continued research on the relationship between students' academic achievement and ELT. Potential future research topics include:

- Expanding the CITS analyses to incorporate Year 2 data for schools that began ELT in 2015–16.
- Conducting CITS analyses on those schools that will begin ELT in the 2017–18 school year.
- Continuing to collect time use data to be able to conduct multiyear analyses on how school trends in time use are related to student performance.
- Conducting analyses comparing ELT schools to determine the highest performing in BPS.
- Conducting analyses comparing ELT schools to determine the schools most improved since beginning ELT.
- Determining whether there are additional measures of student performance that are common across enough schools to be used in analyses.

Attendance Impact Analyses

CITS analysis was performed to see whether a longer day impacted student attendance; analysis used the same ELT implementation cohorts, non-ELT schools from the academic analyses, controls for school demographics, pre-ELT academic trends, and grade levels served.

This analysis showed neither a positive nor a negative impact of expanded time on student attendance. Although the analysis did not demonstrate an increase in attendance, consistent attendance levels between ELT and non-ELT schools is important to note for two reasons. First, students who attend ELT schools are receiving increased academic and enrichment time. This means that students attending a similar number of days of school at an ELT school are actually receiving more instruction than students in a non-ELT school. Second, concerns about start and end times at ELT schools, as well as teacher and student fatigue due to the longer school day, do not appear to have had a negative impact on student attendance, as was feared. However, it is possible there is an increase in the number of students released early to parents who arrive before the end of the ELT school day, and that these early releases are not being accurately tracked or reported. BPS staff are concerned about this possibility and feel it is worth further investigation.

Student and Teacher Climate Descriptive Analyses

AIR and BPS hoped to determine whether implementing ELT was related to how students and teachers experienced the climate in their schools. To attempt to do this, researchers used data from annual BPS district student and teacher climate surveys from 2011–12 through 2015–16; this was the time period during which the survey changes were minimal and comparisons were recommended by district staff. Due to the limited time period in which data are available for this measure, when compared with achievement and attendance data, only descriptive analyses were run using the climate data. The descriptive analyses did not show patterns among responses from staff and students in ELT schools or any differences in patterns between responses from staff and students in ELT and non-ELT schools.

Parent Satisfaction

Parent satisfaction, although critically important to the success of ELT, was the project's most challenging success metric to operationalize. Finding data, such as the district's parent survey was challenging due to low response rates. Researchers were unsuccessful in finding a data set that fit this requirement. This section identifies the complexities of measuring parent satisfaction regarding ELT in the Boston context to date, as well as recommendations for future data collection.

Parent Climate Surveys

Similar to teacher and student climate surveys, the district's parent climate survey asks parents' opinions from perceptions of principal performance to parent engagement. Although these components were relevant to the project, the consistently low response rate of parents on the climate survey deemed the data unusable. From school year 2011–12 through school year 2015–

16, the response rate of the survey has averaged 21.6%. This is too low for the district to be comfortable reporting or sharing results.

School Choice

When selecting prospective schools, a parent considers a multitude of factors to determine the school that is a best fit for both the student's and the family's needs. School start and end times, the length of the school day, and the quality and length of before and after school services are all key factors in school choice decisions and central components of ELT models. Unfortunately, parent satisfaction with ELT cannot be explicitly measured through the school choices that families made at the time of registration or through the schools that students ultimately were assigned to for several reasons.

First, major changes to both Boston's public school choice model and school times have occurred several times since ELT was first implemented. Concerning school choice, Boston is currently finishing its fourth school "choice season" under a new Home-Base Assignment plan. Initiated by shared parent and district concerns regarding the cost and distance of transportation, this plan shifted the choice lists available to families from a "zone" model to customized choice lists based on a family's home address. Under the zone model, the city was divided into three catchment areas, but now parents have their own unique choice list designed to provide both schools in close proximity to home and high-quality as well as citywide options. While the 1988 zone model was a stable geographic catchment area in which parents could choose schools; the home-base model decreased the number of overall school choices and often provides choices of schools previously unavailable to those families who lived near the old zone catchment boundaries. Changes in start and end times to schools within the district, as well as the implementation of the ELT models, have also occurred over time within the district. Therefore, the combination of change in policy and change in school times (before even considering more specific school-based changes or outcomes) makes it challenging to measure the relationship between parent satisfaction and school choice.

A second challenge is that the scale and complexity of the Boston assignment system means that student assignment is not always a direct reflection of either a parent's top preference or preferences at all. First, given the aforementioned home-based plan, the information that parents need to gather to navigate their choice lists is more complex than in the past. Because the current choice plan offers unique lists by address, there is an increased responsibility on individual families to research schools and understand the choices they are making during the registration process. Extensive research also documents how information is stratified by access to technology and social networks, which in Boston results in some families with higher quantity and quality of information regarding school options. Second, given the high demand for admission to some schools, students are not always assigned to their top-choice school because there are not enough seats for all students who want that school. Therefore a student's number in the lottery system, the family's timing of registration, and the limited number of seats in highly selected schools all may result in administrative assignment to a school that a family did not prioritize as a choice. For these reasons, neither family choices nor assignments are valid measures of parent engagement or satisfaction.

While the complexity of the choice system, issues regarding information stratification, and the limited availability of seats in highly selected schools all complicate our measurement of parent satisfaction, a third challenge in measuring parent satisfaction through school choice is the diverse number of factors that families must negotiate when making school choices. As a result, we cannot reliably disentangle their school preferences (and the ways in which they do or do not reflect elements of ELT) from other preferences that parents have when making school choices. A survey of 1,886 parents who registered for school in January 2017 revealed that many factors are important to parents when making their school choices. Looking specifically at ELT-related elements, 60% checked that school start and end times were important to them, and 48% checked before and after school programming when making school choices. However, other popular indicators included proximity to home (72%), reputation (61%), and academic performance (65%). Given that parents' preferences are often reflective of their ideals rather than their actual choice lists, their preferences for ELT variables does not necessarily indicate a preference for ELT explicitly. Further, these preferences cannot reliably indicate whether or not these preferences translated into choices that reflect these factors given the gaps in information that impact parental awareness of school-specific information.

Future Research on Parent Satisfaction and ELT

Parent engagement, voice, and participation in ELT schools are critical to its success. Therefore, future research is needed to provide valuable insights into how parents perceive expanded learning time and how ELT schools can better engage parents and provide desired services. Some potential strategies that researchers and practitioners could use to begin to measure the relationship between ELT and parent satisfaction include:

- Qualitative data (field notes and/or transcripts), work products, and/or parent surveys from ELT design meetings that involve families.
- Surveys of parents in schools that are in the process of deciding whether to implement ELT regarding their interest, awareness, and satisfaction with school changes and opportunities related to ELT programming.
- Pre- and post-surveys from families in the years before and after transition to ELT regarding their interest, awareness, and satisfaction with school changes and opportunities related to ELT programming.
- Intake surveys from families who newly enter ELT schools (either at entrance grades or transfer in) regarding the factors that influenced their assignment there.
- Longitudinal examination of student attrition or yield data at ELT schools, including before their transition to the ELT model.

Appendix A: Impacts on Achievement and Attendance

AIR conducted inferential and descriptive analyses to assess the impact of implementing an expanded learning time (ELT) program on achievement, attendance, and school climate. This appendix provides a more detailed description of the achievement and attendance inference analyses.

To look at the impact of implementing an ELT program on student achievement, American Institutes for Research (AIR) conducted a comparative interrupted time series (CITS) analysis in which we looked at Year 1 and Year 2 impacts on mathematics, English language arts (ELA), and science as measured by grade-level school composite performance index (CPI) scores provided by Massachusetts State.[11] CITS compares outcome scores between ELT and non-ELT schools pre- and post-program implementation, accounting for preintervention outcome trends and school composition. We conducted a confirmatory analysis with all schools with available data across the district. We also conducted a sensitivity analysis using a propensity score matched comparison sample. Additional subgroup analyses were conducted by student demographic groups (e.g., gender, race, English language learners [ELLs], students with disabilities, and economically disadvantaged); we also conducted an exploratory analysis by school type (e.g., turnaround, pilot, innovation, and Schedule A).

Similarly, AIR evaluated the impact of implementing ELT on student attendance as measured by grade-level school attendance rate provided by the state for all students using a CITS analysis. This analysis was conducted for all students in Grades 1 through 12. No further subgroup analysis was conducted for the attendance outcome.

Data

We used publicly available longitudinal school-level achievement, attendance, enrollment, and demographic data for our analyses, accessed from the Massachusetts Department of Elementary and Secondary Education's website. Available outcome data (see Table A1) was disaggregated by grade. Achievement data also were available by student demographic subgroup (see Table A2).

Table A1. Data Availability and Use

Outcome	Years Used
Achievement: MCAS and PARCC (Mathematics, Reading and Science)[a]	2005–06 to 2015–16
Attendance[b]	2005–06 to 2015–16

Note. MCAS = Massachusetts Comprehensive Assessment System; PARCC = Partnership for Assessment of Readiness for College and Careers
[a] For all students and by demographic subgroups.
[b] For all students.

[11] Outcome achievement data was available for Grades 3–8 and 10 for mathematics and ELA subjects, and for Grades 5, 8 and 10 for science.

Table A2. Data Availability and Use

Achievement Outcome by Demographic Subgroups	Years Used
Low Income (or Economically Disadvantaged)[a]	2005–06 to 2015–2016
Gender	2005–06 to 2015–2016
Race/Ethnicity	2005–06 to 2015–2016
English Language Learner	2005–06 to 2015–2016
Special Education	2005–06 to 2015–2016

[a] Achievement data for low income students was available from 2005–06 to 2013–2014 and for economically disadvantaged students from 2014–15 to 2015–2016.

Analytical Approach

Comparative Interrupted Time Series Analyses

The CITS approach used leverages from the varied implementation schedules for ELT schools and examined effects according to years of implementation. The general multiple baseline CITS model used is:

$$Y_{ijt} = \beta_0 + \boldsymbol{\beta}_1 \mathbf{ELT}_j + \beta_2 Year_t + \boldsymbol{\beta}_3 \mathbf{Post}_t + \boldsymbol{\beta}_4 \mathbf{Imp}_t + \boldsymbol{\beta}_5 \mathbf{ELT}_j \times Year_t + \boldsymbol{\beta}_6 \mathbf{W}_{jt} + \nu_j + \upsilon_{jt} + \varepsilon_{ijt} \quad (1)$$

where Y_{ijt} is the average outcome for students in grade *i* in school *j* at time *t*; $\mathbf{ELT}_j$ is a vector of ELT cohort indicators (seven cohorts for achievement and attendance outcomes); $Year_t$ is a continuous year variable that runs from 2006 through 2016; $\mathbf{Post}_t$ is a vector of post-ELT implementation year indicators (one for each year after the first cohort began implementing ELT); $\mathbf{Imp}_t$ is a vector of implementation indicators (one for each year of ELT implementation); $\mathbf{W}_{jt}$ is a vector of school-level characteristics (e.g., school type, demographic composition, enrollment).[12] $\boldsymbol{\beta}_4$ is the vector of impacts for ELT schools, which represent the effects of ELT for each year of implementation (i.e., the effect of implementing ELT for one year, two years … *t* years).

For these CITS analyses, we required at least three years of pre-ELT implementation data to include an ELT school.[13] The comparison group was composed of all non-ELT schools that were still active in 2016 and had outcome data (a total of 63 schools for the achievement analysis and the same number and schools for the attendance analysis with the exception of one fewer school). Accordingly, for student achievement and attendance outcomes, we applied this CITS framework to seven ELT cohorts, seen in Table A3. See Appendix B (Table B1) for achievement results and (Table B8) for attendance results.

[12] School demographic covariates included percentages of ELLs in school, students with disabilities, and economically disadvantages students, and gender and racial composition. The model also included grade fixed effects.

[13] For the achievement analysis, the exception was one ELT school from Cohort 1 that only had two years of pre-ELT implementation data for science given that in the school years 2005–06 and 2006–07 there was no CPI data for science in Grade 10.

Table A3. ELT Cohorts for Achievement and Attendance Outcomes in CITS

School	Began Implementing	Cohort
Mario Umana Academy	2009–10	C1
New Mission High School	2009–10	C1
William Monroe Trotter	2009–10	C1
Blackstone	2010–11	C2
Dearborn	2010–11	C2
John F Kennedy	2010–11	C2
Orchard Gardens	2010–11	C2
The English High	2010–11	C2
Roger Clap	2011–12	C3
Boston Community Leadership Academy	2012–13	C4
Eliot Elementary	2012–13	C4
John W McCormack	2012–13	C4
Washington Irving Middle	2012–13	C4
Mattahunt	2013–14	C5
John Winthrop	2014–15	C6
Joseph Lee	2014–15	C6
Mildred Avenue K-8	2014–15	C6
Sarah Greenwood	2014–15	C6
William Ellery Channing	2014–15	C6
Curtis Guild	2015–16	C7
Harvard-Kent	2015–16	C7
Henry Grew	2015–16	C7
Higginson/Lewis K-8	2015–16	C7
Jackson Mann	2015–16	C7
James J Chittick	2015–16	C7
James Otis	2015–16	C7
Joseph P Manning	2015–16	C7
Mather	2015–16	C7
Maurice J Tobin	2015–16	C7
Michael J Perkins	2015–16	C7
Warren-Prescott	2015–16	C7
~~Snowden International High~~	~~2015-16~~	~~C7~~
~~Dorchester Collegiate Academy~~	~~2015-16~~	~~C7~~

Note. Snowden International High and Dorchester Collegiate Academy were removed from the analyses due to limited time added to the school day and self-identification as non-ELT schools. These schools were also not included as comparison schools.

Constructing Matched Comparison Groups

As part of our sensitivity analyses for achievement impacts, we constructed a matched comparison group for each of the participating ELT cohorts. AIR used propensity score matching using the last four years of preintervention achievement in ELA and mathematics along with school type (i.e., Grades served, 3–5, 3–8, 6–8, and 10) and economically disadvantaged school information to construct the matched comparison groups. Propensity score matching is a statistical technique that estimates the probability of group membership (treatment versus control) based on observed characteristics, and then uses that predicted probability to create a control group similar to the treatment group on these observed characteristics. A nearest neighbor matching approach was used, with a one-to-one matching ratio without replacement, where ELT schools were only paired with schools that offered same grade spans (e.g., Grades 3–5, Grades 3–8, Grades 6–8, and Grade 10). Furthermore, ELT schools were only paired with non-ELT schools that had at least four years of preintervention outcome data (or all available years) and at least three years of postintervention outcome data (or all available years).[14]

Multiple propensity score models were tested resulting in multiple matched comparison groups. To select the optimum model and matched comparison group, we conducted baseline equivalence analyses to assess balance improvement.[15] Specifically, the following measures were used to assess baseline equivalence:

- Average ELA, mathematics, and science CPI scores
- Percent of economically disadvantaged students[16]
- Percent ELLs
- Percent students with disabilities
- Enrollment (school size)

[14] Due to the small number of schools in some of the cohorts (see Table A3), schools in Cohorts 1 and 2 were grouped together in the propensity score matching model. The last four years of preintervention achievement data were determined based on the earliest implementing cohort (e.g. Cohort 1 in this example). Once the model was run and propensity scores were calculated, ELT schools in this grouped cohort were then matched to non-ELT schools based on their propensity scores out of a pool of 63 non-ELT schools (one-to-one school matching with no replacement among schools with same grades offered). Similarly, Cohorts 3 and 4 were grouped together, as well as Cohorts 5 and 6, and the same process was applied whereby each ELT school in the grouped cohort was matched to one of the 63 non-ELT schools. Cohort 7 was matched separately, given the fact it was the largest cohort, and following the same procedure. This approach resulted in some non-ELT schools being matched to multiple ELT schools from different grouped cohorts.

[15] The selected propensity score model included an averaged ELA CPI score across the four years prior to intervention both for mathematics and ELA, and an average annual slope change in CPI across the four years prior to intervention for the same two subjects. Additionally, the model included indicators for Grades served: 3–5, 3–8, 6–8 and 10.

[16] Due to changing variables used for data collection and reporting over the period of data analyzed, the variable used to determine percentage of economically disadvantaged students combined data from the "low income" variable for years 2006 through 2014, and the "economic disadvantage" variable for 2015 and 2016. "Low income" indicates the percent of enrolled students who meet any one of the following definitions of low income: is eligible for free or reduced-price lunch, receives Transitional Aid to Families benefits, or is eligible for food stamps. The inclusion of a student in the "economic disadvantage" group is based on the student's participation in one or more of the following state-administered programs: Supplemental Nutrition Assistance Program (SNAP), Transitional Assistance for Families with Dependent Children (TAFDC), Department of Children and Families' (DCF) foster care program, and MassHealth (Medicaid).

- Percent female
- Percent White, Black, Hispanic
- Grade levels served

See Appendix C for baseline equivalence analysis tables across years, by school-level characteristic for the selected matched comparison group. At a given year, the tables compare the difference in a given school-level characteristic between ELT schools that have not started to implement the program at that time and both (a) all non-ELT schools active by 2016, and (b) all non-ELT matched comparison schools whose matched ELT school has not yet started to implement the program. The study calculated effect-size differences (i.e., differences in standard deviations) to evaluate improvements in baseline equivalence, following the standards of the What Works Clearinghouse (WWC, 2011).[17] For continuous variables, effect size differences are computed using standardized mean differences (Hedges' *g* corrected for small-sample bias). For dichotomous variables, Cox log odds ratios are calculated as suggested by WWC, thus making them comparable with standardized mean differences for continuous variables.

The matched comparison group identified was then used to run a sensitivity analysis to measure the impact of implementing ELT on achievement for all students. Model specification 1 was used with the addition of cohort indicators for matched schools. Despite the smaller statistical power, results obtained were similar in magnitude and statistical significance as the ones observed in the main analysis that used all non-ELT schools active by 2016 as comparison group.[18] See Appendix B (Table B1) for results.

ELT Student and School Subgroup Analyses

AIR examined the effect of ELT on achievement by student subgroups, including economically disadvantaged students, students in special education, ELL status, race/ethnicity, and gender. To examine the effects of these subgroups, we used the same CITS model indicated in model specification (1) but used the outcomes for that specific subgroup. Outcomes for some of the student subgroups were missing for some schools based on the school's student population. See Appendix B (Tables B2, B3, and B4) for results.

AIR also conducted an exploratory analysis on the effect of ELT on achievement by school type (see Table A4) by conducting separate CITS analyses for each subgroup using the following variation of the model specification 1:

[17] What Works Clearinghouse. (2011). *What Works Clearinghouse: Procedures and standards handbook* (Version 3.0). Washington, DC: U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf

[18] For mathematics, the estimated impact of Year 1 in the main analysis was 3.56 CPI points ($p < 5\%$) and in Year 2 of 5.82 CPI points ($p < 1\%$). Similarly, Year 1 impact estimate using the matched comparison group was 3.04 CPI points ($p < 5\%$) and 5.28 CPI points ($p < 5\%$) in Year 2. For reading, Year 1 and Year 2 estimated impacts were 2.10 CPI points ($p < 10\%$) and 3.09 CPI points ($p < 10\%$), respectively in the main analysis; and 1.74 CPI points ($p = 16\%$) and 2.81 CPI points ($p = 11\%$) using the matched comparison group.

$$Y_{ijt} = \beta_0 + \boldsymbol{\beta}_1\mathbf{ELT}(\mathbf{subgroup\_evaluated})_j + \boldsymbol{\beta}_2\mathbf{ELT}(\mathbf{remaining})_j + \beta_3 Year_t + \boldsymbol{\beta}_4\mathbf{Post}_t + \boldsymbol{\beta}_5\mathbf{Imp}(\mathbf{subgroup})_t + \boldsymbol{\beta}_6\mathbf{Imp}(\mathbf{remaining})_t + \boldsymbol{\beta}_7\mathbf{ELT}(\mathbf{subgroup\_evaluated})_j \times Year_t + \boldsymbol{\beta}_8\mathbf{ELT}(\mathbf{remaining})_j \times Year_t + \boldsymbol{\beta}_9\mathbf{W}_{jt} + \nu_j + \upsilon_{jt} + \varepsilon_{ijt} \quad (2)$$

Where $\mathbf{ELT}(\mathbf{subgroup\_evaluated})_j$ is a vector of ELT cohort indicators for ELT schools in the school type subgroup being evaluated; and $\mathbf{ELT}(\mathbf{remaining})_j$ is the same vector of ELT cohort indicators for all other ELT schools without including schools in the subgroup evaluated. Thus, these two sets of vectors are mutually exclusive and their schools are compared with non-ELT schools that fall under the omitted group. Similarly, $\mathbf{Imp}(\mathbf{subgroup\_evaluated})_t$ is a vector of implementation indicators (one for each year of ELT implementation) for schools in the subgroup type evaluated; whereas $\mathbf{Imp}(\mathbf{remaining})_t$ is the same vector but for all other ELT schools. $\boldsymbol{\beta}_5$ is the vector of impacts for ELT schools in the school type subgroup evaluated and represents effects of ELT for each year of implementation (as compared with non-ELT schools); $\boldsymbol{\beta}_6$ is the same vector of impacts but for all other ELT schools (comparing them to non-ELT schools). See Appendix B (Tables B5, B6, and B7) for results.

Table A4. ELT School Type

Type	Number of ELT Schools
Pilot	3
Turnaround	6
Innovation	5
Traditional – Other	2
Traditional – Schedule A	15

Appendix B: Impact Results

The tables in this appendix show impact and exploratory results of the achievement and attendance analyses described in Appendix A.

Table B1. Achievement Impact Results by Subject Outcome

	English Language Arts (ELA)		Mathematics		Science	
	Main Analysis	Sensitivity Analysis	Main Analysis	Sensitivity Analysis	Main Analysis	Sensitivity Analysis
First Year Impact	2.10*	1.74	3.56**	3.04**	0.31	-0.74
	(1.15)	(1.24)	(1.42)	(1.54)	(1.68)	(1.74)
Second Year Impact	3.09*	2.81	5.82***	5.28**	1.84	0.13
	(1.63)	(1.75)	(2.01)	(2.17)	(2.37)	(2.45)
Observations	3,355	2,104	3,345	2,103	1,222	747

Note. Subject outcome impact estimates are measured in composite performance index (CPI) points. The main analysis uses all non-ELT schools that were still active in 2016 and had outcome data as a comparison group. The sensitivity analysis uses the matched-constructed non-ELT group as a comparison group. Observations refer to number of outcome grade-level records. Standard errors in parentheses.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table B2. Achievement Impact Results by Student Subgroup on English Language Arts (ELA)

	All Students	Asian	African American	Hispanic	White	English Language Learners	Students with Disabilities	Economically Disadvantaged	Females	Males
First Year Impact	2.10*	-2.53	3.08**	2.53*	1.42	1.63	0.98	1.34	1.87	1.88
	(1.15)	(2.76)	(1.35)	(1.42)	(2.48)	(2.14)	(2.40)	(1.23)	(1.18)	(1.31)
Second Year Impact	3.09*	0.41	4.52**	3.95*	2.75	3.97	1.93	2.94*	3.39**	2.54
	(1.63)	(4.98)	(1.87)	(2.03)	(3.55)	(3.05)	(3.37)	(1.74)	(1.68)	(1.85)
Observations	3,355	423	2,145	2,372	8,82	1,614	1,754	3,176	2,929	3,079

Note. Subject outcome impact estimates are measured in composite performance index (CPI) points. Observations refer to number of outcome grade-level records. Standard errors in parentheses.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table B3. Achievement Impact Results by Student Subgroup on Mathematics

	All Students	Asian	African American	Hispanic	White	English Language Learners	Students with Disabilities	Economically Disadvantaged	Females	Males
First Year Impact	3.56**	-3.72	3.95**	3.89**	-1.97	3.82*	2.43	2.76*	2.70*	3.27**
	(1.42)	(2.96)	(1.67)	(1.66)	(2.70)	(2.24)	(2.46)	(1.50)	(1.52)	(1.48)
Second Year Impact	5.82***	-3.09	7.44***	7.65***	-1.99	7.19**	2.37	5.81***	5.66***	5.40**
	(2.01)	(5.31)	(2.30)	(2.36)	(3.94)	(3.15)	(3.39)	(2.11)	(2.16)	(2.11)
Observations	3,345	419	2,137	2,367	872	1,613	1,746	3,166	2,920	3,072

Note. Subject outcome impact estimates are measured in composite performance index (CPI) points. Observations refer to number of outcome grade-level records. Standard errors in parentheses.
* p <0.10, ** p < 0.05, *** p < 0.01.

Table B4. Achievement Impact Results by Student Subgroup on Science

	All Students	Asian	African American	Hispanic	White	English Language Learners	Students with Disabilities	Economically Disadvantaged	Females	Males
First Year Impact	3.56**	-3.72	3.95**	3.89**	-1.97	3.82*	2.43	2.76*	2.70*	3.27**
	(1.42)	(2.96)	(1.67)	(1.66)	(2.70)	(2.24)	(2.46)	(1.50)	(1.52)	(1.48)
Second Year Impact	5.82***	-3.09	7.44***	7.65***	-1.99	7.19**	2.37	5.81***	5.66***	5.40**
	(2.01)	(5.31)	(2.30)	(2.36)	(3.94)	(3.15)	(3.39)	(2.11)	(2.16)	(2.11)
Observations	3,345	419	2,137	2,367	872	1,613	1,746	3,166	2,920	3,072

Note. Subject outcome impact estimates are measured in composite performance index (CPI) points. Observations refer to number of outcome grade-level records. Standard errors in parentheses.
* p < 0.10; ** p < 0.05; *** p < 0.01.

Table B5. Achievement Exploratory Results by School Type on ELA

	Innovation Schools	Pilot Schools	Schedule A Schools	Traditional Schools	Turnaround Schools
First Year Impact (Subgroup Evaluated)	0.42	3.18	-0.10	8.09**	6.03**
	(2.90)	(4.56)	(1.52)	(3.59)	(2.70)
Second Year Impact (Subgroup Evaluated)	1.86	6.82	-0.97	8.98**	3.48
	(3.33)	(5.24)	(3.05)	(3.82)	(3.25)
First Year Impact (Remaining)	2.38*	2.05*	4.39***	1.44	1.35
	(1.23)	(1.18)	(1.60)	(1.20)	(1.25)
Second Year Impact (Remaining)	3.32*	2.71	4.95***	2.00	3.10*
	(1.84)	(1.71)	(1.854)	(1.75)	(1.83)
Observations	3,355	3,355	3,355	3,355	3,355

Note. Subject outcome impact estimates are measured in composite performance index (CPI) points. In each column, a different school type subgroup is being evaluated. The impact estimates for the remaining group corresponds to the impact of all ELT schools excluding the subgroup being evaluated. Observations refer to number of outcome grade-level records. Standard errors in parentheses.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table B6. Achievement Exploratory Results by School Type on Mathematics

	Innovation Schools	Pilot Schools	Schedule A Schools	Traditional Schools	Turnaround Schools
First Year Impact (Subgroup Evaluated)	3.94	9.16*	0.12	9.33**	7.72**
	(3.67)	(5.56)	(1.90)	(4.46)	(3.29)
Second Year Impact (Subgroup Evaluated)	3.97	12.46*	-0.92	12.95***	9.69**
	(4.23)	(6.38)	(3.80)	(4.74)	(3.95)
First Year Impact (Remaining)	3.46**	3.19**	7.30***	2.93**	2.73*
	(1.52)	(1.45)	(2.01)	(1.48)	(1.53)
Second Year Impact (Remaining)	6.48***	5.30**	9.12***	4.50**	4.79**
	(2.26)	(2.10)	(2.33)	(2.17)	(2.25)
Observations	3,345	3,345	3,345	3,345	3,345

Note. Subject outcome impact estimates are measured in composite performance index (CPI) points. In each column, a different school type subgroup is being evaluated. The impact estimates for the remaining group corresponds to the impact of all ELT schools excluding the subgroup being evaluated. Observations refer to number of outcome grade-level records. Standard errors in parentheses.
* $p < 0.10$; ** $p < 0.05$, *** $p < 0.01$.

Table B7. Achievement Exploratory Results by School Type on Science

	Innovation Schools	Pilot Schools	Schedule A Schools	Traditional Schools	Turnaround Schools
First Year Impact (Subgroup Evaluated)	-4.20	-1.83	-0.68	3.51	4.28
	(4.42)	(5.78)	(2.30)	(4.91)	(4.11)
Second Year Impact (Subgroup Evaluated)	1.56	2.89	-4.85	8.40	-1.04
	(5.05)	(6.84)	(4.75)	(5.24)	(4.96)
First Year Impact (Remaining)	0.90	0.50	0.92	-0.15	-0.52
	(1.79)	(1.75)	(2.34)	(1.77)	(1.81)
Second Year Impact (Remaining)	1.24	1.46	3.54	0.33	2.49
	(2.65)	(2.53)	(2.71)	(2.59)	(2.64)
Observations	1,222	1,222	1,222	1,222	1,222

Note: Subject outcome impact estimates are measured in composite performance index (CPI) points. In each column, a different school type subgroup is being evaluated. The impact estimates for the remaining group corresponds to the impact of all ELT schools excluding the subgroup being evaluated. Observations refer to number of outcome grade-level records. Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B8. Attendance Impact Results

	Attendance Analysis
First Year Impact	0.04
	(0.05)
Second Year Impact	0.00
	(0.06)
Observations	5,391

Note. Outcome impact estimates are attendance rate points, where 0.01 would refer to 1 percentage point. Observations refer to number of outcome grade-level records. Standard errors in parentheses.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Appendix C: Baseline Equivalence Analysis

The tables in this appendix cover baseline equivalence analysis across years and by school-level characteristic for the selected matched comparison group.

Table C1. Baseline Equivalence Analysis: School Average English Language Arts Composite Performance Index (CPI)

Year	ELT Schools	Non-ELT Schools (All Active BPS Schools by 2016)			Non-ELT Matched Comparison Schools from Selected Model			Number of Schools		
	Mean (1)	Mean (2)	Raw Mean Difference (1) - (2)	Standardized Difference[a]	Mean (3)	Raw Mean Difference (1) - (3)	Standardized Difference[a]	ELT Schools	Non-ELT Schools	Non-ELT Matched Schools
2006	65.64	72.83	-7.19	-0.70 [d]	68.90	-3.26	-0.34 [d]	31	63	24
2007	67.30	75.19	-7.89	-0.80 [d]	71.49	-4.19	-0.47 [d]	31	63	24
2008	66.53	75.21	-8.68	-0.85 [d]	70.45	-3.91	-0.39 [d]	31	63	24
2009	69.50	77.02	-7.53	-0.76 [d]	72.96	-3.46	-0.33 [d]	31	63	24
2010	70.36	77.48	-7.12	-0.76 [d]	73.25	-2.88	-0.31 [d]	28	63	23
2011	71.64	77.54	-5.90	-0.59 [d]	72.58	-0.94	-0.10 [c]	23	63	21
2012	68.12	76.97	-8.86	-0.80 [d]	69.74	-1.63	-0.15 [c]	22	63	20
2013	67.23	77.98	-10.75	-1.00 [d]	68.24	-1.01	-0.10 [c]	18	63	17
2014	68.74	77.61	-8.88	-0.84 [d]	70.14	-1.40	-0.14 [c]	17	63	16
2015	71.06	78.56	-7.51	-0.69 [d]	72.14	-1.08	-0.12 [c]	12	63	12

[a] Effect size (ES) or standardized mean difference.
[b] Based on What Works Clearinghouse (WWC) standards, a standardized difference of ($|0.00| \leq$ ES Difference $\leq |0.05|$) indicates "Satisfies baseline equivalence."
[c] Based on WWC standards, a standardized difference of ($|0.05| <$ ES difference $\leq |0.25|$) indicates "Statistical adjustment required to satisfy baseline equivalence."
[d] Based on WWC standards, a standardized difference of (ES difference $> |0.25|$) indicates "Does not satisfy baseline equivalence."

Table C2. Baseline Equivalence Analysis: School Average Mathematics CPI

Year	ELT Schools	Non-ELT Schools (All Active BPS Schools by 2016)			Non-ELT Matched Comparison Schools from Selected Model			Number of Schools		
	Mean (1)	Mean (2)	Raw Mean Difference (1) - (2)	Standardized Difference[a]	Mean (3)	Raw Mean Difference (1) - (3)	Standardized Difference[a]	ELT Schools	Non-ELT Schools	Non-ELT Matched Schools
2006	56.80	65.74	-8.94	-0.68 [d]	60.86	-4.06	-0.32 [d]	31	63	24
2007	58.91	68.99	-10.08	-0.79 [d]	64.10	-5.19	-0.43 [d]	31	63	24
2008	61.01	69.89	-8.88	-0.70 [d]	63.85	-2.85	-0.22 [c]	31	63	24
2009	60.83	70.44	-9.61	-0.79 [d]	66.26	-5.43	-0.41[d]	31	63	24
2010	65.84	72.40	-6.55	-0.56 [d]	66.71	-0.87	-0.07 [c]	28	63	23
2011	65.67	72.09	-6.41	-0.57 [d]	66.32	-0.65	-0.05 [c]	23	63	21
2012	63.84	71.09	-7.24	-0.63 [d]	63.71	0.14	0.01 [b]	22	63	20
2013	64.41	71.80	-7.40	-0.67 [d]	64.14	0.27	0.03 [b]	18	63	17
2014	67.10	72.75	-5.65	-0.48 [d]	65.19	1.91	0.17 [c]	17	63	16
2015	69.52	71.70	-2.19	-0.19 [c]	68.23	1.29	0.17 [c]	12	63	12

[a] Effect size (ES) or standardized mean difference.
[b] Based on What Works Clearinghouse (WWC) standards, a standardized difference of ($|0.00| \leq$ ES Difference $\leq |0.05|$) indicates "Satisfies baseline equivalence."
[c] Based on WWC standards, a standardized difference of ($|0.05| <$ ES difference $\leq |0.25|$) indicates "Statistical adjustment required to satisfy baseline equivalence."
[d] Based on WWC standards, a standardized difference of (ES difference $> |0.25|$) indicates "Does not satisfy baseline equivalence."

Table C3. Baseline Equivalence Analysis: School Average Science CPI

Year	ELT Schools	Non-ELT Schools (All Active BPS Schools by 2016)			Non-ELT Matched Comparison Schools from Selected Model			Number of Schools		
	Mean (1)	Mean (2)	Raw Mean Difference (1) - (2)	Standardized Difference[a]	Mean (3)	Raw Mean Difference (1) - (3)	Standardized Difference[a]	ELT Schools	Non-ELT Schools	Non-ELT Matched Schools
2006	48.36	59.74	-11.37	-0.86 [d]	53.30	-4.94	-0.37 [d]	31	63	24
2007	48.85	61.03	-12.18	-0.89 [d]	54.66	-5.82	-0.44 [d]	31	63	24
2008	46.53	56.36	-9.83	-0.77 [d]	51.82	-5.29	-0.45 [d]	31	63	24
2009	48.73	60.45	-11.72	-0.94 [d]	54.50	-5.77	-0.48 [d]	31	63	24
2010	51.04	60.08	-9.04	-0.68 [d]	54.80	-3.76	-0.29 [d]	28	63	23
2011	49.18	58.31	-9.13	-0.67 [d]	50.17	-0.99	-0.08 [c]	23	63	21
2012	49.06	58.86	-9.80	-0.77 [d]	52.00	-2.95	-0.26 [d]	22	63	20
2013	50.88	60.86	-9.98	-0.81 [d]	52.98	-2.11	-0.22 [c]	18	63	17
2014	54.72	60.81	-6.09	-0.44 [d]	53.27	1.45	0.13 [c]	17	63	16
2015	52.53	59.38	-6.85	-0.51 [d]	50.70	1.83	0.16 [c]	12	63	12

[a] Effect size (ES) or standardized mean difference.
[b] Based on What Works Clearinghouse (WWC) standards, a standardized difference of ($|0.00| \leq$ ES Difference $\leq |0.05|$) indicates "Satisfies baseline equivalence."
[c] Based on WWC standards, a standardized difference of ($|0.05| <$ ES difference $\leq |0.25|$) indicates "Statistical adjustment required to satisfy baseline equivalence."
[d] Based on WWC standards, a standardized difference of (ES difference $> |0.25|$) indicates "Does not satisfy baseline equivalence."

Table C4. Baseline Equivalence Analysis: Percent Economically Disadvantaged Students (School Averages)

Year	ELT Schools	Non-ELT Schools (All Active BPS Schools by 2016)			Non-ELT Matched Comparison Schools from Selected Model			Number of Schools		
	Mean (1)	Mean (2)	Raw Mean Difference (1) - (2)	Standardized Difference[a]	Mean (3)	Raw Mean Difference (1) - (3)	Standardized Difference[a]	ELT Schools	Non-ELT Schools	Non-ELT Matched Schools
2006	77.91	74.30	3.60	0.28 [d]	77.69	0.21	0.02 [b]	31	63	24
2007	80.22	73.63	6.59	0.48 [d]	76.72	3.50	0.27 [d]	31	63	24
2008	79.01	71.94	7.07	0.54 [d]	75.74	3.27	0.28 [d]	31	63	24
2009	81.18	72.81	8.36	0.59 [d]	76.79	4.39	0.34 [d]	31	63	24
2010	83.01	73.75	9.27	0.64 [d]	77.46	5.56	0.43 [d]	28	63	23
2011	78.92	73.73	5.19	0.35 [d]	77.30	1.63	0.12 [c]	23	63	21
2012	72.68	70.65	2.03	0.14 [c]	73.82	-1.14	-0.09 [c]	22	63	20
2013	73.91	71.88	2.03	0.16 [c]	75.02	-1.11	-0.10 [c]	18	63	17
2014	80.81	77.35	3.47	0.27 [d]	79.91	0.90	0.07 [c]	17	63	16
2015	54.39	47.94	6.45	0.58 [d]	51.79	2.60	0.20 [c]	12	63	12

[a] Effect size (ES) or standardized mean difference.
[b] Based on What Works Clearinghouse (WWC) standards, a standardized difference of ($|0.00| \leq$ ES Difference $\leq |0.05|$) indicates "Satisfies baseline equivalence."
[c] Based on WWC standards, a standardized difference of ($|0.05| <$ ES difference $\leq |0.25|$) indicates "Statistical adjustment required to satisfy baseline equivalence."
[d] Based on WWC standards, a standardized difference of (ES difference $> |0.25|$) indicates "Does not satisfy baseline equivalence."

Table C5. Baseline Equivalence Analysis: Percent English Language Learners (School Averages)

Year	ELT Schools	Non-ELT Schools (All Active BPS Schools by 2016)			Non-ELT Matched Comparison Schools from Selected Model			Number of Schools		
	Mean (1)	Mean (2)	Raw Mean Difference (1) - (2)	Standardized Difference[a]	Mean (3)	Raw Mean Difference (1) - (3)	Standardized Difference[a]	ELT Schools	Non-ELT Schools	Non-ELT Matched
2006	19.91	15.46	4.45	0.33 [d]	16.08	3.84	0.27 [d]	31	63	24
2007	22.37	17.37	5.00	0.33 [d]	17.70	4.68	0.29 [d]	31	63	24
2008	22.68	17.76	4.92	0.33 [d]	18.67	4.01	0.26 [d]	31	63	24
2009	22.14	17.11	5.04	0.34 [d]	18.00	4.15	0.26 [d]	31	63	24
2010	24.03	18.61	5.41	0.36 [d]	19.74	4.29	0.27 [d]	28	63	23
2011	30.13	27.16	2.97	0.17 [c]	28.67	1.46	0.08 [c]	23	63	21
2012	31.94	30.00	1.94	0.11 [c]	30.78	1.16	0.07 [c]	22	63	20
2013	31.76	29.88	1.88	0.11 [c]	31.79	-0.03	-0.00 [b]	18	63	17
2014	31.15	29.49	1.66	0.09 [c]	30.56	0.60	0.03 [b]	17	63	16
2015	31.04	29.76	1.28	0.07 [c]	34.61	-3.57	-0.18 [c]	12	63	12

[a] Effect size (ES) or standardized mean difference.
[b] Based on What Works Clearinghouse (WWC) standards, a standardized difference of ($|0.00| \leq$ ES Difference $\leq |0.05|$) indicates "Satisfies baseline equivalence."
[c] Based on WWC standards, a standardized difference of ($|0.05| <$ ES difference $\leq |0.25|$) indicates "Statistical adjustment required to satisfy baseline equivalence."
[d] Based on WWC standards, a standardized difference of (ES difference $> |0.25|$) indicates "Does not satisfy baseline equivalence."

Table C6. Baseline Equivalence Analysis: Percent Students with Disabilities (School Averages)

Year	ELT Schools	Non-ELT Schools (All Active BPS Schools by 2016)			Non-ELT Matched Comparison Schools from Selected Model			Number of Schools		
	Mean (1)	Mean (2)	Raw Mean Difference (1) - (2)	Standardized Difference[a]	Mean (3)	Raw Mean Difference (1) - (3)	Standardized Difference[a]	ELT Schools	Non-ELT Schools	Non-ELT Matched Schools
2006	19.68	20.70	-1.02	-0.08 [c]	20.99	-1.31	-0.10 [c]	31	63	24
2007	20.96	21.41	-0.45	-0.03 [b]	22.25	-1.29	-0.10 [c]	31	63	24
2008	21.08	21.70	-0.62	-0.05 [b]	22.52	-1.44	-0.12 [c]	31	63	24
2009	20.83	22.41	-1.59	-0.13 [c]	23.34	-2.52	-0.21 [c]	31	63	24
2010	19.31	21.51	-2.20	-0.18 [c]	22.06	-2.75	-0.23 [c]	28	63	23
2011	18.83	21.25	-2.42	-0.19 [c]	21.79	-2.96	-0.22 [c]	23	63	21
2012	17.42	21.67	-4.25	-0.31 [d]	21.83	-4.41	-0.33 [d]	22	63	20
2013	17.42	22.35	-4.93	-0.35 [d]	22.82	-5.41	-0.37 [d]	18	63	17
2014	18.33	22.64	-4.31	-0.30 [d]	23.62	-5.29	-0.34 [d]	17	63	16
2015	17.16	21.62	-4.46	-0.37 [d]	18.18	-1.03	-0.16 [c]	12	63	12

[a] Effect size (ES) or standardized mean difference.
[b] Based on What Works Clearinghouse (WWC) standards, a standardized difference of ($|0.00| \leq$ ES Difference $\leq |0.05|$) indicates "Satisfies baseline equivalence."
[c] Based on WWC standards, a standardized difference of ($|0.05| <$ ES difference $\leq |0.25|$) indicates "Statistical adjustment required to satisfy baseline equivalence."
[d] Based on WWC standards, a standardized difference of (ES difference $> |0.25|$) indicates "Does not satisfy baseline equivalence."

Table C7. Baseline Equivalence Analysis: School Enrollment (Number of Students)

Year	ELT Schools	Non-ELT Schools (All Active BPS Schools by 2016)			Non-ELT Matched Comparison Schools from Selected Model			Number of Schools		
	Mean (1)	Mean (2)	Raw Mean Difference (1) - (2)	Standardized Difference[a]	Mean (3)	Raw Mean Difference (1) - (3)	Standardized Difference[a]	ELT Schools	Non-ELT Schools	Non-ELT Matched Schools
2006	436.37	454.66	-18.29	-0.05 [c]	436.79	-0.43	-0.00 [b]	31	63	24
2007	428.40	453.76	-25.36	-0.07 [c]	435.25	-6.85	-0.03 [b]	31	63	24
2008	414.77	465.12	-50.35	-0.15 [c]	439.25	-24.48	-0.11 [c]	31	63	24
2009	417.53	465.83	-48.30	-0.16 [c]	445.75	-28.22	-0.13 [c]	31	63	24
2010	427.64	472.28	-44.64	-0.15 [c]	420.52	7.12	0.04 [b]	28	63	23
2011	404.13	479.67	-75.54	-0.25 [d]	416.29	-12.16	-0.07 [c]	23	63	21
2012	435.14	513.90	-78.76	-0.26 [d]	470.15	-35.01	-0.19 [c]	22	63	20
2013	422.39	509.13	-86.74	-0.29 [d]	492.41	-70.02	-0.36 [d]	18	63	17
2014	415.71	499.87	-84.17	-0.28 [d]	495.56	-79.86	-0.39 [d]	17	63	16
2015	413.83	506.36	-92.53	-0.30 [d]	505.75	-91.92	-0.47 [d]	12	63	12

[a] Effect size (ES) or standardized mean difference.
[b] Based on What Works Clearinghouse (WWC) standards, a standardized difference of ($|0.00| \leq$ ES Difference $\leq |0.05|$) indicates "Satisfies baseline equivalence."
[c] Based on WWC standards, a standardized difference of ($|0.05| <$ ES difference $\leq |0.25|$) indicates "Statistical adjustment required to satisfy baseline equivalence."
[d] Based on WWC standards, a standardized difference of (ES difference $> |0.25|$) indicates "Does not satisfy baseline equivalence."

Table C8. Baseline Equivalence Analysis: Percent Female Students (School Averages)

Year	ELT Schools	Non-ELT Schools (All Active BPS Schools by 2016)			Non-ELT Matched Comparison Schools from Selected Model			Number of Schools		
	Mean (1)	Mean (2)	Raw Mean Difference (1) - (2)	Standardized Difference[a]	Mean (3)	Raw Mean Difference (1) - (3)	Standardized Difference[a]	ELT Schools	Non-ELT Schools	Non-ELT Matched Schools
2006	47.84	47.41	0.43	0.08 [c]	46.23	1.61	0.32 [d]	31	63	24
2007	48.13	47.43	0.70	0.12 [c]	45.82	2.31	0.40 [d]	31	63	24
2008	48.06	47.12	0.95	0.17 [c]	45.08	2.98	0.55 [d]	31	63	24
2009	48.40	47.37	1.03	0.19 [c]	45.81	2.59	0.50 [d]	31	63	24
2010	47.99	47.43	0.57	0.12 [c]	46.87	1.13	0.23 [c]	28	63	23
2011	47.03	47.82	-0.79	-0.16 [c]	47.32	-0.28	-0.05 [c]	23	63	21
2012	47.10	47.81	-0.71	-0.17 [c]	47.45	-0.35	-0.08 [c]	22	63	20
2013	47.62	47.30	0.32	0.07 [c]	46.24	1.38	0.30 [d]	18	63	17
2014	47.73	47.47	0.26	0.06 [c]	46.30	1.43	0.28 [d]	17	63	16
2015	48.18	47.36	0.82	0.17 [c]	48.67	-0.49	-0.15 [c]	12	63	12

[a] Effect size (ES) or standardized mean difference.
[b] Based on What Works Clearinghouse (WWC) standards, a standardized difference of ($|0.00| \leq$ ES Difference $\leq |0.05|$) indicates "Satisfies baseline equivalence."
[c] Based on WWC standards, a standardized difference of ($|0.05| <$ ES difference $\leq |0.25|$) indicates "Statistical adjustment required to satisfy baseline equivalence."
[d] Based on WWC standards, a standardized difference of (ES difference $> |0.25|$) indicates "Does not satisfy baseline equivalence."

Table C9. Baseline Equivalence Analysis: Percent White Students (School Averages)

Year	ELT Schools	Non-ELT Schools (All Active BPS Schools by 2016)			Non-ELT Matched Comparison Schools from Selected Model			Number of Schools		
	Mean (1)	Mean (2)	Raw Mean Difference (1) - (2)	Standardized Difference[a]	Mean (3)	Raw Mean Difference (1) - (3)	Standardized Difference[a]	ELT Schools	Non-ELT Schools	Non-ELT Matched Schools
2006	13.00	16.21	-3.21	-0.22 [c]	14.15	-1.14	-0.08 [c]	31	63	24
2007	12.52	16.14	-3.62	-0.25 [d]	13.82	-1.30	-0.09 [c]	31	63	24
2008	11.54	15.82	-4.29	-0.31 [d]	14.00	-2.46	-0.17 [c]	31	63	24
2009	11.27	15.65	-4.38	-0.31 [d]	13.58	-2.31	-0.16 [c]	31	63	24
2010	10.55	15.38	-4.83	-0.35 [d]	14.03	-3.48	-0.25 [c]	28	63	23
2011	12.17	15.18	-3.01	-0.21 [c]	13.76	-1.60	-0.11 [c]	23	63	21
2012	10.58	14.23	-3.64	-0.27 [d]	11.94	-1.36	-0.10 [c]	22	63	20
2013	9.74	15.24	-5.50	-0.39 [d]	14.28	-4.54	-0.30 [d]	18	63	17
2014	10.82	15.01	-4.18	-0.30 [d]	15.49	-4.67	-0.31 [d]	17	63	16
2015	14.63	15.44	-0.82	-0.06 [c]	13.43	1.20	0.08 [c]	12	63	12

[a] Effect size (ES) or standardized mean difference.
[b] Based on What Works Clearinghouse (WWC) standards, a standardized difference of ($|0.00| \leq$ ES Difference $\leq |0.05|$) indicates "Satisfies baseline equivalence."
[c] Based on WWC standards, a standardized difference of ($|0.05| <$ ES difference $\leq |0.25|$) indicates "Statistical adjustment required to satisfy baseline equivalence."
[d] Based on WWC standards, a standardized difference of (ES difference $> |0.25|$) indicates "Does not satisfy baseline equivalence."

Table C10. Baseline Equivalence Analysis: Percent African American Students (School Averages)

Year	ELT Schools	Non-ELT Schools (All Active BPS Schools by 2016)			Non-ELT Matched Comparison Schools from Selected Model			Number of Schools		
	Mean (1)	Mean (2)	Raw Mean Difference (1) - (2)	Standardized Difference[a]	Mean (3)	Raw Mean Difference (1) - (3)	Standardized Difference[a]	ELT Schools	Non-ELT Schools	Non-ELT Matched Schools
2006	43.06	41.06	2.00	0.09 [c]	45.46	-2.40	-0.10 [c]	31	63	24
2007	41.81	39.01	2.80	0.13 [c]	43.52	-1.71	-0.08 [c]	31	63	24
2008	40.34	37.31	3.03	0.14 [c]	41.90	-1.56	-0.07 [c]	31	63	24
2009	39.37	35.75	3.62	0.17 [c]	40.68	-1.31	-0.06 [c]	31	63	24
2010	38.11	34.77	3.34	0.16 [c]	40.10	-1.99	-0.09 [c]	28	63	23
2011	38.25	33.98	4.27	0.21 [c]	39.89	-1.64	-0.08 [c]	23	63	21
2012	36.73	32.47	4.26	0.22 [c]	38.30	-1.57	-0.08 [c]	22	63	20
2013	40.15	34.58	5.57	0.26 [d]	37.60	2.55	0.11 [c]	18	63	17
2014	37.53	33.62	3.91	0.19 [c]	35.12	2.41	0.11 [c]	17	63	16
2015	31.44	32.70	-1.26	-0.06 [c]	30.82	0.63	0.03 [b]	12	63	12

[a] Effect size (ES) or standardized mean difference.
[b] Based on What Works Clearinghouse (WWC) standards, a standardized difference of ($|0.00| \leq$ ES Difference $\leq |0.05|$) indicates "Satisfies baseline equivalence."
[c] Based on WWC standards, a standardized difference of ($|0.05| <$ ES difference $\leq |0.25|$) indicates "Statistical adjustment required to satisfy baseline equivalence."
[d] Based on WWC standards, a standardized difference of (ES difference $> |0.25|$) indicates "Does not satisfy baseline equivalence."

Table C11. Baseline Equivalence Analysis: Percent Hispanic Students (School Averages)

Year	ELT Schools	Non-ELT Schools (All Active BPS Schools by 2016)			Non-ELT Matched Comparison Schools from Selected Model			Number of Schools		
	Mean (1)	Mean (2)	Raw Mean Difference (1) - (2)	Standardized Difference[a]	Mean (3)	Raw Mean Difference (1) - (3)	Standardized Difference[a]	ELT Schools	Non-ELT Schools	Non-ELT Matched Schools
2006	37.29	32.72	4.57	0.23 [c]	31.34	5.95	0.29 [d]	31	63	24
2007	39.14	34.81	4.33	0.22 [c]	33.22	5.92	0.29 [d]	31	63	24
2008	41.16	36.92	4.24	0.21 [c]	34.67	6.49	0.32 [d]	31	63	24
2009	42.20	38.62	3.58	0.18 [c]	36.48	5.72	0.27 [d]	31	63	24
2010	44.02	39.51	4.51	0.22 [c]	36.29	7.73	0.36 [d]	28	63	23
2011	41.57	40.48	1.09	0.05 [c]	36.25	5.31	0.27 [d]	23	63	21
2012	44.94	43.40	1.54	0.08 [c]	39.49	5.45	0.29 [d]	22	63	20
2013	41.72	39.52	2.19	0.11 [c]	35.99	5.72	0.26 [d]	18	63	17
2014	41.96	40.67	1.29	0.06 [c]	36.79	5.17	0.23 [c]	17	63	16
2015	42.52	41.06	1.45	0.07 [c]	41.16	1.36	0.06 [c]	12	63	12

[a] Effect size (ES) or standardized mean difference.
[b] Based on What Works Clearinghouse (WWC) standards, a standardized difference of ($|0.00| \leq$ ES Difference $\leq |0.05|$) indicates "Satisfies baseline equivalence."
[c] Based on WWC standards, a standardized difference of ($|0.05| <$ ES difference $\leq |0.25|$) indicates "Statistical adjustment required to satisfy baseline equivalence."
[d] Based on WWC standards, a standardized difference of (ES difference $> |0.25|$) indicates "Does not satisfy baseline equivalence."