MPR Reference No.: 8927-006

MATHEMATICA
Policy Research, Inc.

High School Reform in Boston Public Schools: The Effect of Focus on High Schools on Student Academic Outcomes

Final Report

June 23, 2008

Susanne James-Burdumy
Irma Perez-Johnson
Sonya Vartivarian

Submitted to:

Boston Plan for Excellence
6 Beacon Street, #615
Boston, MA 02108

Submitted by:

Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005

Project Officer:
Stephanie J. Sibley

Project Director:
Susanne James-Burdumy

ACKNOWLEDGMENTS

This report represents a collaborative effort among many individuals. We are grateful to the Boston Public Schools (BPS) for providing the data for our analyses. Completing the report would not have been possible without the contributions of other individuals at Mathematica Policy Research, Inc. Natalya Verbitsky provided expert advice on the specification of statistical models. Dan Finkelstein helped lead the initial data cleaning, which was completed by James Mulligan. James Mulligan, Mindy Hu, Yuhong Zheng, Shinu Verghese, Anna Comerford, and Matthew Jacobus carefully and efficiently completed the programming that supported the analyses presented, and helped prepare the tables and figures presented in the report. Dan Player helped draft the report's technical appendix. Jane Nelson provided excellent word processing and production assistance, as usual. We are also grateful to Peter Schochet and Neil Seftor for their insightful comments as internal reviewers of our analysis plan and draft report, respectively.

We gratefully acknowledge the support and guidance of Stephanie Sibley of the Boston Plan for Excellence (BPE) throughout the study. In addition to Stephanie's careful review, this report benefited from input from other reviewers at BPE and BPS, including Barbara Adams, Irving Scott, Ellen Guiney, and Dr. Carol Johnson. We are thankful for all these contributions, as well as those of others we may have omitted unintentionally.

CONTENTS

Chapter | Page

TABLES

FIGURES

Figure **Page**

Figure **Page**

EXECUTIVE SUMMARY

CONTEXT

The Focus on High Schools (FOHS) initiative is intended to transform the 12 comprehensive high schools of the Boston Public Schools (BPS) into more effective institutions.[1] FOHS is addressing the combined problems of student alienation and poor literacy skills by breaking these high schools down into smaller, more intimate learning communities and reforming their English and Language Arts (ELA) instructional programs. BPS is working collaboratively with four organizations on the FOHS initiative: (1) Boston Plan for Excellence (BPE), (2) Boston Private Industry Council (PIC), (3) Center for Collaborative Education (CCE), and (4) Jobs for the Future (JFF).

The Boston Plan for Excellence contracted with Mathematica Policy Research, Inc. (MPR) to evaluate the FOHS initiative. The study includes two main components. The first is a survey that assessed students' perceptions of their school and ELA classrooms (James-Burdumy and Finkelstein 2006).[2] The second component, on which this report is focused, examined the changes in students' academic outcomes after the implementation of the initiative.

STUDY DESIGN

This study used an interrupted time series (ITS) design to examine the extent to which FOHS reforms are associated with changes in students' academic outcomes. An ITS approach was deemed appropriate for the study since it was not possible to randomly assign students or particular schools to the FOHS reforms and the requisite administrative records data were available to implement an ITS design. The ITS approach allowed us to investigate whether the FOHS initiative had an effect on students' academic outcomes by estimating a model that *projected* the outcomes that *would have been expected* in the absence of the FOHS initiative and comparing these projections to the pattern of outcomes actually observed after the implementation of the FOHS initiative.[3] We also examined whether students of different

[1] When the study began, there were 12 comprehensive high schools: Boston, Brighton, Burke, Charlestown, Dorchester, East Boston, English, Hyde Park, Madison Park Technical-Vocational, Snowden International, South Boston, and West Roxbury. By spring 2003, there were 11 comprehensive high schools and one pilot school. By March 2005 (at the time of the second survey administration), there were seven comprehensive high schools, one pilot school, two education complexes that housed five small, non-pilot schools, and two schools that were in the planning phases of becoming educational complexes. By September 2005, these two new complexes housed seven non-pilot schools.

[2] The survey results showed significant differences in 3 of 12 composite measures of student perceptions examined: (1) reported increases in teachers use of progressive pedagogical methods, (2) decreases in student reports of misbehaviors, and (3) diminished perceptions of student relationships with other students in their schools (James-Burdumy and Finkelstein 2006).

[3] More specifically, for each of the student outcomes examined, we developed a statistical model that *projected* outcomes in the years after the FOHS reforms were introduced based on the pattern of outcomes in the years

subgroups (classified by gender, race, use of English in the home, special education status, and eligibility for free or reduced-price lunch) were differentially affected by the FOHS initiative. Finally, we examined effects at the school level.

Our ITS analyses are based on administrative records data from the 1995-1996 through 2006-2007 school years. This 12-year period includes seven years of data from before the FOHS initiative began (school years 1995-1996 through 2001-2002) and five years after it began (school years 2002-2003 through 2006-2007). The study focused on the following outcomes: MCAS scaled scores and proficiency ratings in mathematics and English Language Arts (ELA); number of days absent, tardy, and present; number of unexcused absences; whether students were promoted to the next grade; whether students were suspended; and number of suspensions.

Note that the findings presented in this report can not establish a *causal* relationship between the FOHS initiative and estimated changes in student outcomes. This is principally because factors other than the FOHS reforms may have also influenced student outcomes at or around the time when these reforms were introduced. Therefore, caution must be exercised when interpreting our estimates of the possible effects of FOHS reforms, since the reforms may not have *caused* the changes observed.

RESULTS

Full Sample

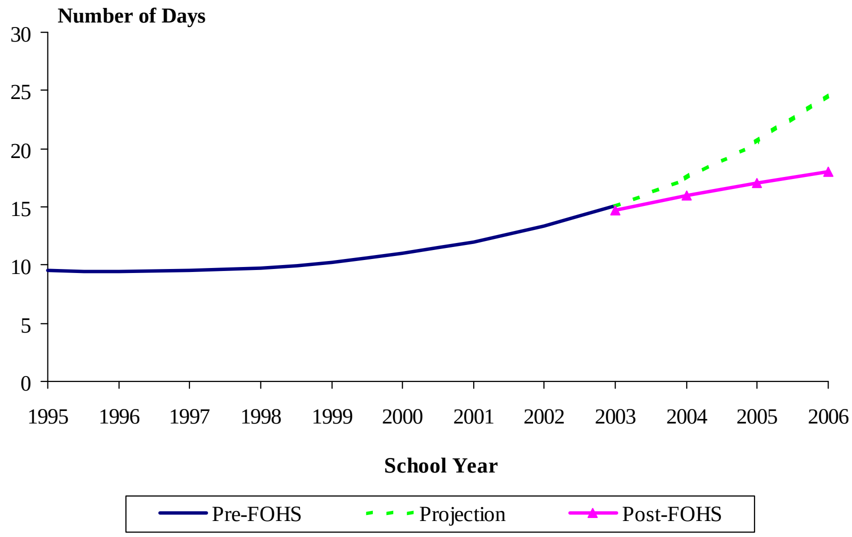
In the full sample, we observed significant changes in all student outcomes after the implementation of FOHS relative to what would have been projected in the absence of the FOHS initiative. The majority of the estimated effects were in a direction consistent with the goals of the FOHS initiative. In particular, after the implementation of FOHS, the number of days students were absent from school declined (Figure ES.1), the number of unexcused absences declined, the number of days tardy declined (Figure ES.2), the percentage of students suspended declined (Figure ES.3), the number of suspensions declined, and the percentage of students promoted to the next grade increased (Figure ES.4), all relative to what would have been projected in the absence of the FOHS initiative.

Other estimated changes in outcomes were not consistent with the goals of the FOHS initiative. In particular, MCAS scaled scores and proficiency rates declined—principally in ELA (Figure ES.5) and more modestly in mathematics—relative to what would have been projected in the absence of the FOHS initiative. In addition, the number of days present declined modestly relative to what would have been projected in the absence of FOHS (Figure ES.6).

(continued)
preceding FOHS and other statistical adjustments (e.g., for changes in the characteristics of students attending BPS schools over time). Each model then compared these predictions to the outcomes actually *observed* after FOHS implementation, also with some statistical adjustments. The difference between these *projections* and *observed* values constitute the ITS estimates of the effects of FOHS reforms on the outcome being examined in each post-FOHS year.

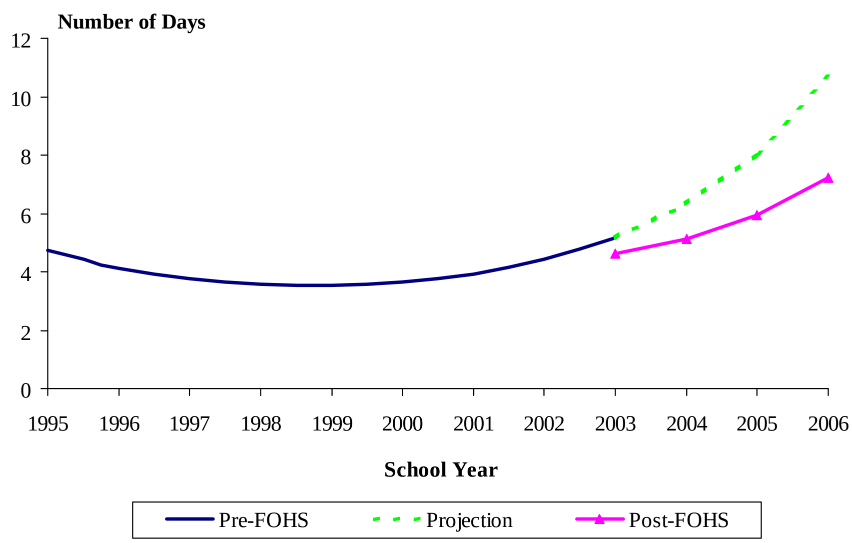
FIGURE ES.1

EFFECT OF FOHS ON TOTAL ABSENCES



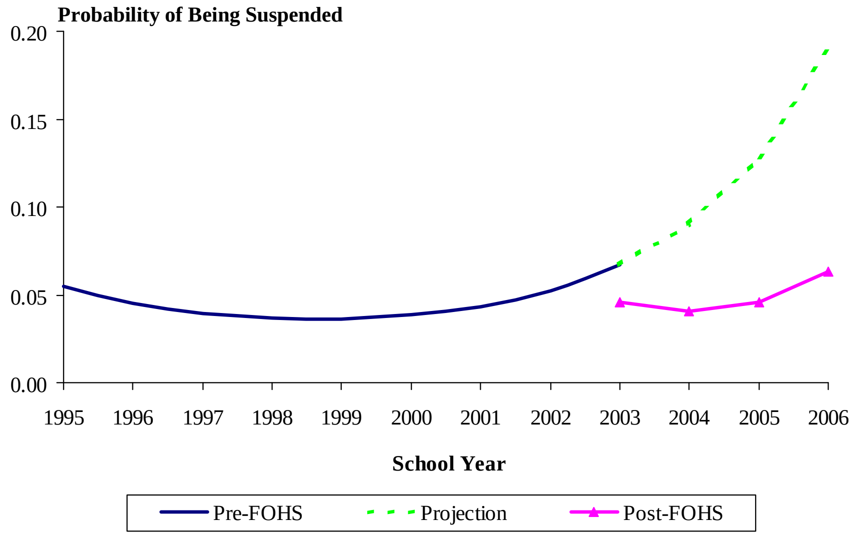
Note: Depicted trajectories are for the reference category group in the corresponding statistical model. The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

FIGURE ES.2
EFFECT OF FOHS ON TARDINESS



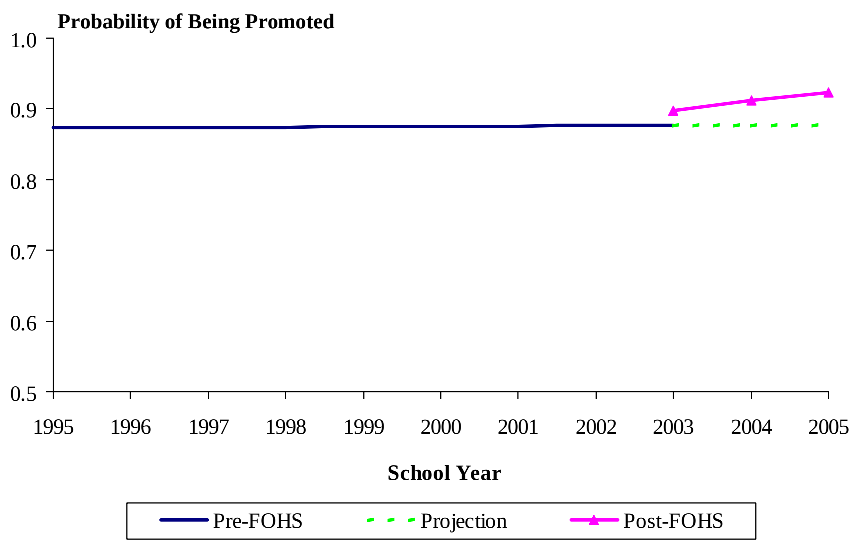
Note: Depicted trajectories are for the reference category group in the corresponding statistical model. The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

FIGURE ES.3
EFFECT OF FOHS ON THE PROBABILITY OF BEING SUSPENDED



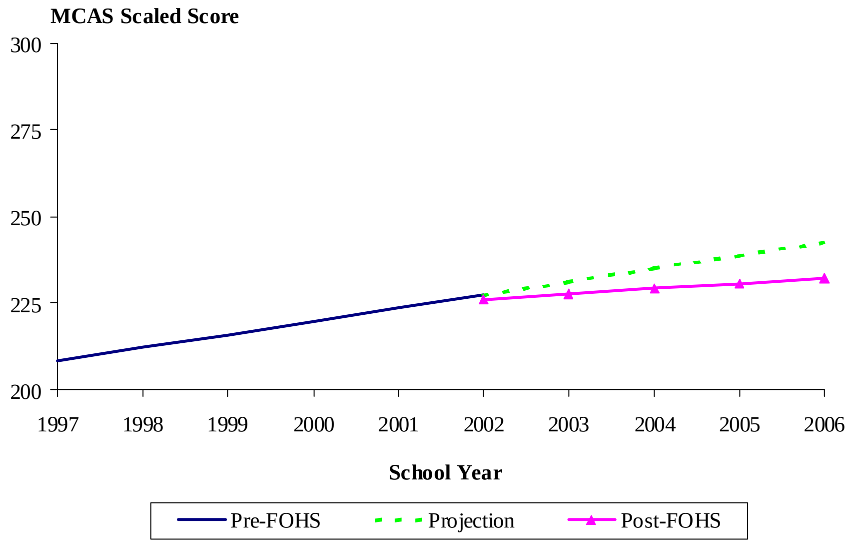
Note: Depicted trajectories are for the reference category group in the corresponding statistical model. The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

FIGURE ES.4
EFFECT OF FOHS ON GRADE PROMOTION



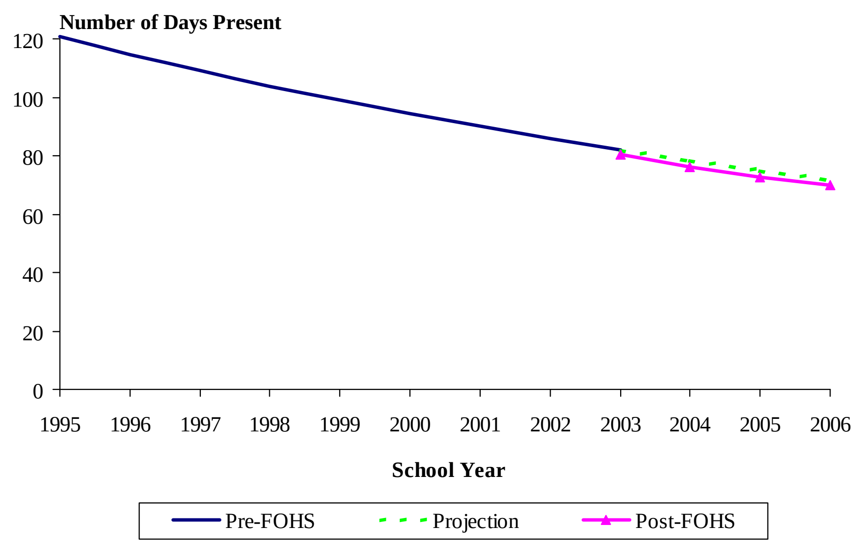
Note: Depicted trajectories are for the reference category group in the corresponding statistical model. The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

FIGURE ES.5
EFFECT OF FOHS ON MCAS TENTH GRADE ELA SCALED SCORES



Note: The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the horizontal axis of the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

FIGURE ES.6
EFFECT OF FOHS ON ATTENDANCE



Note: Depicted trajectories are for the reference category group in the corresponding statistical model. The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

Subgroups

Results from the subgroup analyses showed that some student subgroups experienced significantly different effects than other student subgroups after the implementation of FOHS. We present these findings here by outcome, as this helps to see more clearly when the observed changes in outcomes were experienced uniformly across all or most students subgroups and when the effects were concentrated among a certain subgroup of students.

- ***MCAS scaled scores.*** The reductions in ELA MCAS scores were observed across most student subgroups, except students who spoke English at home, and were significantly larger among black, Hispanic, and special education students. These three student subgroups also experienced significant decreases in their MCAS mathematics scores, while the MCAS mathematics scores of students eligible for free or reduced-price lunch increased. Significant effects on MCAS mathematics scores were not observed for other student subgroups.

- ***MCAS proficiency.*** Students who speak English in the home were significantly more likely to score proficient or higher on the ELA MCAS test, as compared to students who spoke other languages at home. Compared to regular education students, special education students were significantly less likely to score proficient or higher on both the mathematics and ELA MCAS tests. Significant effects on MCAS proficiency rates were not observed for other student subgroups.

- ***Attendance.*** The reductions in number of days present and number of days absent were observed across all student subgroups. The effect on number of days tardy was quite mixed, with reductions in tardiness concentrated among black and Hispanic students.

- ***Suspensions.*** The estimated reductions in student suspensions were somewhat more modest for black students, as compared to white students. Otherwise, there was no variation in effects on suspensions across the student subgroups.

- ***Promotion to Next Grade.*** Most student subgroups did not experience a significant effect on promotions to the next grade. The one exception was special education students, who experienced a significant increase in the percentage of students promoted to the next grade.

Schools

Finally, analyses conducted at the school level suggest that the observed changes in MCAS outcomes were experienced fairly evenly across the 12 high schools that were present at the outset of the study. While there was some variation across schools in the post-FOHS changes observed, many schools did not experience changes that differed significantly from the changes observed in other schools.

CONCLUSION

Overall, significant differences were observed in all of the outcomes examined. The estimated decline in MCAS scores, especially in ELA, relative to what would have been projected in the absence of the FOHS reforms was consistent across most subgroups, which provides suggestive evidence that the FOHS initiative may not have had the intended effect on MCAS scores. The overall reduction in number of days present was also inconsistent with the goals of FOHS. However, a number of other estimated post-FOHS changes were in the direction intended by the FOHS initiative. In particular, we observed a reduction in number of days absent (including unexcused absences), which was observed consistently across the student subgroups. We also observed a reduction in number of days tardy, which was concentrated among black and Hispanic students. Student suspensions were also reduced consistently across all subgroups in the post-FOHS period relative to what would be expected in the absence of FOHS. Finally, we observed an overall increase in promotions to the next grade; however these results were concentrated among special education students.

I. INTRODUCTION

Increasingly, researchers, policymakers, and school leaders have been calling for major reforms in the structure and operations of U.S. high schools. Reform advocates commonly argue that today's large, comprehensive high schools are not adequately preparing students for college, work, or life (American Diploma Project 2004), and that their structure fosters feelings of alienation among teachers and students (National Research Council 2003).

The Focus on High Schools (FOHS) initiative is intended to transform the 12 comprehensive high schools of the Boston Public Schools (BPS) into smaller, more personal, and more effective institutions.[1] FOHS reforms have strived to address the combined problems of student alienation and poor literacy skills in BPS by breaking high schools down into smaller, more intimate learning communities and reforming their English Language Arts (ELA) instructional programs. BPS is working with the following four organizations on the FOHS initiative: (1) Boston Plan for Excellence (BPE), (2) the Boston Private Industry Council (PIC), (3) the Center for Collaborative Education (CCE), and (4) Jobs for the Future (JFF).

BPE contracted with Mathematica Policy Research, Inc. (MPR) to evaluate the FOHS initiative. MPR's evaluation focused on the analysis of quantitative data on students' perceptions and outcomes. (A separate evaluation, focusing on the implementation of the FOHS reforms, was conducted by Education Matters, Inc.)

[1] When the study began, there were 12 comprehensive high schools: Boston, Brighton, Burke, Charlestown, Dorchester, East Boston, English, Hyde Park, Madison Park Technical-Vocational, Snowden International, South Boston, and West Roxbury. By spring 2003, there were 11 comprehensive high schools and one pilot school. By March 2005 (at the time of the second survey administration), there were seven comprehensive high schools, one pilot school, two education complexes that housed five small, non-pilot schools, and two schools that were in the planning phases of becoming educational complexes. By September 2005, these two new complexes housed seven non-pilot schools. Appendix A provides additional details about the FOHS initiative and the evolution of BPS high schools between 1999 and 2005.

MPR's evaluation of the FOHS reforms had two components. The first was a student engagement study, based on surveys administered to 9th- and 11th-grade students in BPS high schools in spring 2003 and spring 2005. The survey study examined changes in students' perceptions of their schools and ELA courses since the outset of the FOHS initiative (James-Burdumy and Finkelstein 2006; Mayer 2003), and found significant differences in 3 of 12 composites examined. An increase in teachers' use of progressive pedagogical methods was consistently reported across most subgroups, providing some evidence that FOHS reforms may have influenced how ELA classes are taught. Decreases in student reports of misbehavior also suggested possible improvements in student conduct. However, students' perceptions of their relationships with other students in their same schools diminished—a result contrary to the goals of FOHS reforms. School-level analyses further suggested that observed changes were not experienced evenly across the 12 high schools present at the outset of the study.

The second component of MPR's FOHS study was a student academic outcomes study. This report presents findings from that component of the study. This study used an interrupted time series (ITS) design to examine the extent to which FOHS reforms are associated with changes in students' academic outcomes. An ITS approach was deemed appropriate for the study since it was not possible to randomly assign students or schools to the FOHS reforms and the requisite administrative records data were available. The ITS approach allowed us to investigate whether the FOHS initiative had an effect on students' academic outcomes by estimating a model that *predicted* the outcomes that *would have been expected* in the absence of the FOHS initiative and comparing these projections to the pattern of outcomes actually observed after the implementation of the FOHS initiative.[1]

[1] More specifically, for each of the student outcomes examined, we developed a statistical model that *predicted* outcomes in the years after the FOHS reforms were introduced based on the pattern of outcomes in the years

Our ITS analyses were designed to address three main research questions:

1. To what extent are there changes in outcomes for ***the overall student population*** following the implementation of FOHS?
2. Are there differential effects ***for subgroups of students***—classified by gender, race, home use of the English language, special education status, and eligibility for free or reduced-price lunch?
3. Are there differential effects across ***study schools*** following the implementation of FOHS?

The ITS analyses are based on administrative records data from the 1995-1996 through 2006-2007 school years. This 12-year period includes 7 years of data from before the FOHS initiative began (from school years 1995-1996 through 2001-2002) and 5 years after it began (from school years 2002-2003 through 2006-2007).

The ITS analyses examine student outcomes over time, including the period before and after the implementation of FOHS reforms. However, because they cannot attribute changes in outcomes to the reforms themselves, caution must be used when interpreting these findings.

A. SMALL HIGH SCHOOL REFORMS

Some educational reformers, concerned about the poor functioning of large, comprehensive high schools, have been advocating for the creation of smaller high schools (see, for example, Coalition of Essential Schools 2008; Meier 1995; Sizer 1984). According to these reformers, small high schools could be an antidote to comprehensive schools, which they claim mainly sort

---

(continued)
preceding FOHS and other statistical adjustments (e.g., for changes in the characteristics of students attending BPS schools over time). Each model then compared these predictions to the outcomes actually *observed* after FOHS implementation, also with some statistical adjustments. The difference between these *predicted* and *observed* values constitute the ITS estimates of the effects of FOHS reforms on the outcome being examined in each post-FOHS year.

students into academic and nonacademic tracks (Powell et al. 1985), as opposed to striving to adequately prepare all students.

Calls to reinvent America's high schools (Gates 2005) have intensified with the demands for higher graduation rates and test scores brought about by the No Child Left Behind Act (NCLB) of 2001 and the advent of state accountability systems. School safety concerns, heightened by the Columbine High School shootings in 1999 and subsequent violent incidents, have also contributed to a sense that contemporary high schools are an institution in crisis (Hendrie 2004).

In recent years, reforms focusing on transforming large, comprehensive high schools into smaller, more personal learning communities have been spurred on by both federal and foundation funding. The federal Small Learning Communities (SLC) Program, reauthorized under NCLB, awards discretionary grants for up to 60 months to local education agencies to support the implementation of SLCs and other activities aimed at improving academic achievement in large public high schools with enrollments of 1,000 or more students (U.S. Department of Education 2008). SLCs include such structures as freshman academies, multigrade academies organized around career interests or other themes, "houses" in which small groups of students remain together throughout high school, and autonomous schools-within-a-school. Allowable SLC activities also include personalization strategies, such as student advisories, family advocate systems, and mentoring programs.

Since 2001, the Bill & Melinda Gates Foundation has committed more than $900 million to high school reforms, most of it in support of efforts aimed at improving high schools through the creation and replication of small schools or small learning communities (Robelen 2006). As of 2004, the Foundation was estimated to have invested $647 million to support the creation of smaller high schools (*Education Week* 2004).

Despite the growing popularity of reforms aimed at downsizing high schools, evidence of the effectiveness of such reforms is mixed and limited. Some studies find that small schools are associated with improved student achievement (Darling-Hammond et al. 2002; Holland 2002; Howley 1989; Lee 2002). A recent study of the New Century High Schools in New York City found that their students graduated on time more often than their peers citywide (Foley et al. 2007). There is also some evidence that smaller schools may promote more equitable gains in achievement (Lee and Smith 1995; Darling-Hammond et al. 2002) and lower dropout rates (Darling-Hammond et al. 2002; Foley et al. 2007; Holland 2002; Pittman and Haughwout 1987).

Other studies on the effects of small-school reforms at the secondary level are less encouraging. For instance, Wasley et al. (2000) and Hess and Cytrynbaum (2002) found that small-school reforms in Chicago led to improvements in student engagement but not in academic achievement. Early evaluation results for the Gates Foundation's small-schools initiative have also been mixed. Mitchell et al. (2005) found evidence of a more positive school climate, more rigorous homework assignments, and higher-quality student work in ELA, but lower-quality work in mathematics. In addition to positive effects on graduation rates, Foley et al. (2007) found that the New Century students earned standard diplomas less often than students from other New York City high schools.

Finally, concerns have been raised about the array of approaches being implemented under the banner of small high schools. Research evidence on the effects of taking a large school and turning it into small learning communities is very limited. To our knowledge, no comparative studies have been conducted on the benefits of converting existing schools into smaller learning communities versus starting new, autonomous, small schools.

B. THE BOSTON PUBLIC SCHOOLS AND FOCUS ON HIGH SCHOOLS

BPS enrolls more than 60,000 students each year, including almost 19,000 students in grades 9 to 12.[3] Over the past decade, BPS has acquired a national reputation as a model for school reform and has become known as one of the highest-performing urban public school systems in the country. BPS students have demonstrated consistent improvement on statewide assessments—the Massachusetts Comprehensive Assessment System (MCAS)—since these tests were first administered in 1998. In 2006, BPS won the distinguished Broad Prize for Urban Education as the best city school district in the nation.

Like all urban school districts, however, the BPS district also faces challenges. More than 70 percent of the district's students qualify for free or reduced-price lunch programs. As of June 2006, almost one in five members of the district's class of 2006 had failed to pass the 10th-grade MCAS mathematics and ELA exams, a requirement for graduation.

Efforts to transform high schools into smaller learning communities and improve literacy instruction have been an explicit component of comprehensive school reform plans in the BPS district. Focus on Children II included as a goal "to reorganize *every* district high school into small learning communities."[4] In a March 2003 message to the National Governors Association, then Superintendent Payzant outlined three fundamental challenges for reforms that focus on high schools: (1) to raise student achievement to proficiency levels, (2) to reduce the dropout rate so that higher test scores are not seen as the result of eliminating low performers, and (3) to

[3] "The Boston Public Schools at a Glance." [http://boston.k12.ma.us/bps/bpsglance.asp]. Accessed February 13, 2008.

[4] "Focus on Children II, Boston's Education Reform Plan: 2001-2006." (Adopted by the Boston School Committee, April 25, 2001), p. 10.

raise the rate of students continuing their education in college. He also described the role of smaller schools in helping to meet these challenges:

> In smaller schools, teachers, and administrators will be able to work more closely together to assure that everyone is pulling in the same direction for every student. Students, especially low performers, will be much less likely to disappear into anonymity. And, as smaller schools define themselves around unifying themes and curricula, the connections to our local colleges and universities will grow even stronger than they have been in the past decade. (Payzant 2003).

In 2001, the Carnegie Corporation of New York launched a national initiative, Schools for a New Society, and awarded to BPS $8.25 million over five years to support FOHS reforms in its 12 comprehensive high schools. The strategies developed were intended to reduce school alienation and improve literacy instruction. In 2003, the Bill and Melinda Gates Foundation augmented that support by awarding BPS $13.6 million for small school creation and development. A second Gates Foundation grant, $7.9 million awarded to BPS in November 2005, supported the development of both small schools and SLCs.

Consistent with the aims of these various grant awards, BPS has pursued several strategies to transform its large, comprehensive high schools (Boston Public Schools 2005). The district has created a number of new, small high schools, each with fewer than 500 students, and their own headmaster and teaching staff. Many of these new, small high schools have been organized around a theme or career interest. BPS has also developed a number of pilot high schools, which are free to determine their own budget, staffing, governance, curricula, assessment, and school calendar. In 2003, BPS began creating small, autonomous schools within the walls of large district high schools. These SLCs are typically overseen by the school's headmaster and share facilities, teachers, policies, and other schoolwide activities. Appendix A provides additional information about the FOHS initiative and the evolution of BPS high schools from 1999 to 2005.

C. ORGANIZATION OF THE REPORT

The next chapter describes the data and analytic methods used to examine the possible effects of FOHS reforms on student outcomes. The last chapter discusses our findings with respect to the overall, districtwide effects of the reforms, followed by discussions of effects for key subgroups of interest and for each of the 12 comprehensive high schools present at the outset of the study.

II. DATA AND ANALYTIC METHODS

The interrupted time series (ITS) component of the Focus on High Schools (FOHS) study used administrative records data from the Boston Public Schools (BPS) for a 12-year period to examine the extent to which FOHS reforms are associated with changes in student academic outcomes. This period includes eight years of data from the period before FOHS reforms (from school years 1995-1996 through 2002-2003) and four years after they began (from school years 2003-2004 through 2006-2007). This chapter describes the data used for this study, the trends evident in these data during the study period, and the analytic methods applied for our ITS analyses.

A. DATA

Our ITS study draws upon data related to students' background characteristics and academic outcomes. MPR received the following data on students' background characteristics from BPS:

- Gender (male, female)
- Race/ethnicity (white, black, Hispanic)
- Language used at home (English, not English)
- Special education status (yes, no)
- Free or reduced-price lunch eligibility (yes, no)

The data files provided by BPS also included student identifiers (that is, unique student IDs) and a school number, to enable the linking of student outcomes over time and the modeling of clustering of students within schools.

Data on student background characteristics were used as regressors in our models, to control for differences in student composition across schools and for changes over time in the characteristics of students attending BPS high schools. These data were also used in subgroup analyses to assess whether changes in student outcomes varied by subgroup.

Our ITS study examined 11 outcomes, including measures of academic achievement and other academic outcomes of interest to the FOHS partners. The outcome measures examined were:

- Massachusetts Comprehensive Assessment System (MCAS) scores in mathematics for 10th-grade students (scaled scores)
- Students scoring at MCAS proficiency or higher levels in mathematics in the 10th grade (yes/no)
- MCAS scores in English Language Arts (ELA) in the 10th grade (scaled scores)
- Students scoring at MCAS proficiency or higher levels in ELA in the 10th grade (yes/no)
- Days absent per academic year (number)
- Unexcused absences per academic year (number)
- Days present per academic year (number)
- Days tardy per academic year (number)
- Ever suspended in a given academic year (yes/no)
- Suspensions per academic year (number)
- Promoted to next grade in a given academic year (yes/no)

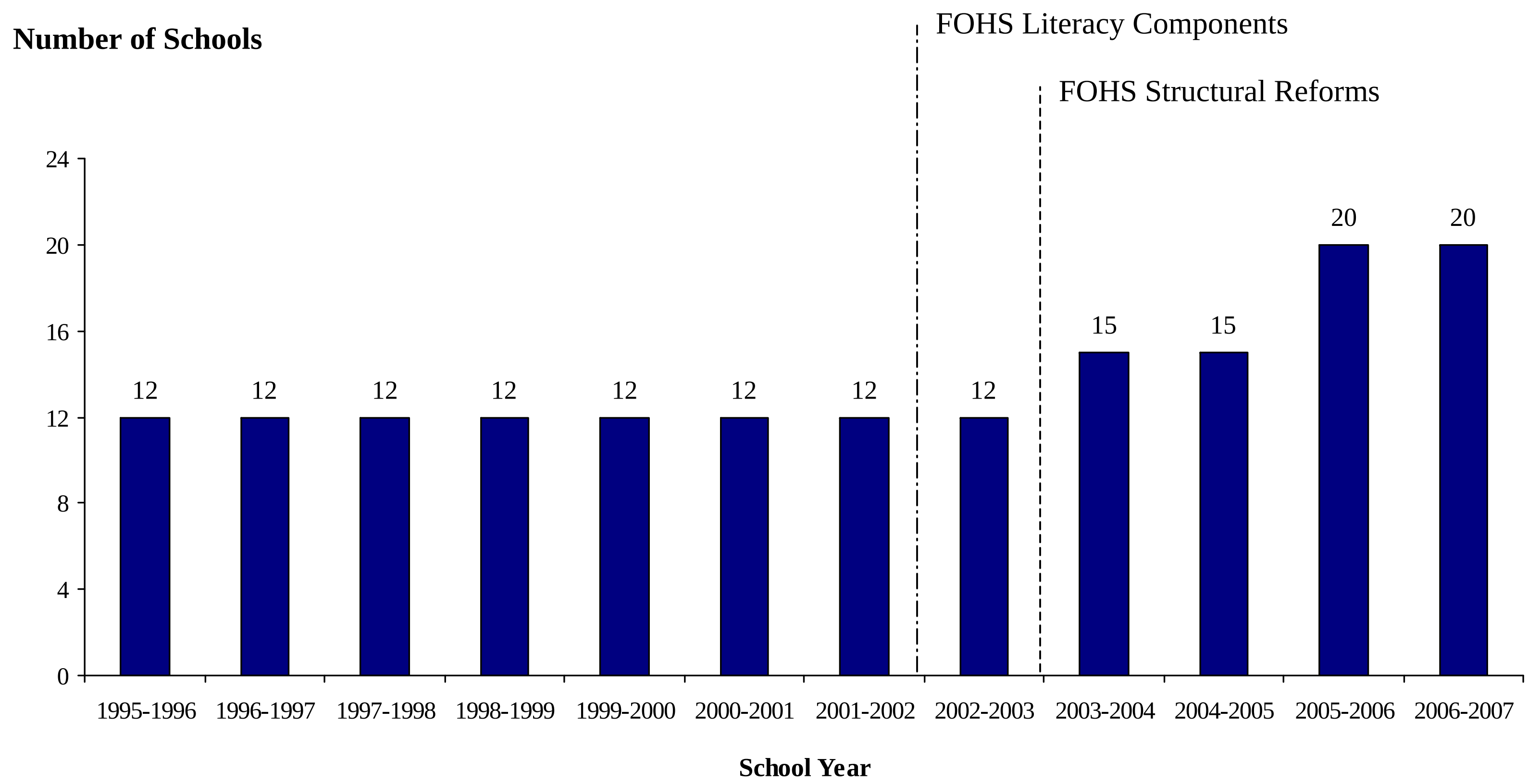
B. DEMOGRAPHIC AND OUTCOME TRENDS IN THE BPS HIGH SCHOOLS DURING THE STUDY PERIOD

Figures II.1 through II.16 show, for each of the academic years included in our study, descriptive statistics for the student demographic variables, outcome variables, number of schools in the study, and average number of students in study schools. Each figure also displays two vertical lines. The first, after the 2001-2002 school year and labeled “FOHS Literacy Components,” signals the introduction of literacy-focused FOHS reforms. The second line, after the 2002-2003 school year and labeled “FOHS Structural Reforms,” marks the start of FOHS reforms aimed at creating small schools or SLCs.

Consistent with the aim of the FOHS reforms to create smaller, high-school learning communities, Figure II.1 shows that the number of high schools increased after the

FIGURE II.1

NUMBER OF SCHOOLS IN FOHS STUDY



implementation of FOHS and Figure II.2 shows that the mean number of students per school declined from 1,339 in 2002-2003 to just under 750 in 2006-2007.

1. Changes in Mean Student Characteristics

The descriptive statistics presented in these figures suggest that the demographic profile of BPS high school students became somewhat more disadvantaged over the study period. Figure II.3 shows that there was no discernable change in the gender of students attending BPS high schools over the years covered by the study. The proportion of minority students, although always high, increased slightly over the study period, from 86 percent in 1995-1996 to 90 percent in 2006-2007 (Figure II.4). The proportion of students eligible for free or reduced-price lunch hovered between 48 and 51 percent in the late 1990s, but ranged between 61 and 65 percent in the mid-2000s (Figure II.5). The proportion of BPS high school students assigned to special education stayed relatively constant over the study period (Figure II.6), while the proportion of students coming from non-English-speaking households declined somewhat (Figure II.7).

2. Changes in Mean Academic Outcomes

Figures II.8 through II.16 display graphically the mean values over our 12-year study period for the outcome measures examined. While these figures may show some differences in the mean values of outcomes during the pre- and post-FOHS reform periods, these differences should not be used to make statements about the possible effects of the FOHS reforms on academic outcomes. This is because these differences fail to take into account changes in the student body composition of BPS high schools over time (like those noted above) and other

FIGURE II.2

MEAN NUMBER OF STUDENTS PER SCHOOL

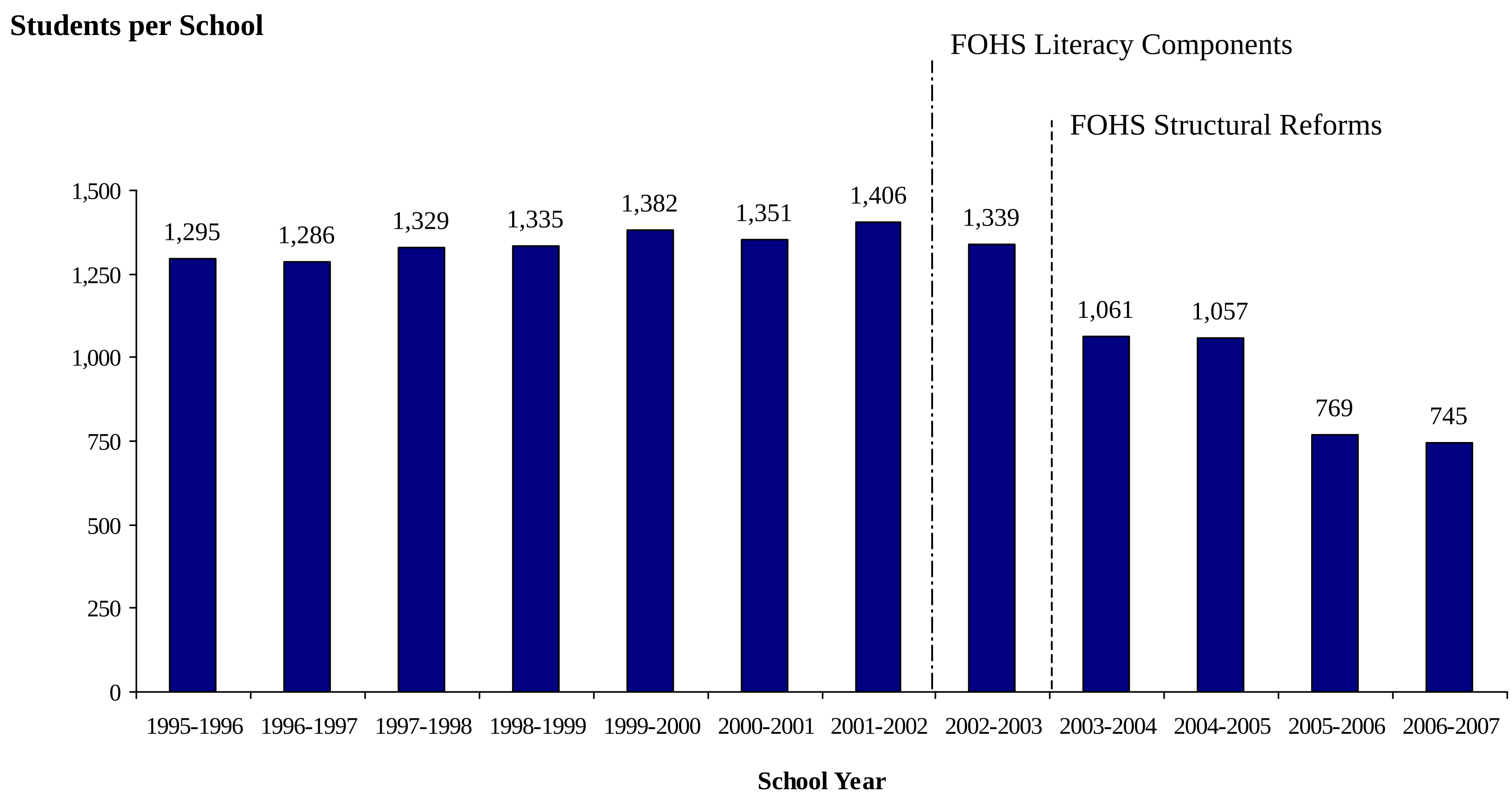


FIGURE II.3

PERCENTAGE OF STUDENTS THAT ARE FEMALE

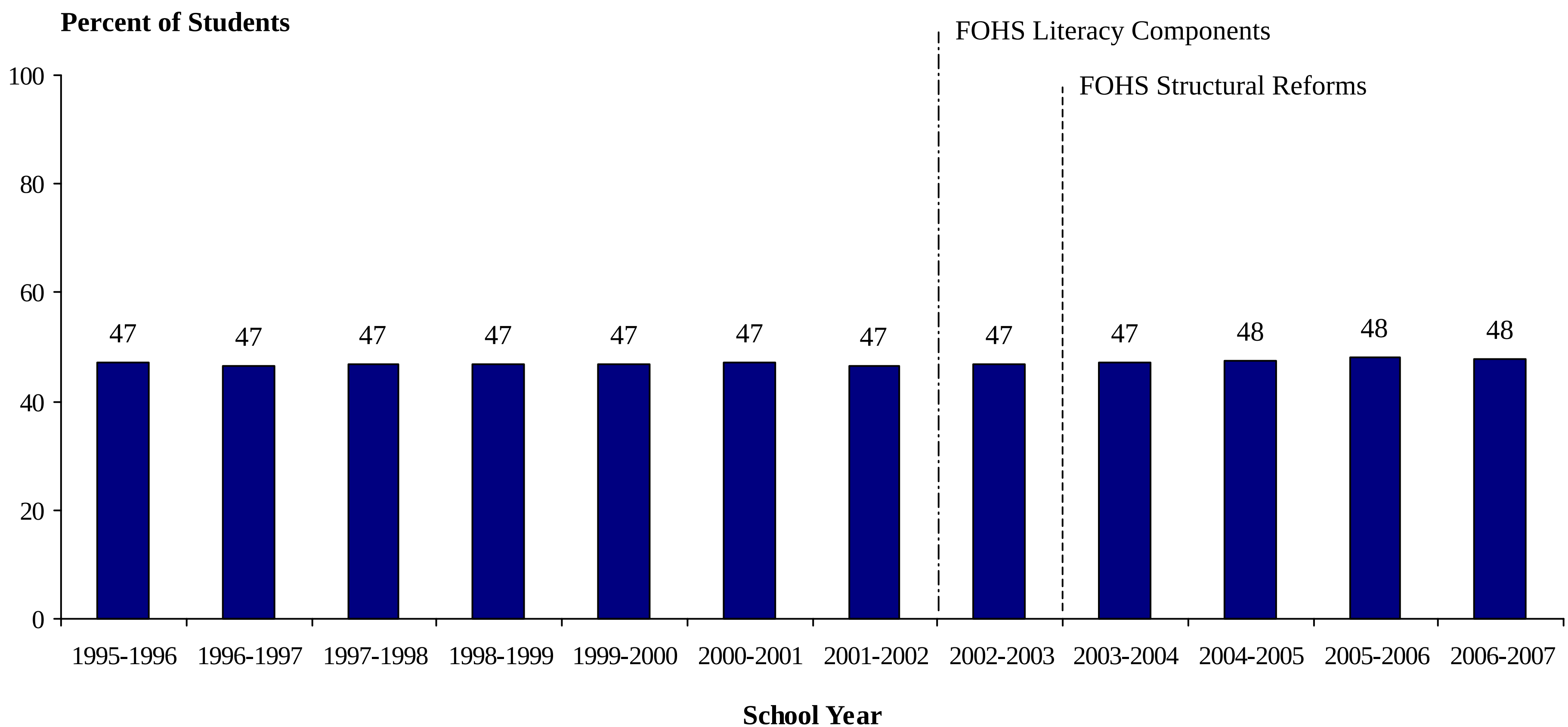


FIGURE II.4

PERCENTAGE OF STUDENTS THAT ARE MINORITY

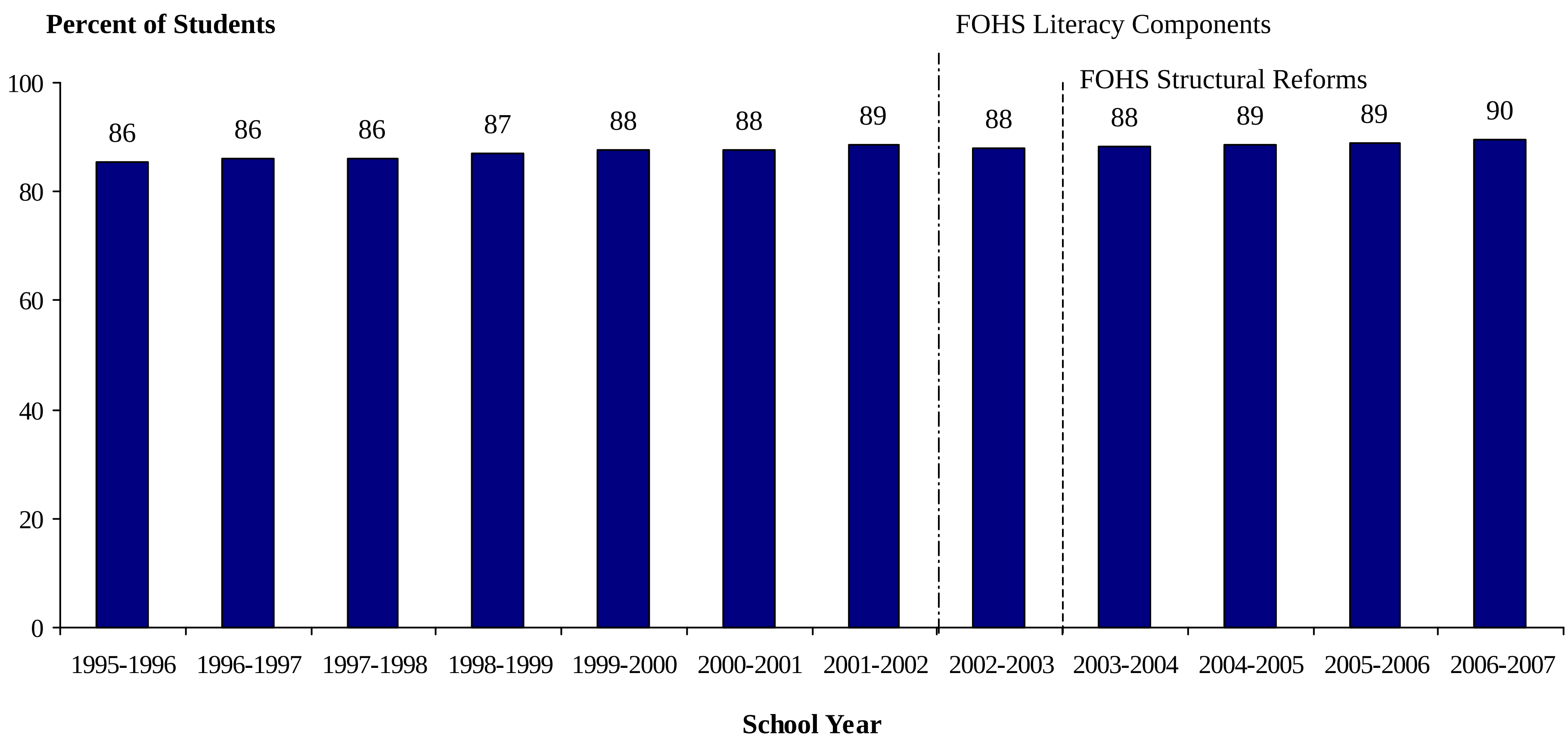


FIGURE II.5

PERCENTAGE OF STUDENTS ELIGIBLE FOR FREE OR REDUCED-PRICE LUNCH

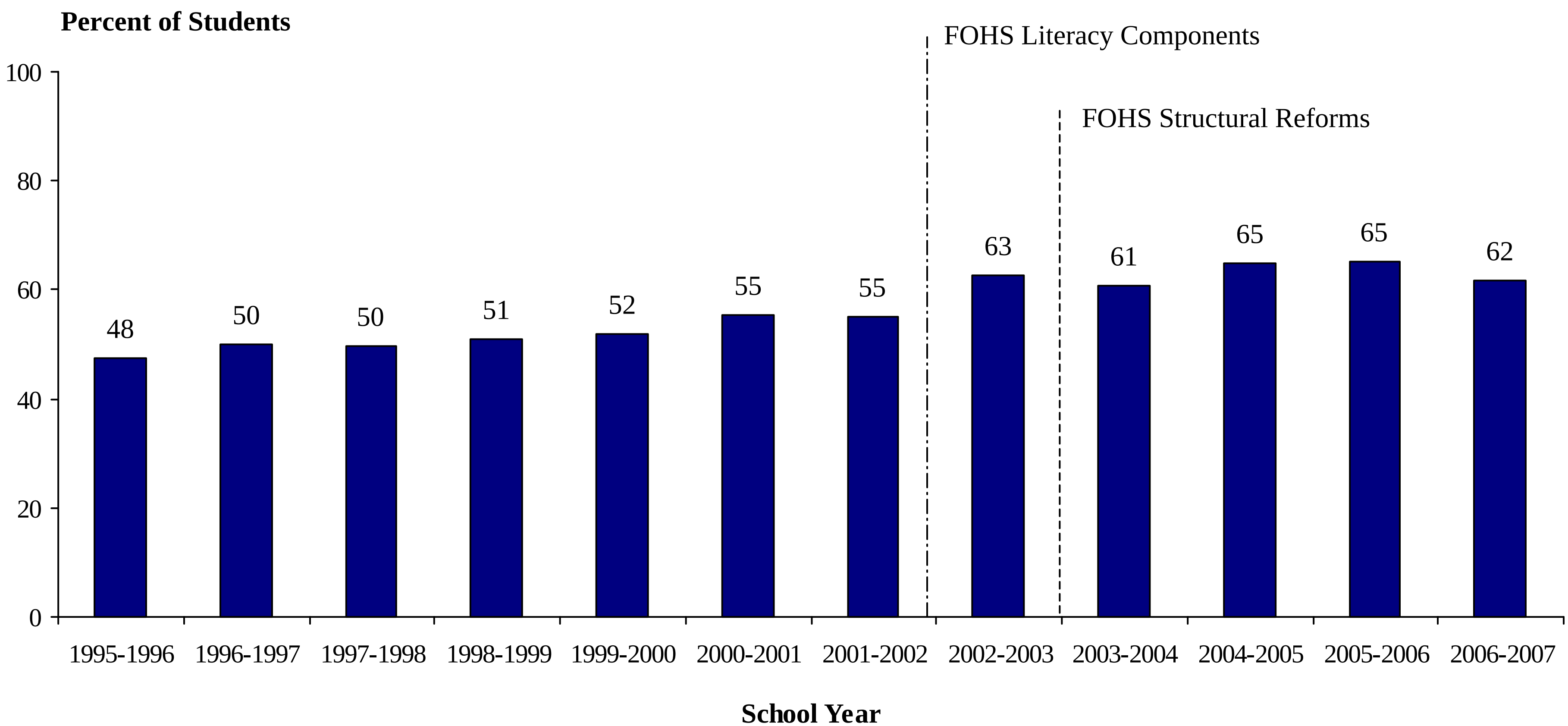


FIGURE II.6

PERCENTAGE OF STUDENTS THAT RECEIVE SPECIAL EDUCATION SERVICES

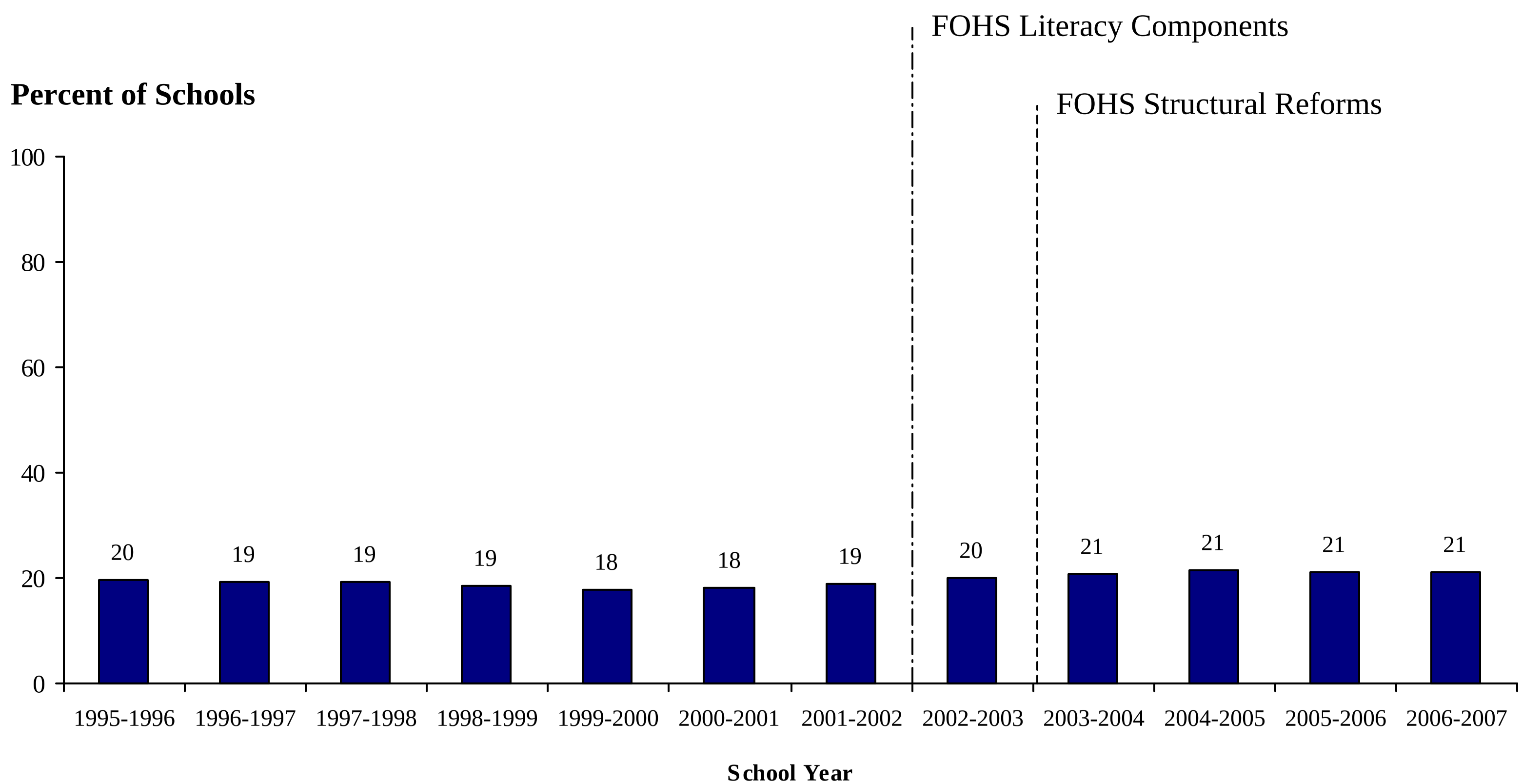
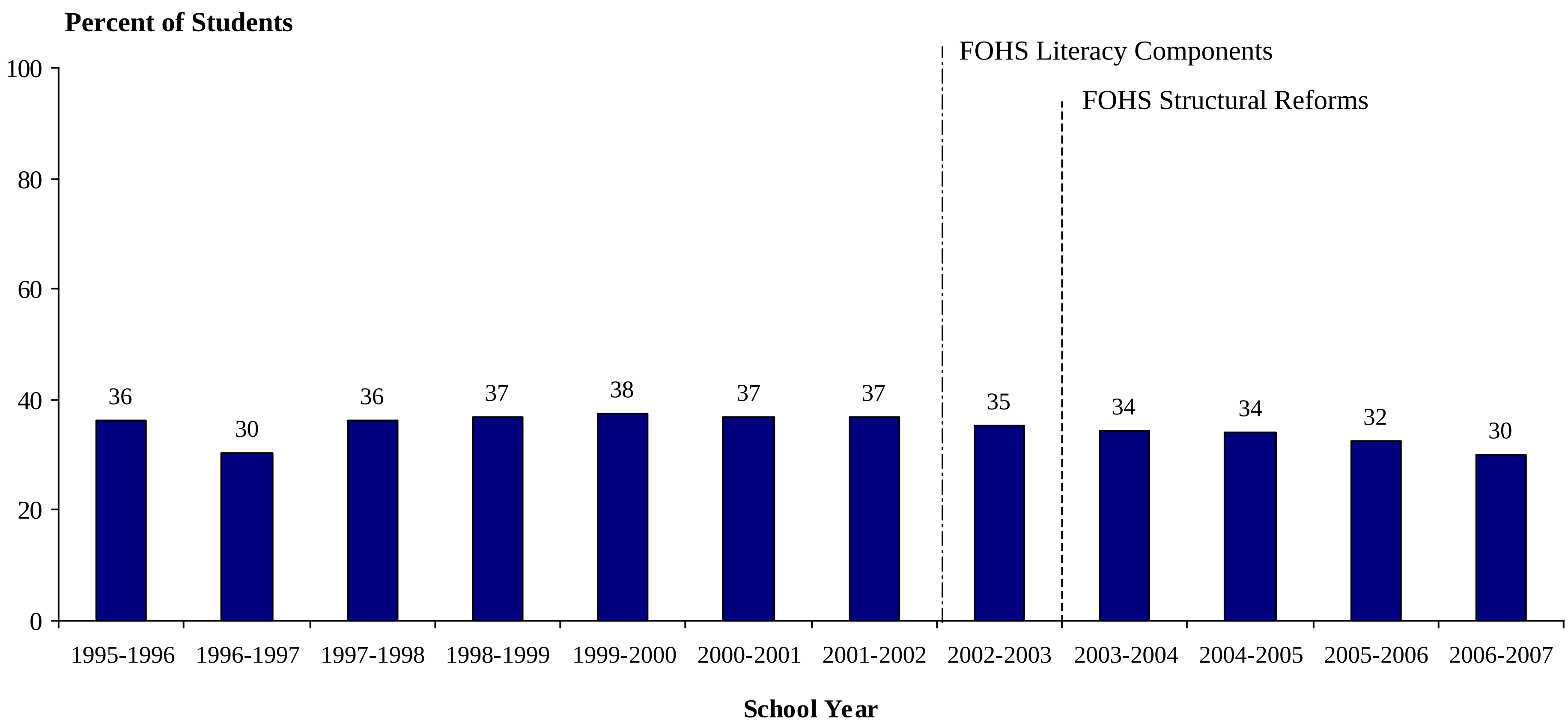


FIGURE II.7

PERCENTAGE OF STUDENTS FROM NON-ENGLISH-SPEAKING HOUSEHOLDS



factors that may have influenced student outcomes.[1] We provide these figures principally as an aid in identifying marked shifts in outcomes over the study period, which may provide BPS with clues as to whether there are other plausible, alternative explanations for any observed changes in outcomes (other than the FOHS reforms).

MCAS Scaled Scores in Mathematics and ELA. As Figures II.8 and II.9 show, mean scaled scores in both subjects increased over the 12-year study period. However, mean scaled scores exhibit a marked jump between 1999-2000 and 2000-2001—13 and 11 points in mathematics and ELA, respectively. Yearly increases in mean scaled scores beyond that point are more modest.

MCAS Proficiency Rates in Mathematics and ELA. The proportions of students scoring at proficiency or higher in mathematics and ELA also display a marked increasing trend over the 12-year study period (Figures II.10 and II.11). Marked jumps in proficiency rates seem to have occurred for ELA in 2002-2003, and for both subjects in 2005-2006.[2]

Student Attendance. Mean student attendance (that is, number of days present) appears to have remained stable over the 12-year study period (Figure II.12).

Student Absences and Tardiness. The mean number of total student absences appears to have remained fairly stable over the 12-year study period (Figure II.13). In contrast, mean

[1] For instance, after the 2005-2006 school year, Massachusetts stopped including the scores of students absent during the regular MCAS testing days from schools' overall performance calculations (Sibley, Adams, and Scott, 2008). Note, however, that our analyses are based on individual data for students (rather than school aggregate data) and do not exclude such students.

[2] In November 2002, Massachusetts' voters approved the Unz initiative, a ballot measure to end bilingual education and replace it instead with English-immersion programs (Sibley 2008). Under the new law, English-language-learners would receive one year of intensive English language instruction, and then be placed in more traditional, English-language classes. This change could have affected proficiency rates.

FIGURE II.8

MEAN MCAS MATHEMATICS SCALED SCORES

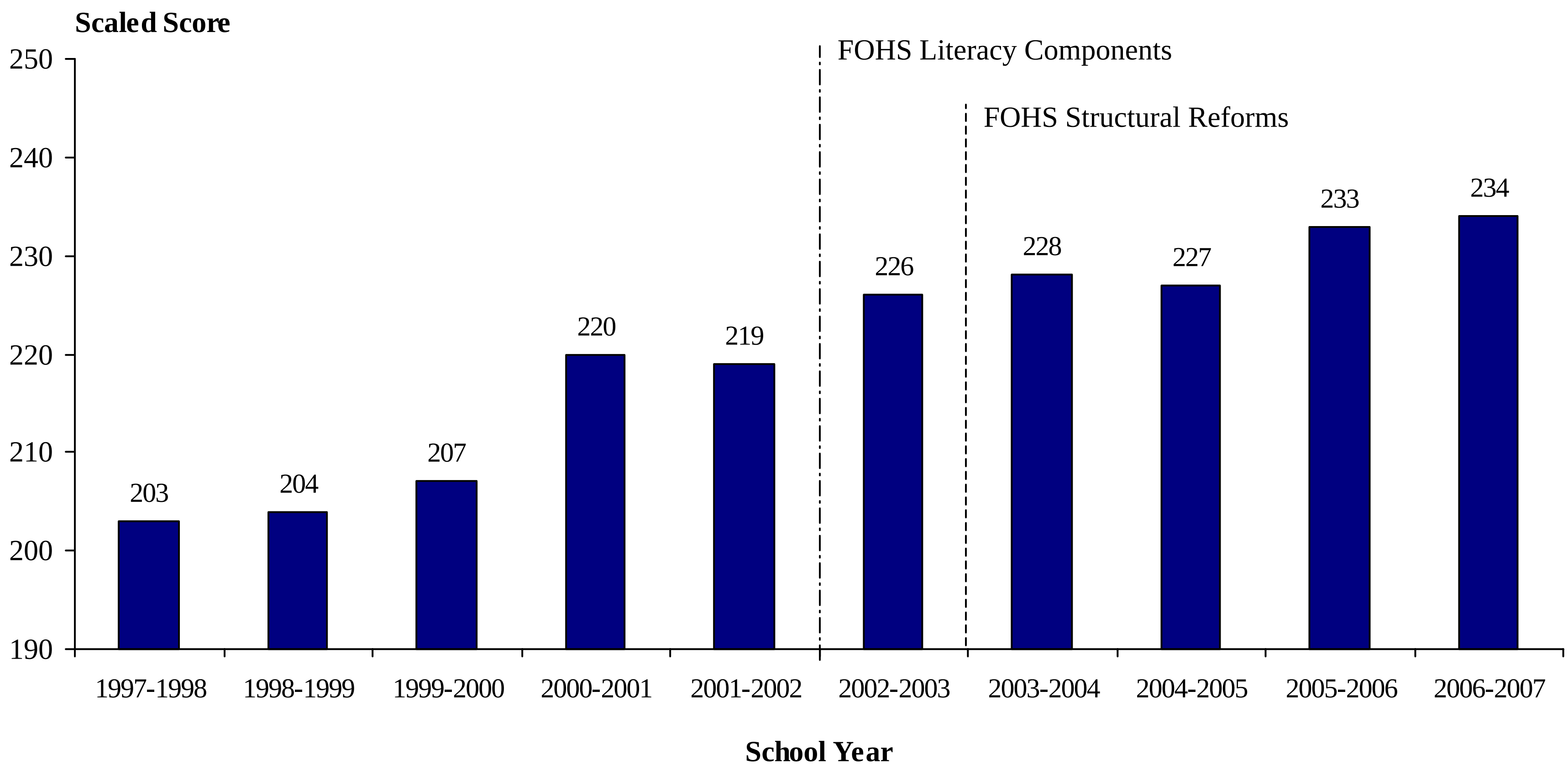


FIGURE II.9

MEAN MCAS ENGLISH LANGUAGE ARTS SCALED SCORES

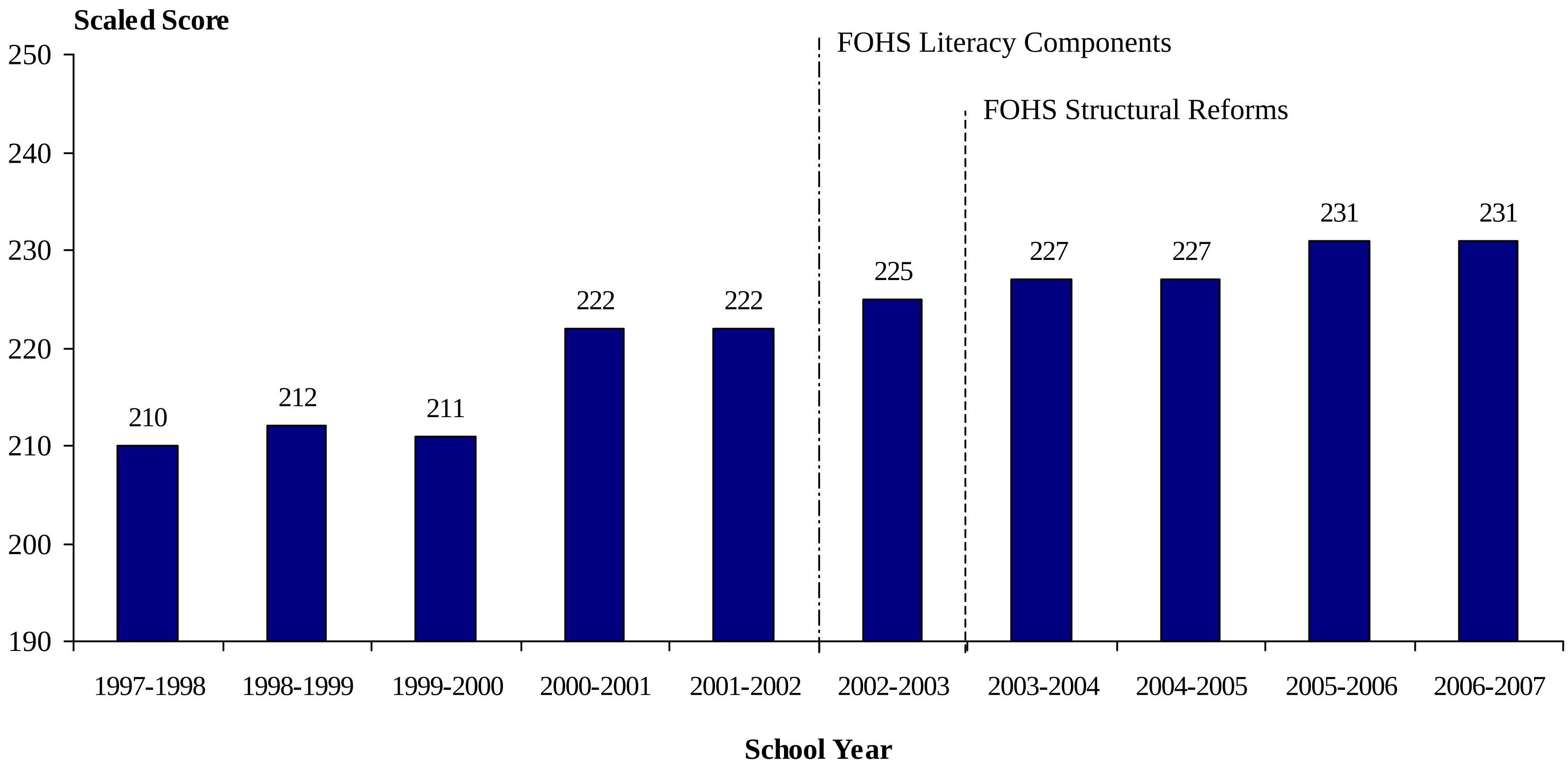


FIGURE II.10

MEAN MCAS 10TH-GRADE MATHEMATICS PROFICIENCY RATES

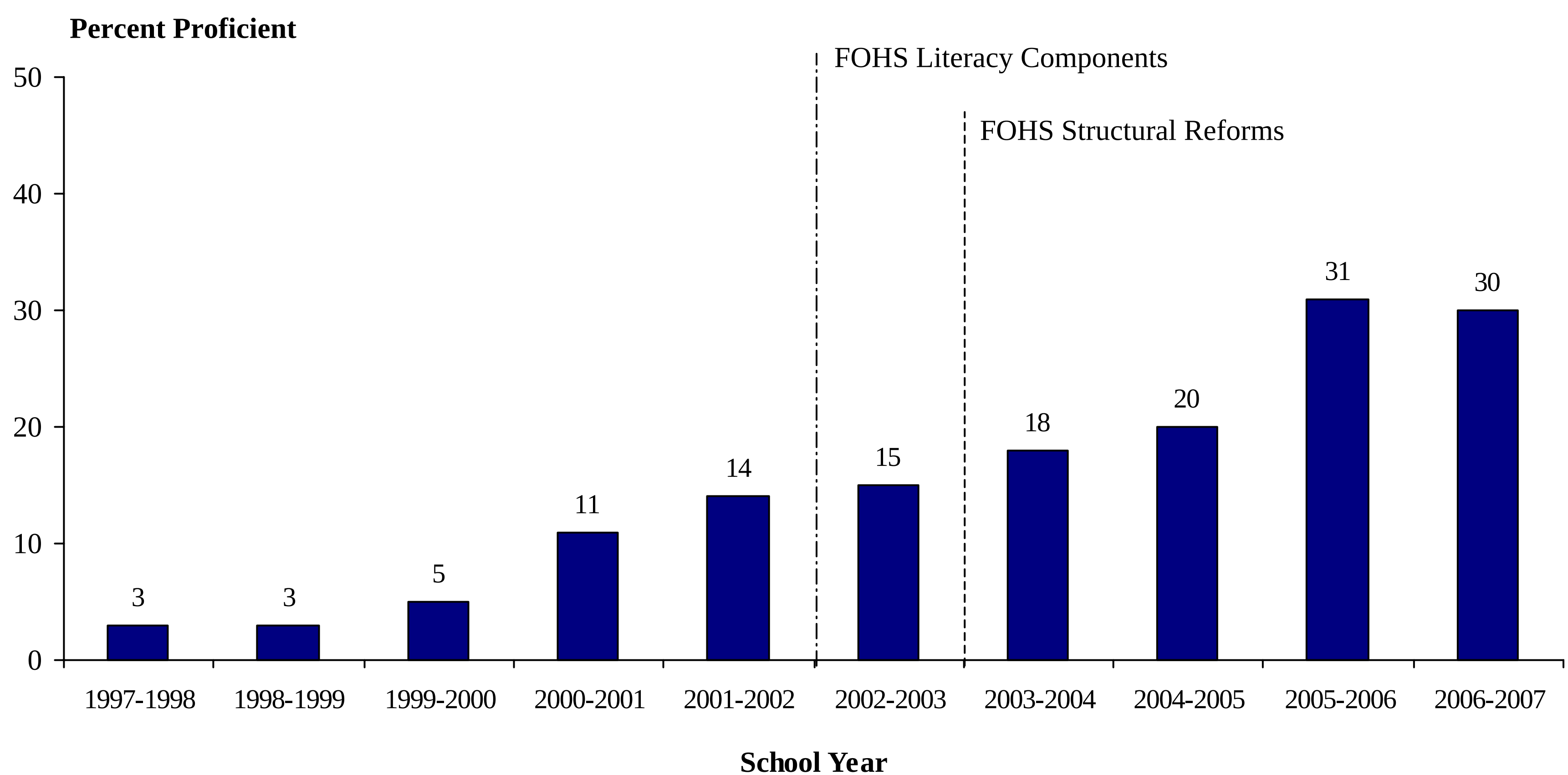


FIGURE II.11

MEAN MCAS 10TH-GRADE ENGLISH LANGUAGE ARTS PROFICIENCY RATES

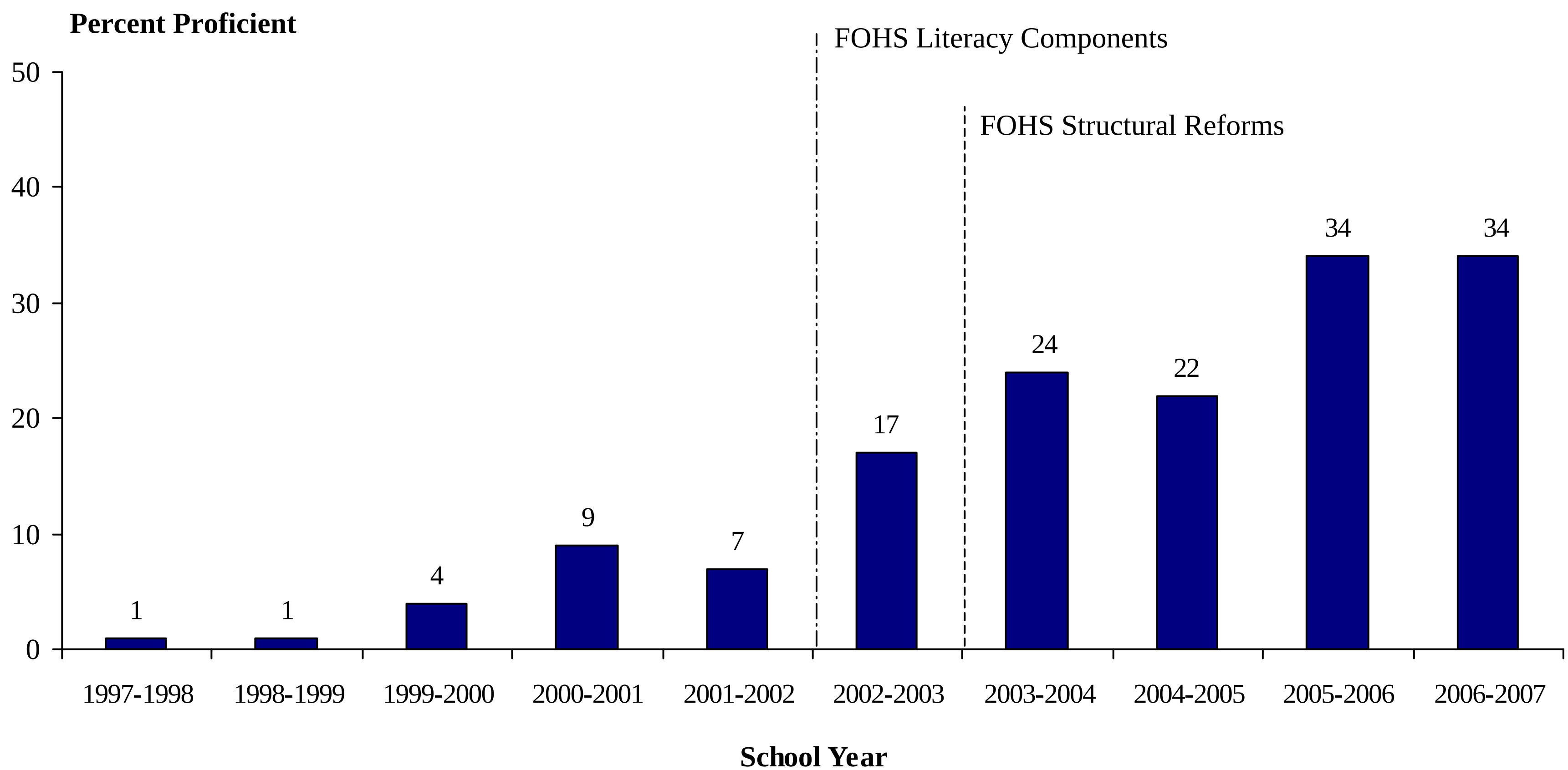


FIGURE II.12

MEAN ATTENDANCE

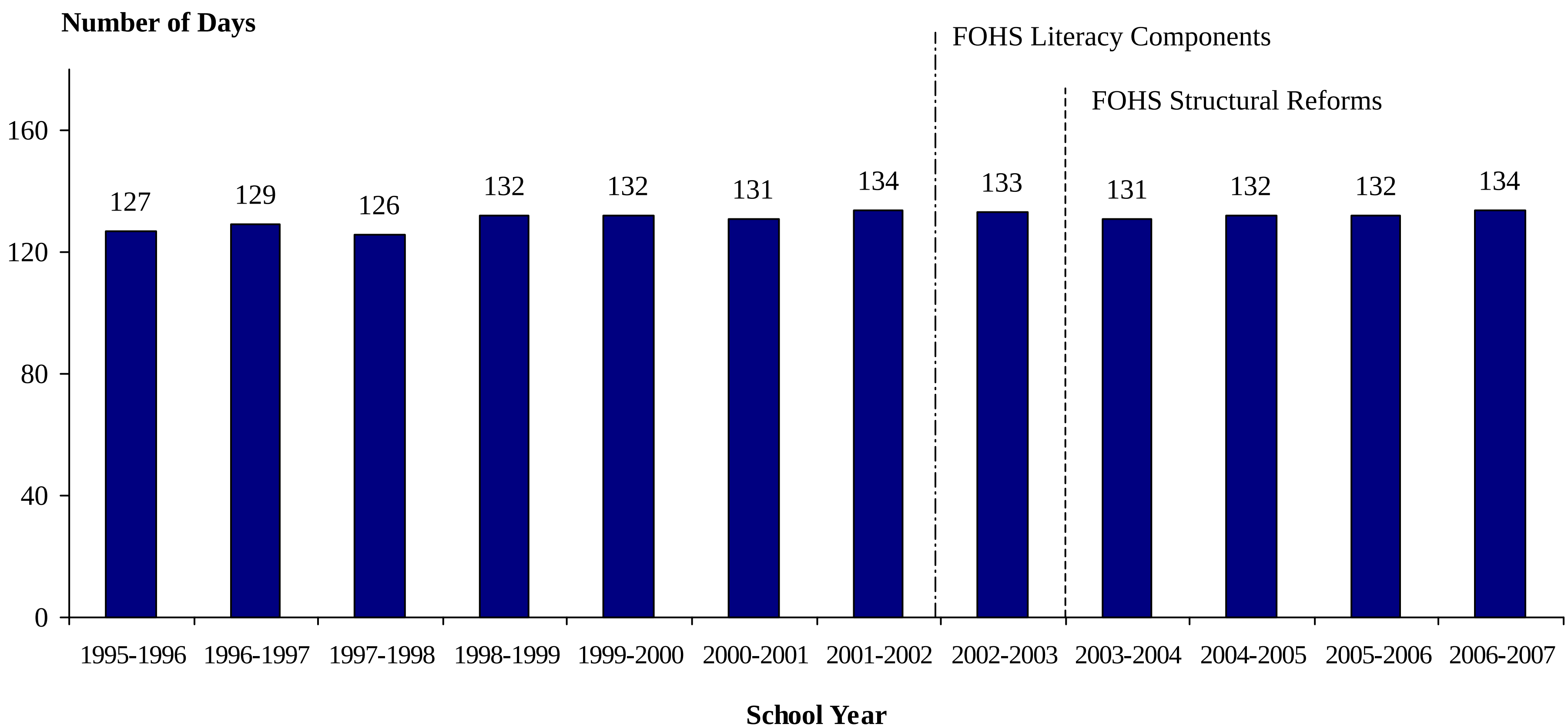
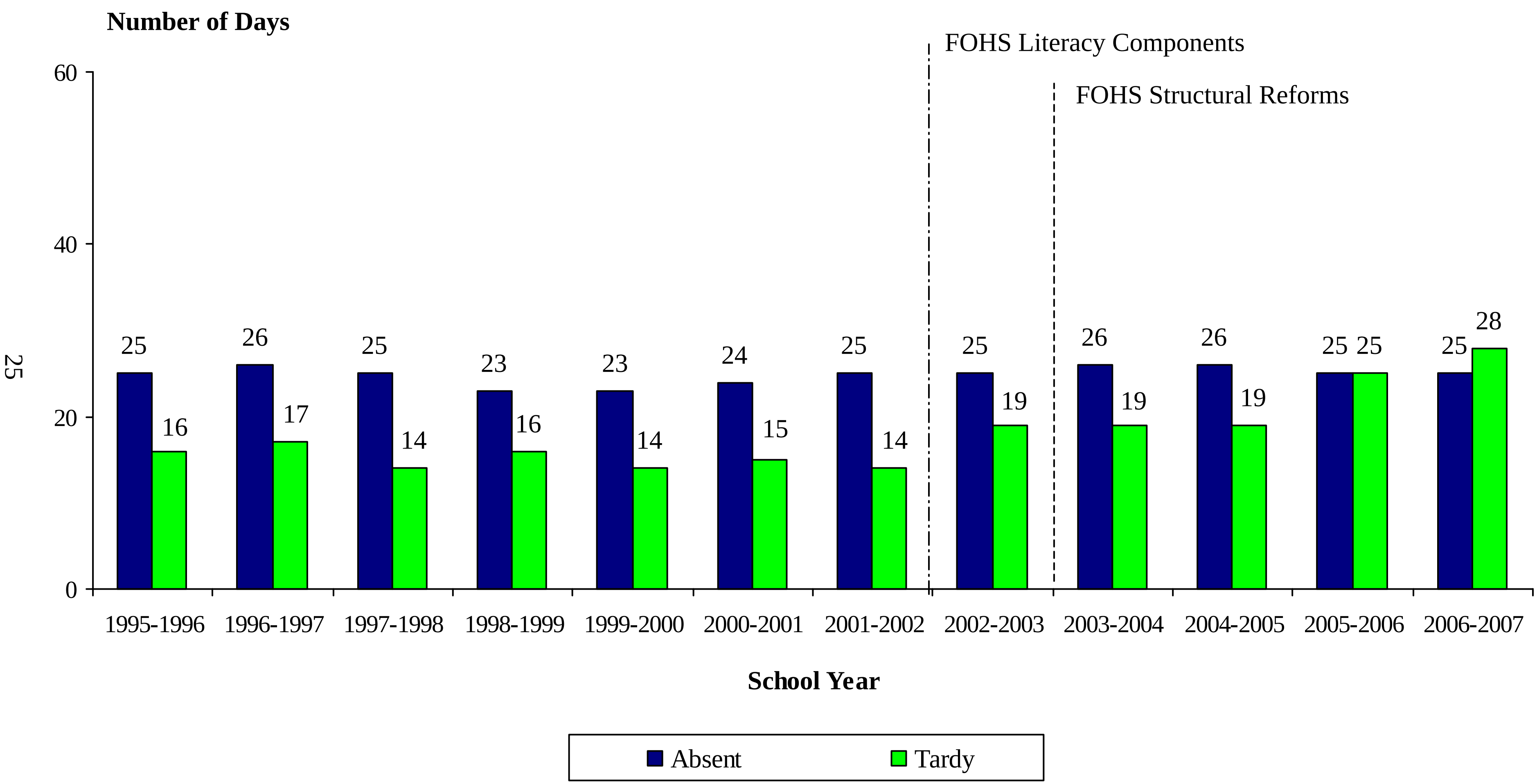


FIGURE II.13

MEAN ABSENCES AND TARDIES



tardiness appears to be higher (mean values between 19 and 28 days tardy) after the 2002-2003 school year.[3]

Suspensions. Figures II.14 and II.15 suggest quite a bit of variability in both the probability of suspension (that is, percent of students ever suspended) and the number of suspensions per student, respectively, over the years shown.

Grade Promotion. The rate of grade promotion exhibits a slight decreasing trend from school years 1995-1996 through 2002-2003. This trend appears to reverse after 2003-2004 (Figure II.16).

C. ANALYTIC METHODS

We used an ITS approach to examine the effect of FOHS on student outcomes. The prerequisites to implementing the ITS method are (1) the presence of a sufficiently large number of observations measured consistently on the same variables over time, and (2) knowledge of the specific point in this series at which the treatment occurred or was introduced (Shadish et al. 2002). If the treatment or intervention being tested—in this case, the FOHS reforms—had an effect, the expectation is that observations after the treatment will exhibit a different pattern than those before the treatment. That is, the *time series* should show an *interruption* at or shortly after the time when the treatment was introduced or delivered, relative to the pattern one would predict based on the pre-treatment observations. An effect of the treatment may be observed in several ways. For example, we may observe a shift in the level of the outcome at the point of the treatment's implementation or a shift in the slope, or rate of change, of the outcome over time

[3] In May 2006, the BPS passed a new tardy policy, which banned schools' practice of locking students out after a certain time (Sibley 2008). Both the expectation of and actual passage of this policy change may have contributed to the observed increase in tardy students observed during the post-FOHS period.

FIGURE II.14

PERCENTAGE OF STUDENTS SUSPENDED

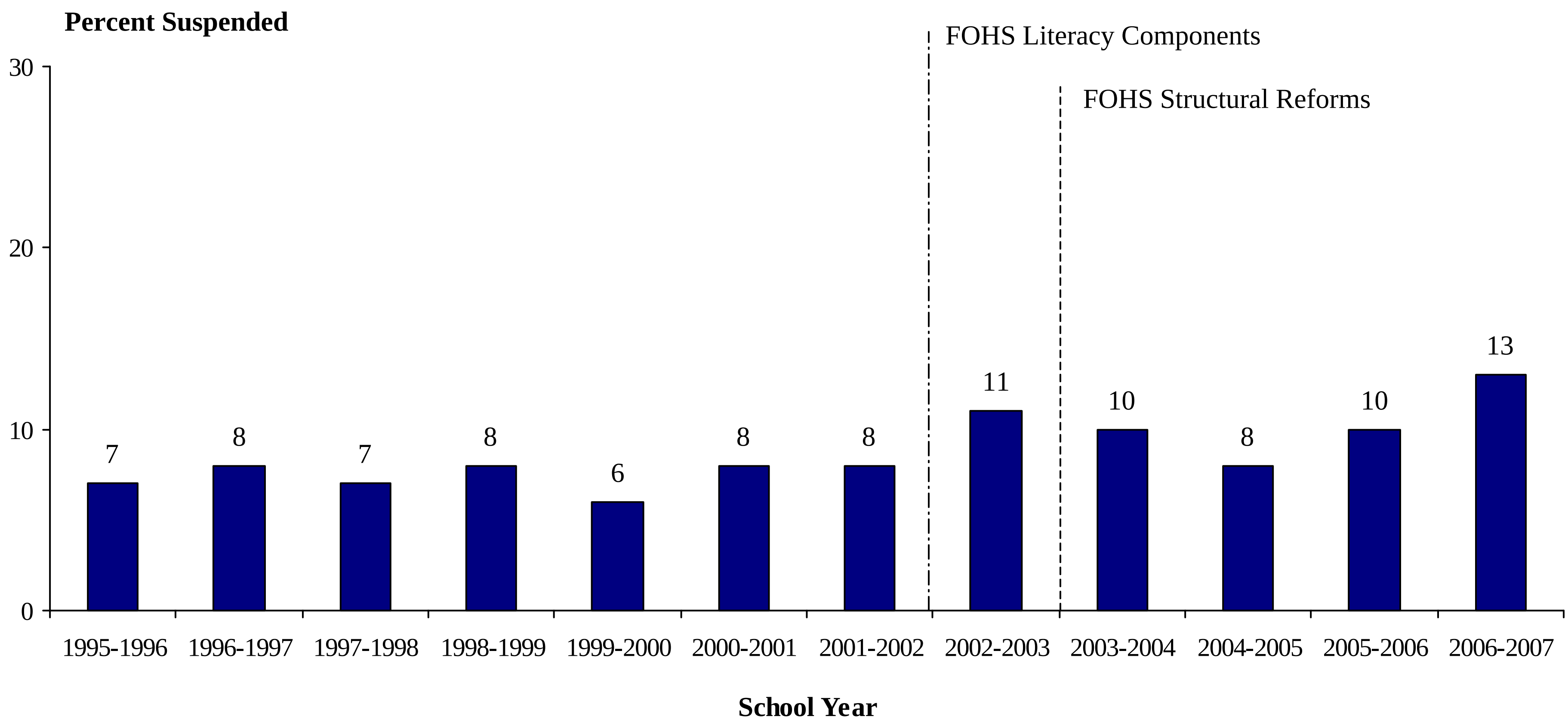


FIGURE II.15

MEAN NUMBER OF SUSPENSIONS

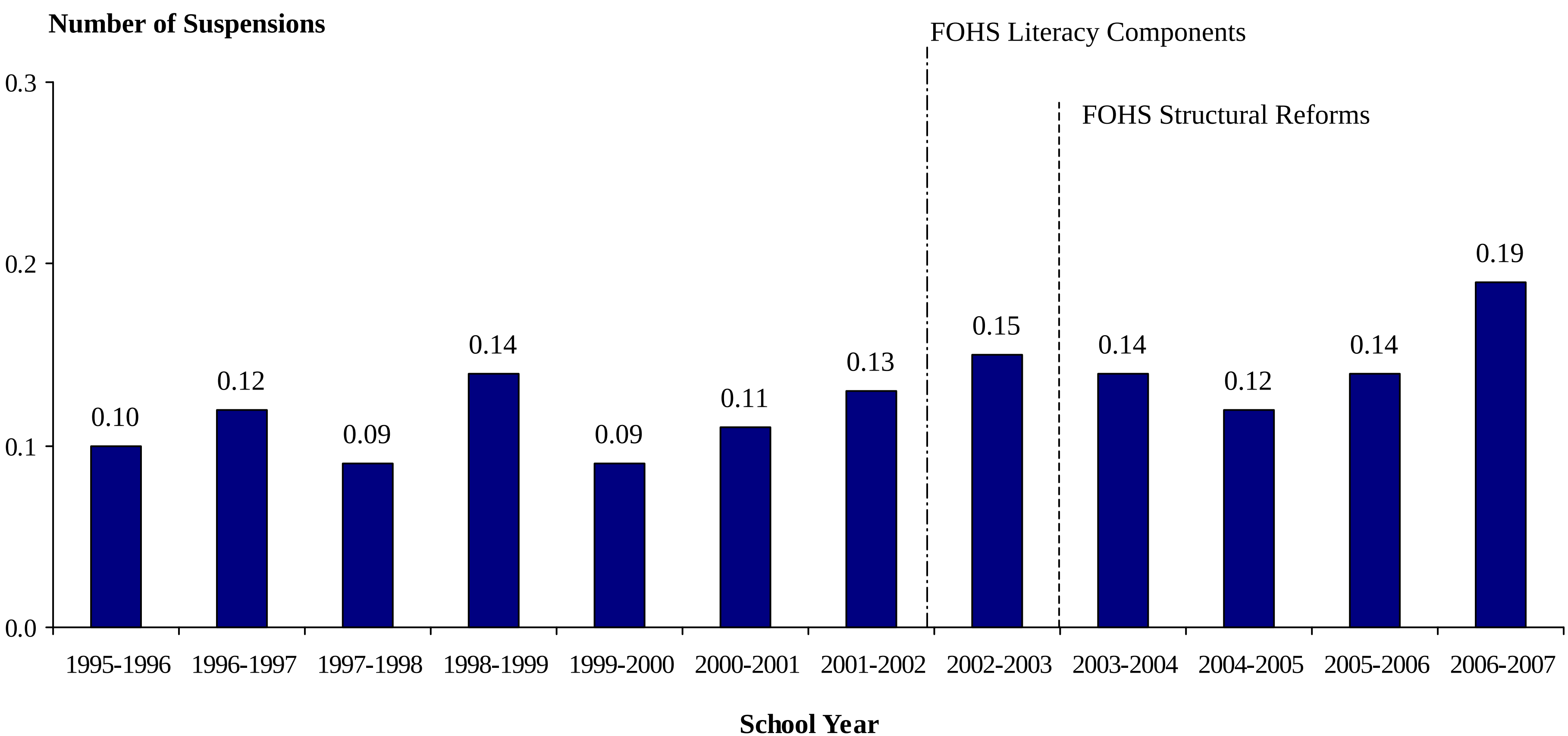
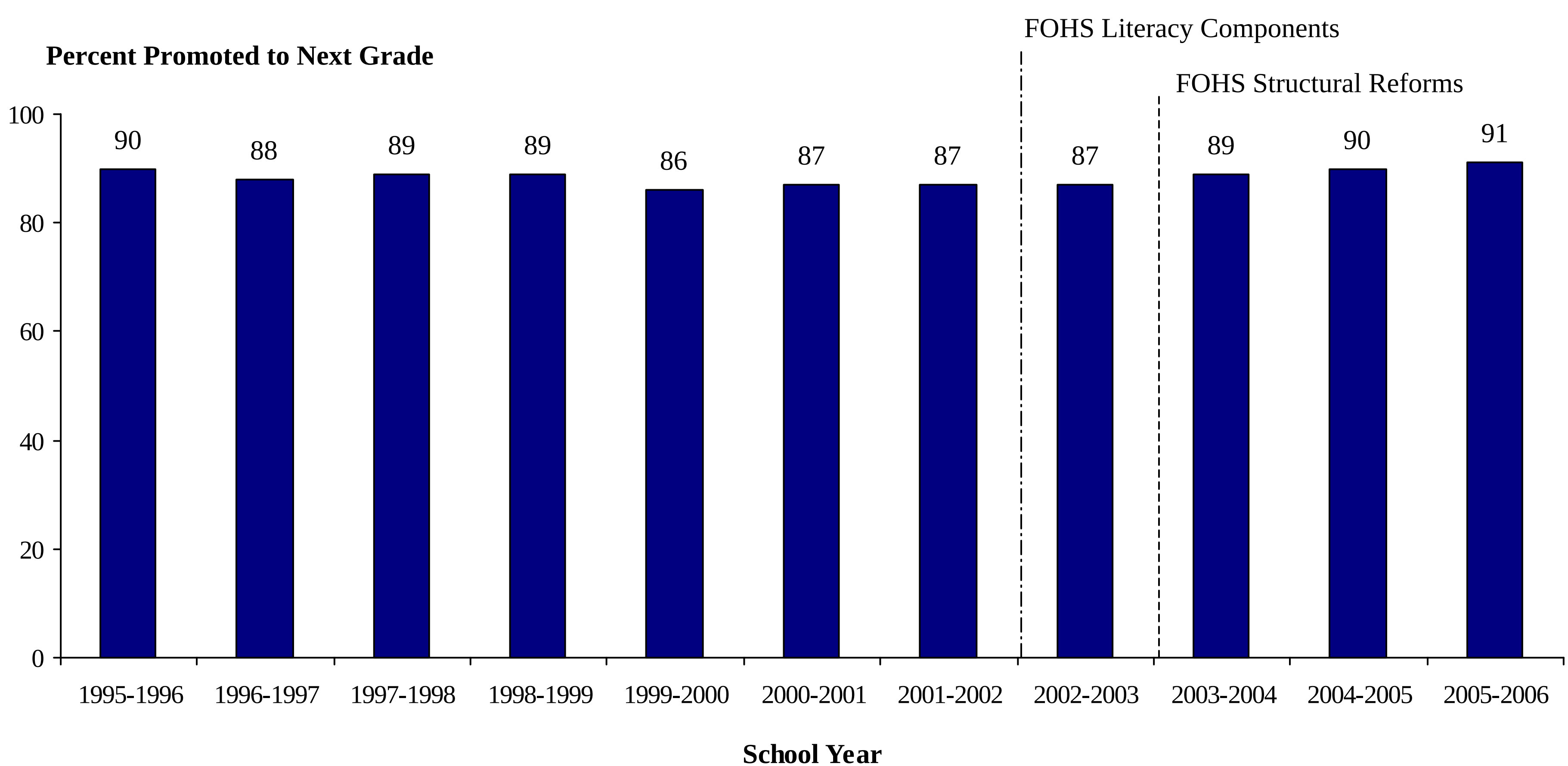


FIGURE II.16

PERCENTAGE OF STUDENTS PROMOTED TO THE NEXT GRADE



after the treatment was implemented. Effects may also manifest immediately or with some delay upon introduction or implementation of the treatment.

Despite their intuitive appeal, ITS designs cannot establish *causal* relationships between treatment and outcome variables. This means that, from our analyses, we cannot definitively conclude that FOHS is the *cause* of any observed changes in academic outcomes. The principal threat to the internal validity of ITS studies is that factors *other than the treatment* may have influenced outcomes at or around the same time as the treatment was introduced and, thus, may be wholly or partially responsible for any observed changes in outcomes (Shadish et al. 2002). Another possible threat results if there are changes in how outcomes are measured over time due to changes in administrative procedures, assessments used, or other factors. If this occurs, these changes may account for any observed changes in outcomes. Changes in the composition of the groups being examined can also be responsible for any observed outcome changes around or after the time of intervention. These possible threats to the internal validity must be considered carefully when interpreting the findings of any ITS study.

In ITS study designs, a "baseline" model—developed based on the pattern of outcomes for observations in the pretreatment period—must be specified. This model is used to predict the expected pattern of outcomes *in the absence of the intervention* during the post-treatment period. The effect of the intervention is then estimated as the difference between this predicted pattern of outcomes and the actual trend in outcomes.

The outcome measures we examined were of two different types. Some of the variables were *longitudinal*, meaning that, for a given student, data may be observed at multiple time points over a student's high school career. Outcome measures that fell into this "repeated measures" framework included, for example, whether a student was ever suspended, the number of absences, and grade promotion in a given academic year. In contrast, the MCAS outcome measures were *cross-sectional*. That is, the scale scores and performance levels in mathematics and ELA were available only for 10th graders each school year.[4]

Our analyses of all of these outcomes required the use of statistical methods that take into account the correlation among observations for students nested within groups (in this case, schools). For longitudinal outcomes, our analyses also had to take into account the correlation among repeated measures on the same individuals.

Multilevel models can be used to analyze data on individuals nested within groups, repeated measures on the same study subjects collected over time, or both types of data (Hox 2000; Raudenbush and Bryk 2002; Singer and Willett 2003). The general idea of multilevel modeling is to think of the lowest-level units as organized into a hierarchy of higher-level units. For example, in our cross-sectional models, students are "clustered" within schools. In our longitudinal models, multiple observations are clustered within a student (over time) and students are clustered within schools.

Appendix B provides additional details on the specification of the multilevel models used in the study. It also provides details on the methods used to generate both subgroup and school-level estimates of the effects of FOHS reforms.

[4] We restricted our analyses of MCAS outcomes to students' first administration of the 10th-grade exams. This was done so that our estimates of the effects of FOHS reforms would not be influenced by students' later attempts at achieving proficiency in the 10th-grade exams (a graduation requirement).

III. FINDINGS

This chapter presents findings from the interrupted time series analyses (ITS) of the effects of Focus on High School (FOHS) reforms on student academic outcomes. We begin with our findings for the full sample of BPS high schools and students. Next, we outline the results of analyses examining the effects of FOHS reforms for key subgroups of students, including by gender, race, English language use at home, receipt of special education services, and eligibility for free or reduced price lunch. Finally, we outline results for the 12 high schools present when the FOHS reforms were introduced. Each section includes a description of how report tables and figures should be interpreted.

Before turning to our discussion of study results, we remind readers of our cautionary notes regarding the interpretation of FOHS effects presented in this report. First, readers should keep in mind that the findings presented in this report cannot establish a causal relationship between the FOHS reforms and any observed changes in students' academic outcomes. A second caution is that our estimates of the effects of FOHS reforms are model-based and these models may become increasingly likely to overstate or understate counterfactual outcomes the further one extrapolates beyond the period for which pre-FOHS outcomes were observed. Thus, our estimates of the possible effects of FOHS reforms on student outcomes—particularly estimates for the later years in the study—should be interpreted with caution.

A. FULL SAMPLE ANALYSIS

This section summarizes our findings regarding changes in academic outcomes for the full sample of BPS students and high schools examined. Table III.1 presents our ITS estimates of the

TABLE III.1

ESTIMATES OF THE EFFECTS OF FOHS REFORMS ON MCAS TENTH GRADE OUTCOMES, BY ACADEMIC YEAR

	2002-2003		2003-2004		2004-2005		2005-2006		2006-2007	
	Effect	p-value	Effect	p-value	Effect	p-value	Effect	p-value	Effect	p-value
MCAS Tenth Grade Outcomes										
Scaled Scores										
Mathematics	0.44	0.46	-2.41***	<0.01	-5.26***	<0.01	-8.12***	<0.01	-10.97***	<0.01
English Language Arts	-1.42***	<0.01	-3.67***	<0.01	-5.92***	<0.01	-8.18***	<0.01	-10.43***	<0.01
Percent Proficient or Better										
Mathematics	0.04	0.15	-0.03	0.44	-0.11	0.06	-0.18**	0.01	-0.21***	<0.01
English Language Arts	-0.10***	<0.01	-0.18***	<0.01	-0.28***	<0.01	-0.37***	<0.01	-0.43***	<0.01

Note: For all outcomes except the MCAS scaled scores, this table shows the effects of FOHS at the reference category. See Appendix B for more details on the reference category.

[a]NA is indicated for 2006-2007 because the administrative records did not include data on promotions in that academic year.

**Test statistically significant at the .05 level.
***Test statistically significant at the .01 level.

changes in MCAS outcomes during the post-FOHS school years *relative to what would have been expected in the absence of these reforms.* Table III.2 presents similar ITS estimates of FOHS effects for other student academic outcomes.[1] Our estimates of FOHS effects for these two types of outcomes are presented in separate tables since FOHS literacy and structural reforms were not all introduced at the same time (see Chapter II).[2]

[1] Our ITS models for MCAS proficiency rates and other student academic outcomes estimate the effects of FOHS reforms using various mathematical transformations depending on the type of outcome examined (for example, log-odds for "count" outcomes such as number of days present). (Appendix B provides details about the various types of ITS models we estimated.) To facilitate interpretation of the overall effects, Tables III.1 and III.2 translate the estimated effects back into the units in which outcomes were reported originally in the BPS administrative records data.

[2] As discussed in Chapter II, FOHS reforms comprised both literacy-focused strategies and small-school "structural" reforms. We expected literacy-focused FOHS reforms, introduced during the 2002-2003 school year, to have their strongest effects on MCAS student achievement outcomes. In contrast, we expected other student academic outcomes examined (for instance, student attendance or suspensions) to be influenced more strongly by small-school FOHS reforms, which began one year later, in 2003-2004. Reflecting these expectations, our ITS models use 2002-2003 as the cutpoint for MCAS outcomes and 2003-2004 as the cutpoint for other student outcomes.

TABLE III.2

ESTIMATES OF THE EFFECTS OF FOHS REFORMS ON OTHER STUDENT OUTCOMES, BY ACADEMIC YEAR

	2003-2004		2004-2005		2005-2006		2006-2007	
	Effect	p-value	Effect	p-value	Effect	p-value	Effect	p-value
Other Student Outcomes								
Number of days present	-1.60***	<0.01	-2.41***	<0.01	-2.47***	<0.01	-1.89***	<0.01
Number of days absent	-0.37***	<0.01	-1.49***	<0.01	-3.47***	<0.01	-6.62***	<0.01
Number of days absent, unexcused	-0.27***	<0.01	-1.48***	<0.01	-4.16***	<0.01	-8.96***	<0.01
Number of days tardy	-0.54***	<0.01	-1.20***	<0.01	-2.11***	<0.01	-3.45***	<0.01
Percent suspended	-0.02***	<0.01	-0.05***	<0.01	-0.08***	<0.01	-0.13***	<0.01
Number of suspensions	-0.02***	<0.01	-0.04***	<0.01	-0.07***	<0.01	-0.09***	<0.01
Percent promoted to next grade	0.02***	<0.01	0.03***	<0.01	0.05***	<0.01	NA[a]	NA[a]

Note: For all outcomes, this table shows the effects of FOHS at the reference category. See Appendix B for more details on the reference category.

[a]NA is indicated for 2006-2007 because the administrative records did not include data on promotions in that academic year.

**Test statistically significant at the .05 level.
***Test statistically significant at the .01 level.

For example, the estimate shown in Table III.2 for the effect of FOHS on grade promotion in 2003-2004 (that is, percent promoted to next grade) indicates that promotion rates for the reference student in that academic year were two percentage points (0.02) higher than one would have projected in the absence of the FOHS reforms. (Appendix B provides additional details on the reference student in our ITS models.) Similarly, the estimated effect on number of days absent in 2003-2004 suggests a reduction in total absences of about 0.4 days.

Figures III.1 through III.11 depict our estimates of the overall effects of FOHS reforms on student outcomes. There are three lines of interest in each figure:

1. The solid line shows the pre-FOHS pattern of the outcome.
2. The dashed *projection* line shows the pattern of outcomes we would expect to observe, given pre-FOHS trends, *in the absence of the FOHS reforms*.
3. The solid line with triangles represents the *actual* pattern of outcomes observed during the period after the FOHS reforms began (after appropriate statistical adjustments).

The distance between the projection (dashed) line and the post-FOHS line in each figure represents the estimated effect of FOHS reforms in a given academic year.

Next, we discuss our results for the overall sample according to the two main types of outcomes examined: (1) student achievement (MCAS) and (2) other student outcomes.

1. Student Achievement

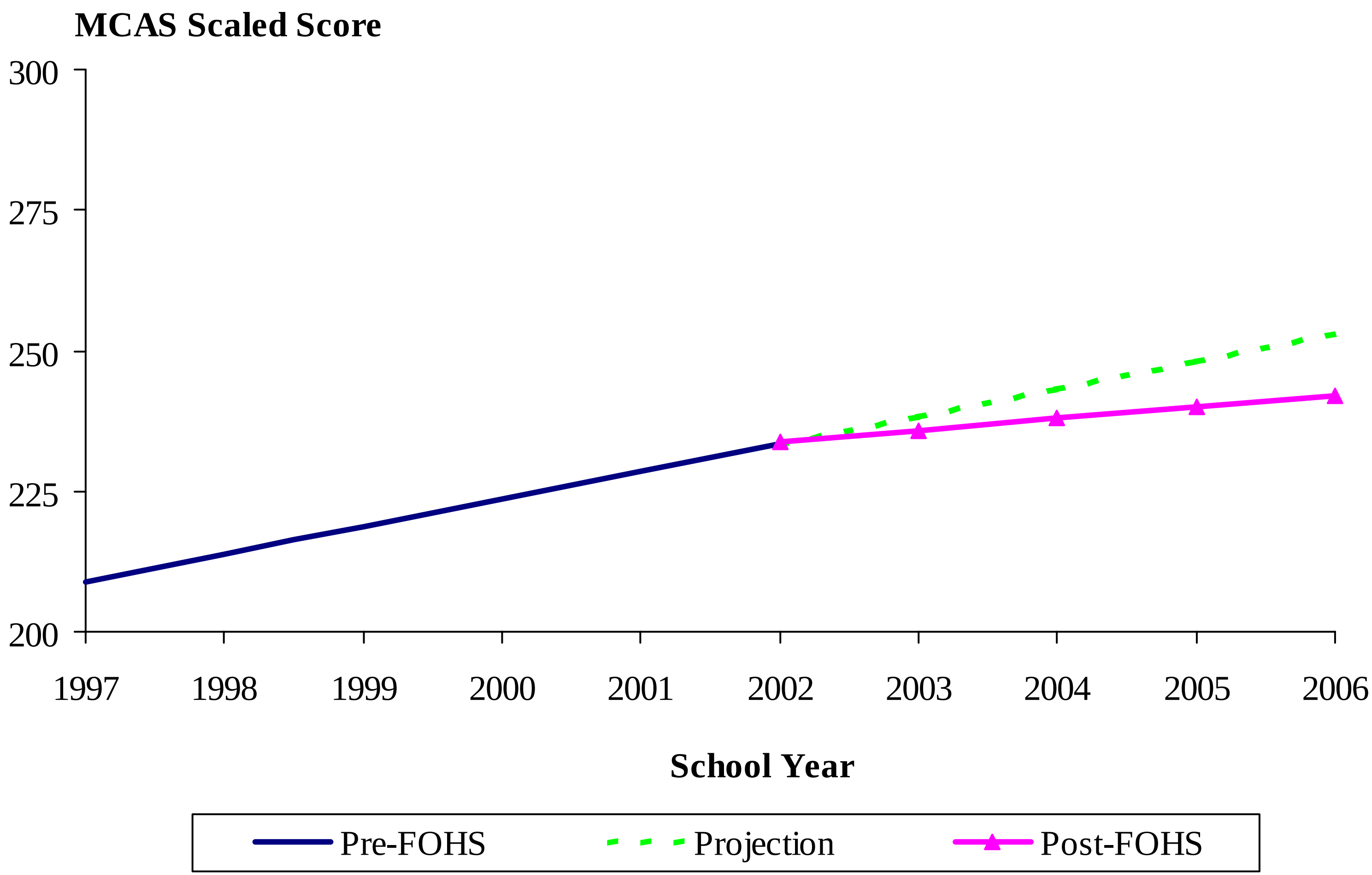
In general, our analyses suggest that student achievement declined somewhat after the introduction of FOHS relative to what one would have predicted in the absence of the reforms.

MCAS Scaled Scores. After statistical adjustments, the MCAS tenth grade scaled scores in mathematics observed during the post-FOHS reform academic years were between 0 and 11 points lower than the scores one would have projected based on pre-FOHS trends (Table III.1 and Figure III.1). The estimated reductions in MCAS tenth grade scaled scores in ELA were similar to the estimated reductions in the mathematics scaled scores, between 1 and 10 points (Figure III.2). Note that all estimated differences, except for Mathematics in 2002-2003 (when FOHS literacy components were introduced), were statistically significant (Table III.1).

Proficiency Rates. Our analyses also suggest post-FOHS declines in the proportion of students scoring at a proficiency level or higher in the MCAS tenth grade exams. In mathematics, MCAS proficiency rates were estimated to be 18 and 21 percent *lower* in 2005-2006 and 2006-2007, respectively, relative to the rates one would have projected in the absence of the FOHS reforms (Table III.1 and Figure III.3). Estimated FOHS effects in earlier years were not statistically significant (that is, were not distinguishable from zero effect). In ELA, larger reductions in proficiency rates were estimated, between 10 and 43 percent (Table III.1 and Figure III.4). In ELA, all estimated differences were statistically significant (Table III.1).

FIGURE III.1

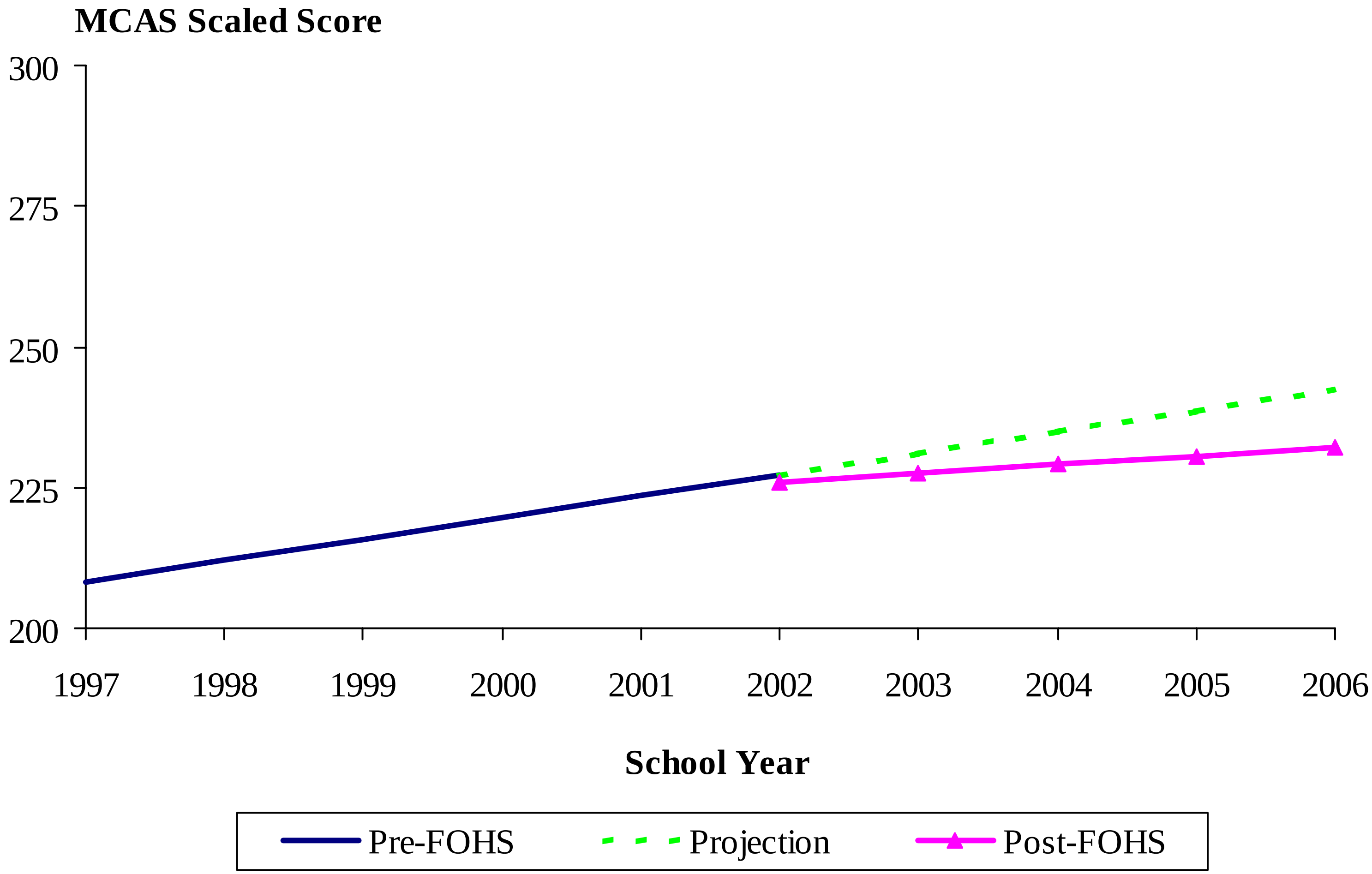
EFFECT OF FOHS ON MCAS TENTH GRADE MATHEMATICS SCALED SCORES



Note: The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the horizontal axis of the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

FIGURE III.2

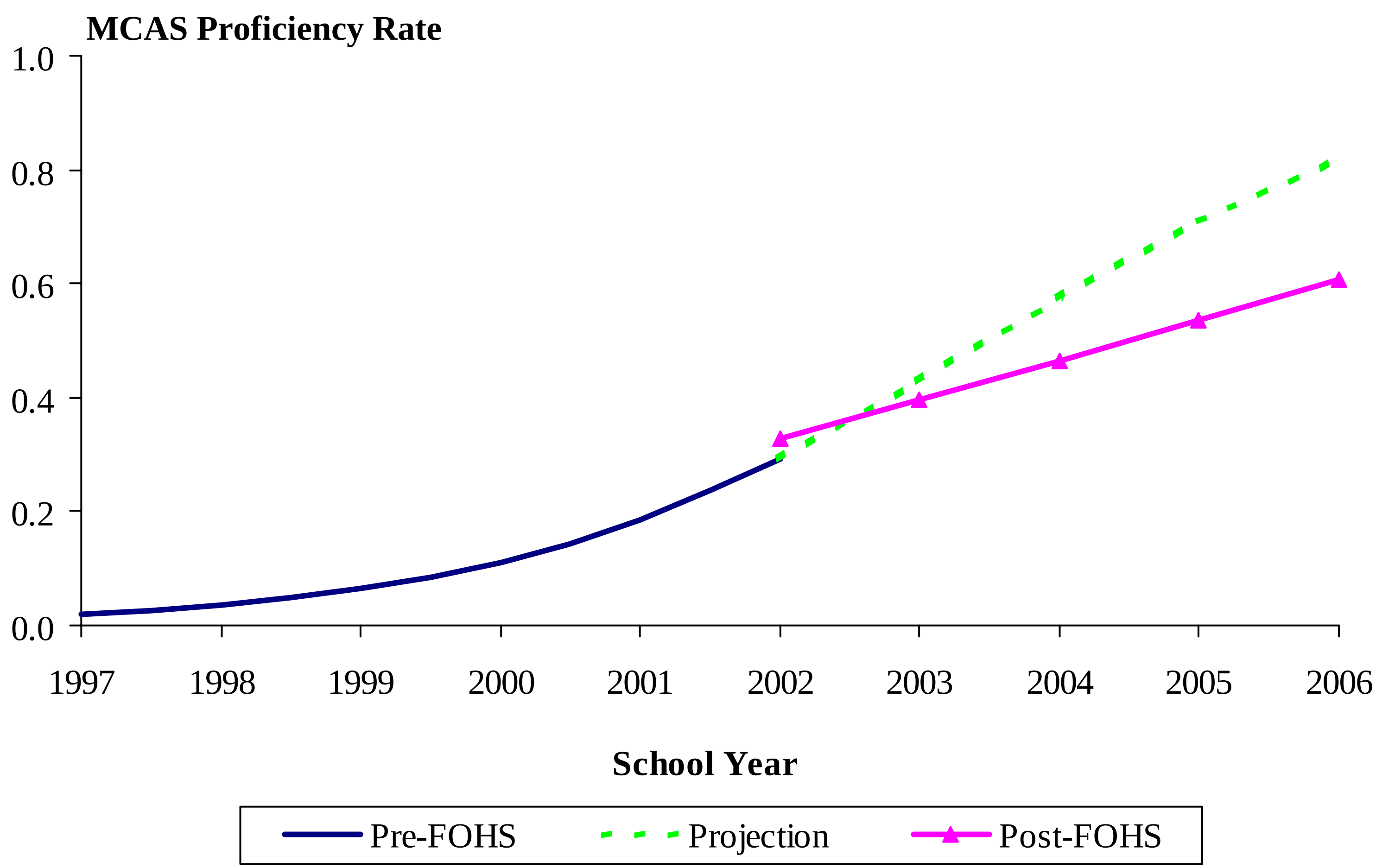
EFFECT OF FOHS ON MCAS TENTH GRADE ELA SCALED SCORES



Note: The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the horizontal axis of the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

FIGURE III.3

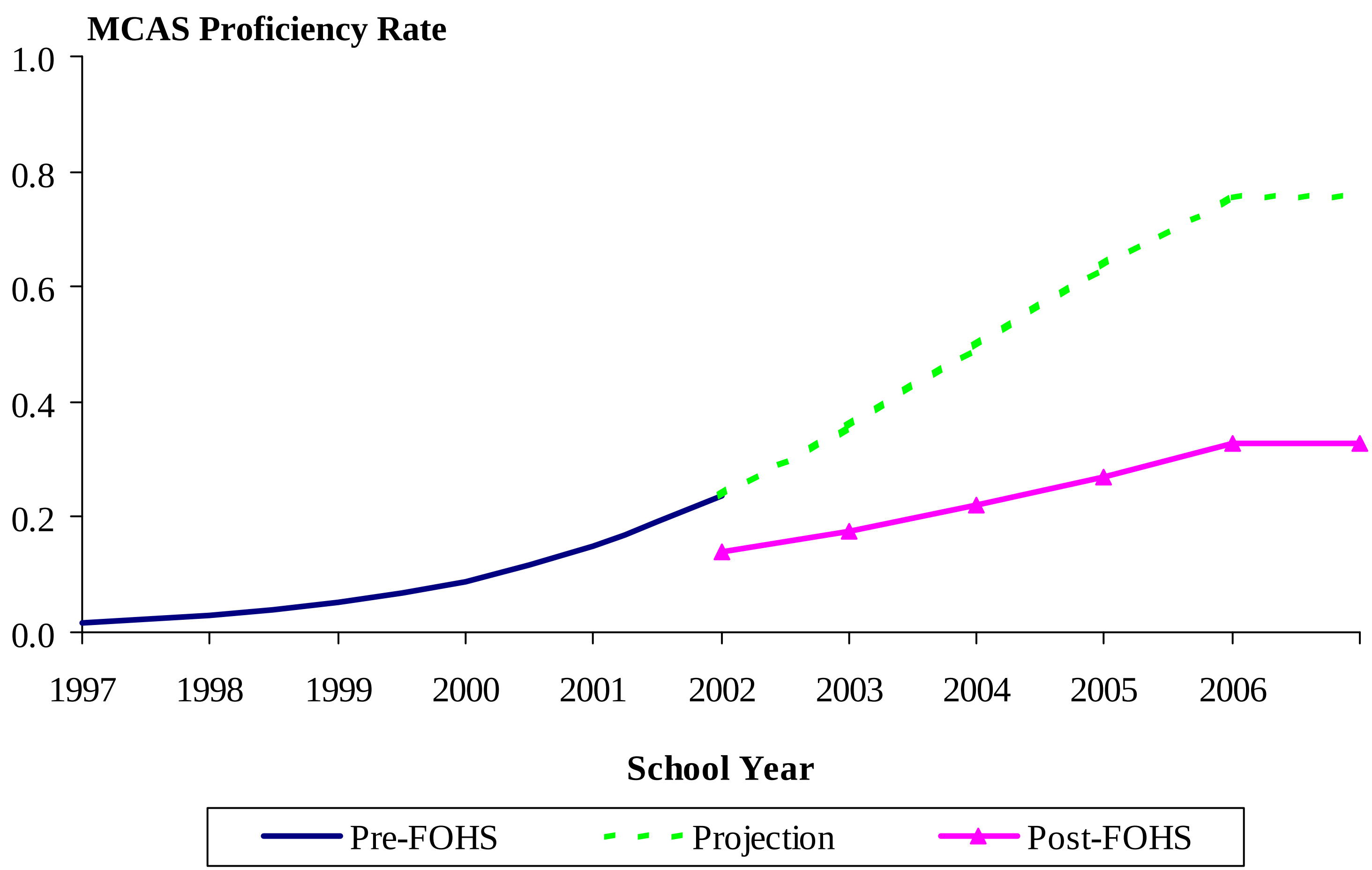
EFFECT OF FOHS ON MCAS TENTH GRADE PROFICIENCY IN MATHEMATICS



Note: The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the horizontal axis of the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

FIGURE III.4

EFFECT OF FOHS ON MCAS TENTH GRADE PROFICIENCY IN ELA



Note: The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the horizontal axis of the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

2. Other Academic Outcomes

In contrast to the achievement results, our ITS analyses suggest improvements in most other student academic outcomes after implementation of FOHS relative to expectations in the absence of these reforms. In this section, we discuss findings regarding the post-FOHS changes in student absences, attendance, tardiness, suspensions, and grade promotion.

Attendance. We examined the effects of FOHS on both *total* absences (excused and unexcused) and *unexcused* absences, and our models suggested changes in both of these outcomes that would be consistent with the goals of the reforms. Total absences were estimated to *decline* by between 0.4 and 6.6 days relative to what one would have predicted in the absence of the FOHS reforms (Table III.2 and Figure III.5). Unexcused absences also were estimated to decline by between 0.3 and 9.0 days (Table III.2 and Figure III.6).[3]

Our analyses also suggested modest declines in student attendance, which would be contrary to FOHS goals. After the FOHS reforms, school attendance declined between 1.6 and 2.5 days relative to what one would have expected in the absence of the FOHS reforms (Table III.1 and Figure III.7). Given the estimated reductions in student absences, this was a puzzling result that we investigated a bit further. These analyses suggested that the reduction in absences was likely driven by students with truncated attendance histories and, hence, should be interpreted with caution.[4]

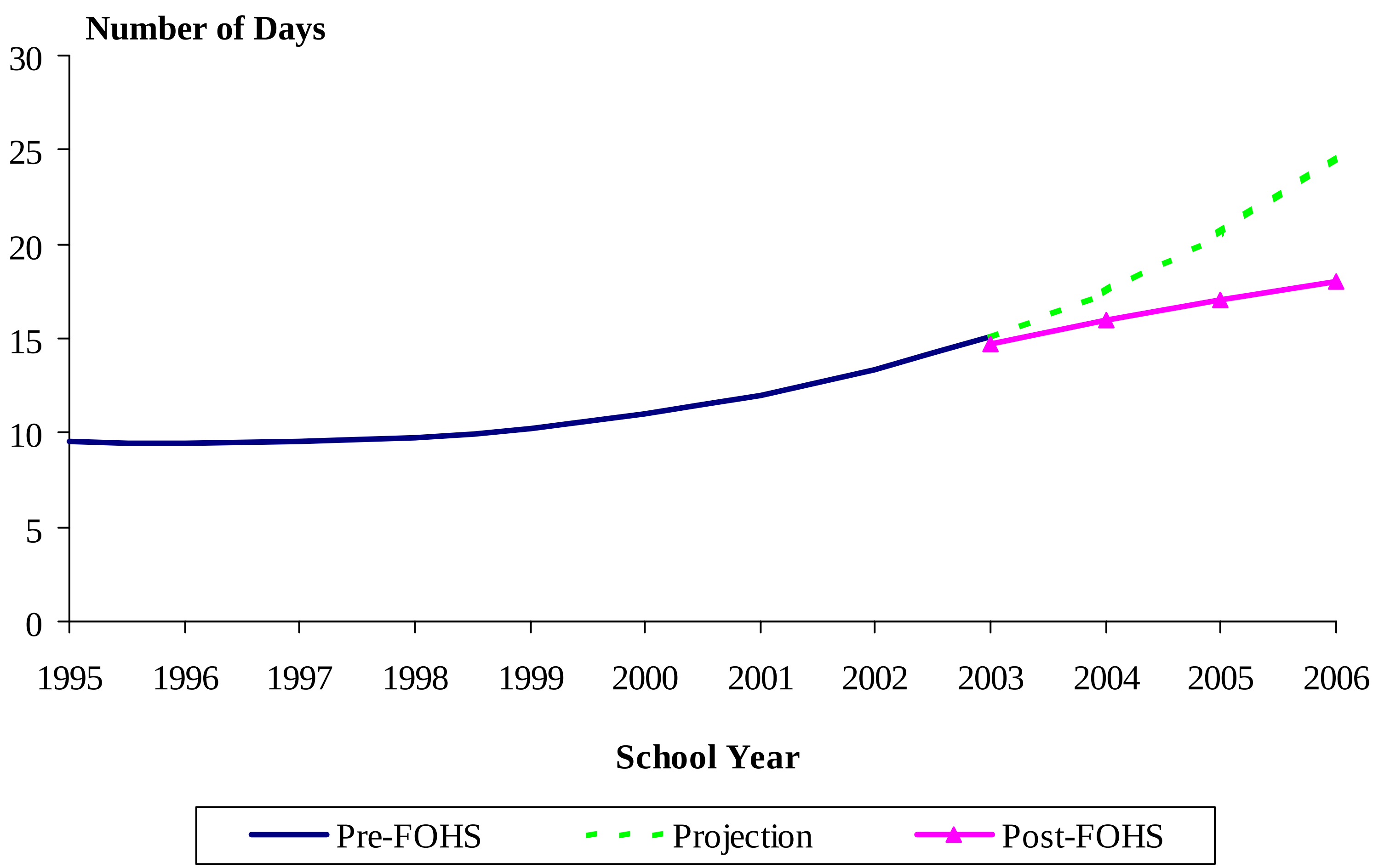
---

[3] Students with number of days absent and number of days present both equal to zero were excluded from our analyses of attendance outcomes (that is, total absences, unexcused absences, and total days attended).

[4] More specifically, we examined whether our estimates of the effects of FOHS reforms on attendance and (total) absences differed between students for whom “days absent” + “days present” = 180 versus students for whom “days absent” + “days present” < 180, given that many of the students in our sample appeared to have truncated attendance records. For students with complete attendance, the estimated effect on “days absent” was opposite what was found with the full sample (that is, an increase in “days absent”), while we still found a decrease in “days present.” These results are more consistent with what one would expect as we observed a reduction in attendance combined with an increase in absences. This suggests that the reduction in absences estimated for the overall sample was likely driven by students with truncated attendance histories.

FIGURE III.5

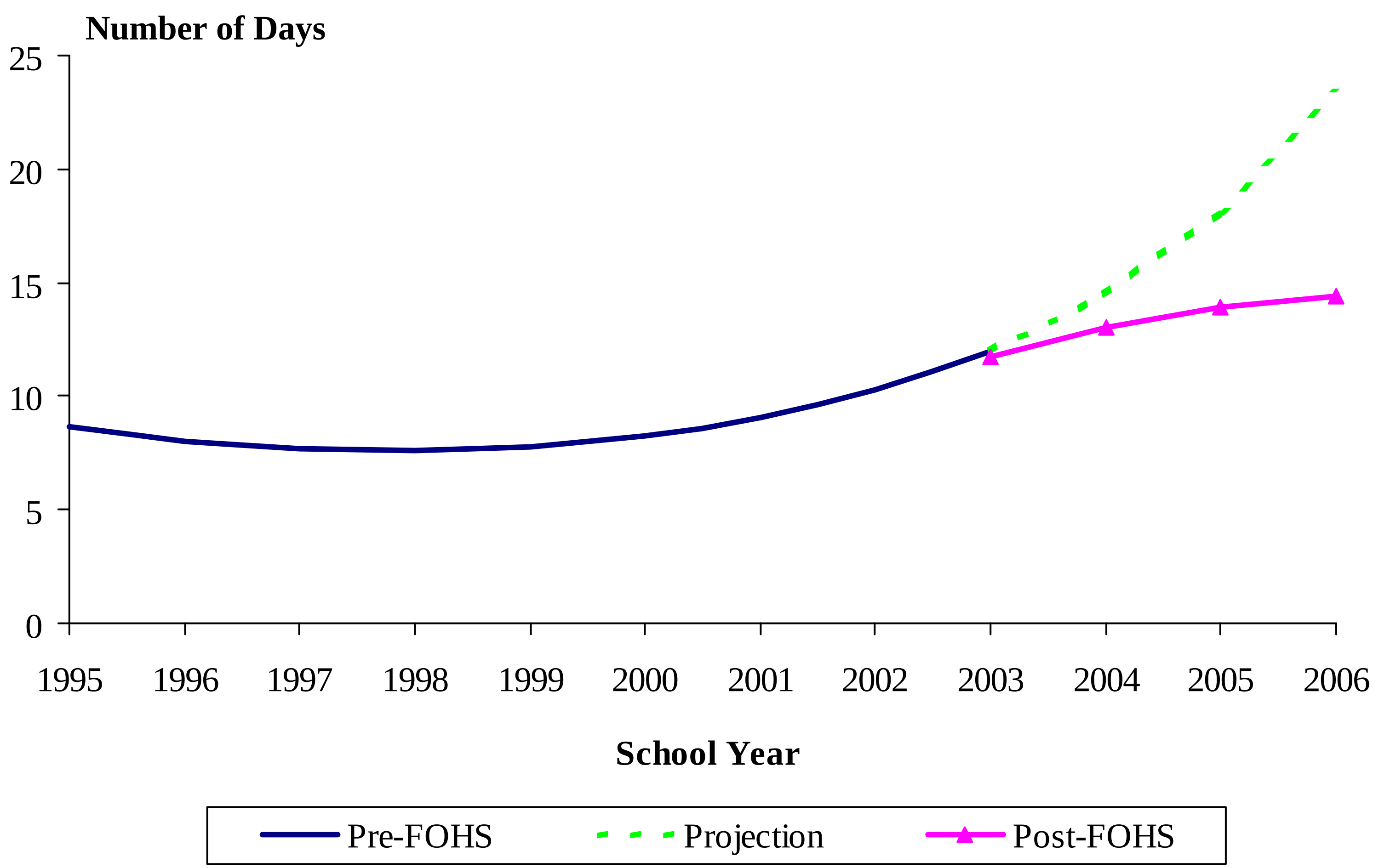
EFFECT OF FOHS ON TOTAL ABSENCES



Note: Depicted trajectories are for the reference category group in the corresponding statistical model. The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

FIGURE III.6

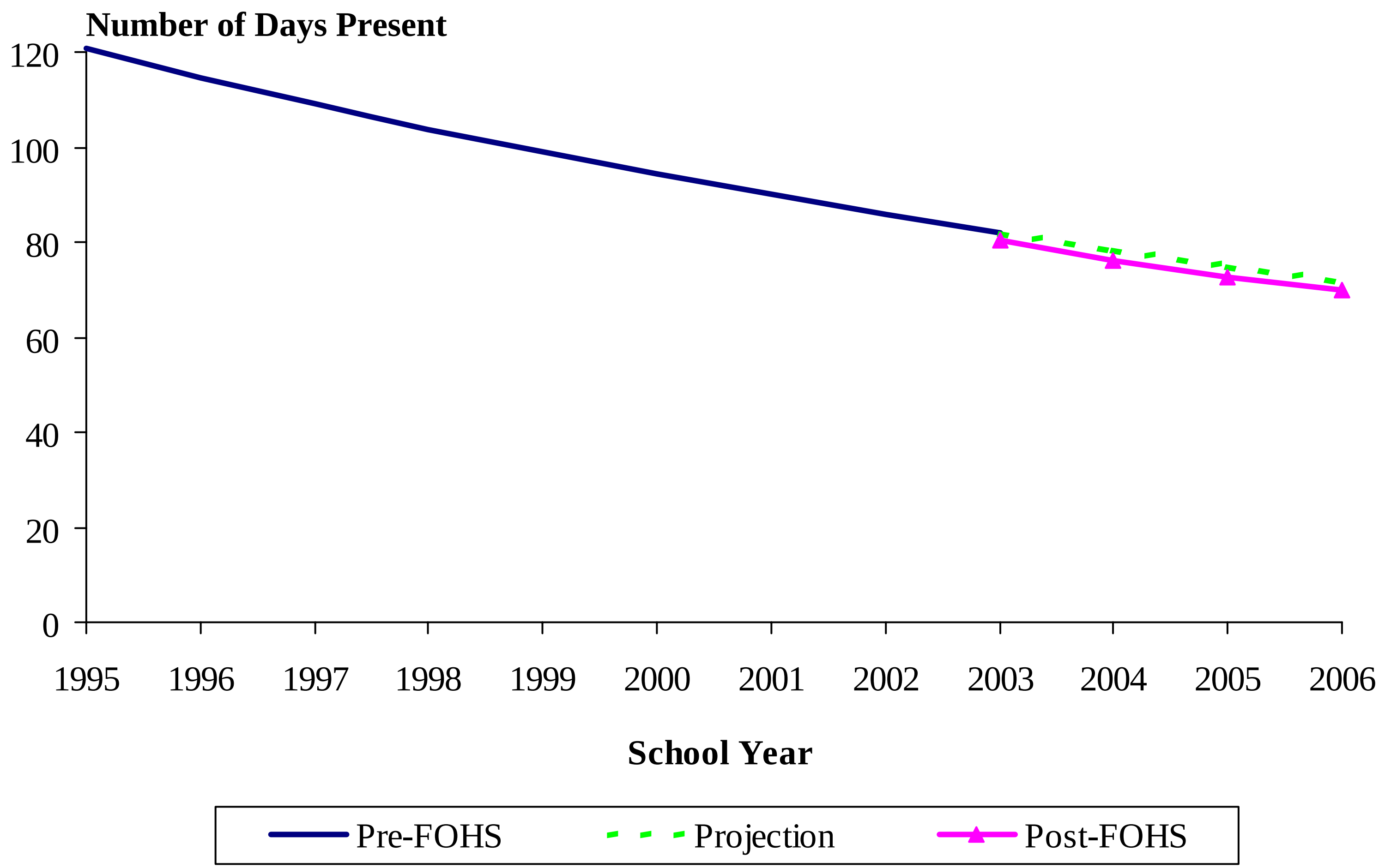
EFFECT OF FOHS ON UNEXCUSED ABSENCES



Note: Depicted trajectories are for the reference category group in the corresponding statistical model. The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

FIGURE III.7

EFFECT OF FOHS ON ATTENDANCE



Note: Depicted trajectories are for the reference category group in the corresponding statistical model. The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

Tardiness. Our analyses suggested reductions in student tardiness during the post-FOHS reform period. The model estimated that students were tardy between 0.5 and 3.5 *fewer* days relative to the number of days projected in the absence of the FOHS reforms (Table III.2 and Figure III.8).

Suspensions. We investigated the effects of FOHS reforms on (1) students' probability of being suspended in a given school year and (2) the total number of suspensions per student. Both sets of models suggested changes in outcomes consistent with the goals of the FOHS reforms. The probability of suspension was estimated to decrease by between 2 and 13 percentage points relative to projections in the absence of the FOHS reforms (Table III.2 and Figure III.9). The number of suspensions per student was also estimated to decrease significantly, although the mean number of suspensions per student was already low (Table III.2 and Figure III.10).

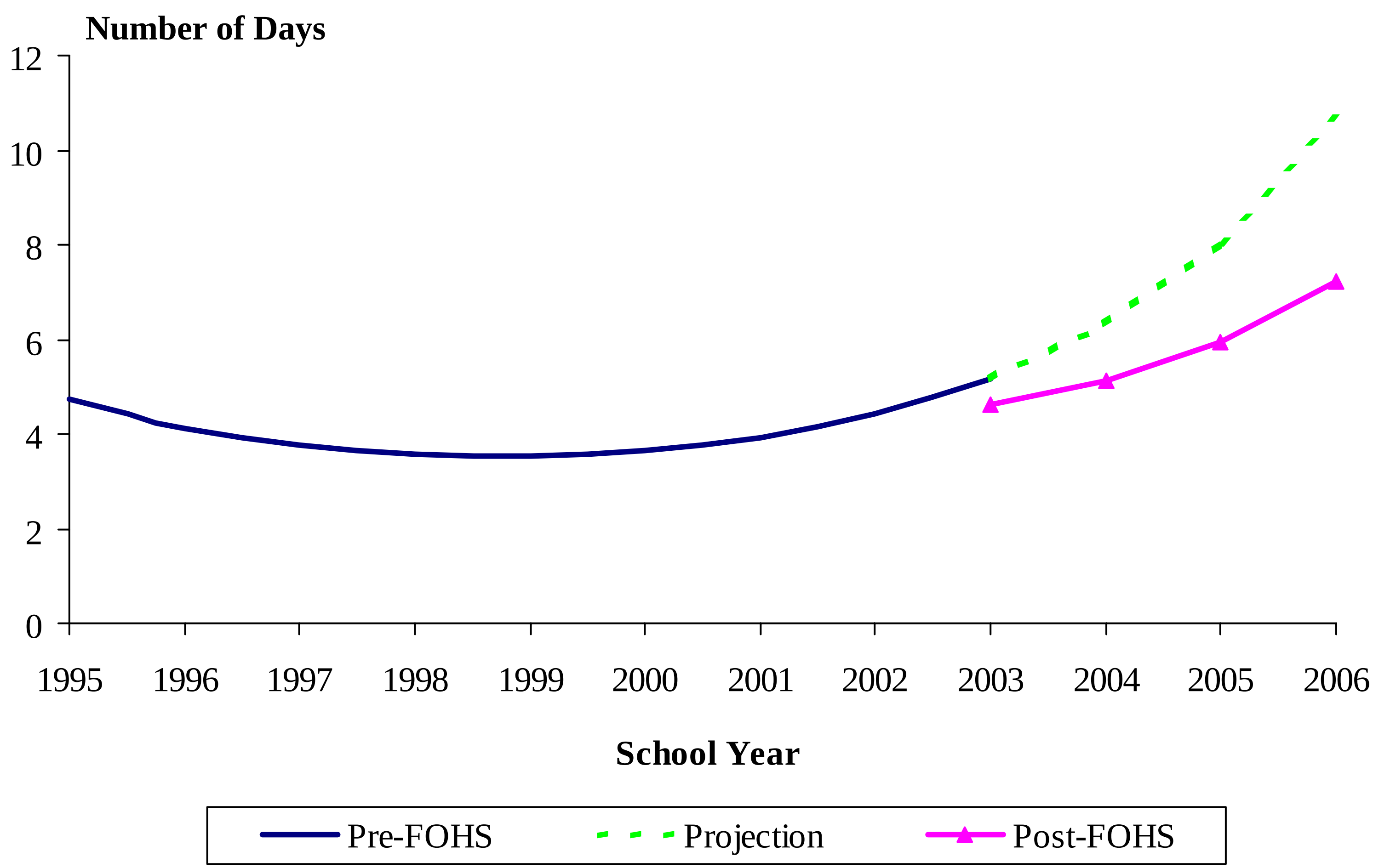
Grade Promotion. Lastly, our models suggested modest increases in students' probability of promotion to their next grade. Students were between 2 and 5 percent more likely to be promoted during the post-FOHS reform period relative to projections in the absence of the FOHS reforms (Table III.2 and Figure III.11).

B. SUBGROUP ANALYSES

In addition to investigating the post-FOHS changes in outcomes for the full sample of students, we also examined outcome changes for key subgroups of students. The subgroup analyses focused on the extent to which outcome changes observed in the overall sample were consistently observed across student subgroups. Our analyses showed that some student subgroups experienced significantly different outcome changes relative to other student subgroups after the implementation of FOHS.

FIGURE III.8

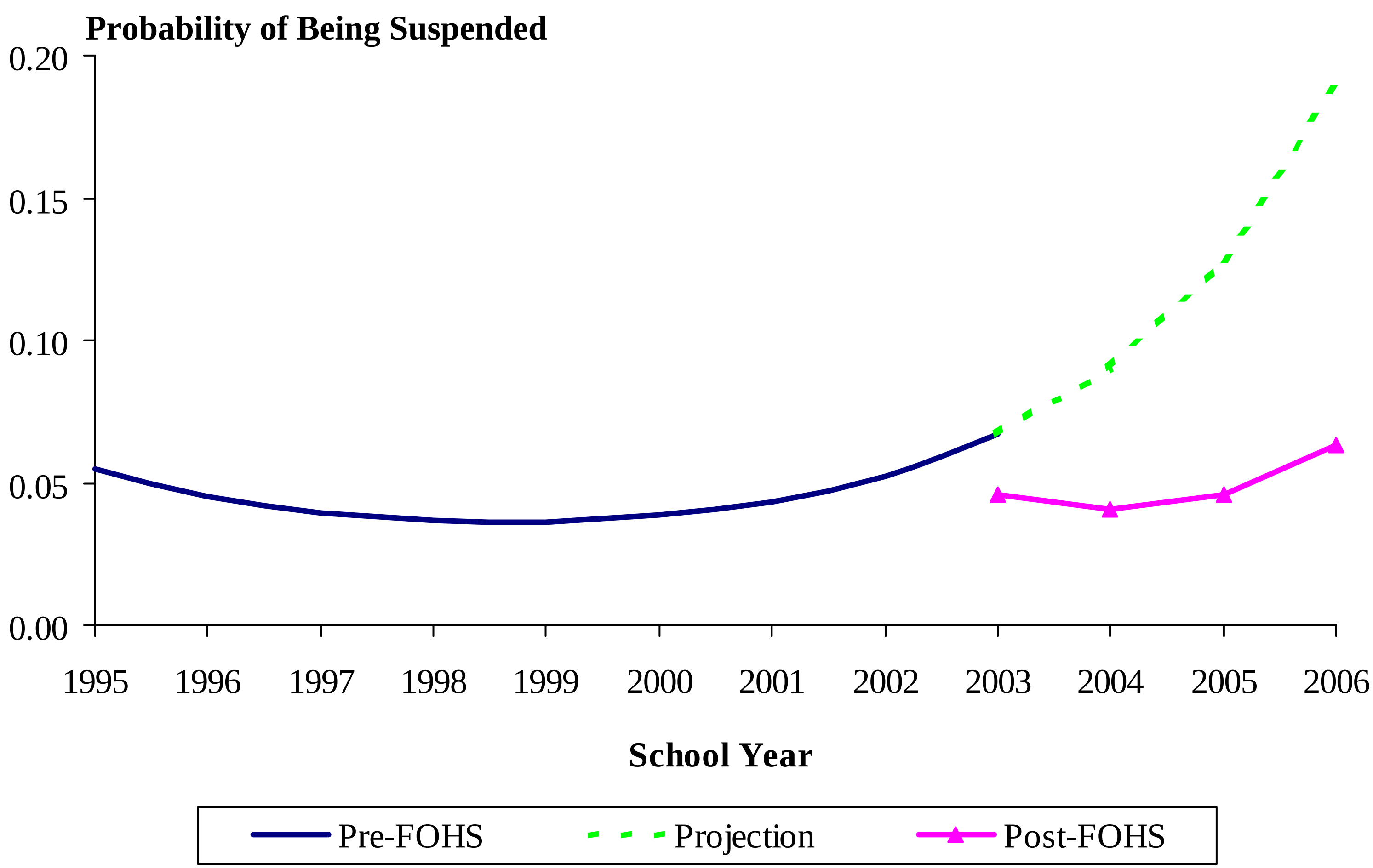
EFFECT OF FOHS ON TARDINESS



Note: Depicted trajectories are for the reference category group in the corresponding statistical model. The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

FIGURE III.9

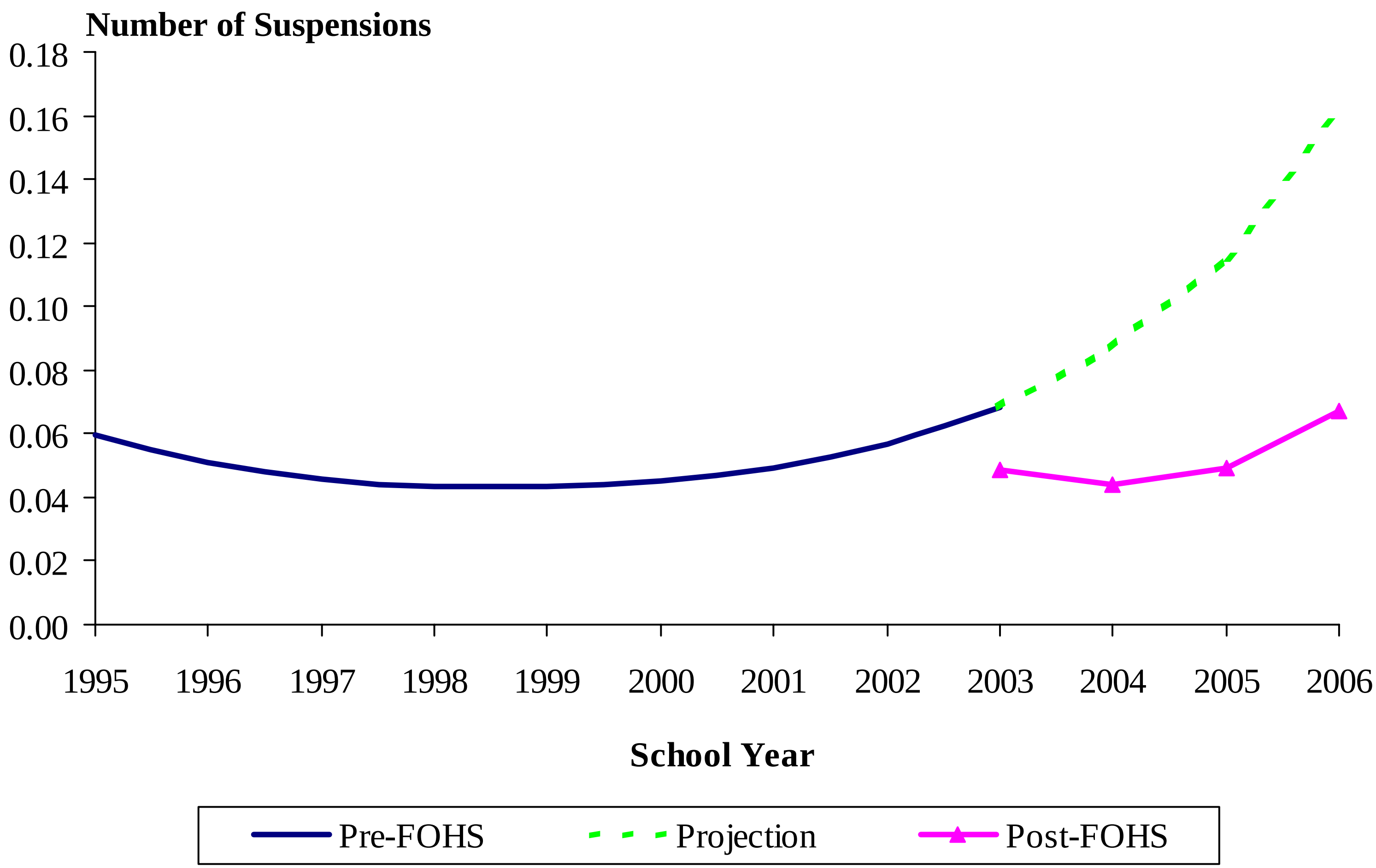
EFFECT OF FOHS ON THE PROBABILITY OF BEING SUSPENDED



Note: Depicted trajectories are for the reference category group in the corresponding statistical model. The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

FIGURE III.10

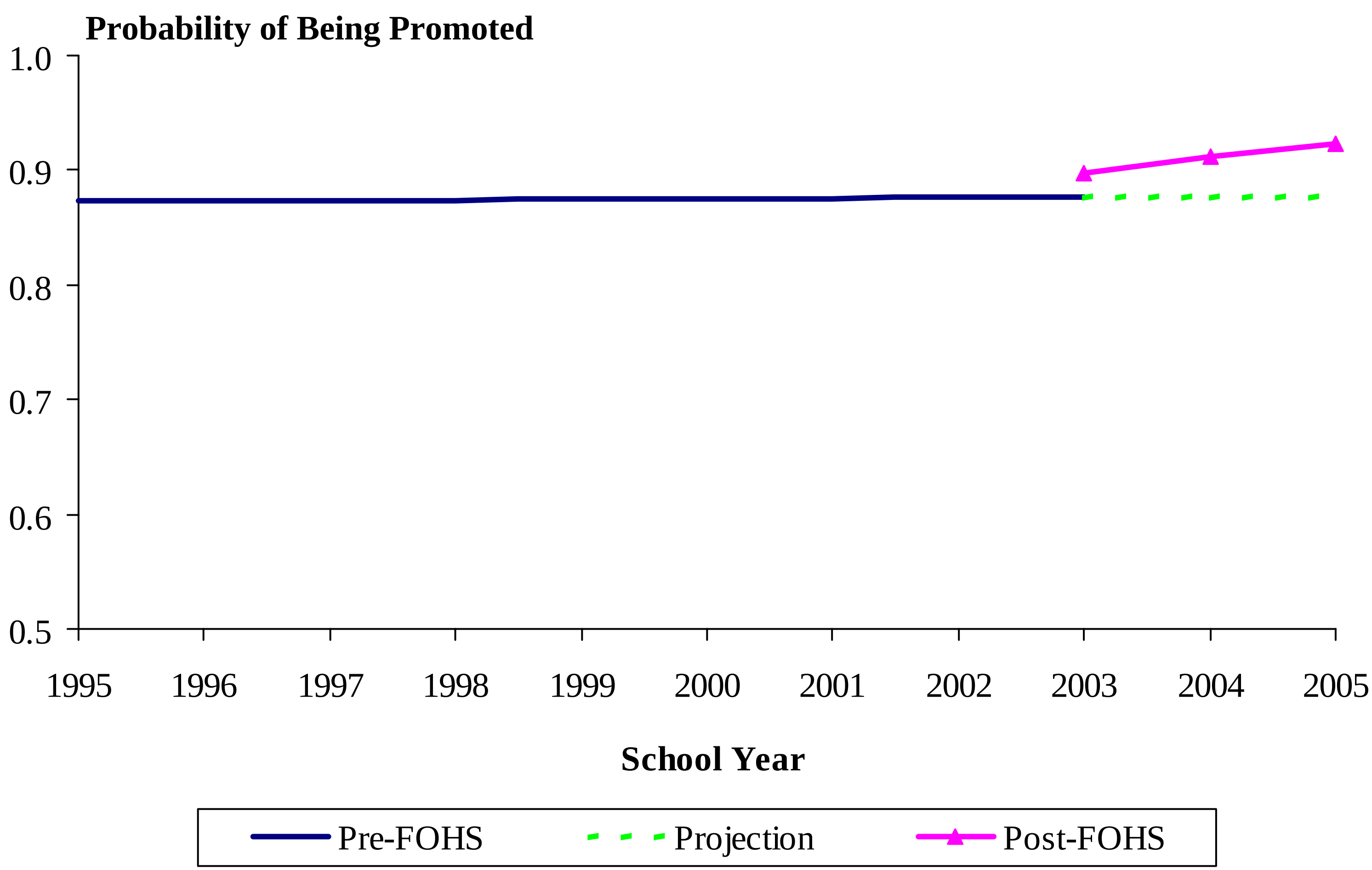
EFFECT OF FOHS ON THE NUMBER OF SUSPENSIONS



Note: Depicted trajectories are for the reference category group in the corresponding statistical model. The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

FIGURE III.11

EFFECT OF FOHS ON GRADE PROMOTION



Note: Depicted trajectories are for the reference category group in the corresponding statistical model. The distance between the projection (dashed) line and the post-FOHS line represents the estimated effect of FOHS reforms. The years in the figure refer to academic years. For example, 1999 refers to the 1999-2000 academic year.

To facilitate the examination of the consistency of the estimated effects across subgroups, Tables III.3 and III.4 present the coefficients generated by our ITS models. The "reference group" column in each table provides the estimates for the omitted group in each subgroup category (for example, males for gender and white students for race-ethnicity). These estimates can be contrasted with those for other student categories within each subgroup to determine if their estimated outcome changes were statistically different and smaller or larger (as compared to the omitted group).

Consider, for example, the estimates for unexcused absences (that is, number of days absent, unexcused) in Table III.3. The -0.05 model coefficient noted for the "reference group" suggests a *significant decrease* in unexcused absences for the omitted category in our subgroup analysis. The model coefficient for females (0.03) then suggests smaller *decreases* in unexcused absences for female students (because the sum of the coefficients for the reference group and females is still negative) and the associated p-value indicates that this difference (females relative to males) is statistically significant.[5] Considering race-ethnicity, the model coefficients indicate that *decreases* in unexcused absences were significantly *larger* for black students as compared to white students (because the model coefficient for black students, -0.09, is statistically significant), but comparable between white and Hispanic students (since the model coefficient for Hispanics, -0.01, is not significant).

1. Gender

Females and males experienced significantly different effects on ELA achievement, number of days present, number of unexcused absences, and number of days tardy. Both females and

[5] Note that our models take into account any pre-existing differences in the levels of outcomes across subgroups (for example, unexcused absences for males vs. females). Thus, they assess whether the FOHS reforms appear to have had a *differential* effect on outcomes for one student subgroup vs. the other.

TABLE III.3

ESTIMATED DIFFERENCES IN THE EFFECTS OF FOHS REFORMS, BY GENDER AND RACE

	Reference Group		Female		Black		Hispanic	
	Model Coefficient	p-value	Model Coefficient	p-value	Model Coefficient	p-value	Model Coefficient	p-value
MCAS Tenth Grade Outcomes								
Scaled Scores								
Mathematics	-0.41	0.75	-0.96	0.09	-8.80***	<0.01	-9.68***	<0.01
English Language Arts	-2.46**	0.04	2.11***	<0.01	-3.24***	<0.01	-2.93***	<0.01
Percent Proficient or Better[a]								
Mathematics	-0.16	0.60	-0.19	0.31	-1.41***	<0.01	-1.70***	<0.01
English Language Arts	-0.52	0.09	0.31**	0.04	-0.24	0.19	-0.17	0.44
Other Student Outcomes								
Number of days present[b]	-0.04***	<0.01	0.02***	<0.01	0.02***	<0.01	0.02***	<0.01
Number of days absent[b]	-0.06***	<0.01	0.00	0.72	-0.05***	<0.01	0.03***	<0.01
Number of days absent, unexcused[b]	-0.05***	<0.01	0.03***	<0.01	-0.09***	<0.01	-0.01	0.52
Number of days tardy[b]	0.07***	<0.01	0.06***	<0.01	-0.32***	<0.01	-0.17***	<0.01
Percent suspended[a]	-0.39	0.07	-0.09	0.42	0.17	0.35	0.02	0.90
Number of suspensions[b]	-0.58***	<0.01	-0.03	0.74	0.36**	0.02	0.17	0.31
Percent promoted to next grade[a]	0.15	0.20	0.02	0.78	-0.12	0.21	-0.09	0.37

Note: See Appendix B for more details on the reference category.

[a]The estimated coefficient is a log-odds.

[b]The estimated coefficient is the log of an event rate.

**Test statistically significant at the .05 level.
***Test statistically significant at the .01 level.

TABLE III.4

ESTIMATED DIFFERENCES IN THE EFFECTS OF FOHS REFORMS, BY HOME LANGUAGE, SPECIAL EDUCATION, AND ELIGIBILITY FOR FREE OR REDUCED-PRICE LUNCH

	Reference Group		English Spoken at Home		Special Education		Free or Reduced-Price Lunch	
	Model Coefficient	p-value	Model Coefficient	p-value	Model Coefficient	p-value	Model Coefficient	p-value
MCAS Tenth Grade Outcomes								
Scaled Scores								
Mathematics	-0.41	0.75	-1.04	0.11	-10.79***	<0.01	2.04***	<0.01
English Language Arts	-2.46**	0.04	5.13***	<0.01	-9.99***	<0.01	0.05	0.93
Percent Proficient or Better[a]								
Mathematics	-0.16	0.60	0.05	0.79	-1.77***	<0.01	0.33	0.11
English Language Arts	-0.52	0.09	1.00***	<0.01	-2.05***	<0.01	-0.08	0.63
Other Student Outcomes								
Number of days present[b]	-0.04***	<0.01	0.01**	0.02	-0.01***	0.01	0.01	0.08
Number of days absent[b]	-0.06***	<0.01	0.03***	<0.01	0.04 ***	<0.01	0.02***	0.01
Number of days absent, unexcused[b]	-0.05***	<0.01	0.03***	<0.01	0.06 ***	<0.01	0.00	0.61
Number of days tardy[b]	0.07***	<0.01	-0.01	0.31	-0.01	0.13	0.06***	<0.01
Any suspensions[a]	-0.39	0.07	-0.13	0.33	0.14	0.21	-0.05	0.66
Number of suspensions[b]	-0.58***	<0.01	-0.07	0.52	0.16	0.09	-0.06	0.48
Percent promoted to next grade[a]	0.15	0.20	0.10	0.20	0.32***	<0.01	0.08	0.26

Note: See Appendix B for more details on the reference category.

[a]The estimated coefficient is a log-odds.

[b]The estimated coefficient is the log of an event rate.

**Test statistically significant at the .05 level.
***Test statistically significant at the .01 level.

males had lower MCAS scaled scores and proficiency rates in ELA after the FOHS reforms, but the declines were smaller for females (Table III.3). Our analyses suggest that both males and females were present for fewer days after the implementation of FOHS (relative to baseline predictions), but the effect on females was smaller than the effect for males. Males and females both had fewer unexcused absences after FOHS, but the decrease among females was smaller. Males and females were both tardy more often after FOHS, but females were tardy even more than males. There were no significant differences between males and females in estimated post-FOHS changes in other outcomes.

2. Race-Ethnicity

Black and white students experienced significantly different effects on MCAS scaled scores in both mathematics and ELA, mathematics proficiency rates, number of days present, number of days absent, number of unexcused absences, number of days tardy, and number of suspensions. While FOHS reforms were estimated to have no effect on the mathematics MCAS scaled scores or proficiency rates of white students, our analyses suggest large decreases for black students (Table III.3). Both white and black students had lower MCAS scaled scores in ELA after FOHS, but estimated declines were larger among black students. Both black and white students were present for fewer days after the implementation of FOHS, but the estimated decrease was smaller for black students. Black and white students were both estimated to have fewer absences (including unexcused absences) after FOHS, but the reductions were significantly larger among black students. Black students were tardy significantly less often after FOHS, while white students were tardy more often. Black and white students both had fewer suspensions after FOHS, but the reduction was smaller for black students. There were no significant differences between black and white students in estimated post-FOHS changes in other outcomes.

Hispanic and white students differed significantly in their estimated post-FOHS changes on MCAS scaled scores in both mathematics and ELA, mathematics proficiency rates, number of days present, number of days absent, and number of days tardy. FOHS reforms were estimated to have no effect on the MCAS mathematics scaled scores or proficiency rates of white students, but (like black students) Hispanic students saw significant declines in both outcomes. Both Hispanic and white students had lower MCAS scaled scores in ELA after FOHS, but decreases were larger among Hispanic students. Both Hispanic and white students were present for fewer days after the implementation of FOHS, but the reduction was smaller for Hispanic students. Hispanic and white students both had fewer absences after FOHS, but the decline was smaller for Hispanic students. Hispanic students were tardy significantly less often after FOHS, while white students were tardy more often. There were no significant differences between Hispanic and white students in post-FOHS changes for other outcomes.

3. Language Spoken at Home

Students who did and did not speak English in their homes experienced significantly different post-FOHS changes on five outcomes. Our models suggested a significant decrease in the MCAS scaled scores in ELA of students who did not speak English at home, but increases among students who spoke English in their homes (Table III.4). The MCAS proficiency rate in ELA of students who spoke English at home also increased, while FOHS reforms appeared to have no effect on the ELA proficiency rates of students who spoke other languages at home. Both groups were present for fewer days after the implementation of FOHS, but the reduction for students who spoke English at home was smaller (Table III.4). Both groups had fewer absences (including unexcused absences) after FOHS, but the declines were smaller for students who spoke English in their homes.

4. Receipt of Special Education Services

Special education students experienced significantly different changes on several outcomes relative to non-special education students. Both groups had lower ELA MCAS scores after the implementation of FOHS, but the decrease was much larger for special education students as compared to non-special education students (Table III.4). Among non-special education students, FOHS reforms were estimated to have no effect on mathematics scaled scores, or on proficiency rates in either mathematics or ELA. Special education students, however, had significant decreases in all three outcomes. Both groups were present for fewer days after the implementation of FOHS, but the estimated decrease was larger for special education students (Table III.4). Both groups had fewer absences after FOHS, but the decrease was smaller for special education students. Special education students had significantly larger improvements in grade promotion during the post-FOHS years as compared to non-special education students (for whom improvements were not statistically significant).

5. Eligibility for the Free or Reduce-Price Lunch Programs

Students eligible for free- or reduced-price lunch (FRPL) experienced significantly different post-FOHS changes on three outcomes as compared to ineligible students. FRPL students had significant improvements in MCAS mathematics scores during the post-FOHS years as compared to non-FRPL students (for whom statistically significant changes were not observed, Table III.4). Both groups had fewer (total) absences after FOHS, but the reduction was smaller for FRPL students. Both groups were tardy more often after FOHS, but FRPL students were tardy significantly more often than non-FRPL students.

C. SCHOOL-LEVEL ANALYSES

This section discusses the findings of analyses examining the effects of FOHS for the 12 schools in existence at the start of the reforms. The multilevel techniques used allowed estimation of school-level effects for MCAS outcomes and all other student outcomes, except number of days tardy. Figures III.12a/b through III.21a present our school-level estimates of the effects of FOHS reforms for these outcomes in the units directly generated by our ITS models. That is, they show estimated effects in conventional units for MCAS mathematics and ELA scaled scores; odds-ratios for MCAS proficiency, grade promotion, and student suspension rates; and ratios for other "count" outcomes (i.e., number of days present, absent, unexcused absences, and number of suspensions). Appendix B provides additional details about the methods used for our school-level analyses.

Our school-level analyses focused on whether outcome change patterns differed or were similar across the 12 original FOHS high schools. This is reflected by the extent of overlap among the 95 percent confidence intervals for school-level estimates of the effects of FOHS, which are shown as vertical lines in the figures. *That is, overlapping vertical lines suggest no significant differences in the school-level effects of FOHS reforms. Similarly, if the vertical lines overlap the horizontal line in a given figure, we know that the estimated school-level effects are statistically indistinguishable from the overall effect discussed earlier in this report.*

We examined this extent of overlap for school-level estimates of the effects of FOHS (1) on outcome levels in the "cutoff" academic year for our analyses and (2) on the rate of annual change in outcomes over later study years, for selected outcomes. The "a" figures in this section show the estimated *initial* effects of FOHS reforms on the outcome of interest in the first year after relevant reforms were introduced. The "b" figures show our estimates of the rate of change in FOHS effects for the outcome of interest for each subsequent post-FOHS year.[6] The horizontal line in each figure indicates the value of the corresponding estimate of FOHS effects for the overall sample (discussed in Section A of this chapter). The vertical lines depict the 95 percent confidence intervals for ITS model estimates of school-level effects.[7]

Below, we discuss the results of our school-level analyses. These analyses suggested that the observed changes in MCAS outcomes were experienced fairly evenly across the 12 high schools present at the outset of the study. While there was some variation across schools in the

---

[6] Outcomes with only an "a" figure (for example, student attendance – Figure III.16a) involved ITS models with both linear and quadratic components for estimation of the school-level effects of FOHS reforms. To ease the interpretation of these complex quadratic models, we present and discuss only differences in school-level estimates of the *initial* effects of the FOHS reforms.

[7] For readers interested in additional technical guidance on how to read the figures in this section of the report, we provide the following more detailed explanation. Consider, for example, the estimates presented in Figures III.12a and III.12b. Interpretation of these estimates is fairly straightforward, since no transformation of the MCAS scaled scores was needed for ITS modeling. The horizontal line in Figure III.12a shows the estimated effect of FOHS reforms on mathematics scaled scores in 2002-2003 for the overall sample—an increase of about 0.4 points (which is consistent with the estimate presented in Table III.1). The 95 percent confidence intervals (i.e., the vertical lines) for 9 of the 12 FOHS high schools—from Charlestown to Brighton—overlap the horizontal line representing the FOHS overall effect. This suggests that the initial FOHS effects on mathematics scaled scores estimated for these schools are not significantly different from the effect estimated for the overall sample. Nor do they differ significantly from one another, since the confidence intervals overlap. Burke High School appeared to experience a larger initial decline in mathematics MCAS scores, while Boston Community Leadership High School seemed to experience a modest increase. However, the 95-percent confidence intervals (vertical lines) for these schools still overlap those for other FOHS schools, suggesting that their FOHS effects were not significantly different from the effects estimated for other schools nor for the overall sample.

Figure III.12b then shows that, on average, across all FOHS schools, mathematics scaled scores were estimated to decrease by about 2.8 points in each subsequent year after FOHS literacy reforms were introduced. (Again, this is represented by the horizontal line in the figure.) Thus, the estimated average effects of FOHS reforms across all schools were declines of about 2.4 points in 2003-2004, 5.2 points in 2004-2005, and so on. (Note that these estimates are also consistent with those presented in Table III.1.) The overlap in vertical lines in Figure III.12b suggests that further post-FOHS declines in MCAS mathematics scaled scores (relative to projections absent the reforms) were similar in magnitude and experienced uniformly across all targeted high schools, except Boston Community Leadership High School. At this high school, further declines appeared more modest.

Figures III.15a and III.15b similarly present estimates of FOHS effects on ELA proficiency, expressed as odds ratios. These estimates can also be readily interpreted. A value of one is equivalent to no change in the odds of proficiency. Values under (over) one signal a reduction (increase) in the odds of proficiency. The magnitude of this decrease (increase) is obtained by subtracting one from the value shown. So, for example, the horizontal line in Figure III.15a, at about 0.5, indicates that we estimated about a 50 percent *reduction* (1 – 0.5 = 0.5) in the odds of proficiency in ELA for the overall sample in 2002-2003. (Note, again, that this is consistent with the estimates shown in Figure III.4.) All but one of the 95 percent confidence intervals for our school-level estimates of the initial effects of FOHS reforms overlap the horizontal line. This suggests that initial FOHS-related changes in ELA proficiency rates were similar across these 11 original study schools. In contrast, at Boston Community Leadership High School, ELA proficiency rates appeared to remain unchanged or increase.

post-FOHS changes observed, most schools experienced changes that did not differ significantly from the changes estimated for the overall sample or from the changes observed in other schools.

Mathematics MCAS Scaled Scores. Our school-level analyses suggested that the changes in mathematics scaled scores observed in the 12 original comprehensive high schools in the 2002-2003 academic year were statistically indistinguishable from one another (Figure III.12a). In the remaining post-FOHS academic years, mathematics MCAS scaled scores appeared to decrease more modestly at Boston Community Leadership High School (Figure III.12b).

ELA MCAS Scaled Scores. Eleven of the 12 original FOHS schools seemed to experience similar changes in their tenth graders' MCAS ELA scaled scores during the 2002-2003 academic year (Figure III.13a), while Boston Community Leadership High School exhibited initial increases. In the remaining post-FOHS academic years, ELA scaled scores appeared to decrease fastest at Brighton High School, and most modestly at South Boston High School. The estimated changes in ELA scaled scores for the remaining FOHS high schools were statistically indistinguishable from one another (Figure III.13b).

Mathematics MCAS Proficiency Rates. Initial changes in students' odds of testing proficient in mathematics in 2002-2003 were statistically indistinguishable across most FOHS high schools (Figure III.14a). At Charlestown High School, students' odds of testing proficient seemed more likely to decline in 2002-2003. In the remaining post-FOHS academic years, MCAS proficiency rates in mathematics appeared to decrease most markedly in Charlestown and Brighton High Schools (Figure III.14b). Their declines differed significantly from the estimated changes in South Boston, Burke, Madison Park, Dorchester, Hyde Park, and Boston Community Leadership High School, where proficiency rates seemed to remain stable.

Figure III.12a
School-level Effect of FOHS on Math Achievement
in 2002-2003

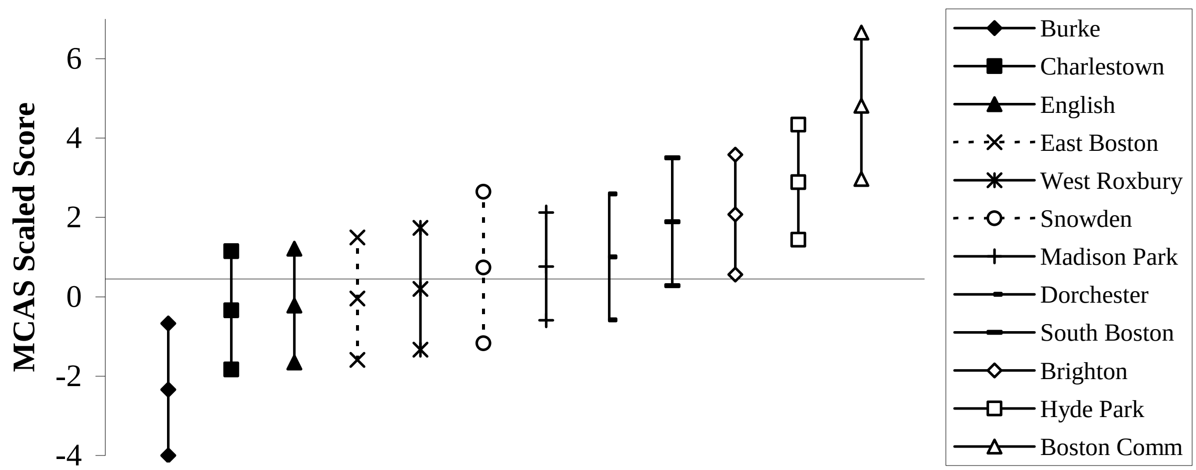


Figure III.12b
School-level Effect of FOHS on Annual Change
in Math Achievement

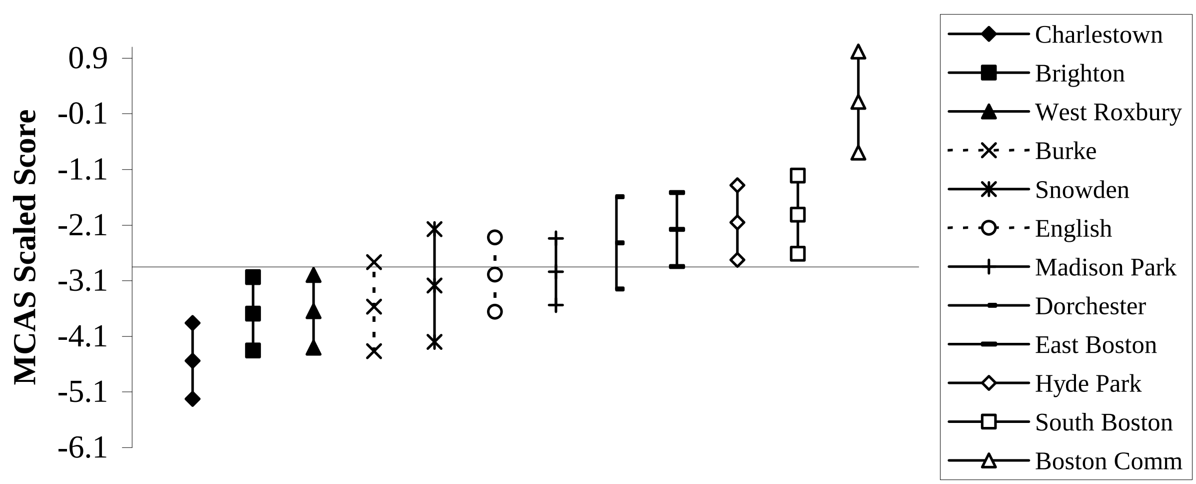


Figure III.13a
School-level Effect of FOHS
on ELA Achievement in 2002-2003

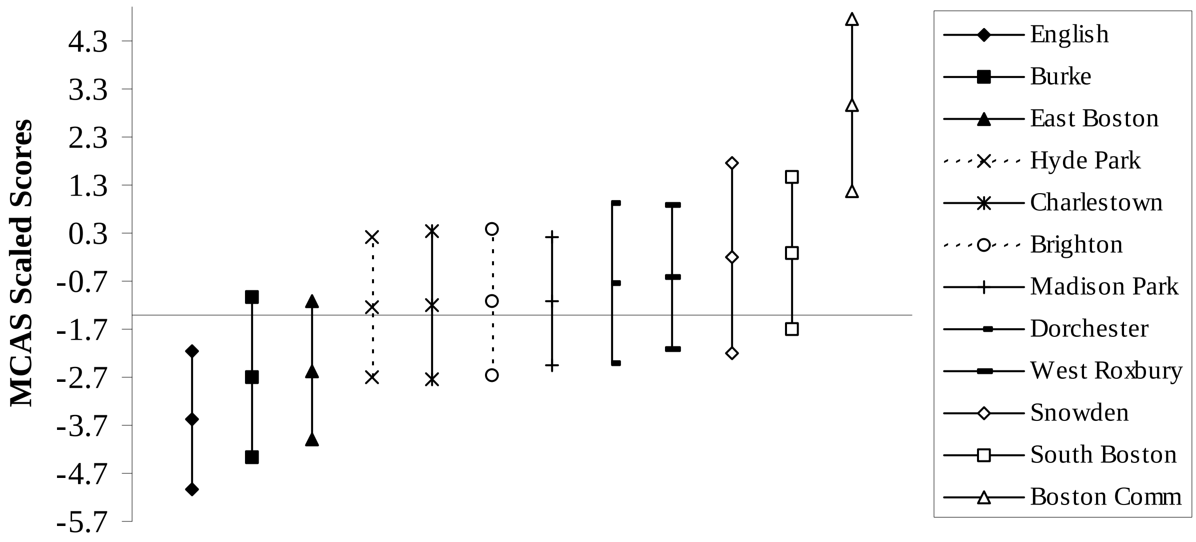


Figure III.13b
School-level Effect of FOHS
on Annual Change in ELA Achievement

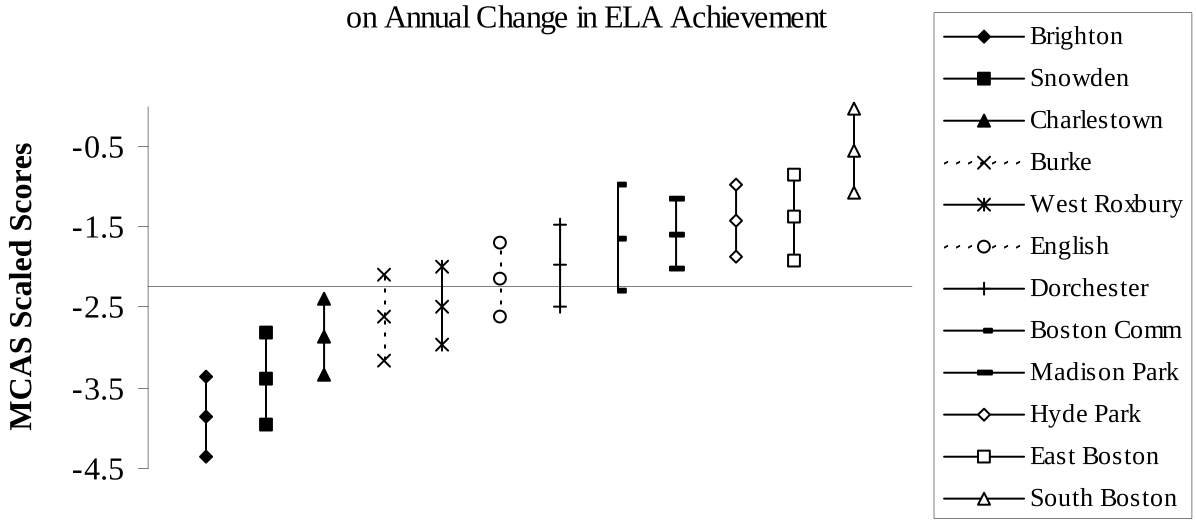


Figure III.14a
School-level Effect of FOHS on Math Proficiency
in 2002-2003

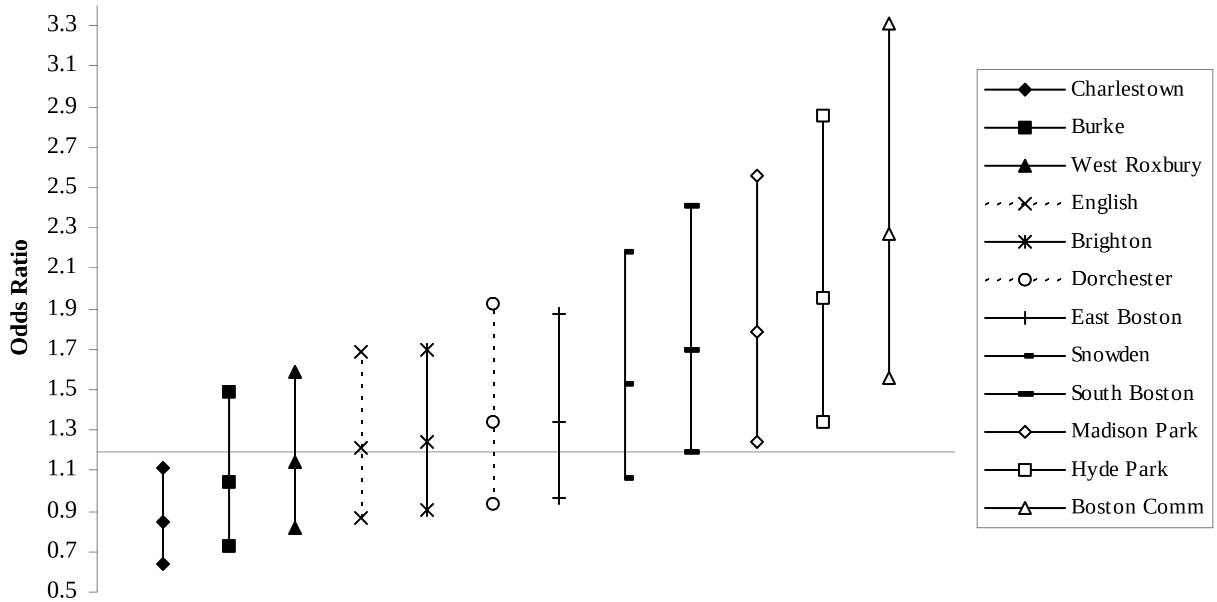
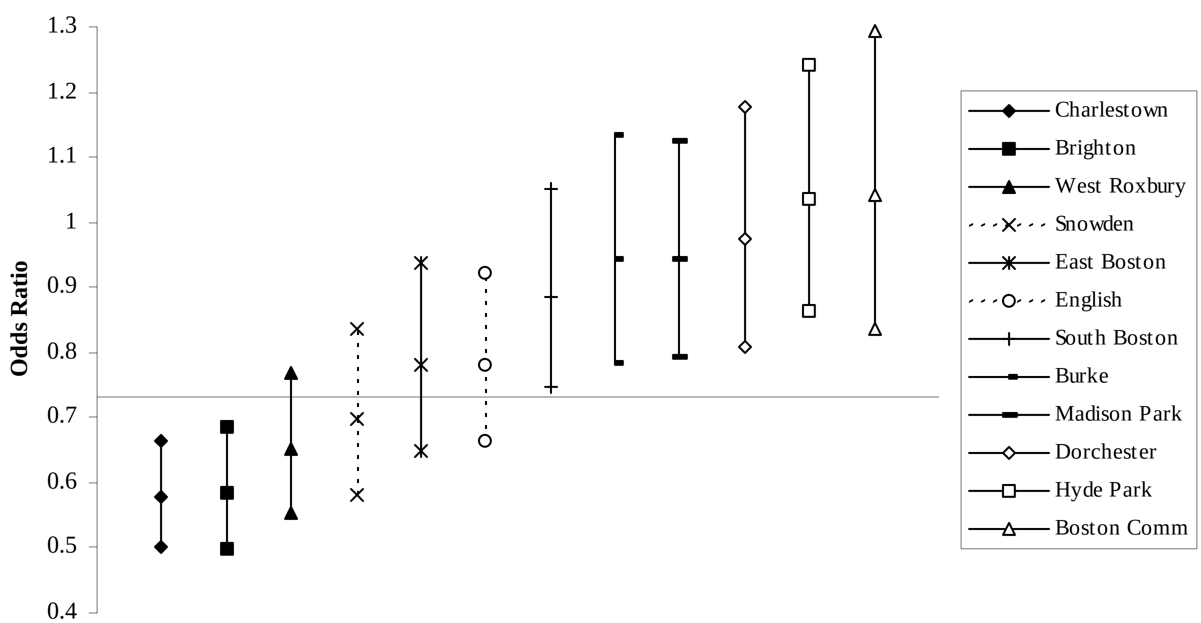


Figure III.14b
School-level Effect of FOHS on Annual Change in Math Proficiency



ELA MCAS Proficiency Rates. Eleven of the 12 original FOHS schools exhibited comparable reductions in their students' odds of testing at proficiency or better in ELA in the 2002-2003 academic year (Figure III.15a). Meanwhile, Boston Community Leadership High School exhibited possible increases. In the remaining post-FOHS years, MCAS proficiency rates in ELA decreased most markedly in Brighton High School (Figure III.15b). The estimated declines in this school differed significantly from the more modest changes estimated for South Boston, East Boston, Madison Park, and Boston Community Leadership High School.

Attendance. Our school level analyses suggested statistically significant and large declines in student attendance at Snowden High School during the 2003-2004 academic year (Figure III.16a), while attendance was estimated to increase modestly or remain stable at English High School. The estimated changes in student attendance for the remaining 10 original FOHS comprehensive high schools were statistically indistinguishable from one another.

Total and Unexcused Absences. During the 2003-2004 academic year, the modest decreases in total student absences at English and Dorchester High Schools were indistinguishable from one another (Figure III.17a). The increases in total absences at Brighton, East Boston, Hyde Park, Burke, and Boston High Schools were also indistinguishable from one another. In 2003-2004, unexcused absences decreased most markedly at Dorchester High School, and indistinguishably at English, Snowden, Madison Park, and Charlestown High Schools (Figure III.18a). In contrast, the increases in unexcused absences at East Boston, Burke, and Hyde Park High Schools were statistically indistinguishable from one another.

Percent Suspended. For the 2003-2004 academic year, the increases in students' probability of suspension at Dorchester and Burke High Schools were statistically indistinguishable from one another (Figure III.19a). The decreases in students' probability of

Figure III.15a
School-level Effect of FOHS on ELA Proficiency
in 2002-2003

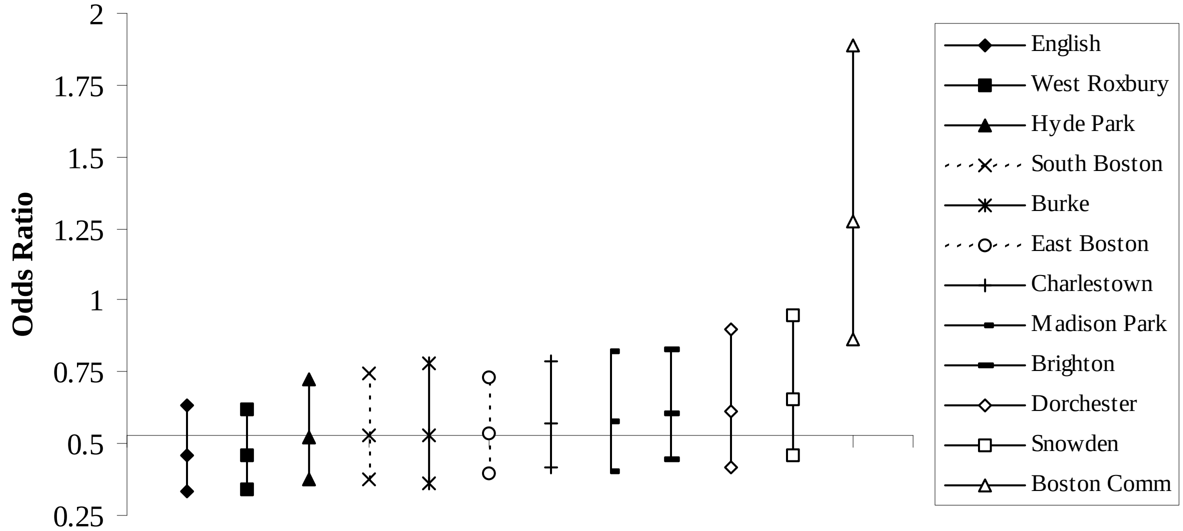


Figure III.15b
School-level Effect of FOHS
on Annual Change in ELA Proficiency

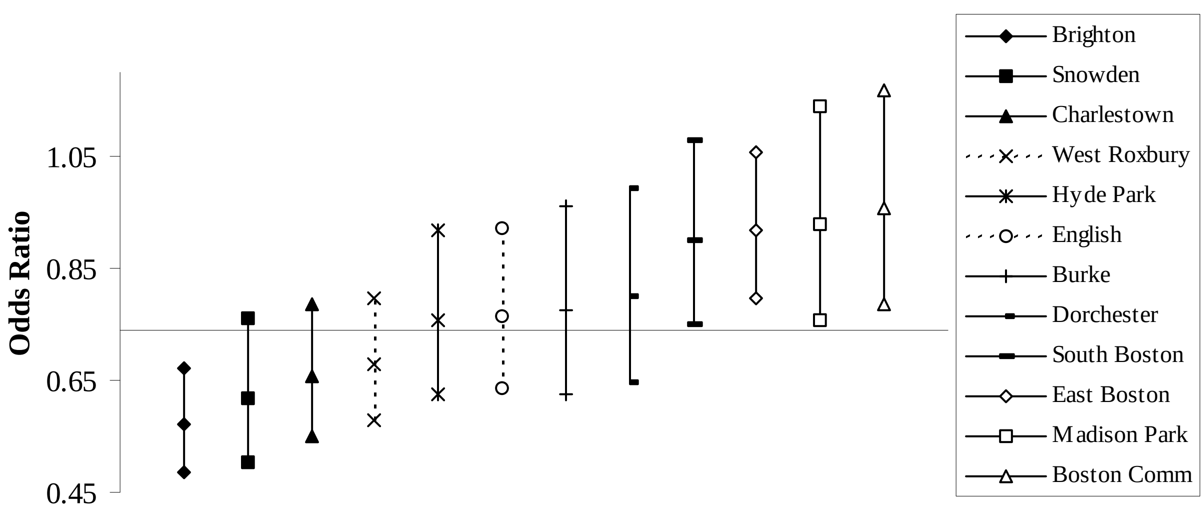


Figure III.16a
School-level Effect of FOHS on
Attendance in 2003-2004

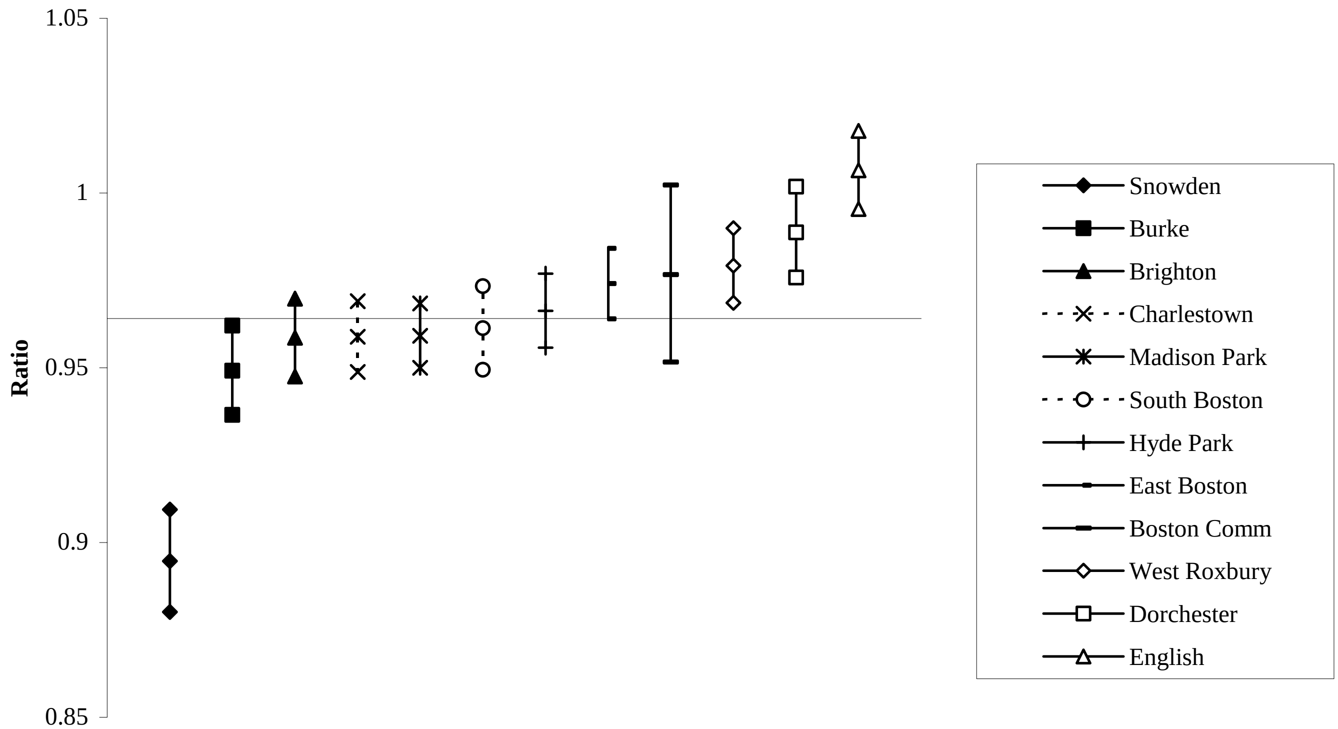


Figure III.17a
School-level Effect of FOHS on
Absences in 2003-2004

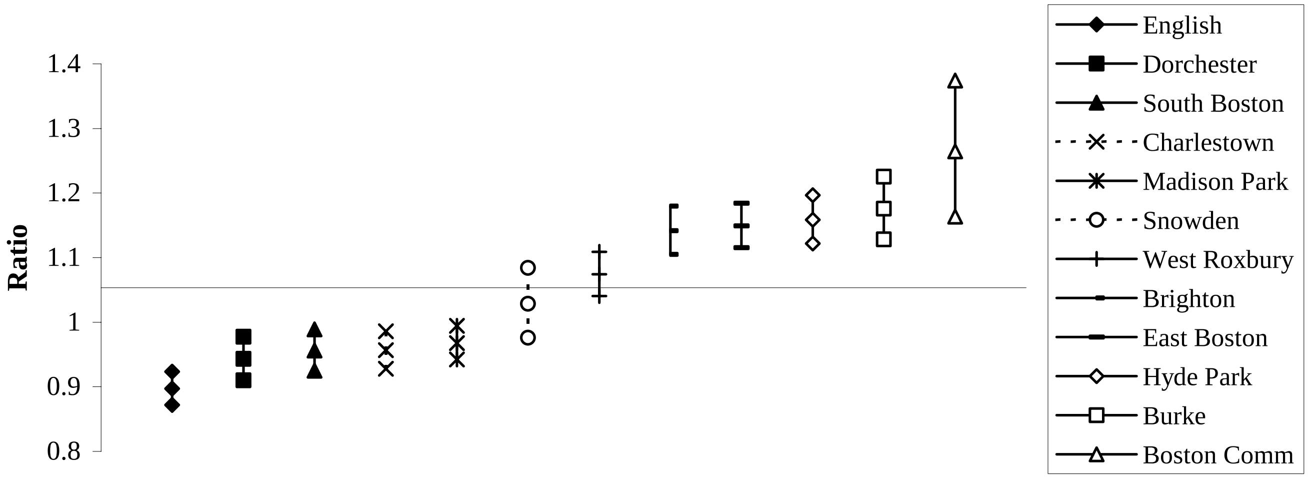


Figure III.18a
School-level Effect of FOHS on
Unexcused Absences in 2003-2004

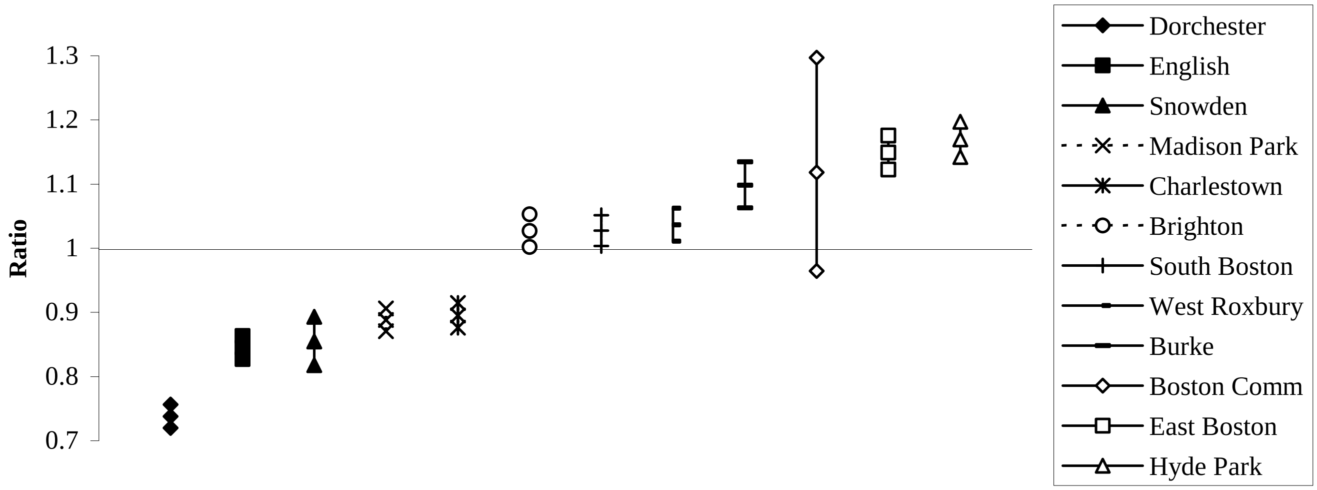
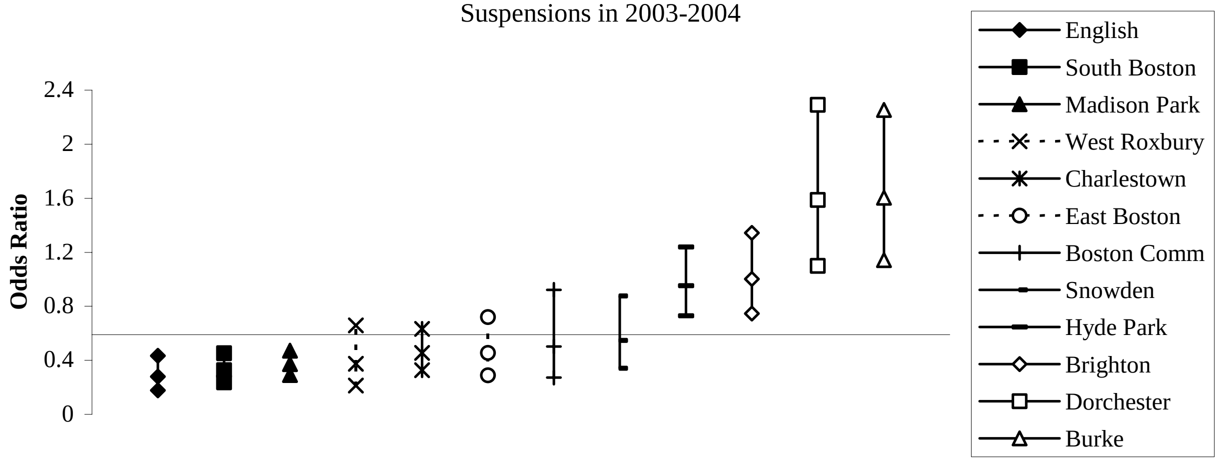


Figure III.19a
School-level Effect of FOHS on
Suspensions in 2003-2004



suspension at English, South Boston, Madison Park, West Roxbury, Charlestown, and East Boston High Schools were also statistically indistinguishable from one another.

Number of Suspensions. Declines in the number of student suspensions during the 2003-2004 academic year were indistinguishable for seven of the 12 original FOHS high schools: English, South Boston, Charlestown, Madison Park, East Boston, West Roxbury, and Snowden (Figure III.20a). Meanwhile, increases in student suspensions were highest at Dorchester High School and indistinguishable for Hyde Park and Brighton High Schools.

Grade Promotion. In 2003-2004, the increases in students' probability of promotion to the next grade at Hyde Park, South Boston, West Roxbury, East Boston, Brighton, and Charlestown High Schools were statistically indistinguishable from one another (Figure III.21a), while the probability of grade promotion decreased at Madison Park High School. At the remaining original FOHS schools, promotion rates were estimated to remain stable.

D. CONCLUSION

Overall, significant differences were observed in all of the outcomes examined. The estimated declines in MCAS scores, especially in English Language Arts, relative to what would have been projected in the absence of FOHS were consistent across most subgroups, which provides suggestive evidence that the FOHS initiative may not have had the intended effect on student achievement. The slight reduction in number of days present for the overall sample was also inconsistent with the goals of the FOHS reforms. However, a number of other estimated post-FOHS changes were in the direction intended by the FOHS initiative. In particular, we observed a reduction in number of days absent (including unexcused absences), which was observed consistently across the student subgroups. We also observed a reduction in number of days tardy, which was concentrated among black and Hispanic students. Student suspensions were also reduced consistently across all subgroups in the post-FOHS period relative to what

Figure III.20a
School-level Effect of FOHS on
Number of Suspensions in 2003-2004

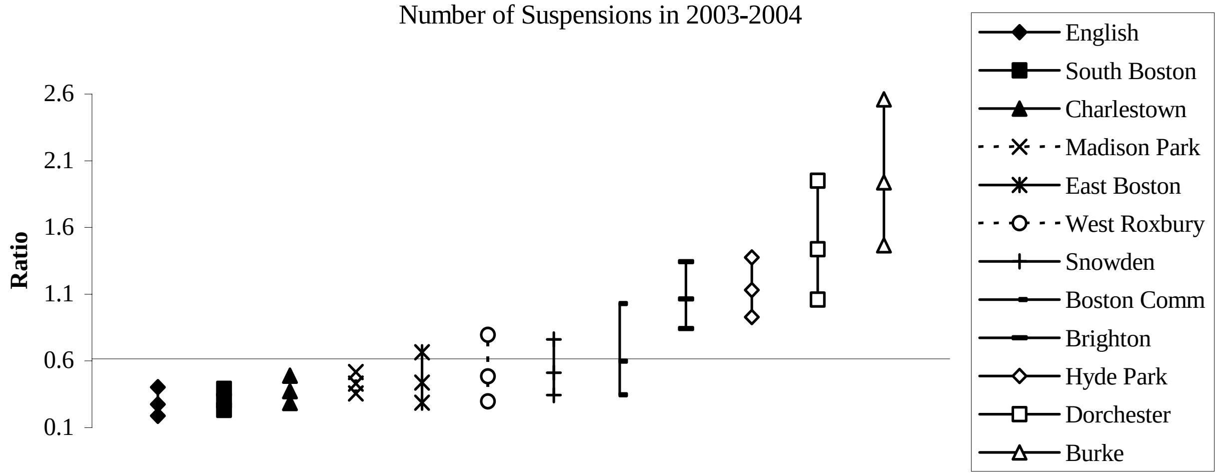
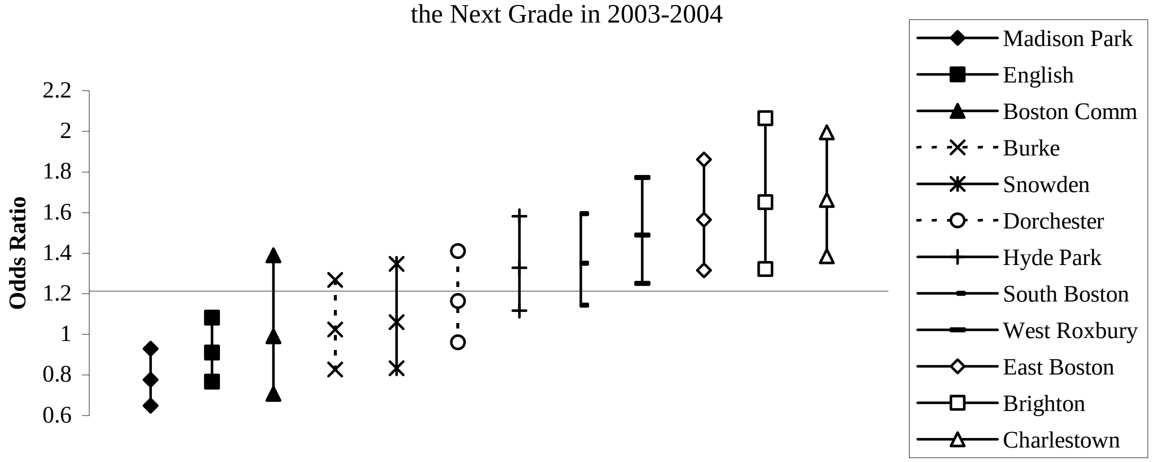


Figure III.21a
School-level Effect of FOHS on Promotion to
the Next Grade in 2003-2004



would be projected in the absence of FOHS. Finally, we observed an overall increase in promotions to the next grade, and these results were concentrated among special education students.

Our study findings are generally consistent with those from prior studies of small school reforms at the secondary level. The limited number of available studies similarly find that small school reforms lead to improvements in varied academic outcomes but do not always translate into improvements in student achievement.

REFERENCES

American Diploma Project. “Ready or Not: Creating a High School Diploma That Counts.” Washington, DC: Achieve, Inc., (2004). [http://achieve.org/files/ADPreport_7.pdf]. Accessed February 15, 2008.

Bill & Melinda Gates Foundation. “All Students College-Ready: Findings from the Foundation’s Education Work 2000-2006.” 2006. [http://www.gatesfoundation.org/nr/downloads/ed/researchevaluation/EducationFindings2000-2006.pdf]. Accessed February 15, 2008.

Bloom, Howard. “Using ‘Short’ Interrupted Time-Series Analysis to Measure the Impacts of Whole School Reforms: With Applications to a Study of Accelerated Schools.” *Evaluation Review*, vol. 27, no. 3, 2003, pp. 3–49.

Boston Public Schools. “Points of Pride.” 2008. [http://www.cityofboston.gov/bps/pdfs/BPSPointsofPride.pdf]. Accessed February 14, 2008.

Boston Public Schools. “Boston Public Schools at a Glance.” 2007. [http://www.cityofboston.gov/bps/pdfs/BPSglance.pdf]. Accessed February 13, 2008.

Boston Public Schools. “A Decade of Transforming the Boston Public Schools.” 2005. [http://www.cityofboston.gov/bps/pdfs/transform.pdf]. Accessed February 14, 2008).

Boston Public Schools. “Focus on Children II, Boston’s Education Reform Plan: 2001-2006.” (Adopted by the Boston School Committee, April 25, 2001). [http://boston.k12.ma.us/teach/FocusOnChildrenII.doc]. Accessed February 12, 2008.

Bryk, Anthony, Valerie E. Lee, and Peter B. Holland. *Catholic Schools and the Common Good.* Cambridge, MA: Harvard University Press, 1993.

Coalition of Essential Schools. “About the Coalition of Essential Schools.” 2008. [www.essentialschools.org/pub/ces_docs/about/about.html]. Accessed February 15, 2008.

Darling-Hammond, Linda, Jacqueline Ancess, and Susanna Wichterle Ort. “Reinventing High School: Outcomes of the Coalition Campus Schools Project.” *American Educational Research Journal*, vol. 39, no. 3, 2002, pp. 639–673.

Education Week. “Major Gates Foundation Grants to Support Small High Schools.” *Education Week*, vol. 23, no. 40, 2004, pp. 28–29.

Foley, Eileen M., Allan Klinge, and Elizabeth R. Reisner. “Evaluation of New Century High Schools: Profile of an Initiative to Create and Sustain Small, Successful High Schools.” Washington, DC: Policy Studies Associates, Inc., 2007.

[http://www.policystudies.com/studies/school/NCHS%20Evaluation%20Final%20Report%2010-07.pdf]. Accessed February 15, 2008.

Gates, Bill. Prepared remarks to the National Education Summit on High Schools, Washington, DC, February 26, 2005. [http://www.achieve.org/node/767].

Hendrie, Caroline. "High Schools Nationwide Paring Down." *Education Week*, vol. 23, no. 40, 2004, pp. 1, 28–30.

Hess, G. Alfred, and Solomon Cytrynbaum. "The Effort to Redesign Chicago High Schools: Effects on Schools and Achievement." In *Reforming Chicago's High Schools: Research Perspectives on School and System Level Change*, edited by Valerie Lee. Chicago: Consortium on Chicago School Research at the University of Chicago, 2002, pp. 19–49.

Holland, Nicole. "Small Schools: Transforming Teacher and Student Experiences in Urban High Schools." In *Reforming Chicago's High Schools: Research Perspectives on School and System Level Change*, edited by Valerie Lee. Chicago: Consortium on Chicago School Research at the University of Chicago, 2002, pp. 89–123.

Howley, Craig B. "Synthesis of the Effects of School and District Size: What Research Says About Achievement in Small Schools and School Districts." *Journal of Rural and Small Schools*, vol. 4, no. 1, 1989, pp. 2–12.

Hox, Joop J. "Multilevel analyses of grouped and longitudinal data." In *Modeling Longitudinal and Multilevel Data: Practical Issues, Applied Approaches, and Specific Examples*, edited by Todd D. Little, Kai U. Schnabel, and Jurgen Baumer. Mahway, NJ: Lawrence Erlbaum Associates, 2000.

James-Burdumy, Susanne, and Daniel Finkelstein. "The Focus on High School Initiative: Changes in Student Views Between 2003 and 2005." Princeton, NJ: Mathematica Policy Research, Inc., 2006.

Lee, Valerie E. "Setting Chicago High School Reform Within the National Context." In *Reforming Chicago's High Schools: Research Perspectives on School and System Level Change*, edited by Valerie Lee. Chicago: Consortium on Chicago School Research at the University of Chicago, 2002, pp. 7–17.

Lee, Valerie E., and Julia B. Smith. "Effects of High School Restructuring and Size Gains in Achievement for Early Secondary School Students." Sociology of Education, vol. 68, no. 4, 1995, pp. 241–270.

McCullagh, Peter, and John A. Nelder. *Generalized Linear Models, Second Edition (Monographs on Statistics and Applied Probability)*. Boca Raton, FL: Chapman & Hall/CRC Press, 1989.

Meier, Deborah. *The Power of Ideas: Lessons for America from a Small School in Harlem.* Boston: Beacon Press, 1995.

Mitchell, Karen, Jamie Shkolnik, Mengli Song, Kazuaki Uekawa, Robert Murphy, Mike Garet, and Barbara Means. "Rigor, Relevance, and Results: The Quality of Teacher Assignments and Student Work in New and Conventional High Schools." Washington, DC, and Menlo Park, CA: American Institutes for Research and SRI International, 2005.

National Research Council. "Engaging Schools: Fostering High School Students' Motivation to Learn." Washington, DC: National Academies Press, 2003.

Neufeld, Barbara, and Anne Levy, with Sara Schwartz Chrismer. "High School Renewal in the Boston Public Schools: Focus on Organization and Leadership." Cambridge, MA: Education Matters, Inc., 2005.

Neufeld, Barbara, and Anne Levy, with Sara Schwartz Chrismer. "Baseline Report: High School Renewal in Boston." Cambridge, MA: Education Matters, Inc., 2004.

Payzant, Thomas W. "Boston Leads the Way in High School Redesign." 2003. [http://boston.k12.ma.us/supt/msgs/msgs_3_05.asp]. Accessed February 13, 2008.

Pittman, Robert B., and Perri Haughwout. "Influence of High School Size on Dropout Rate." *Educational Evaluation and Policy Analysis,* vol. 9, no. 4, 1987, pp. 337–343.

Powell, Arthur G., David K. Cohen, and Eleanor Farrar. *The Shopping Mall High School: Winners and Losers in the Educational Marketplace.* Boston: Houghton-Mifflin, 1985.

Raudenbush, Stephen W., and Anthony S. Bryk. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Second Edition. Thousand Oaks, CA: Sage, 2002.

Shadish, William R., Thomas D. Cook, and Donald T. Campbell. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Boston: Houghton-Mifflin, 2002.

Sibley, Stephanie. Personal email communication, May 7, 2008.

Sibley, Stephanie, Barbara Adams, and Irving Scott. "Feedback on the FOHS Evaluation." BPE memorandum to Susanne James-Burdumy, May 8, 2008.

Singer, Judith D. and John B. Willett. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence.* New York: Oxford University Press, 2003.

Sizer, Theodore R. *Horace's Compromise: The Dilemma of the American High School.* Boston: Houghton Mifflin, 1984.

U.S. Department of Education. "Smaller Learning Communities Program." Program description. [www.ed.gov/programs/slcp/index.html]. Accessed February 13, 2008.

Wasley, Patricia, Michelle Fine, Matt Gladden, Nicole E. Holland, Sherry P. King, Esther Mosak, and Linda C. Powell. *Small Schools, Great Strides: A Study of New Small Schools in Chicago.* New York: Bank Street College of Education, 2000.

APPENDIX A

FOHS HIGHLIGHTS AND BPS HIGH SCHOOLS IN 1999 AND 2005

Focus on High Schools

Highlights

GRANT: Carnegie Corporation of New York's *Schools for a New Society* Initiative

AWARDED TO: Boston Public Schools and Boston Plan for Excellence

GRANT AMOUNT: $8 million over five years

GRANT MANAGER: Boston Plan for Excellence

GRANT PERIOD: July 1, 2001- June 30, 2006

SITES: The district's twelve "non-exam" high schools — Boston High School, Brighton High School, Burke High School, Charlestown High School, Dorchester High School, East Boston High School, English High School, Hyde Park High School, Madison Park Technical-Vocational High School, Snowden International High School, South Boston High School, West Roxbury High School

PURPOSE OF GRANT:

1. intensify professional development for high school teachers in literacy

2. create small learning communities to individualize instruction and reduce student alienation

GRANT OF ABOUT $1.6M/YEAR WILL FUND:

Full-time staff developers with expertise in literacy in twelve high schools

Technical assistance to help large schools reorganize as small learning communities

Pilot efforts at community outreach and parent organizing

Support for new, more academically focused school-business partnerships

Boston's High Schools • 1999 & 2005

ONE GOAL OF HIGH SCHOOL RENEWAL in Boston:

To reduce alienation by creating **schools** *in which students can be known well by their teachers and* **conditions** *in which teachers can teach effectively — and to each student's needs.*

The first step has been to redesign the city's large high schools.

(1) Four large comprehensive high schools have been closed — Dorchester, Hyde Park, South Boston, West Roxbury — and re-opened as several autonomous small schools within the same building, or "education complex."

(2) Other high schools have been reorganized as several small learning communities (SLC) for grades 9-12.

Chart: Boston Plan for Excellence
Source: BPS, January 2005

SEPTEMBER 1999	Size	Type of school
Another Course to College	120	Alternative program, grades 11-12
Boston Arts Academy	160	Pilot, grades 9-11 (gr. 12 in 9-00)
Boston Evening Academy	150	Horace Mann Charter, gr. 9-12
Boston High School	800	Comprehensive, grades 9-12
Boston Latin Academy	1,560	Exam school, grades 7-12
Boston Latin School	2,350	Exam school, grades 7-12
Brighton High School	1,050	Comprehensive, grades 9-12
Burke High School	700	Comprehensive, grades 9-12
Charlestown High School	1,150	Comprehensive, grades 9-12
Community Academy	50	Alternative program, grades 8-12
Dorchester High School	950	Comprehensive, grades 9-12
East Boston High School	1,250	Comprehensive, grades 9-12
English High School	1,300	Comprehensive, grades 9-12
Fenway Middle College High	260	Pilot, grades 9-12
Greater Egleston Community	100	Pilot, 16-21 years
Health Careers Academy	175	Pilot, H. Mann Charter, gr. 9-12
Hyde Park High School	1,100	Comprehensive, grades 9-12
Madison Park Tech-Voc High	1,620	Comprehensive, grades 9-12
Multicultural High School	170	Pilot, grades 9-12
O'Bryant School/Math & Science	1,450	Exam school, grades 7-12
Snowden International School	450	Comprehensive, grades 9-12
South Boston High School	1,020	Comprehensive, grades 9-12
West Roxbury High School	1,310	Comprehensive, grades 9-12

SEPTEMBER 2005	Size		Type of school	Notes
Another Course to College	240		Small School/Pilot, grades 9-12	formerly alternative program
Boston Arts Academy	410	❖	Small School/Pilot, grades 9-12	
Boston Day & Evening Academy	230	❖	Small School/Pilot, 16-23 years	Day Acad'y added 9-04, H. Mann Charter
Boston Community Leadership Academy	400	❖	Small School/Pilot, grades 9-12	formerly Boston High School
Boston Latin Academy	1,590	❖	Exam school, grades 7-12	
Boston Latin School	2,430	❖	Exam school, grades 7-12	
Brighton High School	1,300		Grades 9-12 SLCs	
Burke High School	770		Grades 9-12 SLCs	
Charlestown High School	1,260		Grades 9-12 SLCs	
Community Academy	150		Alternative school, grades 9-12	formerly alternative program
Dorchester HS closed 6-03 and reopened as three small autonomous high schools 9-03				
East Boston High School	1,460		Grades 9-12 SLCs	
English High School	1,270		Grades 9-12 SLCs	
Fenway High School	280	❖	Small School/Pilot, grades 9-12	
Greater Egleston Community	110	❖	Small School/Pilot, 16-21 years	
Health Careers Academy	200	❖	Small School/Pilot, grades 9-12	Horace Mann Charter
Hyde Park HS closed 6-05 and reopened as three small autonomous high schools 9-05				
Madison Park Tech-Voc High	1,660		Comprehensive, grades 9-12	
New Mission High School	250	❖	Small School/Pilot, grades 9-12	formerly Multicultural HS
O'Bryant School/Math & Science	1,220	❖	Exam school, grades 7-12	
Snowden International School	440		Comprehensive, grades 9-12	
South Boston HS closed 6-03 and reopened as three small autonomous high schools 9-03				
West Roxbury HS closed 6-05 and reopened as four small autonomous high schools 9-05				

NEW SCHOOLS OPENED SINCE SEPTEMBER 1999 ❖ **School has special admissions process**

Academy of Public Service	300		Small School, grades 9-12	at Dorchester Education Complex
Boston Adult Tech Academy	230	❖	Alternative school, 20-22 years	
Boston International High	180	❖	Small School, grades 9-12	
Brook Farm Business & Service Career Academy	350		Small School, grades 9-12	at West Roxbury Education Complex
Community Academy of Science & Health	350		Small School, grades 9-12	at Hyde Park Education Complex
Economics & Business Academy	300		Small School, grades 9-12	at Dorchester Education Complex
Engineering School	350		Small School, grades 9-12	at Hyde Park Education Complex
Excel High School	370		Small School, grades 9-12	at South Boston Education Complex
Media Communications Tech	350		Small School, grades 9-12	at West Roxbury Education Complex
Monument High School	370		Small School, grades 9-12	at South Boston Education Complex
Odyssey High School	340		Small School, grades 9-12	at South Boston Education Complex
Parkway Acad'y/Tech & Health	350		Small School, grades 9-12	at West Roxbury Education Complex
Quincy Upper School	430		Small School/Pilot, grades 6-12	
Social Justice Academy	350		Small School, grades 9-12	at Hyde Park Education Complex
TechBoston Academy	300		Small School/Pilot, grades 9-12	at Dorchester Education Complex
Urban Science Academy	350		Small School, grades 9-12	at West Roxbury Education Complex

APPENDIX B

DETAILS OF METHODS FOR ESTIMATING EFFECTS

The Interrupted Time Series (ITS) design used on this study is a quasi-experimental approach for evaluating the effectiveness of an intervention. It is based on the availability of time series data before the intervention was implemented (the pre-intervention period) and after the intervention was implemented (the post-intervention period). In this technique, data from the pre-intervention period are used to predict what would have happened *in the absence of the intervention* during the post-intervention period. The effect of the intervention is then estimated as the difference between this predicted pattern of outcomes and the actual model trend in outcomes in the post-intervention period.

Because an experimental design was not used for the study, care must be exercised when interpreting its findings. In particular, factors that may have occurred at the same time as the FOHS intervention, such as policy changes or changes in unobserved student characteristics, might also contribute to the estimated effects of the FOHS reforms.

In this appendix, we first describe the types of outcomes being examined as part of the study. We begin with these details because the types of outcomes affect the specification of the ITS models, which are described in the second section. The third section describes the process that we used to develop the ITS models. Finally, we provide details on the process used to generate the estimates of the effects.

A. TYPES OF OUTCOMES EXAMINED

The study examined 11 outcome variables, shown in Table B.1. These outcomes can be divided into two main groups: cross-sectional and longitudinal. The key difference between these types of outcomes is in the number of times the outcomes are observed for each student. The outcomes can also be categorized by the way in which they are measured (continuous, binary, or count), which influences the formation of the ITS model. Each outcome is discussed in more detail below.

TABLE B.1

OUTCOMES EXAMINED IN FOHS EVALUATION

Outcomes Examined	Outcome Type	Outcome Measure Type
MCAS mathematics score	Cross-sectional	Continuous
MCAS ELA score	Cross-sectional	Continuous
Proficient on MCAS mathematics	Cross-sectional	Binary
Proficient on MCAS ELA	Cross-sectional	Binary
Promoted to next grade level	Longitudinal	Binary
Suspended at least once this year	Longitudinal	Binary
Number of days present	Longitudinal	Count
Number of days absent	Longitudinal	Count
Number of days tardy	Longitudinal	Count
Number of unexcused absences	Longitudinal	Count
Number of times suspended	Longitudinal	Count

Note: Outcome type refers to the frequency of measurement *for each student*. The cross-sectional outcomes are measured once when each student is in 10th grade, and the longitudinal outcomes are measured each year the student is enrolled in a BPS high school.

1. Cross-Sectional and Longitudinal

A key distinction for the outcomes is the frequency at which they are measured for a given student; some are measured one time for each student (that is, cross-sectional), while others are measured every year in which the student is enrolled in a BPS high school (that is, longitudinal). Below we distinguish the cross-sectional and longitudinal outcomes.

Cross-Sectional Outcomes. Massachusetts Comprehensive Assessment System (MCAS) mathematics and English Language Arts (ELA) achievement scores, which are observed for 10th-grade students for school years 1997-1998 through 2006-2007, are the study's key cross-sectional outcomes. These data are referred to as cross-sectional because, although they are measured on the same set of schools, a different set of students is measured each school year.[1] There are 26,810 and 26,275 10th grade students in our study schools with MCAS scores available for mathematics and ELA, respectively.

[1] Approximately 11 percent of students have taken the MCAS test more than once. To ensure that our results reflect changes in initial MCAS scores (and not any change in test scores due to students simply taking the tests more than once), the MCAS analyses include only the first test administration for each student.

Longitudinal Outcomes. The study's key longitudinal outcomes are number of days absent, present, and tardy; number of suspensions; being promoted to the next grade level; or being suspended at least once during an academic year. These outcomes are referred to as longitudinal because they are measured each year for students from 1995-1996 to 2006-2007 as long as the student is enrolled in the BPS.[2] Administrative records data on these outcomes are available for 69,317 students. A student attending a BPS high school for four years would have four observations, one for each year they are in a BPS high school. Students who repeated a grade may have more than four observations, while students who transferred in or out of BPS or started high school before 1995-1996 or after 2003-2004 may have fewer than four observations.

2. Continuous, Binary, and Count Outcomes

The outcomes described above can be categorized by the scale on which they are measured: (1) continuous, (2) binary, and (3) count. The MCAS scores are measured on a continuous scale, ranging from 200 to 280 (see Massachusetts Department of Education MCAS Website http://www.doe.mass.edu/mcas). Four of the variables examined are binary, consisting of a yes or no outcome: (1) scoring at or above proficiency on the MCAS mathematics test, (2) scoring at or above proficiency on the MCAS ELA test, (3) being promoted to the next grade level, and (4) being suspended at least once during an academic year. This study also examined five count outcomes: (1) number of days present, (2) number of days absent, (3) number of days tardy, (4) number of unexcused absences, and (5) number of times suspended.[3]

---

[2] Data on promotions to the next grade were not available for the 2006–2007 school year.

[3] The data files included roughly 22,000 person-year records (out of 185,067 person-year records) with both number of days absent and number of days present equal to zero. These students were excluded from our longitudinal analyses.

B. MULTILEVEL MODELS

The data in this study are hierarchical in nature, with students nested or clustered within schools. Because students attending the same school are exposed to the same environment, student outcomes within the same school may be more similar than those of students from different schools. This potential correlation must be taken into account in the estimation of the effect of FOHS on student outcomes in order to produce accurate standard errors of such effects.

Multilevel models (Raudenbush and Bryk 2002) can account for correlated observations and for the nested structure of the data. The simplest version of a multilevel model is a two-level model, where a set of observations is clustered within a set of larger units. For example, in this study, students (level 1) are clustered within schools (level 2). In this case, level 1 models student outcomes *within* a school, while level 2 models differences *across* schools. A more complex version of the model involves three levels, where a set of observations are clustered within a set of units, which themselves are clustered within a larger set of units. For example, for the longitudinal outcomes on this study, multiple observations (level 1) are clustered by student (level 2), and students are clustered within schools (level 3).

In this section, we outline the two- and three-level models used on the study. We start first with the two-level models and then turn to the more complex three-level models.

1. Two-Level Models

This section is divided into two subsections. The first section presents the two-level model that was used for this study's continuous, cross-sectional outcomes, and the second section presents the two-level model that was used for this study's binary, cross-sectional outcomes.

a. Two-Level Model for Continuous Outcomes

Level-1 (Student-level equation)

$$\begin{aligned} Y_{ij} = {} & \beta_{0j} + \beta_{1j}(CPERIOD)_{ij} + \beta_{2j}(TRT)_{ij} + \beta_{3j}(CPERIOD)_{ij} * (TRT)_{ij} \\ & + \beta_{4j} FEMALE_{ij} + \beta_{5j} BLACK_{ij} + \beta_{6j} HISPANIC_{ij} \\ & + \beta_{7j} ESH_{ij} + \beta_{8j} SED_{ij} + \beta_{9j} FRPL_{ij} \\ & + \varepsilon_{ij} \end{aligned} \tag{1.1}$$

where:

Y is the outcome variable for student *i* in school *j*,

CPERIOD is the number of years before or after the onset of the FOHS initiative in 2003-2004,

TRT is a dummy variable equal to one for years after the beginning of the FOHS initiative,

FEMALE is a dummy variable equal to one for female students,

BLACK is a dummy variable equal to one for black students,

HISPANIC is a dummy variable equal to one for Hispanic students,

ESH is a dummy variable equal to one if students speak English in their homes,

SED is a dummy variable equal to one if students are classified as special education students,

FRPL is a dummy variable equal to one if students are eligible for the free- or reduced-price lunch program,

GRD10, *GRD11*, *GRD12* are dummy variables equal to one if students were enrolled in grades 10, 11, or 12, respectively, when they entered the sample.

Note that the omitted categories define the reference group in our analyses (that is, ninth graders who are male and white, speak a language other than English at home, are not in Special Education, and are not eligible for the free or reduced lunch programs).

Level-2 (School-level equations)

$$\begin{aligned}
&\beta_{0j} = \gamma_{00} + \gamma_{01}(\%\text{FRPL})_j + \gamma_{02}(\%\text{MIN})_j + u_{0j} \\
&\beta_{kj} = \gamma_{k0} + u_{kj} \text{ for } k = 0, 1, 2 \text{ and } 3 \qquad (1.2) \\
&\beta_{kj} = \gamma_{k0} \text{ for } k > 3.
\end{aligned}$$

where *%FRPL* is the percentage of students in each school who are eligible for the free- or reduced-price lunch program and *%MIN* is the percentage of students in each school who are minority students.

Combined Model

By substituting the terms in Eq. (1.2) into Eq. (1.1), we obtain the combined model as follows[4]:

$$\begin{aligned}
Y_{ij} = {} & \gamma_{00} + \gamma_{01}(\%\text{FRPL})_j + \gamma_{02}(\%\text{MIN})_j \\
& + \gamma_{10}(CPERIOD)_{ij} + \gamma_{20}(TRT)_{ij} + \gamma_{30}(CPERIOD)_{ij} * (TRT)_{ij} \\
& + \gamma_{40} FEMALE_{ij} + \gamma_{50} BLACK_{ij} + \gamma_{60} HISPANIC \qquad (1.3) \\
& + \gamma_{70} ESH_{ij} + \gamma_{80} SED_{ij} + \gamma_{90} FRPL_{ij} \\
& + [e_{ij} + U_j]
\end{aligned}$$

The level-1 equation models how outcomes vary with student characteristics, over time, and with exposure to the FOHS interventions. The model controls for student-level characteristics as measured at baseline (which is defined as the time at which each student entered the study sample) including gender, race (through two dummy indicators for black and Hispanic), whether

[4] The last summation in the combined model is the random error in the model and contains many terms because the error in the level-2 model is multiplied by each of the parameters in the level-1 model. It is denoted as $[E_{ij} + U_j]$ for ease of display.

a student speaks English at home, each student's special education status, and grade (through three dummies for grade 10, 11, and 12) in the student-level equation.

As shown in equation (1.2), the parameters (or coefficients) in the level-1 model are shown as outcomes in the level-2 models. The level-2 equations model how these coefficients vary across the population of schools. At level 2, we control for school level characteristics by including a school's grand mean centered percentage of minority students (*%MIN*) and grand mean centered percentage of students eligible for free or reduced-price meals (*%FRPL*) in the intercept equation for β_{0j}.[5] We allow for the intervention effect in school year 2003-2004 and the effect of the intervention on growth to vary across schools by treating these parameters as random effects in the level-2 model. By doing so, we are able to produce school-level estimates for these continuous outcomes.

b. Two-Level Model for Binary Outcomes

Binary outcomes indicate the occurrence of an event. For example, promotion to the next grade involves a variable coded as 1 if a student was promoted and 0 if the student was not promoted. Such outcomes are not continuous, and should not be treated as such because doing so can lead to (1) predicted values that are negative, (2) random effects that violate the normality assumption at level-1, and (3) a violation of the assumption of homogeneous variance at level-1 since the variance depends on the expected value of a binary outcome (see Raudenbush and Bryk 2002). To appropriately handle binary outcomes, we turn to the generalized linear modeling

[5] Variables are grand-mean centered by rescaling the variables so that they are measured as differences from the overall mean. Therefore, a school with *%FRPL* exactly equal to the overall mean will have a value of zero. Schools with a greater *%FRPL* than the mean will have positive values and those with less will have negative values.

framework that can be found in McCullough and Nelder (1989) and as formulated in Raudenbush and Bryk (2002).

Binary outcomes follow a Bernoulli distribution, where the probability of an event occurring (for example, the probability of being promoted to the next grade) is equal to p_{ij} and the expected value and variance of the Bernoulli outcome are as follows:

$$E\left\{Y_{ij} \mid p_{ij}\right\} = p_{ij} \text{ and } \mathrm{V}\left\{Y_{ij} \mid p_{ij}\right\} = p_{ij}(1 - p_{ij})$$

Unlike with continuous outcomes, where the predicted value (the expected value of the outcome) is linearly related to the covariates of the model as in Eq. (1.1), for binary outcomes, a transformation of the predicted value of the outcome is related to the covariates in the level-1 model in (1.1).[6] The transformation we use for binary outcomes is the logit function, which is the log of the odds of the outcome occurring (for example, the odds of being promoted):

$$\log(\frac{p_{ij}}{1 - p_{ij}}) \qquad (1.4)$$

The predicted value of the outcome Y_{ij} is linearly related to the predictors on the right hand side of Eq. (1.1) through this function (see Raudenbush and Bryk 2002). This results in the following level-1 equation for binary outcomes:

[6] For further discussion of general discussion of generalized linear models, see McCullough and Nelder (1989), for example.

Level-1 (Student-level equation)

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_{0j} + \beta_{1j}(CPERIOD)_{ij} + \beta_{2j}(TRT)_{ij} + \beta_{3j}(CPERIOD)_{ij} * (TRT)_{ij}$$
$$+\beta_{4j}FEMALE_{ij} + \beta_{5j}BLACK_{ij} + \beta_{6j}HISPANIC_{ij} \qquad (1.5)$$
$$+\beta_{7j}ESH_{ij} + \beta_{8j}SED_{ij} + \beta_{9j}FRPL_{ij}$$

The specification of level-2 equations for binary outcomes is the same as that for continuous outcomes and is listed in Eq. (1.2). As with Eq. (1.3), the combined model can be created by substituting the level-2 equation into the level-1 equation. Because the predicted value in a binary hierarchical linear model is a log of the odds of that event occurring, we can recover the odds of the event occurring by exponentiation of the predicted value given the covariates in the model. Note that if the probability of an event occurring is less than 0.5, the odds ratio will be less than 1; if the probability of an event occurring is equal to 0.5, the odds are equal to 1; and if the probability of an event occurring is greater than 0.5, the odds will be greater than 1.

2. Three-Level Models

The previous section presented the models used to estimate the effects of FOHS on cross-sectional study outcomes. This section presents the three-level models that were used to estimate the effects of FOHS on the study's longitudinal outcomes. This section focuses on presenting models for binary and count outcomes, as there are no continuous, longitudinal outcomes being examined in the study.

The three-level model for longitudinal outcomes is similar to the two-level model for cross-sectional data, except that we have level-1 repeated observations that are nested within students at level-2, and students are nested within schools at level-3. At level-1, the three-level model estimates a baseline trajectory in outcomes for each of the students in the sample. The parameters in the level-1 model are outcomes at level-2, and these outcomes can vary across the population of students within schools, as a function of demographic characteristics, such as gender and minority status. The parameters in level-2 are then outcomes at level-3 that potentially vary across the population of schools as a function of school level characteristics.

a. Three-level Model for Binary Outcomes

Outcomes that are binary (for example, promoted or not) have a level-1 equation similar to the binary cross-sectional model presented above for level-1, but in this model the student-level covariates, such as gender and the student's baseline special education status, are now at level-2.

Level-1 (Equation for repeated student observations)

$$\log\left(\frac{p_{tij}}{1-p_{tij}}\right) = \pi_{0ij} + \pi_{1ij}(CPERIOD)_{tij} + \pi_{2ij}(TRT)_{tij} + \pi_{3ij}(CPERIOD)_{tij} * (TRT)_{tij} \qquad \textbf{(1.6)}$$

Level-2 (Student-level equations)

$$\begin{aligned}
\pi_{0ij} = \beta_{00j} &+ \beta_{01j}BLACK_{ij} + \beta_{02j}HISPANIC_{ij} + \beta_{03j}ESH_{ij} + \beta_{04j}SED_{ij} \\
&+ \beta_{05j}FRPL_{ij} + \beta_{06j}FEMALE_{ij} \\
&+ \beta_{07j}GRADE10_{ij} + \beta_{08j}GRADE11_{ij} + \beta_{09j}GRADE12_{ij} \qquad \textbf{(1.7)} \\
&+ \varepsilon_{kij} \\
\pi_{kij} = \beta_{k0j}&,\ k = 1, 2, 3.
\end{aligned}$$

Level-3 (School-level equations)

$$\begin{aligned}
\beta_{00j} &= \gamma_{000} + \gamma_{001}(\text{PFRPL})_j + \gamma_{002}(\text{PMIN})_j + u_{00j} \\
\beta_{0kj} &= \gamma_{0k0} \text{ for } k = 1, \ldots, 6 \qquad \textbf{(1.8)} \\
\beta_{k0j} &= \gamma_{k00} \text{ for } k = 1, 2, 3
\end{aligned}$$

Combined Model

Substituting Eqs. (1.7) and (1.8) into (1.6) yields the following combined model.[7]

$$
\begin{aligned}
\log\left(\frac{p_{tij}}{1-p_{tij}}\right) = {} & \gamma_{000} + \gamma_{001}\%FRPL_{j} + \gamma_{002}\%MIN_{j} \\
& +\gamma_{010}BLACK_{ij} + \gamma_{020}HISPANIC_{ij} + \gamma_{030}ESH_{ij} + \gamma_{040}SED_{ij} + \gamma_{050}FRPL_{ij} \\
& +\gamma_{060}FEMALE_{ij} + \gamma_{070}GRADE10_{ij} + \gamma_{080}GRADE11_{ij} + \gamma_{090}GRADE12_{ij} \\
& +\gamma_{100}(TRT)_{tij} + \gamma_{200}(CPERIOD)_{tij} + \gamma_{300}(CPERIOD)_{tij} * (TRT)_{tij} \\
& +[E_{ij} + U_{j}]
\end{aligned} \qquad \textbf{(1.9)}
$$

b. Three-Level Model for Count Outcomes

Outcomes that are counts (for example, number of days absent) follow a Poisson distribution with expected value and variance equal to each other and as follows:

$$E\{Y_{ij} \mid \lambda_{ij}\} = \lambda_{ij} \text{ and } \mathrm{V}\{Y_{ij} \mid \lambda_{ij}\} = \lambda_{ij},$$

where the predicted value is an event rate. Like binary outcomes, count outcomes must be treated differently than continuous outcomes when considering an appropriate level-1 model. Count outcomes are similar to binary outcomes in that the predicted value of the outcome is linearly related to the covariates in the model through a transformation, but for the count outcome, the transformation is the log function, resulting in the following level-1 equation:

Level 1 (Equation for repeated student observations)

$$\log(\lambda_{tij}) = \pi_{0ij} + \pi_{1ij}(CPERIOD)_{tij} + \pi_{2ij}(TRT)_{tij} + \pi_{3ij}(CPERIOD)_{tij} * (TRT)_{tij} \qquad \textbf{(1.10)}$$

---

[7] The last summation in the combined model is the random error in the model and contains many terms (for example,
$U_{j} = u_{00j} + u_{01j}FEMALE_{ij} + u_{02j}BLACK_{ij} + ... + u_{06j}FRPL_{ij} + u_{10j}CPERIOD_{tij} + u_{20j}TRT_{tij} + u_{30j}TRT_{tij} * CPERIOD_{tij}$
and $E_{ij} = e_{0ij} + e_{1ij}CPERIOD_{tij} + e_{2ij}TRT_{tij} + e_{3ij}TRT_{tij} * CPERIOD_{tij}$), but is denoted as $[E_{ij} + U_{j}]$ for ease of display.

The setup of level-2 and level-3 equations for count outcomes is the same as that for binary outcomes in the three-level model specified above in Eq. (1.8-1.9).

C. MODEL BUILDING

1. Functional Form

In selecting the appropriate functional form for our level-1 models, we examined a linear and a quadratic trend. All longitudinal outcomes, with the exception of promotion to the next grade, supported a quadratic trend, whereas MCAS outcomes supported a linear trend.[8]

2. Random Effects

The instantaneous effect of FOHS and the effect of FOHS on growth or change in the outcomes examined were modeled as random effects for the MCAS outcomes. For longitudinal outcomes, such random effects were not estimable in our models.

The model summaries are found in Table A.2. Continuous outcomes of MCAS scores and the binary outcomes of MCAS proficiency levels were modeled as two-level models and were determined to have linear trends. The binary outcome of whether the student was promoted was modeled as a three-level model, because it was longitudinal, and was also determined to have a linear trend. The other longitudinal outcomes were also modeled as three-level models and were determined to follow quadratic trends.

[8] A quadratic model was indicated by a significant interaction of the squared time by treatment interaction.

TABLE B.2

MODEL SUMMARIES

Outcome	Levels	Type	Trend
MCAS scaled mathematics score; MCAS scaled ELA score	2-level	Continuous	Linear
MCAS percent proficiency in mathematics; MCAS percent proficiency in ELA	2-level	Binary	Linear
Promoted	3-level	Binary	Linear
Suspended	3-level	Binary	Quadratic
Number of days absent, unexcused absences, present and tardy; Number of times suspended	3-level	Count	Quadratic

D. GENERATING ESTIMATES OF THE EFFECTS OF FOHS REFORMS

1. Overall Effects

The overall FOHS effects shown in Tables III.1 and III.2 are found from the appropriate linear contrast of fixed effects (that is, γ coefficients), according to the outcome type. The estimates for continuous outcomes represent simple differences between the predicted outcomes in each of the post-FOHS reform academic years relative to expected outcomes *in the absence of the FOHS reforms*. For example, in Table III.1, the –2.41 estimate shown for the MCAS 10th grade mathematics scaled scores in 2003-2004 indicates that scale scores were about 2.4 points lower than one would have predicted based on the trends evident during the *pre*-FOHS reform period. The associated p-value (<0.01) and asterisks (***) indicate that this estimated difference *is* statistically significant.[9]

[9] Note that these estimates take into account differences in the characteristics of students attending different BPS high schools, as well as changes in the characteristics of students attending BPS high schools over time (to the extent possible based on the available administrative records data).

2. Academic Year Effects

In addition to the overall FOHS effect, academic-year effects for the overall sample were estimated and tested through a general linear hypothesis test of the fixed effects. This results in estimated FOHS effects for the overall sample for 2002-2003 (MCAS outcomes only), 2003-2004, 2004-2005, 2005-2006 and 2006-2007. These academic-year effects for the overall sample are shown in Tables III.1 and III.2.

3. School-Level Effects

School-level estimates were obtained for all outcomes, except days tardy. For these outcomes, we were able to specify a random intervention effect in the appropriate "cutoff" school year (i.e., 2002-2003 for MCAS outcomes and 2003-2004 for other outcomes) as well as a random effect of the intervention on the linear and quadratic growth parameters, as appropriate. The specification of these random effects permitted the calculation of school-level estimates, which are empirical Bayes (EB) estimates (Raudenbush and Bryk 2002). These school-level estimates, along with their confidence intervals, are reported in Figures III.12 through III.21.

4. Subgroup Effects

Subgroup analyses were based on models that had the same functional form as the full analysis model for a particular outcome and included all interactions between subgroup variables and the intervention effect, as well as three-way interactions between the subgroup variables, the intervention effect, and time. Results from the subgroup analyses are presented in Tables III.3 and III.4.