WWC Single Study Review

A review of the design and summary of findings for an individual study

ies INSTITUTE OF EDUCATION SCIENCES

April 2016 Revised*

WWC Review of the Report "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice"[1,2]

The findings from this review do not reflect the full body of research evidence on Boston charter schools.

What is this study about?

The study measured the effects of attending Boston charter high schools on students' reading and math achievement, high school graduation, and college enrollment and persistence. Six Boston charter schools that included one or more high school grades participated in the study.

Researchers analyzed cohorts of students who applied to the study schools between 2002–09, a period when there were more applicants than the study schools could admit. When the number of applicants exceeded enrollment limits, a lottery was used to randomly select students who would be offered admission. Students who were not selected were not offered admission to a charter school. Some students who were offered admission did not ultimately enroll in a charter school, and some students who were not initially offered admission through the lottery ultimately enrolled in a charter school. The analytic sample contained up to 3,920 students who applied to a charter school.

Study authors used statistical techniques to estimate the differences in outcomes for students who would enroll in a charter school if offered admission but would not enroll if they were assigned to the comparison group. These students are referred to as compliers, and the average effect of taking up the intervention among compliers is called the complier average causal effect (CACE).[3]

The study measured the CACE of attending a charter school on student achievement in math and reading, rates of taking Advanced Placement (AP) exams and the SAT, performance on AP exams and the SAT, high school graduation within 4 and 5 years, college enrollment, and college persistence.[4]

WWC Rating

The research described in this report meets WWC group design standards without reservations

The study design is based on randomized offers of admission to charter schools. Using specialized statistical techniques, the study authors estimated the effects of enrollment in a charter school among students who would enroll in a charter school if offered admission and would not enroll in a charter school if they were not offered admission. This estimate is known as the complier average causal effect (CACE).

The *meets WWC group design standards without reservations* rating pertains to all analyses of outcomes measured in high school, such as high school graduation rates and student performance on Advanced Placement (AP) exams and the SAT. Some analyses of college enrollment and persistence do not meet WWC group design standards because the relationship between charter school enrollment and the lottery outcomes ("offered enrollment on the day of the lottery" and "ever offered enrollment") is not sufficiently strong for the analytic samples. The WWC requires a sufficiently strong relationship between these variables because the strength of the relationship influences the precision of the statistical techniques that are used to calculate CACE estimates.[5]

What did the study find?

The study authors found that enrollment in Boston charter schools increased student achievement in math and reading, the rates of taking AP exams, performance on AP exams and the SAT, and college enrollment at 2-year colleges within 6 months of high school graduation. The study also found that enrollment in Boston charter schools decreased the fraction of students graduating within 4 years, but there were no statistically significant impacts on the fraction of students graduating within 5 years. The study also reported that enrollment in Boston charter schools improved performance on AP exams in Science, Calculus, and US History. Finally, the study reported impacts of Boston charter school enrollment on some key outcomes for student subgroups defined by gender, eligibility for free and reduced-price lunch, and special education status. For several of these subgroups, the study found statistically significant improvements in student achievement in math and reading and the rates of taking AP exams. The study findings are shown in Appendices C and D of this report. The WWC determined that the presentation of the findings in the study was credible.

Features of Boston Charter Schools

Charter schools are public schools that have been established on the basis of a contract, or charter, held by a private board of directors. They are exempt from many state and district regulations that govern traditional public schools, including those involving staffing, curriculum, and budget decisions. Charter school attendance is free of charge. If more students want to enroll in a charter school than the charter school can enroll, an admissions lottery must be held.

Massachusetts' urban charter schools typically use a *No Excuses* pedagogy, which emphasizes discipline, reading and math skills, extended instruction time, and selective teacher hiring.[6] Like most Boston charter schools, the Boston charter high schools included in this study largely utilize the *No Excuses* approach.

Appendix A: Study details

Angrist, J. D., Cohodes, S. R., Dynarski, S. M., Pathak, P. A., & Walters, C. R. (2014). *Stand and deliver: Effects of Boston's charter high schools on college preparation, entry, and choice*. Cambridge, MA: Department of Economics Working Paper, Massachusetts Institute of Technology.[1]

Setting

The study was conducted in Boston, Massachusetts.

Study sample

The study collected applicant lists from fall 2002 through fall 2009 from six charter schools that had more applicants than they could enroll and used lotteries to determine who would be offered enrollment. The six participating schools were Match, Codman Academy, City on a Hill, Boston Preparatory, Academy of the Pacific Rim, and Boston Collegiate. The grade level of students participating in the lotteries differed by school: ninth-grade lotteries were held for Match, Codman Academy, and City on a Hill; sixth-grade lotteries were held for Boston Preparatory and Academy of the Pacific Rim; and a fifth-grade lottery was held for Boston Collegiate. Applicant lists were then matched to administrative records for all Massachusetts public school students. Applicants were excluded from the analytic sample if they (1) were disqualified from the lottery (usually because they were in the wrong grade), (2) were siblings of other charter school students, (3) applied late, or (4) were from out of the area. After these sample restrictions, there were 4,711 charter applicants in the sample. Analytic samples for outcomes assessed in high school range from over 2,300 students (scores on the SAT) to 3,920 students (started the twelfth grade on time, and graduated from high school within four years).

In the analytic sample for the tenth grade English language arts [ELA] scores on the Massachusetts Comprehensive Assessment System [MCAS] assessment, 54% of the students were female; 61% were Black, 31% were Hispanic, and 3% were Asian.

Intervention group

The intervention group consisted of students who were admitted to one of the six charter schools at the time of the admission lottery or afterward. The intervention was defined as enrollment in one of six charter schools in Boston that used lotteries for admission between fall 2002 and fall 2009.

Comparison group

The comparison group consisted of students who were never admitted to one of the six charter schools. Students in the comparison condition participated in an admissions lottery, but did not attend any of the six charter schools at any point in their academic careers.

Outcomes and measurement

Eligible outcomes included mathematics and ELA scores on the tenth-grade MCAS; MCAS performance category based on test scores ("Needs Improvement" or higher, "Proficient" or higher, "Advanced" or higher); competency determination and eligibility for the John and Abigail Adams scholarship based on MCAS scores; rates of high school graduation within 4 and 5 years; whether at least one AP exam was taken; the number of AP exams taken; the number of AP exams taken for which the score was 2 or higher, 3 or higher, or 4 or higher; whether the SAT was taken; Math, Verbal, and Writing scores on the SAT; and quartiles of Math, Verbal, and Writing scores on the SAT. The study also examined measures of college enrollment (enrollment within 6 and 18 months of expected high school graduation) and college persistence (enrolled for one, three, or five semesters within 18, 30, and 42 months of expected high school graduation). For a more detailed description of these outcome measures, see Appendix B. Findings for outcomes that the study designated as key outcomes are shown in Appendix C. Findings for other outcomes were considered supplemental and are shown in Appendix D.

Support for implementation	This is not applicable for this intervention.
Reason for review	This study was identified for review by receiving media attention.

Appendix B: Outcome measures for each domain

Mathematics achievement	
Massachusetts Comprehensive Assessment System (MCAS), tenth-grade Mathematics score	The MCAS is a state assessment that measures student performance in mathematics, English language arts (ELA), and science and technology/engineering. Scores were converted to *z*-scores to measure student achievement relative to the average statewide performance within a particular grade, year, and subject.
Received a score of 2 or higher on at least one Advanced Placement (AP) Calculus exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP Calculus exam and received a score of 2 or higher. This measure is equal to 0 if the student did not take an AP Calculus exam, or if the student received a score of 1. Students' most recent AP exam scores were used.
Received a score of 3 or higher on at least one AP Calculus exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP Calculus exam and received a score of 3 or higher. This measure is equal to 0 if the student did not take an AP Calculus exam, or if the student scored below 3. Students' most recent AP exam scores were used.
Received a score of 4 or 5 on at least one AP Calculus exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP Calculus exam and received a score of 4 or 5. This measure is equal to 0 if the student did not take an AP Calculus exam, or if the student scored below 4. Students' most recent AP exam scores were used.
SAT Math score	The SAT is a standardized college admissions exam that measures student performance. The SAT Math score in this study ranged from 200 to 800 points. Students' most recent SAT scores were used.
SAT Math score was above the bottom quartile	The SAT is a standardized college admissions exam that measures student performance. This is a binary measure that is equal to 1 if the student took the SAT and received a score on the SAT Math section that was above the bottom quartile of SAT Math scores for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student scored in the bottom quartile. Students' most recent SAT scores were used.
SAT Math score was above the median	The SAT is a standardized college admissions exam that measures student performance. This is a binary measure that is equal to 1 if the student took the SAT and received a score on the SAT Math section that was above the median SAT Math score for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student scored below the median. Students' most recent SAT scores were used.
SAT Math score was in the top quartile	The SAT is a standardized college admissions exam that measures student performance. This is a binary measure that is equal to 1 if the student took the SAT and received a score on the SAT Math section that was in the top quartile of SAT Math scores for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student was not in the top quartile. Students' most recent SAT scores were used.
General literacy achievement	
MCAS, tenth-grade ELA score	The MCAS is a state assessment that measures student performance in mathematics, ELA, and science and technology/engineering. Scaled scores were converted to *z*-scores to measure student achievement relative to the average statewide performance within a particular grade, year, and subject.
Received a score of 2 or higher on at least one AP English exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP English exam and received a score of 2 or higher. This measure is equal to 0 if the student did not take an AP English exam or if the student received a score of 1. Students' most recent AP exam scores were used.
Received a score of 3 or higher on at least one AP English exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP English exam and received a score of 3 or higher. This measure is equal to 0 if the student did not take an AP English exam, or if the student scored below 3. Students' most recent AP exam scores were used.
Received a score of 4 or 5 on at least one AP English exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP English exam and received a score of 4 or 5. This measure is equal to 0 if the student did not take an AP English exam, or if the student scored below 4. Students' most recent AP exam scores were used.
SAT Verbal score	The SAT is a standardized college admissions exam that measures student performance. The SAT Verbal score in this study ranged from 200 to 800 points. Students' most recent SAT scores were used.
SAT Verbal score was above the bottom quartile	The SAT is a standardized college admissions exam that measures student performance. This is a binary measure that is equal to 1 if the student took the SAT and received a score on the SAT Verbal section that was above the bottom quartile of SAT Verbal scores for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student scored in the bottom quartile. Students' most recent SAT scores were used.

SAT Verbal score was above the median	The SAT is a standardized college admissions exam that measures student performance. This is a binary measure that is equal to 1 if the student took the SAT and received a score on the SAT Verbal section that was above the median SAT Verbal score for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student did not score above the median. Students' most recent SAT scores were used.
SAT Verbal score was in the top quartile	The SAT is a standardized college admissions exam that measures student performance. This is a binary measure that is equal to 1 if the student took the SAT and received a score on the SAT Verbal section that was in the top quartile of SAT Verbal scores for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student did not score in the top quartile. Students' most recent SAT scores were used.
SAT Writing score	The SAT is a standardized college admissions exam that measures student performance. The SAT Writing score in this study ranged from 200 to 800 points. Students' most recent SAT scores were used.
SAT Writing score was above the bottom quartile	The SAT is a standardized college admissions exam that measures student performance. This is a binary measure that is equal to 1 if the student took the SAT and received a score on the SAT Writing section that was above the bottom quartile of SAT Writing scores for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student scored in the bottom quartile. Students' most recent SAT scores were used.
SAT Writing score was above the median	The SAT is a standardized college admissions exam that measures student performance. This is a binary measure that is equal to 1 if the student took the SAT and received a score on the SAT Writing section that was above the median SAT Writing score for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student scored below the median. Students' most recent SAT scores were used.
SAT Writing score was in the top quartile	The SAT is a standardized college admissions exam that measures student performance. This is a binary measure that is equal to 1 if the student took the SAT and received a score on the SAT Writing section that was in the top quartile of SAT Writing scores for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student did not score in the top quartile. Students' most recent SAT scores were used.
History achievement	
Received a score of 2 or higher on at least one AP US History exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP US History exam and received a score of 2 or higher. This measure is equal to 0 if the student did not take an AP US History exam, or if the student received a score of 1. Students' most recent AP exam scores were used.
Received a score of 3 or higher on at least one AP US History exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP US History exam and received a score of 3 or higher. This measure is equal to 0 if the student did not take an AP US History exam, or if the student received a score below 3. Students' most recent AP exam scores were used.
Received a score of 4 or 5 on at least one AP US History exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP US History exam and received a score of 4 or 5. This measure is equal to 0 if the student did not take an AP US History exam, or if the student received a score below 4. Students' most recent AP exam scores were used.
Science achievement	
Received a score of 2 or higher on at least one AP Science exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP Science exam and received a score of 2 or higher. This measure is equal to 0 if the student did not take an AP Science exam, or if the student received a score of 1. Students' most recent AP exam scores were used.
Received a score of 3 or higher on at least one AP Science exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP Science exam and received a score of 3 or higher. This measure is equal to 0 if the student did not take an AP Science exam, or if the student received a score below 3. Students' most recent AP exam scores were used.
Received a score of 4 or 5 on at least one AP Science exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP Science exam and received a score of 4 or 5. This measure is equal to 0 if the student did not take an AP Science exam, or if the student received a score below 4. Students' most recent AP exam scores were used.
General academic achievement	
Eligible for John and Abigail Adams scholarship	This binary measure is recorded as 1 if a student was eligible for the John and Abigail Adams scholarship, and 0 otherwise. Students were eligible for the Adams scholarship if they were in the high school graduating class of 2015 or earlier, were scored as "Advanced" or "Proficient" on the tenth-grade MCAS exams in mathematics and ELA, were scored as "Advanced" in at least one of the tenth-grade MCAS exams in mathematics and ELA, and had combined scores on the MCAS exams in mathematics and ELA that were in the top quartile of students in the graduating class in their district.

MCAS performance category of "Needs Improvement" or higher	The MCAS is a state assessment that measures student performance in mathematics, ELA, and science and technology/engineering. This is a binary measure that is equal to 1 if the student's scaled scores on the ELA and mathematics sections were at or above 220 on any attempt, and 0 otherwise.
MCAS performance category of "Proficient" or higher	The MCAS is a state assessment that measures student performance in mathematics, ELA, and science and technology/engineering. This is a binary measure that is equal to 1 if the student's scaled scores on the mathematics and ELA sections were at or above 240 on any attempt, and 0 otherwise.
MCAS performance category of "Advanced" or higher	The MCAS is a state assessment that measures student performance in mathematics, ELA, and science and technology/engineering. This is a binary measure that is equal to 1 if the student's scaled scores on the mathematics and ELA sections were at or above 260 on any attempt, and 0 otherwise.
Meets Competency Determination	The Competency Determination is a high school graduation requirement in Massachusetts based on the MCAS. From 2006–09, students were required to receive scaled scores of 220 or higher in the mathematics and ELA sections of the MCAS to graduate from high school. From 2010–13, the required range of scores was changed to 240 or higher in the mathematics and ELA sections of the MCAS. This is a binary measure that is equal to 1 if the student met the applicable Competency Determination requirement, and 0 otherwise.
Received a score of 2 or higher on at least one AP exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP exam and received a score of 2 or higher. This measure is equal to 0 if the student did not take an AP exam, or if the student received a score of 1 on all AP exams taken. Students' most recent AP exam scores were used.
Received a score of 3 or higher on at least one AP exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP exam and received a score of 3 or higher. This measure is equal to 0 if the student did not take an AP exam, or if the student received a score below 3 on all AP exams taken. Students' most recent AP exam scores were used.
Received a score of 4 or 5 on at least one AP exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took an AP exam and received a score of 4 or 5. This measure is equal to 0 if the student did not take an AP exam, or if the student received a score below 4 on all AP exams taken. Students' most recent AP exam scores were used.
SAT Composite score	The SAT is a standardized college admissions exam that measures student performance. The SAT Composite score in this study is the sum of the SAT Math, SAT Verbal, and SAT Writing scores and ranged from 800 to 2,400 points. Students' most recent SAT scores were used.
SAT Composite score was above the bottom quartile	The SAT is a standardized college admissions exam that measures student performance. The SAT Composite score in this study is the sum of the SAT Math, SAT Verbal, and SAT Writing scores. This is a binary measure that is equal to 1 if the student took the SAT, and the SAT Composite score was above the bottom quartile of SAT Composite scores for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student's SAT Composite score was in the bottom quartile. Students' most recent SAT scores were used.
SAT Composite score was above the median	The SAT is a standardized college admissions exam that measures student performance. The SAT Composite score in this study is the sum of the SAT Math, SAT Verbal, and SAT Writing scores. This is a binary measure that is equal to 1 if the student took the SAT, and the SAT Composite score was above the median SAT Composite score for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student's SAT Composite score was below the median. Students' most recent SAT scores were used.
SAT Composite score was in the top quartile	The SAT is a standardized college admissions exam that measures student performance. The SAT Composite score in this study is the sum of the SAT Math, SAT Verbal, and SAT Writing scores. This is a binary measure that is equal to 1 if the student took the SAT, and the SAT Composite score was in the top quartile of SAT Composite scores for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student's SAT Composite score was below the top quartile. Students' most recent SAT scores were used.
SAT Reasoning score	The SAT is a standardized college admissions exam that measures student performance. The SAT Reasoning score in this study is the sum of the SAT Math and SAT Verbal scores and ranged from 400 to 1,600 points. Students' most recent SAT scores were used.
SAT Reasoning score was above the bottom quartile	The SAT is a standardized college admissions exam that measures student performance. The SAT Reasoning score in this study is the sum of the SAT Math and SAT Verbal scores. This is a binary measure that is equal to 1 if the student took the SAT, and the SAT Reasoning score was above the bottom quartile of SAT Reasoning scores for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student's Reasoning score was in the bottom quartile. Students' most recent SAT scores were used.

SAT Reasoning score was above the median	The SAT is a standardized college admissions exam that measures student performance. The SAT Reasoning score in this study is the sum of the SAT Math and SAT Verbal scores. This is a binary measure that is equal to 1 if the student took the SAT, and the SAT Reasoning score was above the median SAT Reasoning score for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student's Reasoning score was below the median. Students' most recent SAT scores were used.
SAT Reasoning score was in the top quartile	The SAT is a standardized college admissions exam that measures student performance. The SAT Reasoning score in this study is the sum of the SAT Math and SAT Verbal scores. This is a binary measure that is equal to 1 if the student took the SAT, and the SAT Reasoning score was in the top quartile of SAT Reasoning scores for students in Massachusetts. This measure is equal to 0 if the student did not take the SAT, or if the student's Reasoning score was below the top quartile. Students' most recent SAT scores were used.
Completing school	
Graduated from high school within 4 years	This is a binary measure that is equal to 1 if the student graduated by the projected graduation year, and 0 otherwise. The projected graduation year is based on the grade to which the student applied in the lottery; for instance, the projected graduation year for an applicant to the fifth grade would be 8 years in the future (i.e., including the fifth grade).
Graduated from high school within 5 years	This is a binary measure that is equal to 1 if the student graduated by the year following the projected graduation year, and 0 otherwise. The projected graduation year is based on the grade to which the student applied in the lottery; for instance, the year following the projected graduation year for an applicant to the fifth grade would be 9 years in the future (i.e., including the fifth grade).
Repeated the twelfth grade	This is a binary measure that is equal to 1 if the student repeated the twelfth grade one or more times, and 0 otherwise.
Started the twelfth grade on time	This is a binary measure that is equal to 1 if the student was in the twelfth grade by the projected graduation year, and 0 otherwise. The projected graduation year is based on the grade to which the student applied in the lottery; for instance, the projected graduation year for an applicant to the fifth grade would be 8 years in the future (i.e., including the fifth grade).
College enrollment	
Enrolled in any college within 6 months	This is a binary measure that is equal to 1 if the student enrolled in any college within 6 months (one semester) after the projected graduation year, and 0 otherwise. The projected graduation year is based on the grade to which the student applied in the lottery; for instance, the projected graduation year for an applicant to fifth grade would be 8 years in the future (i.e., including fifth grade).
Enrolled in a 2-year college within 6 months	This is a binary measure that is equal to 1 if the student enrolled in a 2-year college within 6 months (one semester) after the projected graduation year, and 0 otherwise. The projected graduation year is based on the grade to which the student applied in the lottery; for instance, the projected graduation year for an applicant to the fifth grade would be 8 years in the future (i.e., including the fifth grade).
Enrolled in a 4-year college within 6 months	This is a binary measure that is equal to 1 if the student enrolled in a 4-year college within 6 months (one semester) after the projected graduation year, and 0 otherwise. The projected graduation year is based on the grade to which the student applied in the lottery; for instance, the projected graduation year for an applicant to the fifth grade would be 8 years in the future (i.e., including the fifth grade).
Enrolled in a 4-year public college within 6 months	This is a binary measure that is equal to 1 if the student enrolled in a 4-year public college within 6 months (one semester) after the projected graduation year, and 0 otherwise. The projected graduation year is based on the grade to which the student applied in the lottery; for instance, the projected graduation year for an applicant to the fifth grade would be 8 years in the future (i.e., including the fifth grade).
Enrolled in a 4-year private college within 6 months	This is a binary measure that is equal to 1 if the student enrolled in a 4-year private college within 6 months (one semester) after the projected graduation year, and 0 otherwise. The projected graduation year is based on the grade to which the student applied in the lottery; for instance, the projected graduation year for an applicant to the fifth grade would be 8 years in the future (i.e., including the fifth grade).
Enrolled in a 4-year public college in Massachusetts within 6 months	This is a binary measure that is equal to 1 if the student enrolled in a 4-year public college in Massachusetts within 6 months (one semester) after the projected graduation year, and 0 otherwise. The projected graduation year is based on the grade to which the student applied in the lottery; for instance, the projected graduation year for an applicant to the fifth grade would be 8 years in the future (i.e., including the fifth grade).

Enrolled in any college within 18 months	This is a binary measure that is equal to 1 if the student enrolled in any college within 18 months (three semesters) after the projected graduation year, and 0 otherwise. The projected graduation year is based on the grade to which the student applied in the lottery; for instance, the projected graduation year for an applicant to the fifth grade would be 8 years in the future (i.e., including the fifth grade).
Enrolled in a 4-year college within 18 months	This is a binary measure that is equal to 1 if the student enrolled in a 4-year college within 18 months (three semesters) after the projected graduation year, and 0 otherwise. The projected graduation year is based on the grade to which the student applied in the lottery; for instance, the projected graduation year for an applicant to the fifth grade would be 8 years in the future (i.e., including the fifth grade).
College preparation	
Number of AP exams taken	AP exams are end-of-course exams that are scored from 1 to 5. This measure counts the number of AP exams that were taken. This measure is equal to 0 if the student did not take an AP exam.
Number of AP Science exams taken	AP exams are end-of-course exams that are scored from 1 to 5. This measure counts the number of AP Science exams that were taken. This measure is equal to 0 if the student did not take an AP Science exam.
Took at least one AP Calculus exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took at least one AP Calculus exam. This measure is equal to 0 if the student did not take an AP Calculus exam.
Took at least one AP English exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took at least one AP English exam. This measure is equal to 0 if the student did not take an AP English exam.
Took at least one AP exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took at least one AP exam. This measure is equal to 0 if the student did not take an AP exam.
Took at least one AP US History exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took at least one AP US History exam. This measure is equal to 0 if the student did not take an AP US History exam.
Took at least one AP Science exam	AP exams are end-of-course exams that are scored from 1 to 5. This is a binary measure that is equal to 1 if the student took at least one AP Science exam. This measure is equal to 0 if the student did not take an AP Science exam.
Took SAT	The SAT is a standardized college admissions exam that measures student performance. This is a binary measure that is equal to 1 if the student took the SAT. This measure is equal to 0 if the student did not take the SAT.

Table Notes: In addition to the outcomes in this table, the study assessed measures of college enrollment within 18 months of the projected high school graduation for any colleges, 2-year colleges, 4-year public colleges, 4-year private colleges, and 4-year Massachusetts public colleges. The study also assessed measures of college persistence (enrolled for three or five semesters) within 18, 30, and 42 months of the projected high school graduation by college type (any college, 2-year college, 4-year college, 4-year public college, 4-year private college, 4-year Massachusetts public college). These measures are not included in this single study review because lottery outcomes (the instrumental variable) was not a sufficiently strong predictor of charter school enrollment (the measure of compliance) for these analytic samples; consequently, these analyses do not meet WWC group design standards.

Appendix C: Study findings for each domain

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Mathematics achievement								
Massachusetts Comprehensive Assessment System (MCAS), tenth-grade Mathematics score	All students	3,629 students	0.36 (na)	–0.23 (0.91)	0.59	0.59	+22	< .01
Domain average for mathematics achievement						**0.59**	**+22**	**Statistically significant**
General literacy achievement								
MCAS, tenth-grade English language arts score	All students	3,685 students	0.12 (na)	–0.29 (0.83)	0.41	0.41	+16	< .01
Domain average for general literacy achievement						**0.41**	**+16**	**Statistically significant**
General academic achievement								
Received a score of 3 or higher on at least one Advanced Placement (AP) exam	All students	3,672 students	0.21 (0.40)	0.08 (0.28)	0.12	0.34	+13	< .05
SAT Composite score	Students who took the SAT	2,378 students	1,372.50 (223.00)	1,268.20 (250.30)	104.30	0.45	+17	< .01
Domain average for general academic achievement						**0.39**	**+15**	**Statistically significant**
Completing school								
Graduated from high school within 4 years	All students	3,920 students	0.54 (0.50)	0.69 (0.46)	–0.15	–0.30	–12	< .01
Graduated from high school within 5 years	All students	3,208 students	0.78 (0.41)	0.79 (0.41)	0.00	–0.01	0	> .10
Domain average for completing school						**–0.15**	**–6**	**Statistically significant**
College preparation								
Took at least one AP test	All students	3,672 students	0.58 (0.49)	0.28 (0.45)	0.30	0.62	+23	< .01
Took the SAT	All students	3,672 students	0.72 (0.45)	0.64 (0.48)	0.08	0.18	+7	> .10
Domain average for college preparation						**0.40**	**+16**	**Statistically significant**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on individual outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study's domain average was determined by the WWC. Some statistics may not sum as expected due to rounding. na = not available.

Study Notes: Corrections for multiple comparisons were needed for the general academic achievement domain and the completing school domain in this table but did not affect whether any of the contrasts were found to be statistically significant. The *p*-values presented here were reported in the original study, and the effect sizes presented for the MCAS scores in this table are equal to the impact estimates reported in the study. This study is characterized as having a statistically significant positive effect in mathematics, general literacy achievement, general academic achievement, and college preparation because the effect for at least one measure within each domain is positive and statistically significant, and no effects within these domains are negative and statistically significant, accounting for multiple comparisons. This study is characterized as having a statistically significant negative effect on completing school because the effect for at least one measure within the domain is negative and statistically significant, and no effects within this domain are positive and statistically significant, accounting for multiple comparisons. For more information, please refer to the WWC Standards and Procedures Handbook (version 3.0), p. 26.

Appendix D.1: Supplemental findings by domain for all students

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Mathematics achievement								
SAT Math score was above the bottom quartile	All students	3,672 students	0.49 (0.50)	0.32 (0.47)	0.17	0.36	+14	< .05
SAT Math score was above the median	All students	3,672 students	0.29 (0.45)	0.13 (0.33)	0.17	0.40	+16	< .01
SAT Math score was in the top quartile	All students	3,672 students	0.11 (0.31)	0.04 (0.20)	0.07	0.24	+9	< .05
Received a score of 2 or higher on at least one Advanced Placement (AP) Calculus exam	All students	3,672 students	0.15 (0.36)	0.03 (0.16)	0.13	0.42	+16	< .01
Received a score of 3 or higher on at least one AP Calculus exam	All students	3,672 students	0.13 (0.33)	0.02 (0.14)	0.11	0.39	+15	< .01
Received a score of 4 or 5 on at least one AP Calculus exam	All students	3,672 students	0.05 (0.21)	0.01 (0.10)	0.04	0.20	+8	> .05
General literacy achievement								
Received a score of 2 or higher on at least one AP English exam	All students	3,672 students	0.17 (0.38)	0.11 (0.31)	0.07	0.19	+8	> .10
Received a score of 3 or higher on at least one AP English exam	All students	3,672 students	0.07 (0.26)	0.04 (0.18)	0.04	0.16	+6	> .10
Received a score of 4 or 5 on at least one AP English exam	All students	3,672 students	0.02 (0.15)	0.01 (0.12)	0.01	0.06	+2	> .10
SAT Verbal score was above the bottom quartile	All students	3,672 students	0.41 (0.49)	0.27 (0.45)	0.14	0.29	+11	< .01
SAT Verbal score was above the median	All students	3,672 students	0.20 (0.40)	0.11 (0.32)	0.09	0.23	+9	< .05
SAT Verbal score was in the top quartile	All students	3,672 students	0.05 (0.21)	0.03 (0.18)	0.01	0.06	+2	> .10
SAT Writing score was above the bottom quartile	All students	3,672 students	0.39 (0.49)	0.29 (0.45)	0.10	0.21	+8	< .10
SAT Writing score was above the median	All students	3,672 students	0.18 (0.39)	0.10 (0.31)	0.08	0.22	+9	< .05
SAT Writing score was in the top quartile	All students	3,672 students	0.07 (0.25)	0.03 (0.17)	0.04	0.17	+7	> .10
History achievement								
Received a score of 2 or higher on at least one AP US History exam	All students	3,672 students	0.12 (0.32)	0.03 (0.17)	0.09	0.32	+12	< .05

Received a score of 3 or higher on at least one AP US History exam	All students	3,672 students	0.08 (0.27)	0.02 (0.14)	0.06	0.25	+10	< .01
Received a score of 4 or 5 on at least one AP US History exam	All students	3,672 students	0.03 (0.18)	0.01 (0.10)	0.02	0.14	+5	> .10
Science achievement								
Received a score of 2 or higher on at least one AP Science exam	All students	3,672 students	0.10 (0.30)	0.04 (0.18)	0.07	0.25	+10	< .05
Received a score of 3 or higher on at least one AP Science exam	All students	3,672 students	0.07 (0.25)	0.02 (0.14)	0.05	0.21	+8	< .05
Received a score of 4 or 5 on at least one AP Science exam	All students	3,672 students	0.01 (0.09)	0.01 (0.10)	0.00	–0.03	–1	> .10
General academic achievement								
Massachusetts Comprehensive Assessment System (MCAS) performance category of "Needs Improvement" or higher	All students	3,608 students	0.99 (0.10)	0.98 (0.16)	0.01	0.11	+4	> .10
MCAS performance category of "Proficient" or higher	All students	3,608 students	0.71 (0.45)	0.54 (0.50)	0.17	0.37	+14	< .05
MCAS performance category of "Advanced" or higher	All students	3,608 students	0.24 (0.43)	0.08 (0.26)	0.16	0.43	+16	< .01
Meets Competency Determination	All students	3,608 students	0.89 (0.32)	0.74 (0.44)	0.15	0.41	+16	< .05
Eligible for Adams scholarship	All students	3,608 students	0.44 (0.50)	0.20 (0.40)	0.24	0.52	+20	< .01
Received a score of 2 or higher on at least one AP exam	All students	3,672 students	0.32 (0.47)	0.16 (0.36)	0.16	0.38	+15	< .05
Received a score of 4 or 5 on at least one AP exam	All students	3,672 students	0.10 (0.29)	0.05 (0.21)	0.05	0.19	+7	> .10
SAT Reasoning score was above the bottom quartile	All students	3,672 students	0.43 (0.50)	0.27 (0.44)	0.17	0.35	+14	< .01
SAT Reasoning score was above the median	All students	3,672 students	0.23 (0.42)	0.10 (0.30)	0.12	0.32	+13	< .01
SAT Reasoning score was in the top quartile	All students	3,672 students	0.05 (0.22)	0.03 (0.18)	0.02	0.08	+3	> .10
SAT Composite score was above the bottom quartile	All students	3,672 students	0.41 (0.49)	0.27 (0.44)	0.15	0.31	+12	< .01
SAT Composite score was above the median	All students	3,672 students	0.20 (0.40)	0.09 (0.29)	0.11	0.29	+12	< .01
SAT Composite score was in the top quartile	All students	3,672 students	0.04 (0.18)	0.03 (0.16)	0.01	0.05	+2	> .10
Completing school								
Started the twelfth grade on time	All students	3,920 students	0.76 (0.43)	0.78 (0.41)	–0.02	na	na	> .10

College preparation								
Number of AP exams taken	All students	3,672 students	1.62 (1.48)	0.59 (1.17)	1.04	0.75	+27	< .01
Took at least one AP Calculus exam	All students	3,672 students	0.29 (0.46)	0.07 (0.26)	0.22	0.56	+21	< .01
Took at least one AP English exam	All students	3,672 students	0.25 (0.43)	0.16 (0.37)	0.08	0.20	+8	> .01
Took at least one AP US History exam	All students	3,672 students	0.22 (0.42)	0.04 (0.20)	0.18	0.51	+20	< .05
Took at least one AP Science exam	All students	3,672 students	0.41 (0.49)	0.10 (0.30)	0.31	0.71	+26	< .01
Number of AP Science exams taken	All students	3,672 students	0.41 (0.44)	0.12 (0.37)	0.29	0.70	+26	< .01
College enrollment								
Enrolled in any college within 6 months	All students	3,205 students	0.53 (0.50)	0.50 (0.50)	0.03	na	na	> .10
Enrolled in a 2-year college within 6 months	All students	3,205 students	0.02 (0.13)	0.12 (0.33)	–0.11	na	na	< .05
Enrolled in a 4-year college within 6 months	All students	3,205 students	0.51 (0.50)	0.37 (0.48)	0.13	na	na	< .05
Enrolled in a 4-year public college within 6 months	All students	3,205 students	0.28 (0.45)	0.14 (0.35)	0.14	na	na	< .01
Enrolled in a 4-year private college within 6 months	All students	3,205 students	0.22 (0.42)	0.23 (0.42)	–0.01	na	na	> .10
Enrolled in a 4-year public college in Massachusetts within 6 months	All students	3,205 students	0.24 (0.43)	0.12 (0.33)	0.12	na	na	< .05

Table Notes: The supplemental findings presented in this table are additional findings that meet WWC design standards without reservations, but do not factor into the determination of the study rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on individual outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). Effect size estimates are not available for contrasts where sample sizes for the intervention and comparison groups are unknown. The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding. na = not available.

Study Notes: The *p*-values presented here were reported in the original study. A correction for multiple comparisons was needed for each outcome domain across the supplemental findings in Appendices D.1 and D.2. The study presented impact estimates of key outcomes for the full sample, boys, girls, special education students, non-special education students, students whose baseline MCAS scores were below the median, students whose baseline MCAS scores were above the median, students who were eligible for free and reduced-price lunch, and students who were not eligible for free and reduced-price lunch. Baseline MCAS scores were measured prior to charter application, in either the fourth or eighth grade. The correction for multiple comparisons across the supplemental findings in Appendices D.1 and D.2 resulted in WWC-computed critical *p*-values of .033 for enrollment in a 4-year college within 6 months (Appendix D.1); .007 for repeating the twelfth grade (Appendix D.2), .011 for graduating from high school within 4 years for boys (Appendix D.2), .014 for graduating from high school within 4 years for students who were eligible for free and reduced-price lunch (Appendix D.2), .034 for SAT Composite scores among students who were eligible for free and reduced-price lunch (Appendix D.1), .025 for whether SAT Writing scores were above the median (Appendix D.1); .028 for average SAT Writing scores (Appendix D.1); .031 for whether SAT Verbal scores were above the median (Appendix D.1); .033 for average SAT Verbal scores (Appendix D.1); and .033 for whether AP US History scores were 2 or higher (Appendix D.1). Because the *p*-values for these impact estimates were higher than the WWC-computed critical *p*-values, the WWC does not find the results for these particular outcomes to be statistically significant.

Appendix D.2: Supplemental findings by domain of subgroups of students

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Mathematics achievement								
SAT Math score	Students who took the SAT	2,378 students	491.20 (92.40)	439.10 (98.30)	52.10	0.55	+21	< .01
Massachusetts Comprehensive Assessment System (MCAS), tenth-grade Mathematics score	Boys	1,661 students	0.29 (0.92)	–0.27 (0.97)	0.56	0.56	+21	< .01
MCAS, tenth-grade Mathematics score	Girls	1,968 students	0.41 (0.89)	–0.20 (0.86)	0.62	0.62	+23	< .01
MCAS, tenth-grade Mathematics score	Non-special education students	2,978 students	0.48 (0.86)	–0.10 (0.85)	0.57	0.57	+22	< .01
MCAS, tenth-grade Mathematics score	Students whose baseline MCAS scores were below the median	1,728 students	–0.21 (0.86)	–0.81 (0.75)	0.60	0.60	+23	< .01
MCAS, tenth-grade Mathematics score	Students whose baseline MCAS scores were above the median	1,743 students	0.83 (0.65)	0.30 (0.70)	0.54	0.54	+20	< .01
MCAS, tenth-grade Mathematics score	Students who were eligible for free and reduced-price lunch	2,644 students	0.27 (0.91)	–0.27 (0.89)	0.54	0.54	+21	< .01
MCAS, tenth-grade Mathematics score	Students who were not eligible for free and reduced-price lunch	985 students	0.68 (0.88)	–0.13 (0.96)	0.80	0.80	+29	< .01
General literacy achievement								
SAT Verbal score	Students who took the SAT	2,378 students	442.40 (82.50)	416.40 (91.00)	26.00	0.30	+12	< .05
SAT Writing score	Students who took the SAT	2,378 students	438.90 (80.70)	412.70 (89.30)	26.20	0.31	+12	< .05
MCAS, tenth-grade English language arts (ELA) score	Boys	1,690 students	0.03 (0.79)	–0.44 (0.87)	0.46	0.46	+18	< .01
MCAS, tenth-grade ELA score	Girls	1,995 students	0.21 (0.81)	–0.16 (0.78)	0.37	0.37	+14	< .01
MCAS, tenth-grade ELA score	Non-special education students	3,017 students	0.24 (0.72)	–0.13 (0.73)	0.38	0.38	+15	< .01

MCAS, tenth-grade ELA score	Students whose baseline MCAS scores were below the median	1,771 students	–0.32 (0.81)	–0.76 (0.79)	0.44	0.44	+17	< .01
MCAS, tenth-grade ELA score	Students whose baseline MCAS scores were above the median	1,771 students	0.51 (0.59)	0.14 (0.60)	0.37	0.37	+14	< .01
MCAS, tenth-grade ELA score	Students who were eligible for free and reduced-price lunch	12,687 students	0.07 (0.80)	–0.33 (0.82)	0.40	0.40	+15	< .01
MCAS, tenth-grade ELA score	Students who were not eligible for free and reduced-price lunch	998 students	0.37 (0.81)	–0.17 (0.86)	0.54	0.54	+21	< .01
General academic achievement								
SAT Reasoning score	Students who took the SAT	2,378 students	933.60 (155.60)	855.50 (173.90)	78.10	0.48	+18	< .01
Received a score of 3 or higher on at least one Advanced Placement (AP) exam	Boys	1,720 students	0.25 (0.44)	0.08 (0.26)	0.18	na	na	< .01
Received a score of 3 or higher on at least one AP exam	Girls	1,952 students	0.16 (0.37)	0.09 (0.29)	0.07	na	na	> .10
Received a score of 3 or higher on at least one AP exam	Non-special education students	2,976 students	0.24 (0.43)	0.10 (0.30)	0.15	na	na	< .05
Received a score of 3 or higher on at least one AP exam	Students whose baseline MCAS scores were below the median	1,667 students	0.04 (0.20)	0.01 (0.11)	0.03	na	na	> .10
Received a score of 3 or higher on at least one AP exam	Students who were eligible for free and reduced-price lunch	2,703 students	0.16 (0.37)	0.07 (0.25)	0.09	na	na	> .10
SAT Composite score	Girls who took the SAT	1,401 students	1,368.50 (222.60)	1,263.40 (235.00)	105.10	na	na	< .01
SAT Composite score	Non-special education students who took the SAT	2,047 students	1,402.00 (215.00)	1,296.10 (236.30)	105.90	na	na	< .01
SAT Composite score	Students who took the SAT and were eligible for free and reduced-priced lunch	1,710 students	1,326.00 (217.80)	1,240.60 (224.10)	85.40	na	na	< .05
Completing school								
Repeat the twelfth grade	Students who started the twelfth grade on time	2,415 students	0.19 (0.40)	0.07 (0.26)	0.12	na	na	< .05

Graduated from high school within 4 years	Boys	1,823 students	0.39 (0.49)	0.60 (0.49)	–0.21	na	na	< .05
Graduated from high school within 4 years	Girls	2,097 students	0.67 (0.47)	0.76 (0.43)	–0.09	na	na	< .10
Graduated from high school within 4 years	Non-special education students	3,182 students	0.61 (0.49)	0.71 (0.45)	–0.11	na	na	< .05
Graduated from high school within 4 years	Students whose baseline MCAS scores were below the median	1,771 students	0.37 (0.48)	0.65 (0.48)	–0.28	na	na	< .01
Graduated from high school within 4 years	Students whose baseline MCAS scores were above the median	1,771 students	0.78 (0.41)	0.81 (0.40)	–0.03	na	na	> .10
Graduated from high school within 4 years	Students who were eligible for free and reduced-price lunch	2,876 students	0.54 (0.50)	0.68 (0.47)	–0.14	na	na	< .05
Graduated from high school within 4 years	Students who were not eligible for free and reduced-price lunch	1,044 students	0.54 (0.50)	0.71 (0.46)	–0.16	na	na	> .10
Graduated from high school within 5 years	Boys	1,476 students	0.72 (0.45)	0.72 (0.45)	0.00	na	na	> .10
Graduated from high school within 5 years	Girls	1,732 students	0.82 (0.39)	0.84 (0.36)	–0.02	na	na	> .10
Graduated from high school within 5 years	Non-special education students	2,634 students	0.83 (0.38)	0.81 (0.39)	0.02	na	na	> .10
Graduated from high school within 5 years	Students whose baseline MCAS scores were below the median	1,443 students	0.67 (0.47)	0.79 (0.41)	–0.12	na	na	> .10
Graduated from high school within 5 years	Students who were eligible for free and reduced-price lunch	2,356 students	0.76 (0.43)	0.78 (0.42)	–0.02	na	na	> .10
College preparation								
Took at least one AP test	Boys	1,720 students	0.48 (0.50)	0.21 (0.40)	0.27	na	na	< .01
Took at least one AP test	Girls	1,952 students	0.64 (0.48)	0.34 (0.48)	0.30	na	na	< .01
Took at least one AP test	Non-special education students	2,976 students	0.61 (0.49)	0.32 (0.47)	0.29	na	na	< .01

Took at least one AP test	Students whose baseline MCAS scores were below the median	1,667 students	0.45 (0.50)	0.17 (0.37)	0.28	na	na	< .01
Took at least one AP test	Students who were eligible for free and reduced-price lunch	2,703 students	0.59 (0.49)	0.27 (0.44)	0.32	na	na	< .01
Took the SAT	Boys	1,720 students	0.60 (0.49)	0.55 (0.50)	0.05	na	na	> .10
Took the SAT	Girls	1,952 students	0.82 (0.39)	0.71 (0.45)	0.11	na	na	> .10
Took the SAT	Non-special education students	2,976 students	0.77 (0.42)	0.68 (0.47)	0.10	na	na	> .10
Took the SAT	Students whose baseline MCAS scores were below the median	1,667 students	0.63 (0.48)	0.55 (0.50)	0.07	na	na	> .10
Took the SAT	Students who were eligible for free and reduced-price lunch	2,703 students	0.69 (0.46)	0.62 (0.49)	0.09	na	na	> .10
College enrollment								
Enrolled in any college within 18 months	Students whose baseline MCAS scores were below the median	1,168 students	0.66 (0.47)	0.52 (0.50)	0.14	na	na	> .10
Enrolled in a 4-year college within 18 months	Students whose baseline MCAS scores were below the median	1,168 students	0.53 (0.50)	0.30 (0.46)	0.22	na	na	< .05

Table Notes: The supplemental findings presented in this table are additional findings that meet WWC design standards without reservations, but do not factor into the determination of the study rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on individual outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). Effect size estimates are not available for contrasts where sample sizes for the intervention and comparison groups are unknown. The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding. na = not available.

Study Notes: The *p*-values presented here were reported in the original study. A correction for multiple comparisons was needed for each outcome domain across the supplemental findings in Appendices D.1 and D.2. The study presented impact estimates of key outcomes for the full sample, boys, girls, special education students, non-special education students, students whose baseline MCAS scores were below the median, students whose baseline MCAS scores were above the median, students who were eligible for free and reduced-price lunch, and students who were not eligible for free and reduced-price lunch. Baseline MCAS scores were measured prior to charter application, in either the fourth or eighth grade. The correction for multiple comparisons across the supplemental findings in Appendices D.1 and D.2 resulted in WWC-computed critical *p*-values of .033 for enrollment in a 4-year college within 6 months (Appendix D.1); .007 for repeating the twelfth grade (Appendix D.2), .011 for graduating from high school within 4 years for boys (Appendix D.2), .014 for graduating from high school within 4 years for students who were eligible for free and reduced-price lunch (Appendix D.2), .034 for SAT Composite scores among students who were eligible for free and reduced-price lunch (Appendix D.1), .025 for whether SAT Writing scores were above the median (Appendix D.1); .028 for average SAT Writing scores (Appendix D.1); .031 for whether SAT Verbal scores were above the median (Appendix D.1); .033 for average SAT Verbal scores (Appendix D.1); and .033 for whether AP US History scores were 2 or higher (Appendix D.1). Because the *p*-values for these impact estimates were higher than the WWC-computed critical *p*-values, the WWC does not find the results for these particular outcomes to be statistically significant.

Endnotes

[*] A quick review of this study was released in July 2013, and a single study review was released in March 2015. In April 2016, the WWC revised the study rating in response to an independent review by a quality review team and additional information from the authors, which showed that the study had low attrition. The WWC has determined that the study *meets WWC group design standards without reservations*. This report replaces the initial assessment in the quick review and the prior single study review. The reported analyses in this single study review are only for those eligible outcomes that *meet WWC group design standards without reservations*, and do not necessarily apply to all results presented in the study.

[1] Angrist, J. D., Cohodes, S. R., Dynarski, S. M., Pathak, P. A., & Walters, C. R. (2014). *Stand and deliver: Effects of Boston's charter high schools on college preparation, entry, and choice*. Cambridge, MA: Department of Economics Working Paper, Massachusetts Institute of Technology. The authors also provided an updated version of the paper that has been accepted for publication in the *Journal of Labor Economics* (2016). This review used sample sizes from the updated version of the paper. An earlier version was used to complete a quick review of this study in July 2013: Angrist, J. A., Cohodes, S. R., Dynarski, S. M., Pathak, P. A., & Walters, C. D. (2013). *Charter schools and the road to college readiness: The effects on college preparation, attendance and choice*. Boston, MA: The Boston Foundation and NewSchools Venture Fund.

[2] Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the authors) to determine the study rating. The review reports the WWC's assessment of whether the study *meets WWC group design standards without reservations* and summarizes the study findings following WWC conventions for reporting evidence on effectiveness. This study was reviewed using the single study review protocol (version 2.0).

[3] The CACE is the average effect of taking up the intervention among compliers—those who would take up the intervention if assigned to the intervention group and who would not take up the intervention if assigned to the comparison group. In some disciplines, the CACE is also referred to as the local average treatment effect (LATE). The CACE is a different measure of program effectiveness than the intent-to-treat (ITT) effect, which provides an estimate of the offer of the intervention. In studies that present both a CACE estimate and an ITT estimate, the WWC will review both estimates. The authors did not analyze ITT estimates and therefore, only the CACE estimates are reviewed in this report. More information on the WWC guidance to review CACE estimates is available here: http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=260

[4] There were 21 outcomes included in the study that are not described in this WWC report. These outcomes are related to college enrollment within 18 months of the projected high school graduation and college persistence (enrolled for three or five semesters) within 18, 30, and 42 months of the projected high school graduation, by college type (any college, 2-year college, 4-year college, 4-year public college, 4-year private college, 4-year Massachusetts public college). See the table notes in Appendix B for more information. Additionally, 34 subgroup contrasts included in the study are not described in this WWC report. These subgroup estimates are not included in the SSR because lottery outcomes were not determined to be a sufficiently strong predictor of charter school enrollment (i.e., these outcomes did not meet the WWC's CACE standards for sufficient instrument strength); consequently, these excluded analyses *do not meet WWC group design standards*. See the table notes in Appendix B for more information.

[5] Compliers are defined based on what their behavior would be if they were assigned to the intervention group and what their behavior would be if they were assigned to the comparison group. Because students can only be assigned to one research group, student behavior under the other research group is not known. Specifically, the compliers in the intervention group cannot be distinguished from individuals who would always take up the intervention (known as always-takers) and similarly, the compliers in the comparison group cannot be distinguished from individuals who would never take up the intervention (known as never-takers). CACEs are typically estimated with instrumental variable (IV) estimators that use study participants' randomly assigned status as an instrument for their take-up of the intervention. The IV estimator uses the portion of the variation in take-up that is induced by the random assignment process to estimate the impacts of taking up the intervention on outcomes. Only CACE estimates where the relationship between the instrument and take-up is sufficiently strong for the analytic samples can *meet WWC group design standards without reservations*.

[6] Thernstrom, A. M., & Thernstrom, S. (2003). *No excuses: Closing the racial gap in learning.* New York: Simon & Schuster.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2016, April). *WWC review of the report: Stand and deliver: Effects of Boston's charter high schools on college preparation, entry, and choice.* Retrieved from http://whatworks.ed.gov

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.

Clustering adjustment If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.

Confounding factor A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

Design The design of a study is the method by which intervention and comparison groups were assigned.

Domain A domain is a group of closely related outcomes.

Effect size The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

Eligibility A study is eligible for review if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

Equivalence A demonstration that the analytic sample groups are similar on observed characteristics defined in the review area protocol.

Improvement index Along a percentile distribution of individuals, the improvement index represents the gain or loss of the average individual due to the intervention. As the average individual starts at the 50th percentile, the measure ranges from –50 to +50.

Multiple comparison adjustment When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.

Quasi-experimental design (QED) A quasi-experimental design (QED) is a research design in which study participants are assigned to intervention and comparison groups through a process that is not random.

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which eligible study participants are randomly assigned to intervention and comparison groups.

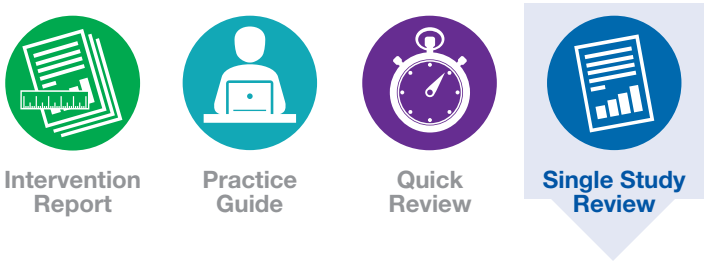
Single-case design (SCD) A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

Standard deviation The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample are spread out over a large range of values.

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < .05$).

Substantively important A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the WWC Procedures and Standards Handbook (version 3.0) for additional details.



A **single study review** of an individual study includes the WWC's assessment of the quality of the research design and technical details about the study's design and findings.

This single study review was prepared for the WWC by Mathematica Policy Research under contract ED-IES-13-C-0010.