Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 24-1087

Effects of Early College on Educational Attainment for All in Massachusetts

Pierre M. Lucien
University of Oxford

Ariel Lindorff
University of Oxford

Steve Strand
University of Oxford

Evaluations of Early College, a type of intervention that enables simultaneous enrollment in secondary and post-secondary courses in the United States, consistently find positive effects on educational attainment across racial and socioeconomic groups. Unlike Early College initiatives in other states, Massachusetts launched Early College in Fall 2018, enabling a within-school as well as a whole-school intervention, in which each participating school may enroll some or all of its students in Early College, with guiding principles of equitable access, guided academic pathways, student support, and connection to career. This study uses propensity score matching to evaluate the impact of participating in Massachusetts Early College on students' educational attainment. Positive effects found on college enrollment, with statistically significant positive interactions between the treatment and being socioeconomically disadvantaged, and on college persistence, with statistically significant positive interactions between the treatment and being Latinx, suggest the intervention may help promote equitable access to higher education in Massachusetts. Massachusetts is a local-control state, where public school governance is legally delegated to district and school boards located in the communities they serve, as opposed to the state government, making it more difficult to have state-wide interventions. The flexibility of Massachusetts Early College renders it more easily replicated in local-control states, than the whole school models previously studied.

VERSION: November 2024

Suggested citation: Lucien, Pierre, Ariel Lindorff, and Steve Strand. (2024). Effects of Early College on Educational Attainment for All in Massachusetts. (EdWorkingPaper: 24 -1087). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/95rt-8k07

Effects of Early College on Educational Attainment for All in Massachusetts

Pierre M. Lucien
University of Oxford

Ariel Lindorff
University of Oxford

Steve Strand
University of Oxford

October 2024

Abstract

Evaluations of Early College, a type of intervention that enables simultaneous enrollment in secondary and post-secondary courses in the United States, consistently find positive effects on educational attainment across racial and socioeconomic groups. Unlike Early College initiatives in other states, Massachusetts launched Early College in Fall 2018, enabling a within-school as well as a whole-school intervention, in which each participating school may enroll some or all of its students in Early College, with guiding principles of equitable access, guided academic pathways, student support, and connection to career. This study uses propensity score matching to evaluate the impact of participating in Massachusetts Early College on students' educational attainment. Positive effects found on college enrollment, with statistically significant positive interactions between the treatment and being socioeconomically disadvantaged, and on college persistence, with statistically significant positive interactions between the treatment and being Latinx, suggest the intervention may help promote equitable access to higher education in Massachusetts. Massachusetts is a local-control state, where public school governance is legally delegated to district and school boards located in the communities they serve, as opposed to the state government, making it more difficult to have state-wide interventions. The flexibility of Massachusetts Early College renders it more easily replicated in local-control states, than the whole school models previously studied.

Introduction

Although contemporary education systems in the United States often claim to be founded on the notion that all people are educable and deserve an opportunity to learn (Domina et al., 2017), there is a long history of inequity in education access, quality, and attainment based on race and socioeconomic status in the US (Gamoran, 2015). These educational inequities are not relics of the past but remain. For example, bachelor's degree attainment rates for Black and Latinx adults in the US are 14 and 21 percentage points lower than the rates for White adults, respectively (U.S. Census Bureau, 2023). Likewise, 14 percent of students from low socioeconomic backgrounds and 29 percent of students from middle socioeconomic backgrounds attained a bachelor's degree within 8 years of high school completion, compared to 60 percent of high socioeconomic background students (National Center for Education Statistics, 2015). Furthermore, first-generation college students were found to be half as likely to obtain a degree as students whose parents attended college (Redford & Hoyer, 2017). These inequities persist not only temporally but also geographically, as they are seen throughout the country, including the setting of this study, Massachusetts.

Massachusetts ranks at or among the top US states, when comparing primary, secondary, and tertiary education outcomes, on average. Looking at the results from the National Assessment of Educational Progress (NAEP)—the largest nationally representative and continuing assessment of student performance in core subjects (National Center for Education Statistics, 2018)— Massachusetts ranked first, second or third in 4th and 8th grade reading and math scores in 2022, the most recent year in which those data were collected (The Nation's Report Card, 2023). Likewise, Massachusetts has higher shares of residents over the age of 25 who have associate degrees or higher, and who have bachelor's degrees or higher, than any other US state (U.S. Census Bureau, 2021).

However, when the data is disaggregated, great variations are apparent in achievement and attainment by race and socioeconomic status. In Massachusetts, Black and Latinx students averaged 21 and 23 points lower than White students, respectively, on the 2022 8th grade NAEP reading tests, while students who were eligible for the National School Lunch Program averaged 26 points lower than students who were not eligible (National Center for Education Statistics, 2022). Furthermore, only 31 percent of Black students and 36 percent of Latinx students meet college-readiness benchmarks in reading and math compared to 65 percent of White students in Massachusetts (The College Board, 2022). At the college level, inequity persists in Massachusetts, and there is variation in college readiness, access, and success by race (Massachusetts Department of Higher Education, 2016); the Latinx/White Gap in college enrollment (-21 percentage points) is nearly twice that of the national average (-11 percentage points). Performance gaps between White students and Black and Latinx students in post-secondary math and reading in Massachusetts are also higher than the national average. Six years after enrolling at a public college or university, less than 33 percent of Black and Latinx students have earned a degree (Massachusetts Department of Higher Education, 2016).

Higher educational attainment has been consistently found to predict higher earnings, better health, and civic engagement, among other positive outcomes (Petronijevic & Oreopoulos, 2013; Lobo & Burke-Smalley, 2018; Hummer and Lariscy, 2011; Kubota et al., 2017; Levin-Waldman, 2013). As state demographics shift, with the share of the population that are Black or Latinx increasing (Massachusetts Education Equity Partnership, 2018), promoting educational attainment for all is an increasingly urgent matter for Massachusetts, so that it may remain a leading state educationally and economically. In response, the Massachusetts Department of Elementary and Secondary Education (DESE) and the Massachusetts Department of Higher Education (DHE) launched Early College in Autumn 2018 to promote equitable educational attainment.

This study seeks to investigate the effectiveness of the Early College program in promoting educational attainment for all in the state of Massachusetts. More specifically, the research questions are:

1. What is the effect of participating in Early College on students' likelihood of being college ready, where college readiness is approximated by the completion of MassCore (a recommended program of study that aligns high school coursework with college and workforce expectations)?
2. What is the effect of participating in Early College on students' likelihood of enrolling in college within 6 months of on-time secondary school completion?
3. What is the effect of participating in Early College on students' likelihood of persisting to a second year in college, after having enrolled within 6 months of on-time secondary school completion?
4. Do these effects vary by racial and economic background?

Although effects of Early College participation on earning a college degree is of high interest, the data will not be available in time for this study, so this will be investigated in future research.

Literature Review

Effects of Early College on College Access and Success

With the persistent and pervasive disparities in educational attainment by economic status and race in the United States, school districts, state education agencies, and charter management networks continue to explore policies and practices that may close the gaps. Early College is one such intervention. Although descriptive and qualitative studies have been conducted on Early College in other states, much of the impact evaluation of Early College is limited to two series of evaluations by the American Institutes for Research (AIR) and the University of North Carolina's SERVE Center, discussed below.

AIR conducted a longitudinal randomized controlled trial to evaluate the impacts of Early College with a sample of 2,458 students from 10 schools, in five states, that used lotteries in their admissions processes. By comparing the outcomes across the groups offered and not offered enrollment, they estimated the effects of program participation over the course of 3 studies. Regarding college readiness, Berger et al. (2013) found that Early College participants were 5 percentage points more likely to graduate high school. In terms of college access, Haxton et al. (2016) found that being admitted to an Early College high school made a 9-percentage point positive difference in college enrollment, and this effect did not differ significantly by demographic characteristics including racial and socioeconomic background. For college success, Song and Zeiser (2019) found that 6 years after expected high school graduation, Early College participants were more likely than their non-Early College counterparts to earn a certificate, associate degree or a bachelor's degree by 11.9 percent.

Using data from 12 Early College high schools in the state of North Carolina, The SERVE Center likewise conducted a longitudinal randomized controlled trial by comparing outcomes for lottery applicants who were offered enrollment with the outcomes for lottery applicants who were not offered enrollment and attended high school elsewhere. For the first SERVE center study, Edmunds et al. (2017) found that six years after enrolling in the Early College program in grade 9, Early College students enrolled in college at a rate 15.6 percentage points higher than students in the control group. Program participation made slightly less of a positive difference for students underrepresented in higher education (14.3 percent). In a follow-up study, Edmunds et al. (2020) found that 6 years after grade 12 -- when students generally graduate high school -- 44.3 percent of the Early College participants, compared to 33 percent of the control group, earned a post-secondary credential or degree.

Consistent with experimental literature, a non-experimental study also suggests that Early College impacts students of all racial backgrounds positively. Britton et al. (2022) used survival analysis to measure the relationship between demographic factors and college persistence of Early College participants, while controlling for academic and behavioral risk factors. They found that participating Black and Latinx students are as likely as White students to persist in college, when baseline achievement and attendance are controlled. Thus, we argue that participating in the Early College intervention has the potential to promote equity in educational attainment, if there is a higher concentration of students underrepresented in higher education enrolled, relative to the rest of the state, as is the case in Massachusetts.

Theory of Change

College course taking in high school, a prominent feature of the Massachusetts Early College program, is indeed a way to promote educational attainment. Shields et al (2021) found that earning college credits in high school has a large, positive, statistically significant effect on the probability of graduating from high school, immediately enrolling in college, and avoiding enrollment in remedial courses during the first year of college in Massachusetts' neighboring state of Rhode Island. Altogether, the literature suggests that cutting costs for families, better preparing students through rigorous courses, helping them gain momentum via credit accumulation, giving students favorable perception in the admissions process, and exposing students to norms and expectations early are the mechanisms by which taking college courses in high school, such as through the Early College program, promotes post-secondary enrollment, persistence, and completion.

Even after entry into an institution of higher learning, high costs prevent college students from completing their degrees. An estimated 1.6 million bachelor's degrees were not earned among college-qualified high school graduates with socioeconomically disadvantaged backgrounds in the 1990's due to costs (Advisory Committee on Student Financial Assistance, 2006). It is within reason then for those who wish to promote college enrollment and completion for said students to cut the costs. Among many others, one way to do that is to offer college credit-bearing courses at no additional costs to families when students are in high school. The Center for College Affordability and Productivity (2010) found that the total cost of a college education could potentially be reduced by up to 12.5% by offering a program like Early College.

Moreover, students exposed to rigorous curricula during their secondary schooling are better prepared to succeed at post-secondary institutions, due to the increased effort demanded (Shireman, 2004). Shireman (2004) further explains that "in the high schools that serve low-income and minority youth, there often is little attention to rigorous, standards-based instruction" (p.3). Using data from the National Education Longitudinal Study, which randomly sampled nearly 25,000 students from about 1,000 nationally representative schools, Akerhielm et al. (1998) found the relationship between taking advanced courses and pursuing higher education to be greater for students of low-income backgrounds. From that same survey, Stearns et al. (2010) found that Black and Latinx students benefit most from taking more rigorous academic courses, relative to other racial groups. Culver et al (2019) found that students who take rigorous courses—which they define as courses that promote a higher-order understanding of concepts and skills, as opposed to just more work—are more likely to think critically, value skills and ideas from the discipline, and appreciate the challenge of the course content. Culver et al (2019) further assert that in-class rigor especially benefits the critical thinking skills of students underrepresented in higher education. Course rigor is thus the principal theory of change.

Barnett (2016) draws from various empirical papers to conclude that college level courses taken in high school are particularly effective at promoting college enrollment and completion, as they serve as preparation for postsecondary success. For example, Karp et al. (2008) ran OLS and logistic regressions, controlling for observable preexisting characteristics, to conclude that students who dually enroll in high school and college courses are more likely to graduate from high school, attend college, and persist. Using the National Education Longitudinal Study, An (2013) found dual enrollment participation to increase the probability of attaining any postsecondary degree and a bachelor's degree by 8 percentage points and 7 percentage points, respectively. Credit accumulation in high school is valuable in and of itself. Adelman's (2006) series of logistic regressions found that there is a minimum number of credits earned during the first term and first year below which completion becomes less likely. In other words, college credit accumulation in high school gives students momentum to get to and through college (Leinbach & Jenkins, 2008).

Furthermore, Klopfenstein and Thomas (2009) found from their literature review that advanced course taking in secondary school, including college level courses, is perceived by college admissions officers as an indicator of student motivation. Consequently, colleges use advanced course enrollment and performance as a proxy for college readiness, viewing said students favorably in the admissions process (Kretchmar & Farmer, 2013). The last mechanism by which taking college courses in high school promotes post-secondary enrollment, persistence, and completion is by exposing students to the cultural aspects of college prior to their enrollment. Per Smith & Wertlieb (2005), introducing students to norms and expectations that they will encounter in postsecondary institutions facilitates their transition from high school.

Theoretical Framework

This study is framed by Human Capital Theory, which posits that any deliberate acquisition of useful skills and knowledge that yields economic value, including direct expenditures on education, is an investment (Schutz, 1961; Becker, 1962). Given the economic returns of educational attainment (Lobo & Burke-Smalley, 2018) and the effects of Early College on educational attainment (Edmunds et al., 2017; Haxton et al., 2016) this paper argues that Massachusetts made a human capital investment by offering the Early College, so that its citizens may acquire useful knowledge and skills that will produce an economic return. Determining the effects of Early College on college readiness, access, and success is analogous to a return-on-investment analysis. Schultz (1961) further attributes differences in earnings by race to differences in human capital investment by racial groups. By targeting Black and Latinx students in recruitment for Early College, Massachusetts also seeks to address this supposed racial gap in human capital investment. Interactions between race and program participation included in the paper will help determine their effectiveness on that front.

Additionally, this study is structured by Perna and Thomas' (2006) Framework for Reducing the College Success Gap and Promoting Success for All. According to Perna and Thomas (2006), the four key transitions in a success process are college readiness, college enrollment, college achievement, and post-college attainment. Furthermore, there are four layers of context to understanding student success—internal, family, school, and sociopolitical. The framework maps on to this research project in direct and substantive ways. Specifically, the outcomes of interest—college readiness as approximated by MassCore completion, college enrollment, and college persistence—align with three of the four transitions in Perna and Thomas' (2006) student success process. Covariates for regression models that estimate the propensity scores, the probability of receiving treatment assignment, as well as the post-matching regression models that estimate effects of program participation correspond to Perna and Thomas' (2006) layers of context.

Intervention Description

The increasing use of the term Early College in American education in the last twenty years or so can be traced back and attributed to the 2002 Bill and Melinda Gates Early College High School Initiative, which aimed to promote college access for students historically underrepresented in higher education by enabling them to earn a high school diploma while simultaneously building credits toward a post-secondary credential (Berger et al., 2010). High schools in the United States have offered dual enrollment, or the opportunity to simultaneously take college courses, for many decades, but these opportunities were typically reserved for academically advanced or gifted students (Berger et al., 2010).

Given the above-noted inequities in an otherwise high performing state, Massachusetts launched its Early College initiative in 2018 to promote equitable access to higher education. Unlike Early College in other states, Massachusetts Early College can involve participation from only some students in participating schools, or the whole school. Additionally, schools offering the program are asked to present a plan for outreach and recruitment specifically for racial minority and socioeconomically disadvantaged students (DESE & DHE, 2017). Furthermore, programs must demonstrate that students are not excluded based on prior academic performance, that students will be onboarded to ensure they ready for the coursework, and that students will be provided ongoing support to help them succeed in the courses they take (DESE & DHE, 2017). These features distinguish Massachusetts Early College from other Early College interventions evaluated in the existing literature.

Participating schools must partner with a local higher education institution (henceforth referred to as “college”) to offer college-level courses. Per the Early College designation criteria, which determine which high school and college partnerships get to offer Early College, participating in Massachusetts Early College means that a student has the option to enroll in at least 12 credits by high school graduation, starting as early as 9th grade (DESE & DHE, 2017). 12 credits are approximately four college-level courses (typically 3 credits), and one tenth (1/10) of the typical 120-credit requirement to graduate with a bachelor’s degree in the United States. These 12 credits must all be taken in a pathway, meaning that the 4 courses must feed into a subject-specific concentration, and should ideally be sequential so that the student is progressing toward a degree at the partner institution of higher education (DESE & DHE, 2017). Furthermore, courses must be transferable to public college or university in the state.

Courses are offered during the school day to minimize the logistical burden on students and their families. Furthermore, the high school and college partners must collaborate to schedule college courses such that they present negligible to no interference with students’ high school experience. Courses may be offered at the high school or at the college. In instances where a course is offered at the college, transportation must be provided for students and their departure and arrival must fit within their high school schedule. In all instances, instructors must be qualified to teach college courses (DESE & DHE, 2017). In addition to taking 12 college credits toward a degree, student participants receive academic advising from an Early College counselor at the high school. This is a critical component of the initiative, as program participants are likely to be those underrepresented in higher education, including first generation collegegoers, and as such are less likely to receive similar guidance at home (DESE & DHE, 2017).

Lastly, the high school and college partnerships that offer Early College must have in place adequate student support. This may include “on-ramping”, which involves provision of preparatory courses in advance as well as scaffolding services (e.g. tutoring) to ensure that students have an opportunity to fully participate and benefit from Early College (DESE & DHE, 2017). The designation criteria provided by the state, which serve as the intervention description for this study, leave room for variation across schools and thus introduces some uncertainty. It would be useful to be able to categorize and measure aspects of implementation across participating schools, such as details of college advising student support offerings, but data on these aspects are not readily available and so are beyond the scope of this study’s secondary analysis.

It should be noted that variation across schools in their implementation in terms of what proportion of students participate, what courses and academic pathways are offered, what recruitment strategies are used, types and dosage of supports provided, and profiles of the college or university partners—complicate the evaluation of the initiative as a whole but were put in place by design: “Flexibility will allow for learning during the first years of this designation process. While each element of the designation criteria was developed with intention and purpose, we understand that parts of this process will evolve, and it is likely that we will revisit and revise the criteria” (DESE & DHE, 2017, p.6).

Data and Methods

Data Sources

This study draws from four sources of administrative data that capture information about students enrolled in the Massachusetts public secondary school system—the Student Information Management System (SIMS), the Massachusetts Comprehensive Assessment System (MCAS), the School Safety and

Discipline Report (SSDR), and the National Student Clearinghouse (NSC) Student Tracker. Per the Massachusetts DESE website (2023), SIMS is a student-level data system that enables the collection and analysis of information to meet federal and state reporting requirements, and to inform decisions making, policy and practice. SIMS contains fifty-two data elements, or variables, which detail each student's demographic profile as well as their enrollment and attendance. MCAS is a series of statewide assessments that help parents, students, educators, and policymakers determine where districts, schools, and students are meeting expectations and where they need additional support (DESE, 2023). SSDR collects student removals from class or school, including suspensions and expulsions (DESE, 2023).

NSC is a nonprofit organization that provides reporting, verification, and research services to educational institutions in the United States. The NSC network comprises nearly all colleges and universities in the United States, representing 97 percent of postsecondary education enrollment in the country (Dunbar & Shapiro, 2016). NSC provides a service called the Student Tracker for High Schools, which enables all secondary schools in the United States to follow their graduates' transition to college by querying participating institutions' postsecondary enrollment and degree records (Dunbar & Shapiro, 2016). This is the data that is used to determine whether graduates from Massachusetts public secondary schools enrolled and persisted in college.

Sample Selection

Early College is a school level designation, meaning that the state designates which schools may offer the program. Even though student selection into the program is informed by the same guiding principles throughout the state, the specific processes are not uniform across schools. Per the Designation Criteria, participating schools must demonstrate evidence of building "strong and efficient pathways for student groups who may have been traditionally underserved in higher education" (DESE & DHE, 2017, p.7). More specifically, schools had to present an effective plan for outreach and recruitment of students who are historically underrepresented in higher education, including students of color and socioeconomically disadvantaged students (DESE & DHE, 2017) to receive designation. To meet those ends, the schools are further asked to provide "tuition-free participation, open enrollment without regard to prior academic performance, student supports to promote success, multiple entry points for students, and student supports to prepare students for entry into the program" (DESE & DHE, 2017, p. 9). If they operate within those guidelines, schools can select students however else they think best and retain their designation. Furthermore, there is "no minimum initial cohort size requirement" (DESE, 2017, p.7), so a school may enroll their entire student body in the program or very few students. For example, 42 Massachusetts high schools offered EC to 12th graders in 2022. In terms of count of 12th grade EC enrollment in the 42 schools, the range was from 1 student to 226 students, with a mean of 49 students. In terms of share of the school enrolled in EC, the range was from 1.5% to 100% of the 12th graders in the school, when rounded, with a mean of 21%.

It should be noted that there were two interventions launched as part of the Massachusetts High Quality College and Career Pathway programs in 2018, Early College and Innovation Career Pathways. During the years of the study, the two interventions are mostly offered at different schools, but there is some overlap, which results in 243 Innovation Career Pathways students attending schools that also offer Early College. Those 243 students were excluded from the sample. Therefore, the treated students used in this study are the 5,177 Early College participants in grade 12 from the 2018-19 through the 2021-22 school years. The untreated comparison students are drawn from the 27,577 remaining 12th grade students who attend a school that offers Early College but do not participate in the program during that same time (Table 1), minus the 243 who participate in Innovation Career Pathways. The analysis starts with the 2018-19 school year because it is the first year for which there are grade 12 students in the programs. From the matching process used in this study, described in greater detail in the upcoming methods section, 84 percent of the treatment observations were kept, while 16 percent were pruned, due to not having a sufficiently approximate match in terms of maximum distance from their propensity score in standard deviations. This means that the final sample included 4,358 Early College participants and 4,358 untreated comparison students, totaling 8,716 students.

Program Participation

There were 44 unique high schools in the state of Massachusetts that offered the Early College program to 12th graders at any point during the 2018-19 to the 2021-22 academic years. There were 19 schools that offered the program in 2018-19, the first year in the study, which increased to 42 by 2021-22, the last year in the study, due to additional programs being designated each year and a couple ceasing to offer the program. There were 28 high schools participating in 2019-20 and 31 in 2020-21. There were 361 Early College participants in the 12th grade in 2018-19, making up 7.2 percent of 12th graders at those schools. In 2019-20, there were 1,144 Early College participants in the 12th grade, and they made up 15.9 percent of 12th graders in the Early College offering schools. In 2020-21, there were 1,702 Early College participants in the 12th grade, with a school participation rate of 18.6 percent. Lastly, there were 1,970 Early College participants in the 12th grade in 2021-22, with a participation rate of 17.3 percent. On aggregate, participants made up 15.8 percent of 12th graders in their schools and 1.8 percent of 12th graders in the state during those 4 years (Table 1).

Table 1. Early College Participation in 12th Grade

	Student Group			
School Year	Early College	Rest of School	Rest of State	Total
2019	361	4649	66975	71985
2020	1144	6049	64470	71663
2021	1702	7437	63423	72562
2022	1970	9442	60609	72021
Total	5177	27577	255477	288231

Early College is typically offered at schools with higher rates of Black, Latinx, and socioeconomically disadvantaged students than the state. There are a small number of schools where this is not the case, but they focus on serving those groups. Mutually exclusively, 59 percent of 5,177 grade 12 Early College participants were Black or Latinx from the 2018-19 to 2021-22 academic years, while 57 percent of their 27, 577 school peers, and 24 percent of their 285,250 state peers identified similarly. Nearly 50 percent of 12th grade Early College participants were from socioeconomically disadvantaged families, compared to 56 percent of the rest of their schools, and 28 percent of the rest of the state. Demographically, participation in the program within schools offering Early College seems broadly representative of those schools (Table 2).

Table 2. Participation of Underrepresented Students in EC

	Percent Black	Percent Latinx	Percent Low Income
Early College	16%	43%	49%
Rest of School	11%	46%	56%
Rest of State	9%	16%	28%

However, Early College participants are more likely to be proficient or advanced in English and Math than their school peers. Looking at the 8th grade standardized test scores for the 12th graders from the

2018-19 to the 2021-22 academic years, 45 percent of Early College participants were proficient or advanced in Math, compared to 28 percent of their school peers and 48 percent of their state peers. 53 percent of Early College participants were proficient or advanced in English Language Arts (ELA), compared to 38 percent of their school peers and 63 percent of their state peers. In terms of baseline achievement, participation in the program within schools is more reflective of the state than of the schools.

Table 3. MCAS Performance Levels

	Proficient or meeting expectations in Math	Proficient or meeting expectations in ELA	Advanced or exceeding expectations in Math	Advanced or exceeding expectations in ELA
Early College	33%	43%	12%	10%
Rest of School	21%	32%	7%	6%
Rest of State	31%	48%	17%	15%

Dependent Variables

Given the outcomes of interest—college readiness, college access, and college success—the dependent variables are MassCore completion, college enrollment within six months of graduating high school, and persistence to a second year in college. MassCore is a recommended program of study that aligns high school coursework with college and workforce expectations. This entails performing satisfactorily in four years of English, four years of mathematics, three years of a lab-based science, three years of history, two years of the same world language, and one year of the arts (DESE, 2022). For this study, the completion of MassCore serves as a proxy for college readiness. It should be noted that the criteria for meeting MassCore reflects the national literature on college readiness. For example, DeAngelo and Franke (2016, p.1596) define college ready students as having "completed four years of English, three years of math, two years of a foreign language, one year each of biological and physical sciences, plus an additional year of one or the other (in total three years of science), one year of history/government, and one year of art" in high school. For all four cohorts included in this study, the MassCore completion information is available in the Student Information Management System (SIMS) while the college enrollment information is available in the detailed National Student Clearinghouse (NSC) report. College success, defined as second year persistence in this study, is also drawn from the annual college enrollment data from NSC. Second year persistence is an intermediate measure, as college completion is the ultimate measure of college success, however such data is not available in time to be included.

Independent Variables

The primary independent variable is program participation, for which there is an indicator variable in SIMS. Additionally, a correct specification of covariates is crucial to propensity score calculation. In education, covariates representing students' prior achievement, gender, race, disability status, and socioeconomic background, can play a role in determining selection into interventions and influence the outcomes of interest (Fisher, 2019). As such, the Student Information Management System (SIMS) covariates that will be used in this study are race, gender, disability status and socioeconomically disadvantaged status, as well as English learner status.

Race is a seven-value nominal variable indicating whether a student is Asian, Black, Latinx, Multiracial, Native American, Pacific Islander, or White. For the purposes of this study, which investigates the effect of program participation overall and particularly for Black and Latinx students, racial groups of Asian, Multiracial, Pacific Islander, and Native American were collapsed as Other Race. Black, Latinx, and Other

Race are compared against White as the reference group. Findings for Other Race are not substantively meaningful, as the small groups collapsed may have quite different historical patterns of performance, but collapsing was a statistical necessity, as the individual groups were too small to draw stable inferences. It should also be noted that the term racial minority students in this paper refers to students who identified as Black or Latinx.

Gender is a categorical variable indicating if a student is male, female, or non-binary. Although nonbinary is a gender option, it is not present the sample, thus gender is operationally a binary variable in this study. Socioeconomically disadvantaged status is a binary variable indicating whether a student's "household is eligible for free or reduced-price lunch, or if they receive Transitional Aid to Needy Families benefits" (DESE, 2023, p.26). Disability status is a binary variable indicating when a student meets criteria to receive disability support services and accommodations. English learner students are defined as "children who have indicated a language other than English on the Home Language Survey, are less than proficient on an English language proficiency assessment and are unable to perform ordinary classroom work in English" (DESE, 2023, p.32).

After retrieving demographic and program participation information, the data was merged with math and reading standardized test scores from grade 8. Grade 8 scores are being used in lieu of grade 10 scores because some Early College programs start preparatory activities in grade 9, which disqualifies grade 10 test scores as baseline. Massachusetts administered 3 different standardized tests during the years subjects of the study were in grade 8—Partnership for Assessment of Readiness for College and Careers (PARCC), Legacy MCAS, and Next Generation MCAS—all of which were graded on different scales. While the test scores are on different scales, the test results all come with four performance levels—failing, needs improvement, proficient, and advanced for PARCC and Legacy MCAS; and not meeting expectations, partially meeting expectations, meeting expectations, and exceeding expectations for Next Generation MCAS. Performance levels across tests were numerically coded one through four, with missingness given the value of 0, to be included as five factor variables for the calculations of propensity scores.

Lastly, Chronically Absent in Grade 8 and Suspended in Grade 8 were added as baseline measures. Chronically absent in Grade 8 is a binary variable indicating whether a student missed 10% or more school days in the 8th grade. For each student, the number of days they enrolled in and the number of days they attended school on an annual basis, are included in SIMS. By subtracting days attended from days enrolled, one may calculate how many days a student was absent each year. Dividing days absent by days enrolled yields the absence rate. Any student whose absence rate was equal to or greater than 0.1, or 10%, during the 8th grade year was given the value of 1 in the binary indicator Chronically Absent in Grade 8. Suspended in Grade 8 is a binary variable indicating whether a student was suspended at any time in the 8th Grade. Suspension data for all Massachusetts public schools is stored in the School Safety and Discipline Report (SSDR). SSDR collects information on all rule infringements that occur on school property and the corresponding disciplinary action taken against the student offender, including suspensions. Any student with a value greater than 0 for the days missed due to suspension during the 8th grade year was given the value of 1 in the binary indicator Suspended in Grade 8.

Methods

This paper gauges the effects of the Early College program on participants using propensity score matching. The propensity score is the conditional probability of being assigned a treatment given a vector of observed covariates (Rosenbaum & Rubin, 1983). Matching means that each member of the treatment group is then paired with a non-treated person in the population with an equal or proximate propensity score. When the matching process is successful, the resulting data set will be balanced, whereas the treatment group would be matched to a comparison group that is similar on observable characteristics. The average difference between the balanced treatment and control units from matched pairs with the same propensity scores is an estimate of the average treatment effect on the treated, or ATT (Guo & Fraser, 2015).

This research project chooses to match for several different reasons. Because regression analyses can be conducted post-match, matching enables one to leverage the advantages of both matching and regression analysis. Stuart (2010) asserts that matching should not be seen in conflict with regression modeling, rather the two methods are complementary and are best when combined. This study does indeed conduct regression analyses after matching. More importantly, matching methods address areas of the covariate distribution where there is not sufficient overlap between the treatment and control groups (Stuart, 2010). Imbalance between treatment and control groups could otherwise be described as systematic differences, which render a researcher less able to make causal claims. Indeed, Littnerova et al (2013) argue that the benefit of propensity score matching over multivariable adjustment is "the separation of confounding factors adjustment and analysis of the treatment effect steps" (p. 384).

It should be noted that the propensity score calculation introduced by Rosenbaum and Rubin (1983) may employ various analytic approaches. Upon calculating propensity scores, a researcher may use propensity score matching, stratification on the propensity score, inverse probability of treatment weighting, and covariate adjustment using the propensity score (Guo & Fraser, 2015). Because matching eliminates a greater proportion of systematic differences in baseline characteristics between treatment and comparison groups than the three other methods (Austin, 2011) and the Rosenbaum Bounds Sensitivity Analysis, discussed later in the paper (Table 6), determined that the matching model used is not highly sensitive to an unobserved confounder (DiPrete & Gangl, 2004), matching was preferred.

The prevailing approach to calculating propensity scores in education research, and social sciences broadly, is logistic regression (Guo & Fraser, 2015), with treatment assignment as the binary outcome and a set of pertinent covariates as predictors (D'Agostino, 1998). From the logistic regression model, predicted probabilities for being assigned to the treatment group, referred to as propensity scores, are generated for both the treatment group and potential comparison group members. The step following the calculation of propensity scores is to match treated participants to non-treated ones of a similar propensity score, as an approximation of comparability. After the logistic regression—which had Early College participation as the outcome variable—and the subsequent matching process, a sample including Early College students and their corresponding untreated match was generated.

This study matches using nearest neighbor with a caliper, as opposed to other approaches, such as Mahalanobis metric matching and optimal matching. Nearest neighbor is understood as pairing each treatment participant to a control participant with the nearest propensity score (Stone & Tang, 2013). When the nearest neighbor is far, the treated may be matched with a control that is quite different, in terms of propensity score. To overcome shortcomings of erroneously choosing a distant match, a researcher may impose a prespecified maximum distance between propensity scores, known as a caliper (Guo & Fraser, 2015). Although Rosenbaum and Rubin (1983) recommend a caliper that is .25 standard deviations of the logit of the propensity score, the reduced one of .2 standard deviations was selected, to minimize the mean square error of the resultant estimated treatment effect (Austin, 2011).

Nearest neighbor matching with a caliper is particularly favorable for this study because it "permits subsequent multivariate analysis of almost any kind that allows researchers to evaluate causal effects as they do with randomized experiments" (Guo & Frazer, 2015, p. 148). More specifically, a researcher may "perform multivariate analyses and undertake covariate adjustments for matched sample as is done in randomized experiments," (Guo & Frazer, 2015, p. 153) to estimate average treatment effects. It should also be noted that this study matches one-to-one, as opposed to many-to-one. That is one comparison student per program participant. One of the benefits of one-to-one matching is that it enables a researcher to estimate the effect of treatment by directly comparing outcomes between treated and untreated subjects in the matched sample. Furthermore, Austin (2010, p.1094) conducted 1,000 simulations and found that "bias increased as *M* increased with *M*:1 matching." In other words, increasing the number of untreated subjects matched to each treated subject increases the bias in the estimated treatment effect. Additionally, one-to-one matching allows for the sensitivity test conducted and discussed in the robustness checks section (Table 6).

It should also be noted that although the matching ensures the treatment and comparison students are similar across covariates on average in the entire sample, it leaves room for variation between pairs. Consequently, a matched treatment and comparison pair of students with similar propensity scores may differ on multiple characteristics. Because of this, a researcher may consider stratifying the match by variables believed to need exact matching, such as test score proficiency levels and school year in this study. In our case, this results in 64 different groups, considering 4 school years, and 4 proficiency levels, for 2 subjects. Having this many groups presents two critical challenges. Firstly, Guo and Frazer (2015) conclude their text by explaining that a common pitfall in propensity score matching is "failure to justify and present the model predicting propensity scores" (p. 383). Stratifying in the manner described, would result in 64 separate regressions to calculate propensity scores, many of which do not have enough power and do not statistically significantly predict program participation, due to the smaller samples. Secondly, stratifying complicates the interpretation of the result of sensitivity analysis, which is most helpful when the effect of the unknown confounder is compared to covariates in the model predicting program participation. With both considerations, matching was done without stratification. Odds ratios for the propensity score calculating logistic regression can be seen in Table 4.

Table 4. Propensity Score Calculating Logistic Regression

	Logistic Regression Odds Ratios	Standard Error
Black	2.629***	(0.155)
Other Race	1.372***	(0.079)
Latinx	1.818***	(0.081)
Economically. Disadvantaged/Low Income	0.915*	(0.035)
Male	0.541***	(0.020)
Disability	0.390***	(0.032)
English Learner	0.638***	(0.068)
Chronically Absent in G8	0.607***	(0.041)
Suspended in G8	0.620***	(0.053)
Missing ELA Test	1.700**	(0.287)
Needing Improvement/Partially Meeting Expectations in ELA	1.337***	(0.106)
Proficient/Meeting Expectations in ELA	1.668***	(0.142)
Advanced/Exceeding Expectations in ELA	1.728***	(0.181)
Missing Math Test	1.528*	(0.287)
Needing Improvement/Partially Meeting Expectations in Math	1.269***	(0.086)
Proficient/Meeting Expectations in Math	1.789***	(0.133)
Advanced/Exceeding Expectations in Math	2.311***	(0.203)
School Year=2019	0.397***	(0.028)
School Year=2021	1.460***	(0.075)
School Year=2022	1.324***	(0.066)
Observations	25599	

Exponentiated coefficients; Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

When matching on propensity scores, the concern is whether covariate balance between the treated and untreated is achieved after the application of propensity scores. As such, a good propensity score calculating logistic regression is not necessarily one that best predicts treatment assignment (Shiba & Kawahara, 2021). In other words, reporting measures of model fit that assess the predictive performance of a propensity model is of limited value (Westreich et al., 2011).

Matching Within the Early College Network

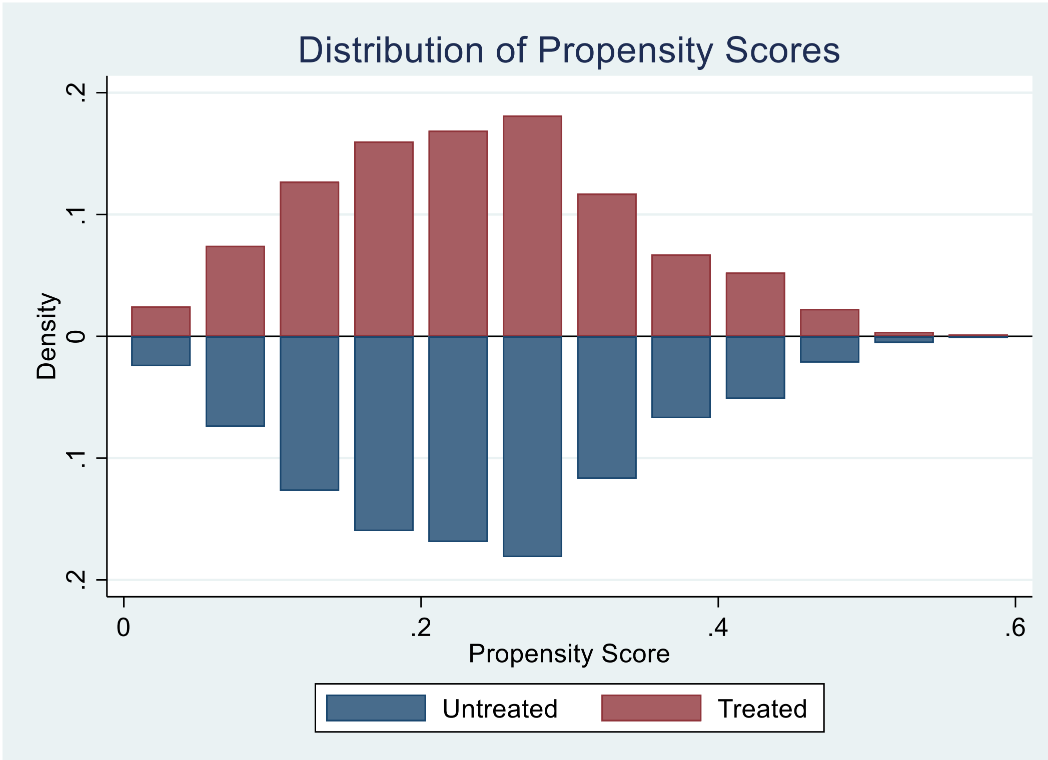
Ideally, Early College participants would be matched with comparable students within their own school. However, the share of students participating in the program varied by school, such that some schools have less than 5 percent of their 12th graders participating while others have more than 50 percent of their 12th grade students participating. In the schools with higher shares of students participating, there would not be enough non-participating students with whom to match. Thus, program participants were matched with students within the network of schools that offer Early College. Given that schools self-selected to apply, drawing only from the Early College designated schools, as opposed to the rest of the state, alleviates concerns around the possibility that participating schools and non-participating schools are systematically different, which may lead to the erroneous conclusion that school effects are Early College program effects. School fixed effects are included in the post-match regression analysis to address variation that may exist between the Early College-offering schools.

Diagnostic Tests

Balance between Treatment and Comparison Groups

After calculating propensity scores, one must confirm that there is common support, or investigate whether there is overlap in the range of propensity scores across treatment and comparison groups (Garrido et al., 2014). Garrido et al. (2014) further assert that inferences about the effect of an intervention cannot be made for treated observations for whom there are no comparison observations with similar propensity scores. Common support may be assessed by examining a graph of propensity score distribution across treatment and comparison groups. As shown in Figure 1, there is nearly an identical distribution as well as overlapping of propensity scores in the treated and comparison groups.

Figure 1. Distribution of Propensity Score across Treatment and Comparison Groups



Furthermore, a balance table was constructed and shows that there are no statistically significant differences, in terms of the rates at which different attributes are represented, between the treatment and comparison group (Table 5). At an 84 percent match rate, 4,358 of the 5,177 Early College participants that were in grade 12 from the 2018-19 to the 2021-22 academic years were matched with 4,358 students from Early College offering schools. Note that the two reference groups from the regressions—White for race and Failing/Not Meeting Expectations for test score level—are omitted.

Table 5. Post-match Early College Balance Table.

Variable	Treated	Comparison	t-statistic	p-value
Black	15%	14%	0.73	0.463
Other Race	13%	13%	0	1
Latinx	42%	42%	-0.07	0.948
Female	64%	64%	0.22	0.824
Economically. Disadvantaged/Low Income	48%	48%	-0.19	0.847
English Learner	3%	3%	1.12	0.263
Disability	4%	4%	-0.1	0.916
Missing ELA Test	1%	1%	0.7	0.484
Needing Improvement/Partially Meeting Expectations in ELA	30%	30%	-0.12	0.907
Proficient/Meeting Expectations in ELA	51%	52%	-0.47	0.637
Advanced/Exceeding Expectations in ELA	12%	11%	0.24	0.814
Missing Math Test	1%	1%	1.07	0.284
Needing Improvement/Partially Meeting Expectations in Math	37%	37%	0.09	0.929
Proficient/Meeting Expectations in Math	39%	39%	-0.18	0.86
Advanced/Exceeding Expectations in Math	15%	15%	-0.42	0.672
Chron. Abs. G8	7%	6%	0.48	0.631
Suspended in G8	4%	4%	1.07	0.286
Total Students	4,358	4,358	—	—

Sensitivity Analysis

A common strategy for addressing hidden bias from unobserved characteristics that may affect treatment assignment as well as the outcome is the Rosenbaum Bounds Sensitivity Analysis, which allows a researcher to calculate how strongly an unobserved confounder must affect treatment assignment to be able to undermine the results from a matching analysis (DiPrete & Gangl, 2004). One may determine the strength an unmeasured influence must have in order to render the estimated treatment effect inconclusive by looking at the Rosenbaum bounds on treatment effects at different levels of gamma (DiPrete & Gangl, 2004). According to the Rosenbaum bounds test on this propensity score match, as displayed in Table 6, the critical level of gamma at which we would have to question our conclusion of a positive effect is between 1.6 and 1.7, when sigma becomes greater than .05. In other words, it would require a hidden bias of gamma between 1.6 and 1.7 to render spurious the conclusion of a positive effect of program participation on college enrollment.

Table 6. Rosenbaum Bounds for Delta

Gamma	Sig +
1	0
1.1	0
1.2	2.20E-16
1.3	3.50E-11
1.4	2.20E-07
1.5	0.000112
1.6	0.007711
1.7	0.108771
1.8	0.455049
1.9	0.828073
2	0.974551

These gamma values, which are in odd ratios units, correspond to the odds ratios of the covariates from the logistic regression from which the propensity scores were calculated. 1.6 is higher than all but 7 of the 20 covariates in the propensity score calculating model— including Black, Latinx, Proficient in ELA, Advanced in ELA, Proficient in Math, and Advanced in Math (See Table 4). Given that EC specifically targets Black and Latinx students in recruitment, per the equitable access guiding principle in the designation criteria (DESE & DHE, 2017), and the observed pattern of EC enrolling higher achieving students, as described in the Program Participation section of this paper, it is expected that these variables would have the highest odd ratios in the model predicting program participation. Sitting in the range of 1.6 to 1.7 odds ratio, the unobservable confounder would have to change the odds of participating in the program as much as variables expected to strongly predict program participation. Thus, the researchers interpret the results to suggest that this match is not particularly sensitive to omitted variables. The seventh covariate that has an odds ratio higher than or equal to 1.7 is Missing ELA Test Score. Less than one percent of 12th graders are missing ELA test scores in the state of Massachusetts during the years of this study, so the odds ratio of 1.7 for this particular covariate is not strongly considered.

Matched vs. Unmatched Records

Under the .2 caliper specification chosen by the researcher, 84 percent of the treatment observations were kept, while 16 percent were pruned. In the context of propensity score matching, pruning can be described as a negotiation between imprecise matching and incomplete matching (Sainani, 2012). More precise matches can mean a smaller sample and a larger sample can mean less precise matching, i.e. a trade-off between internal and external validity. As shown in Table 7, the two largest statistically significant differences in the matched and pruned records are the higher rates at which the pruned records are English learners and advanced in math, at 26 and 17 percentage point differences, respectively. Henry et al. (2016) find English proficiency to have a statistically significant positive relationship with mathematics scores, so it is sensible that English learners who score highly in math are scarce and thus difficult to match. This potentially presents a limitation on the external validity of this research, such that its findings have limited applicability to English Learners who score highly in Math. Other differences are either much smaller or not statistically significant.

Table 7. Matched vs. Unmatched Early College Treatment Records.

Variable	Matched	Pruned	t-statistic	p-value
Black	15%	22%	-5.32	0
Other Race	13%	8%	4.09	0
Latinx	42%	44%	-0.59	0.556
Female	64%	65%	-0.48	0.632
Economically. Disadvantaged/Low Income	48%	55%	-3.59	0
English Learner	3%	29%	-29.24	0
Disability	4%	2%	3.04	0.002
Missing ELA Test	1%	8%	-3	0.003
Needing Improvement/Partially Meeting Expectations in ELA	30%	44%	-1.5	0.135
Proficient/Meeting Expectations in ELA	51%	40%	1.1	0.271
Advanced/Exceeding Expectations in ELA	12%	8%	0.55	0.585
Missing Math Test	1%	0%	0.48	0.63
Needing Improvement/Partially Meeting Expectations in Math	37%	32%	0.55	0.581
Proficient/Meeting Expectations in Math	39%	24%	1.48	0.138
Advanced/Exceeding Expectations in Math	15%	32%	-2.45	0.014
Chron. Abs. G8	7%	8%	-0.48	0.63
Suspended in G8	4%	2%	0.95	0.34
Total Students	4,358	819	—	—

Results/Findings

Effects of EC on College Readiness

Various post-match regression models estimate that EC participation has a positive effect on college readiness, as approximated by MassCore completion (Table 8). The first model, which is a logistic regression including all the covariates in the match, estimates that EC participants have 1.95 times the odds of being college ready, relative to their comparison group. The second regression model is a post-match school fixed effects model, which estimates that EC participants have 1.37 times the odds of being college ready, relative to the comparison group. The big difference in the first and second model suggests that there is great variation in impact by school. This is possibly because MassCore is a recommended, but not a mandated program of study by the state, although the MassCore Curriculum is a feature of the Early College program. Some schools do not require it for high school graduation. The third model incorporates interaction terms between values of race and the treatment. Black and Latinx have significant interaction terms having odds ratios of 1.96 and 2.12, respectively. This means that although Early College participation increases the odds of completing MassCore for all students on average, the increased odds are greater for Black and Latinx students. This interaction is graphed in probability units in Figure 2. Model three also includes a significant interaction between program participation and being low-income. The odds ratio for that interaction is .77, which means that the positive effects of program participation are slightly less pronounced for low-income students, per Figure 3. The similar effect sizes from models two and three, suggest that the variation in effect by school may be due to variation in share of program participants who are Black or Latinx at different schools. No significant interactions were found between treatment and race or treatment and income status in the school fixed effects model (Appendix 1A).

Table 8. Effects on MassCore Completion/College Readiness

	Post-match Regression	(SE)	With School Fixed Effects	(SE)	With Race and Low-Income Interactions	(SE)
Early College	1.948***	(0.104)	1.370***	(0.109)	1.341**	(0.151)
Black	0.738***	(0.067)	0.925	(0.130)	0.557***	(0.066)
Other	0.380***	(0.032)	0.931	(0.120)	0.310***	(0.035)
Latinx	0.664***	(0.047)	0.653***	(0.071)	0.486***	(0.045)
Low Income	0.780***	(0.045)	0.711***	(0.059)	0.868	(0.064)
Male	0.902	(0.050)	0.776**	(0.060)	0.901	(0.050)
Disability	0.767*	(0.095)	0.548***	(0.095)	0.765*	(0.095)
English Learner	0.915	(0.151)	0.835	(0.198)	0.921	(0.153)
Chronically Absent in G8	0.549***	(0.054)	0.455***	(0.065)	0.546***	(0.054)
Suspended in G8	0.788	(0.102)	0.683*	(0.123)	0.786	(0.102)
Missing Test Score (ELA)	1.038	(0.269)	1.176	(0.425)	1.059	(0.275)
Partially Meeting Expectations (ELA)	0.946	(0.115)	1.228	(0.211)	0.950	(0.116)
Proficient/Meeting Expectations (ELA)	1.190	(0.155)	1.278	(0.236)	1.197	(0.156)
Advanced/Exceeding Expectations (ELA)	1.383*	(0.224)	1.413	(0.324)	1.394*	(0.226)
Missing Test Score (Math)	0.788	(0.246)	0.708	(0.308)	0.777	(0.243)
Partially Meeting Expectations (Math)	1.120	(0.117)	1.258	(0.183)	1.121	(0.117)
Proficient/Meeting Expectations (Math)	1.025	(0.117)	1.766***	(0.287)	1.024	(0.118)
Advanced/Exceeding Expectations (Math)	0.813	(0.108)	1.970***	(0.376)	0.811	(0.109)
School Year =2019	1.666***	(0.209)	1.010	(0.180)	1.678***	(0.211)
School Year =2021	0.955	(0.075)	0.793*	(0.090)	0.956	(0.075)
School Year =2022	0.966	(0.074)	0.939	(0.109)	0.966	(0.074)
EC & Black Interaction					1.959***	(0.355)
EC & Other Race Interaction					1.618**	(0.271)
EC & Latinx Interaction					2.116***	(0.298)
EC & Low-Inc. Interaction					0.774*	(0.087)
Constant	4.039***	(0.651)	194.016***	(201.204)	4.720***	(0.793)
Observations	8716		7572		8716	

Exponentiated coefficients; Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 2. Early College and Black Interaction for College Readiness

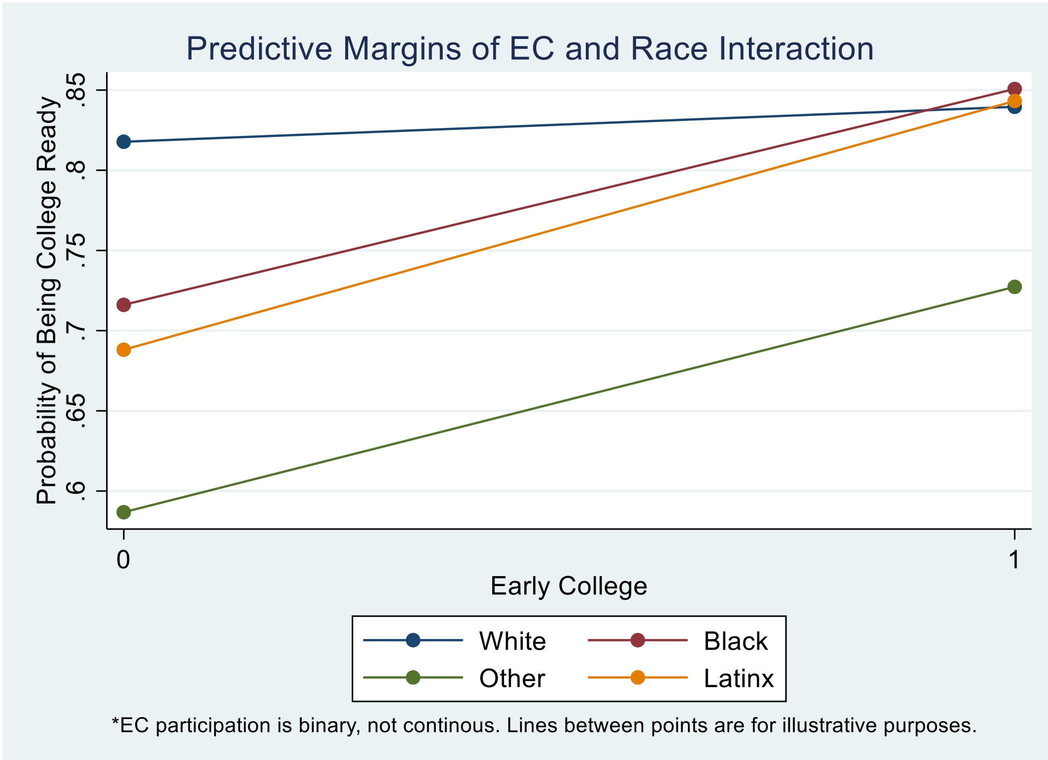
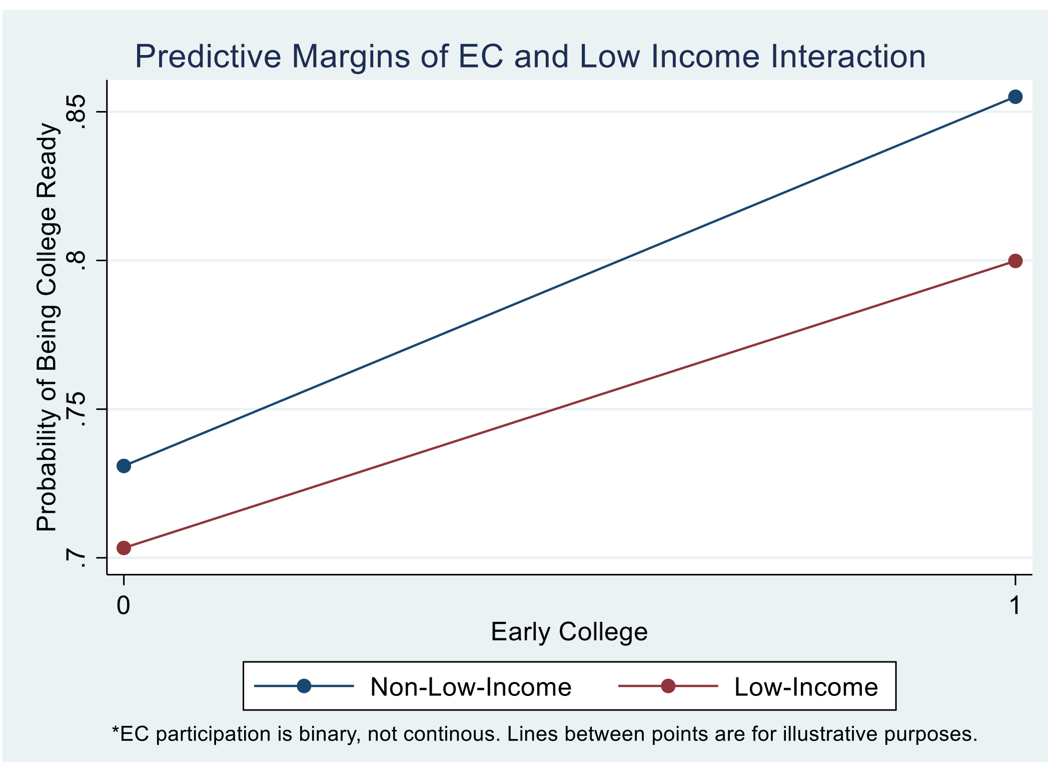


Figure 3. Early College and Low-income Interaction for College Readiness



Effects of EC on College Access

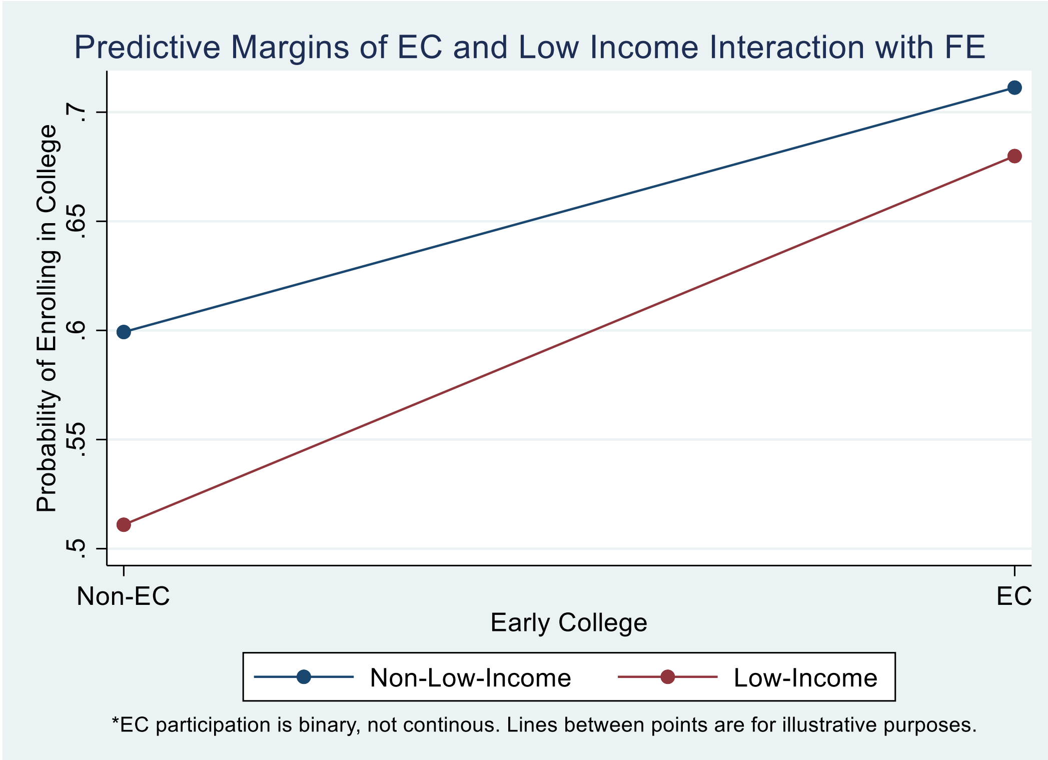
Early College is likewise estimated to have positive effects on college enrollment, by various regression models (Table 9). Per the post-match logistic regression containing all covariates from the match, EC participants have 1.69 times the odds of enrolling in college the year after completing high school, relative to the comparison group. Including school fixed effects in the second model increased the estimate to 2.02 times the odds of enrolling in college for EC participants, relative to the comparison group, meaning that variation in impact by school may mask an even greater impact. The third regression model features an interaction between being a participant and being low-income, which remained significant with school fixed effects. It estimates an impact of increasing the odds of enrolling in college by 1.78 times for participants overall, and an odds ratio coefficient of 1.28 for the interaction of being an EC participant and having a socioeconomically disadvantaged background. This means that although Early College participation increases the odds of enrolling in college for all students on average, the increased odds are greater for low-income students. This interaction is graphed in probability units in Figure 4. There were no significant interactions between program participation and race values of Black or Latinx (Appendix 1B).

Table 9. Effects on College Enrollment/College Access

	Post-match Regression	(SE)	With School Fixed Effects	(SE)	With School Fixed Effects and Low-Income Interaction	(SE)
Early College	1.689***	(0.080)	2.022***	(0.110)	1.778***	(0.133)
Black	1.007	(0.080)	1.143	(0.104)	1.148	(0.105)
Other	0.987	(0.082)	0.964	(0.087)	0.967	(0.087)
Latinx	0.524***	(0.032)	0.566***	(0.040)	0.566***	(0.040)
Low Income	0.700***	(0.036)	0.741***	(0.039)	0.660***	(0.046)
Male	0.701***	(0.035)	0.717***	(0.037)	0.716***	(0.037)
Disability	0.671***	(0.079)	0.623***	(0.076)	0.624***	(0.076)
English Learner	0.734	(0.116)	0.721*	(0.118)	0.718*	(0.117)
Chronically Absent in G8	0.515***	(0.050)	0.558***	(0.056)	0.558***	(0.056)
Suspended in G8	0.571***	(0.071)	0.565***	(0.072)	0.566***	(0.072)
Missing Test Score (ELA)	1.015	(0.235)	0.908	(0.215)	0.907	(0.215)
Partially Meeting Expectations (ELA)	0.979	(0.109)	0.970	(0.111)	0.970	(0.111)
Proficient/Meeting Expectations (ELA)	1.252	(0.147)	1.222	(0.148)	1.222	(0.148)
Advanced/Exceeding Expectations (ELA)	1.487**	(0.220)	1.421*	(0.217)	1.420*	(0.217)
Missing Test Score (Math)	1.632	(0.463)	1.665	(0.500)	1.677	(0.504)
Partially Meeting Expectations (Math)	1.484***	(0.141)	1.407***	(0.138)	1.407***	(0.138)
Proficient/Meeting Expectations (Math)	2.069***	(0.214)	1.883***	(0.202)	1.882***	(0.202)
Advanced/Exceeding Expectations (Math)	3.286***	(0.412)	2.816***	(0.369)	2.814***	(0.369)
School Year =2019	1.637***	(0.177)	1.930***	(0.224)	1.929***	(0.224)
School Year =2021	0.951	(0.068)	0.929	(0.069)	0.930	(0.069)
School Year =2022	0.926	(0.065)	0.816**	(0.062)	0.818**	(0.062)
EC and Low-Income Interaction					1.277*	(0.125)
Constant	1.261	(0.184)	0.763	(0.380)	0.783	(0.391)
Observations	8716		8707		8707	

Exponentiated coefficients; Standard errors in parentheses
$^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

Figure 4. Early College and Low-income Interaction for College Enrollment (School Fixed Effects)



Effects of EC on College Success

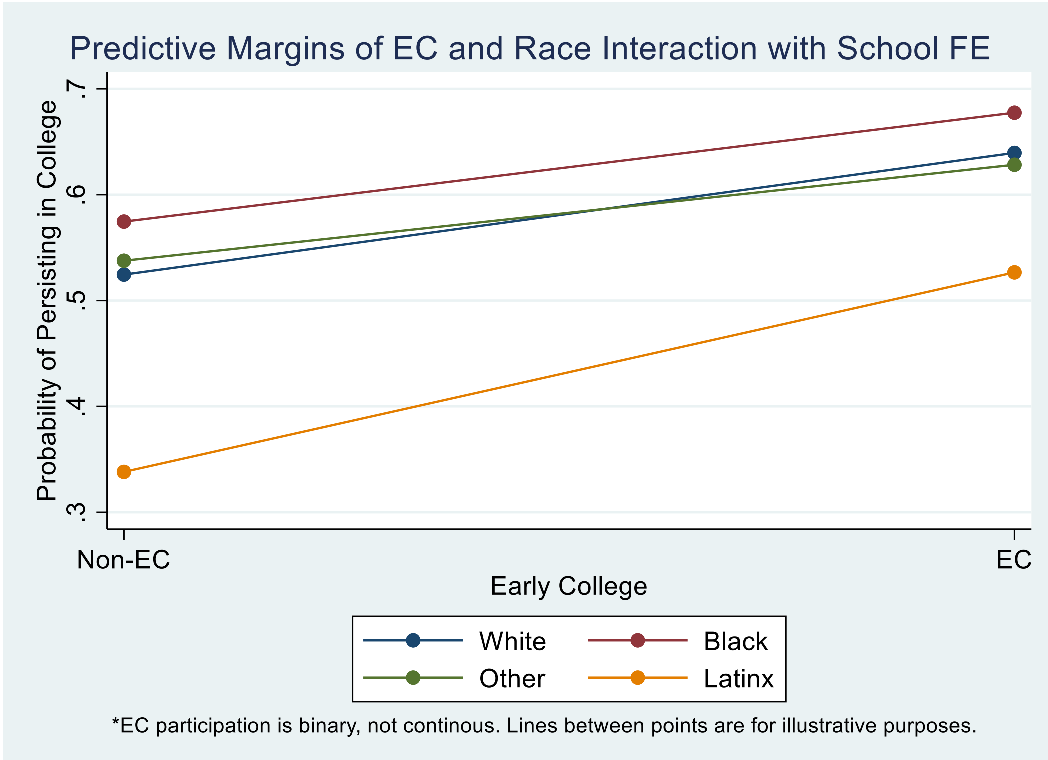
With a smaller sample, due to needing an additional year of data to calculate, Early College made an estimated positive difference on second-year persistence (Table 10). Looking at the post-match regression with the match covariates, EC participants had 1.7 times the odds of persisting to a second year in college, relative to the comparison students. Including school fixed effects in the second regression model increased the estimated impact to 2.01 times the odds of persisting to a second year in college for EC participants, relative to the comparison group. Like the impact on college enrollment, this means that variation in impact by school may mask a greater effect size for EC participation. The third model includes an interaction between program participation and being Latinx, with an odds ratio of 1.42, meaning that although EC participation increased the odds of persisting in college for all students on average, the increased odds are greater for Latinx students, as shown in probability units in Figure 5. There were no significant interactions between program participation and income group in the post-match regression with school fixed effects (Appendix 1C).

Table 10. Effects on Second-year Persistence/College Success

	Post-match Regression	(SE)	With School Fixed Effects	(SE)	With School Fixed Effects and Race Interaction	(SE)
Early College	1.697***	(0.102)	2.014***	(0.141)	1.759***	(0.213)
Black	1.097	(0.112)	1.257*	(0.145)	1.272	(0.189)
Other	0.986	(0.103)	1.004	(0.114)	1.065	(0.162)
Latinx	0.448***	(0.034)	0.478***	(0.042)	0.406***	(0.046)
Low Income	0.748***	(0.048)	0.794***	(0.053)	0.794***	(0.053)
Male	0.638***	(0.040)	0.645***	(0.042)	0.646***	(0.042)
Disability	0.516***	(0.091)	0.494***	(0.089)	0.492***	(0.089)
English Learner	0.814	(0.184)	0.811	(0.189)	0.821	(0.190)
Chronically Absent in G8	0.436***	(0.063)	0.462***	(0.069)	0.460***	(0.068)
Suspended in G8	0.539***	(0.096)	0.546***	(0.099)	0.545***	(0.099)
Missing Test Score (ELA)	1.785	(0.554)	1.688	(0.535)	1.719	(0.545)
Partially Meeting Expectations (ELA)	1.208	(0.215)	1.190	(0.218)	1.187	(0.218)
Proficient/Meeting Expectations (ELA)	1.609**	(0.294)	1.580*	(0.297)	1.574*	(0.296)
Advanced/Exceeding Expectations (ELA)	2.228***	(0.462)	2.154***	(0.459)	2.154***	(0.460)
Missing Test Score (Math)	1.918*	(0.562)	1.848*	(0.575)	1.845*	(0.574)
Partially Meeting Expectations (Math)	1.623***	(0.206)	1.499**	(0.197)	1.496**	(0.196)
Proficient/Meeting Expectations (Math)	2.474***	(0.332)	2.219***	(0.310)	2.220***	(0.311)
Advanced/Exceeding Expectations (Math)	4.098***	(0.614)	3.403***	(0.536)	3.405***	(0.538)
School Year =2019	1.349**	(0.138)	1.557***	(0.173)	1.570***	(0.175)
School Year =2021	1.068	(0.078)	1.014	(0.078)	1.010	(0.078)
EC and Latinx Interaction					1.416*	(0.213)
Constant	0.495***	(0.103)	0.308***	(0.102)	0.364**	(0.125)
Observations	5396		5390		5390	

Exponentiated coefficients; Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 5. Early College and Latinx Interaction for College Persistence



Robustness Check

As a robustness check, the researcher limited the data to the 2020-21 and 2021-22 academic years, when all the 12th graders would have only taken the Next Generation MCAS in 8th grade, in order to use the continuous test scores as opposed to the test performance levels in the regression models that calculate propensity scores as well as those that estimate the main effects. All other aspects of the models remained the same. As shown in Table 11, this approach estimates that Early College Participants have 1.41 higher odds of being college ready, 1.59 higher odds of enrolling in college, and 1.34 higher odds of persisting to a second year in college, relative to non-participants in their schools, excluding those who participate in Innovation Career Pathways. All models across the two approaches, continuous test scores or factor test performance levels, estimate that Early College participation has a positive effect on all three outcomes. Notably, the 1.59 odds estimated effect on college enrollment from this main effect model using continuous test scores is similar in magnitude to the 1.69 odds estimated effect from the main effect model in Table 9, which uses test performances levels.

Table 11. Estimated Effects Using Continuous Test Scores from 2020-21 and 2021-22

	Effects on College Readiness	(SE)	Effects on College Access	(SE)	Effects on College Success	(SE)
Early College	1.413***	(0.088)	1.593***	(0.090)	1.336***	(0.112)
Black	0.751**	(0.079)	1.107	(0.101)	1.091	(0.147)
Other	0.349***	(0.034)	0.826*	(0.078)	0.660**	(0.093)
Latinx	0.587***	(0.050)	0.580***	(0.042)	0.475***	(0.050)
Low Income	0.761***	(0.052)	0.677***	(0.041)	0.661***	(0.060)
Male	0.981	(0.065)	0.667***	(0.040)	0.609***	(0.055)
Disability	0.755	(0.110)	0.875	(0.121)	0.786	(0.184)
English Learner	1.001	(0.198)	0.988	(0.191)	0.821	(0.284)
Chronically Absent in G8	0.592***	(0.068)	0.515***	(0.058)	0.515***	(0.096)
Suspended in G8	0.699*	(0.106)	0.638**	(0.096)	0.499**	(0.125)
Next Gen ELA	1.009***	(0.002)	1.008***	(0.002)	1.014***	(0.003)
Next Gen Math	0.993**	(0.002)	1.015***	(0.002)	1.017***	(0.003)
School Year 2022	1.034	(0.065)	0.918	(0.052)	1.000	(.)
Observations	6050		6050		2750	

Exponentiated coefficients; Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Discussion

The findings of this paper add to the broader discussion on Early College. As previously mentioned, Edmunds et al. (2017) indicated that an area for further research is the investigation of the efficacy of Early College strategies within traditional comprehensive high schools. This study does indeed respond to that call, by evaluating the impact of Massachusetts Early College, which is mostly a program within a school model, with an average of 21% of 12th grade students at any designated school participating and only one school having 100% of its students participate. The 2.02 odds ratio from the post-match multivariate logistic regression with school fixed effects that estimates the impact of EC participation on college enrollment translates to a 14-percentage point positive difference when using a linear probability model. At a 14-percentage point positive impact estimate on college enrollment, the findings of this paper are consistent in direction and magnitude with the 15.6-percentage point positive impact estimate found by Edmunds et al. (2017; 2020), as part of their longitudinal randomized controlled trial in North Carolina.

Albeit to a lesser extent, the 14-percentage point positive impact estimate is also relatively consistent with Haxton et al. (2016), who found Early College to have a 9-percentage point positive impact estimate on college enrollment from a longitudinal randomized controlled trial. Thus, this propensity score matching non-experimental causal study—in addition to the experimental studies of Edmunds et al. (2017; 2020), Berger et al. (2013), and Haxton et al. (2016)—provides further evidence that Early College promotes educational attainment. While the estimated effects are positive and consistent with the experimental literature, the mechanism by which Massachusetts Early College is promoting college readiness, access, and success is not evaluated, as part of this study. We posit that while the implementation of Massachusetts Early College as a policy decision has had positive effects on average, further studies should be conducted, to identify best practices to inform the improvement of existing Early College programs and the expansion of the intervention.

Beyond the effectiveness of the program on average, this paper particularly argues that Early College is an intervention to promote more equitable educational attainment. In models estimating effects on college access, including one with school fixed effects, there is an interaction between program participation and having a socioeconomically disadvantaged background, which result in increased odds of said students enrolling in college. In a society where socioeconomically disadvantaged students are underrepresented in their attendance at colleges and universities, relative to their share of the population, this intervention may be a sensible way to address this inequity in Massachusetts, and potentially beyond. Likewise, in models estimating effects on college success, there is an interaction between program participation and being Latinx, which result in increased odds of Latinx students persisting in college.

Although the within-school feature is what distinguishes Massachusetts Early College from the whole school Early College models in other states, a point of discussion is the variation in the count and share of 12th graders who participate in EC at each EC-offering school. With one school having as few as 1.5% of 12th graders participating and another having 100% of 12th graders participating, some interesting questions emerge. While variation by school is addressed in the regression models that estimate impact with school fixed effects, what are the implications for implementation? Can a school deliver Early College as well for 5% of the cohort as it does 100 %? In light of these questions, further research – and the Massachusetts DESE and DHE – might usefully consider issues around implementation and program quality for schools. It should be noted that the number of students across the commonwealth enrolled in Early College during the years of the study is relatively low. Per Table 1, although there were 5,177 12th grade students participating in Early College between 2018-19 and 2021-22, they made up just under 2% of 12th graders in the state of Massachusetts.

Massachusetts is a local-control state, where public school governance is legally delegated to district and school boards located in the communities they serve, as opposed to the state government, making it more difficult to have state-wide interventions. Therefore, this impact evaluation should be of particular interest to the northeastern states where local-control is prominent. The flexibility of Massachusetts Early College renders it more easily replicated in such states, than the whole school models of southern or western states, where state-directed control is the norm.

Although the discussion emphasizes college enrollment, there are two additional outcomes investigated in this study—college readiness and college success. Greater effects of EC on college readiness for Black and Latinx students may be about the relationship between EC and the measure of college readiness used in the study, as opposed to college readiness itself, which is unknown. College readiness is approximated by MassCore completion, which is the taking and completion of certain courses that align with college and workforce expectations. There is a world in which a student may have completed MassCore but is not college ready, and vice versa. Likewise, college success is approximated by second-year persistence, which is an intermediary measure of the ultimate measure of college success, earning a degree. Nevertheless, this study finds that Early College participants have higher odds of completing MassCore, enrolling in college the year after graduating high school, and persisting to a second year in college than non-participants with similar observable characteristics and baseline achievement.

With the framing of offering Early College as a human capital investment, as defined by Shultz (1961) and Becker (1962), and the years of schooling as an outcome measure (Sweetland, 1996), the positive effects of program participation on college enrollment suggest that the return is very promising. It is the positive effects on college persistence that makes this return concrete, as it indicates an additional year of schooling. Using the Mincer equation, which estimates the average monetary returns of one additional year of education, the World Bank estimates that each additional year of education produces a private rate of return to schooling of about 5–8% per year, with the highest returns being for tertiary education (Patrinos, 2016). This means that every year of learning generates about a 5-8% increase in annual income for the individual. For students who enrolled in college and persisted to second year as a result of participating in this intervention, this theoretical return on investment may be very real.

Limitations

There are two concerns that arise from using performance levels across three tests. Firstly, the four performance levels of one test do not directly correspond to the four performance levels of the other tests. For example, a student who is proficient (PARCC/Legacy MCAS) does not exactly correspond to a student who is meeting expectations (Next Generation MCAS). The multivariate logistic regression model that calculates propensity scores as well as the regressions that estimate effect sizes all include cohort fixed effects, via the school years, so comparisons across years, and therefore across tests, is not a concern. While PARCC and Legacy MCAS tests were administered in the same years, the Massachusetts Department of Elementary and Secondary Education provided a conversion table by which all PARCC scores were converted to Legacy MCAS scores, after which their performance level was determined, so the performance levels for PARCC test takers correspond with performance levels of Legacy MCAS test takers.

The second concern is that the performance levels are not granular enough to sufficiently capture variation that may exist between treated and untreated subjects. For example, a student on the high end of proficient (PARCC/Legacy MCAS) or meeting expectations (Next Generation MCAS) may have performed substantially differently from a student on the low end of that same performance level. Although they were speaking about test scores as an outcome variable, as opposed to a predictor variable, May et al. (2009) argue in their Institute for Education Sciences (IES) technical methods report that while scale scores provide greater precision, performance levels are not just categorized continuous scores, but rather judgments about what cutoff points indicate substantively meaningful attainment of knowledge or skills, and thus may be used in impact evaluations, so long as caution is exercised when interpreting results. Nevertheless, the robustness check indicates that a model using continuous test scores yields similar results to those using factor test performance levels, when looking at college enrollment. While MassCore completion and college persistence are proxies for college readiness and success, college enrollment is a direct measure, so it is of principal interest in this study.

It would be remiss for a researcher using propensity score matching not to acknowledge its limitations. There are three standout limitations of this analysis—imbalance on unobserved characteristics, pruning, and uncertainty that it is the intervention and not something else at the school causing the differences in outcomes—all of which have been variously addressed above. Methods that seek to mitigate bias by controlling for confounding variables, such as propensity score matching, do well in balancing observed baseline covariates between groups, but they do not necessarily balance unmeasured characteristics and confounders (Nuttall & Houle, 2008). In other words, propensity score matching is limited in that unmeasured or otherwise unincluded confounding variables—e.g. a highly motivated parent—may be present, leading to biased estimators of treatment effects. This is addressed with a sensitivity analysis.

Additionally, researchers who use propensity score matching for policy analysis and program evaluation must often negotiate a trade-off between imprecise matching and incomplete matching (Sainani, 2012). Propensity scores can also be understood as a measure of the distance between two participants in a sample or population, which are used to determine whether an individual is a good match for another (Huber et al., 2017). As discussed, a researcher decides the caliper, or minimum distance between treatment and control participants to be matched. The shorter the distance, the more precise the match, but the less likely that a match will be found—incomplete matching. The longer the distance, the less precise the match, but the higher likeliness of finding a match—imprecise matching.

The consequence of incomplete matching is that treatment participants for whom a sufficiently close control participant could not be found are dropped. Why would a researcher move forward with an incomplete match? Dropping records, referred to as pruning in the literature, by eliminating poorly matched treatment participants, ensures balanced propensity score distributions (Ripollone, 2018). The downside to pruning, however, is that it progressively leads to a lack of generalizability of findings. Conversely, while increasing the minimum distance between treatment and control participants to be matched safeguards against pruning and gives researchers a large sample, imprecise matching may lead to residual confounding. This paper addresses pruning by comparing the kept and the pruned records.

Conclusion

This study finds Early College to be both generally effective, with odds of enrolling in college that are 1.9 times higher for program participants relative to non-participants, and to promote equity by particularly benefitting low-income participants via interactions. Other impact evaluations of Early College have similarly found this type of intervention to be effective. The within-school option of Massachusetts Early College—as opposed to entirely whole-school programs—distinguishes the Massachusetts program from otherwise similar programs that have been studied elsewhere, giving states that are considering Early College, especially those with local-control, another model to consider. Edmunds et al. (2017) indeed suggest that implementation of Early College within traditional comprehensive high schools need to be evaluated to determine whether it works in these settings.

There remain areas to be further researched. Although the effects are positive on average, there is great variation in their magnitude across the schools, evidenced by school fixed effects models. Because the guiding principles and designation criteria are flexible, it is not clear what is the driver behind the efficacy of Early College. This paper is therefore limited in its ability to highlight which and what dosage of the components of the Massachusetts Early College intervention works best. Furthermore, as of the writing of this paper, the college completion rate, which is typically measured five to six years after students enter college, is not yet available for evaluation, but further analysis is planned to focus on this outcome.

References

Adelman, C. (2006). The Toolbox Revisited: Paths to Degree Completion from High School Through College. Washington, DC: U.S. Department of Education. https://files.eric.ed.gov/fulltext/ED490195.pdf

Advisory Committee on Student Financial Assistance (2006). Mortgaging Our Future: How Financial Barriers to College Undercut America's Global Competitiveness. https://files.eric.ed.gov/fulltext/ED529499.pdf

Akerhielm, K., Berger, J., Hooker, M., & Wise, D. (1998). Factors Related to College Enrollment. Final Report. Prepared for U.S. Department of Education by Mathtech, Inc. https://files.eric.ed.gov/fulltext/ED421053.pdf

An, B. P. (2013). The Impact of Dual Enrollment on College Degree Attainment: Do Low-SES Students Benefit? Educational Evaluation and Policy Analysis, 35(1), 57–75. http://www.jstor.org/stable/23356970

Austin P. C. (2010). Statistical criteria for selecting the optimal number of untreated subjects matched to each treated subject when using many-to-one matching on the propensity score. American journal of epidemiology, 172(9), 1092–1097. https://doi.org/10.1093/aje/kwq224

Austin P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate behavioral research, 46(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Austin P. C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharmaceutical statistics, 10(2), 150–161. https://doi.org/10.1002/pst.433

Barnett, E. (2016). Building Student Momentum from High School to College. Jobs For the Future. Ready or Not, It’s Time to Rethink the Twelfth Grade. https://files.eric.ed.gov/fulltext/ED564836.pdf

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. Journal of Political Economy, 70(5), 9–49. http://www.jstor.org/stable/1829103

Berger, A., Turk-Bicakci, L., Garet, M., Song, M., Knudson, J., Haxton, C., Zeiser, K., Hoshen, G., Ford, J., Stephan, J., Keating, K., & Cassidy, L. (2013). Early College, Early Success: Early College High School Initiative Impact Study . (). American Institutes for Research: https://www.air.org/sites/default/files/downloads/report/ECHSI_Impact_Study_Report_Final1_0.pdf

Britton, T. A., Symns, M. O., & Paul, V. (2022). Early Birds: An Exploration of Early College Initiative High Schools in New York City and College Persistence. Journal of College Student Retention: Research, Theory & Practice, 24(2), 548-574. https://journals.sagepub.com/doi/epub/10.1177/1521025120924782

Culver, K.C., Braxton, J. & Pascarella, E. (2019). Does teaching rigorously really enhance undergraduates’ intellectual development? The relationship of academic rigor with critical thinking skills and lifelong learning motivations. Higher Education 78, 611–627. https://doi.org/10.1007/s10734-019-00361-z

D'Agostino, R. B. (1998). Propensity Score Methods for Bias Reduction in the Comparison of a Treatment to a Non-Randomized Control Group. Statistics in Medicine, 17(19), 2265–2281. https://doi.org/10.1002/(sici)1097-0258(19981015)17:19<2265::aid-sim918>3.0.co;2-b

DeAngelo, L., & Franke, R. (2016). Social Mobility and Reproduction for Whom? College Readiness and First-Year Retention. American Educational Research Journal, 53(6), 1588–1625. http://www.jstor.org/stable/44245965

DiPrete, T. A., & Gangl, M. (2004). Assessing Bias in the Estimation of Causal Effects: Rosenbaum Bounds on Matching Estimators and Instrumental Variables Estimation with Imperfect Instruments. Sociological Methodology, 34(1), 271-310. https://doi.org/10.1111/j.0081-1750.2004.00154.x

Domina, T., Penner, A., & Penner, E. (2017). Categorical Inequality: Schools As Sorting Machines. Annual review of sociology, 43(1), 311–330. https://doi.org/10.1146/annurev-soc-060116-053354

Dunbar, A., & Shapiro, D. (2016). The National Student Clearinghouse as an Integral Part of the National Postsecondary Data Infrastructure. Envisioning the National Postsecondary Data Infrastructure in the 21st Century. https://nscresearchcenter.org/wp-content/uploads/NSC-as-an-Integral-Part-of-the-National-Postsecondary-Data-Infrastructure.pdf

Edmunds, J. A., Arshavsky, N., Lewis, K., Thrift, B., Unlu, F., Furey, J., (2017) Preparing Students for College: Lessons Learned From the Early College. NASSP Bulletin 2017, Vol. 101(2) 117 –141. https://files.eric.ed.gov/fulltext/EJ1145050.pdf

Edmunds, J. A., Unlu, F., Glennie, E., Bernstein, L., Fesler, L., Furey, J., & Arshavsky, N. (2017). Smoothing the Transition to Postsecondary Education: The Impact of the Early College Model. Journal of Research on Educational Effectiveness, 10(2), 297-325. https://www.tandfonline.com/doi/epdf/10.1080/19345747.2016.1191574?needAccess=true

Edmunds, J., Unlu, F., Furey, J., Glennie, E., & Arshavsky, N. (2020). What Happens When You Combine High School and College? The Impact of the Early College Model on Postsecondary Performance and Completion. Educational Evaluation and Policy Analysis, 42(2), 257-278. https://journals.sagepub.com/doi/epub/10.3102/0162373720912249

Fisher, A. E. (2019). Students Identities and Teacher Expectations: A Factorial Experiment at the Intersection of Race, Gender, and Ability. Theses and Dissertations--Educational, School, and Counseling Psychology. 85. https://uknowledge.uky.edu/edp_etds/85/

Gamoran, A, (2015). The Future of Educational Inequality in the United States: What Went Wrong, and How Can We Fix It? American Educational Research Association. https://wtgrantfoundation.org/wp-content/uploads/2015/11/The-Future-of-Educational-Inequality-Adam-Gamoran.pdf

Garrido, M. M., Kelley, A. S., Paris, J., Roza, K., Meier, D. E., Morrison, R. S., & Aldridge, M. D. (2014). Methods for Constructing and Assessing Propensity Scores. Health Services Research, 49(5), 1701-1720. https://doi.org/10.1111/1475-6773.12182

Guo, S., & Fraser, M. W. (2015). Propensity Score Analysis: Statistical Methods and Applications. SAGE Publications.

Haxton, C., Song, M., Zeiser, K., Berger, A., Turk-Bicakci, L., Garet, M. S., Knudson, J., & Hoshen, G. (2016). Longitudinal Findings From the Early College High School Initiative Impact Study. Educational Evaluation and Policy Analysis, 38(2), 410–430. http://www.jstor.org/stable/44984545

Henry, D. L., Nistor, N. & Baltes, B. (2014). Examining the Relationship Between Math Scores and English Language Proficiency. Journal of Educational Research and Practice, 4(1), 11–29. https://files.eric.ed.gov/fulltext/EJ1118452.pdf

Huber, S., Dietrich, J. F., Nagengast, B., & Moeller, K. (2017). Using propensity score matching to construct experimental stimuli. Behavior research methods, 49(3), 1107–1119. https://doi.org/10.3758/s13428-016-0771-8

Hummer, R.A., Lariscy, J.T. (2011). Educational Attainment and Adult Mortality. In: Rogers, R., Crimmins, E. (eds) International Handbook of Adult Mortality. International Handbooks of Population, vol 2. Springer, Dordrecht. https://www.researchgate.net/profile/Joseph-Lariscy/publication/216839888_Educational_Attainment_and_Adult_Mortality/links/0c96053175fb694d4b000000/Educational-Attainment-and-Adult-Mortality.pdf

Karp, M., Calcagno, J., Hughes, K., Jeong, D., & Bailey, T. (2008). Dual Enrollment Students in Florida and New York City: Postsecondary Outcomes. New York, NY: Columbia University, Teachers College Community College Research Center. https://ccrc.tc.columbia.edu/media/k2/attachments/dual-enrollment-student-outcomes.pdf

Klopfenstein, K., & Thomas, M. K. (2009). The Link Between Advanced Placement Experience and Early College Success. Southern Economic Journal, 75, 873-891. https://doi.org/10.1002/j.2325-8012.2009.tb00935.x

Kretchmar, J., & Farmer, S. (2013). How Much is Enough? Rethinking the Role of High School Courses in College Admission. Journal of College Admission. https://files.eric.ed.gov/fulltext/EJ1011884.pdf

Kubota, Y., Heiss, G., MacLehose, R. F., Roetker, N. S., & Folsom, A. R. (2017). Association of Educational Attainment With Lifetime Risk of Cardiovascular Disease: The Atherosclerosis Risk in Communities Study. JAMA Internal Medicine, 177(8), 1165-1172. 10.1001/jamainternmed.2017.1877

Leinbach, D., & Jenkins, D. (2008). Using Longitudinal Data to Increase Community College Student Success: A Guide to Measuring Milestone and Momentum Point Attainment. New York, NY: Columbia University, Teachers College, Community College Research Center. https://files.eric.ed.gov/fulltext/ED499922.pdf

Levin-Waldman, O. M. (2013). Income, civic participation and achieving greater democracy. The Journal of Socio-Economics, 43, 83-92. https://doi.org/10.1016/j.socec.2013.01.004

Littnerova, S., Jarkovsky, J., Parenica, J., Pavlik, T., Spinar, J., & Dusek, L. (2013) Why to use propensity score in observational studies? Case study based on data from the Czech clinical database AHEAD 2006–09. Cor et Vasa, 55(4), 383-390. https://www.sciencedirect.com/science/article/pii/S0010865013000441

Lobo, B. J., & Burke-Smalley, L. A. (2018). An empirical investigation of the financial value of a college degree. Education Economics, 26(1), 78-92. https://www.tandfonline.com/doi/epdf/10.1080/09645292.2017.1332167?needAccess=true-

Massachusetts Department of Elementary and Secondary Education (2023). Massachusetts Comprehensive Assessment System: Test Design and Development. https://www.doe.mass.edu/mcas/tdd/default.html

Massachusetts Department of Elementary and Secondary Education (2023). School Safety and Discipline Report (SSDR) Data Handbook Version 23.0. https://www.doe.mass.edu/infoservices/data/ssdr/ssdr-datahandbook.docx

Massachusetts Department of Elementary and Secondary Education (2023). Student Information Management System (SIMS) Data Handbook Version 30.0. https://www.doe.mass.edu/infoservices/data/sims/sims-datahandbook.docx

Massachusetts Department of Elementary and Secondary Education. (2022). MassCore: Questions and Answers. https://www.doe.mass.edu/ccte/ccr/masscore/faq.docx#:~:text=MassCore%20was%20developed%20to%20increase,of%20admission%20to%20private%20colleges

Massachusetts Department of Higher Education & Massachusetts Department of Elementary and Secondary Education (2017). Early College Program Designation Process and Criteria. https://www.doe.mass.edu/ccte/pathways/early-college/designation-criteria.pdf

Massachusetts Department of Higher Education (2016). The Degree Gap—Honing In on College Access, Affordability & Completion in Massachusetts. https://www.mass.edu/visionproject/_documents/2016%20The%20Degree%20Gap%20-%20Vision%20Project%20Annual%20Report.pdf

Massachusetts Education Equity Partnership (2018). #1 For Some: Opportunity & Achievement in Massachusetts. https://masseduequity.org/wp-content/uploads/2023/08/Number-1-for-Some-9.25-18.pdf

May, H., Perez-Johnson, I., Haimson, J., Sattar, S., & Gleason, P. (2009). Technical Methods Report: Using State Tests in Education Experiments: A Discussion of the Issues. https://ies.ed.gov/ncee/pdf/2009013.pdf

National Center for Education Statistics (2015). Postsecondary Attainment: Differences by Socioeconomic Status. https://nces.ed.gov/programs/coe/pdf/coe_tva.pdf

National Center for Education Statistics (2022). The Nation's Report Card: 2022 Reading State Snapshot Report. https://nces.ed.gov/nationsreportcard/pubs/stt2022/2023010.aspx

National Center for Education Statistics. (2018). An Overview of NAEP. U.S. Department of Education Institute of Education Sciences. https://nces.ed.gov/nationsreportcard/subject/about/pdf/naep_overview_brochure_2018.pdf

Nuttall, G. A., & Houle, T. T. (2008). Liars, damn liars, and propensity scores. Anesthesiology, 108(1), 3–4. https://doi.org/10.1097/01.anes.0000296718.35703.20

Patrinos, H. A. (2016). Estimating the return to schooling using the Mincer equation. IZA World of Labor. https://wol.iza.org/articles/estimating-return-to-schooling-using-mincer-equation/long

Perna, L. W., & Thomas, S. L. (2006). A Framework for Reducing the College Success Gap and Promoting Success for All. National Symposium on Postsecondary Student Success: Spearheading a Dialog on Student Success. http://repository.upenn.edu/gse_pubs/328

Petronijevic, U., & Oreopoulos, P. (2013). Making College Worth It: A Review of Research on the Returns to Higher Education. National Bureau of Economic Research Working Papers. http://www.nber.org/papers/w19053

Redford, J., & Hoyer, K. M. (2017). First-generation and continuing-generation college students: A comparison of high school and postsecondary experiences. https://nces.ed.gov/pubs2018/2018009.pdf

Ripollone, J. E., Huybrechts, K. F., Rothman, K. J., Ferguson, R. E., & Franklin, J. M. (2018). Implications of the Propensity Score Matching Paradox in Pharmacoepidemiology. American journal of epidemiology, 187(9), 1951–1961. https://doi.org/10.1093/aje/kwy078

Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika, 70(1), 41–55. https://www.jstor.org/stable/2335942

Sainani, K. L. (2012). Propensity Scores: Uses and Limitations. American Academy of Physical Medicine and Rehabilitation, 4(9), 693–697. https://doi.org/10.1016/j.pmrj.2012.07.002

Schultz, T. W. (1961). Investment in Human Capital. The American Economic Review, 51(1), 1–17. http://www.jstor.org/stable/1818907

Shiba, K., & Kawahara, T. (2021). Using Propensity Scores for Causal Inference: Pitfalls and Tips. Journal of epidemiology, 31(8), 457–463. https://doi.org/10.2188/jea.JE20210145

Shields, K. A., Bailey, J., Hanita, M., & Zhang, X. (2021). The effects of accelerated college credit programs on educational attainment in Rhode Island. Washington, DC: U.S. Department of Education, Institute of Education Sciences. https://files.eric.ed.gov/fulltext/ED612887.pdf

Shireman, R. (2004). Rigorous Courses and Student Achievement in High School: An Options Paper for the Governor of California. Research & Occasional Paper Series. Center for the Study of Higher Education at UC Berkeley.
https://cshe.berkeley.edu/sites/default/files/publications/2004_rigorous_courses_and_student_achievement_in_high_school.pdf

Smith, J., & Wertlieb, E. (2005). "Do First-Year College Students' Expectations Align with Their First-Year Experiences?" Journal of Student Affairs Research and Practice, 42(2), 153–174.
https://www.researchgate.net/publication/254612661_Do_First-Year_College_Students'_Expectations_Align_with_their_First-Year_Experiences

Song, M., & Zeiser, K. (2019). Early College, Continued Success: Longer-Term Impact of Early College High Schools. American Institutes for Research. https://www.air.org/sites/default/files/Early-College-Continued-Success-Longer-Term-Impact-of-ECHS-September-2019.pdf

Stearns, E., Potochnick, S., Moller, S., & Southworth, S. (2010). High School Course-Taking and Post-Secondary Institutional Selectivity. Research in Higher Education, 51(4), 366–395. https://www.jstor.org/stable/40606368

Stone, C. A. & Tang, Y. (2019). Comparing Propensity Score Methods in Balancing Covariates and Recovering Impact in Small Sample Educational Program Evaluations. Practical Assessment, Research, and Evaluation, 18(1). https://doi.org/10.7275/qkqa-9k50

Stuart E. A. (2010). Matching methods for causal inference: A review and a look forward. Statistical science : a review journal of the Institute of Mathematical Statistics, 25(1), 1–21. https://doi.org/10.1214/09-STS313

Sweetland, S. R. (1996). Human Capital Theory: Foundations of a Field of Inquiry. Review of Educational Research, 66(3), 341-359. https://journals.sagepub.com/doi/epdf/10.3102/00346543066003341

The Center for College Affordability and Productivity (2010). 25 Ways to Reduce the Cost of College: A Policy Paper from the Center for College Affordability and Productivity. https://files.eric.ed.gov/fulltext/ED536144.pdf

The College Board (2022). Massachusetts SAT Suite of Assessments Annual Report. https://reports.collegeboard.org/media/pdf/2022-massachusetts-sat-suite-of-assessments-annual-report.pdf

The Nation's Report Card (2023) Data Tools: State Profiles. https://www.nationsreportcard.gov/profiles/stateprofile?chort=1&sub=MAT&sj=AL&sfj=NP&st=MN&year=2005R3

U.S. Census Bureau (2023). Census Bureau Releases New Educational Attainment Data. https://www.census.gov/newsroom/press-releases/2023/educational-attainment-data.html#:~:text=the%20Hispanic%20population.-,From%202012%20to%202022%2C%20the%20percentage%20of%20adults%20age%2025,20.9%25%20for%20the%20Hispanic%20population

U.S. Census Bureau. (2021). Educational Attainment. https://data.census.gov/table?t=Educational%20Attainment&g=040XX00US01,02,04,05,06,08,09,10,12,13,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,44,45,46,47,48,49,50,51,53,54,55,56,78

Westreich, D., Cole, S. R., Funk, M. J., Brookhart, M. A., & Stürmer, T. (2011). The role of the c-statistic in variable selection for propensity score models. Pharmacoepidemiology and drug safety, 20(3), 317–320. https://doi.org/10.1002/pds.2074