Interim Outcomes Report

Boston Coaching for Completion (BosC4C)

Final

April 12, 2019

Prepared for:
Elizabeth Pauley
Antoniya Marinova
Juan Cantu
John Griffin

The Boston Foundation
75 Arlington Street, 10th Floor
Boston, MA 02116

Written by:
Tamara Linkow
Erin Bumgarner
Hayley Didriksen
Kelly Lack
Austin Nichols
Emily Dastrup
Samuel Dastrup
Beth Gamse

Abt Associates
10 Fawcett Street
Cambridge, MA 02138

List of Exhibits

Acknowledgements

The Success Boston Study Team would like to thank the students, coaches, non-profit organizational leaders and staff, higher education partners and staff, and the Boston Foundation staff. Their many contributions to the study are deeply appreciated. We are the beneficiaries of their generosity of time and spirit, which were instrumental in the preparation of this report.

The report would not have been possible without the contributions of many staff at Abt Associates including RJ dela Cruz, Bry Pollack, and Julie Strawn, as well as our colleagues Jane Furey and Fatih Unlu.

Executive Summary

Access to middle class jobs increasingly requires a college degree or credential. College graduates earn more, are less likely to suffer job losses in a recession, and are projected to have superior long-term labor market prospects (U.S. Census Bureau 2017; McFarland et al. 2018). Individuals with postsecondary education have competitive advantage in the labor market: even when a job does not explicitly require a degree, a candidate with a degree will tend to be hired over an equally qualified candidate without one. Over the next decade, the number of jobs requiring a bachelor's degree will continue to increase at a higher pace than any other educational level (Bureau of Labor Statistics 2018). At the same time, Massachusetts is faced with an aging workforce where nearly half of the labor market is 45 or older (Commonwealth Corporation 2014). In Boston, the six-year college graduation rate for the city's 2011 public high school graduates who enrolled in college was 52 percent (McLaughlin and Van Eaton 2018).[1] This rate improves upon the 39 percent seven-year rate for 2000 graduates, yet likely is not sufficient to meet the predicted demand for a college-educated workforce.[2]

Low-income students, in particular, along with first-generation students and racial/ethnic minorities, are underrepresented in postsecondary education (Bailey and Dynarski 2011; McFarland et al. 2018). Boston's public school graduates are disproportionately from groups underrepresented in postsecondary education. Like students from low-income backgrounds and racial/ethnic minority groups elsewhere, they may face social, academic, logistical, and financial barriers to succeeding in college; may lack the supports needed to overcome barriers; and may struggle with managing key deadlines, such as financial aid and course registration.

One-on-one coaching from experienced counselors when students are completing their senior year in high school and beginning college can help them succeed (Arnold et al. 2009; Castleman, Arnold, and Wartman 2012; Bettinger and Baker 2014; Castleman, Page, and Schooley 2014; Avery, Howell, and Page 2014; Carrell and Sacerdote 2013; Scrivener and Weiss 2009; Stephan and Rosenbaum 2013; Sum et al. 2013).

A core strategy of the city-wide **Success Boston** initiative is one-on-one transition coaching. Together, the Boston Foundation (TBF), the City of Boston, Boston Public Schools (BPS), University of Massachusetts Boston, Bunker Hill Community College, other regional colleges and universities, uAspire, the Boston Private Industry Council, and other local nonprofit organizations are working together through Success Boston to reduce barriers students' face in transitions to and success in college. The coaching model is implemented across a network of nonprofit organizations (that provide coaching) in partnership with institutions of higher education. Transition coaching offers students sustained, proactive, and responsive support in their first two years after high school.

---

[1] The college graduation rate includes completion of certificates, Associate degrees, and Bachelor degrees.

[2] A 2008 report, *Getting to the Finish Line: College Enrollment and Graduation, a Seven-Year Postsecondary Longitudinal Study of the Boston Public Schools Class of 2000 Graduates* (Sum et al. 2008), found that 64 percent of nearly 3,000 Boston Public Schools Class of 2000 graduates enrolled in a postsecondary institution within the first seven years of high school graduation, yet only 35.5 percent of college enrollees had earned a certificate, a two-year degree, or a four-year degree. That figure was later revised to 39 percent.

Beginning with the high school Class of 2009, Success Boston provided transition coaching to Boston high school graduates, many of whom are from groups traditionally underrepresented in college. Since the fall of 2015, and funded in part through a Social Innovation Fund grant from the federal Corporation for National and Community Service, the Boston Coaching for Completion (BosC4C) scale-up has broadened the reach of transition coaching, by increasing the number of students served threefold from around 600 to about 2,000 students annually. To maintain continuity with prior evaluation reports and resonate with students, colleges, nonprofit organizations, and the community, this report refers to the transition coaching program as **Success Boston Coaching (SBC)**.

Selected Study Findings

SBC coached students are 5% more **likely** than are non-coached peers to persist into the second year of college

After two years of college, coached students **accumulated** 7% more **credits** than did non-coached students

Coached students are 8% **more likely** to renew their FAFSA than are non-coached students

About the Evaluation

Prior studies provide evidence that SBC can be effective at increasing student persistence in college (Sum et al. 2013; Linkow et al. 2017b), suggesting potential for SBC to boost college graduation rates for BPS high school graduates. This evaluation examines both the implementation and impact of SBC after the scale-up.

Research Questions. In particular, the evaluation is designed to answer four main research questions about implementation and impact:

1. What is the effect of SBC on student success in college?
2. How, if at all, do the impacts of SBC vary by student characteristics and features of the coaching?
3. How is SBC implemented across partner organizations and partner colleges?
4. What resources are necessary to implement SBC?

This report, the second of three reports to be released over the course of the multi-year evaluation, focuses on the effectiveness of coaching on students' persistence and achievement early in college, answering the first and second research questions. It focuses on the BPS graduating Classes of 2015 and 2016 who entered college in either the fall of 2015 or 2016.[3] This report also includes updated information about the resources necessary to implement coaching in the 2016-17 and 2017-18 academic years. The third, and final, report will examine the impacts of coaching on college completion.

[3] The third and fourth research questions are addressed in an earlier implementation-focused report, *Success Boston Coaching for Completion 2015-16 Implementation Report* (Linkow et al. 2017a).

Data Sources. For this interim outcomes report, the study team collected data from BPS, the Massachusetts Department of Elementary and Secondary Education, and 11 partner colleges.

Analysis. The report assesses whether and how transition coaching delivered through SBC by eight nonprofit coaching organizations affects students' short-term college outcomes. It examines impacts for two specific cohorts of students: those who graduated from BPS in 2015 and 2016 and would have entered college in the fall of 2015 or 2016. These two cohorts have been out of high school for three and two years, respectively. Using a quasi-experimental design, the report compares outcomes for the students who participated in SBC ("treatment" group) to outcomes of a group of similar students ("comparison" group) who did not participate. As such, the report provides evidence that observed differences in outcomes between the two groups are due to participation in Success Boston coaching. Based on what we learned, several themes have emerged.

About the Findings

The analyses estimate that SBC students have better early college outcomes than do their carefully matched peers not participating in SBC. Specifically, SBC students are

- more likely to persist into their *second* year of college (82.6 percent for SBC students in the treatment group vs. 78.3 percent for the non-coached students in the comparison group).
- enrolled full-time for more semesters (67.3 percent of semesters vs. 63.6 percent).
- accumulating more college credits (of the credits necessary to graduate at their college, SBC students completed 45.5 percent vs. 42.5 percent for non-coached students in the comparison group).
- more likely to complete FAFSA renewals for their second year of college (83.9 percent vs. 77.7 percent).

The magnitude of the impacts is smaller than that reported about earlier cohorts of SBC students and other similar programs targeting improved college student outcomes (see Bettinger and Baker 2014; Castleman and Page 2015; Linkow et al. 2017a; Oreopoulos and Petronijevic 2016). These other studies have observed increases of 10-15 percent in college persistence and achievement. Overall, the effects of the scaled-up SBC program can be characterized as positive and moderate in magnitude.

Exploratory analyses estimate that, generally, coaching is particularly impactful for students who experience more frequent interactions with coaches. Specifically, SBC students who experience 14 or more coaching interactions and have more interactions about academics and career in particular, have more positive college outcomes than SBC students who experience fewer interactions. Those students who experienced

- more coach-student interactions had higher rates of persistence into the second and third years of college, have higher cumulative GPAs, and are more likely to renew their FAFSA than are students who had experienced fewer interactions; and
- more frequent coaching interactions related to academic and career topics were more likely than students with less exposure to such topics to persist into the second and third years and to renew their FAFSA.

Lastly, the cost analyses for the 2016-17 academic year shows that annual per student cost of SBC was about $3,000. The cost is substantially lower than the $5,700 annual per student cost observed in 2015-16. The higher per student costs in 2015-16 reflect smaller student caseloads and additional start-up costs associated with the first year of the scale-up (Linkow et al. 2017b).

Final Report to Come

The study's final report, scheduled for release in 2022, will track the same two cohorts, plus an additional cohort of students who entered college in the fall of 2017, for up to six years after their college entrance. That report will examine whether the positive impacts on early outcomes found here—which are predictive of college completion (Bettinger and Baker 2014; Scrivener and Weiss 2009; Stephan and Rosenbaum 2013)—do in fact result in more students completing a college degree or credential within that time frame.

1. Introduction

The **Success Boston** initiative is a city-wide collaborative of the Boston Foundation (TBF), City of Boston, Boston Public Schools (BPS), University of Massachusetts Boston, Bunker Hill Community College, other regional colleges and universities, uAspire, the Boston Private Industry Council, and other local nonprofit organizations. The partners focus on boosting college persistence and ultimately improving college completion rates for Boston's public school graduates through a purposeful combination of program, policy, and practice-based activities. Success Boston prioritizes low-income, first-generation students of color, focusing on helping students *get ready* for college, *get into* college; *get through* college and *get connected* to a career upon college graduation.

The Boston Foundation is the convening backbone organization of the Success Boston initiative. In particular, the foundation provides funding and other resources to nonprofit organizations engaged in the initiative's core intervention: **one-on-one *transition coaching.*** Coaching is provided to students during their first two years in college, in this report hereafter known as the **Success Boston Coaching (SBC)** program.

1.1 About Success Boston Coaching (SBC)

Beginning with the high school Class of 2009, Success Boston provides transition coaching to Boston area high school graduates to help reduce barriers to college success, particularly for students from groups traditionally underrepresented in college. The transition coaching model offers students sustained, proactive, and responsive support provided by nonprofit organizations. Importantly, SBC serves students who enroll in two-year and four-year colleges.

In 2015-16, 2016-17, and 2017-18, the Boston Foundation helped to fund eight nonprofit organizations to coach more than 3,000 first- and second-year students enrolled in more than 50 Boston area colleges.

Since the fall of 2015, and funded in part through a Social Innovation Fund (SIF) grant from the federal Corporation for National and Community Service, the **Boston Coaching for Completion scale-up** broadened the reach of the SBC program from several hundred to 1,000 Boston young adults per cohort. SBC is designed to support recent high schools graduates entering college who are most likely to leave before completion—that is, first-generation college and low-income students of color. Through this program, students can access one-on-one coaching at the start of the Fall semester or shortly thereafter and continue through their first two years of college.

In academic year 2016-17, eight nonprofit coaching organizations provided coaching through Success Boston: Boston Private Industry Council, Bottom Line, College Bound Dorchester, Freedom House, Hyde Square Task Force, Sociedad Latina, the Steppingstone Foundation, and West End House.[4] SBC coaches from these organizations worked with BPS students to prepare them to become independent college students able to navigate their way to college graduation. They did so through providing support

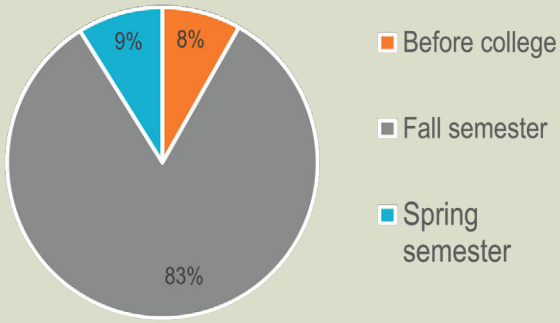
[4] In 2015-16, an additional nonprofit organization provided coaching through Success Boston; this organization did not continue to provide coaching in 2016-17, and therefore students served by this organization in 2015-16 are excluded from the impact evaluation.

on life skills, study skills, help-seeking skills, and academic skills. SBC coaches helped students develop meaningful relationships, clarify goals, access networks and resources, understand college culture, and make college life feasible. SBC coaches also provided job and career mentoring.

In addition, throughout the academic year, uAspire, a national nonprofit organization focused on increasing knowledge and resources to make college affordable, provided the SBC students with direct support in filling out financial aid forms. uAspire also administers a text message program for SBC students that sends financial aid–related information and reminders. Students receive automated text messages throughout their first year of college, with the option to reply back to receive help from a uAspire staff member. In 2016-17, some two-thirds of SBC students (67 percent) received the text message reminders from uAspire; on average students received eight messages during their first year of college and through following summer. uAspire also provided periodic professional development to SBC coaches on financial aid related topics and processes.

Exhibit 1-1: Start of coaching activities

Recruitment of college-intending students can begin as early as the end of high school and continues into the start of college. Coaching activities then begin for most students (83 percent) during the fall semester of college.



Source: Program data 2015-16, 2016-17

To recruit students into SBC, these eight nonprofit coaching organizations used multiple strategies, including referrals from high school guidance counselors and other community organizations; nonprofit organizations' middle school and high school programming pipelines; word of mouth; and outreach on college campuses. Students reported in a 2015 survey that they learned about SBC through presentations from the nonprofit organizations at their high school as well as from conversations with individuals ranging from a nonprofit coach, an afterschool or summer program staff member, a high school or college staff person, to a friend or neighborhood acquaintance. The local area colleges and universities also referred students to the nonprofit partner organizations.

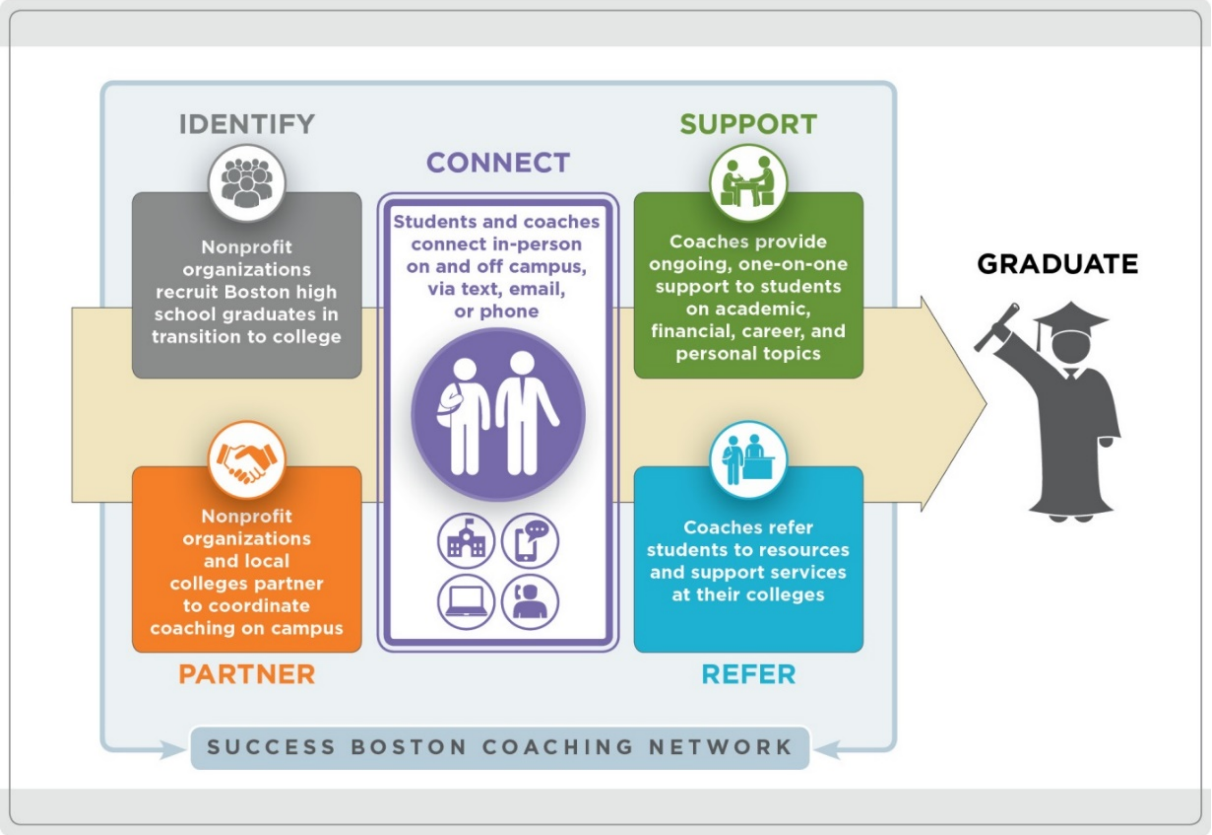
Not surprisingly, the use of such varied strategies, occurring at different points in time, means that students are recruited to participate in SBC as early as the end of high school, continuing during the summer before college enrollment and into of the start of college. It also means that each cohort entering college in the fall of a given year includes students with different relationships to the nonprofit organizations, motivations to attend college, and predispositions to reach out for support. The varied recruitment strategies and staggered timing reduce the chance that unmeasured characteristics related to both participation and outcomes cause any impacts that are found. Once students are recruited, the coaching activities typically start during the first fall semester of college (see Exhibit 1-1).

It should be noted that coaches reported delays in recruitment in 2015-16, and consequently in the start of coaching activities, as compared to subsequent years. This may reflect recruitment challenges experienced as part of the scale-up effort, which required the nonprofits to recruit a substantially greater number of students to the 2015-16 cohort, between April and September, than in previous years. Nonprofits also received grant award notifications slightly later than in past years, delaying the start of recruitment

efforts. Recruitment for SBC continued well into the fall 2015 college semester, which in turn may have limited the amount of time coaches had available to conduct coaching activities. In fact, just 29 percent of first-year college students met with their coach during September 2015, substantially fewer than the almost two-thirds of first-year college students who did so September 2016.

Exhibit 1-2 displays how the organizations, partner colleges, and Success Boston network collaborate to provide transition coaching to support students on the path to college graduation. Each nonprofit organization ***identifies*** and recruits BPS high school graduates through the various means described above to participate in the SBC transition coaching program. Local colleges ***partner*** with the nonprofits and coaches to coordinate coaching activities on their campuses. Once students are confirmed as coaching participants, coaches ***connect*** with them through multiple modes—in person or via text, email, or phone—to help students navigate the college-going process. Through one-on-one meetings, coaches provide ongoing ***support*** to students across a range of topics and ***refer*** students to supports on their campuses.

Exhibit 1-2: Success Boston Coaching model



Coaches support students as needed throughout the academic year, offering proactive and on-demand guidance to students to help them stay on track toward graduation.[5] The SBC network, overseen by TBF,

[5] For a detailed description of how SBC was implemented in the 2015-16 academic year, see the report *Success Boston Coaching for Completion 2015-16 Implementation Report* (Linkow et al. 2017a) available here: https://www.nationalservice.gov/sites/default/files/evidenceexchange/FR_Success%20Boston%20Implementation%20Report_2017.pdf

facilitates communication across organizations; it also provides coaches access to specialized training about financial aid from uAspire as well as access to training on other topics.

The estimated total cost to implement SBC in 2015-16 was $5,214,547, or $5,762 per student (Linkow et al., 2017a).[6] That cost represents multiple inputs, including both direct services to students and program administration. Because 2015-16 was the first year of scale-up, the total cost reflects both start-up and operating expenses An updated cost analysis (included in Appendix D) for the 2016-17 academic year found that the overall cost to implement SBC increased to $6,167,058, but the per student cost decreased to $3,037.

Since the 2009 start of SBC, 10 cohorts of BPS high school graduates have received coaching. Earlier evaluations have demonstrated that coaching increases student persistence and academic achievement in college (Sum, Khatiwada, and Palma 2014; Linkow et al. 2017b). There is also descriptive evidence that for some students, coaching may lead to higher college completion rates. That is, comparing coached and non-coached students from the BPS Class of 2009, the six-year graduation rates for students who enrolled in two-year colleges favor SBC students: 35 percent of SBC students versus 24 percent of non-participating students completed within six years. Further, overall completion rates for Black SBC students were higher than the completion rates of students who did not participate in coaching through Success Boston: 53 percent versus 41 percent (McLaughlin et al. 2016).

1.2 Evaluation Research Questions

Given earlier evidence and TBF's continued investment in BPS students' postsecondary success, TBF contracted with Abt Associates to conduct a comprehensive evaluation of the transition coaching program. The evaluation is designed to answer four main research questions about implementation and impact:

1. What is the effect of coaching—above and beyond the services received by students in the comparison group—on the following key outcomes: annual persistence, grade point average (GPA), academic standing, Free Application for Federal Student Aid (FAFSA) completion, and postsecondary completion?

2. What is the nature of variation in the impacts of coaching for student outcomes, and how is observed variation associated with student characteristics, such as gender and race/ethnicity, and with features of the coaching, including dosage and topics covered?

3. How is SBC implemented across partner organizations (i.e., what is the dosage, caseload, method of delivery, etc.) and partner colleges (i.e., how is coaching integrated into campus support

[6] Costs for 2015-2016 differ somewhat in this report from those reported in the 2015-16 Implementation Report (Linkow et al. 2017a) for two primary reasons. First, we recalculated 2015-2016 total costs excluding the costs of a nonprofit organization that participated in SBC for one year both to facilitate comparisons across years and to align the cost and impact analyses (the impact analysis excludes students enrolled in this nonprofit. Second, costs are multiplied by a 1.058 inflation factor (based onthe combined inflation rates from 2016 to 2018 from Consumer Price Index for the Boston metro area) to allow comparisons across years.

service delivery, what coordination exists between coaches and campus staff)? How do the coaching models vary across partner organizations? How do the coaching and support services provided vary by college?

4. What resources are necessary to implement SBC and to achieve desired student outcomes?

1.3 About This Report

Findings on these research questions are included in three separate reports. The **first report**, produced in March 2017, focused on Research Question 3 to examine how nonprofit partner organizations implemented coaching for 2015 high school graduates (Linkow et al. 2017a) and Research Question 4 to understand resources required to implement the SIF-funded expansion of the coaching program in 2015-16. The **last report**, due out in 2022, will explore long-term outcomes and variation in impacts as of the 2019-20 academic year (Research Questions 1 and 3).

This is the **second report** of the three; it

- focuses on early program impacts (observed as of the fall of the 2018-19 academic year) on student outcomes (Research Question 1) as well as on how program impacts vary by certain student characteristics and features of coaching (Research Question 2);
- estimates the per student costs for the 2016-17 and 2017-18 academic years;
- explores impacts for the high school Classes of 2015 and 2016, who have been out of high school for three and two years, respectively; and
- uses a quasi-experimental design (matched comparison group) to form a strong counterfactual: similar students who did not receive the SBC intervention.

Before we turn to study specifics, we summarize relevant literature about what we might expect to learn about the impact of transition coaching (Chapter 2). Next, we review the study's design, analysis approach, data sources, and measures (Chapter 3). Chapter 4 summarizes the impact analyses separately for each student outcome, including persistence, achievement, and financial aid. For each outcome, we describe the average or overall impact as well as whether and how effects vary by student characteristics. Chapter 5 examines how the impacts of SBC are related to features of the coaching itself. The report concludes with a discussion and recommendations (Chapter 6).

2. Improving College Enrollment and Completion

Today, earning a college degree or credential is seen as crucial for future well-being. College graduates earn more, are less likely to suffer job losses in a recession, and are projected to have superior long-term labor market prospects (U.S. Census Bureau 2017; McFarland et al. 2018). Nationally, more than three of 10 jobs already require postsecondary education (BLS 2018), and more than six of 10 current jobs are filled by candidates with postsecondary education (Carnevale, Smith, and Strohl 2013). These figures reflect the competitive advantage of postsecondary education: even when a job does not explicitly require a degree, a candidate with a degree will tend to be hired over an equally qualified candidate without one. Over the next decade, the number of jobs requiring a bachelor's degree will continue to increase at a higher pace than any other educational level (Bureau of Labor Statistics 2018). At the same time, Massachusetts is faced with an aging workforce where nearly half of the labor market is 45 or older (Commonwealth Corporation 2014). In Boston, 2,397 (or 70.5 percent) of the city's 2011 public high school graduates enrolled in college and 52 percent of them had graduated from college six years later (McLaughlin and Van Eaton 2018). This rate improves upon the 39 percent seven-year rate for 2000 graduates, yet likely is not sufficient to meet the predicted demand for a college-educated workforce.[7]

2.1 Barriers to College Enrollment and Completion

Despite an overall increase both nationally and locally in college-going rates in recent decades, students from low-income backgrounds and racial/ethnic minority groups are less likely to attend, persist, and complete college than their peers (e.g., U.S. Department of Education 2016; Haskins 2008; Bailey and Dynarski 2011). Low-income students in particular, along with first-generation college students, ethnic minorities, and males, have been found to be underrepresented in postsecondary education (Arnold et al. 2009; Arnold, Lu, and Armstrong 2012; Harper 2006; Harper and Griffen 2011; Tym, McMillion, Barone, and Webster 2004). Despite a gradually narrowing gap in enrollment rates, across the income distribution, only 46 percent of those from the lowest income quartile attended a postsecondary institution compared with 78 percent of those from the top income quartile. College completion rates among low-income students paint an even bleaker picture: only 15 percent of youth from the lowest income quartile attain a bachelor's degree, compared with 60 percent of those from the top income quartile (Cahalan et al. 2018).

Success Boston's recent *Staying the Course* report highlights similar trends for Boston. College access and success cut along both racial/ethnic and gender lines among Boston high school graduates: Boston's White and Asian students were more likely than their Black and Hispanic peers to enroll in college and to earn a college credential, and its female students across all racial groups were more likely to graduate from college than male students (McLaughlin and Van Eaton 2018).

In today's knowledge-based economy, disparities in college enrollment and completion rates for male students, students of color, and low-income students place them at a distinct disadvantage in the

[7] A 2008 report, *Getting to the Finish Line: College Enrollment and Graduation, a Seven-Year Postsecondary Longitudinal Study of the Boston Public Schools Class of 2000 Graduates* (Sum et al. 2008), found that 64 percent of nearly 3,000 Boston Public Schools Class of 2000 graduates enrolled in a postsecondary institution within the first seven years of high school graduation, yet only 35.5 percent of college enrollees had earned a certificate, a two-year degree, or a four-year degree. That figure was later revised to 39 percent.

workforce because college education can be a gateway to the middle class for low-income students (e.g., Ayala and Striplen 2002; Haskins 2008; Pfeffer and Hertel 2015). In Boston, the education gap has consequential effects on median annual earnings: adults 25 years of age and older with a Bachelor's degree earn, on average, $54,768—nearly twice what high school graduates earn (U.S. Census Bureau 2017). Adults with associate's degrees also earn higher wages on average and are less likely to be unemployed than individuals with only a high school diploma (U.S. Census Bureau 2017; Vuolo, Mortimer, and Staff 2016). In fact, individuals with an Associate's degree earn more than $259,000 over the course of their career than do high school graduates (Klor de Alva and Schneider 2013). Associate's degree holders focused on occupational and technical skills (e.g., healthcare, high-end manufacturing), in particular, have even greater earning potential than other Associate's degree and some Bachelor's degree recipients (Klor de Alva and Schneider 2013).

A college degree represents an opportunity for socioeconomic mobility; when children born into the lowest 20 percent of the income distribution receive a college degree, their chances of escaping the bottom tier increase by more than 50 percent (Isaacs, Sawhill, and Haskins 2008), reflecting the well-documented significant and positive economic returns to a Bachelor's degree (Aud et al. 2012; Carnevale, Rose, and Cheah 2011). A college degree is also related to improved social and health outcomes (Baum, Ma, and Payea 2013; Hout 2012; Meara, Richards, and Cutler 2008).

Low college attendance and completion rates among students from low-income backgrounds and racial/ethnic minority groups are attributed in part to informational and support gaps for these students both before and once they enroll in college (Arnold et al. 2009; Avery and Kane 2004; Avery, Howell, and Page 2014; Bozick and DeLuca 2011; Roderick et al. 2008). In a phenomenon called "summer melt," low-income college-intending high school graduates fail to matriculate to the college of their choice in the fall following their senior year of high school (Arnold et al. 2009; Castleman, Arnold, and Wartman 2012; Castleman, Page, and Schooley 2014). Summer melt has been attributed to gaps in support available to students during the summer, particularly as they encounter difficulties navigating financial aid options and completing numerous time-sensitive administrative tasks, such as course registration and FAFSA completion (Arnold et al. 2009; Castleman, Arnold, and Wartman, 2012; Castleman and Page 2015; Castleman, Page, and Schooley 2014).

Even once enrolled, students from groups traditionally underrepresented in college, in particular, may lack access to professional guidance to help them navigate the financial aid process (Arnold et al. 2009; Bettinger et al. 2012; Roderick et al. 2008) or to prompt them to meet unfamiliar deadlines (Hoxby and Turner 2013; Ross, White, Wright, and Knapp 2013). Students from low-income backgrounds and racial/ethnic minority groups are more likely to be first-generation college students (NCES 2012) whose parents and peers are unfamiliar with the challenges students can face when entering college (Castleman and Page 2013; Stephens et al. 2015).

Students who are new to college may also experience a range of academic challenges, including unanticipated course difficulty, uncertainty about how to select the appropriate courses to meet degree completion requirements, and time allocation management across classes. Further, too many first-generation college students enter higher education underprepared for college-level academic demands, which can affect their capacity to persist and complete college degrees (Greene and Winters 2005).

Moreover, though college advisor systems have been found to be beneficial for students in need of academic remediation (Bahr 2008; Swecker, Fifolt, and Searby 2013), academic advisors may have limited time to provide the level of support students need, particularly students attending two-year and

four-year public institutions. A survey of college academic advisors found that the median caseload of a full-time academic advisor is 441 advisees at public community colleges and 260 advisees at public four-year colleges (Carlstrom and Miller 2013). A separate study, based on a national survey of college counseling center directors, found that the counselor-to-advisees ratio is 1 to 1,500 for 55 percent of community colleges (Gallagher 2010).

Transition coaching is a promising intervention to help students manage the financial, administrative, and academic obstacles they may face; in such coaching, designated coaches work with high school graduates as they enter and adjust to college. In particular, SBC aims to bridge the gap for students who may not have sufficient resources and supports during the transition from high school to college.

Transition coaches can help students assess how they can reduce the gap between the cost of college and what they can afford. Direct supports from coaches may offer a means to address students' misconceptions or gaps in knowledge about college financial aid, help students complete lengthy and complex financial aid forms, and remind students of key due dates. Bettinger et al. (2012) found that personal assistance, combined with information about financial aid, substantially increased FAFSA submissions and ultimately the likelihood of college attendance, persistence, and aid receipt.

Coaching programs represent a promising solution for students who lack access to information, guidance, and general support from their family and social networks (Avery and Kane 2004; Bettinger, Boatman, and Long 2013; Deming and Dynarski 2009; Roderick et al. 2008).

Coaching can also help students struggling to stay on task in their courses through identifying additional supports that can promote persistence and graduation (Bettinger and Baker 2014; Bettinger, Boatman, and Long 2013; Castleman and Page 2015; Karp 2011; Johnson and Rochkind 2009; Oreopoulos and Petronijevic 2016). For example, a recent study that examined the relationship between academic advising and retention of first-generation college students found that academic advising can consistently and effectively connect these students to academic resources on campus. In fact, this analysis found that the odds of a first-generation college student remaining enrolled at a given college increased 13 percent for every meeting with an advisor (Swecker, Fifolt, and Searby 2013).

2.2 Research on Impact of Transition Coaching

Recent rigorous research on transition coaching examines the summer between students' senior year of high school and freshman year of college. Although SBC coaching services generally begin during the fall of students' first college year, the similarities between SBC and summer-based coaching activities suggest meaningful insights on the potential effects of coaching

2.2.1 Evidence for Programs Similar to SBC

Several recent studies, including randomized control trials, viewed as the gold standard in social policy research, find that coaching interventions significantly increase students' college matriculation and persistence. They provide particularly strong evidence regarding the impact of coaching on students' college outcomes.

- In a study of the Beacon Mentoring Program at South Texas College, students in mathematics classes were randomly assigned either to no mentor or to a mentor who encouraged them to use tutoring and other campus services and who offered one-on-one support if needed. The

program increased students' use of the campus tutoring center and reduced the likelihood that they would withdraw from the course (Visher, Butcher and Cerna 2010).

- A small-scale pilot study examined outcomes for students in urban Big Picture high schools who were randomly assigned either to receive systematic outreach and assistance from transition coaches over the summer or to a business-as-usual condition. Coached students' enrollment in four-year colleges was 14 percentage points higher than control group enrollment rates, corresponding to a 1.5 times higher likelihood of program students keeping their postsecondary plans (Arnold et al. 2009; Castleman, Arnold, and Wartman 2012). However, as Castleman and colleagues cautioned, the unique features of Big Picture schools, including individualized attention students receive from coaches, may mean that these findings do not generalize to other coaching interventions in urban high schools.
- The Big Picture schools pilot study design was subsequently replicated under different conditions: students in two districts (Boston and Fulton County, Georgia) were randomly assigned to summer outreach and coaching. Coaching increased college enrollment among program students in both districts. In particular, coaching increased the probability of college enrollment by 3 percentage points, leading to a 20-percent reduction in summer melt. The impacts of coaching were more pronounced for Boston students and for low-income students in both sites (with increased probabilities of between eight and 12 percentage points). In addition, the study found that summer counseling led to increased rates at which students persisted in college through their sophomore year (Castleman, Page, and Schooley 2014).
- Another study examined the effect of Inside Track, which provides virtual one-on-one coaching targeting students currently attending college. Inside Track coaches regularly contact their students to provide help and support as the students start their college careers and continue through their first year of college. Freshman students attending eight different postsecondary institutions, including two- and four-year schools, who were randomly assigned to receive targeted coaching were 15 percent more likely to have persisted in college 18 to 24 months later than those who did not receive the coaching (Bettinger and Baker 2014).
- A study of the Opening Doors program in Ohio randomly assigned students either to a regular college counselor or to a program counselor. Program students were expected to meet with their Opening Doors counselor at least twice each semester for two semesters to discuss academic progress and resolve any issues that might affect their schooling. Each program counselor worked with far fewer students than did the regular college counselors, which allowed for more frequent, intensive contact. The program improved academic outcomes during students' second semester in the study; however, the program did not significantly increase the average number of credits that students earned after the program ended or over the study's three-year follow-up period (Scrivener and Weiss 2009).
- A study of a peer coaching program at the University of Toronto randomly assigned first-year students to upper-year undergraduate coaches. The peer coaches met regularly with students to provide one-on-one support (either in person or via Skype) on a variety of college-related topics. Students who received coaching had significantly higher average grades and overall GPAs: approximately a 5 percentage point increase in average course grades and a 0.35 standard deviation increase in GPA versus students who did not receive coaching (Oreopoulos and Petronijevic 2016).

- A study of the Accelerated Study in Associate Programs (ASAP), operated by the City University of New York, randomly assigned eligible community college students to participate in ASAP. Students participating in ASAP received intensive advising from an ASAP-dedicated adviser with a small caseload, career information from career and employment services staff, and dedicated tutoring services. ASAP also required students to attend full time, and included free tuition, textbooks, and transportation. Participation in ASAP increased students' likelihood of receiving a degree by 18 percentage points and transferring to four-year institutions by nearly 8 percentage points during the study's three-year follow-up period (Scrivener et al. 2015). Even after six years after the start of the program, ASAP continues to increase graduation rates; 51 percent of program group students versus 41 percent of control group students had earned degrees (Gupta 2017).
- An evaluation of the Valley Initiative for Development and Advancement (VIDA) program, which sought to help students graduate with an associate's degree or industry-recognized certificate in a high-demand occupation, randomly assigned students to participate in the program. Students in the VIDA program participated in weekly one-and-one and group counseling session tailored to students' personal and academic needs. In addition, VIDA participants received assistance with tuition and school-related expenses. The VIDA program significantly increased the total number of college credits earned within a 24-month follow-up period; for more than three-quarters of program participants within 36 months of follow-up, effects on credits and credentials continued to increase (Rolston, Copson, and Gardiner 2017).

Though impacts on student outcomes for these studies are generally positive, the magnitude of the impacts and the outcomes on which impacts are detected vary. Castleman, Page, and Schooley (2014) attribute differences in impact to several factors related to the intervention itself, including differences in the rates of student communication with advisors; amount of attention each student received from his/her counselor; and prior experience coaches had with supporting students' college enrollment tasks. Two additional factors—differences in students' access to other supports and supplemental resources, particularly financial support, and differences in percentage of low-income students in each sample—may also have played a role in the variation in magnitude of impacts across the studies cited above.

2.2.2 Research on Success Boston Coaching

Prior research focused specifically on SBC provides promising evidence of the benefits to coaching. In 2014, the Center for Labor Market Studies at Northeastern University used a matched comparison group design to compare outcomes for BPS 2009 graduates who did and did not participate in SBC (Sum, Khatiwada, and Palma 2014). The study found preliminary evidence of a positive and statistically significant effect on college persistence. SBC students were more likely to persist in each of the first four years of college than their non-coached peers, with estimated effects of 17 percentage points in the first year, 18 percentage points in the second year, 15 percentage points in the third year, and 12 percentage points in the fourth year. Though SBC students outperformed each of their comparison group counterparts, persistence rate impacts varied slightly by gender and ethnicity, and impacts were generally greater for Black students. Specifically, SBC impacts on persistence were larger for Black students than their Hispanic peers (25 percent vs. 13 percent). Sum, Khatiwada, and Palma (2014) conducted analyses that controlled for student demographics, students' 10th-grade Massachusetts Comprehensive Assessment System (MCAS) English/language arts scores, and the type of colleges students initially attended; they

found positive and statistically significant effects on college outcomes for BPS 2009 graduates as of 2013.

Success Boston's 2016 *Reaching for the Cap and Gown* report (McLaughlin et al. 2016) provides a descriptive examination of college enrollment and completion for participants in SBC for the BPS Class of 2009, comparing them with non-participating students. The report found that the coached and non-coached students who enrolled in four-year colleges immediately following high school generally completed college at similar rates of about 60 percent, quite similar to the national six-year completion rate of 62 percent for students entering four-year colleges in the fall of 2009 (Shapiro et al. 2015). SBC students have an edge when they attend two-year colleges, however: 35 percent of Success Boston students and 24 percent of non-participating students completed a degree or credential within six years. The *Reaching for the Cap and Gown* report also examined outcomes at the seven top-enrolling colleges and universities (as had Sum and colleagues in their 2013 and 2014 reports), and it found that nearly half (49 percent) of SBC students at these colleges completed a degree, compared with 38 percent of non-coached students. Further, the overall completion rates for Black SBC students—who represented more than one third (36 percent) of SBC students—were higher than the completion rates of Black students who did not participate in coaching through Success Boston: 53 percent versus 41 percent.

Most recently, the Success Boston Coaching interim outcomes report (Linkow et al. 2017b) examined impacts for two earlier cohorts of students who graduated from BPS in 2013 and 2014 and would have entered college in the fall of 2013 and 2014, respectively. Using a rigorous quasi-experimental design, the report compared outcomes for the group of students who participated in Success Boston Coaching versus those of a group of similar students who did not participate. As such, the report provides evidence that observed differences in outcomes between the two groups are due to participation in Success Boston coaching.

Those analyses estimated that SBC students (the "treatment group" in the evaluation) had better early college outcomes than did their carefully matched peers not participating in SBC ("comparison group"). Specifically, SBC students were more likely to persist into their *second* year of college (83 percent for the treatment group vs. 75 percent for the comparison group), more likely to persist into their *third* year of college (75 percent vs. 62 percent), and more likely to complete FAFSA renewals for their second year of college (85 percent vs. 78 percent) (Linkow et al. 2017b).

The current study builds on the 2014 Center for Labor Market Studies study described above. It uses a more rigorous design that matches students more systematically, uses more extensive baseline characteristics in the matching process, and includes not just one but several cohorts of students. The study also examines additional student outcomes, including academic achievement, FAFSA renewal, and college graduation rates. And it investigates how differences in key programmatic features affect student outcomes. Chapter 3 describes the study design in greater detail.

3. Evaluation Design

In this chapter, we begin with an overview and then describe the evaluation design in more detail. The chapter outlines the quasi-experimental approach we used to estimate program impacts. Then it describes the study sample—both the program students and the non-coached matched students who make up the "comparison group." Next, the chapter describes our approach to exploratory analysis—that is, how we examine *variation* in impacts, namely the relationship between *impacts* on students and *variation* in impacts according to student characteristics and features of coaching. The chapter then summarizes the data sources, outcomes, and measures of student characteristics and features of coaching used to explore how program impacts vary according to those characteristics and features. It concludes by reviewing limitations of the evaluation design.

3.1 Study Design

The study uses a quasi-experimental design in which outcomes are compared for students who participated in SBC and students who did not. This report examines impacts for students who graduated from BPS in 2015 and 2016; who have been out of high school for three and two years, respectively; and who entered college in the falls of 2015 and 2016. The study created a comparison group of students who are as similar to the treatment students as possible, using an approach called *local and focal matching*.

The approach is "local" in that each SBC student is matched with one (and possibly multiple) non-SBC students from the same high school graduating class, from high schools with similar characteristics, and enrolled in the same college. It is "focal" because students are matched based on similar baseline characteristics (e.g., gender, race/ethnicity, high school academic achievement, socioeconomic status) that are empirically linked both to the outcomes of interest and potentially to receipt of coaching.

3.1.1 How We Identified Students in the Sample

Students participating in the SBC program are identified in the program's administrative database. All students who appear in the database are considered SBC students for purposes of the evaluation. This inclusive definition means that all students who were initially recruited into the SBC program, and therefore appear in the program database, are eligible to be in the evaluation sample even though some did not have a single recorded interaction with a coach.[8] A total of 2,053 students are identified as SBC students in the 2015 or 2016 college-entering cohorts. The comparison group is identified from 24,396 high school graduates in 2015 and 2016 from BPS and surrounding districts.

To be eligible for the evaluation sample, students had to

- enroll in college in the fall after high school graduation;
- enroll in a college in which at least one SBC student and at least one potential comparison student were enrolled in that given year; and

[8] In the 2016-17 academic year, 2 percent (37 students) of the 2015 and 2016 cohorts had no coaching interactions recorded in the program database.

- have no missing information on key baseline characteristics (free/reduced-price lunch status, high school GPA) used in the matching process.

After applying the eligibility criteria, there were 1,480 SBC students (674 from the 2015 cohort and 806 from the 2016 cohort) and 8,978 non-SBC students (4,384 from the 2015 cohort and 4,594 from the 2016 cohort) who could be included in the evaluation sample (see Exhibit 3-1). The students eligible for the comparison group came from BPS, 21 surrounding districts, 9 Commonwealth charter schools and 2 in-district charter schools.[9] Students included in the evaluation sample were selected through the local and focal matching process referenced above and described in more detail below.

Exhibit 3-1: Students eligible for inclusion in the sample

	2015 cohort		2016 cohort	
	Coached Students	Potential Comparison Students	Coached Students	Potential Comparison Students
Total students	927	12,113	1,126	12,364
Ineligible students				
Graduated in 2013	15	N/A	9	N/A
Graduated in 2014	36	N/A	19	N/A
Graduated in 2015	N/A	N/A	47	N/A
Graduated in 2016	3	N/A	N/A	N/A
Not found in any administrative HS datasets*	101	1	101	1
Ineligible because did not enroll in college in fall after HS graduation	45	4,657	76	5,017
Ineligible because missing key baseline data	2	51	0	262
Ineligible because did not enroll in a college with at least 1 coached and 1 potential comparison student	51	3,020	68	2,490
Eligible for matching	674	4,384	806	4,594

* 2 potential comparison students (one in the 2015 cohort and one in the 2016 cohort) are excluded because they graduated from out-of-state private high schools.

3.1.2 How We Test Program Impacts Using a Matched Comparison Group

Overview of the Matched Comparison Group Design

In social science research, an *experimental* design is considered the gold standard approach for testing program impacts. Experimental designs use random assignment to form treatment and comparison groups and then compare outcomes for the two groups to test whether the treatment group has different (presumably better) outcomes than the comparison group. Because the groups are formed by random assignment, they are expected to be statistically equivalent with respect to all relevant characteristics, so that any differences in outcomes can be attributed to the program of interest and not to some other

[9] "In-district" charter schools (also refered to as "Horace Mann" charter schools) differ from Commonwealth charter schools in that their charters must be approved by the local school committee. In some cases in-district charter schools also require that the charter be approved by the local teacher's union in addition to the the Board of Elementary and Secondary Education. For more information, please see: http://www.doe.mass.edu/charter/new/2015-2016QandA.pdf

characteristic(s) that might have influenced both participation in the program and outcomes (e.g., various dimensions of academic or non-academic achievement levels, including non-cognitive skills).

Because an experimental design was not possible for SBC, given partner organizations' capacity and the size of the potential participant population, we used one of the strongest *quasi-experimental* designs available—a design that allows us to account for as many of the student background characteristics as possible, to help ensure that the treatment and comparison group students are statistically similar *before* participation in SBC. For example, one potential difference in background characteristics between program participants and non-coached students could be academic readiness for college. Students who participate in SBC do so voluntarily; they may simply be more academically prepared to attend college than students who do not sign up for coaching. Differences such as these (also called "confounders") present an important methodological challenge, because differences in preparation (or other background characteristics) could explain why we might see differences in student outcomes for treatment and comparison group students, rather than the reason being participation in the program alone.

We address this methodological challenge by using a quasi-experimental method that compares SBC students with a comparison group of similar students and that can account for as many of these confounders as possible. Guided by current methodological research on best practices for such studies, we use a specific type of matching process to construct the strongest comparison group possible (Bifulco 2012; Clair, Cook, and Hallberg 2014; Steiner, Cook, and Shadish 2011).[10] Our approach has two features: matches are *local* (treatment and comparison students are from the same setting, to the extent possible) [11] and also *focal* (treatment and comparison students are carefully matched such that they are equivalent on baseline characteristics believed to predict both selection into the program and outcomes of interest).

For this evaluation, we implemented local and focal matching by (1) defining "matching blocks"— that is, unique combinations of cohorts (2015 or 2016 high school graduation years) and postsecondary institutions; and (2) matching each SBC student with one and possibly multiple non-SBC students in his/her block who share similar baseline characteristics. Further, heeding prior research, these baseline characteristics are both empirically linked to the study's key outcomes and also potentially linked to receipt of coaching. These matching criteria yield a large number of matching characteristics, which we translate into estimated *propensity scores*, or the probability of participating in SBC. We describe the matching process below, and provide additional information on local and focal matching in Appendix A.

Estimation of the Propensity Scores

One of the simplest ways to match treatment and comparison groups would be to form matched pairs with the same baseline characteristics (i.e., exact matching). For example, we could match female students in

[10] This line of research uses within-study comparisons (also called *design replication studies*) to inform best quasi-experimental design practices by replacing the randomly determined comparison group in an experiment with comparison groups constructed from units that did not participate in the original experiment, using different quasi-experimental design methods. These studies compare the impact estimates based on comparison groups constructed by quasi-experimental designs versus the experimental impact estimates and try to generalize the specific features of the quasi-experimental designs that replicate (or come closest) to the experimental results.

[11] The matching blocks include students from BPS and nearby districts with similar characteristics to BPS.

the treatment group with female students in the comparison group. Though straightforward, this approach becomes infeasible as the number of characteristics used in the matching increases. Instead, we use propensity score matching, because it allows us to account for a diverse set of background characteristics and experiences within a single measure.

More specifically, a *propensity score* is a number that represents the likelihood of receiving the treatment, based on a student's background characteristics and experiences (Rosenbaum and Rubin 1983). For this study, drawing from MA DESE student-level data, the propensity score represents the likelihood that an individual student participates in SBC, based on the following baseline characteristics:

- **Student demographics**: age, race/ethnicity, gender, free/reduced-price lunch status, disability status, and English language learner (ELL) status
- **Student high school achievement**: GPA, SAT scores, 10th-grade MCAS scores, and number of advanced courses taken in high school
- **Student behavioral measures**: absenteeism and number of suspensions
- **Characteristics of high schools**: college-going rate, average MCAS math and English scores

For a complete list of student and high school characteristics used in the propensity score model, see Exhibit A-4 in Appendix A.

We selected the specific variables listed above based on a comprehensive literature review and on information from coaching organizations about criteria they use when selecting and/or targeting students for their programs. (Exhibit A-3 and Exhibit A-4 in Appendix A summarize key features from the literature review.)

Using this set of characteristics, a propensity score was estimated for each student in the matching blocks, including treatment students and potential comparison group students. Propensity scores can range from 0 to 1, with numbers closer to 1 representing a greater likelihood that a student received the SBC treatment.

Conducting Matching

Once propensity scores were estimated, the next step involved matching SBC students in each matching block with potential comparison group students in the same block. Among the various matching methods, we used *radius matching*, by matching each treatment student with all potential comparison students whose propensity scores were within the pre-specified range ("caliper") of his/her score (± 0.4 of the standard deviation of the propensity scores) in his/her block. We chose this method as our primary method because it balances two important aspects of matching: closeness of the matches and size of the matched groups. Using a caliper ensures that each treatment student is matched with comparison students with sufficiently similar propensity scores. Including all comparison units within the caliper maximizes the size of the analytic sample and statistical power.

Assessing Baseline Balance

After matching SBC students to non-coached students, we checked to see whether the two groups were balanced (i.e., whether the treatment group was similar to the comparison group on background characteristics). Following Steiner et al. (2010) and Rubin (2001), we assessed the similarity ("baseline balance") between the treatment and matched comparison students using the standardized difference in the means of the matching characteristics between treatment and comparison students. We required that the difference be less than 0.15 (or 15 percent of a) standard deviation (SD) in absolute value, which is a

more stringent requirement than one imposed by the U.S. Department of Education's What Works Clearinghouse, which requires baseline differences between quasi-experimental treatment and comparison groups to be less than 0.25 SD to meet Clearinghouse evidence standards (WWC 2014).

Matching and checking of baseline balance continued until satisfactory balance was achieved. When balance was not achieved, the corresponding propensity score model was re-specified (e.g., by including interaction terms or higher-order terms) and the matching and baseline assessment processes were repeated. In order to achieve balance on key demographics, we also impose exact matching by gender (female or not) and by one race category (Black or not). When satisfactory balance was achieved for all matching covariates, the resulting comparison group was treated as final.

Because we are looking at several different outcomes, and *all* the possible data for every single student were not consistently available, the different outcome findings are based on slightly different analytic samples. The primary reason for missing data on certain outcomes (i.e., cumulative GPA) is that such data are provided by particular colleges. Eleven colleges provided data for the evaluation. About half of the students (49 percent) who met the eligibility criteria to be included in the evaluation sample originally enrolled in one of these 11 colleges in the fall immediately after their high school graduation.

To be thorough, we conducted matching and assessed baseline balance separately for each outcome. Exhibit 3-2 shows the averages of the matching characteristics for treatment and potential comparison students and the standardized differences for the two groups *before* matching; it also displays the same statistics *after* matching for one of our primary outcomes, persistence into the second year. Pre-matching differences for some variables were above our desired threshold of 0.15 SD; for example, −0.54 SD for percentage *White* and −1.13 SD for *high school GPA*. However, the last column shows that matching reduced all pre-matching differences over the 0.15 threshold without distorting the balance for the variables balanced prior to matching. By confirming that the two groups are similar on observable characteristics such as these, we can rule out the possibility that these characteristics themselves account for any observed differences in outcome between SBC and comparison group students. The after-matching perfect balance (0.00 SD) on *female* and *Black* characteristics reflects our exact matching on those characteristics.

Exhibit 3-2: Baseline equivalence of treatment versus comparison students for the largest sample

	Before Matching			After Matching		
	Treatment Mean	Comparison Mean	SD	Treatment Mean	Comparison Mean	SD
Student Characteristic						
Age (yrs.)	18.21	18.04	0.16	18.20	18.21	-0.03
Female (%)	0.60	0.54	0.18	0.60	0.60	0.00
English language learner (%)	0.10	0.05	0.17	0.10	0.10	-0.01
Free/reduced-price lunch status (%)	0.75	0.56	0.33	0.74	0.74	-0.01
Student has a high−incidence disability (%)	0.08	0.05	0.08	0.08	0.08	0.01
Student has a low-incidence disability (%)	0.02	0.03	-0.07	0.02	0.03	-0.05
High school GPA (scale of 1-4) (mean)	2.46	2.89	-0.46	2.52	2.54	-0.02
Student took an advanced course (%)	0.57	0.55	0.16	0.56	0.56	0.01
Number of advanced courses taken (mean)	1.07	1.31	-0.01	1.08	1.10	-0.01
SAT score (scale of 400-1600) (mean)	1304.04	1436.20	-0.24	1309.08	1293.43	0.06
10th-grade ELA MCAS standardized scale score	-0.29	0.07	-0.23	-0.28	-0.33	0.05
10th-grade Math MCAS standardized scale score	-0.04	0.18	-0.12	-0.05	-0.03	-0.02
School days student was present (mean)	93.08	93.88	-0.06	93.08	93.05	0.00

	Before Matching			After Matching		
	Treatment Mean	Comparison Mean	SD	Treatment Mean	Comparison Mean	SD
Number of suspensions (mean)	0.18	0.34	-0.09	0.19	0.18	0.01
Race/Ethnicity (%)						
Black	0.44	0.24	0.41	0.42	0.42	0.00
White	0.07	0.35	-0.54	0.08	0.08	0.00
Asian/Pacific Islander	0.15	0.18	-0.07	0.16	0.17	-0.03
Hispanic	0.32	0.21	0.21	0.33	0.31	0.03
Native American	0.00	0.00	0.01	0.00	0.00	-0.01
Other/Multiracial	0.01	0.02	-0.02	0.01	0.01	-0.01
High School Characteristics						
College-going rate (%)	0.68	0.67	0.13	0.67	0.67	-0.01
10th-grade Math MCAS standardized scale score (mean)	-0.42	-0.14	-0.48	-0.40	-0.38	-0.04
10th-grade ELA MCAS standardized scale score (mean)	-0.18	-0.04	-0.21	-0.17	-0.16	-0.03
GPA (mean)	2.25	2.73	-1.13	2.32	2.37	-0.11

ELA=English language arts; GPA=grade point average; MCAS=Massachusetts Comprehensive Assessment System; SAT=Student Achievement Test; SD=standard deviation.
Source: Program database; Boston Public Schools; Massachusetts Department of Elementary and Secondary Education.
Note: Based on the analytic sample for persistence into the second year.

3.1.3 How We Estimate the Average Impact of the Program for the Full Sample

To address the primary research question about the impact of SBC on all students, we estimate a linear regression model that includes indicators for the matching blocks (defined based on student cohorts and postsecondary institutions) and the 19 matching characteristics listed in Exhibi 3-2. To maximize statistical power, the model pools the two cohorts (students entering college in the falls of 2015 and 2016). The model is estimated separately for each outcome measure with the corresponding matched treatment and comparison groups.[12] The model includes, as covariates, all matching characteristics used to construct the corresponding comparison group to increase precision of the impact estimates and be doubly robust (Bang and Robins 2005; Tan 2006).[13] The models did not explicitly adjust the standard errors for the clustering of students within postsecondary institutions because we anticipated that such clustering was captured by the matching block indicators.[14]

We address the issue of *multiple hypothesis* testing (or the increased likelihood of finding a spurious effect as the number of tests increases) by (1) placing outcomes into larger groups ("domains"); (2)

[12] The model for persistence into the third year includes only the Fall 2015 cohort because this outcome was available only for that cohort at the time this report was prepared.

[13] Using the baseline characteristics in the matching process and also using them as covariates in the estimation of impacts yields a consistent estimator if *either* model is correct. That is, if the weights implied by matching are wrong, but the regression model is right, the estimator is unbiased but inefficient; if the regression model is wrong but matching is correct, the estimate has excess variance but is consistent. Thus the combination is deemed to give the analyst two chances to get the "right" model specification (once in the propensity model and once in the impact model for the outcome measure). Therefore, these estimators are called "doubly robust," in the sense that they are robust to either of two types of mistakes.

[14] We tested the validity of this assumption by estimating hierarchical linear models that nest students within colleges. These models yielded virtually identical estimates.

specifying the primary (confirmatory) outcomes in each domain; and (3) adjusting the statistical significance of impact estimates for primary outcomes using the Benjamini-Hochberg method for domains with multiple primary outcomes (Benjamini and Hochberg 1995; Schochet 2008). We conducted sensitivity tests with different covariate sets; all of which yielded similar results (see Appendix B for more information).

3.1.4 How We Conduct Exploratory Analyses of Program Impacts

The third research question pertains to *student characteristics and features of coaching* that may be related to potential variation in the impact of SBC. We examined student baseline characteristics (e.g., students' demographic attributes and high school academic performance) and coaching features such as frequency of coach-student interactions and duration of one-on-one interactions. These variables explore variation in the strength of the impacts; for example, the impacts of SBC may be *greater* when students experience more coaching interactions. To simplify the analyses and ease the interpretation of results, we transformed each continuous or categorical variable into a binary variable representing two subgroups that differed by the value of that characteristic or feature (e.g., fewer vs. more coaching interactions). We then calculated separate impact estimates for the two subgroups and assessed the magnitude and statistical significance of the *difference* in the subgroup-specific impact estimates.

We consider these to be *exploratory* analyses because (1) the subgroup analyses have less statistical power than full-sample analyses; and (2) the programmatic features are post-treatment measures that may reflect program impacts (e.g., frequency of coaching interactions only occurs post-treatment and may be a function of the effectiveness of coaching itself). Thus, the differences in effects for the corresponding subgroups may not be fully attributable to those features of coaching. Further, given the number of subgroups explored, it is possible that any statistically significant impacts found may be due to chance variation and are not true impacts. See Appendix B for an in-depth description of the analytic approaches used in these analyses.

3.2 Data Sources

The analyses rely on data from multiple sources: National Student Clearinghouse, Massachusetts Department of Elementary and Secondary Education, colleges in which students enrolled, and the SBC program database.

Massachusetts Department of Elementary and Secondary Education (MA DESE) provided student data for the entire state, including high school academic measures, behavior, and demographic information (e.g., SAT, 10th-grade MCAS scores; coursework; absences and suspensions; race/ethnicity, gender).

The **National Student Clearinghouse (NSC)** is a nonprofit organization that regularly collects enrollment and graduation information from colleges across the country. As of the fall of 2017, the NSC included student-level data on 96.8 percent of U.S. colleges and 98.6 percent of Massachusetts colleges. Using NSC data allows us to access data for all students regardless of whether or not they transfer colleges. For this evaluation, data on students' college enrollment and graduation come from the NSC, by way of BPS and MA DESE.

College administrative data were collected from 11 colleges with more than 10 enrolled SBC students annually or strong partnerships with the Success Boston initiative. These administrative data on students include individual-level student records on college enrollment and persistence, academic achievement,

and FAFSA renewal. Of the eligible treatment students, 74 percent were enrolled in these institutions immediately after graduating high school:

- Benjamin Franklin Institute of Technology,
- Bridgewater State University,
- Bunker Hill Community College,
- Framingham State University,
- Massachusetts Bay Community College (MassBay),
- Northeastern University,
- Roxbury Community College,
- Salem State University,
- Suffolk University,
- University of Massachusetts Boston (UMass Boston), and
- Wentworth Institute of Technology.

SBC program data, stored in Success Boston's Salesforce™ database,[15] include program participation records for individual students who participate in SBC. In this case, the SBC program data include information about each coach-student interaction for all students served by each nonprofit coaching organization. Each individual student record has information on the following data elements: nonprofit coaching organization, assigned coach, whether the student is actively receiving coaching or has never enrolled, at which college the student is enrolled, the type of support provided to a student during a coaching session (i.e., academic, personal and emotional, financial, etc.), the duration of a given coaching session, and whether the coach provided direct or indirect support (i.e., help completing a task versus referral to campus service).

3.3 Measures

In this section, we first describe the outcomes and then define the measures used in our exploratory analyses to test whether program impacts varied depending on student characteristics and features of coaching.

3.3.1 Outcome Measures

The outcomes for this evaluation fall into three domains: persistence, achievement, and financial aid. These outcomes are operationalized below. For domains with multiple outcomes measures, we further distinguish between *primary* and *exploratory* outcomes.

- **Primary outcomes** are those most closely related to the theory of change, which hypothesizes that the elements of one-on-one coaching that together address logistical,

[15] Salesforce™ is a cloud-based client relationship management database used widely in both for-profit and nonprofit sectors. Internally, SBC program data is known as "Salesforce data."

academic, financial, and emotional support topics can improve persistence (and, ultimately, completion) rates for traditionally underrepresented college students.

- **Exploratory outcomes** are also informed by the theory of change, as they may help explain why or why not impacts are detected on the primary outcomes, the most important of which is college completion (e.g., full-time status is not an outcome in and of itself, but it is useful to examine because it is related to persistence and, ultimately, completion).

For the purposes of standardizing the amount of time for both 2015 and 2016 high school graduating classes, we restricted outcomes to the two-year period following high school graduation, with one exception: we measured persistence into the third year, about two and a half years after high school graduation. Exhibit 3-3 lists details about each of the outcome measures. Wherever possible, data from colleges supplements NSC data, to ensure that we limit the number of students for whom outcome data are missing.[16]

Exhibit 3-3: Outcome domains and measures in this report

Domain	Outcome Measure	Primary or Exploratory	Years Post–High School	Sample Size	Data Source
Persistence	Persistence into the second year	P	1.5	10,371	NSC, college administrative data
	Persistence into the third year	P	2.5	5,016	NSC, college administrative data
	Full-time status	E	2	4,498	College administrative data
Achievement	Cumulative GPA	P	2	4,219	College administrative data
	Semesters enrolled in non-credit-bearing courses	E	2	4,498	College administrative data
	Credit accumulation	E	2	4,674	College administrative data
Financial aid	FAFSA renewal	P	2	4,259	College administrative data

GPA=grade point average; NSC=National Student Clearinghouse.

Persistence Outcomes

This study uses two primary outcome measures to assess the persistence domain: persistence into the second year and persistence into the third year of college. One exploratory outcome is included: full-time status.

Primary Outcomes

Persistence into the second year is a measure of whether students are enrolled in college in the fall (on October 1) of their second year after high school, according to NSC data or college data. Persistence into the second year is a binary outcome (yes/no), in which a student enrolled in the fall semester immediately following high school graduation *and* the fall semester of the subsequent academic year is coded as

[16] Though NSC includes data on almost every college nationally (96.8 percent) and in Massachusetts (98.6 percent), individual student records from any given college may be missing because students may block the release of their records under the Family Educational Rights and Privacy Act (FERPA) or because of matching errors due to misspellings of student names. NSC reports that, on average, 4.3 percent of students block the release of their records. Research on the extent of matching errors shows the NSC algorithm to be robust to student name variants (Dynarski, Hemelt, and Hyman 2013).

persisting. For example, a 2015 high school graduate would be counted as persisting into the second year if enrolled in college in the fall of 2015 *and* in the fall of 2016. All students enrolled in the fall of the second year, regardless of where they enrolled, are counted as persisting. This means that students who transferred colleges between fall 2015 and fall 2016 are considered to have persisted into the second year. Additionally, any students who graduated before the second fall are included as positive outcomes on this measure. Students whose records were not present in either dataset in the fall of their second year after high school graduation, and had not previously graduated from a college, are assumed not to have been enrolled in college at that time point.

Persistence into the third year is a measure of whether students are enrolled in college in the fall (on October 1) of their third year after high school, according to NSC data or college data. Consistent with the approach described above, persistence into the third year is a binary outcome (yes/no) in which a student who enrolled in the fall immediately following high school graduation and the fall semester of the *second* subsequent academic year is coded as persisting.[17] For example, for a 2015 high school graduate to be counted as persisting into the third year, she would be enrolled in the fall of 2015 and again in the fall of 2017. Similar to persistence into the second year, all students enrolled in the fall of the third year, regardless of where they enrolled, are counted as persisting. This means that students who transferred colleges between fall 2015 and fall 2017 are considered to have persisted into the third year. Additionally, any students who graduated before the third fall are included as positive outcomes on this measure. Students whose records were not present in either dataset in the fall of their third year after high school graduation, and had not previously graduated from a college, are assumed not to have been enrolled in college at that time point.

Exhibit 3-4 displays how each persistence measure is defined. It is important to note that persistence into the third year could be measured only for the high school Class of 2015, as data about third-year persistence for the 2016 graduates were not available in time for inclusion in this report.

Exhibit 3-4: Annual persistence measures

College-Entering Cohort	Persisted into the Second Year	Persisted into the Third Year
2015	Enrolled in fall 2015 and fall 2016	Enrolled in fall 2015 and fall 2017
2016	Enrolled in fall 2016 and fall 2017	NYA

NYA=data not yet available

Exploratory Outcome

Semesters enrolled full-time refers to the percentage of semesters in which students were enrolled full-time, relative to the four semesters of full-time enrollment possible since high school graduation. This measure is based on data provided by 11 partner colleges, and therefore the outcome is estimated only for students who attended those 11 colleges.[18] Values range from 0 to 1, where 0 indicates a student was never enrolled full-time and 1 represents a student who has always been a full-time student.

[17] Students who did not enroll in college in the second year after high school can still be considered to have persisted into the third year. Therefore, this outcome does not measure continuous enrollment to the third year.

[18] Sixteen students enrolled elsewhere for any portion of the two years following high school graduation are coded as missing data for this measure.

Achievement Outcomes

The evaluation focuses on one primary outcome measure of student achievement: cumulative GPA. Two exploratory outcomes are included as well: semesters enrolled in non-credit-bearing courses and credit accumulation. Again, these outcomes are estimated only for students who attended the 11 colleges providing data; students enrolled elsewhere for any portion of the two years following high school graduation are coded as missing data.

Primary Outcome

Cumulative GPA is a continuous measure that ranges from 0 to 4, where A=4.0, B=3.0, C=2.0, and D=1.0, reflecting data current as of the most recent semester completed.

Exploratory Outcomes

Semesters enrolled in non-credit-bearing courses describes the number of fall and spring semesters in which college students had enrolled in non-credit-bearing courses (i.e., courses that did not count toward graduation). This measure ranges from 0 to 4, representing the total number of semesters in which it was (theoretically) possible for students to enroll in such courses.

Credit accumulation is a count of the total number of credits successfully completed, as of the students' most recent semester. The final score represents a proportion of the total number of credits completed, divided by the total number of credits needed to graduate. The number of credits needed to graduate varies from school to school; we followed the graduation requirements set by each institution.

Financial Aid Outcomes

FAFSA renewal indicates whether students completed and submitted a Free Application for Federal Student Aid renewal form in their first year of college. Only data for the first year of college was considered, in order to standardize the amount of time for the 2015 and 2016 cohorts. This measure is binary such that 0 corresponds to students *not* having renewed their FAFSA application, and 1 corresponds to students having renewed the application. This outcome is estimated for students who attended the nine colleges providing data for the evaluation;[19] students enrolled outside of these colleges for any period of time during the two years following high school graduation are coded as missing data for this measure.

3.3.2 Student Characteristics and Features of Coaching

We conducted several exploratory analyses to gain a deeper understanding about whether, and if so, how, observed impacts of SBC vary as a function of particular student characteristics or features of coaching. These analyses are exploratory for two reasons: (1) They investigate impacts on subsets of the sample; because the overall study sample has been divided into subgroups, the statistical analyses may be less able to detect educationally meaningful program impacts than analyses based on the full sample. (2) The

[19] Massachusetts Bay Community College did not provide FAFSA renewal data for either the 2016-17 or 2017-18 academic year. Suffolk University provided data for the 2016-17 academic year but not for the 2017-18 academic year.

programmatic features occur only after students have enrolled and begun to participate in the coaching program; features may reflect students' experiences with coaching or other events, which makes it difficult to distinguish between differences in outcomes driven by the features of the coaching experienced by students (variation in program impacts) versus other factors that do not reflect true program impacts.[20]

Student Characteristics

We dichotomized variables to maximize the sizes of the subgroups and, therefore, statistical power and to generate easily interpretable comparisons between subgroups. We explored four student characteristics. For each characteristic, we formed two subgroup categories. Created from MA DESE data, student characteristics explored are (1) *gender* (categorized as male or female); (2) *underrepresented minority* (categorized as traditionally underrepresented in postsecondary education for Black, Hispanic, Native American, or Other/Multiracial students, or categorized as not traditionally underrepresented for White and Asian/Pacific Islander students); (3) *student GPA* (categorized as high or low, based on the median of the GPA distribution, where higher is greater than 3.00 and lower is equal to or less than 3.00); and (4) *type of college* in which a student first enrolls (two-year or four-year institution).

These specific characteristics were selected because previous research indicates they are related to college completion, the ultimate goal of SBC. For example, female students complete college at higher rates than male students (Shapiro et al. 2015); underrepresented minority group students complete college at lower rates than students not underrepresented in postsecondary education (U.S. Department of Education 2016; Haskins 2008; Bailey and Dynarski 2011); students with higher GPAs complete college at higher rates than students with lower GPAs (Belfield and Crosta 2012); and students first enrolling at four-year institutions complete a college degree at higher rates than students first enrolling at two-year institutions (Shapiro et al. 2015).

Features of Coaching

We explored four features of coaching: (1) *content focus*, (2) *frequency*, and (3) *duration*. These features reflect findings described in the previous implementation study (Linkow et al. 2015), which highlighted them as important features of coaching that vary across students and organizations. Data for each of these variables come from the SBC program database. For each feature of coaching, we divided students into

[20] Students may experience coaching differently for multiple reasons—many of which can plausibly be accounted for by the matching process used in this study. However, because the matching uses data collected *before* coaching starts, it is still possible that either or both of the following could be related to the features of coaching *and* the outcomes of interest: (a) student characteristics that were not captured in the matching (e.g., responsiveness to coaching); or (b) students' experiences that occur after matching (e.g., success with college coursework). That possibility means we cannot causally link differences in features of coaching to differences in outcomes. For example, a student may struggle in his first year in college, for reasons completely separate from coaching and/or the specific characteristics used in the matching, and he may seek out coaching support more often than another student who is not struggling. In his case, more coaching would be related to poor college outcomes. On the other hand, a student who is succeeding academically, also for reasons that have nothing to do with coaching and characteristics used in the matching, may seek out coaching more frequently to learn how she can do even better. In his case, more coaching would be related to more positive outcomes. In both cases, students are self-selecting into more frequent coaching, yet their outcomes are caused by factors unrelated to the effectiveness of coaching.

two groups, high and low, based on the median value. Exhibit 3-5 displays the low and high cutoffs for each feature.

- ***Content focus*** measures how many coach interactions a student experiences that were focused on one of four topic areas: academic, financial aid, career planning, or managing life responsibilities. This feature of coaching captures interactions across a two-year time period (June 15, 2015, through June 15, 2017, for the 2015 cohort; and June 15, 2016, through June 15, 2018, for the 2016 cohort), which spans two academic years and includes the summer in between. Four count variables were created, one for each of the topic areas.
- ***Frequency*** measures the number of interactions across all modes (in-person, text, email, phone, and social media) between students and their coach across the same two-year time period as described above.
- ***Duration*** measures the average length (in minutes) of all one-on-one interactions (in person or phone) over the same two-year time period as described above.

Exhibit 3-5: Subgroup cutoffs for features of coaching

Feature	Lower Range	Higher Range
Content focus		
Academic focus	0-9 meetings	> 9 meetings
Financial aid focus	0-2 meetings	> 2 meeting
Career focus	0-1 meetings	> 1 meeting
General life focus	0-1 meetings	> 1 meetings
Frequency	0-14 meetings	> 14 meetings
Duration	0-32.5 minutes	> 32.5 minutes

3.4 Limitations

The study faces some methodological limitations: (1) data availability, (2) matching students across high schools and school districts, and (3) its use of a quasi-experimental design rather than an experimental design.

First, data are available only on certain outcomes for students in the evaluation sample who were enrolled in the 11 colleges that provided data. Students who did not enroll in these 11 colleges or transferred from one of the 11 to college not on that list are excluded from the models. Because of this, we allowed the comparison group to shift with availability of outcome data, and we show in Appendix Exhibit A-7 that baseline equivalence is achieved for each outcome. The sensitivity analysis presented in Appendix Exhibit B-2 shows that the impacts on persistence outcomes, which are based on the most complete samples (from NSC data), tell a similar story as those impacts found when limiting the sample to only students enrolled in the 11 colleges submitting data for the evaluation.

A second limitation is that students within a given college were matched across high schools and school districts. Because sample sizes were too small to allow for matching students from the same high school attending the same college, matching is within colleges and accounts for features of high schools. Even with this expanded pool of potential comparison students, in some cases it was not possible to find a similar comparison student to match to a coached student (see Appendix Exhibit A-5 for details about the treatment group match rate by outcome). Thus, our impact estimates are estimates of the effect of SBC on students who could be matched—these are more likely to be students at colleges with more students from BPS and more students like the typical SBC student. The matching process addresses both differences in

college experiences and high school characteristics to eliminate historical and locational differences (bias) in students' previous educational experiences.

Third, it is possible that the local and focal matching approach did not sufficiently control for potentially confounding factors because we were not able to use an experimental (random assignment) design. To the extent that the distribution of all important confounders is equalized across participants and the matched comparison group, our quasi-experimental design should produce impact estimates with minimal bias and good power, relative to other quasi-experimental designs.

4. How Does Participation in Coaching Affect Early College Outcomes?

This chapter presents the results from the study's impact analyses, which assess whether participation in coaching leads to better early college outcomes for coached students. The SBC program offers students one-on-one support from a coach about diverse transition and college success topics, including life skills, study skills, help-seeking skills, accessing financial aid, time-management strategies, setting academic and career goals, understanding college culture, and balancing school, work, and life. We focus specifically on program impacts for fall 2015 and 2016 college entrants, who graduated from high school three and two years ago, respectively. The findings are organized according to three outcome domains: persistence, achievement, and financial aid.

Key Findings

The analyses estimate that SBC students have more-positive early college outcomes than their peers not participating in SBC. Specifically, SBC students are

- more likely to persist into their *second* year of college (82.6 percent for the SBC treatment group vs. 78.3 percent for the non-coached comparison group);
- enrolled full-time for more semesters (67.3 percent vs. 63.6 percent of four possible semesters);
- accumulating more college credits (of the credits necessary to graduate at their college, completed 45.5 percent vs. 42.5 percent)
- more likely to complete FAFSA renewals for their second year of college (83.9 percent vs. 77.7 percent).

Our analytic approach pools the estimated impact of SBC on student persistence, academic achievement, and financial aid across both cohorts of students. Results for one outcome (persistence into the third year of college) are presented only for the fall 2015 cohort, as data for the 2016 cohort were not available in time for this report. To measure outcomes consistently across these two cohorts, all measures (except persistence into the third year) are calculated for the first two years of expected college enrollment. The exhibits that follow present, for each outcome, the treatment group mean (in purple) below the comparison group mean (in blue). The comparison group is weighted (i.e., regression adjusted) based on baseline student characteristics to represent the mean outcomes for the treatment group had they not received the intervention. As such, the adjusted comparison group means represent the mean outcomes that would have been observed for the treatment group in the absence of SBC.

The following sections detail results for primary and exploratory outcomes for each domain.

4.1 Persistence

A key expectation of the SBC model is that students are more likely to persist in college as coaches help them navigate and manage the academic, financial, and social-emotional challenges typically faced by beginning college students. To test this hypothesis, this evaluation uses two primary outcome measures to assess persistence: persistence into the second year of college and persistence into the third year of college. These persistence outcomes measure whether students who enrolled in college after their high school graduation returned to college in the fall of succeeding academic years. Persistence into the *second year* of college indicates whether students in the 2015 and 2016 cohorts enrolled in the fall semester of 2016 and 2017, respectively. Persistence into the *third year* of college reflects whether students in the 2015 cohort also enrolled in the fall semester of 2017.

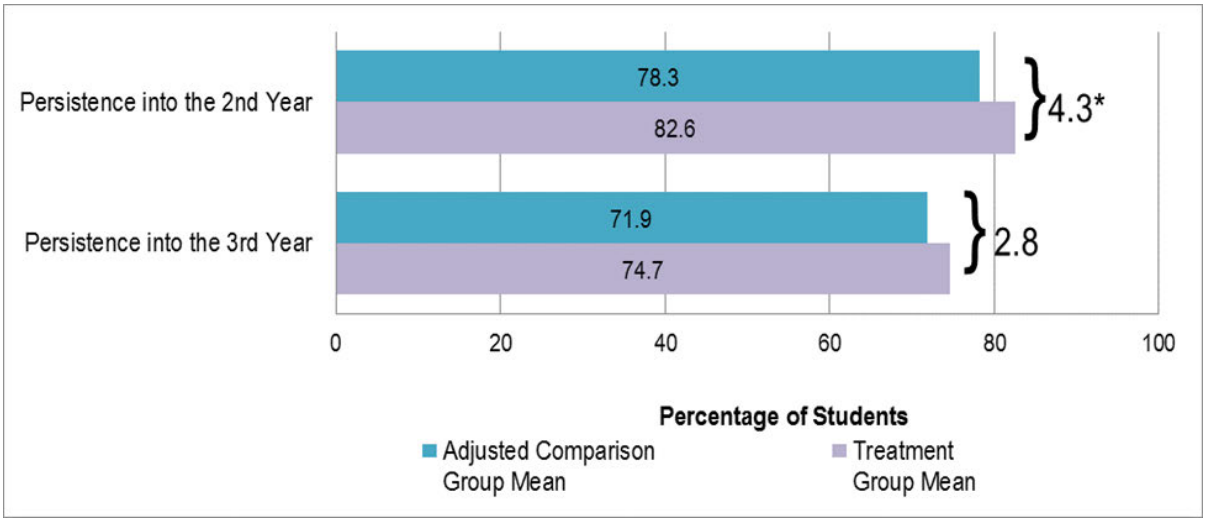
We examined the exploratory outcome of semesters enrolled full-time to assess more fine grained aspects of students' continued persistence toward degree completion. This outcome refers to the proportion of the first four semesters in which students were enrolled full-time after high school graduation.

4.1.1 Persistence into the Second and Third Years of College

The top panel of Exhibit 4-1 below shows that 82.6 percent of SBC students in the treatment group and 78.3 percent of non-coached students (comparison group) persisted into the second year of college. The 4.3 percentage point impact of SBC on persistence is positive and statistically significant. This percentage point change is the equivalent of a five percent increase in persistence rates. Both groups persisted into the second year at higher rates than the 73.9 percent national average for students starting college in the fall of 2016 (NSC 2018).

The bottom panel in Exhibit 4-2 shows that the impact of SBC on persistence into the third year is not significant. (This third-year outcome is measured only for the 2015 cohort, so it should be interpreted with caution due to the smaller sample size.)[21]

Exhibit 4-1: Impact of SBC on persistence into second and third years of college



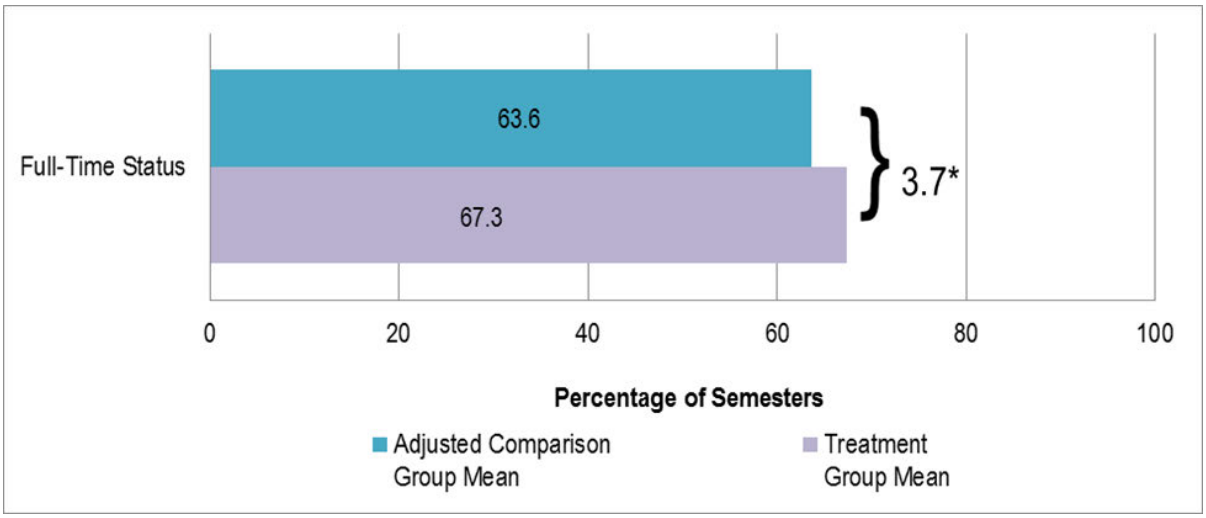
Source: National Student Clearinghouse data from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education.
Note: N=5,863 for overall sample (N=1,234 for treatment and N=4,629 for comparison) for persistence into the second year. N=2,719 for overall sample (N=561 for treatment and N=2,158 for comparison) for persistence into the third year. Students are considered to persist if they enroll year to year or if they complete a certificate. (Completers by second year in college T=0, C=6. Completers by third year in college T=8, C=29.) Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
* Impact is significant at the 5 percent level. Significance determined with the Benjamini-Hochberg correction.
Exhibit Reads: The impact of coaching on persistence into the second year of college is statistically significant at 4.3 percentage points. 82.6 percent of students who participated in coaching persisted into the second year of college, whereas 78.3 percent of students who did not participate in coaching persisted into the second year of college. The impact of coaching on persistence into the third year of college is not statistically significant.

[21] Because two primary outcomes within the single domain of persistence are tested, statistical significance tests use the Benjamini-Hochberg correction.

4.1.2 Semesters Enrolled Full-Time

Because they likely accumulate college credits faster, students who are enrolled full-time may be more likely to complete a degree within a shorter time period than those enrolled part-time. SBC has a positive and statistically significant impact of 3.7 percentage points on the percentage of semesters in which students are enrolled full-time (Exhibit 4.2). This percentage point change is the equivalent of a six percent increase in semesters enrolled full-time.

Exhibit 4-2: Impact of SBC on semesters enrolled full-time



Source: College administrative data.
Note: N=3,748 for overall sample (N=864 for treatment and N=2,884 for comparison). Calculated as the percentage of semesters in which students are enrolled full-time out of the total possible number of semesters they could have been enrolled (four semesters). Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
* Impact is significant at the 5 percent level.
Exhibit Reads: The impact of coaching on proportion of semesters spent enrolled full-time is statistically significant at 3.7 percentage points. For students who participated in coaching, the mean percentage of semesters spent enrolled full-time was 67.3 percent, whereas the mean was 63.6 percent for students who did not participate in coaching.

The positive findings across all persistence outcomes are noteworthy for several reasons. First, these corroborate prior findings in the Success Boston interim impact report (Linkow et al. 2017b) as well as research conducted by Sum, Khatiwada, and Palma (2014), which found evidence of a positive and statistically significant effect of Success Boston Coaching on college persistence. Second, the findings are consistent with evaluations of other coaching interventions shown to improve students' likelihood of persisting in college (e.g., Bettinger and Baker 2014). And third, persistence is an important success indicator; students must persistent in order to complete.

4.2 Achievement

Another key hypothesis of the SBC evaluation is that coaching supports can help students manage various academic challenges that they may face in college, ranging from difficult coursework to course selection and time management. By helping students access available campus supports and cope with stressors related to the academic demands of college, coaches can potentially help students improve their academic achievement.

The analyses described below examine the impact of SBC on one primary outcome measure of academic achievement: cumulative GPA. In addition, we explore impacts for two exploratory outcomes: semesters enrolled in non-credit-bearing courses and credit accumulation.

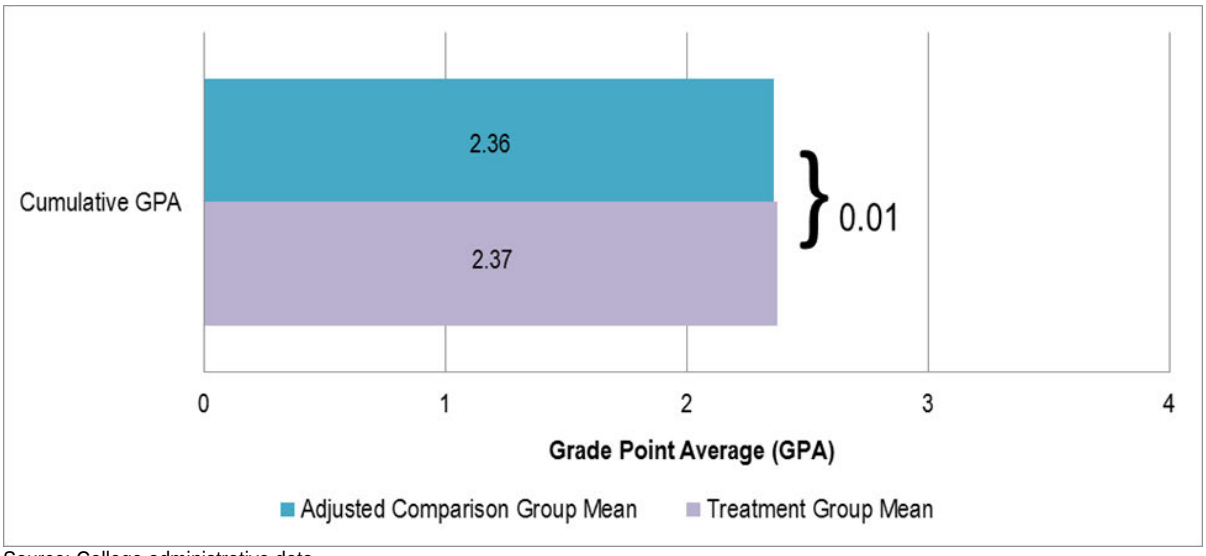
The first outcome, cumulative GPA, is drawn from students' most recent semester enrolled and corresponds with the following values: A=4.0, B=3.0, C=2.0, D=1.0. The second outcome examines the number of semesters for which students enrolled in non-credit-bearing courses.

To explore how students' academic achievement profiles suggest they are moving along the path toward completion, the third outcome investigates credit accumulation. Because the number of credits needed to graduate varies from school to school, we define credit accumulation as the total number of credits successfully completed, divided by the total number of credits needed to graduate at that student's college.[22]

4.2.1 Cumulative GPA

Exhibit 4-3 illustrates that SBC students in the treatment group have comparable cumulative GPAs, on average, relative to non-coached students in the comparison group. Students in both the treatment and comparison groups had GPAs that fell in the C/C+ range (approximately 2.4 points on a four-point scale).

Exhibit 4-3: Impact of SBC on cumulative GPA



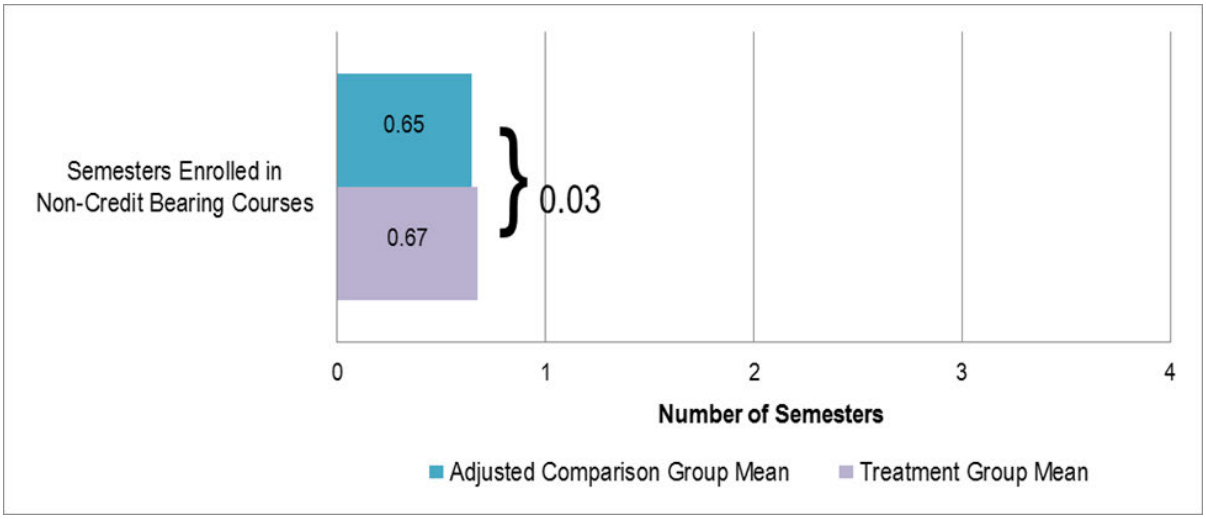
Source: College administrative data.
Note: N=3,553 for overall sample (N=814 for treatment and N=2,739 for comparison). Taken as cumulative GPA from the most recent GPA for which we had data available. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
Exhibit Reads: There is no impact of coaching on cumulative GPA. For students who participated in coaching, the mean cumulative GPA was 2.37 versus 2.36 for students who did not participate in coaching.

[22] When credits necessary to graduate vary by major or school within a college, credits necessary to graduate from the most common major or largest school are used.

4.2.2 Semesters Enrolled in Non-Credit-Bearing Courses

There is no statistically significant impact of SBC on the number of semesters during which students were enrolled in non-credit-bearing courses (i.e., "developmental courses" that confer no credit toward completion of a degree or certificate). Both SBC students and non-coached students spend less than one semester enrolled in non-credit-bearing courses (Exhibit 4-4). Because this outcome measures a potential adverse outcome,[23] a negative estimated impact would mean that SBC students are spending *less* time in non-credit-bearing courses, which in turn, would mean a positive outcome for the SBC program.

Exhibit 4-4: Impact of SBC on semesters enrolled in non-credit-bearing courses



Source: College administrative data.
Note: N=3,748 for overall sample (N=864 for treatment and N=2,884 for comparison). Actual credits needed for graduation vary by college. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
Exhibit Reads: There is no impact of coaching on the number of semesters enrolled in non-credit-bearing courses. For students who participated in coaching, the mean number of semesters enrolled in non-credit-bearing courses was 0.67 versus 0.65 semesters for students who did not participate in coaching.

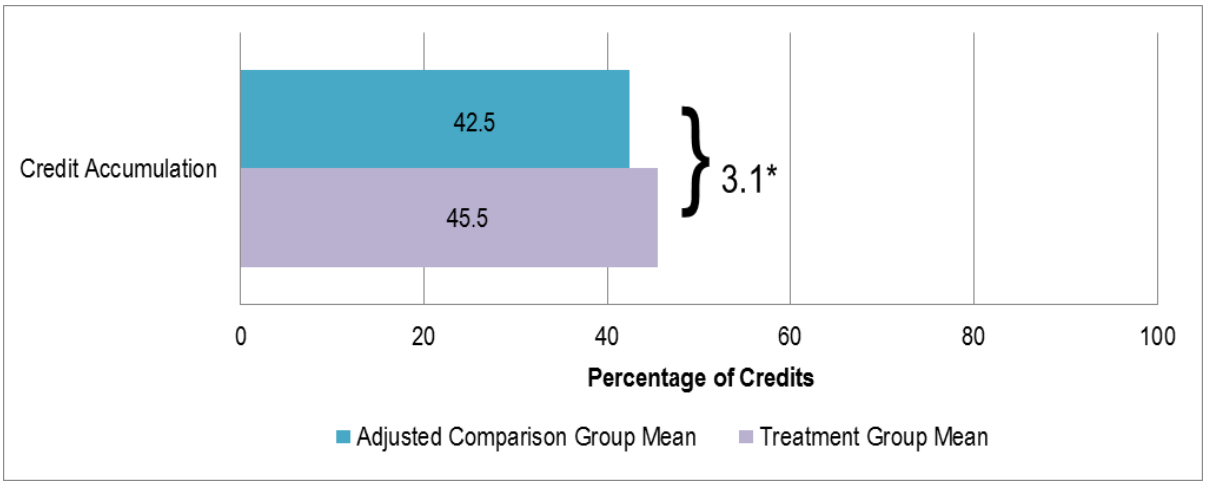
4.2.3 Credit Accumulation

Exploratory analysis of credit accumulation demonstrates that SBC students are completing more credits toward graduation than are their peers in the comparison group (Exhibit 4-5). Two years after enrolling in college, SBC students and non-coached students had completed 45.5 percent and 42.5 percent, respectively, of the credits needed to graduate at their college. The impact is estimated to be 3.1 percentage points. This percentage point change is the equivalent of a seven percent increase in credits accumulated. SBC students may have accumulated slightly more credits toward graduation; however, like their non-coached peers, SBC students , on average, have not accumulated sufficient credits to graduate within 100 percent time (i.e., "normal time" which is two years for those pursuing Associate's degrees and four years for those pursuing Bachelor's degrees). The Success Boston initiative focuses on graduation within 150% of normal time (three years for two-year colleges and six years for four-year

[23] Bailey, Jeong, and Cho (2010) estimate that one third of community college students who are referred to a developmental math sequence fail to complete the sequence, and consequently do not complete a degree or credential.

colleges). The study will continue to monitor students' progress over the next several years to examine completion within 150% of normal time.

Exhibit 4-5: Impact of SBC on credit accumulation



Source: College administrative data.
Note: N=3,806 for overall sample (N=870 for treatment and N=2,936 for comparison). Actual credits needed for graduation vary by college. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
* Impact is significant at the 5 percent level.
Exhibit Reads: The impact of coaching on percentage of credits accumulated toward graduation is statistically significant at 3.1 percentage points. SBC students' mean percentage of credits accumulated toward graduation was 45.5 percent, compared with 42.5 percent for students who did not participate in coaching

4.3 Financial Aid

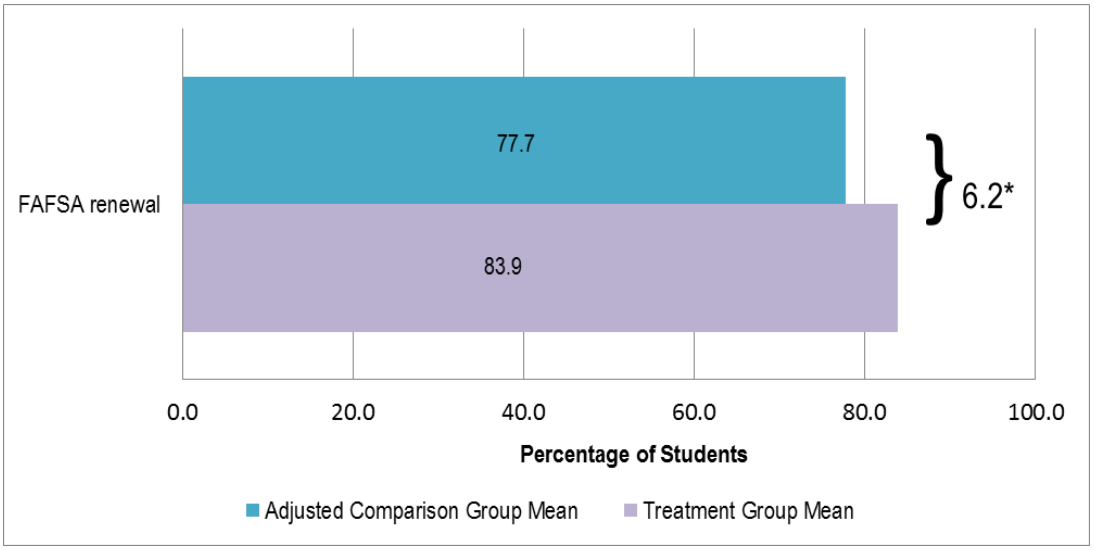
One of the goals of SBC is to help students navigate the financial aid process and presumably help reduce the gap between the cost of college and what students can actually afford. In fact, uAspire offers training for SBC coaches on a range of financial aid and planning topics and hosts FAFSA completion events for SBC students. In 2016-17, some 67 percent of SBC students received the text message nudges from uAspire advisors. To this end, the SBC evaluation examines Free Application for Federal Student Aid renewal, measuring whether students complete and submit a FAFSA renewal application for their second year of college, unlocking access to federal student aid.

4.3.1 FAFSA Renewal

A large majority of both SBC and non-coached students renewed their FAFSA for their second year of college, reflecting the high proportion of students who are eligible for federal financial aid in the study sample (74 percent of SBC students and 75 percent of those not coached were eligible for free/reduced-price lunch in high school).

SBC students in the treatment group renewed their FAFSA at a higher rate than the students in the comparison group, and the difference in renewal rates (6.2 percentage points) is positive and statistically significant (Exhibit 4-6). Among the treatment group students, 83.9 percent renewed their FAFSA for their second year of college versus 77.7 percent of comparison students. This percentage point change is the equivalent of an eight percent increase in FAFSA renewal rates.

Exhibit 4-6: Impact of SBC on FAFSA renewal



Source: College administrative data.
Note: N=3,597 for overall sample (N=821 for treatment and N=2,776 for comparison). Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
* Impact is significant at the 5 percent level.
Exhibit Reads: The impact of coaching on FAFSA renewal is statistically significant at 6.2 percentage points. Among students who participated in coaching, 83.9 submitted FAFSA renewal forms for their second college year, compared with 77.7 percent of students who did not participate in coaching.

4.4 Learning Points

The one-on-one transition supports provided by SBC across a wide-ranging set of topics are hypothesized to improve student outcomes along several important dimensions, including how long students persist in college, their academic achievement while in college, and their awareness that applications are required in order to access available federal financial aid. Analysis of early outcomes explored several mechanisms through which SBC can increase the college completion rates of Boston students.

Relying on a rigorous, quasi-experimental design the results presented in this chapter demonstrate that SBC has statistically significant, positive impacts on students across three domains: persistence, achievement, and financial aid. Results for two out of four primary outcomes indicate positive statistically significant impacts, and impacts for two of three exploratory outcomes are also positive and statistically significant.

Exhibit 4-7 below summarizes the magnitude of impacts of SBC on each outcome. The magnitude is expressed as the percent change for each outcome (i.e., the percent increase divided by the adjusted comparison mean). Findings indicate that SBC students persist into the second year of college at a rate that is five percent higher than that of non-coached students. They also renewed their FAFSA at a higher rate (eight percent) than the students in the comparison group. The impacts on semesters enrolled full-time and credit accumulation represent six and seven percent increases, respectively.

Exhibit 4-7: Summary of impact results

Domain	Outcome Measure	Percent Change
Primary Outcome		
Persistence	Persistence into the second year	+5%*
	Persistence into the third year	+4%
Achievement	Cumulative GPA	0%
Financial Aid	FAFSA renewal	+8%*
Exploratory Outcome		
Persistence	Full-time status	+6%*
Achievement	Semesters in non-credit-bearing courses	+5%
	Credit accumulation	+7%*

+ indicates a positive result; – indicates a negative result.
* Statistically significant impact.

5. How Are Impacts Related to Student Characteristics?

Overall, SBC has positive and significant impacts on measures of students' college persistence, academic achievement in college, and access to financial aid, as discussed in Chapter 4. This chapter explores how the average impacts on primary outcomes hold for particular subgroups of students. For example, SBC may have different impacts on students in two-year colleges than those in four-year colleges. This chapter examines how program impacts are related to four potential student characteristics, listed in Exhibit 5-1.

Key Findings

The results from these exploratory analyses indicate that compared with their non-coached peers, SBC students who fall into one or more of the subgroups listed below have more-positive outcomes.

- Males, students from racial/ethnic groups traditionally underrepresented in postsecondary education, those with low GPAs in high school, and students first enrolled in four-year colleges persist into the second year at a higher rate than do non-coached peers.
- Females, students from an underrepresented minority group or not from an underrepresented group, those with low GPAs in high school, and students first enrolled in a two-year college renew their FAFSA at rate that is higher than their non-coached female peers.

However, coaching is not more effective for one subgroup compared to the other. That is, for example, there is no statistically significant difference in program impacts between males and females.

Our analytic approach estimates separate impacts for the two subgroups of a given characteristic, and then assesses the magnitude and statistical significance of the *difference* between those two impact estimates. In other words, we begin by looking *within* individuals subgroups to estimate impacts (comparing, for example, female SBC treatment students to female comparison students) and then looking at differences *between* subgroups (to compare, for example, the impact for female students to that for male students). When the difference *between* the impacts for any two subgroups is not statistically significant, then the difference between the individual subgroup impacts (females versus males, for example)can be considered indistinguishable from zero. That is, in cases where the difference between the two subgroup impacts is not statistically significant, the results do not provide enough evidence to conclude that coaching is more effective at getting one group to persist than it is for the other group, and we could have found the opposite pattern of impacts by chance (if the apparent difference in impacts was in fact due to chance).

Exhibit 5-1: Student characteristics and subgroups

Student Characteristic	Subgroups
Gender	Female
	Male
Underrepresented minority status	Underrepresented minority
	Not underrepresented minority
High school academic achievement	Higher GPA
	Lower GPA
Type of college	Two-year
	Four-year

The exhibits below illustrate the findings. Each of Exhibits 5-2 to 5-5 displays the impact estimates of SBC separately according to the student characteristic subgroups. (See Appendix C for the full results of these analyses.)

These results are considered exploratory because they are based on subsets of the full sample; the smaller sample sizes mean the estimates are less precise and therefore limit our ability to detect statistical significance. Note that for these (and other) exploratory analyses, we do not apply corrections for multiple hypothesis tests.

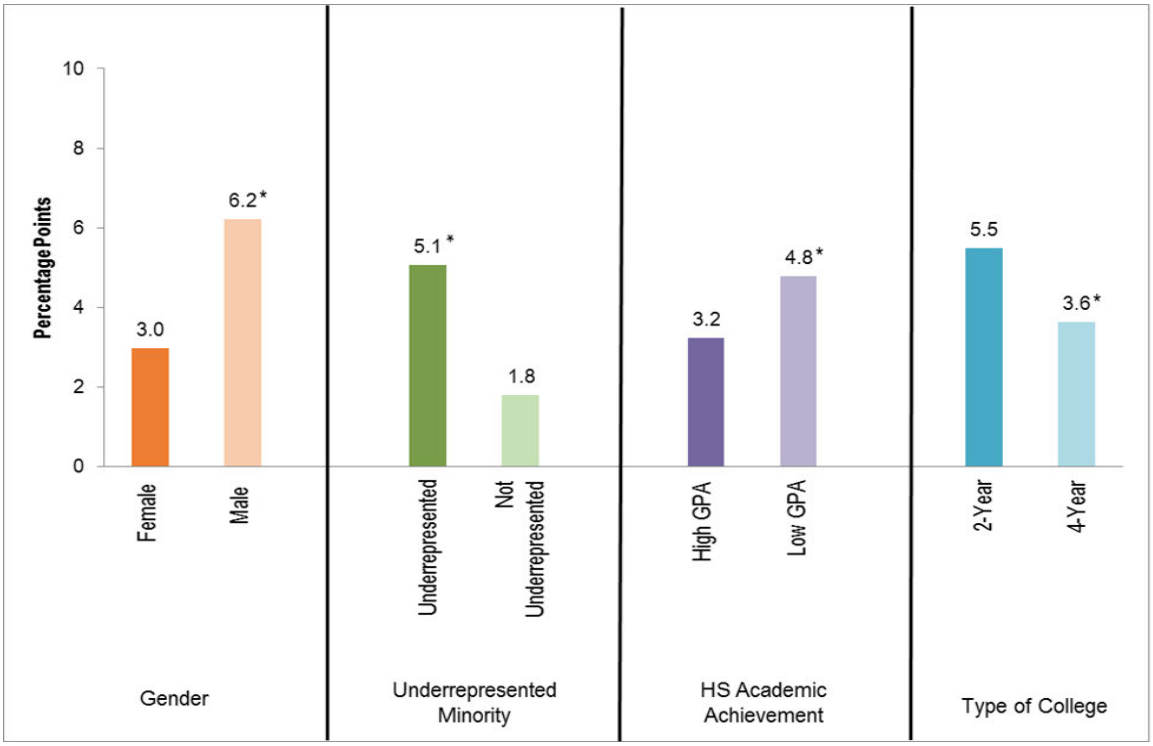
The exhibits in this chapter use an asterisk (*) to indicate a statistically significant impact (where observed) for a particular subgroup of students; it uses the dagger symbol (†) to indicate for which characteristics from Exhibit 5.1 the subgroup effects vary statistically. For example, an asterisk (*) on *female* would indicate that SBC (treatment group) female students persisted at a rate that is statistically different from that of non-SBC (comparison group) female students. The dagger on (†) on *gender* would indicate that the impact of SBC on females is statistically different from that on males.

5.1 Persistence

Exhibit 5-2 below displays the impact estimates of SBC on persistence into the *second* year of college as a function of key student characteristics. When looking *within* subgroups, results show SBC students persist at a higher rate than comparison students if they are male, from an underrepresented minority group, have low GPAs in high school, and first enrolled in a four-year college. However, when looking *across* subgroups, the difference in impact estimates are non-significant and are indistinguishable from zero. That is, for example, male SBC students persist at a higher rate when compared to male comparison students. However, the impact for males is considered similar in magnitude to the impact for females.

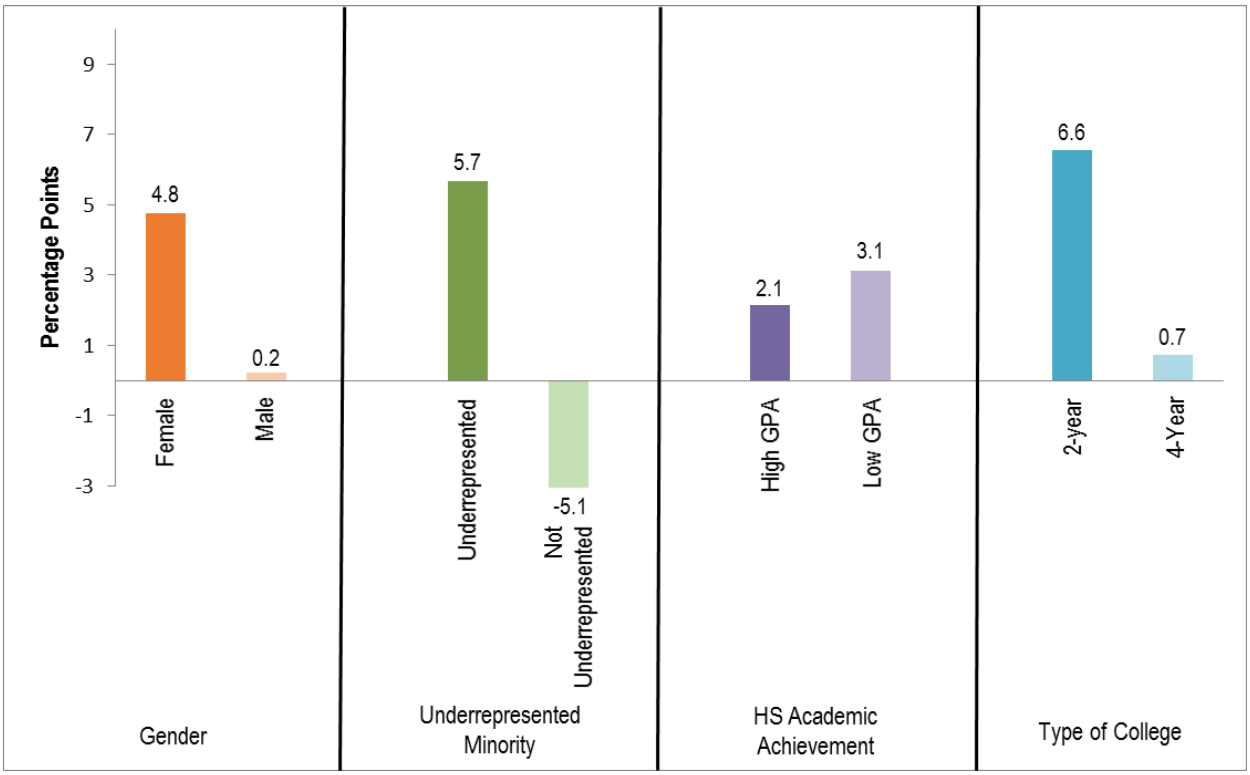
Impact estimates on persistence into the *third* year of college are not statistically significant for any of the student subgroups (Exhibit 5-3).

Exhibit 5-2: Impact of SBC on persistence into second year of college, by student characteristics



Sources: National Student Clearinghouse data from Boston Public Schools and MA DESE; college administrative data.
Note: For more information about sample sizes, please see Appendix C, Tables C-1 through C-4. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
* Impact is significant on subgroup at the 5 percent level.
Exhibit Reads: There is a statistically significant impact of coaching on persistence into the second year of college for males, underrepresented minority students, students with low GPAs in high school, and students attending four-year colleges. There is no statistically significant variation detected in the impacts within any of the student characteristics.

Exhibit 5-3: Impact of SBC on persistence into third year of college, by student characteristics

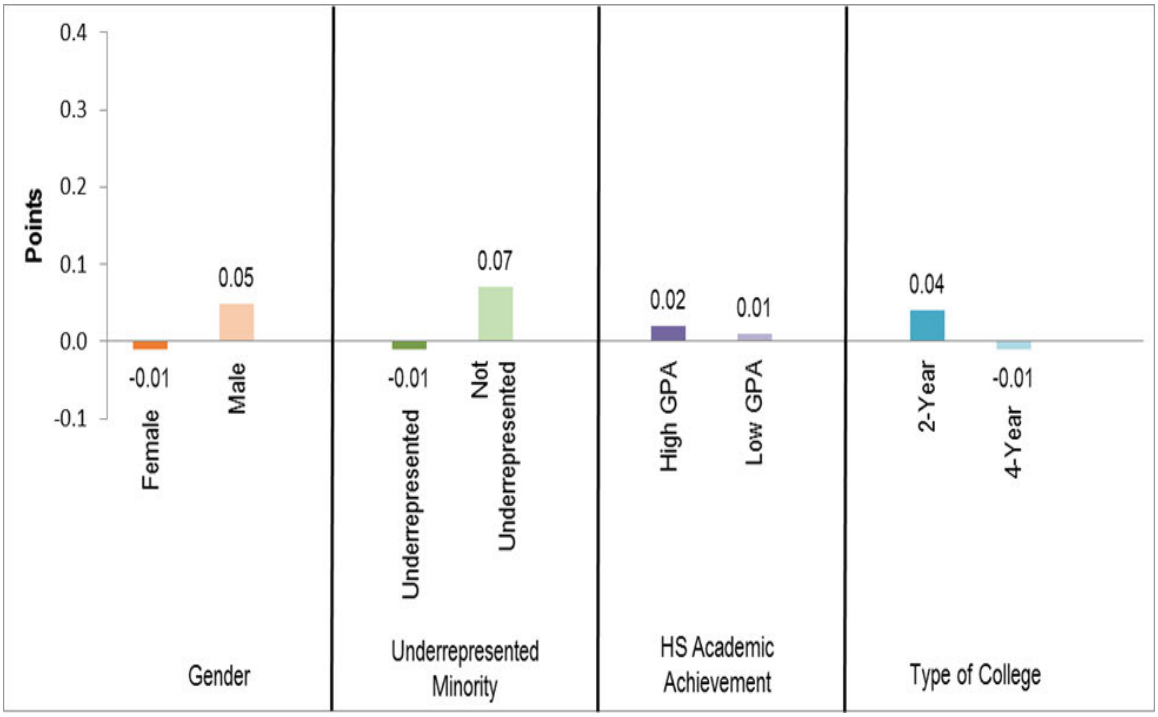


Sources: National Student Clearinghouse data from Boston Public Schools and MA DESE; college administrative data.
Note: For more information about sample sizes, please see Appendix C, Tables C-1 through C-4. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
Exhibit Reads: There is no statistically significant impact of coaching on persistence into the third year of college for any of the student characteristic subgroups. There is no statistically significant variation detected in the impacts within any of the student characteristics.

5.2 Achievement

Exhibit 5-4 displays impact estimates of SBC on cumulative GPA by student characteristics, comparing outcomes for SBC students and non-coached students. Similar to the overall results, there are no statistically significant impacts for any subgroups or significant differences between student subgroups.

Exhibit 5-4: Impact of SBC on cumulative GPA, by student characteristics

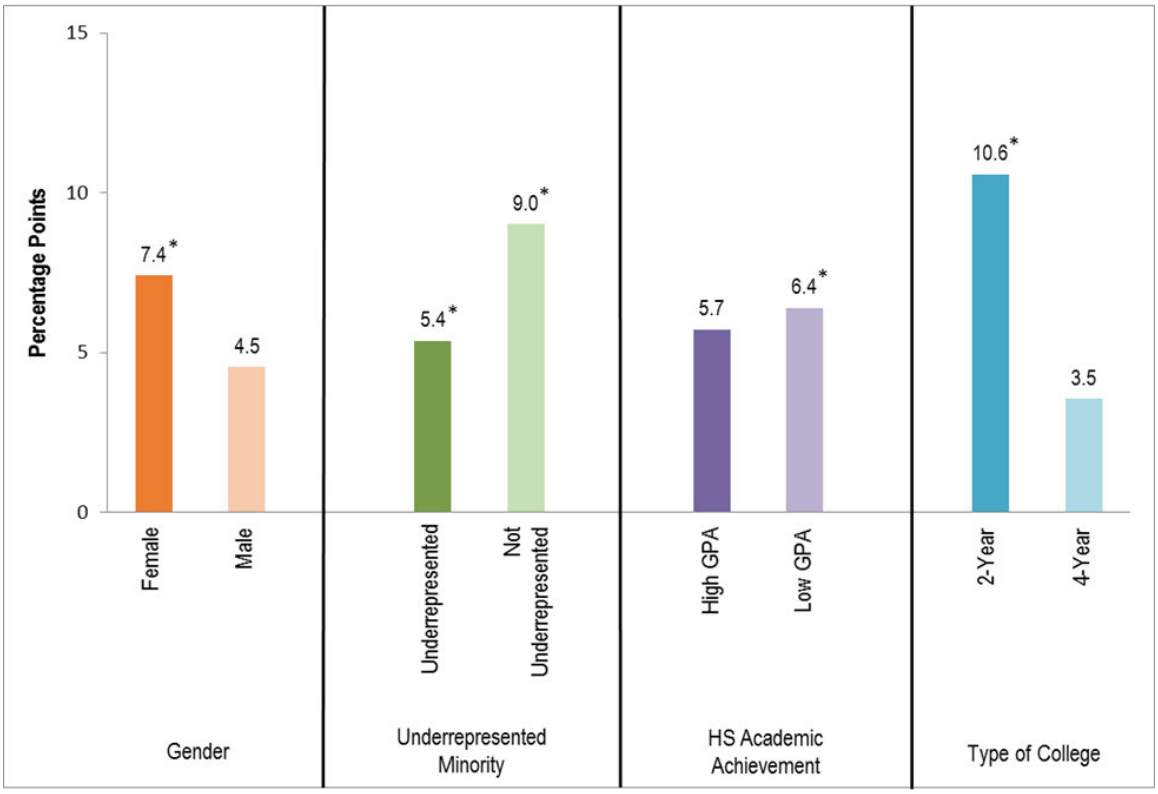


Sources: National Student Clearinghouse data from Boston Public Schools and MA DESE; college administrative data.
Note: For more information about sample sizes, please see Appendix C, Tables C-1 through C-4. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
* Impact is significant on subgroup at the 5 percent level.
Exhibit Reads: There is no statistically significant impact of coaching on cumulative GPA for any of the student characteristic subgroups or significant differences between student subgroups.

5.3 Financial Aid

Exhibit 5-5 displays the impact estimates of SBC on FAFSA renewal for SBC students compared with non-coached students, according to the four student subgroups. When looking *within* subgroups, results show SBC students renew their FAFSA at a higher rate than comparison students if they are female, from an underrepresented minority group or not from an underrepresented group, have low GPAs in high school, and first enrolled in a two-year college. However, when looking *across* subgroups, the difference in impact estimates are non-significant and indistinguishable from zero. That is, for example, SBC students with higher GPAs persist at a higher rate when compared to comparison students with higher GPAs. However, the impact for students with higher GPAs is considered similar in magnitude to the impact for students with lower GPAs.

Exhibit 5-5: Impact SBC on FAFSA renewal, by student characteristics



Sources: National Student Clearinghouse data from Boston Public Schools and MA DESE; college administrative data.
Note: For more information about sample sizes, please see Appendix C, Tables C-1 through C-4. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
* Impact is significant on subgroup at the 5 percent level.
Exhibit Reads: There is a statistically significant impact of coaching on FAFSA renewal rates for females, students not from underrepresented minorities, students from underrepresented minorities, students with low GPAs in high school, and students attending two-year colleges. There is no statistically significant variation detected in the impacts within any of the student characteristics.

5.4 Learning Points

An important question for the study is to understand not only whether SBC participation improves students' college-related outcomes, but for whom. Looking within the subgroups (at coached male students relative to non-coached males, for example), there are some groups for whom coaching has a measurable positive impact relative to the outcomes of their peers who do not participate in coaching. Yet, there is not evidence that coaching is more effective for some students and less effective for others. Further, given the number of subgroups examined and the lack of significant variation between subgroups (between females and males, for example), it is possible that some of the significant impacts on individual subgroups could be due to chance.

6. How Are Impacts Related to Features of Coaching?

This chapter examines how features of coaching are related to the impacts of SBC. The analyses examine whether impacts are concentrated within students who experience more (or less) of a given feature. For example, are the impacts *greater* when students experience more coaching interactions?

The chapter explores three coaching features: **frequency** of coach-student interactions, **duration** of one-on-one coach-student interactions, and **frequency of particular topics** (academic, financial aid, managing life responsibilities, and career) discussed during coach-student interactions.[24] The exploratory analyses discussed in this chapter use the quasi-experimental approach described in Chapter 3, although the analytic models differ slightly, because information on features of coaching is available only for treatment students (those receiving SBC) and because the features of coaching may be related to program participation and outcomes.

To simplify the analyses and ease the interpretation of the results, we transformed each continuous variable into a binary variable such that the values above the median are considered *high* and those below are considered *low* (e.g., one subgroup comprising students who experience a lower frequency of interactions, and the other comprising students who experience a higher frequency). We estimate separate impacts for the two subgroups, and then assess the magnitude and statistical significance of the *difference* between the subgroup-specific impact estimates.

Key Findings

The results from these exploratory analyses indicate that the impacts of coaching are concentrated among students who experience more frequent coach interactions (defined in these analyses as 14 or more coaching interactions). The results also suggest that the impacts of coaching are sometimes concentrated among students who experience longer coaching interactions ("high duration") and shorter coaching interactions ("low duration").

Additionally, in general, SBC students have more positive college outcomes when they experience more interactions addressing any of the topic areas (i.e., academic, financial aid, managing life responsibilities, career). The results suggest that greater exposure to some specific coaching features *strengthens* particular impacts of SBC. Specifically,

- Students were more likely to *persist into the second year* of college if they had more exposure to any one of the topic areas.
- Students were more likely to *persist into the third year* of college if they had more exposure to academic or career topics.
- Students were more likely to *renew their FAFSA* if they had more exposure to academic, financial aid, or career topics.

These analyses are considered exploratory, both because they are based on subsets of the full sample and consequentially have less statistical power to detect differences than do full-sample analyses and because these coaching features are post-treatment measures that may reflect program impacts. For example, how often students choose to meet with their coach may differ as a function of unobserved factors such as students' receptivity to coaching. Thus, the differences in effects for the corresponding subgroups may not be fully attributable to the features of coaching themselves.

[24] Because of relatively small student sample sizes, the study could not generate reliable impact estimates by nonprofict coaching organization.

Findings within this chapter are organized according to the three outcome domains: persistence, achievement, and financial aid. Within each outcome domain, we present results for the primary outcomes (see Appendix C for the full results of these analyses). Note that for these (and other) exploratory analyses, we do not apply corrections for multiple hypothesis tests.

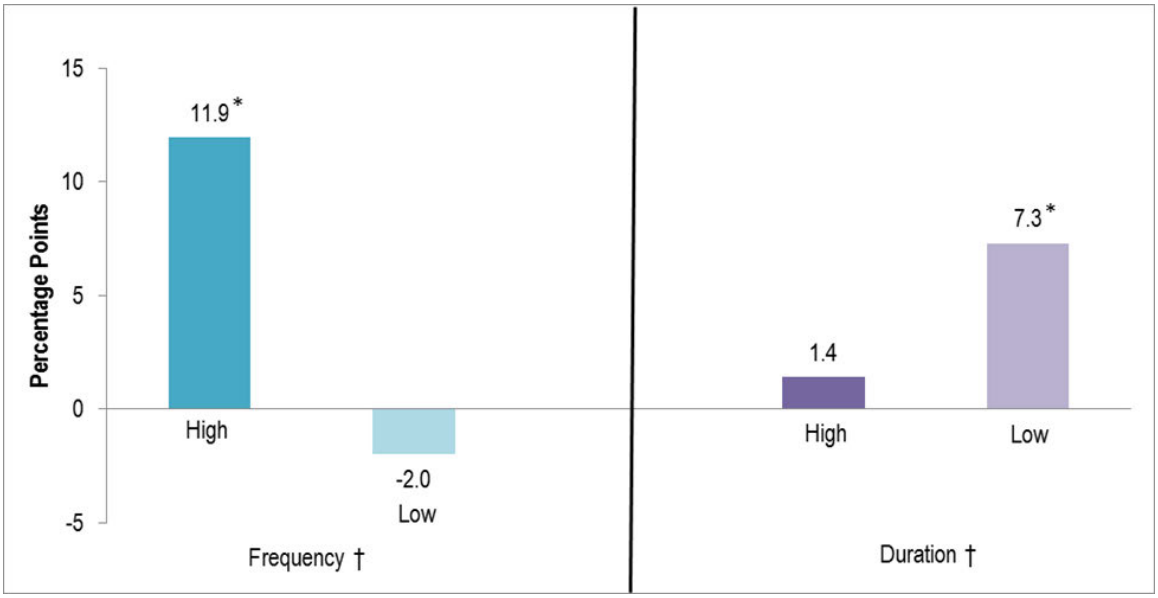
Following the conventions in Chapter 5, we use an asterisk (*) to indicate a statistically significant impact (where observed) for a particular subgroup of students; the dagger symbol (†) indicates for which coaching features (e.g., *frequency*, *duration*) subgroup effects vary statistically.

6.1 Persistence

Exhibits 6-1 and 6-2 display the impact estimates on persistence into the *second year* of college for SBC students compared with non-coached students, by features of how SBC students experience coaching. SBC students are more likely than comparison students to persist into their second year of college if they experience

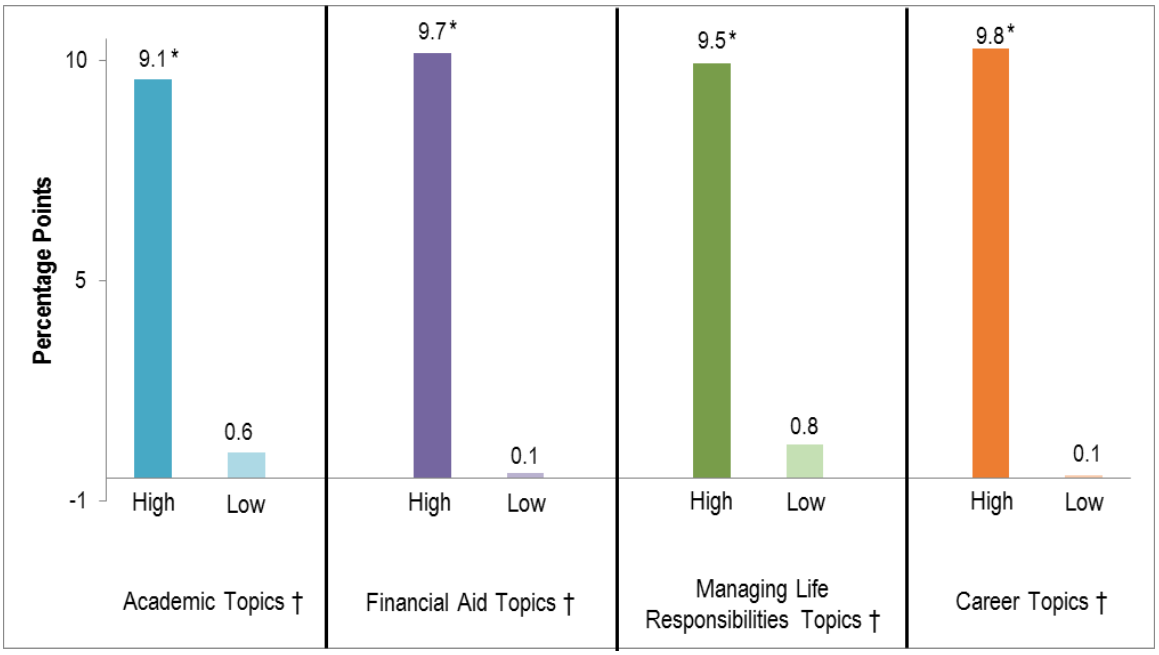
- higher numbers of coach-student interactions overall;
- shorter one-on-one interactions (in person or phone); or,
- higher numbers of coach-student interactions that cover academic, financial aid, managing life responsibilities, and career topics.

Exhibit 6-1: Impact of SBC on persistence into the second year of college, by frequency and duration of coach-student interactions



Sources: National Student Clearinghouse from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data; Salesforce data collected from the coaching organizations.
Note: For more information about sample sizes, please see Exhibits C-9 and C-10 in Appendix C. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
* Impact is significant on subgroup at the 5 percent level.
† Feature of coaching is significant at the 5 percent level.
Exhibit Reads: There are statistically significant impacts of coaching on persistence into the second year of college for students in the high frequency of interactions subgroup (that is, more than 14 meetings) and those in the low duration of one-on-one interactions subgroup (that is, 32.5 minutes or less). There is variation in the impact of coaching on persistence into the second year by frequency of interactions and by duration of interactions..

Exhibit 6-2: Impact of SBC on persistence into the second year of college, by frequency of topics addressed during coaching interactions



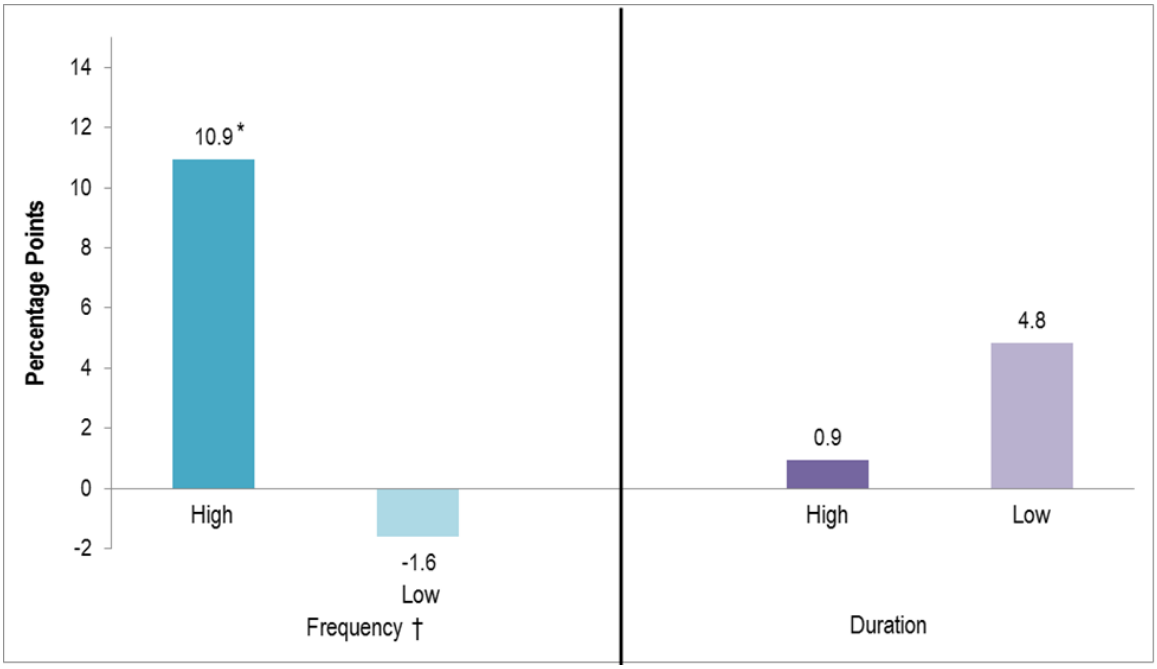
Sources: National Student Clearinghouse from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data; Salesforce data collected from the coaching organizations.
Note: For more information about sample sizes, please see Exhibits C-5 through C-8 in Appendix C. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.* Impact is significant on subgroup at the 5 percent level.
† Feature of coaching is significant at the 5 percent level.

Exhibit Reads: Across the topic areas, there is a statistically significant impact of coaching on persistence into the second year of college for students who experienced more interactions that covered the given topic (students in the "high" subgroups – that is, more than nine meetings for academic topics, two meetings for financial aid topics, one meeting for managing life responsibilities topics, and one meeting for career topics). For each of the four topic areas, there is variation in the impact of coaching on persistence into the second year by frequency of interactions focusing on that topic.

Exhibits 6-3 and 6-4 display the impact estimates on persistence into the *third year* of college for SBC students compared with students who did not participate in SBC, by features of coaching experienced by SBC students. SBC students are more likely than comparison students to persist into their third year of college if they experience

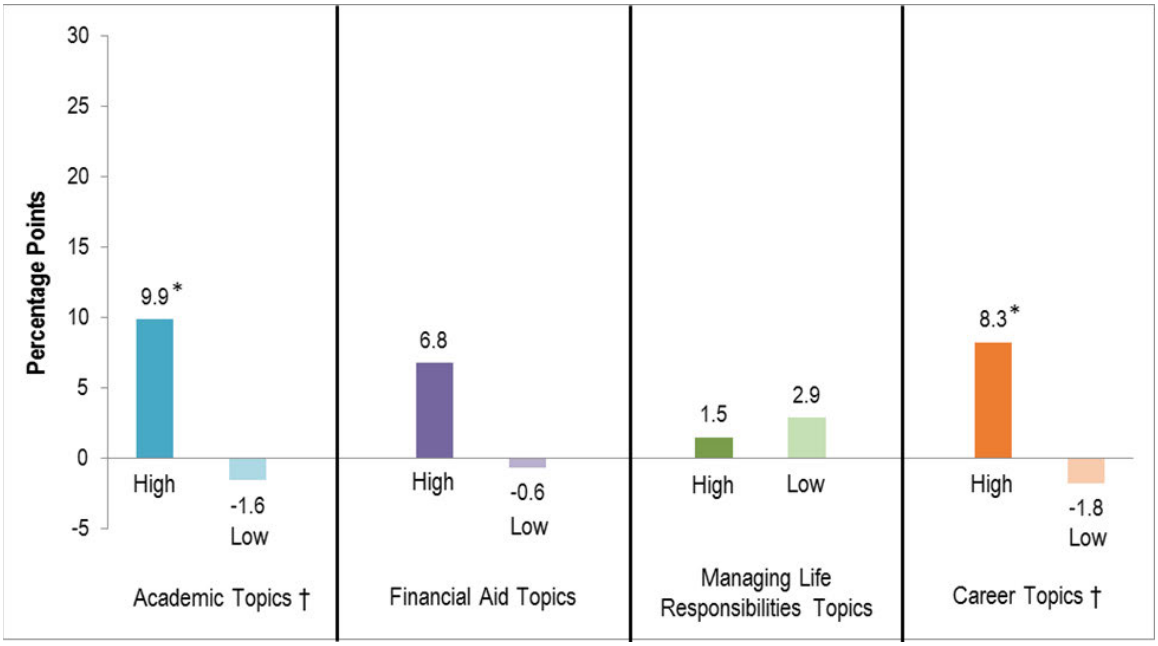
- more coach interactions overall; or
- more coach interactions that cover academic or career topics.

Exhibit 6-3: Impact of SBC on persistence into the third year of college, by frequency and duration of coach-student interactions



Sources: National Student Clearinghouse from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data; Salesforce data collected from the coaching organizations.
Note: For more information about sample sizes, please see Exhibits C-9 and C-10 in Appendix C. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.* Impact is significant on subgroup at the 5 percent level.
† Feature of coaching is significant at the 5 percent level.
Exhibit Reads: There are statistically significant impacts of coaching on persistence into the third year of college for students in the high frequency of interactions subgroup (that is, more than 14 meetings). There is variation in the impact of coaching on persistence into the third year by frequency of interactions.

Exhibit 6-4: Impact of SBC on persistence into the third year of college, by frequency of topics addressed during coaching interactions



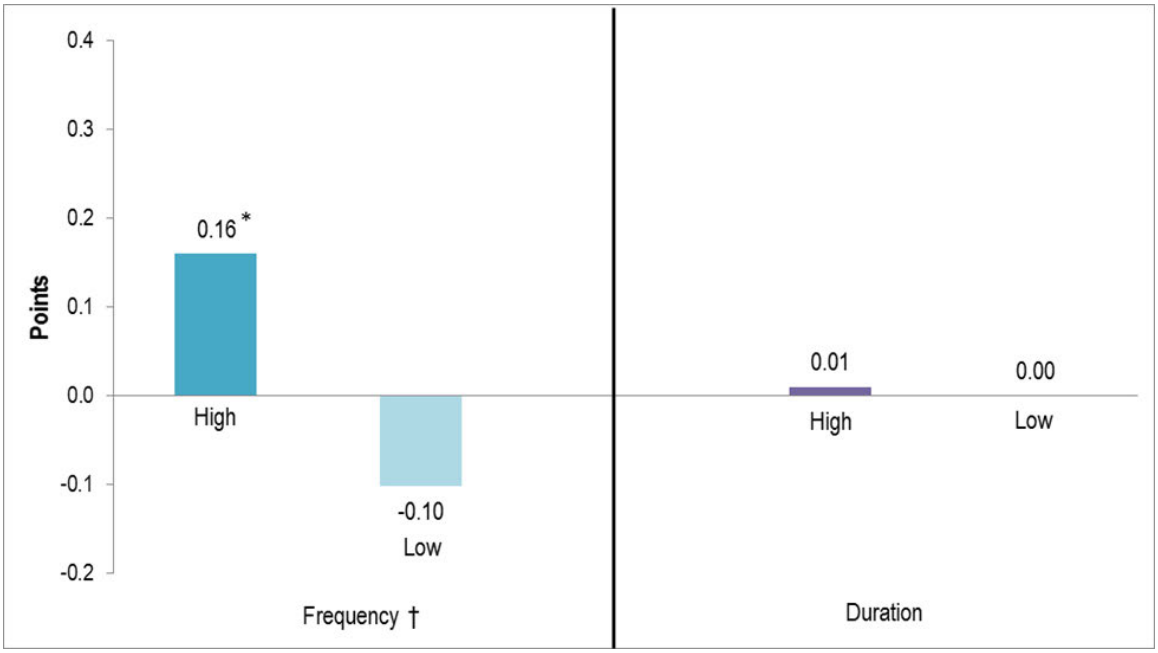
Sources: National Student Clearinghouse from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data; Salesforce data collected from the coaching organizations.
Note: For more information about sample sizes, please see Exhibits C-9 and C-10 in Appendix C. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
* Impact is significant on subgroup at the 5 percent level.
† Feature of coaching is significant at the 5 percent level.

Exhibit Reads: For academic topics and career topics there are statistically significant impacts of coaching on persistence into the third year for students who experience more interactions that covered the given topic (students in the "high" subgroups -- that is, more than nine meetings for academic topics and one meeting for career topics). There is variation in the impact of coaching on persistence into the third year by frequency of interactions focusing on academic topics and career topics.

6.2 Achievement

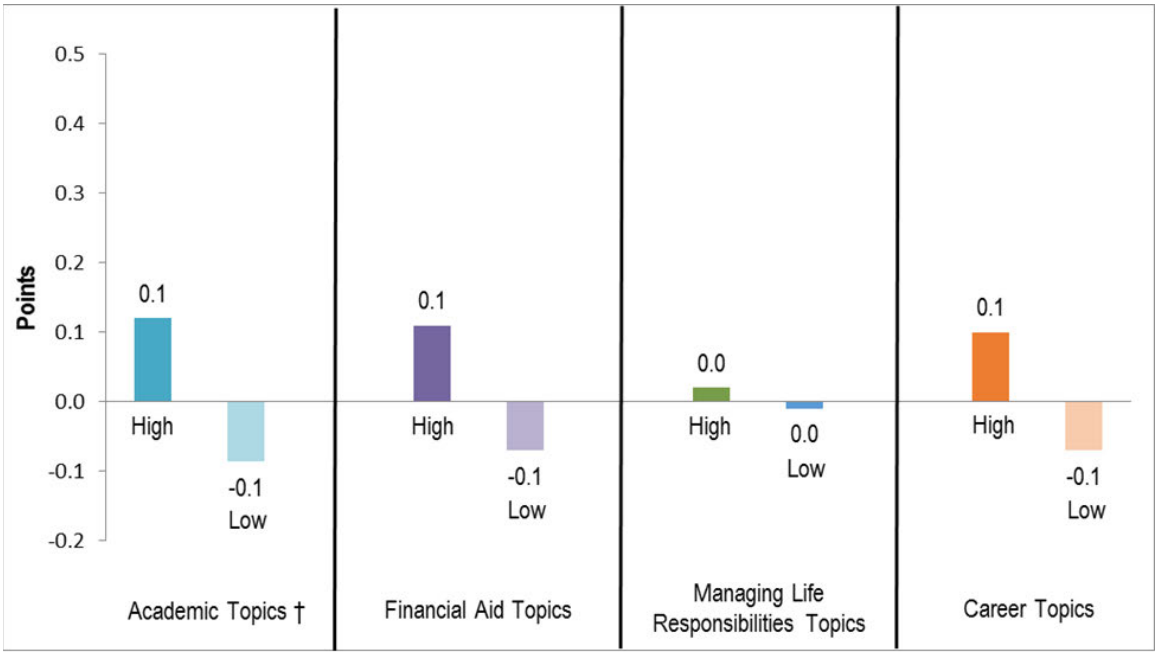
Exhibits 6-5 and 6-6 display the impact estimates on cumulative college GPA for SBC students compared with students who did not participate in SBC, by features of how SBC students experience coaching. SBC students are more likely than comparison students to achieve higher cumulative GPAs if they experience more coach interactions overall.

Exhibit 6-5: Impact of SBC on cumulative GPA, by frequency and duration of coach-student interactions



Sources: College administrative data; Salesforce data collected from the coaching organizations.
Note: For more information about sample sizes, please see Exhibits C-9 and C-10 in Appendix C. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
* Impact is significant on subgroup at the 5 percent level.
† Feature of coaching is significant at the 5 percent level.
Exhibit Reads: There are statistically significant impacts of coaching on cumulative GPA for students in the high frequency of interactions subgroup (that is, more than 14 meetings). There is variation in the impact of coaching on cumulative GPA by frequency of interactions.

Exhibit 6-6: Impact of SBC on cumulative GPA, by frequency of topics addressed during coaching interactions



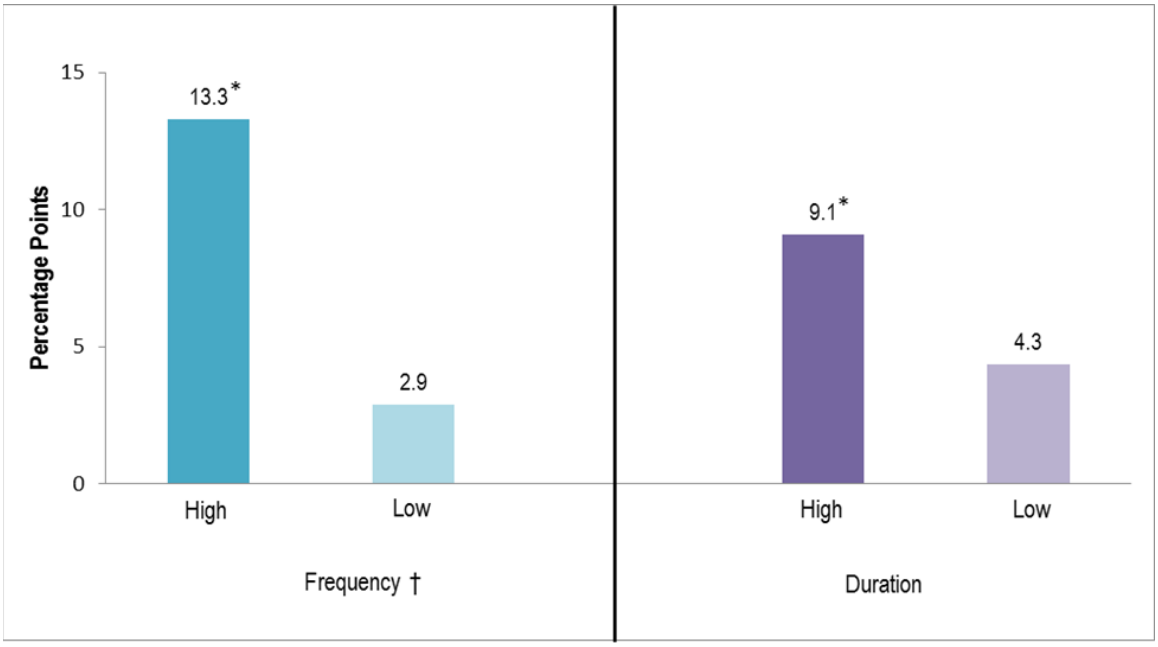
Sources: College administrative data; Salesforce data collected from the coaching organizations.
Note: For more information about sample sizes, please see Exhibits C-9 and C-10 in Appendix C. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
† Feature of coaching is significant at the 5 percent level.
Exhibit Reads: Across all topic areas, there are no statistically significant impacts of coaching on cumulative GPA for any interaction subgroup. There is variation in the impact of coaching on cumulative GPA by frequency of interactions focusing on academic topics.

6.3 Financial Aid

Exhibits 6.7 and 6.8 display the impact estimates on FAFSA renewal for SBC students compared with students who did not participate in SBC, by features of how SBC students experience coaching. SBC students are more likely than comparison students to renew their FAFSAs if they experience

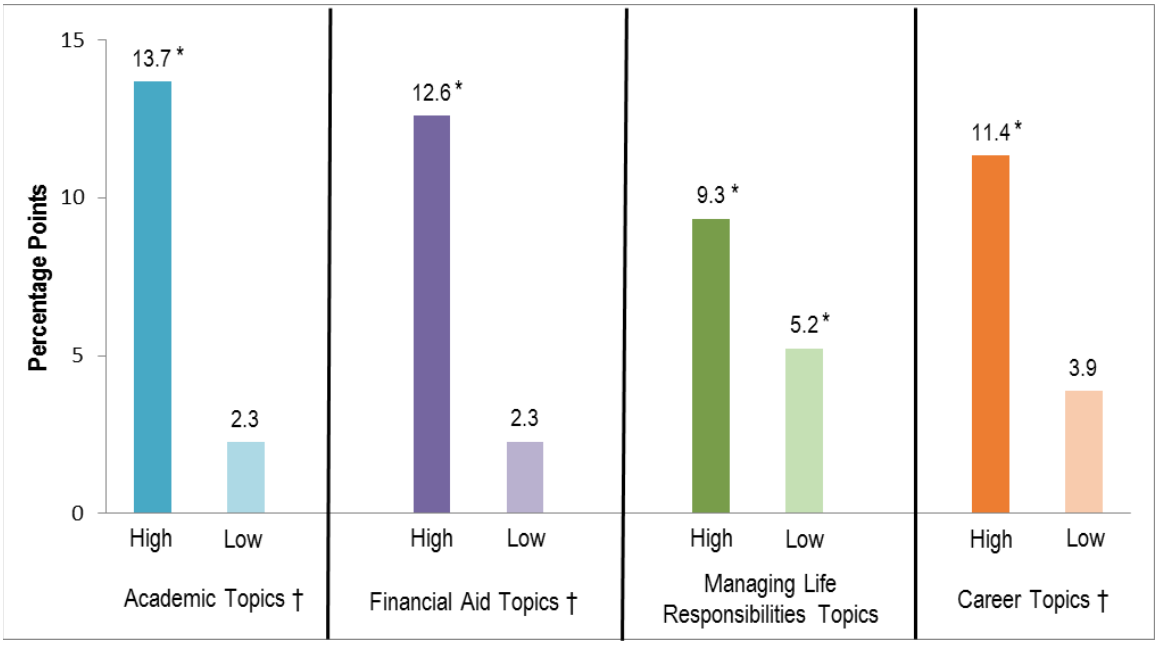
- more coach interactions overall;
- longer one-on-one interactions (in person or phone);
- higher numbers of coach-student interactions that cover academic, financial aid, managing life responsibilities, and career topics; or
- lower numbers of coach-student interactions related to managing life responsibilities.

Exhibit 6-7: Impact of SBC on FAFSA renewal, by frequency and duration of coach-student interactions



Sources: College administrative data; Salesforce data collected from the coaching organizations
Note: For more information about sample sizes, please see Exhibits C-9 and C-10 in Appendix C. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
* Impact is significant on subgroup at the 5 percent level.
† Feature of coaching is significant at the 5 percent level.
Exhibit Reads: There are statistically significant impacts of coaching on FAFSA renewal for students in the high frequency of interactions subgroup (that is, more than 14 meetings) and those in the high duration of one-on-one interactions subgroup (that is, more than 32.5 minutes). There is variation in the impact of coaching on cumulative GPA by frequency of interactions.

Exhibit 6-8: Impact of SBC on FAFSA renewal, by frequency of topics addressed during coaching interactions



Sources: National Student Clearinghouse from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data; Salesforce data collected from the coaching organizations.
Note: For more information about sample sizes, please see Exhibits C-9 and C-10 in Appendix C. Adjusted comparison group means, impacts (treatment mean minus adjusted comparison mean), and statistical significance are drawn from the study's regression model.
* Impact is significant on subgroup at the 5 percent level.
† Feature of coaching is significant at the 5 percent level.
Exhibit Reads: Across the topic areas, there is a statistically significant impact of coaching on FAFSA renewal for students who experience more interactions that cover the given topic (students in the "high" subgroups – that is, more than nine meetings for academic topics, two meetings for financial aid topics, one meeting for managing life responsibilities topics, and one meeting for career topics). There are also statistically significant impacts for students who experience fewer interactions covering general life topics (one meeting or less). There is variation in the impact of coaching on persistence into the third year by interactions focusing on academic topics, financial aid topics, and career topics.

6.4 Learning Points

Taken together, the results from examining variation in the effects of SBC and how the program affects particular student subgroups (as defined by their coaching experiences) can inform decisions about program implementation—specifically, about practices that seem to be more consistently related to positive student outcomes.

One takeaway is that more frequent coaching is associated with better outcomes. SBC students with higher frequency of coaching interactions experience larger program impacts on all outcomes than do SBC students who experience fewer interactions. The subgroup findings are less consistent when presented by length of interactions. SBC students are more likely to persist into the third year of college if they experience *shorter* coaching interactions. Yet, SBC students are more likely to renew their FAFSA if they experience *longer* coaching interactions. These results may suggest that frequent communications, even if brief, could be a powerful lever for improving student outcomes.

Lastly, SBC generally has more positive impacts across the primary outcomes for students who experience more coaching interactions on any of the topic areas. This trend was most salient for persistence into the second year and FAFSA renewal. These findings might suggest that frequency of

coaching interactions is the underlying mechanism, rather than the specific topic covered, when it comes to outcomes such as persistence and FAFSA renewal. And for other outcomes such as GPA, there is no overall impact, nor are there impacts nested within particular subgroups of coached students.

However, because students who experience more coaching interactions may differ in various ways from those who experience fewer interactions, this exploratory analysis cannot address whether outcomes of the latter group would have been better had they experienced more coaching interactions. In particular, it could be the case that those who experience more coaching interactions are precisely those who are most responsive to coaching, and those who experience fewer coaching interactions are disproportionately those who would not benefit from more coaching.

7. Discussion

Success Boston coaches support students to successfully navigate college, through both referrals to campus resources and direct support, across areas that research suggests can be helpful in improving student outcomes. Coaches connect students to resources such as tutoring and financial aid services, help them plan their coursework and identify a major, work with them to understand financial aid processes and options, provide them with time management and study strategies, and develop positive relationships with students (Linkow et al. 2017b).

Perhaps uniquely so, the SBC program integrates these different features into a single intervention that is tailored to individual student needs. Core to the SBC model are responsiveness to student needs and customization of support to address each student's individual needs for support and contact. Coaches vary the frequency, modes of communication, and focus of supports to "meet students where they're at" (Linkow et al. 2017b). Though one coach may provide support across multiple topics, any individual student may receive support only on some, not all topics, depending on that student's needs and interests.

Summary of SBC Program Impacts

Success Boston Coaching demonstrated significant program impacts on the following four student outcomes (**bold** indicates primary outcome):

- **Persistence into second year of college**
- Full-time status
- Credit accumulation
- FAFSA renewal

The flexible, multi-faceted support provided to students through SBC has positive effects on students across multiple early college outcomes. SBC improves students' outcomes across all three outcome domains examined (persistence, achievement, and financial aid). Impacts are consistent across genders, student racial/ethnic groups, the continuum of high school GPA averages, and type of college students attended.

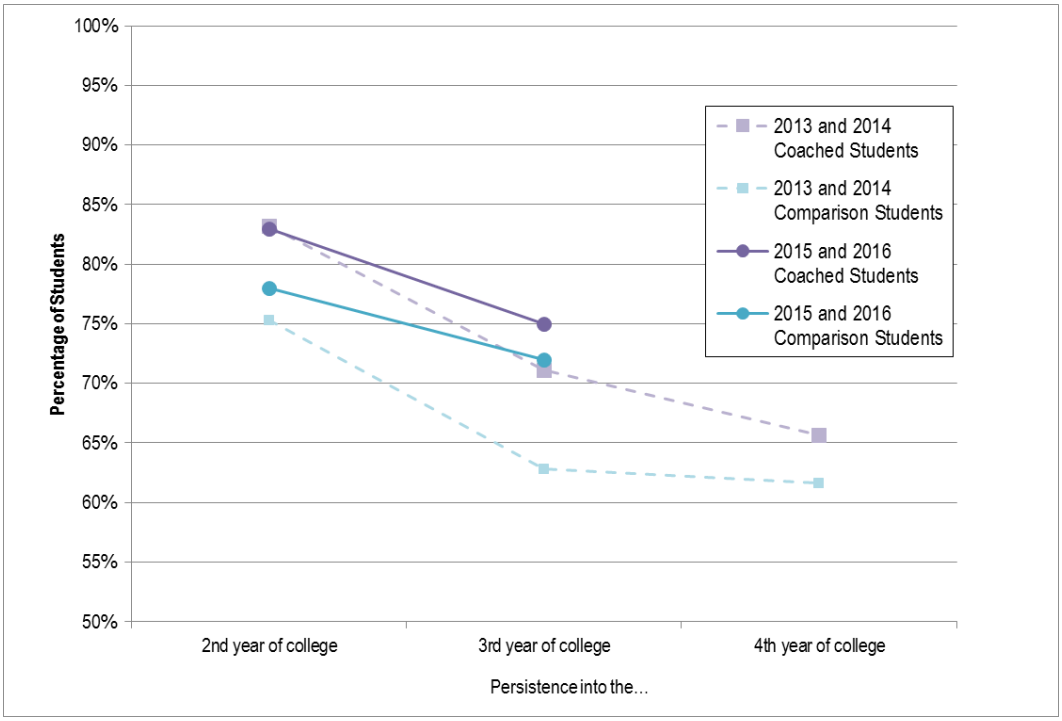
The magnitude of the SBC impacts is smaller than that reported about earlier cohorts of SBC students (see Linkow et al. 2017b) and other similar programs targeting improved college student outcomes (see Bettinger and Baker 2014; Castleman and Page 2015; Oreopoulos and Petronijevic 2016). These other studies have observed effects on college persistence and achievement equivalent to about 10-15 percent increases over the control group means. Overall, the effects of the scaled-up SBC program can be characterized as positive and moderate in magnitude.

The expansion of SBC did not come at the expense of how often coaches interact with students. Students in the scaled-up program typically interacted with their coaches between one and two times a month during the academic year, and each in-person meeting was about 30 minutes; these patterns are similar to those observed in the earlier cohorts. However, the impacts found for the program at scale are smaller than those seen for earlier cohorts, as seen on Exhibit 7.1. Overall, coached and non-coached students from the 2015 and 2016 cohorts are persisting at slightly higher rates than students from the 2013 and 2014 cohorts.

However, the difference between the coached and non-coached students in the 2015 and 2016 cohorts are smaller than that for the earlier cohorts. It may be that the shrinkage in the impact of SBC is caused by improvements in the outcomes of comparison group students. There is some evidence that the colleges participating in SBC, in particular those that enroll large proportions of SBC students, have made changes

in recent years to improve their student support services. For example, some of the colleges have established their own cadre of "coaches" to support students coming from outside of Boston; other colleges have centralized their non-academic support services to improve accessibility to students. In addition, as part of the overall Success Boston collective focus on student success, colleges have implemented new programs and strategies that are designed to improve student outcomes. Colleges' strategies include all Boston students receiving a coach at the University of Massachusetts Boston and the Life Map program at Bunker Hill Community College, among others. These college changes could be improving outcomes for all students.

Exhibit 7-1: Trends in student persistence by cohort and coached status



Source: Linkow et al. 2017b; National Student Clearinghouse data from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education.

Even with overall increases in student outcomes, SBC continues to add value to participating students. Similar to previous results, there is evidence that the impacts of coaching vary in important ways: there are statistically significant impacts across the primary outcomes on students who experience more coaching interactions overall and more interactions that cover academic and career topics. However, longer one-on-one interactions do not necessarily translate into larger impacts, perhaps suggesting that shorter, more frequent interactions may be more helpful to students than a few long interactions.

Together, the impact results suggest that the SBC model has meaningful and significant promise. A model in which professional coaches deliver sustained, proactive, and responsive support within a larger education network may not necessarily need to be "high touch" to accomplish the overarching program goal: to move more students more quickly and effectively along the pathway to college completion. The proactive, sustained, and adaptive advising that is SBC can improve the college outcomes of students who are typically underserved and underrepresented in postsecondary education.

One of the key goals of the Success Boston initiative is to demonstrate that transition coaching can meaningfully improve students' college-related outcomes, and it is the case that at least some of the partner colleges are altering the services they offer to incoming students. The findings reported here reflect impacts observed for the seventh and eighth cohorts of students participating in SBC (2015 and 2016). As the program has matured, the model has been refined, coordination with colleges has improved, and some of the components and practices of SBC coaches have spilled over into the support services provided by the colleges themselves. Should student support services similar to those provided by the SBC nonprofit organizations continue to increase in partner colleges, it is possible that the contrast between SBC students and the comparison students would diminish further for later cohorts. However, if the contrast comes with improved outcomes for all students, then on the whole, the initiative may have achieved its intended systemic change.

7.1 Limitations

This evaluation uses a quasi-experimental research design to estimate the effects of SBC on various student outcomes. Methodological research about within-study comparisons suggests that such designs can generate causal impact estimates and can approximate results from randomized control trials. The analyses reported here have been informed by such methodological research. Nevertheless, we should note that there could be confounding factors not fully accounted for by the local and focal matching strategy this evaluation employed, because such factors are not observable or are measured with error or with proxies. To the extent that the impacts reflect, in part, the influence of unmeasured individual student characteristics on both participation in coaching and outcomes in college, the matched comparison group does not eliminate bias that a random assignment (experimental) process would.

Nonetheless, the SBC recruitment processes and the sample construction each help reduce the likelihood of confounding factors. Students are recruited in SBC through a variety of avenues: referrals from high school guidance counselors and other community organizations, nonprofit organizations' middle school and high school programming pipelines, word of mouth, and sometimes from college referrals. Such recruitment pathways combine to create successive cohorts of SBC students who differ with respect to levels of motivation, attachments to the nonprofit organizations, and willingness to seek support. Also, the sample was constructed using the most inclusive definition of an SBC student. Specifically, for the purposes of the evaluation, all students initially recruited into the SBC program and who appeared in the program database—including those few without a single recorded interaction with a coach—are considered SBC students. Therefore, the analytic sample includes students who may have had second thoughts about participating in SBC and those with lower predispositions to take up the supports SBC offers.

7.2 Future Steps

The results reported here demonstrate that SBC is having the desired effects on students from Boston high schools. SBC participants are persisting longer, achieving higher academic progress, and taking the necessary steps to maintain funding for college—all good signs of progress toward college completion. However, the low number of credits accumulated toward graduation for students at two-year colleges may be cause for concern. The study will continue to monitor students' progress over the next several years, culminating in a 2022 final report describing impacts once study students have had the opportunity to complete five and six years of college. That report will help address questions about whether the positive impacts of coaching reported here persist and translate to more students earning college credentials.

Appendix A. Propensity Score Matching Process

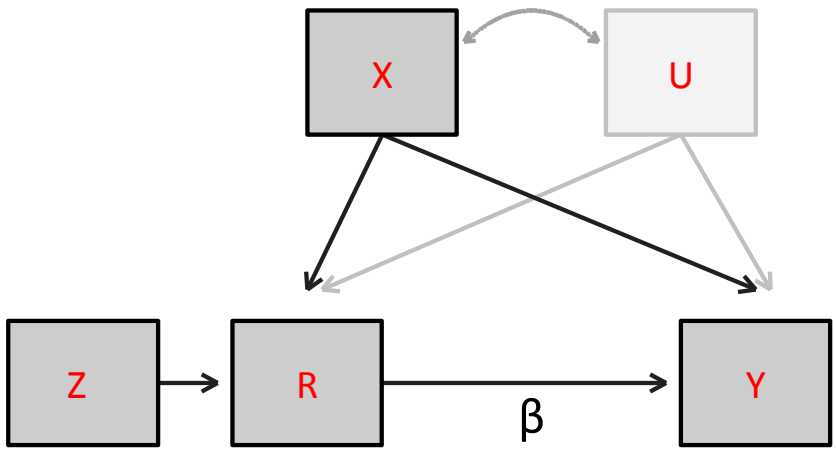
A simple comparison of the postsecondary outcomes of Massachusetts students who receive SBC coaching versus Massachusetts students who did not receive SBC coaching would likely provide a misleading picture of the effect of the SBC coaching program, because such a comparison would not take key information into account. First, these two types of students may have different individual and family characteristics, such as academic achievement (test scores) or parental involvement, which may be directly related both to differences in their interest in participating in the SBC coaching program and to their postsecondary outcomes. We refer to characteristics that affect both selection into the program and postsecondary outcomes as *confounding factors*, and these specific types of characteristics as *individual self-selection factors*.

A second type of confounding factor can arise when participating and non-participating students may have been raised in different neighborhoods and had different high school experiences. For instance, some coaching recipients may have had less academic support during high school, and that lack of support may have led them to seek help from an external organization. We refer to such confounders as *historical and locational factors*.

Another complication for the comparison of SBC and non-SBC students is that they may enroll in different colleges, which means exposure to such college-specific factors as selectivity of the institutions, quality of faculty and instruction, and peers' motivation and performance—any of which might influence students' outcomes in different ways. These are not confounders by definition (because they are observed *after* selection into coaching and they could not have determined whether a student participates in SBC or not), but they could still bias the estimated effects of SBC unless they are accounted for. Unlike the first two types of confounders, the influence of these college-specific factors occurs *at the same time* as the SBC coaching program is providing services; therefore, we refer to these as *contemporaneous sources of bias.*

Exhibit A-1 shows a stylized causal diagram of the nature of a set of confounders X, which affect both the receipt of coaching (receipt of treatment) R and the outcome Y, when we want to measure the direct impact of R on Y (denoted as β).

Exhibit A-1: Stylized causal diagram



If there are other confounders that are unobserved (U), the correction for bias in estimates of the effect β resulting from adjusting for X may be incomplete. One might also use factors Z that influence receipt of coaching R but have no direct impact on the outcome Y, called *excluded instruments*, in an instrumental variables (IV) estimator. But in many cases, an IV estimator relying on factors Z will have unacceptably high variance, on top of which it is very hard to claim with confidence that any observed variable satisfies both of these requirements (i.e., influence R but have no direct impact on Y) unless the variable is randomly assigned.

Ideally, we would like to randomly assign R; but in the absence of random assignment, we wish to adjust for as many factors in X as we can, and hope either that variables in U have small correlations with R and Y or that they are highly correlated with X so that adjusting for X eliminates bias due to confounders U, as well.

Given that is was not feasible to conduct an experimental design (or randomized control trial) that would yield two groups of students balanced on all observable and unobservable confounders, we use a quasi-experimental design that (1) compares SBC students with a comparison group of similar students and (2) can account for as many of the observable confounders as possible. Guided by the current methodological research on best quasi-experimental design practices, we constructed such a comparison group using a matching process that had two features:[25] matches were *local* (the comparison cases drawn from the same settings as the treatment cases to the extent possible) and also *focal* (matching was done using baseline characteristics that we believe to predict both selection into treatment and the outcome).

For this evaluation, we developed a quasi-experimental design using local and focal matching; we match each SBC student with at least one and possibly multiple non-SBC students from the same cohort, graduated from high schools with similar characteristics, and enrolled in the same college (local matching), who also share similar baseline characteristics that are empirically linked to our outcomes of interest and also potentially to receipt of SBC coaching (focal matching). Given the large number of matching characteristics, we implemented matching using estimated propensity scores. These scores represent the conditional probability of getting SBC coaching (given covariates), thereby incorporating all the relevant influence of the confounders on selection into treatment in one variable.[26] The assumption

[25] Within-study comparisons (or *design replication studies*) examine quasi-experimental designs by replacing the randomly determined control group in an experiment with comparison groups constructed from units that did not participate in the original experiment using different quasi-experimental methods. Such studies compare the impact estimates yielded by these additional comparison groups with the experimental impact estimates; the studies try to generalize the specific features of the results yielded by quasi-experimental designs that replicate (or come closest) to the experimental results. Results of several well-designed within-study comparisons suggest that matching methods are promising, and the results replicated experimental benchmarks (Bifulco 2012; Clair, Cook, and Hallberg 2014; Steiner, Cook, and Shadish 2011), especially when pre-treatment variables matched on are analogous to key outcomes (Smith and Todd 2005).

[26] One way to conduct matching is to form matched pairs that have the same baseline characteristics, which is also known as *exact matching*. Though this approach may be desirable, it becomes infeasible as the number of baseline variables used in the matching increases. This "curse of dimensionality" problem is solved by performing the matching on a function of the baseline variables, instead of targeting exact matches on all matching variables. Rosenbaum and Rubin (1983) use the probability of being assigned to treatment given covariates as this function, which they call the *propensity score.*

then is that factors Z affect receipt of coaching R conditional on X, but we need not observe Z. For example, some students may, through happenstance, hear about coaching and become more open to participating, and these students will be more likely to participate, even conditional on all X variables or the propensity score that captures the influence of X variables.

The next section of the appendix describes the matching process and construction of the comparison group in detail. Section A.1 explains our local and focal matching approach in more depth; Section A.2 presents how the propensity scores were estimated; Section A.3 provides details on the matching process; and Section A.4 shows how we assessed the quality of the matches.

A.1 Implementation of Local and Focal Matching

Local Matching

The postsecondary outcomes of interest for this evaluation (including persistence in college, and eventually, attainment of a postsecondary credential) are directly dependent on the extent to which students' high schools prepare them for college-level coursework; the difficulty of coursework; accessibility of student support at different colleges; and students' interactions with college teaching staff, administrators, and peers. Therefore, in this context, "local matching" would ideally be implemented by matching SBC students with non-SBC students who both attend the same college and graduate from the same high school in the same year (i.e., the matching process would be conducted separately, using "matching blocks" of unique combinations of high school, college, and cohort groups). Matching on high school attempts to account for historical and locational differences between the SBC and non-SBC students, whereas matching on college controls for contemporaneous sources of bias, which are defined as college-related factors that are independent of the SBC program, differ across colleges, and potentially affect outcomes of interest (e.g., difficulty of coursework). Finally, matching on SBC cohort would account for differences in the overall characteristics of each cohort and the potential differences in the selection processes employed by the coaching organizations and changes in the college-related factors from one year to the next.

Unfortunately, small cell sizes made exact matching on high schools and colleges untenable: in a number of high school/college combinations, there are no potential comparison students with whom treatment students might be matched; in other combinations, there are only one or two comparison students for many treatment students. Given our focus on postsecondary outcomes, we tried to address this problem by privileging students' postsecondary institutions as our primary matching block and pooling high schools into groups of schools with similar characteristics. However, this approach did not solve the issue, and there were still a number of high-school-group-by-college blocks that lacked a sufficient number of potential comparison students to implement the other important aspect of our matching strategy, focal matching.

The matching process we ultimately implemented entails matching within college-by-cohort blocks using propensity scores that are conditional on high school characteristics (e.g., school-level averages of math and English language arts MCAS scores, GPA, and college-going rate) as a proxy for exact matching on high schools. With matching within the college-by-cohort blocks, we aim to control for the college-related contemporaneous sources of bias; by matching on the high school characteristics, we aim to control for the historical and locational sources of bias.

Because the majority (56 percent) of BPS graduates who enroll in college in the first fall after high school graduation participate in SBC, we selected comparison students from BPS high schools and among other Massachusetts districts with similar characteristics to BPS.[27]

Focal Matching

Focal matching entails matching SBC students with non-SBC students who have similar values for individual self-selection confounders—that is, student-level factors related to the outcomes of interest and to the pairing of SBC students with specific coaching organizations. As mentioned above, we matched SBC and non-SBC students using propensity scores, which represent students' probability of receiving SBC coaching and are calculated as a function of the selection confounders.

When calculating propensity scores, a tension exists between including too many variables and including too few. On the one hand, it is tempting to use every student characteristic available to calculate a propensity score, such that treatment and comparison groups will be balanced on the greatest number of possible confounders. On the other hand, the more variables incorporated into a propensity score, the greater the likelihood that some may not be as balanced as would be using a more parsimonious set of matching variables. Focusing on a smaller set of particularly important variables therefore increases the efficiency of the propensity score to construct matched treatment and comparison groups most balanced on those student characteristics that pose the greatest threat to the calculation of unbiased treatment impacts.

We conducted a thorough literature review to determine pre-treatment (baseline) factors that were shown to be related to our outcomes of interest. Exhibits A-2 and A-3 present the results of this review. We also collected information from coaching organizations about criteria they use when selecting and/or targeting students for their programs. Most organizations indicated that they do not follow a strict selection process based on observable student characteristics when recruiting students.[28] To avoid missing some important confounders, we decided in the estimation of the propensity scores to use all of the relevant variables yielded by the literature review and available in the administrative datasets.[29] Exhibit A-4 lists these variables.

---

[27] Districts that provided comparison students included Avon, Braintree, Brockton, Cambridge, Chelsea, Everett, Fitchburg, Lawrence, Lowell, Malden, Medford, Milton, Norwell, Norwood, Quincy, Randolph, Revere, Somerville, West Bridgewater, Weymouth, and Worcester. Charter schools that provided comparison students include Boston Collegiate Charter School, Boston Green Academy Horace Mann Charter School, Boston Preparatory Charter Public School, City on a Hill Public Charter School Circuit Street, Codman Academy Charter Public School, Edward M. Kennedy Academy for Health Careers, MATCH Charter Public School, Mystic Valley Regional Charter School, Phoenix Charter Academy, Pioneer Charter School of Science, and South Shore Charter Public School.

[28] One coaching organization indicated that it had eligibility criteria that included high school GPA and socioeconomic status indicators.

[29] Some of the variables yielded by the literature review pertain to students' experiences in postsecondary institutions (e.g., employment and on-campus residence). Because these variables are post treatment and may be influenced by coaching, we did not include them in the matching process.

Exhibit A-2: Postsecondary education outcomes literature reviewed

Source	Characteristic Discussed
Advisory Committee on Student Financial Assistance (2013)	ACT scores; full-time status; SES/affordability; spending per student
Allen (1999)	Motivation to finish school
Bridgeman, McCamley-Jenkins, and Ervin (2000)	SAT scores
Brown and Lee (2005)	Race/ethnicity
Buchmann and DiPrete (2006)	Gender
Cabrera, Nora, Castaneda (1992)	SES
Camara and Echternacht (2000)	High school GPA; SAT scores
Community College Leadership Forum (2010)	Faculty hiring practices; full-time status; spending per student
DeAngelo et al. (2011)	Age; first-generation college-goer status; gender; institutional commitment (level of confidence in, and satisfaction with, institutional choice); parental education; postsecondary student achievement; race/ethnicity; SES/affordability; students' residency (on or off campus, near or far from campus)
DeBerard, Spielmans, and Julka (2004)	Early college performance; gender; high school GPA; SAT scores; social support
Dennis, Phinney, and Chuateco (2005)	High school GPA
Feldman (1993)	High school GPA; full-time status; race/ethnicity
Fletcher (2010)	Race/ethnicity
Flores, Batalovo, and Fix (2012)	English language learner status
Frazier et al. (2007)	Learning differences
Gramling (2013)	Full-time status; high school GPA; race/ethnicity; SES/affordability
Harklau et al. (1999)	English language learner status
Horn and Kojaku (2001)	Difficulty of high school curricula
Ishitani and DesJardins (2002)	Financial aid; parental education; SES
Kao and Thompson (2003)	Race/ethnicity
Lotkowski, Robbins, and Noeth (2004)	Academic skills, confidence, goals; ACT scores; financial aid; high school GPA; institutional commitment (level of confidence in, and satisfaction with, institutional choice); institutional selectivity; motivation to finish school; SES; social integration
Osbcorne (2002)	Gender; race/ethnicity
Seidman (2005)	Early college performance; financial aid; gender; high school GPA; on-campus employment; parental education; race/ethnicity; SAT scores; SES; social integration; students' residency (on or off campus, near or far from campus)
Steele (2003)	Race/ethnicity
Trainin and Swanson (2005)	Learning differences
Vogel and Adelman (1992)	Learning differences
Wagner et al. (2005)	Learning differences and other disabilities
Zwick and Skylar (2005)	High school GPA; race/ethnicity; SAT scores

SES=socioeconomic status.
Note: SES measured by free/reduced-price lunch status.

Exhibit A-3: Summary of postsecondary education outcomes literature

Characteristic Discussed	Associated with Outcome		
	Annual Persistence	Academic Achievement	Postsecondary Completion
Academic skills, confidence, goals	X	X	
ACT scores	X	X	X
Age			X
Difficulty of high school curricula (at an individual student level; e.g., number of honors courses taken, etc.)	X	X	
Early college performance	X	X	
English language learner status	X	X	X
Financial aid	X	X	X
Full-time status	X		X
Gender	X	X	X
High school GPA	X	X	X
Institutional selectivity	X		
Learning differences	X	X	X
Motivation to finish school	X	Mixed	X
On-campus employment	X		
Parental education / first-generation college-goer status	X	X	X
Postsecondary student achievement			X
Race/ethnicity	X	X	X
SAT scores	X	X	X
SES	X	X	X
Social integration/support	X	X	
Spending per student			X
Students' residency (on or off campus, near or far from campus)	X		X

SES=socioeconomic status.

Exhibit A-4: Matching characteristics

Variable	Domain	Data Source
Age	Demographics	MA DESE
Gender		
Learning differences		
Race/ethnicity		
SES		
Ever designated as English language learner		
High school suspensions and detentions	Behavioral indicators	MA DESE
High school attendance		
High school GPA	High school performance	MA DESE
SAT scores		
10th-grade MCAS scores		
Advanced course taking in high school		

SES=socioeconomic status; MA DESE=Massachusetts Department of Elementary and Secondary Education; MCAS=Massachusetts Comprehensive Assessment System.

We addressed missing values for the matching characteristics (with the exception of free/reduced-price lunch status—a proxy for SES—and high school GPA, for which there were no missing data among

students in the analysis sample) using the "dummy variable method" (Rosenbaum and Rubin 1984; Stuart 2010).[30]

A.2 Estimation of Propensity Scores

We estimated propensity scores via two logistic regression models specified as below:

1. Logistic model that includes all covariates listed above, plus high school-level average GPA, 10th-grade MCAS scores, and college-going rate, estimated with all SBC students from the 2015 cohort (treatment students) and non-SBC students from the 2015 cohort who enrolled in the same colleges as the treatment students (potential comparison students).

2. Logistic model that includes all covariates listed above, plus high school-level average GPA, 10th-grade MCAS scores, and college-going rate, estimated with all SBC students from the 2016 cohort (treatment students) and non-SBC students from the 2016 cohort who enrolled in the same colleges as the treatment students (potential comparison students).

We estimated separate models for the 2015 and 2016 cohorts in order to capture potential changes in the selection processes employed by the coaching organizations between the two years.

A.3 Conducting Matching and Assessing Quality of the Matches

Matching Methods

We acknowledge that there are many variants of propensity score matching that differ by whether matching is conducted with replacement, how many comparison units are matched with each treatment unit, and whether common support is enforced for each treatment unit (Caliendo and Kopeinig 2008; Smith and Todd 2005; Stuart 2010). We implemented *radius matching*, which entails matching each treatment student with all potential comparison students whose propensity scores are within the pre-specified caliper of his/her score (± 0.4 of the standard deviation of the propensity scores) in his/her block. We also imposed *exact matching* on two variables, *female* and *Black*, to improve balance on those characteristics. We conducted matching with replacement, and matching weights captured the number of comparison units each treatment unit was matched with and vice versa. Treatment students who did not have any potential comparison students within their propensity score caliper were unmatched and excluded from the estimation of SBC effects.

We chose this method as our primary method because it balances the two important aspects of matching: closeness of the matches and the size of the matched groups. Using a caliper ensures that each treatment student is matched with comparison students with sufficiently similar propensity scores and that treatment students without any such matches are excluded. Including all comparison units within the caliper maximizes the size of the analytic sample and statistical power. The results reported in the main text (Chapters 4, 5, and 6) are obtained with the matched groups yielded by this method.

[30] For variables included in the propensity score estimation models, missing rates ranged from 0 percent to 9 percent. The dummy variable method entails replacing the missing values with the sample means and including a dummy variable indicating such values. As Stuart (2010) points out, propensity scores calculated in this manner would match both on observed covariate values and on missing data patterns.

Exhibit A-5 shows the sizes of the matched treatment and comparison groups for each outcome measure. Across the seven measures, between 13 and 17 percent of the SBC students were unmatched for not having a sufficiently similar comparison student. Coached students are more likely to be from groups traditionally underrepresented in college. This made it more difficult to identify adequate comparison students for all coached students. However, to maintain the study's internal validity, it was necessary to include only the coached students for whom we could identify statistically similar comparison students.

Exhibit A-5: Sample sizes for each outcome

	Persistence into Second Year of College	Persistence into Third Year of College	Full-Time Enrollment	Cumulative GPA	Credit Accumulation	Non-Credit-Bearing Semesters	FAFSA Renewal
Matched treatment students (n)	1,234	561	864	814	870	864	821
Non-matched treatment students (n)	246	113	139	135	145	139	123
Non-matched treatment students (%)	17%	17%	14%	14%	14%	14%	13%
Matched comparison group (n)	4,629	2,158	2884	2739	2936	2884	2776
Non-matched potential comparison students (n)	4,270	2,185	550	471	637	550	464
Non-matched potential comparison students (%)	48%	50%	16%	15%	18%	16%	14%

A.4 Matching Diagnostics

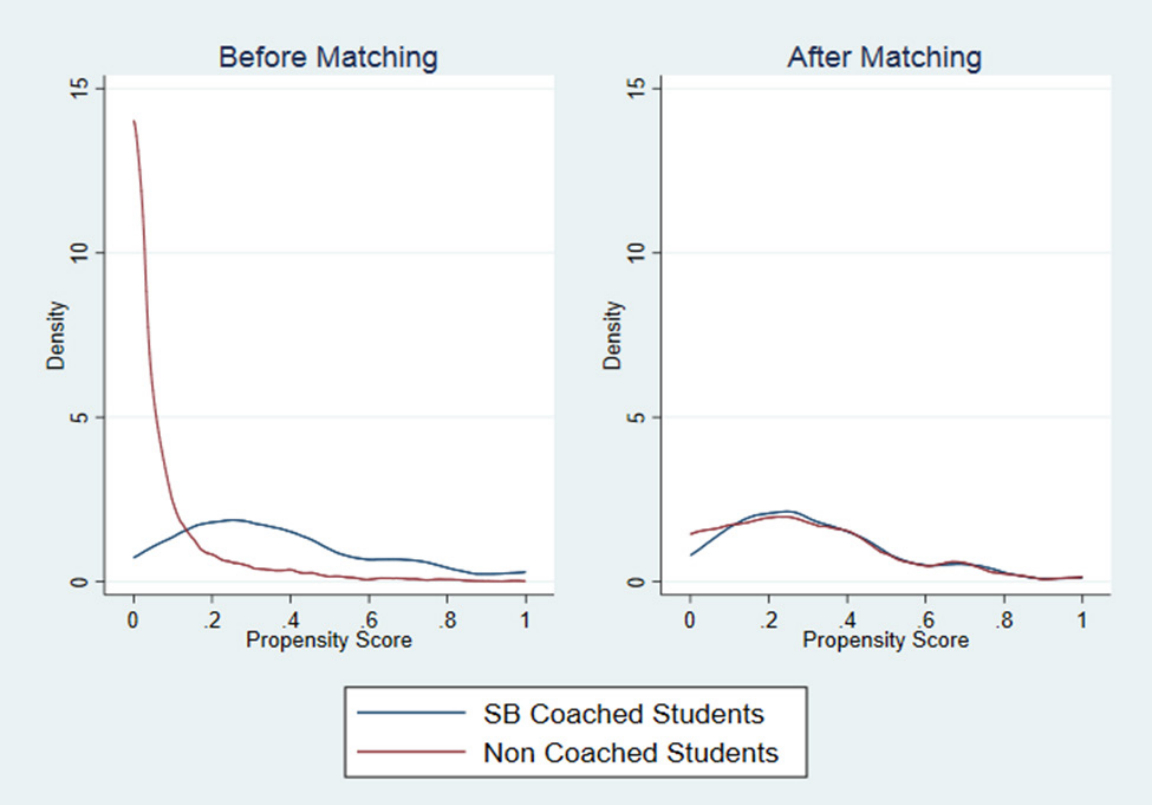
The most important step in matching is to examine to what extent matching worked, by checking the balance of the matched treatment and comparison groups. As explained in more detail below, we assessed the balance of the match by examining the distribution of the propensity scores in the matched treatment and comparison groups and also assessing the standardized difference of each matching variable between the two groups.

We used an iterative process to pick the final matched groups. This process entailed (1) fitting the propensity score model with the matching covariates as described in Section A.2; (2) conducting matching as described in Section A.3.1; and (3) assessing baseline balance. If balance was satisfactory, we deemed the matched groups as final and used them in the estimation of effects. If balance was not satisfactory, we modified the matching mechanism (e.g., by requiring exact matching for the terms with the unbalanced matching variances), and then repeated the whole process until satisfactory balance was achieved. We conducted this process separately for each outcome measure.

Exhibits A-6 and A-7 provide evidence for the balance of the final matched groups for the outcome measure *persistence into second year of college*. The balance of the matched groups for the other outcomes was similar.

The left panel of Exhibit A-6 shows that before matching, distributions of propensity scores for treatment students and potential comparison students were somewhat different, with the latter being more skewed to the right than the former. The right panel shows that matching yields matched treatment and comparison groups with overlapping propensity score distributions.

Exhibit A-6: Distributions of propensity scores



The literature on propensity score matching suggests that having similar propensity score distributions within the matched groups is a necessary but not sufficient condition for having balanced groups (King and Nielsen 2016; Morgan and Winship 2014). Following Rosenbaum and Rubin (1985) and What Works Clearinghouse (2014), we explicitly assessed to what extent matching improved the covariate balance. This we did by examining the standardized differences in the means of each matching covariate between the treatment students and potential comparison students prior to matching and between the matched groups after matching.

We calculated the standardized differences ("effect sizes") as follows: For each matching covariate, we first fit a weighted regression model that used the matching covariate as the dependent variable and the treatment group indicator and indicators for matching blocks (for local matching) as independent variables. The standardized difference was then calculated as the ratio of the coefficient on the treatment indicator to the pooled standard deviation of the matching covariate across the treatment students and potential comparison students. To establish baseline balance between the treatment students and matched comparison students, we required the standardized differences to be less than 15 percent of a standard deviation in absolute value[31] for all matching variables.

Exhibit A-7 shows the standardized baseline differences before and after matching for each outcome measure with radius matching.

As an example, let's examine the differences for *persistence into second year of college*. The first column in the exhibit shows that the pre-matching differences for some variables are notably large: at the student level, −0.54 standard deviation (SD) for *White* and −0.46 SD for *high school GPA*; at the school level, −0.48 SD for *high school-level average English MCAS score* and −1.13 SD for *high school-level average GPA*. The second column shows that matching reduced all of the pre-matching differences that were larger than the 0.15 SD threshold we established (see Section 3.1.2) without distorting the balance for the variables that had been balanced prior to matching. Of the 24 matching variables, the post-matching differences were smaller than 0.05 SD (in absolute value) for 20 variables, between 0.05 and 0.10 SD for three other variables, and between 0.10 and 0.15 SD for the remaining one variable.

The balance estimates for the other six outcomes on Exhibit A-7 were also all below 0.15 SD. Based on these results, we deemed that the matched treatment and comparison groups were balanced and used them in the estimation of SBC effects.

Exhibit A-7: Standardized baseline differences by outcome

Matching Variable	Standardized Difference Before Matching	Standardized Difference After Matching
Outcome: Persistence into Second Year of College		
Demographics		
Age	0.164	−0.026
Female	0.175	0.000
Ever English language learner	0.325	−0.010

[31] Note that this 0.15 criterion is more stringent than what is used by the What Works Clearinghouse, which requires the baseline differences between quasi-experimental treatment and comparison groups be less than 0.25 SD to meet WWC evidence standards.

Matching Variable	Standardized Difference Before Matching	Standardized Difference After Matching
Free/reduced-price lunch eligible	0.168	−0.007
High-incidence disability	0.081	0.008
Low-incidence disability	−0.066	−0.051
Black	0.415	0.000
White	−0.536	−0.002
Asian / Pacific Islander	−0.066	−0.029
Hispanic	0.210	0.028
Other/Multiracial	−0.021	−0.009
Native American	0.014	−0.011
Achievement in High School		
SAT (2400)	−0.242	0.059
MCAS English score	−0.230	0.054
MCAS Math score	−0.118	−0.019
GPA	−0.459	−0.021
Number of advanced courses	−0.008	−0.015
Took any advanced course	0.157	0.012
Behavioral		
Absenteeism	−0.059	0.004
Number of suspensions	−0.087	0.008
High School Characteristics		
High school average Math MCAS score	−0.480	−0.036
High school average English MCAS score	−0.209	−0.032
High school average GPA	−1.135	−0.106
High school college-going rate	0.132	−0.009
Outcome: Persistence into Third Year of College		
Demographics		
Age	0.201	−0.028
Female	0.179	0.000
Ever English language learner	0.230	0.034
Free/reduced-price lunch eligible	0.222	−0.063
High-incidence disability	0.045	−0.030
Low-incidence disability	−0.090	−0.050
Black	0.453	0.000
White	−0.514	−0.020
Asian / Pacific Islander	−0.036	−0.041
Hispanic	0.135	0.069
Other/Multiracial	−0.028	−0.039
Native American	−0.027	−0.069
Achievement in High School		
SAT (2400)	−0.218	0.079
MCAS English score	−0.249	0.072
MCAS Math score	−0.172	−0.059
GPA	−0.524	−0.048
Number of advanced courses	0.016	−0.037
Took any advanced course	0.221	−0.002
Behavioral		
Absenteeism	−0.059	−0.064
Number of suspensions	−0.106	0.050
High School Characteristics		
High school average Math MCAS score	−0.550	−0.035
High school average English MCAS score	−0.331	−0.036
High school average GPA	−1.344	−0.102
High school college-going rate	0.101	0.008

Matching Variable	Standardized Difference Before Matching	Standardized Difference After Matching
Outcome: Full-Time Status		
Demographics		
Age	0.175	−0.023
Female	0.188	0.000
Ever English language learner	0.306	−0.016
Free/reduced-price lunch eligible	0.173	−0.040
High-incidence disability	0.079	0.036
Low-incidence disability	−0.059	−0.028
Black	0.396	0.000
White	−0.539	−0.006
Asian / Pacific Islander	−0.135	−0.100
Hispanic	0.241	0.080
Other/Multiracial	0.005	0.047
Native American	0.008	−0.050
Achievement in High School		
SAT (2400)	−0.251	0.106
MCAS English score	−0.198	0.104
MCAS Math score	−0.155	−0.044
GPA	−0.436	-0.070
Number of advanced courses	0.055	0.010
Took any advanced course	0.172	0.011
Behavioral		
Absenteeism	−0.097	−0.038
Number of suspensions	−0.113	−0.011
High School Characteristics		
High school average Math MCAS score	−0.642	−0.066
High school average English MCAS score	−0.409	−0.086
High school average GPA	−1.099	−0.136
High school college-going rate	−0.008	−0.029
Outcome: GPA		
Demographics		
Age	0.175	−0.035
Female	0.194	0.000
Ever English language learner	0.290	−0.036
Free/reduced-price lunch eligible	0.178	−0.046
High-incidence disability	0.073	0.026
Low-incidence disability	−0.067	−0.035
Black	0.408	0.000
White	−0.535	−0.004
Asian / Pacific Islander	−0.137	−0.104
Hispanic	0.225	0.080
Other/Multiracial	0.021	0.062
Native American	0.013	−0.051
Achievement in High School		
SAT (2400)	−0.248	0.118
MCAS English score	−0.199	0.107
MCAS Math score	−0.147	−0.035
GPA	−0.435	−0.067
Number of advanced courses	0.069	0.021
Took any advanced course	0.177	0.020
Behavioral		
Absenteeism	−0.095	−0.037
Number of suspensions	−0.112	−0.004

Matching Variable	Standardized Difference Before Matching	Standardized Difference After Matching
High School Characteristics		
High school average Math MCAS score	−0.665	−0.070
High school average English MCAS score	−0.439	−0.096
High school average GPA	−1.110	−0.136
High school college-going rate	−0.038	−0.031
Outcome: Semesters Enrolled in Non-Credit-Bearing Courses		
Demographics		
Age	0.175	−0.023
Female	0.188	0.000
Ever English language learner	0.306	−0.016
Free/reduced-price lunch eligible	0.173	−0.040
High-incidence disability	0.079	0.036
Low-incidence disability	−0.059	−0.028
Black	0.396	0.000
White	−0.539	−0.006
Asian / Pacific Islander	−0.135	−0.100
Hispanic	0.241	0.080
Other/Multiracial	0.005	0.047
Native American	0.008	−0.050
Achievement in High School		
SAT (2400)	−0.251	0.106
MCAS English score	−0.198	0.104
MCAS Math score	−0.155	−0.044
GPA	−0.436	−0.070
Number of advanced courses	0.055	0.010
Took any advanced course	0.172	0.011
Behavioral		
Absenteeism	−0.097	−0.038
Number of suspensions	−0.113	-0.011
High School Characteristics		
High school average Math MCAS score	−0.642	−0.066
High school average English MCAS score	−0.409	-0.086
High school average GPA	−1.099	−0.136
High school college-going rate	−0.008	−0.029
Outcome: Credit Accumulation		
Demographics		
Age	0.181	−0.023
Female	0.182	0.000
Ever English language learner	0.298	−0.020
Free/reduced-price lunch eligible	0.175	−0.038
High-incidence disability	0.081	0.034
Low-incidence disability	−0.059	−0.030
Black	0.392	0.000
White	−0.539	−0.003
Asian / Pacific Islander	−0.135	−0.102
Hispanic	0.248	0.079
Other/Multiracial	0.002	0.046
Native American	0.008	−0.050
Achievement in High School		
SAT (2400)	−0.247	0.108
MCAS English score	−0.204	0.106
MCAS Math score	−0.155	−0.045
GPA	−0.446	−0.074

Matching Variable	Standardized Difference Before Matching	Standardized Difference After Matching
Number of advanced courses	0.058	0.007
Took any advanced course	0.175	0.008
Behavioral		
Absenteeism	−0.099	−0.044
Number of suspensions	−0.114	−0.012
High School Characteristics		
High school average Math MCAS score	−0.652	−0.068
High school average English MCAS score	−0.415	−0.087
High school average GPA	−1.121	−0.140
High school college-going rate	−0.007	−0.034
Outcome: FAFSA Renewal		
Demographics		
Age	0.184	−0.023
Female	0.184	0.000
Ever English language learner	0.297	−0.017
Free/reduced-price lunch eligible	0.170	−0.056
High-incidence disability	0.086	0.064
Low-incidence disability	−0.068	−0.043
Black	0.376	0.000
White	−0.532	−0.002
Asian / Pacific Islander	−0.136	−0.107
Hispanic	0.255	0.086
Other/Multiracial	0.003	0.038
Native American	0.008	−0.052
Achievement in High School		
SAT (2400)	−0.236	0.106
MCAS English score	−0.185	0.116
MCAS Math score	−0.147	−0.050
GPA	−0.439	−0.074
Number of advanced courses	0.067	0.024
Took any advanced course	0.188	0.020
Behavioral		
Absenteeism	−0.095	−0.035
Number of suspensions	−0.109	−0.008
High School Characteristics		
High school average Math MCAS score	−0.637	−0.068
High school average English MCAS score	−0.410	−0.095
High school average GPA	−1.091	−0.140
High school college-going rate	0.001	−0.032

Appendix B. Details About the Estimation of Effects and Sensitivity Analyses

B.1 Analytic Approach for Estimating the Average Impact of the Program

To address the primary research question about the impact of SBC on all students, we estimated the following model with the full analytic sample (all SBC students and matched comparison students from the two cohorts with valid data):

(Eq. 1) $Y_{ij} = \pi_0 + \pi_1 T_{ij} + \sum_{b=1}^{B-1} \pi_{(1+b)} I_{ij}^{b} + \sum_{n=1}^{N} \pi_{(B+n)} X_{ij}^{n} + \varepsilon_{ij}$

where:

Y_{ij} = outcome measure for student *i* in matching block *j*.

T_{ij} = treatment indicator for student *i* in block *j*, which equals 1 if student *i* is an SBC student and 0 otherwise.

I_{ij}^{b} = indicator variable for the b[th] matching block for student *i*. It equals 1 if student *i* is a member of the b-[th] block and 0 otherwise. A matching block was defined by the college and cohort.

X_{ij}^{n} = n[th] matching characteristic or covariate for student *i* in block *j*. Similar to the propensity score models, missing values of the covariates were addressed using the dummy variable method.[32]

ε_{ij} = random error term for student *i* in school *j*, which is assumed to be normally distributed with mean 0 and variance of σ_{ε}^{2}.

We estimated this model separately for each outcome measure using the matching weights specific to each outcome measure. Because treatment students and potential comparison students with missing outcome data were not included in the matching process, they were not included in the estimation of the effects. In the estimated mode, the coefficient estimate on the treatment indicator, π_1, was interpreted as the average impact of participating in SBC coaching.

Two aspects of the model in Equation 1 are worthy of further explanation. First, the model does not include a separate random error term for college to capture potential clustering of outcome measures within colleges, because we anticipate that such clustering (i.e., the dependence of outcomes of students from the same college) will be fully explained by the matching block indicators already included in the

[32] Free/reduced-price lunch and GPA baseline covariates are identified as primary by the U.S. Department of Education's What Works Clearinghouse, and therefore were not imputed using the dummy variable method. Students missing values on either of these two covariates are dropped from the analysis.

model.[33] Similarly, the model does not include a separate indicator for students in the 2015 or 2016 cohort because the block indicators are cohort specific.

Second, the independent variables of the model included the matching characteristics that were available for all students and used in the matching process. The purpose of this was to increase the precision of the effect estimates (because these covariates were expected to explain some of the residual variance of the outcome measures) and be doubly robust (Bang and Robins 2005; Tan 2006).[34,35] Section B.3 presents results from alternative specifications that did not control for the matching covariates.

B.2 Analytic Approach for Exploratory Subgroup Analyses

As described in Chapters 5 and 6, we examined (1) pre-treatment student characteristics that were exogenous (not related to program participation or effects) and available for both the treatment students and comparison students (gender, race/ethnicity, high school GPA, type of college, and the interaction between race/ethnicity and gender); and (2) features of coaching such as aspects of the student and coach interactions that were potentially endogenous (may be related to program participation and effects) and only available for treatment students. The analytic approach and interpretation of the resulting relationships between program effects and student characteristics and features of coaching differed for each set of analyses, and it is described in detail below.

Specifically, we examined the extent to which program effects were related to exogenous student characteristics using a slightly modified version of the impact model in Equation 1 to include the interaction of the treatment indicator T_{ij} and the characteristic that is being tested. To simplify the analyses and ease the interpretation of the results, we transformed each continuous and categorical variable into a binary variable. Specifically, when examining race/ethnicity, we created a binary variable *underrepresented students in postsecondary education*, which was set to 1 for Black, Hispanic, Native American, and Other/Multiracial students, and 0 for the remaining students. When examining high school GPA, the binary variable *higher high school GPA* was set to 1 for students whose high school GPA was greater than 3.00 on a four-point scale, and 0 for the remaining students.

We specified the modified version of the impact model that included the interaction term as follows:

$$\text{(Eq. 2)} \quad Y_{ij} = \pi_0 + \pi_1 T_{ij} + \pi_2 T_{ij} M_{ij} + \sum_{b=1}^{B-1} \pi_{(2+b)} I_{ij}^{b} + \sum_{n=1}^{N} \pi_{(1+B+n)} X_{ij}^{n} + \varepsilon_{ij}$$

[33] We tested the validity of this assumption by estimating hierarchical linear models that nest students within colleges. The variance of the college random effect was essentially zero for all outcome measures, and the hierarchical linear models yielded very similar estimates to the single-level model in Equation 1.

[34] Using the baseline characteristics in the matching process and *also* using them as covariates in the estimation of impacts is deemed to give the analyst two chances to get the "right" model specification (once in the propensity model and once in the impact model for the outcome measure). Therefore, these estimators are called "doubly robust."

[35] We considered using a second set of covariates that were measured post high school and potentially associated with the outcomes of interest, such as the location of students' residencies during college (on or off campus, near or far from campus) and whether they held an on-campus job. We decided not to use them, as we were not confident they were exogenous (not influenced by participating in SBC coaching).

In Equation 2, M_{ij} denotes the binary student characteristic. As an example, assume that M_{ij} was an indicator for female students (set to 1 if student i was female and to 0 if student i was male). In this case, the estimate of π_1 captures the effect estimate for male students, and the estimate of π_2 captures the difference in the estimated effects between females and males. The effect estimate for females can be calculated by adding the two coefficients.

We had to adopt a different strategy for the second type of analyses, because the features of coaching were not defined for the matched comparison students. Similar to the exogenous student characteristics, we first created binary indicators for each endogenous feature of coaching. For example, the value of the *higher academic focus* variable was set to 1 for the treatment students who had more than four meetings with their coach and 0 for the remaining students. We then estimated the model in Equation 1 with the treatment students with a particular value of the binary variable and their matched comparison students (with modified weights for the matched comparison students), which yielded the effect estimate specific to that subgroup (e.g., "higher academic focus") and estimated another model for the remaining treatment students and their matched comparison students (e.g., "lower academic focus"). We tested the statistical significance of the difference between the subgroups effects via a Wald test.

The features of coaching are endogenous and measured post treatment; therefore, they may have been influenced by the program effects. For example, students who were in the "higher academic focus" subgroup may have had academic struggles initially and lower academic outcomes; consequently, their interactions with the coach may have had a higher academic focus than the students who did not have such struggles and had better outcomes. In this case, the estimated effects may be lower for the students in the "higher academic focus" subgroup than for the rest of the sample, but this difference should not be attributed to the focus of coaching. Therefore, we recommend caution with the causal interpretation of the analyses conducted with endogenous variables.

Chapters 5 and 6 of the report showed the estimated subgroup effects, their statistical significance, and the differences between the subgroup effects and the statistical significance of the differences. Appendix C shows more details for these results, including standard errors of the subgroup effects and sample sizes.

B.3 Robustness Checks and Sensitivity Analyses

Recall that the results presented in Chapter 4 for the full sample were produced by the impact model that used all of the matching covariates, with the matched comparison group yielded by radius matching. We conducted additional analyses testing the robustness of these results to alternative model specifications and sample definitions. This subsection summarizes the results of these sensitivity analyses.

The first analysis assessed the robustness of the reported results to covariates used in the impact model given in Equation 1. Specifically, we estimated two alternative versions of this model: (1) no matching covariates or matching blocks; and (2) with matching blocks but no matching covariates.

Results are presented in Exhibit B-1. The first panel in this exhibit repeats the results from our preferred specification (from Exhibits 4-1 to 4-6), whereas the second and third panels use the alternative specifications described above. Exhibit B-1 shows that the magnitudes of the effect estimates were not very sensitive to whether the model included matching blocks or to the set of covariates that were controlled for. The inclusion of additional covariates helps with the precision of effect estimates—standard errors of the preferred specification were 2-40 percent lower than those from the model that did not control for any covariates or matching blocks.

Exhibit B-1: Robustness checks, by included covariates

Outcome	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect Size	Standard Error	Sample Size
Matching Blocks and All Matching Covariates[a] (Preferred Specification)						
Persistence						
Persistence into second year of college	83%	78%	4.26*	0.11	1.58	5,863
Persistence into third year of college	75%	72%	2.83	0.06	2.64	2,719
Full−time status	67%	64%	3.70*	0.11	1.51	3,748
Achievement						
Cumulative GPA	2.37	2.36	0.01	0.01	0.05	3,553
Semesters enrolled in non-credit-bearing courses	0.67	0.65	0.03	0.03	0.03	3,748
Credit accumulation	46%	42%	3.06*	0.12	1.34	3,806
Financial Aid						
FAFSA renewal	84%	78%	6.22**	0.15	2.01	3,597
No Covariates (Alternative Specification 1)						
Persistence						
Persistence into second year of college	83%	79%	4.01*	0.10	1.71	5,863
Persistence into third year of college	75%	73%	1.69	0.04	2.95	2,719
Full−time status	67%	64%	3.27~	0.09	1.8	3,748
Achievement						
Cumulative GPA	2.37	2.37	0	0.00	0.06	3,553
Semesters enrolled in non-credit-bearing courses	0.67	0.67	0.01	0.01	0.05	3,748
Credit accumulation	46%	43%	2.49	0.09	1.55	3,806
Financial Aid						
FAFSA renewal	84%	78%	5.70**	0.14	2.06	3,597
No Covariates, Controlling for Matching Blocks (Alternative Specification 2)						
Persistence						
Persistence into second year of college	83%	79%	4.01*	0.10	1.63	5,863
Persistence into third year of college	75%	73%	1.69	0.04	2.93	2,719
Full−time status	67%	64%	3.27*	0.09	1.63	3,748
Achievement						
Cumulative GPA	2.37	2.37	0	0.00	0.06	3,553
Semesters enrolled in non-credit-bearing courses	0.67	0.67	0.01	0.01	0.04	3,748
Credit accumulation	46%	43%	2.49~	0.09	1.47	3,806
Financial Aid						
FAFSA renewal	84%	78%	5.70**	0.14	1.98	3,597

[a]: The covariates include all 19 student and high school characteristics listed in Exhibit A-7.
~ Indicates statistical significance at the 10 percent level.
* Indicates statistical significance at the 5 percent level.
** Indicates statistical significance at the 1 percent level.

Finally, Exhibit B-2 shows the effect estimates for two outcomes obtained from NSC with the reduced sample that includes only students from the 11 colleges in the college administrative dataset. These estimates are similar to the estimates obtained with the full sample shown in Exhibit B-1. However, although the magnitudes of the impacts on persistence into the second year are similar in the reduced sample and the full sample, they do differ—that is, unlike in the full sample, in the smaller, college administrative sample, the impact on persistence into the second year of college is not statistically

significant at the 5 level (rather, it is only significant at the 10 level). In both samples, the impact on persistence into the third year of college is positive but not statistically significant. The smaller size of the college administrative sample could be the reason why the impact on persistence into the second year is not statistically significant at the 5 percent level with this limited sample. Smaller sample sizes tend to have larger standard errors, reducing our ability to detect significant impacts.

Exhibit B-2: National Student Clearinghouse outcomes for college administrative data sample

Outcome	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Standard Error	Sample Size
Persistence (Reduced Sample)					
Persistence into second year of college	81%	78%	3.28~	1.98	3,781
Persistence into third year of college	72%	73%	-0.86	3.19	1,977

~ Indicates statistical significance at the 10 percent level.

Appendix C. Variation across Student Characteristics and Features of Coaching, All Outcomes

Exhibit C-1: Impacts by gender

Outcome	Female Students						Male Students						Difference
	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect Size	Standard Error	Sample Size	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect Size	Standard Error	Sample Size	
Persistence													
Persistence into second year of college	86%	83%	2.97	0.08	1.96	3,156	78%	72%	6.21*	0.14	2.60	2,707	−3.24
Persistence into third year of college	81%	76%	4.77	0.12	3.42	1,426	66%	65%	0.22	0.00	4.14	1,293	4.55
Achievement													
Cumulative GPA	2.52	2.54	−0.01	-0.01	0.07	1,850	2.18	2.13	0.05	0.04	0.08	1,703	−0.06
Financial aid													
FAFSA renewal	88%	81%	7.42*	0.20	2.51	1,880	78%	74%	4.55	0.11	3.18	1,717	2.87

Source: National Student Clearinghouse data from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data.
* Indicates statistical significance at the 5 percent level

Exhibit C-2: Impacts by race/ethnicity

Outcome	Underrepresented Minority Students						Not Underrepresented Minority Students						Difference
	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect Size	Standard Error	Sample Size	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect Size	Standard Error	Sample Size	
Persistence													
Persistence into second year of college	81%	75%	5.06*	0.12	1.95	3,220	89%	87%	1.79	0.05	2.37	2,643	3.27
Persistence into third year of college	74%	68%	5.68	0.12	3.24	1,510	78%	83%	−5.14	-0.14	4.47	1,209	10.82
Achievement													
Cumulative GPA	2.25	2.26	−0.01	-0.01	0.07	1,960	2.79	2.72	0.07	0.07	0.08	1,593	−0.08
Financial Aid													
FAFSA renewal	82%	76%	5.37*	0.13	2.43	2,005	92%	83%	9.01*	0.25	2.74	1,592	−3.64

Source: National Student Clearinghouse data from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data.
* Indicates statistical significance at the 5 percent level.

Exhibit C-3: Impacts by high school academic achievement

Outcome	Students with a Higher GPA						Students with a Lower GPA						Difference
	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect Size	Standard Error	Sample Size	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect Size	Standard Error	Sample Size	
Persistence													
Persistence into second year of college	91%	88%	3.23	0.10	2.21	2,513	79%	74%	4.79*	0.11	2.07	3,350	−1.55
Persistence into third year of college	87%	85%	2.14	0.06	4.62	1,054	70%	67%	3.13	0.07	3.16	1,665	−0.99
Achievement													
Cumulative GPA	2.82	2.80	0.02	0.02	0.09	1,412	2.22	2.19	0.01	0.01	0.07	2,141	0.01
Financial Aid													
FAFSA renewal	90%	84%	5.72	0.15	3.08	1,434	82%	75%	6.40*	0.15	2.48	2,163	−0.68

Source: National Student Clearinghouse data from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data.
* Indicates statistical significance at the 5 percent level

Exhibit C-4: Impacts by college type

Outcome	Students Inititally Enrolled at a **Two**−Year College						Students Inititally Enrolled at a Four−**Year** College						Difference
	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect Size	Standard Error	Sample Size	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect Size	Standard Error	Sample Size	
Persistence													
Persistence into second year of college	66%	60%	5.48	0.11	3.48	1,753	91%	88%	3.63*	0.11	1.57	4,110	1.85
Persistence into third year of college	62%	55%	6.55	0.13	4.79	856	82%	81%	0.72	0.02	3.14	1,863	5.83
Achievement													
Cumulative GPA	2.07	2.03	0.04	0.04	0.09	1,285	2.58	2.59	-0.01	-0.01	0.06	2,268	0.05
Financial Aid													
FAFSA renewal	77%	66%	10.58*	0.23	3.83	1,239	88%	85%	3.54	0.10	2.13	2,358	7.03

Source: National Student Clearinghouse data from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data.
* Indicates statistical significance at the 5 percent level

Exhibit C-5: Impacts by number of coaching interactions focused on academic topics

Outcome	Students who Experienced a Higher Number of Coaching Interactions Focused on Academic Topics						Students who Experienced a Lower Number of Coaching Interactions Focused on Academic Topics						Difference
	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect SIze	Standard Error	Sample Size	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect SIze	Standard Error	Sample Size	
Persistence													
Persistence into second year of college	86%	77%	9.07*	0.22	2.19	3,626	80%	80%	0.59	0.02	1.89	4,620	8.48*
Persistence into third year of college	75%	66%	9.89*	0.21	4.05	1,518	75%	76%	−1.56	-0.04	2.90	2,263	11.45*
Achievement													
Cumulative GPA	2.42	2.30	0.12	0.11	0.07	2,228	2.35	2.43	−0.09	-0.08	0.07	2,753	0.21*
Financial Aid													
FAFSA renewal	89%	75%	13.67*	0.33	2.74	2,191	81%	78%	2.27	0.06	2.37	2,945	11.40*

Source: National Student Clearinghouse data from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data; program database data collected from the coaching organizations.
* Indicates statistical significance at the 5 percent level

Exhibit C-6: Impacts by number of coaching interactions focused on financial aid topics

Outcome	Students who Experienced a Higher Number of Coaching Interactions Focused on Financial Aid Topics						Students who Experienced a Lower Number of Coaching Interactions Focused on Financial Aid Topics						Difference
	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect SIze	Standard Error	Sample Size	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect SIze	Standard Error	Sample Size	
Persistence													
Persistence into second year of college	87%	77%	9.68*	0.24	1.98	3,990	79%	79%	0.12	0.00	2.00	4,447	9.56*
Persistence into third year of college	78%	71%	6.82*	0.15	3.40	1,948	71%	72%	−0.64	-0.01	3.28	2,063	7.46
Achievement													
Cumulative GPA	2.41	2.3	0.11	0.10	0.07	2,461	2.35	2.42	−0.07	-0.07	0.06	2,842	0.18
Financial Aid													
FAFSA renewal	90%	78%	12.60*	0.31	2.49	2,486	80%	77%	2.28	0.06	2.49	2,889	10.32*

Source: National Student Clearinghouse data from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data; program database data collected from the coaching organizations.
* Indicates statistical significance at the 5 percent level

Exhibit C-7: Impacts by number of coaching interactions focused on managing life responsibilities

Outcome	Students who Experienced a Higher Number of Coaching Interactions Focused on General Life Topics						Students who Experienced a Lower Number of Coaching Interactions Focused on General Life Topics						Difference
	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect SIze	Standard Error	Sample Size	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect SIze	Standard Error	Sample Size	
Persistence													
Persistence into second year of college	79%	70%	9.45*	0.21	2.52	3,175	85%	84%	0.77	0.02	1.78	4,947	8.68*
Persistence into third year of college	67%	66%	1.51	0.03	4.21	1,495	80%	77%	2.92	0.07	3.03	2,315	−1.41
Achievement													
Cumulative GPA	2.23	2.21	0.02	0.02	0.08	2,136	2.49	2.5	−0.01	-0.01	0.06	2,988	0.03
Financial Aid													
FAFSA renewal	83%	74%	9.34*	0.22	3.07	2,128	85%	80%	5.23*	0.13	2.38	3,060	4.11

Source: National Student Clearinghouse data from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data; program database data collected from the coaching organizations.
* Indicates statistical significance at the 5 percent level

Exhibit C-8: Impacts by number of coaching interactions focusing on career topics

Outcome	Students who Experienced a Higher Number of Coaching Interactions Focused on Career Topics						Students who Experienced a Lower Number of Coaching Interactions Focused on Career Topics						Difference
	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect SIze	Standard Error	Sample Size	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect SIze	Standard Error	Sample Size	
Persistence													
Persistence into second year of college	85%	76%	9.77*	0.23	2.15	3,892	81%	81%	.08	0.00	1.92	4,502	9.69*
Persistence into third year of college	77%	69%	8.26*	0.18	3.68	1,649	73%	75%	−1.75	-0.04	3.12	2,061	10.00*
Achievement													
Cumulative GPA	2.43	2.33	0.10	0.09	0.07	2,337	2.33	2.4	−0.07	-0.06	0.07	2,850	0.17
Financial Aid													
FAFSA renewal	87%	76%	11.35*	0.28	2.78	2,341	82%	78%	3.88	0.10	2.41	2,937	7.47*

Source: National Student Clearinghouse data from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data; program database data collected from the coaching organizations.
* Indicates statistical significance at the 5 percent level

Exhibit C-9: Impacts by frequency of coaching interactions

Outcome	High Frequency						Low Frequency						Difference
	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect SIze	Standard Error	Sample Size	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect SIze	Standard Error	Sample Size	
Persistence													
Persistence into second year of college	88%	76%	11.94*	0.29	2.05	3,951	79%	80%	−1.96	-0.05	2.01	4,580	13.91*
Persistence into third year of college	77%	67%	10.94*	0.24	4.10	1,631	73%	75%	−1.62	-0.04	3.04	2,210	12.56*
Achievement													
Cumulative GPA	2.45	2.29	0.16*	0.15	0.07	2,283	2.32	2.42	−0.1	-0.09	0.07	2,933	0.26*
Financial Aid													
FAFSA renewal	89%	75%	13.29*	0.32	2.76	2,281	81%	79%	2.87	0.07	2.46	3,029	10.42*

Source: National Student Clearinghouse data from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data; program database data collected from the coaching organizations.
* Indicates statistical significance at the 5 percent level

Exhibit C-10: Impacts by duration of one-on-one coaching interactions

Outcome	High Duration						Low Duration						Difference
	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect SIze	Standard Error	Sample Size	Treatment Group Mean	Adjusted Comparison Group Mean	Program Impact	Effect SIze	Standard Error	Sample Size	
Persistence													
Persistence into second year of college	85%	84%	1.41	0.04	1.80	4,481	80%	73%	7.29*	0.17	2.27	4,062	−5.89*
Persistence into third year of college	76%	75%	0.94	0.02	3.08	1,905	74%	69%	4.82	0.10	3.53	2,178	−3.88
Achievement													
Cumulative GPA	2.49	2.48	0.01	0.01	0.07	2,697	2.24	2.24	0.00	0.00	0.07	2,765	0.01
Financial Aid													
FAFSA renewal	88%	78%	9.08*	0.23	2.46	2,703	80%	76%	4.34	0.10	2.68	2,788	4.74

Source: National Student Clearinghouse data from Boston Public Schools and Massachusetts Department of Elementary and Secondary Education; college administrative data; program database data collected from the coaching organizations.
* Indicates statistical significance at the 5 percent level

Appendix D. Resources Used to Implement Coaching

This appendix describes the costs of implementing SBC. We report annual costs for the 2016-17 academic years and compare them to the 2015-16 costs (previously reported in Linkow et al. 2017b).

For the purposes of this report, we define *costs* as the monetized value (in dollar terms) of the resources required to implement transition coaching.

We purposefully include resources for the nonprofit organizations that provided the coaching, as well as the resources provided by TBF and other network-wide community partners (i.e., colleges, uAspire, and BPS) whose contributions included thought leadership, coordination, and managing the SIF grant and Success Boston Coaching program. All of those resources combined represent the actual costs of operating the SBC program. We characterize these resources in this chapter by aggregating their monetized value, and then dividing the total amount by the number of students served.

We report the combined overall costs of the SBC program in the first section of this appendix. In the second section, we characterize variation in costs across the eight nonprofit organizations. We review learning points and discuss costs in the third section, and close with an overview of our cost data collection and analysis methodology.[36]

Key Findings

- The combined monetized amount of resources used to implement coaching increased from $5,214,547 in 2015-16 to $6,167,058 in 2016-17.
- The total cost per student in 2015-16 was $5,762. The cost per coached student fell to $3,037 in 2016-17. These lower costs per student were primarily due to the addition of the second cohort beginning the two-year program while the first cohort continued to receive their second year of coaching. The 2015-16 academic year was also the first fiscal year in which both TBF and participating nonprofit organizations operated SBC using SIF grant funds, and therefore included additional administrative costs associated with program start-up.
- The total number of students enrolled in SBC was 2,030 in 2016-17, an increase of almost 125 percent over the number of coached students in 2015-16, which was 905. Meanwhile, the total resources used to implement the program increased by 18 percent from 2015-16 to 2016-17.
- Given that the SBC coaching model hinges on one-on-one support provided by coaches to students, the cost of hiring the personnel who provide that support represents the largest input, or 41 percent of total costs. In monetary terms, this translates into an investment in staff supporting SBC students of $1,227 per student (in 2016-17).

Overall Costs of SBC

This section presents the overall cost of transition coaching in each of the first two years of the program's implementation. Costs are broken out for the two main resource categories, **administrative functions** and

[36] Costs for 2015-2016 differ somewhat in this report from those reported in the 2016 BosC4C Implementation Report (Linkow et al., 2016) for two primary reasons. First, we recalculated 2015-2016 total costs excluding the costs of a nonprofit that did not continue in the BosC4C program for subsequent years. This allows costs in subsequent years to be more comparable and makes the cost analysis approach parallel to the outcome analysis, which does not include coached students in this nonprofit in the analysis sample. Second, 2015-16 costs are multiplied by a 1.058 inflation factor (from the Consumer Price Index for the Boston metro area) to allow comparisons across years. Similarly, we apply a 1.032 inflation factor to all 2016-17 costs.

direct services to students, as well as into the subcategories listed in Exhibit D-5 in the methodology section below. Cost figures have been rounded to nearest whole dollar amounts and adjusted for inflation (to 2018 dollars).

The total value of resources used to implement the SBC coaching grew from an estimated $5,214,547 (adjusted for inflation to 2018 dollars) in the 2015-16 school year to $6,167,085 in the 2016-17 school year.

During 2015-16, the eight nonprofit programs delivering coaching collectively enrolled 905 students.[37] This represents the first year of the scaled up program, comprising only the 2015 cohort of students. A slightly larger 2016 cohort was enrolled in 2016-17. Because coaching is available to students for two years, 2016-2017 sees 905 students from the 2015 cohort receive their second year of coaching and 1,125 students from the 2016 cohort receive their first year of coaching. Thus, the total number of coached students more than doubles from 2015-16 to 2016-17, from 905 to 2,030 students.

Dividing the total resources used to provide SBC coaching by the number of coached students results in an average cost per student of $5,762 in the 2015-16 first year of the program. In 2016-17, whereas the number of coached students increases in size by 125 percent, resources used for the program increase by 18 percent. The resulting average cost per coached student in 2016-17 is $3,038.

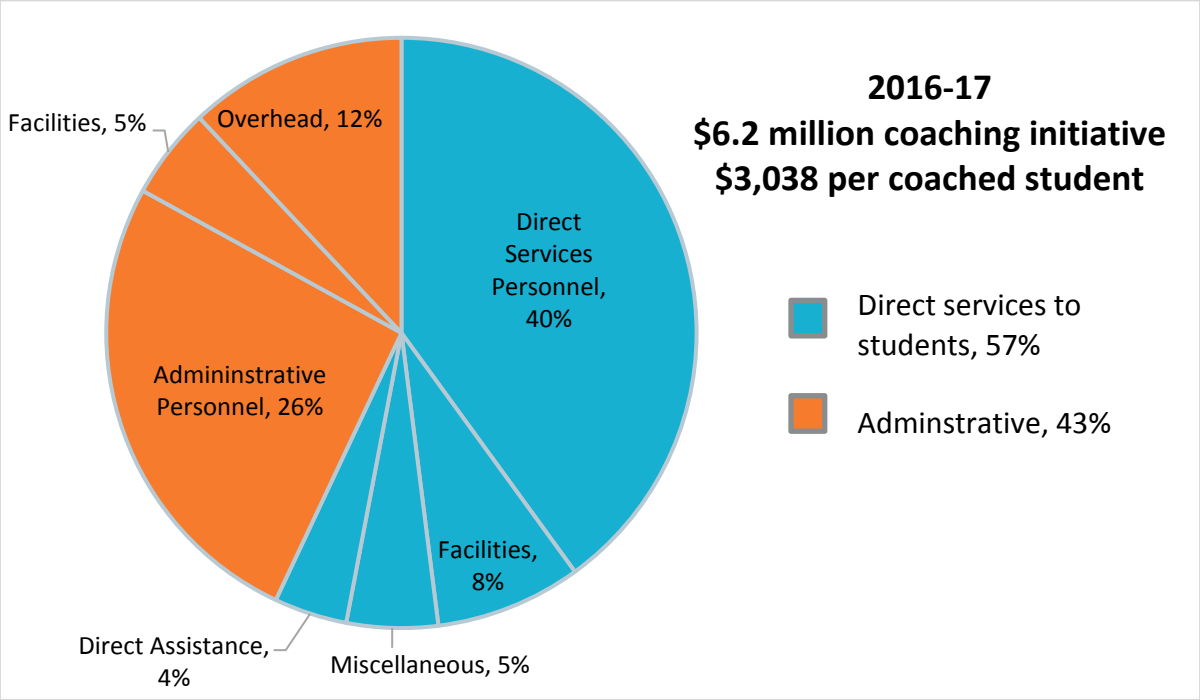
Exhibit D-1: Summary of total costs in first two years of SBC program across organizations

Year	Total Resources	Coached Students	Average Cost per Student	% Administrative, Direct Service
2015-16	$5,214,547	905	$5,762	54% direct services to students 46% administrative
2016-17	$6,167,085	2,030	$3,037	57% direct services to students 43% administrative

The final column of Exhibit D-1 shows that the average share of resources used in direct service coaching activities—which we define in the cost study as activities including student interaction—increases from 54 percent in 2015-16 to 57 percent in 2016-17. Additional detail on how we categorize the comprehensive list of inputs into two groups is provided later, in the methodology section. A more detailed breakdown for the 2016-17 year only is presented in Exhibit D-2.

[37] The fiscal year periods (April 1–March 30) for which we collect cost data substantially overlap but do not perfectly align with the academic years over which outcome data are collected (August 1–June 1). The cost data collection window was driven by the SIF grant fiscal year and to facilitate real-time data collection. However, because a full year of costs are captured, these costs are representative of the annual cost of the SBC initiative.

Exhibit D-2: Importance of resource types in costs of transition coaching



Source: SBC cost data.

Direct Services Subcategories. Within the direct services category, **compensation of personnel** who provide direct services to students (primarily SBC coaches) represents the largest cost subcategory (40 percent of total costs). The next largest subcategory is the cost of **facilities** in which coaching interactions occur (8 percent). These facilities costs include space used by coaches onsite at nonprofits, such as offices and common space where they interact with students. Most frequently, when coaches interact with students on campuses, they make use of available public space—typically common areas or empty classrooms—and we do not consider the cost of this space in calculating the cost of the coaching program, because the space would not otherwise be used for some other coaching purpose. However, coaches from five nonprofits use dedicated space on a college campus for coaching interactions on a regular basis—usually a classroom for up to a few hours a week. The space is donated as an in-kind resource by the college; we do take the value of those in-kind resources into account in estimating the direct services facilities cost.

The smallest subcategory is **direct assistance** (4 percent). This subcategory includes limited direct cash or cash-equivalent assistance to students such as transportation subsidies and includes scholarships that are reserved specifically for a small number of coached students.

Administrative Subcategories. Within the administrative category, again **personnel** costs represent the largest share of resources, comprising 26 percent of total costs. These staff train and manage coaches and otherwise administer the coaching program. Organizational **overhead** costs account for 12 percent and include professional services (e.g., accounting and finance), technology and insurance costs, and national organization fees. Finally, the occupancy costs associated with **facilities** used by administrative staff and for administrative functions represent 5 percent of total costs.

Variation in Costs across Nonprofit Organizations

Not surprisingly, the costs of coaching vary across organizations for multiple reasons, including organization size, number of students with whom coaches work, and differences in the amount of time coaches spend with each student and where coaching occurs. In particular, at the organization level, per-coach caseloads range from 12 to 65. To some extent, the availability of resources also contributes to cost differences. Some organizations are housed in larger facilities with more space for both administration and programming; others rely on smaller spaces, as well as on community resources for activities such as career nights or tutoring and mentoring programs. Program size and coaches' level of experience are also related to costs per student. Organizations serving larger numbers of students tend to have lower per-student costs. Differences in administrative cost share do not appear to drive the differences in costs, as discussed in greater detail below.

Exhibit D-3 presents summary statistics to describe the variation in program-level costs, total number of coached students, and the percentage of costs allocated to direct services across the eight nonprofit programs for each of the first two years of the program scale-up.[38] In the second year, there was a smaller range in cost per student across the nonprofit organizations. The maximum cost per student decreased by more than one half from 2015-16 to 2016-17.

Exhibit D-3: Summary of variation in individual nonprofits' program-level costs in first two years of SBC

	Academic Year	
Program Characteristic	2015-16	2016-17
Cost per coached student	Min: $2,944 Median: $5,438 Max: $15,537	Min: $1,697 Median: $2,289 Max: $7,735
# of coached students	Min: 34 Median: 69 Max: 386	Min: 94 Median: 148 Max: 890
Costs allocated to direct services	Min: 42% Median: 56% Max: 73%	Min: 45% Median: 61% Max: 72%

Source: SBC cost data.

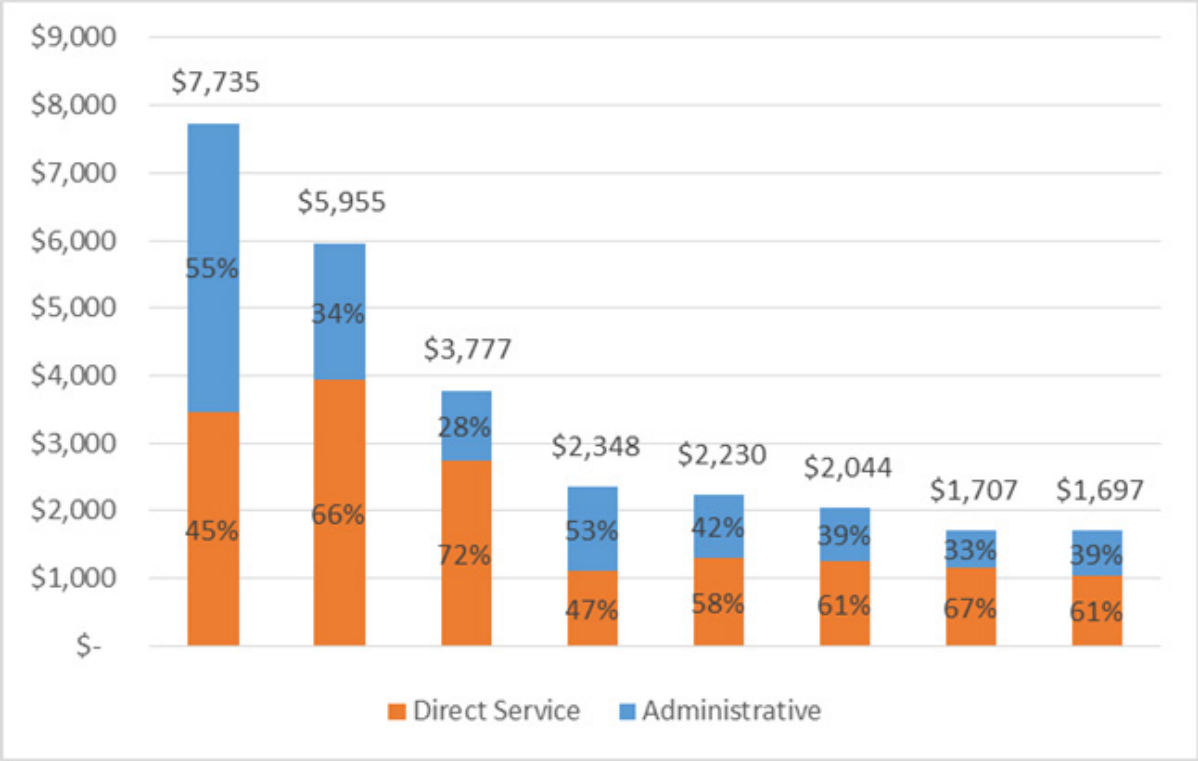
Exhibit D-4 below provides additional detail on the variation in costs per coached student and in the mix of resources used to support coaching activities. Costs are shown for the 2016-17 year. Five of the eight programs have costs below $2,400 per coached student. The number of students served is related to per-student costs in that the two organizations with the most coached students are among the three lowest in per-student costs in each year.

Key drivers of administrative costs include the amount of time senior staff spend training and monitoring coaches and how much experience the organization has with coaching. Also related to administrative costs are whether the organization has other programming that shares overhead resources, its experience

[38] To encourage candid participate in the cost analysis, we do not link organizations by name to costs reported in this section.

with federal grant compliance, and whether the organization pays membership fees to a larger or parent organization.

Exhibit D-4. Costs per coached student of coaching for the eight nonprofit organizations, 2016-17



Source: SBC cost data.

Two additional types of resources that contributed to costs are in-kind contributions from the community at large and resources provided by college partners. Although they represent a small share of costs (detailed in the next section below), nonprofit partners considered both of these types of resources to be important to their approach, and therefore the resource types are included in the direct services to students category. In-kind contributions are used by all but one program to provide transition coaching. Examples of in-kind contributions include community volunteers providing tutoring, reduced-fare passes provided by the transit authority, career night presentations or mock job interviews by community professionals, space and food for annual events, and AmeriCorps volunteers who provide low-cost staffing.

Five of the eight nonprofit organizations rely on campus space provided by colleges, and we include that space as a program cost. All nine nonprofits engage with colleges to coordinate coaching activities with general college services such as academic advising. In these five instances, however, colleges provide the nonprofit coaching program with dedicated space—cubicles or scheduled time in offices or classrooms—so the value of this resource is included in program costs. TBF also provided colleges with funding to support liaisons to the nonprofit organizations. These costs are included and are described below.

Costs to the Boston Foundation and Network-Wide Community Partners

TBF and network-wide community partners—colleges, uAspire, and the BPS—support the work of SBC coaches and the nonprofits in two forms: through **direct mechanisms**, via trainings and supports for coaches, and through **less direct means**, including thought leadership, program management and

coordination, and strategic direction. The value of TBF's and the network-wide partners' efforts was $1.16 million in 2015-16, which represented 22 percent of total program costs, or $1,285 of the $5,762 cost per coached student. The overall value of resources decreased by about 6 percent to $1.09 million in 2016-17, which represented 18 percent of SBC's costs, or $537 of the $3,038 cost per coached student. These resources are reflected in Exhibits D-1 and D-2 above, which shows the combined resources that support SBC coaching; they are not reflected in the organization costs shown in Exhibits D-3 and D-4.

Personnel. Multiple staff from TBF contribute to the management and administration of the SBC program and the SIF grant. The Success Boston director, the program director, the program officer, and program associate together are responsible for managing SBC program implementation and grant administration. TBF reported that approximately 10 percent of the Success Boston program associate's time went to supporting student recruitment in each year. Program finances were managed by TBF's chief financial officer, its finance director, and its finance coordinator. TBF employs a communications team of four, who allocated a portion of their time to publicizing SBC and its events, and a fundraiser, who allocated approximately one third time. Taken together, these staff salaries represent between $420,000 and $450,000 in each of the first two years.

TBF also used SIF funds to provide $200,000 to fund college liaison positions at five partner colleges in each year, positions designed to provide the nonprofit organizations with a dedicated point of contact to coordinate coaching activities on campus. TBF also allocated $250,000 each year for program partner uAspire, which offers services specifically focused on supporting SBC, including training for SBC coaches, hosting FAFSA completion events for SBC students, and mounting a text message program for SBC students.

Overhead, Facilities, Materials. In addition to personnel compensation, TBF had other costs in each year; these include overhead (about $130,000), facilities ($43,000), and materials ($15,000). Overhead costs include space, phones, and information technology. For example, TBF considerably reconfigured its program database, which it had earlier developed to track students and coach-student interactions, for the SBC program.

Launch vs. Ongoing Administration. Some of the costs borne by TBF in the initial years of the program represent those associated with launching (or expanding) an initiative. Consequently, some of the upfront costs may continue to decline (particularly on a per-student basis) in future years. Revamping the program database, learning about the SIF-specific grant reporting requirements, setting up procedures to monitor the nonprofit organizations' spending and grant compliance, and establishing procedures for preparing SIF financial reporting are all initial costs that might be expected to decrease over time. However, some level of administrative burden is expected to accompany public funding sources, even with familiarity gained from experience with reporting.[39]

Cost Analysis Methodology

It is important to note that our analysis approach was purposefully designed to produce a comprehensive cost measure, one that includes *all* the resources used to implement transition coaching. By deliberately casting a

[39] Though this administrative burden contributes to costs, increased accountability and reporting efforts may also improve program effectiveness.

broad net, we can more accurately characterize the full scope of the program—including costs not captured within the framework of SIF grant reporting. The cost analysis translates the full set of resources used to implement coaching into a quantitative metric to complement the SBC activities and support summaries in earlier chapters.

Categories and Subcategories. To implement the approach, the study team first clearly defined all activities that constitute the coaching program. We focused specifically on those resources that support coaching, as distinguished from resources used to support other initiatives. For example,

- Some SBC coaches are employed in organizations that operate multiple programs; by defining the activities specific to SBC coaching, we can separate out the activities and salaries spent on SBC coaching versus those allocated to other activities and programs. We do so by prorating such coaches' salaries to reflect the share of their time spent on SBC coaching more accurately.

For the purpose of identifying which resources to include in the cost analysis, we considered those activities and supports that only SBC students could access because they were participants in the SBC coaching program, and not the kinds of resources available to *all* students at a given campus. For example,

- SBC students could be referred by a coach to their college's advising services. However, that advising is generally available to all its students, not just students enrolled in the SBC coaching program. We distinguish between the SBC coach's time to make such referrals, which is included as a program resource (and is typically recorded as a coaching interaction in the program database) versus the time spent by college advisors to provide the referred services, which is not included as a SBC resource, even though SBC coaching may have led to increased use of the college-provided services.
- Many of the nonprofit organizations offer high school programming in addition to SBC coaching. Although high school programs can and do serve as a pipeline for connecting students to SBC, the staff time/resources for the high school programming is not considered a necessary resource for implementing SBC, because students are not yet enrolled in SBC when the high school programming occurs.
- Recruitment efforts for SBC specifically, however—that is, the time coaches spend actively recruiting students to participate in SBC (e.g., via presentations at high schools, coordination with colleges) are captured as a program cost.

The cost analysis categorizes the specific types of costs associated with program implementation into two primary categories: **direct services to students** and **administrative**. These two broad categories are then further broken into subcategories, including personnel, organizational overhead, facilities, program activities, and direct assistance, to show how resources are allocated. Resources are classified into categories and subcategories both to assist in collecting data and to facilitate cost analysis. Breaking out costs by type of resource helps us understand how the coaching program is structured, and it helps us appreciate the specific "inputs" that together comprise the coaching program.

Cost categories and examples of resources within them are presented in Exhibit D-5.

Exhibit D-5. Cost categories and resource examples

Category	Subcategory	Examples of Resources
Administrative	Personnel	• Executive director • Program director • Finance/accounting staff • Coach supervisor
	Overhead	• General overhead • Technology • Insurance • Office materials
	Facilities (space used for administrative functions)	• Rent and rental value of owned space used for administrative activities • Utilities
Direct services to students	Personnel	• Coaches • Recruitment staff • Volunteers • Dedicated college staff
	Facilities (space used for program activities)	• Rent and rental value of owned nonprofit space used for providing services • Dedicated space on campus provided by colleges • Utilities
	Miscellaneous	• Special events • Workshops • Materials
	Direct assistance	• Scholarships specifically designated for coached students • Transportation subsidy • Cash or near-cash (e.g., school supplies)

Data Collection. Data collection and analysis occurred at two time periods. The initial data collection covering the 2015-16 academic year was conducted in the late spring and early summer of 2016. The second data collection updated the initial effort and added data for the 2016-17 academic year. We followed the same data collection approach for the follow-up data collection. To allow costs to be compared across years, we adjusted all dollar values for inflation to be in 2018-equivalent dollars. We used the Consumer Price Index for the Boston-Cambridge-Newton, MA-NH metropolitan area to make the adjustment.[40]

To collect cost data, the study team took advantage of existing records with cost information. Data collection began with a review of administrative records generated as part of SIF grant reporting. We used quarterly invoices to TBF for the fiscal year ending March 31, 2016, for each nonprofit organization for data collection for the 2015-16 academic year. Information contained in the invoices allowed the study team to prepopulate a cost summary worksheet template to generate a preliminary list of each organization's costs. We also reviewed notes from organization leader interviews for each nonprofit

[40] Retrieved from https://data.bls.gov/timeseries/CUURS11ASA0 in November 2018.

organization. The team noted relevant activities, events, and program elements on the cost summary template. We used information in the quarterly invoices and from first-round data collection to prepopulate two-year cost summary worksheets.

After these pre-population exercises, we conducted telephone interviews with nonprofit organization staff familiar with coaching activities and with SBC financial statements as well as interviews with TBF staff themselves.[41] We conducted interviews with each of the eight nonprofit organizations in each of the initial and follow-up data collections. During those interviews, we reviewed the organization's prepopulated cost worksheet with the interviewee. All eight organizations actively engaged in helping the study team identify any missing resources, update changes in missing resources between the first and second interviews, classify resources as devoted to either administrative or coaching activities, and determine a monetary value of each resource.

Sometimes resources were identified as being used both for administrative functions and for coaching activities. Such resources were typically personnel responsible for managing the program as well as interacting with students, including program directors and senior coaching personnel. Facilities were also used both by administrative staff and for coaching activities. In these cases, the study team and organization staff discussed how these resources were generally used, and we collaboratively determined how to allocate costs most appropriately across the categories.

As necessary, items were identified for follow-up so organization staff could confirm figures and/or check with colleagues. Finally, the list of activities, events, and tasks identified from notes of early interviews was reviewed to verify that all activities and associated resources had been identified.

Data Analysis. After conducting interviews with all organizations, data were organized and processed and then reviewed for consistency across organizations. At this stage, we focused on calculating estimated dollar cost values for resources that did not involve direct expenditures such as occupancy costs, donated goods, and volunteer time.

[41] The study team emphasized that the data were being collected as part of the evaluation, not as part of SIF grant reporting, and that the data would not be subject to either an audit or a compliance review. This was done to encourage open dialogue that would, we hoped, result in a more complete and accurate picture of actual costs as implemented, regardless of funding source.

Imputing costs of volunteer time

To value volunteer time, we asked the nonprofit organizations what it would cost them to hire staff to provide the services delivered by their volunteers. So a professional giving an overview of career opportunities in a particular field has a higher associated cost than a college student peer tutor.

Responses were similar across organizations for similar types of volunteers, and the study team used consistent, conservative rates across all organizations.

Imputing occupancy costs

Imputing a value for occupancy costs of owned space is critical for determining costs for organizations operating programs in space they own. By coming up with a value for owned space, we can compare these programs' costs with the costs of leasing space.

To estimate these costs, the study team collected information from the organizations about the size of the space used for coaching, and then identified current market rental rates for similar properties nearby. To be more conservative, we used a low-end value of these rates to impute the unobserved occupancy cost. The imputed costs for the two organizations that own their space were comparable with the occupancy costs for those nonprofit organizations that lease facilities.

Seven organizations relied on volunteer time in at least one of the three years examined, and two organizations own the facility out of which their coaching program operates and therefore do not have lease expenses. These resources make up a small fraction of overall costs; however, they do represent a tangible input valued by organizations as important resources for providing transition coaching activities. Our approach to imputing values for volunteer time and occupancy costs is described in the textbox above.

To calculate total cost per student, the number of SBC coached students at each organization was merged to the cost data. After a final quality review of data collected for each organization, the study team summed costs across all initiative partners—the nonprofit organizations, TBF, and network-wide partners. These costs were then divided by the total number of students coached and analyzed.

Summary

The study team estimates the cost to implement SBC in 2015-16 was $5.2 million, which increased to $6.2 million by 2016-17. With the number of coached students increasing from 905 to 2,030, the costs per coached student decreased from $5,762 in 2015-16 to $3,038 in 2016-17. Because the SBC coaching model hinges on one-on-one support provided by coaches to students, expectedly the cost of hiring the personnel who provide direct services to students represents the largest input, at approximately 40 percent of total costs (in all years). In monetary terms, this translates into an investment in staff directly supporting SBC students' academic success of $1,227 per student in 2016-17.

This cost analysis is unique in that it monetizes a complex, collaborative, comprehensive program. Given the multifaceted role of coaches and the combined efforts of the higher education community and TBF, there are few comparable programs across the country, none of which has assigned comprehensive costs to date. Oreopoulos and Petronijevic (2016) include a cost estimate of a peer coaching program, which totals approximately $700 per student. However, unlike SBC, which employs full-time coaches, their estimate is based on a coaching program employing college students as coaches who work just seven

hours per week. Moreover, their cost estimate includes only the cost per student based on the cost of labor (hourly wage of the peer coaches). It does not account for additional costs associated with implementing such a coaching program (e.g., facility costs, staff trainings and benefits).

It is important to understand the overall program costs, particularly administrative costs. The changes in observed costs from the first year to the subsequent two years measured highlight that the program had a higher-cost (per student) start-up phase. Fiscal year 2015-16 was the first in which both TBF and participating nonprofit organizations operated the Success Boston Coaching initiative using SIF grant funds. The SIF funding came with reporting requirements, the burden of which added to administrative costs, and some organizations noted increased demands on staff time to become familiar with the grant management process.

Another avenue for per-student cost reductions in future years may well be further increases in the number of students served. Leaders of several of the nonprofit organizations noted that many administrative functions would not be affected by an increase in program scale; in some cases, coach caseloads were below potential capacity. Serving more students with the same resources would mean lower per-student costs.

Per-student costs and the relative importance of different cost categories varied substantially across organizations. This is not surprising, as program implementation (e.g., frequency of coach-student interactions, average caseloads) also varied substantially across the nonprofit organizations and coaches, as did the numbers of students served and the size of the nonprofit organizations. Furthermore, the costs to provide SBC coaching may decrease as the initiative matures: organizations will become more familiar with the grant administration process, and the nonprofit organizations new to coaching will develop and codify their program practices.

References

Arnold, K. D., E. C. Lu, and K. J. Armstrong. 2012. "The Ecology of College Readiness." *ASHE Higher Education Report* 38 (Special Issue): 1–138.

Arnold, K., S. Fleming, M. DeAnda, B. Castleman, and K. L. Wartman. 2009. "The Summer Flood: The Invisible Gap among Low-Income Students." *Thought & Action*, Fall, 23.

Aud, S., W. Hussar, F. Johnson, G. Kena, E. Roth, E. Manning, and J. Zhang. 2012. *The Condition of Education 2012*. NCES Report 2012-045. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Avery, C., and T. J. Kane. 2004. "Student Perceptions of College Opportunities. The Boston COACH Program." In *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It*, edited by C. Hoxby, 355–394. Chicago: University of Chicago Press.

Avery, C., J. Howell, and L. Page. 2014. *A Review of the Role of College Counseling, Coaching, and Mentoring on Students' Postsecondary Outcomes.* New York: College Board Research.

Ayala, C., and A. Striplen. 2002. "A Career Introduction Model for First-Generation College Freshmen Students" (ERIC ED469996). In *Thriving in Challenging and Uncertain Times*, edited by Garry R. Walz, Richard Knowdell, and Chris Kirkman, 57–62. Greensboro, NC: ERIC Clearinghouse on Counseling and Student Services.

Bahr, P. R. 2008. "Does Mathematics Remediation Work? A Comparative Analysis of Academic Attainment among Community College Students." *Research in Higher Education* 49 (5): 420–450.

Bailey, M. J., and S. M. Dynarski. 2011. "Gains and Gaps: Changing Inequality in U.S. College Entry and Completion." In *Whither Opportunity? Rising Inequality and the Uncertain Life Chances of Low-Income Children*, edited by G. J. Duncan and R. J. Murnane, 117–132. New York: Russell Sage.

Bailey, T., D. W. Jeong, and S. Cho. 2010. "Referral, Enrollment, and Completion in Developmental Education Sequences in Community Colleges." *Economics of Education Review* 29 (2): 255–270.

Bang, H., and J. M. Robins. 2005. "Doubly Robust Estimation in Missing Data and Causal Inference Models." *Biometrics* 61: 962–972.

Baum, S., J. Ma, and K. Payea. 2013. *Education Pays: The Benefits of Higher Education for Individuals and Society*. New York: College Board.

Belfield, C. R., and P. M. Crosta. 2012. "Predicting Success in College: The Importance of Placement Tests and High School Transcripts." CCRC Working Paper No. 42. New York: Community College Research Center, Columbia University.

Benjamini, Y., and Y. Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society, Series B (Methodological)*, 289–300.

Bettinger, E. P., A. Boatman, and B. T. Long. 2013. "Student Supports: Developmental Education and Other Academic Programs." *The Future of Children* 23 (1): 93–115.

Bettinger, E. P., and R. B. Baker. 2014. "The Effects of Student Coaching: An Evaluation of a Randomized Experiment in Student Advising." *Educational Evaluation and Policy Analysis* 36 (1): 3–19.

Bettinger, E., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu. 2012. "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment." *Quarterly Journal of Economics* 127: 1205–1242.

Bifulco, R. 2012. "Can Nonexperimental Estimates Replicate Estimates Based on Random Assignment in Evaluations of School Choice? A Within-Study Comparison." *Journal of Policy Analysis and Management* 31 (3): 729–775.

Bird, K., and B. L. Castleman. 2016. "Here Today, Gone Tomorrow? Investigating Rates and Patterns of Financial Aid Renewal among College Freshmen." *Research in Higher Education* 57 (4): 395–422.

BLS (Bureau of Labor Statistics). 2017. "Employment Projections." Washington, DC: Bureau of Labor Statistics. Retrieved from https://data.bls.gov/projections/occupationProj.

BLS (Bureau of Labor Statistics). 2018. "Employment Projections." Washington, DC: Bureau of Labor Statistics. Retrieved from https://data.bls.gov/projections/occupationProj.

BLS (Bureau of Labor Statistics). 2018. Occupational projections and worker characteristics. Retrieved from: https://www.bls.gov/emp/tables/occupational-projections-and-characteristics.htm

Bozick, R., and S. DeLuca. 2011. "Not Making the Transition to College: School, Work, and Opportunities in the Lives of American Youth." *Social Science Research* 40: 1249–1262.

Cahalan, M., L. Perna, W., Yamashita, M., Wright, J. and S. Santillan. 2018. 2018 Indicators of Higher Education Equity in the United States: Historical Trend Report. Washington, DC: The Pell Institute for the Study of Opportunity in Higher Education, Council for Opportunity in Education (COE), and Alliance for Higher Education and Democracy of the University of Pennsylvania (PennAHEAD).Carlstrom, A., and M. A. Miller, eds. 2013. *NACADA National Survey of Academic Advising*. Monograph No. 25. Manhattan, KS: National Academic Advising Association.

Caliendo, M., and S. Kopeinig . 2008. "Some Practical Guidance for the Implementation of Propensity Score Matching." *Journal of Economic Surveys* 22 (10): 31–72.

Carnevale, A. P., S. J. Rose, and B. Cheah. 2011. *The College Payoff: Education, Occupations and Lifetime Earnings.* Washington, DC: McCourt School of Public Policy, Center on Education and the Workforce. Retrieved from https://cew.georgetown.edu/report/the-college-payoff/.

Carnevale, A. P., N. Smith, and J. Strohl. 2013. *Recovery: Job Growth and Education Requirements through 2020.* Washington, DC: Center on Education and the Workforce, Georgetown University.

Carrell, S. E., and B. Sacerdote. 2013. "Late Interventions Matter Too: The Case of College Coaching New Hampshire." NBER Working Paper No. 19031. Cambridge, MA: National Bureau of Economic Research.

Castleman, B. L., K. Arnold, and K. L. Wartman. 2012 "Stemming the Tide of Summer Melt: An Experimental Study of the Effects of Post-High School Summer Intervention on Low-Income Students' College Enrollment." *Journal of Research on Educational Effectiveness* 5 (1): 1–17.

Castleman, B. L., and L. C. Page. 2013. "A Trickle or a Torrent? Understanding the Extent of Summer 'Melt' among College-Intending High School Graduates." *Social Science Quarterly* 95 (1): 202–220.

Castleman, B. L., and L. C. Page. 2015. "Summer Nudging: Can Personalized Text Messages and Peer Mentor Outreach Increase College Going among Low-Income High School Graduates?" *Journal of Economic Behavior & Organization* 115: 144–160. Retrieved from http://www.sciencedirect.com/science/article/pii/S0167268114003217.

Castleman, B. L., L. C. Page, and K. Schooley. 2014. "The Forgotten Summer: Does the Offer of College Counseling after High School Mitigate Summer Melt among College-Intending, Low-Income High School Graduates?" *Journal of Policy Analysis and Management* 33 (2): 320–344.

Clair, T., T. Cook, and K. Hallberg. 2014. "Examining the Internal Validity and Statistical Precision of the Comparative Interrupted Time Series Design by Comparison with a Randomized Experiment." *American Journal of Evaluation* 35 (3): 311–327.

Commonwealth Corporation (2016). Closing the Skills Gap: Meeting the Demand for Skills in a Growing Economy. Retrieved from:http://commcorp.org/wp-content/uploads/2017/01/Resources_LMI-Data-Report_Final_Jan-2017.pdf

Deming, D., and S. M. Dynarski. 2009. "Into College, Out of Poverty? Policies to Increase the Postsecondary Attainment of the Poor." NBER Working Paper No. 15387. Cambridge, MA: National Bureau of Economic Research.

Dynarski, S. M., S. W. Hemelt, and J. M. Hyman. 2013. "The Missing Manual: Using National Student Clearinghouse Data to Track Postsecondary Outcomes." NBER Working Paper No. 19552. Cambridge, MA: National Bureau of Economic Research.

Gallagher, R. P. 2010. *National Survey of Counseling Center Directors: 2010.* Monograph 85. Alexandria, VA: International Association of Counseling Centers.

Greene, J. P., and M. A. Winters. 2005. "Public High School Graduation and College-Readiness Rates: 1991-2002." Education Working Paper No. 8. Atlanta, GA: Center for Civic Innovation.

Gupta, H. 2017. The Power of Fully Supporting Community College Students: The Effects of the City University of New York's Accelerated Study in Associate Programs After Six Year. New York: MDRC. Retrieved from https://www.mdrc.org/publication/power-fully-supporting-community-college-students#graduation.

Harper, S. R. 2006. *Black Male Students at Public Flagship University in the US: Status, Trends, and Implications for Policy and Practices.* Washington, DC: Joint Center for Political and Economic Studies.

Harper, S. R., and K. A. Griffen. 2011. "Opportunity Beyond Affirmative Action: How Low-Income and Working Class Black Male Achievers Access Highly Selective High-Cost Colleges and Universities." *Harvard Journal of African American Public Policy* 17 (1): 43–60.

Haskins, R. 2008. "Education and Economic Mobility." In *Getting Ahead or Losing Ground: Economic Mobility in America*, edited by J. B. Isaacs, I. V. Sawhill, and R. Haskins, 91-104. Washington, DC:

The Brookings Institution. Retrieved from http://brookings.edu/research/reports/2008/02/economic-mobility-sawhill.

Hout, M. 2012. "Social and Economic Returns to College Education in the United States." *Annual Review of Sociology* 38: 379–400.

Hoxby, C., and C. Turner. 2013. "Expanding College Opportunities for High-Achieving, Low Income Students." SIEPR Discussion Paper No. 12-014. Stanford, CA: Stanford Institute for Economic Policy Research.

Isaacs, J. B., I. V. Sawhill, and R. Haskins. 2008. *Getting Ahead or Losing Ground: Economic Mobility in America*. Washington, DC: The Brookings Institution. Retrieved from http://brookings.edu/research/reports/2008/02/economic-mobility-sawhill.

Johnson, J., and J. Rochkind. 2009. *With Their Whole Lives Ahead of Them: Myths and Realities about Why So Many Students Fail to Finish College. A Report for the Bill and Melinda Gates Foundation.* New York, NY: Public Agenda. Retrieved from http://www.publicagenda.org/pages/with-their-whole-lives-ahead-of-them.

Karp, M. M. 2011. "Toward a New Understanding of Non-Academic Student Support: Four Mechanisms Encouraging Positive Student Outcomes in the Community College." CCRC Working Paper No. 28. Assessment of Evidence Series. New York: Community College Research Center, Columbia University.

King, Gary, and Richard Nielsen. 2016 "Why Propensity Scores Should Not Be Used for Matching." Working paper. Retrieved December 7, 2016, from http://j.mp/1sexgVw.

Klor de Alva, J., and M. Schneider. 2013. *What's the Value of an Associate's Degree? The Return on Investment for Graduates and Taxpayers.* Retrieved from https://www.air.org/sites/default/files/Value_of_an_Associate_Degree_10.13.pdf.

Linkow, T., K. Dadisman, B. Gamse, H. Didriksen, G. Schwartz, M. Hillard, and M. Karuu. 2015. *Degrees of Coaching. Success Boston's Transition Coaching Model*. Boston, MA: The Boston Foundation.

Linkow, T., H. Didriksen, B. Gamse, E. Bumgarner, M. Velez, S. Dastrup, S. Ballinger, S. Burke, M. Meneses, and R. J. de la Cruz. 2017a. *Success Boston Coaching for Completion 2015-16 Implementation Report*. Boston, MA: The Boston Foundation. Retrieved from https://www.nationalservice.gov/sites/default/files/evidenceexchange/FR_Success%20Boston%20Implementation%20Report_2017.pdf.

Linkow, T., B. Gamse, F. Unlu, E. Bumgarner, H. Didriksen, J. Furey, M. Meneses, M. Sami, and A. Nichols. 2017b. *The Power of Coaching: Interim Report on the Impact of Success Boston's Transition Coaching on College Success*. Boston, MA: The Boston Foundation.

McFarland, J., B. Hussar, X. Wang, J. Zhang, K. Wang, A. Rathbun, A. Barmer, E. Forrest Cataldi, and F. Bullock Mann. 2018. *The Condition of Education 2018.* NCES 2018-144. Washington, DC: National Center for Education Statistics, U.S. Department of Education. Retrieved from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018144

McLaughlin, N. Sullivan, A. Van Eaton, T. Linkow, H. Didriksen, B. Gamse, and J. Furey. 2016. *Reaching for the Cap and Gown: Progress toward Success Boston's College Completion Goals for Graduates of the Boston Public Schools*. Boston, MA: The Boston Foundation.

McLaughlin, J., and A. Van Eaton. 2018. *Staying the Course. Six-Year College Enrollment and Completion Experiences of BPS Class of 2011 Graduates*. Boston, MA: The Boston Foundation. http://www.successboston.org/-/media/tbf/reports-and-covers/2018/success-boston-april3.pdf.

Meara, E. R., S. Richards, and D. M. Cutler. 2008. "The Gap Gets Bigger: Changes in Mortality and Life Expectancy, by Education, 1981–2000." *Health Affairs* 27 (2): 350–360.

Morgan, Stephen L., and Christopher Winship. 2014. *Counterfactuals and Causal Inference*. New York: Cambridge University Press.

NSC (National Student Clearinghouse). 2018. *Snapshot Report: First-Year Persistence and Retention.* Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved September 17, 2018, from https://nscresearchcenter.org/snapshotreport33-first-year-persistence-and-retention/.

NCES (National Center for Education Statistics). 2012. Education Longitudinal Study of 2002 (ELS:2002), Third Follow-Up, 2012 [Restricted-use data file].

NCES (National Center for Education Statistics). 2014. *Profile of Undergraduate Students: 2011-2012.* Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved December 7, 2016, from http://nces.ed.gov/pubs2015/2015167.pdf.

Oreopoulos, P., and U. Petronijevic. 2016. "Student Coaching: How Far Can Technology Go?" NBER Working Paper No. 22630. Cambridge, MA: National Bureau of Economic Research.

Pfeffer, F. T., and F. R. Hertel. 2015. "How Has Educational Expansion Shaped Social Mobility Trends in the United States?" *Social Forces* 94 (1): 143–180.

Roderick, M., J. Nagaoka, V. Coca, E. Moeller, K. Roddie, J. Gilliam, and D. Patton. 2008. *From High School to the Future: Potholes on the Road to College*. Chicago, IL: Consortium on Chicago School Research.

Rolston, H., Copson, E., and K. Gardiner. 2017. *Valley Initiative for Development and Advancement: Implementation and Early Impact Report Pathways for Advancing Careers and Education (PACE).* Report prepared by Abt Associates for the U.S. Department of Health and Human Services.

Rosenbaum, P., and D. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika*, 70(1), 41-55.Rosenbaum, P., and D. Rubin. 1984. "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *Journal of the American Statistical Association* 79 (387): 516–524.

Rosenbaum, P., and D. Rubin. 1985. "Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score." *The American Statistician.* 39(1), 33-38.
Ross, R., S. White, J. Wright, and L. Knapp. 2013. *Using Behavioral Economics for Postsecondary Success*. Washington, DC: Ideas 42.

Rubin, D. B. 2001. "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation." *Health Services & Outcomes Research Methodology* 2: 169–188.

Schochet, P. Z. 2008. *Technical Methods Report: Guidelines for Multiple Testing in Impact Evaluations.* Washington, DC: Mathematica Policy Research, Inc. http://ies.ed.gov/ncee/pdf/20084018.pdf.

Scrivener, S., and M. J. Weiss. 2009. *Opening Doors: More Guidance, Better Results? Three-Year Effects of an Enhanced Student Services Program at Two Community Colleges.* New York: MDRC.

Scrivener, S., M. J. Weiss, A. Ratledge, T. Rudd, C. Sommo, and H. Fresques. 2015. *Doubling Graduation Rates: Three-Year Effects of CUNY's Accelerated Study in Associate Programs (ASAP) for Developmental Education Students.* New York: MDRC.

Shapiro, D., A. Dundar, P. K. Wakhungu, X. Yuan, A. Nathan, and Y. Hwang. 2015. *Completing College: A National View of Student Attainment Rates – Fall 2009 Cohort.* Signature Report No. 10. Herndon, VA: National Student Clearinghouse Research Center.

Smith, J. A., and P. E. Todd. 2005. "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics* 125: 305–353.

Steiner, P. M., T. D. Cook, and W. R. Shadish. 2011. "On the Importance of Reliable Covariate Measurement in Selection Bias Adjustments Using Propensity Scores." *Journal of Educational and Behavioral Statistics* 36 (2): 213–236.

Steiner, P. M., T. D. Cook, W. R. Shadish, and M. H. Clark. 2010. "The Importance of Covariate Selection in Controlling for Selection Bias in Observational Studies." *Psychological Methods* 15 (3): 250–267.

Stephan, J. L., and J. E. Rosenbaum. 2013. "Can High Schools Reduce College Enrollment Gaps with a New Counseling Model?" *Educational Evaluation and Policy Analysis* 35 (2): 200–219.

Stephens, N. M., T. N. Brannon, H. R. Markus, and J. E. Nelson. 2015. "Feeling at Home in College: Fortifying School-Relevant Selves to Reduce Social Class Disparities in Higher Education." *Social Issues and Policy Review* 9 (1): 1–24.

Stuart, Elizabeth A. 2010. "Matching Methods for Causal Inference: A Review and a Look Forward." *Statistical Science* 25 (1): 1–21.

Sum, A., N. Fogg, I. Khatiwada, J. McLaughlin, S. Palma, N. Sullivan, and C. Smith. 2008. *Getting to the Finish Line: College Enrollment and Graduation, a Seven-Year Postsecondary Longitudinal Study of the Boston Public Schools Class of 2000 Graduates.* Boston: Northeastern University and Boston Private Industry Council.

Sum, A., I. Khatiwada, W. McHugh, and S. Palma. 2013. *Getting Closer to the Finish Line: The College Enrollment and Completion Experiences of Graduates of the Boston Public Schools.* Boston: The Boston Foundation. Retrieved from http://www.tbf.org/~/media/TBFOrg/Files/Reports/GettingCloser_011513.pdf.

Sum, A., I. Khatiwada, and S. Palma. 2014. *The College Persistence and Graduation Experiences of Success Boston Coaching Participants from the BPS Graduating Class of 2009.* Boston: Center for

Labor Market Studies, Northeastern University. Retrieved from https://repository.library.northeastern.edu/downloads/neu:m0407544b?datastream_id=content.

Swecker, H. K., M. Fifolt, and L. Searby. 2013. "Academic Advising and First-Generation College Students: A Quantitative Study on Student Retention." *NACADA Journal* 33 (1): 46–53.

Tan, Z. 2006. "A Distributional Approach for Causal Inference Using Propensity Scores." *Journal of American Statistical Association* 101: 1619–1637.

Tym, C., R. McMillion, S. Barone, and J. Webster. 2004. *First-Generation College Students: A Literature Review.* Round Rock, TX: Research and Analytic Services, Trellis Company.

U.S. Census Bureau. 2017. Current Population Survey (CPS), "Annual Social and Economic Supplement, March 2017."

U.S. Department of Education, National Center for Education Statistics. 2016. *The Condition of Education 2016, Immediate College Enrollment Rate.* NCES Report 2016-144. Retrieved from http://nces.ed.gov/programs/coe/indicator_cpa.asp

Visher M., K. F. Butcher, and O. S. Cerna. 2010. *Guiding Math Students to Campus Services: An Impact Evaluation of the Beacon Program at South Texas College.* New York: MDRC.

Vuolo, M., J. T. Mortimer, and J. Staff. 2016. "The Value of Educational Degrees in Turbulent Economic Times: Evidence from the Youth Development Study." *Social Science Research* 57: 233-252.

What Works Clearinghouse. 2014. *What Works Clearinghouse Standards Handbook, Version 4.0.* Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf