

Race, Friends, and College Readiness: Evidence from the High School Longitudinal Study

Steven Elias Alvarado · Brian P. An

Published online: 13 February 2015
© Springer Science+Business Media New York 2015

Abstract Close friends are likely to transmit influence on students' educational attitudes and decisions that are independent of students' own background abilities and motivations. However, previous research suggests that close friends may have uneven effects on educational outcomes by race and ethnicity. We analyze the impact of close friends who are college bound on students' college readiness using new and restricted panel data from the High School Longitudinal Study (2009–2011). Descriptive analyses suggest that having a college-bound friend is positively associated with college readiness and that these impacts are felt by racial and ethnic subgroups in separate and unique ways. Results from propensity score models suggest that while having a college-bound friend generally yields positive effects on all students, it has a more consistent effect on white students' college readiness compared with Asians, blacks, and Latinos. A formal sensitivity analysis suggests that these treatment effects are robust to the confounding influence of an unobserved confounder.

Keywords Friends · Peer effects · College readiness · Propensity score matching

Introduction

Adolescents' educational and career paths may be altered by the company they keep vis-a-vis their friends' influence on their coursework patterns, educational aspirations, achievement, and self-perceptions. Moreover, friends are especially salient during the formative teenage years when one's location in the social hierarchy often supersedes other spheres of influence such as parents and teachers (Coleman 1961). While the literature on "peer effects" and educational outcomes has been very active over the past few decades, we still know little about how close friends (i.e., those with whom students share trust and reciprocity) impact students' college readiness during the early formative years of high school when students are making critical decisions regarding educational trajectories (Sacerdote 2011). To fill this gap in the literature, this paper examines whether friends influence (1) students' expectations for post-secondary attainment and (2) their course-taking patterns during the early years of high school.

Studying *close friends'* influence on students' behavior may uncover important pathways toward academic preparation for college. Hallinan and Williams' (1990) seminal work on friendship introduced Parson's (1963) theoretical model of influence to this literature and has since helped move the literature to consider the impact of more proximal social agents on students' behavior. It makes sense then that the literature on peer effects suggests members of one's social network who are more proximal (i.e., those whom are more trusted and are members of an 'inner ring') exert the strongest influence on students' educational behaviors than so-called peers who may be less trustworthy (Burke and Sass 2013; Halliday and Kwak 2012; Mora and Oreopoulos 2011). This theoretical perspective implies that

S. E. Alvarado (✉)
Cornell University, 336 Uris Hall, Ithaca, New York 14853,
USA
e-mail: alvarado@cornell.edu

B. P. An
University of Iowa, 493 Lindquist Center North, Iowa City,
IA 52242, USA

solidarity is a necessary medium upon which influence flows and conditions students' influence on one another. This perspective also informs all aspects of our empirical investigation of the influence of *close friends* in this paper.

Meanwhile, students form their post-secondary trajectories by first passing through a predisposition stage in which students discuss future plans with significant others (e.g., parents, teachers, guidance counselors, and friends) and hone their academic interests through course taking during the early years of high school (Hossler and Gallagher 1987). If close friends influence students' attitudes, opinions, beliefs, and behaviors, then we may assume close friends also influence students' plans regarding preparation for college during high school. However, to date, little research has examined the process of influence on college readiness during the early years of high school. As such, this paper is among the first to coalesce these two previously separated, but influential, spheres of adolescence in a single study of how close friends impact students' early preparation for college.

To do so, the current paper capitalizes on new, unused, and restricted data from the High School Longitudinal Study of 2009 (HSL:09) to examine the effect of close college-bound friends (CBF) on students' college readiness. These nationally representative panel data lend themselves to a causally sequenced analysis of how college-bound friends influence white, Asian, black, and Latino students' academic preparation for college. In addition, we apply a counterfactual causal framework to estimate effects for having close college-bound friends on students' educational expectations and on their course-taking patterns. The research questions we investigate are as follows: (1) Do CBF increase students' expectations about college attainment; (2) do CBF increase students' likelihood of taking advanced coursework; and (3) is the influence of CBF moderated by race and ethnicity? The HSL:09 is uniquely qualified to answer these questions because it provides key data on students' self-reported close friends' commitment to college measured in the fall of the ninth grade and students' self-reported educational expectations and course-taking patterns measured in the spring of the 11th grade.

Literature Review

The Impact of Friends on Schooling

According to Parsons (1963), influence is directly proportional to how much one needs information that one party can offer to another party, and it is something that impacts an individual's attitudes, opinions, and behaviors by affecting his or her beliefs (Hallinan and Williams 1990).

The twin cornerstones of Parsons' conceptualization of influence are the following: (1) the person receiving information will trust the provider when the chances of being deceived are low, as in relationships built upon solidarity and friendship; and (2) influence is proportional to the degree of trust between individuals so that students' sensitivity or vulnerability to influence increases as the trustworthiness of friends increases.

In a school setting, students' need for information regarding what courses to take, the difficulty of those courses, and the teachers who teach those courses are essential information all students need and should seek early in the school year in order to make decisions that affect their immediate goals. Access to this key information may also impact their preparation for college. For example, knowing which courses to do well in to impress potential letter writers or understanding which courses act as gatekeepers into honors or advanced tracks can have lasting effects on students' educational careers. Of course, students pull from many sources for this information (e.g., teachers, counselors, and parents) but often rely on friends due to their accessibility and, critically, their trustworthiness on these and other matters that often also involve interpersonal fidelity that solidify strong bonds. However, in contrast to Parsons' poignant theoretical perspective and with some notable exceptions, researchers have usually relied upon peers to measure social spheres of influence, a group whose trustworthiness is often nebulous at best and whose influence may be low (Sacerdote 2011).

Friendship Ties and the Problem of Self-Selection

The theoretical model we adapt from Parsons (1963) and which Hallinan and Williams (1990) elucidated in reference to the peer effects literature establishes that influence has an elastic nature that bends proportional to the weight of the level of trustworthiness among friends. That is, closer friends who share mutual trust and understanding will likely impart greater influence on one another compared with casual friends or peers.

Often, students are attracted to one another based on shared interests, achievement, reciprocity, status, and ascribed characteristics (Hallinan and Williams 1987). Students may find stronger, more cohesive, and longer-lasting bonds among those who share similar interests because they appear more attractive than alternative members of one's social circle who do not share as many interests, real or perceived (Levinger 1976). For example, students who expect to go to college, who are interested in getting good grades, or who share similar gender or race and ethnic heritage may be more likely to become friends and impart more influence on one another than students who do not share these similarities. Researchers have found these

shared ascribed and achieved characteristics of students are the strongest factors influencing friendship formation (Hallinan and Williams 1987, 1989).

Family-level processes may also influence friendship formation and college readiness. For example, scholars have found that socioeconomic advantage leads to higher rates of persistence in high school math courses (Crosnoe and Schneider 2010), which in turn may influence whether students have college-bound friends and whether they eventually become college ready.

Structural institutions, most notably schools, may also play an important role in friendship formation through organizing the school curriculum into spheres of learning (Gamoran 1989; Hallinan and Williams 1989; Kubitschek and Hallinan 1998). Schools often structure the formation of friendships through the practice of academic “tracking,” which facilitates the creation of similar interests by limiting mobility between separate tracks during high school. For example, on the one hand, students in advanced academic tracks are more likely to develop interests in advanced courses and college preparation because these students are conditioned to prepare for rigorous college curricula. On the other hand, students not in academic tracks may not become as interested in advanced courses or preparing for college as those in academic tracks because of the lack of such pressures to meet college prerequisites. While there is variation within these tracks, the variation between them is likely much more palpable. The main point is students in the same track are more likely to develop trust and solidarity and are therefore more likely to become friends than students in different tracks (Hallinan and Sorensen 1985). Organizational characteristics of schools are therefore another important factor governing the creation of trust and reciprocity among students.

These ascribed and structural constraints on friendship formation also imply that students self-select into friendship networks based on latent, and developed, shared interests. For example, students interested in doing well in school may seek out others whom they see as someone who can provide both information and strategies regarding courses and with whom they share a trusting relationship. Similarly, students nested within a given track will also likely approach other students in the same track for information. But, students may do so not because they are randomly selecting friends, but rather because something *compels* them to (e.g., shared interests and goals).

The challenge in this example presents itself when one attempts to tie the influence of friends to students’ outcomes. When students self-select into friendship circles, it becomes difficult to disentangle the “effect” of friendship circles from the “effect” of students’ own latent interests and abilities (Manski 1993; McPherson et al. 2001). Famously, Hauser (1970) outlined the difficulty of estimating

causal effects for social contexts on students’ outcomes. In this paper, we estimate effects for CBF on college readiness using propensity score models that balance a rich set of theoretically driven observed characteristics for student who in reality had CBF and those who did not. In this way, we account for friendship selection using both individual-level and school-level characteristics. Further, we assess the sensitivity of the influence of close friends using a formal test that tells us how strong an unobserved confounder would have to be to undermine our effects. In doing so, we fill a gap in the literature by improving the understanding of whether close friends impact students’ post-secondary expectations and college readiness during high school.

College Readiness

Friends are important sources of influence on all types of outcomes, but two especially salient ones are students’ early expectations for their educational attainment after high school and students’ coursework preparation for meeting these post-secondary expectations during the first few years of high school. The influence of friends may enter the college-going pipeline at this point by providing information about course requirements, course weighting (e.g., honors and Advanced Placement), and the resources schools provide to meet college eligibility (Hill 2008; Sadler and Tai 2007). All of these vital pieces of information are likely to come from trusted sources, such as close friends. Friends may also share sentiments about attending college, such as which types of schools to aim for and what classes to take in order to apply, without necessarily having to account for the heavy burden of affordability, prestige, and other considerations one makes when deciding whether or not to actually apply or to enroll. That is, friends may make an even greater impact on college readiness than on actually applying to or enrolling in college because the cost of exchanging and adopting information from friends is so much lower at this point. Moreover, all of these data on college preparation are likely shared among friends beginning early on in high school. However, little research in the peer effects literature has examined the impact of close friends on these important early educational processes thus far.

Compared with other post-secondary outcomes such as college enrollment, college preparation is an understudied topic that may reveal something different about the role of friends in educational attainment. In particular, preparing for college lays the foundation upon which students make decisions about whether or not to pursue post-secondary education further down the pipeline (Stearns et al. 2010). Furthermore, while scholars often examined peer effects on test scores in elementary and secondary school using proxy

peer measures such as school-level mean achievement characteristics of all students within the same school (Hanushek et al. 2003; Hoxby 2000; Vigdor and Nechyba 2007) or SAT scores and high school rank of randomly assigned roommates in college (Sacerdote 2001; Zimmerman 2003), they often tend to focus on classmates and other types of peers but not friends. However, only a couple of studies (e.g., Alvarado and Turley 2012; Riegle-Crumb et al. 2006) have examined how actual friends impact students' college readiness during high school.

To this end, AddHealth data have been especially useful because of the In-School Survey's ability to match friends attending the same school together via questions that identified up to five of focal students' best male and female friends. For example, Riegle-Crumb et al. (2006) used AddHealth transcript data and found that female students' academically oriented female friends increased the odds of taking advanced coursework in the 11th and 12th grades. These authors' illuminating study established that friends may in fact influence students' college readiness using longitudinal data that allowed them to analyze this association in a causal sequence. However, their study design was still unable to capture the influence of friends on college readiness that occurs during the bulk of the period when course selection and performance matters for college applications (i.e., between ninth and 11th grades) and was limited in its ability to account for the threat of selection bias stemming from unobserved confounding. We therefore aim to fill an important gap in the literature in the current paper by studying the influence of close friends on college readiness earlier in the educational pipeline and by accounting for the influence of unobserved selection bias on our estimates.

Racial and Ethnic Variation

The elements that go into college choice and their effects vary for members of different racial and ethnic groups (Freeman 1999; Hurtado et al. 1997; Jackson 1990; Perna 2000; St. John 1991). So far, only a few studies have addressed racial and ethnic variation in the influence of peers and friends on individuals' outcomes (Alvarado and Turley 2012; Kao 2004; Way and Chen 2000), none of which have examined college readiness in high school.

Studies that do examine variation in these effects by race and ethnicity find mixed results. For example, Crosnoe et al. (2003) used AddHealth data and found that academically oriented high school friends protect students from academic problems. Still, while they found that this effect did not vary between whites and blacks, they did not study Latinos or Asians. Similarly, Arcidiacono and Nicholson (2005), using characteristics of peers in a medical school class, reported that the influence of peers on choice of specialty did not vary by race in medical school and

Cheng and Starks (2002) reported that the effect of friends' aspirations on students' likelihood of dropping out of high school was similar for all racial groups. In contrast, some recent research suggests that race and ethnicity may moderate the influence of friends on high school students' college preparation. For example, Alvarado and Turley (2012) found the influence of college-oriented friends on college application was less powerful among Latino compared with white high school students. Using Texas Higher Education Opportunity Project (THEOP) data that asked students to state how many college-oriented friends they had, these authors argued that differences in the importance of family ties may be one possible explanation for why friends mattered less for Latinos than they did for whites. Overall, however, our understanding of racial and ethnic heterogeneity seems to be underdeveloped in the peer effects literature and studies that include national samples of whites, Asians, blacks, and Latinos are rare.

While the literature on heterogeneous racial and ethnic peer effects is a nascent field, we may expect members of minority groups to be less sensitive to the influence of friends for various reasons. For example, minority students may not be as influenced by high-achieving peers if they perceive the academically oriented behavior of these friends to be of limited utility due to discrimination in the wider society (Fordham and Ogbu 1986). Instead of improving minorities' likelihood of following a college track, structural discrimination may impede minority students' identification with success in the classroom and lead them to eschew friends who are academically oriented and college bound. A second possible reason why minorities may be less sensitive to the influence of friends is the strength of kinship ties. Among blacks, strong kinship ties have been a tenet of family researchers for decades (Stack 1974). Indeed, scholars have found black youth spent significantly more time around family members compared with friends (Larson et al. 2001). This emotional and physical closeness to kin may lead minority youth to become increasingly codependent on kin for support and may attenuate the influence of friends. For example, Giordano et al. (1993) found that, compared with white youth, black youth placed less importance on having close friends and expressed lower levels of intimacy with the friends they had and that black youth expressed having much more intimacy with kin compared with friends.

Similarly, familial obligations may also primarily govern Latino students' educational decisions (Desmond and Turley 2009; Suarez-Orozco and Suarez-Orozco 1995). Among high-achieving students, Latinos are the least likely group to enroll in four-year colleges and universities (Hurtado et al. 1997), a disparity that may be affected by familial obligations. Latinos are also the most likely group to attend two-year community colleges (Aud et al. 2010), suggesting they

may lower their educational expectations based on familial obligations and adjust their course-taking behaviors accordingly to less demanding academic tracks in high school. The weight of responsibilities toward the family among Latinos who are immersed in academically oriented friendships may suggest that these students are influenced less by their college-bound friends compared with whites.

The Current Study

We examine effects of close friends on college readiness. We follow previous studies and define college readiness as the level of preparation students require to complete entry-level post-secondary courses without the need of remediation (Conley 2012). Researchers and policy makers have defined readiness as either cognitive (e.g., achievement and course-work) or non-cognitive (e.g., goal commitment, socialization, and effort). Furthermore, they have considered accelerated programs, such as dual enrollment and advanced placement (AP), as viable programs that promote college readiness (Conklin and Sanford 2007; Struhl and Vargas 2012; Texas P-16 Council 2007). For example, dual enrollment programs allow students to take college courses while in high school (An 2013b). Not only do students participate in college coursework, some researchers advocate dual enrollment as a channel to socialize high school students into becoming college students (An 2015; Karp 2012). Each of these programs is a key predictor of college readiness (Adelman et al. 1999; An 2013a; b; Sadler and Tai 2007). For instance, An (2013b) found students who participated in dual enrollment are 6 percent points lower in their probability to take a remedial course in college than similar students who did not participate in dual enrollment. We also incorporate another element to college readiness, students' educational expectations, which are key predictors of post-secondary enrollment (Bates and Anderson, in press) and may condition students' academic course-taking patterns.

Based on previous studies, we expect close friends to impart influence on all students' college expectations and college readiness. However, we also expect race and ethnicity to moderate the CBF effect on these outcomes because of minority students' strong kin ties in adolescence. That is, we expect black and Latino students to benefit from CBF in terms of college readiness. However, based on previous studies, we also expect CBF to have weaker effects among black and Latino students compared to white and Asian students.

Data and Methods

The High School Longitudinal Study of 2009 (HSL:09) is an ongoing nationally representative survey of approximately

25,210 ninth grade students nested within 944 public and private high schools in the fall of 2009. The HSL:09 is especially useful in answering the questions we have raised related to students' academic preparation for college and their engagement with STEM because of its rich data on friendship contexts and its explicit focus on math and science engagement in high school. HSL:09 school and student samples are nationally representative and also state representative for a subset of 10 states. We merged the student-level data with data at the school and home levels to provide context and to capture important sociodemographic and structural characteristics of the students' learning environment. The parent was self-selected, and therefore not nationally representative, using the criterion that the responding parent should be the one most knowledgeable about the ninth-grader's current situation.

We use the restricted data file from the baseline wave that was completed in the fall of 2009 when students were in the first semester of ninth grade to measure the "treatment variable" (i.e., having CBF), and we use the restricted data file from the first follow-up that was completed in the spring of 2012 when students were in the second semester of 11th grade to measure all outcomes. Rather than drop cases with missing values on predictor variables, we imputed missing data that were assumed missing at random using Stata's 'ice' command for both predictors and outcomes and we excluded observations that were missing data on the outcome variables from all analyses (Royston 2005). Previous research recommends imputing for missing on both predictors and outcomes but then removing cases with missing data on the outcomes in the final analysis (Allison 2002; von Hippel 2007a). Recent researchers argue that imputed outcomes are needed to impute the predictors, but the outcomes in and of itself add no new information (von Hippel 2007b). Further, researchers recommend removing observations with missing outcome values after imputation and before running analyses to reduce noise in estimates and because including these observations adds little to regression estimates (von Hippel 2007a). The range of missing data for all variables either we directly use in the analysis or we use to create the variables in the analysis was 0–9.55 percent. The missing data range from 0 to 9.55 percent, indicating low levels of missingness overall.

College-Bound Friends

The dummy "treatment" variable is students having a CBF. We coded CBF to equal 1 if the student's closest friend plans to go to college and if the student indicated that he or she talked to his or her friends about going to college, 0 otherwise. The HSL simply asks a general question about whether or not the student's closest friend plans to go to college, without being more specific about

2-, 4-year, public, private, or any other subsequent level of detail about type of college. Furthermore, the HSLS does not directly ask whether students talked to their closest friend about college, just whether they talked to their friends in general about it. Therefore, we must assume that our CBF variable reasonably indicates the student talked to his or her closest friend, who was college bound and falls into their circle of friends, about college. In this manner, we can identify the influence that *close friends*, rather than peers, who are college bound have on students' college readiness. This also improves the precision of contextual effects because our CBF variable, for all intents and purposes, ensures that the environmental influences we estimate are stemming from individuals who are nested closely within the focal students' social network.

College Readiness

The four binary-dependent variables of college readiness are all measured in the spring of 11th grade. The first outcome corresponds to students' assuredness of their college completion, while the last three correspond to students' academic behaviors undergirding these expectations for post-secondary attainment. They include (1) whether the student expects to earn at least a BA degree taken from the question "As things stand now, how far in school do you think you will get?", (2) whether students have taken any dual enrollment course through the spring of 11th grade, (3) whether students have taken any AP course through the spring of 11th grade, and (4) whether students have taken any AP STEM course through the spring of 11th grade.¹ Students' course work includes courses they were enrolled in during the spring 2012 first follow-up interview.

Covariates

Table 1 summarizes the CBF variable, covariates, and outcomes in terms of means and standard deviations. We include a rich array of covariates that influence friendship networks, students' educational expectations, and their course-taking patterns. All variables are coded as binary indicators except where noted by a scale or the word "percent."

¹ AP STEM courses include any of the following: AP Math, AP Science, AP Biology, AP Chemistry, AP Physics, and AP Computer Science. The intercorrelation for the outcome variables is low (0.24) and the Cronbach's alpha is also low (0.56). Both of these statistics suggest that combining the outcomes leads to an unreliable measure of a single construct and separating the outcomes is likely capturing multiple dimensions of college readiness.

For instance, at the student level, we begin by including students' sex and immigrant status. We further account for whether students' favorite school subjects include science, math, or computer science to gauge their likelihood of becoming friends with other students with college ambitions. Extracurricular academic enrichment programs may also promote friendship among college-bound students so we account for whether students were involved in STEM enrichment programs.² Similarly, students' academic track often influences their friendship selection (Hallinan and Sorensen 1985; Kubitschek and Hallinan 1998). Therefore, we included indicators of academic achievement in the eighth grade such as whether students completed an advanced science or math course or if they had received an A or B grade in their most advanced science or math course in the eighth grade. We also account for the influence of motivation on friendship selection, which we measure through an indicator for whether students have an educational or career plan. Because science and math course-taking behavior in the ninth grade may also influence friendship patterns, we include measures of participation in these subjects in ninth grade. Self-perceptions, as measured through students' internalized and externalized identity, may also have an impact on how open students are to social interaction among college-bound students. Therefore, we account for internalized and externalized science and math identity. Finally, we account for math scores in the ninth grade.

Because family-level processes influence friendship formation and college readiness, we included resources such as family income, poverty status, parent's education, parent's socioeconomic status, and the number of persons in the household. To partially account for direct parental influences on children's educational careers, we included an indicator for parents' educational expectations for their children. Because parents' STEM interests may impact their children's interest in STEM and these interests may compound to influence students' academic social circles, we included an indicator on whether either parent received any degree in STEM.

School-level organizational and procedural features such as tracking, human capital resources, and curricula are important contextual resources that may impact students' selection of friends as well as their college readiness. To address remaining socioeconomic, structural, and human capital heterogeneity at the school-level, we controlled for the percent of students receiving free or reduced priced lunch, the percent of students who are enrolled in Advanced Placement courses, and the number of full-time

² STEM enrichment programs include: Mathematics, Engineering, Science Achievement (MESA), Upward Bound, Talent Search, GEAR Up, and Advancement Via Individual Determination (AVID).

Table 1 Descriptive statistics for HSLs sample

Variables	Whites (N = 11,850)		Asians (N = 1,670)		Blacks (N = 2,220)		Latinos (N = 3,320)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
<i>Treatment</i>								
College-bound friend	0.51	0.50	0.56	0.50	0.47	0.50	0.46	0.50
<i>Outcomes</i>								
Expects to earn at least a BA	0.71	0.45	0.85	0.36	0.69	0.46	0.60	0.49
Took any dual enrollment by spring 2012	0.23	0.42	0.25	0.43	0.22	0.42	0.21	0.40
Took any AP by spring 2012	0.35	0.48	0.60	0.49	0.26	0.44	0.29	0.46
Took an AP STEM course by spring 2012	0.18	0.38	0.39	0.49	0.13	0.34	0.13	0.34
<i>Students characteristics</i>								
Female	0.49	0.50	0.49	0.50	0.48	0.50	0.50	0.50
Immigrant	0.26	0.44	0.57	0.50	0.36	0.48	0.43	0.50
Favorite subject is science, math, or computer science	0.27	0.44	0.36	0.48	0.28	0.45	0.26	0.44
Participated in a STEM engagement program	0.09	0.28	0.18	0.39	0.22	0.41	0.13	0.34
Advanced science course in eighth grade	0.05	0.21	0.08	0.27	0.06	0.24	0.06	0.24
A or B in most advanced science course in eighth grade	0.80	0.40	0.87	0.34	0.66	0.47	0.66	0.47
Advanced math course in eighth grade	0.77	0.42	0.83	0.37	0.73	0.44	0.69	0.46
A or B in most advanced math course in eighth grade	0.77	0.42	0.87	0.34	0.64	0.48	0.66	0.47
Educational/career plan in fall of ninth grade	0.59	0.49	0.68	0.47	0.74	0.44	0.62	0.49
Student took science fall of ninth grade	0.85	0.36	0.86	0.34	0.77	0.42	0.79	0.41
Student took math fall of ninth grade	0.92	0.28	0.90	0.30	0.87	0.34	0.87	0.34
Sees self as a science person	2.63	0.90	2.46	0.86	2.79	0.91	2.81	0.88
Others see him/her as a science person	2.71	0.85	2.53	0.82	2.82	0.86	2.89	0.82
Sees self as a math person	2.50	0.95	2.18	0.84	2.48	0.95	2.57	0.94
Others see him/her as a math person	2.50	0.91	2.07	0.83	2.47	0.91	2.58	0.89
Math score in fall of ninth grade	51.97	9.69	58.07	10.63	46.41	9.26	48.21	9.38
<i>Family background</i>								
Family income (1 = 0–34k; 2 = 35–74k; 3 = 75–114k; 5 = 115–194k; 6 = 195+ k)	2.57	1.20	2.50	1.28	1.85	1.04	1.83	1.04
In poverty	0.15	0.36	0.22	0.42	0.39	0.49	0.42	0.49
Parent's education (1 = less than HS; 2 = HS; 3 = some college; 4 = BA +)	2.82	0.76	3.04	0.84	2.62	0.78	2.28	0.86
Parent's education expectation of child (1 = less than HS; 2 = HS; 3 = some college; 4 = BA +)	3.31	0.67	3.63	0.56	3.41	0.70	3.31	0.74
Parents socioeconomic status	0.19	0.74	0.37	0.89	-0.17	0.73	-0.39	0.73
Number of household members	4.30	1.29	4.49	1.36	4.33	1.51	4.61	1.44
At least one parents highest degree was STEM	0.14	0.34	0.33	0.47	0.11	0.31	0.08	0.28
<i>High school characteristics</i>								
Percent of students receiving reduced/free lunch	29.50	22.52	29.03	25.05	43.21	28.15	42.21	27.96
Percent of students enrolled in advanced placement courses	15.43	13.16	21.33	14.80	15.12	12.75	16.15	13.75
Number of full-time certified math teachers	8.79	5.97	12.06	7.07	9.94	6.43	10.95	7.40
Number of full-time certified science teachers	7.99	5.45	11.58	6.79	8.77	5.68	9.73	6.74
School offers advanced STEM courses	0.89	0.32	0.95	0.22	0.89	0.31	0.90	0.30
School requires specific math courses	0.90	0.30	0.90	0.30	0.93	0.25	0.91	0.29
School requires specific science courses	0.87	0.34	0.87	0.34	0.89	0.32	0.86	0.35
School STEM climate	2.86	1.41	3.56	1.34	3.07	1.39	3.13	1.41
School uses a tracking policy	0.45	0.50	0.46	0.50	0.44	0.50	0.49	0.50
Program that encourages underrepresented students in STEM	0.28	0.45	0.34	0.47	0.37	0.48	0.34	0.47

Source: US Department of Education, National Center for Educational Statistics: HSLs 2009 and 2011

certified math and science teachers. School rigor indicators include whether schools require specific math or science courses for graduation. To capture the level of effort schools exert in helping students in STEM, we also control for whether a school offers STEM resources for students by combining various indicators into a composite variable called “School STEM climate.”³ The Cronbach’s alpha for this scale is 0.79, which suggests good underlying correlation. We also included indicators for whether schools have explicit tracking policies and whether they have programs that encourage underrepresented students to participate in STEM classes and programs. The latter gave us a sense of both the school’s overall interest in STEM and its commitment to assisting racial and ethnic minorities.

Propensity Score Matching Methodology

In an attempt to estimate effects using non-experimental data, we invoked a counterfactual causal framework wherein the effect is defined as the difference in outcome between the scenario in which an individual receives a treatment and the counterfactual scenario in which the same, or similar, individual does not (Morgan and Winship 2014; Winship and Morgan 1999). Here, the treatment is having a CBF and it is measured in the fall of ninth grade. We create our approximation of an “apples to apples” comparison by estimating students’ propensity to have a CBF, regardless of if they actually did in reality, conditional on achieving sufficient covariate balance using observed characteristics of children, their families, and the schools they attend.

Specifically, we used propensity score matching (PSM), as developed by Rosenbaum and Rubin (1983b), Rubin (1974, 1977, 1978, 1980), which is widely considered a suitable alternative for estimating effects in the absence of randomized data (Becker and Ichino 2002; Caliendo and Kopeinig 2008; Ho et al. 2007; Imai et al. 2008; Imbens 2004; Stuart and Rubin 2008). This technique matches subjects on available observable characteristics, creating

two groups that are similar on observed covariates. We compared students who had a CBF with the control group of students who did not and estimated the average treatment effect on the treated (ATT) for each outcome. The ATT is useful inasmuch as we are concerned with the effect of having a CBF for those students who, in reality, did have CBF. The strength of matching lies in its ability to reduce the role of observed covariates on any remaining differences between students who had a CBF and students who did not if having CBF depends exclusively on observed variables (D’Agostino 1998). That is, propensity score matching removes most of the bias due to observed covariates conditional on the assumption of “ignorable” treatment assignment. Ignorable treatment assignment holds where conditional on the observed covariates, selection into treatment is unrelated to unmeasured variables that affect the outcome. In what follows, we will test the robustness of this assumption in a formal sensitivity analysis.

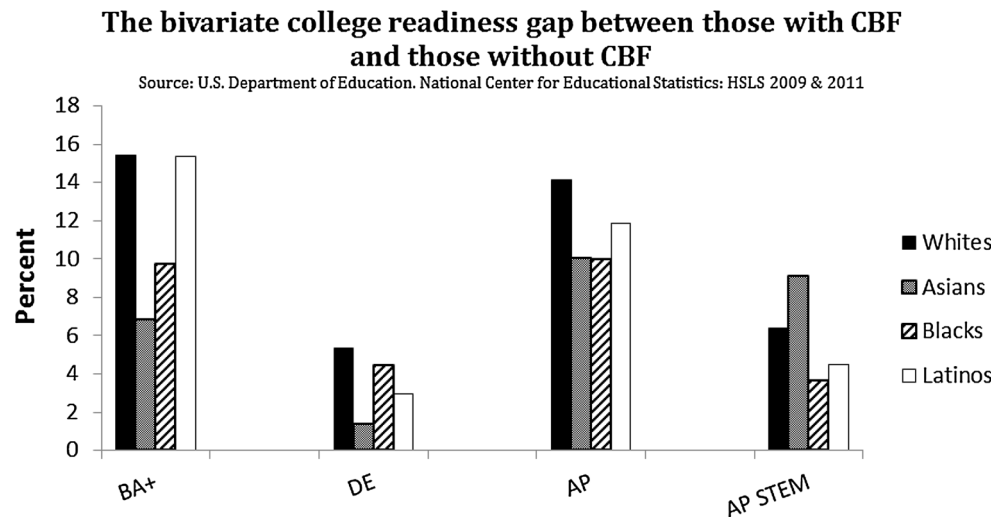
The first step in propensity score matching is to select the covariates upon which the matching will be based (Augurky and Schmidt 2001). We compiled a vector of theoretically relevant variables based on the literature that would conceivably predict having CBF.⁴ Table 1 lists the vector of covariates used for matching save for the outcomes. The list of matching variables is purposefully comprehensive since including as many potentially relevant covariates as possible typically will not reduce the quality of the matches, but excluding potentially relevant covariates runs the risk of creating bias in the estimation of CBF effects (Rubin and Thomas 1996). The selected covariates maximize the similarity between students with and without CBF. We entered all covariates into the selection model as main effects.

The second step in propensity score matching is to estimate the predicted probability of having CBF, which is known as the propensity score. We calculated propensity scores using a logit model (Leuven and Sianesi 2003) and matched students using kernel matching. Kernel matching ensures that students with and without CBF with similar propensity scores are matched within a given bandwidth, in this case 0.09 (Heckman et al. 1998; Stuart and Rubin 2008). The kernel approach is a nonparametric matching estimator that uses weighted averages of all individuals in the control group to construct the counterfactual outcome (Caliendo and Kopeinig 2008). All analyses are restricted to observations that fell in the region of common support to minimize the possibility of bad matches.

³ These school-level STEM resources are as follows: whether the school has a special focus in math or science, whether the school partners with MESA or similar STEM enrichment program, whether the school partners with a college or university that offers a math or science summer program, whether the school sponsors a science or math summer and after school programs, whether the school pairs students with a science or math mentor, whether the school holds math or science fairs, workshops or competitions, whether the school brings in guest speakers to talk about math or science, whether the school takes students on math- or science-relevant field trips, whether the school tells students about math or science contests, websites, blogs, or other programs, whether the school or district offers incentives to attract full-time high school science or math teachers, and whether the school offers AP Calculus (AB and BC), Computer Science (A and AB), Biology, Chemistry, and Physics on-site.

⁴ In order to control for heterogeneity in selection into the treatment and outcome that lies between schools, we included school-level variables in the matching model. This adjustment allows us to partially account for the clustered nature of the data and accurately specify the propensity score model (Arpino and Mealli 2008).

Fig. 1 Bivariate association between college-bound friends and college readiness. All differences that are displayed are significant at the 0.05 α level



The third step in propensity score matching is to evaluate whether the groups being compared have equal (or sufficiently balanced) distributions of relevant observed variables (Dehejia 2005). Balance testing is a critical step because causal inferences can be made only when sufficient balance is achieved on observed covariates across students with and without CBF. Therefore, the main purpose and utility of propensity score matching is to maximize covariate balance across these two groups (Augurzky and Schmidt 2001). Although there are a variety of methods for testing balance (t tests, Chi-squared tests, and F tests), we evaluated it by inspecting standardized bias scores. This approach avoids the “balance test fallacy,” where, for example, randomly deleting observations can improve balance. While hypothesis tests refer to populations, balance is a sample property that gives us a sense of how successfully the PSM technique has matched students with and without CBF based on observed characteristics (Imai et al. 2008). Using Stata’s `pstest` command (Leuven and Sianesi 2003), the standardized bias test calculates the difference of the sample means in the treated and non-treated subsamples as a percent of the square root of the average of the sample variances in the treated and non-treated groups (Rosenbaum and Rubin 1985). Caliendo and Kopeinig (2008) provide a formulaic definition of the standardized bias test.⁵ We evaluate balance in two ways: (1) by evaluating pre- and post-match improvements in balance for each variable in the matching model and (2) by evaluating the omnibus pre- and post-match improvements in balance for all of the variables combined. We report the latter in our results.

⁵ The `pstest` command relies on the average treatment effect on the treated (ATT) estimation based on the `psmatch2` command.

The final step in PSM is to compare the outcomes of the respondents in the different groups. We compared students with CBF to students who were similar on observed characteristics but did not have CBF. In doing so, we examine the effect of CBF in a nonparametric model that makes no assumptions about the functional form of the relationship between our covariates and outcomes. In particular, we calculated the average ATT, which represents the difference in outcomes between the groups being compared in the metric of percentages.

A key limitation of the propensity score matching technique lies in its inability to reduce the bias in estimated CBF effects that stems from unobserved covariates. Unobserved variables that affect both having CBF and the outcome threaten our ability to make strong causal inferences (Stuart and Rubin 2008). As a result, we conducted a formal sensitivity analysis of our statistically significant ATTs, which gauges the robustness of our estimates and increases confidence that these estimates represent “real” effects.

Sensitivity Analysis

Sensitivity analyses allow us to gauge the strength of estimates in the face of bias introduced through a hypothetical unobserved binary variable, U , that impacts having CBF (Rosenbaum 2002; Rosenbaum and Rubin 1983a). Rosenbaum (2002) argued that by generating a hypothetical unobserved variable and manipulating its effect on having CBF, researchers can place bounds (“Rosenbaum bounds”) for significance levels and confidence intervals around the CBF effect and assess how strong the unmeasured variable must be before the CBF effect is undermined (DiPrete and Gangl 2004). We used an extension of

Rosenbaum bounds, called Mantel–Haenszel bounds, that specifically addresses unobserved heterogeneity (i.e., “hidden bias”) when using binary outcomes (Becker and Caliendo 2007).

We used Stata’s `mhbounds` command for our sensitivity analysis that tests how the CBF effect changes based on specific violations to the ignorability assumption. The sensitivity analysis incrementally manipulates the odds ratio of having CBF, γ (Γ), until the original ATT is no longer statistically significant. That is, we continue to increase the odds of having CBF attributed to U until we “kill” our observed statistically significant ($p < 0.05$) ATT from the propensity score model, which we evaluate by examining the corresponding p -value associated with each increase in Γ . Our increments for Γ are in the metric of odds ratios and are 0.05 in size, ranging between 1.00 and 2.00. This method allows us to pinpoint which specific failure (within 0.05 increments) of the ignorability assumption implied by the particular configuration of the parameter, Γ , renders our results statistically insignificant. We can then compare the strength of association at the “kill” point of U to a conceptually important observed variable. This will help us gain a clearer understanding of the type of unobserved confounder necessary to undermine our results.

Results

Descriptive Results

Table 1 summarizes the CBF variable, the outcomes, and the covariates in terms of means and standard deviations. Asian (0.56) students reported the highest incidence of having college-bound friends followed by whites (0.51), blacks (0.47), and Latinos (0.46). Among the outcomes, Asians (85 percent) are most likely to state that they expect to earn at least a BA followed by whites (71 percent), blacks (69 percent), and Latinos (60 percent). All students seem to be equally as likely to take dual enrollment courses. However, Asian students (60 percent) are the most likely to have taken any AP course followed by whites (35 percent), Latinos (29 percent), and blacks (26 percent). Asian students (39 percent) are also most likely to have taken an AP STEM course followed by whites (18 percent) and blacks and Latinos (both 13 percent). Socioeconomic gaps between Asians and whites, on the one hand, and blacks and Latinos, on the other hand, are also made clear in these descriptive statistics. Asians and whites have higher family incomes and parental education than blacks and Latinos. Meanwhile, blacks and Latinos have higher rates of poverty than either Asians or whites. Asians and

whites also show higher rates of having at least one parent with a STEM degree than blacks and Latinos.

At the school level, blacks and Latinos are surrounded by lower SES student bodies than are Asians and whites but seem to be on par with Asians and whites on other school resources. Because the focus of this paper is on the relationship between college-bound friends and college readiness outcomes, we now turn to Fig. 1 to describe how CBF relates to college readiness.

Figure 1 shows the bivariate associations between having a college-bound friend and college readiness by racial and ethnic groups. We only display effects that are statistically significant at the $p < 0.05$ level. On the horizontal axis are the five outcomes that we study and on the vertical axis is the percentage point difference between students with a CBF and those without a CBF. That is, Fig. 1 shows the advantage associated with CBF on expecting to earn at least a BA, taking any dual enrollment courses, taking any AP courses, and taking any AP STEM courses. Among all students, the bivariate associations between CBF and educational expectations are generally uneven, suggesting that CBF may have heterogeneous effects by race and ethnicity. CBF appears to benefit whites and Latinos equally in their expectation to attain at least a BA. Asians and blacks, meanwhile, yield less powerful associations between CBF and educational expectations.

Figure 1 also summarizes bivariate associations between CBF and academic course-taking behaviors. The bivariate association between CBF and dual enrollment suggests a positive impact for CBF on all students, albeit to a lesser degree than the benefits for educational expectations. Moreover, white students appear to benefit the most from CBF on dual enrollment followed by blacks, Latinos, and then Asians. In contrast to dual enrollment, the association between CBF and having taken an AP course by the spring of 11th grade is much stronger for all racial and ethnic groups. College-bound friends benefit whites the most, while Latinos trail behind. The association between CBF and taking any AP appears to be about equal between blacks and Asians. Finally, while the association between CBF and AP STEM course taking is weaker than the association between CBF and any AP course taking among all students, blacks and Latinos trail behind Asians and whites in terms of how much they seem to be benefiting from CBF.

In short, the impact of CBF on college readiness appears to be uneven by race and ethnicity. Still, we must remember that non-random selection into social circles may drive these associations. To examine this possibility, we now turn to the results from the propensity score matching models.

Table 2 Treatment effects of a college-bound friend on college readiness by race and ethnicity

	Expects to earn at least a BA	Took any dual enrollment by spring 2012	Took any AP by spring 2012	Took an AP STEM course by spring 2012
<i>All students</i>				
CBF	0.06	0.02	0.08	0.04
Effect size (odds ratio)	1.38	1.14	1.39	1.25
Standard error	0.01	0.01	0.01	0.01
<i>T</i> -statistic	7.72	3.32	10.08	5.84
On support untreated	7,225	7,028	7,638	8,158
On support treated	7,841	7,523	8,170	8,473
Mean bias pre-matching	11.10	11.11	10.73	11.28
Mean bias post-matching	3.10	3.05	3.09	3.13
<i>Whites</i>				
CBF	0.06	0.03	0.08	0.03
Effect size (odds ratio)	1.39	1.19	1.41	1.23
Standard error	0.01	0.01	0.01	0.01
<i>T</i> -statistic	5.89	3.54	8.17	4.12
On support untreated	4,454	4,352	4,705	5,015
On support treated	5,006	4,804	5,184	5,369
Mean bias pre-matching	12.36	12.32	11.92	12.47
Mean bias post-matching	2.94	2.83	2.86	2.96
<i>Asians</i>				
CBF	0.04	−0.01	0.05	0.06
Effect size (odds ratio)	1.45	0.96	1.25	1.27
Standard error	0.02	0.03	0.03	0.03
<i>T</i> -statistic	1.97	−0.33	1.86	2.19
On support untreated	532	526	594	621
On support treated	722	713	804	821
Mean bias pre-matching	7.97	7.84	7.03	7.85
Mean bias post-matching	2.65	2.68	2.56	2.83
	<i>Expects to earn at least a BA</i>	<i>Took any dual enrollment by spring 2012</i>	<i>Took any AP by spring 2012</i>	<i>Took an AP STEM course by spring 2012</i>
<i>Blacks</i>				
CBF	0.04	0.03	0.07	0.04
Effect size (odds ratio)	1.23	1.16	1.39	1.36
Standard Error	0.02	0.02	0.02	0.02
<i>T</i> -statistic	1.77	1.25	3.09	2.26
On support untreated	891	854	927	990
On support treated	827	779	846	884
Mean bias pre-matching	10.37	10.69	10.39	10.48
Mean bias post-matching	2.97	3.12	3.06	3.20
<i>Latinos</i>				
CBF	0.07	0.00	0.06	0.02
Effect size (odds ratio)	1.39	1.02	1.32	1.20
Standard Error	0.02	0.02	0.02	0.01
<i>T</i> -statistic	3.79	0.22	3.39	1.70
On support untreated	1,348	1,296	1,412	1,532
On support treated	1,260	1,211	1,316	1,378

Table 2 continued

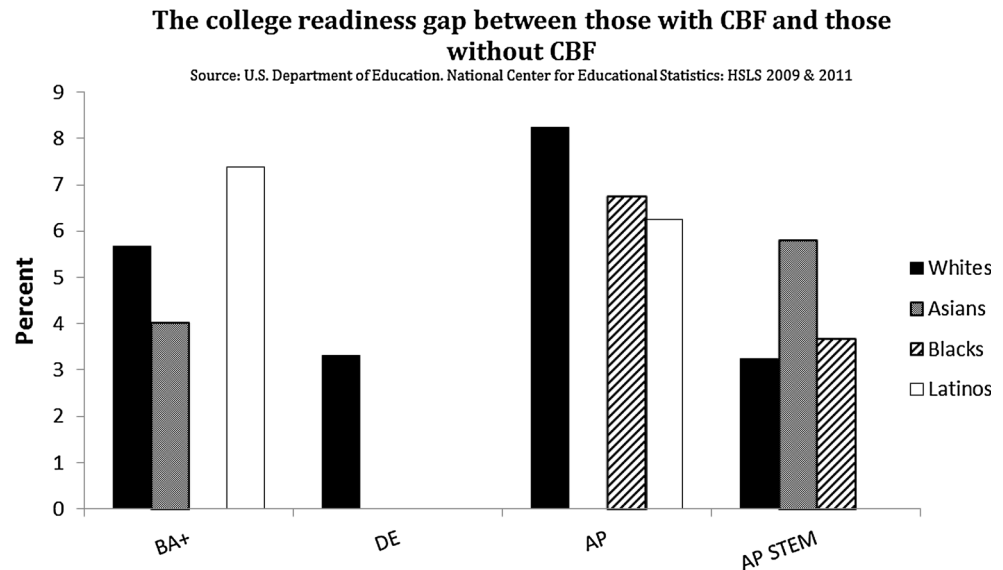
	<i>Expects to earn at least a BA</i>	<i>Took any dual enrollment by spring 2012</i>	<i>Took any AP by spring 2012</i>	<i>Took an AP STEM course by spring 2012</i>
Mean bias pre-matching	11.74	11.81	11.23	11.87
Mean bias post-matching	3.50	3.45	3.57	3.50

Bold ATTs denote statistically significant effects ($p < 0.05$)

All models include student characteristics, family background, and high-school characteristics (see Table 1) and exclude observations that were missing values on outcome variables

Source: U.S. Department of Education. National Center for Educational Statistics: HSLs 2009, 2011

Fig. 2 Propensity score matching CBF effects for college-bound friends on college readiness outcomes. All differences that are displayed are significant at the 0.05 α level



Propensity Score Matching Results

The analytical procedure of the propensity score matching model included three steps. First, we conducted an examination of CBF effects on each of the college readiness outcomes for all students. Second, we stratified the data by race and ethnic subgroup and, when necessary, conducted formal post hoc tests for the moderating effect of race and ethnicity on the relation between the CBF and each of the outcomes. Third, we conducted a formal sensitivity analysis to check the robustness of the estimated CBF effects to bias stemming from a simulated unobserved binary confounder.

Table 2 summarizes the results from the propensity score models for all students and for each racial and ethnic subgroup. We included the estimated ATT for CBF followed by the effect size in the metric of odds ratios, the standard error, t-statistic, sample sizes on the area of common support, and reported mean pre- and post-matching bias. The common support region is the area where the distribution of observed variables for student

with CBF is as similar as possible to those without CBF. Therefore, one can conceptually imagine that the only difference among individuals who share the same space in the common support region is whether they have a CBF or not. Therefore, it is important to match on the common support region (Caliendo and Kopeinig 2008) and students with CBF whose propensity scores exceed the largest propensity score for students without a CBF are left unmatched. All numbers in bold indicate statistically significant CBF effects at the 95 percent level of confidence.

The first panel indicates that CBF positively affects each outcome of college readiness for all students. CBF effects are most powerful, however, in increasing the probability students expect to earn at least a BA degree and the probability of them taking any AP course in the early years of high school. These main effects, however, may hide important differences in effects by race and ethnicity. The second panel summarizes CBF effects for white students. Results for white students mirror results for all students where CBF positively affects each of the college readiness outcomes. Moreover, for white students, CBF has the most

Table 3 Magnitude of the unobserved binary confounder's effect on selection of college-bound friends that renders our ATT null ($p < 0.05$)

	Expects to earn at least a BA Γ (odds ratios)	Took any dual enrollment by spring 2012 Γ (odds ratios)	Took any AP by spring 2012 Γ (odds ratios)	Took an AP STEM course by spring 2012 Γ (odds ratios)
All students	1.85–1.90	1.20–1.25	1.70–1.75	1.45–1.50
Whites	1.95–2.00	1.25–1.30	1.75–1.80	1.40–1.45
Asians	1.25–1.30	–	–	1.15–1.20
Blacks	–	–	1.30–1.35	1.25–1.30
Hispanics	1.55–1.60	–	1.45–1.50	–

Mantel–Haenszel bounds for treatment effects of college-bound friend on college readiness

Binary unobserved confounder effect on selection into the treatment (Γ) is in the metric of odds ratios in increments of 0.05 between 1.00 and 2.00

Source: U.S. Department of Education. National Center for Educational Statistics: HSLs 2009, 2011

powerful effects on BA expectations and any AP course taking.

The third panel summarizes results for Asian students. Here, we see that CBF effects begin to depart from the trend for the total sample and for whites alone. Among Asians, CBF only affects educational expectations (0.04) and AP STEM course taking (0.06). CBF does not impact dual enrollment course taking or any AP course taking among Asians.

The fourth panel summarizes CBF effects for blacks and demonstrates a positive effect for CBF on taking any AP course (0.07) and on taking an AP STEM course (0.04). Moreover, the effect of CBF on taking any AP course is almost equal among blacks and whites. Among blacks, CBF does not impact educational expectations or the probability of taking dual enrollment.

The fifth panel, which summarizes CBF effects for Latinos, demonstrates that CBF increases educational expectations (0.07) and the probability of taking any AP course by the spring of 11th grade (0.06). CBF does not affect dual enrollment or AP STEM course taking among Latinos. Moreover, the effect of CBF on educational expectations appears to be strongest among Latinos compared to any other racial and ethnic group.

Figure 2 graphs the propensity score matching model results for all racial and ethnic groups for a more direct comparison of the relative sizes of the CBF effects. We only display effects that are statistically significant at the $p < 0.05$ level. The graph demonstrates that white students benefit from CBF on all measures of college readiness, while all other groups yield more inconsistent effects. Black students are the only group for whom CBF does not impact educational expectations. CBF also consistently does not impact dual enrollment among all students, save for whites. Generally, CBF also has the strongest positive effect on students' probability of taking an AP course, although the effect of CBF appears to have a slightly

stronger effect on Latinos' educational expectations compared with their probability of taking any AP course. Finally, CBF positively affects the probability of taking AP STEM courses for all students except for Latinos. Although we confirmed our expectation that CBF has a positive effect on college readiness, the contours of these findings suggest that CBF does not impact all racial and ethnic groups equally. Whites unilaterally benefited from CBF on all college readiness outcomes, while the benefits for Asians, blacks, and Latinos were less consistent.

Lastly, we conducted a formal post hoc examination to test for the moderating effect of race and ethnicity on the CBF effects of CBF using a conventional t-test at the 95 percent confidence level (results available upon request). These tests revealed that all of the significant CBF effects for whites, Asians, blacks, or Latinos were statistically different from one another. In other words, all of the differences in CBF effects were statistically significant, leading us to conclude that race and ethnicity act as a moderator for these effects.

Sensitivity Analysis

Table 3 summarizes the results for the sensitivity analysis. "Appendix" provides a discussion of our rationale for conducting a sensitivity analysis. In Table 3, we report the range of gammas where the statistically significant ATT became statistically insignificant (i.e., the "kill zones") due to the unobserved confounder, U . The gammas (Γ) are presented as odds ratios ranging between 1.00 and 2.00 in increments of 0.05. To be clear, Mantel–Haenszel bounds assume that the unobserved confounder is perfectly correlated with the outcome, suggesting that these "kill zones" represent conservative bounds on our CBF effect. That is, they represent scenarios where the effect of U may be much stronger than we might expect *a priori* (DiPrete and Gangl 2004).

Moving across the rows, among white students, U would need to increase the odds of having college-bound friends by 95–100 percent to undermine the statistically significant effect of CBF on educational expectations. An unobserved confounder with this influence is quite large. Similarly, U would have to increase the odds of having college-bound friends by 25–30 percent to undermine the statistically significant ATT on the taking a dual enrollment for whites. For taking any AP course and for taking an AP STEM course, U would have to increase the odds of CBF by between 75 and 80 percent and 40 and 45 percent, respectively, to undermine the CBF effects for whites. These represent substantial levels of unobserved bias necessary to undermine our CBF effects.

Among Asians, U would have to increase the odds of CBF by between 25 and 30 percent to undermine the CBF effect on educational expectations and would have to increase the odds of CBF by between 15 and 20 percent to undermine the CBF effect on taking an AP STEM course. Among blacks, U would have to increase the odds of CBF by between 30 and 35 percent to undermine the effect on taking any AP course and would have to increase the odds of CBF by between 25 and 30 percent to undermine the effect on taking an AP STEM course. Finally, among Latinos, U would have to increase the odds of CBF by between 55 and 60 percent to undermine the effect on educational aspirations and would have to increase the odds of CBF by between 45 and 50 percent to undermine the effect on taking any AP course.

Although these results appear that it would take a fairly substantial omitted variable to undermine our results, it is unclear from this information alone whether the associations of U are substantively large. To give these results from the sensitivity analysis substantive meaning, we compared these confounders to conceptually important observed covariates in our propensity score model to draw comparisons with characteristics for which we have already accounted. To do so, we first analyzed the logistic regression models that predicted having CBF (results available upon request). Second, we searched for statistically significant covariates that had similar odds ratios as our “killer” confounders.

Among whites, U would have to have a stronger impact on selection into CBF (i.e., having a college-bound friend) than any of the covariates that we include in the PSM model (except for sex) to undermine the effect on educational expectations. Those observed characteristics include previous achievement, parents’ SES (e.g., income and education), and school resources. The confounder would also need to be slightly stronger than eighth grade science and math achievement in addition to being much stronger than ninth grade math scores, family SES, and school resources to undermine the CBF effect on dual enrollment for

whites. For taking any AP course and for taking an AP STEM course among whites, U would have to be much stronger than any observed covariate (save for sex) to undermine the CBF effects.

Among Asians, U would have to exert a stronger effect on CBF than family SES or school resources to undermine the CBF effect on educational expectations. The closest observed covariate to this would be prior math achievement, which is still not as strong as U . To undermine the CBF effect on taking an AP STEM course, U would have to have a stronger impact on CBF than prior science achievement and would have to be much stronger than family SES and school resources.

Among blacks, U would have to exert an impact on CBF that was much greater than prior achievement, family SES, and school resources to undermine the CBF effects on taking any AP course and taking an AP STEM course. Finally, among Latinos, U would have to exert an effect on having CBF that was beyond that for any of the observed covariates in our PSM model to undermine the CBF effects on educational expectations and any AP course.

Although we cannot rule out the possibility that unobserved confounders may undermine our results, we are confident that we have exhausted the available resources to try to find such a confounder. Our conclusion after this search is that although possible, it is unlikely that a plausible and theoretically relevant variable that we have not already included in our PSM model would undermine any of our results. Therefore, we are confident we have uncovered real CBF effects for college-bound friends on college readiness outcomes.

Summary and Discussion

Friends may have an enormous impact on children’s schooling trajectories. We have studied the relationship between having college-bound friends and college readiness outcomes early in the high school career because we view the early formation of social capital as salient for students’ educational success. Specifically, we have examined whether close college-bound friends affect educational expectations, dual enrollment, AP course taking, and AP STEM course taking and whether those effects vary for students from different racial and ethnic backgrounds. We applied propensity score models and sensitivity analyses to new restricted data from the HSLs to disentangle the effect of college-bound friends from observed and unobserved sources of bias. As expected, we conclude that college-bound friends generally have positive effects on students’ college readiness. However, we also conclude that these effects are not equally distributed across racial and ethnic subgroups. Instead, CBF most consistently affects white

students' college readiness while having uneven effects on Asians, blacks, and Latinos.

These results adhere to previous findings that have examined racial and ethnic differences in students' sensitivity to college-bound friends. Alvarado and Turley (2012) used Texas data and found race and ethnicity moderated the effect of college-bound friends on college application decisions. However, that study was not nationally representative, did not examine the impact of friends early on in the educational pipeline, and did not examine effects for blacks or Asians. The current study fills these gaps in our understanding of the impact of CBF on students' academic preparation for college in high school. In contrast to those earlier findings, we do not find that whites always experience the strongest effects for CBF. Instead, although whites exhibit the strongest effects on AP course taking and exhibit the only effects on dual enrollment, Latinos exhibit the strongest effects on educational expectations and Asians exhibit the strongest effects on AP STEM course taking. Interestingly, CBF has an even stronger influence on AP STEM course taking for Asians than for whites (0.03), suggesting that Asian students' high school social networks may place them in the most advantaged position to engage in post-secondary STEM subjects and set them up for an advantaged position in the labor market.

Exposure to CBF increases Latinos' educational expectations and their likelihood to take an AP course. One optimistic interpretation of the results is that Latino students' exposure to CBF results in their higher likelihood of translating higher educational expectations into higher advanced course taking. Based on our results, it would behoove educators and policy makers to increase college-going attitudes and behaviors among Latino students because doing so may have spillover effects on Latino social networks. Moreover, our findings suggest that policies and teaching strategies that engage Latino students with other students who are college bound can increase their college readiness and may increase their representation in post-secondary institutions. Although Latinos are not more likely to take AP STEM courses, they may be better off because of CBF in the long-term, especially when college admissions officers examine their advanced course work.

Blacks, meanwhile, exhibit positive effects for CBF on AP course taking and AP STEM course taking in spite of not experiencing gains on their educational expectations due to CBF. The effect of CBF on black students' probability of taking an AP STEM course is even stronger than the CBF effect for whites. CBF consistently improves black students' advanced course-taking behavior in high school, and the findings for blacks suggest that social capital investments early in high school can result in significant gains in the likelihood of following an advanced course trajectory. Our results suggest that exposing black

students to college-bound social networks can indeed have positive results on their college readiness and may increase their representation in post-secondary institutions. Taken together with the results for Latinos, it appears that the results suggest that policies that promote the creation and maintenance of social ties between minorities' and college-bound students, who may also be minorities, may have a substantial payoff in terms of educational attainment for these underrepresented students. Nevertheless, we must note that taking courses may not necessarily translate into doing well in these courses or other gains that white students achieve in rigorous academic courses. Indeed, previous work has found that racial and ethnic academic achievement gaps are strongest among students taking the most advanced courses (Riegle-Crumb and Grodsky 2010).

Although we found uneven CBF effects by race and ethnicity, we emphasize the CBF effect on AP course taking because it represents a crucial activity in the early high school years and positively affects black and Latino students' actual chances of enrolling in college. This is because AP courses involve a GPA boost, regardless of students' scores on AP examinations, and signals to college admissions counselors that these students are seriously committed to their studies.

Overall, our results contribute to the understanding of how close friends, rather than peers writ large, can impact student outcomes (Hallinan and Williams 1990; Parsons 1963). As expected from previous findings, close friends yield consistent effects on students' outcomes. Dual enrollment was the sole outcome that yielded positive effects only among whites. While we do not have an empirically compelling reason for this, we speculate that non-white students' ninth grade social networks and friends may not have as much information regarding dual enrollment as those of whites and therefore may not pass along information regarding dual enrollment.

However, similar to all analyses that rely on observational data, our results are subject to limitations. First, neither the panel structure nor the estimation strategy can completely account for factors that may undermine our results. Although we assess the sensitivity of our estimates to an unobserved binary confounder and conclude that they are robust to very strong sources of bias, unobserved confounders may linger. Second, the HSLs questions force us to assume that students discuss college plans with their closest friends. Third, we lack transcript data that have been a great resource for previous scholars and can more accurately measure friends' level of academic orientation (Riegle-Crumb et al. 2006). For instance, Crosnoe et al. (2003) and Riegle-Crumb et al. (2006) both used AddHealth data that take information directly from friends who also participated in AddHealth. Fortunately, this limitation will be alleviated in next round of the HSLs as transcript data will be available for all students.

A final limitation is that, given the nascent research on CBF on college readiness, it remains unclear whether the magnitude of effects in our study is large, moderate, or modest. As more social scientists continue to conduct research on this important topic, we as a research community would gain a better sense of the substantive importance of CBF on college readiness. In the meantime, we can compare our results to another study's results that examined factors associated with AP participation. Although not perfect, this comparison helps anchor the effects of CBF on college readiness to other important covariates related to participation in accelerated programs. To accomplish this task, we converted the CBF effects into odds ratios (see Table 2). We then compared the CBF effects in our study to effects related to participation in accelerated programs from other studies, but we used logit coefficients instead of odds ratios. In particular, Klopfenstein (2004) estimated factors associated with student AP participation in the state of Texas. For whites, blacks, and Latinos, in our study the effect of CBF on taking an AP course through the spring of 11th grade is 42–53 percent of the effect size of ever low income on AP participation found in Klopfenstein's (2004) study.

Nevertheless, our findings suggest that college-bound friends can plausibly improve students' college readiness. Further, our findings imply that these effects are more consistently influential for whites than for other racial and ethnic groups. Programs that promote social capital formation among high school students would likely prove beneficial to all students, but those that promote college preparation among blacks and Latinos may have important impacts on their likelihood of applying to and enrolling in college in greater numbers.

Appendix: Rationale for Sensitivity Analysis

The propensity score model is useful over traditional regression analysis because it can give us a sense of the extent to which we are meeting the assumption of ignorability. That is, we can formally assess how balanced observed characteristics are between students with and without CBF. When balance is high, selection bias (on observables) is low, and we can almost say that our estimated ATT mimics an experimental setting where covariates are essentially randomly distributed across students with and without CBF. Unfortunately, it is close to impossible to achieve this assumption. Therefore, we assessed our deviation from this ideal by using the `pstest` command in Stata, which yielded mean standardized bias estimates before and after matching. Our post-matching bias ranged from 2.56 (Asians; any AP) to 3.57 (Latinos; any AP), which are below the 5 percent bias threshold that is generally considered sufficient (Caliendo and Kopeinig

2008). The PSM model appears to have successfully reduced the bias on observed characteristics relative to traditional regression, but did not do so completely. This leaves open the possibility that unobserved covariates related to both CBF and the outcome may still bias our estimated CBF effects. Therefore, we conducted a sensitivity analysis that allowed us to assess how large an unobserved confounder, U , and its associated selection bias, must be in order to undermine our results.

References

- Adelman, C., Riley, R. W., McGuire, C. K., Lacampagne, C., & Dorfman, C. H. (1999). *Answers in the toolbox: Academic intensity, attendance patterns, and bachelor's degree attainment*. Jessup, MD: U.S. Department of Education.
- Allison, P. D. (2002). *Missing data*. Thousand Oaks: Sage Publications.
- Alvarado, S. E., & Turley, R. N. L. (2012). College-bound friends and college application choices: Heterogeneous effects for Latino and white students. *Social Science Research*, 41(6), 1451–1468. doi:10.1016/j.ssresearch.2012.05.017.
- An, B. P. (2013a). The impact of dual enrollment on college degree attainment: Do low-SES students benefit? *Educational Evaluation and Policy Analysis*, 35(1), 57–75. doi:10.3102/0162373712461933.
- An, B. P. (2013b). The influence of dual enrollment on academic performance and college readiness: Differences by socioeconomic status. *Research in Higher Education*, 54(4), 407–432. doi:10.1007/s11162-012-9278-z.
- An, B. P. (2015). The role of academic motivation and engagement on the relationship between dual enrollment and academic performance. *Journal of Higher Education*, 86(1), 98–126.
- Arcidiacono, P., & Nicholson, S. (2005). Peer effects in medical school. *Journal of Public Economics*, 89(2–3), 327–350.
- Arpino, B., & Mealli, F. (2008). The specification of the propensity score in multilevel observational studies Munich Personal RePEc Archive Paper No. 17407, posted 20.
- Aud, S., Fox, M. A., & Kewal-Ramani, A. (2010). *Status and trends in the education of racial and ethnic groups*. (NCES 2010-015). Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.
- Augurzky, B., & Schmidt, C. (2001). *The propensity score: A means to an end*. Discussion Paper No. 271.
- Bates, L. A., & Anderson, P. D. (in press). Do expectations make a difference? A look at the effect of educational expectations and academic performance on enrollment in post-secondary education. *Race and Social Problems*.
- Becker, S. O., & Caliendo, M. (2007). Sensitivity analysis for average treatment effects. *Stata Journal*, 7(1), 71–83.
- Becker, S. O., & Ichino, A. (2002). Estimation of average treatment effects based on propensity scores. *The STATA Journal*, 2(4), 358–377.
- Burke, M. A., & Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of Labor Economics*, 31(1), 51–82. doi:10.1086/666653.
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31–72.
- Cheng, S., & Starks, B. (2002). Racial differences in the effects of significant others on students' educational expectations. *Sociology of Education*, 75(4), 306–327.

- Riegle-Crumb, C., Farkas, G., & Muller, C. (2006). The role of gender and friendship in advanced course taking. *Sociology of Education*, 79(3), 206–228.
- Riegle-Crumb, C., & Grodsky, E. (2010). Racial-ethnic differences at the intersection of math course-taking and achievement. *Sociology of Education*, 83(3), 248–270. doi:10.1177/0038040710375689.
- Rosenbaum, P. (2002). *Observational studies*. New York: Springer.
- Rosenbaum, P., & Rubin, D. (1983a). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society Series B*, 45(2), 212–218.
- Rosenbaum, P., & Rubin, D. B. (1983b). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.
- Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician*, 39(1), 33–38.
- Royston, P. (2005). Multiple imputation of missing values: Update of ice. *The Stata Journal*, 5(4), 527–536.
- Rubin, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66, 688–701.
- Rubin, D. (1977). Assignment to treatment group on the basis of a covariate. *Journal of Educational and Behavioral Statistics*, 2, 1–26.
- Rubin, D., & Thomas, N. (1996). Matching using estimated propensity scores: Relating theory to practice. *Biometrics*, 52(1), 249–264.
- Rubin, D. (1978). Bayesian inference for causal effects: The role of randomization. *Annals of Statistics*, 6(1), 34–58. doi:10.1214/aos/1176344064
- Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *The Quarterly Journal of Economics*, 116(2), 681–704.
- Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? *Handbook of the Economics of Education*, 3(3), 249–277. doi:10.1016/S0169-7218(11)03004-8.
- Sadler, P. M., & Tai, R. H. (2007). Weighting for recognition: Accounting for advanced placement and honors courses when calculating high school grade point average. *National Association of Secondary School Principals*, 91(1), 5–32.
- St. John, E. P. (1991). What really influences minority attendance? sequential analyses of the high school and beyond sophomore cohort. *Research in Higher Education*, 32(2), 141–158.
- Stack, C. (1974). *All our kin*. New York: Harper & Row.
- Stearns, E., Potochnick, S., Moller, S., & Southworth, S. (2010). High school course-taking and post-secondary institutional selectivity. *Research in Higher Education*, 51(4), 366–395. doi:10.1007/s11162-009-9161-8.
- Struhl, B., & Vargas, J. (2012). *Taking college courses in high school: A strategy for college readiness: The college outcomes of dual enrollment in texas*. Boston: Jobs for the Future.
- Stuart, E., & Rubin, D. B. (2008). Best practices in quasi-experimental designs: Matching methods for causal inference. In J. Osborne (Ed.), *Best practices in quantitative methods* (pp. 155–176). Thousand Oaks: Sage Publications.
- Suarez-Orozco, C., & Suarez-Orozco, M. (1995). *Transformations: Migration, family life, and achievement motivation among latino adolescents*. Stanford, CA: Stanford University Press.
- Texas P-16 Council. (2007). *Study on dual credit programs in texas: A report to the 80th legislature from the texas P-16 council*. Austin, TX: Texas Education Agency.
- Vigdor, J. L., & Nechyba, T. S. (2007). Peer effects in North Carolina public schools. In L. Woessmann & P. E. Peterson (Eds.), *Schools and the equal opportunity problem* (pp. 73–101). Cambridge: The MIT Press.
- von Hippel, P. T. (2007a). Regression with missing Ys: An improved strategy for analyzing multiply imputed data. *Sociological Methodology*, 37, 83–117.
- von Hippel, P. T. (2007b). Regression with missing Ys: An improved strategy for analyzing multiply imputed data. *Sociological Methodology*, 37, 83–117.
- Way, N., & Chen, L. (2000). Close and general friendships among African American, Latino, and Asian American adolescents from low-income families. *Journal of Adolescent Research*, 15(2), 274–301.
- Winship, C., & Morgan, S. L. (1999). The estimation of causal effects from observational data. *Annual Review of Sociology*, 25, 659–706.
- Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *The Review of Economics and Statistics*, 85(1), 9–23.