Article

Journal of Sports Economics
2019, Vol. 20(1) 115-153

DOI: 10.1177/1527002517716975
journals.sagepub.com/home/jse
SAGE

Do Sports Crowd Out Books? The Impact of Intercollegiate Athletic Participation on Grades

Michael A. Insler[1] and Jimmy Karam[2]

Abstract

We investigate the influence of intercollegiate athletic participation on grades using data from the U.S. Naval Academy. Athletic participation is an endogenous decision with respect to educational outcomes. To identify a causal effect, we develop an instrument via the Academy's random assignment of students into peer groups. Instrumental variable (IVs) estimates suggest that sports participation modestly reduces recruited athletes' grades. This finding has implications beyond college, as we also show that grades—not athletic participation—are most strongly associated with postcollegiate outcomes such as military tenure and promotion rates.

Keywords

higher education, human capital, sports, athletic participation, academic achievement

Introduction

In 2007, the National Collegiate Athletic Association (NCAA) launched a widely viewed advertising campaign[1] on national television with the tagline: "There are over 380,000 student-athletes, and most of us go pro in something other than sports." In this campaign, three different television commercials suggest that college

[1] Department of Economics, U.S. Naval Academy, Annapolis, MD, USA
[2] U.S. Navy (Retired), Poway, CA, USA

Corresponding Author:
Michael A. Insler, Department of Economics, U.S. Naval Academy, 589 McNair Rd., Mail Stop 10D, Annapolis, MD 21402, USA.
Email: insler@usna.edu

athletics offer compelling "off the field" benefits to student-athletes, such as pride, confidence, and—in some cases—the basic opportunity to obtain a college education. More recently, various NCAA regulations and bylaws have come under scrutiny. In March 2014, the National Labor Relations Board determined that Northwestern University football players who receive scholarships are employees under the National Labor Relations Act and thus have the right to unionize.[2] A U.S. District Court ruling in August 2014 permits universities to offer trust funds of up to US$5,000 per year to football players in top 10 conferences and all Division I men' s basketball players.[3] These developments have called into question the central role of "amateurism" in NCAA athletics, generating a national debate regarding student-athletes' needs, rights, and—more generally—their college experiences.

There is extensive research on the role of higher education in the U.S. labor market. Economists have studied how college achievement is influenced by many aspects of the educational experience, including major selection, school quality, teacher quality, and peer effects. There is limited empirical research, however, on the influence of athletic participation on students' outcomes while in college. As the NCAA asserts in the quote above, very few student-athletes become professional athletes after college. It follows that most former NCAA athletes must leverage their college education upon entering the labor force, and their academic achievement at college is an important signal of their human capital. Previous work has shown, for example, that college grade point average (GPA) impacts earnings (Jones & Jackson, 1990; Wise, 1975), particularly for undergraduates who do not obtain graduate degrees (Gemus, 2010). Given these relationships and given the sheer number of student-athletes in the United States' system of higher education, it is important to investigate the link between academics and athletics. Our objective is to estimate the impact of intercollegiate athletic participation on students' academic achievement.

Athletic participation is a time-intensive undertaking. The NCAA limits participants to 20 "countable" hours per week in play or practice while in season and 8 hours per week while out of season.[4] But some teams may bypass these guidelines via captains' meetings (which can effectively be full-scale practice sessions), exceptions for such sports as baseball and football, or lack of enforcement of these rules. Do student-athletes' grades suffer, on average, due to their sports commitments? Or does their participation provide some academically helpful structure to their daily lives? While the direction of the true effect may differ from student to student, we aim to estimate its net average impact on the population of intercollegiate student-athletes. It is challenging to identify a causal channel because of the interdependence between athletics and academics. For instance, an ordinary least squares (OLS) regression of individual students' academic achievement (i.e., GPA) on a binary athletic participation indicator and observable controls (i.e., academic ability, high school preparation, etc.) may yield biased estimates. Students' decisions to participate in athletics may respond to their academic outcomes (simultaneity bias).[5] There may be unobserved differences between athletes and nonathletes, perhaps in their social skills or cognitive ability—beyond what is observable via SAT scores, high school achievement, and so on—that affect academic achievement (omitted variable bias). Athletes may depart from the university at a different rate than nonathletes (attrition bias). The presence of any of these factors would compromise standard OLS analysis.

We employ instrumental variables (IVs) estimation to circumvent these issues, using data from U.S. Naval Academy (USNA) students during the period 1991-2009. Identification of causal effects via IV requires a characteristic that is correlated with students' decisions to compete in varsity athletics but otherwise unconnected to their academic outcomes, conditional on observable characteristics. A unique feature of residential life at the Academy provides such a variable: Upon enrollment at USNA, the Admissions Office randomly assigns each student into a "company" that forms his or her primary peer group.[6] For each student, our instrument is specified as his or her number of companymates who were *recruited for a varsity sport* (hereafter, "recruits") prior to enrolling at USNA. Thus, the instrument captures the "intensity of athleticism" of a student's primary peer group, which should contribute to his or her sports participation decision. A key point is that at USNA, athletic participation is not compulsory for recruits once they arrive on campus, and all USNA students—regardless of their athletic status—receive free tuition, room, and board. It is also crucial that the instrument is based on a pre-USNA (i.e., pretreatment) characteristic, athletic recruitment. We cannot use, for example, the *current* athletic status of students' peers as an instrument because, although those peers are randomly allocated, own participation decisions may contemporaneously affect companymates' athletic participation (in which case, the IV itself would be endogenous). Another potential concern is that recruited companymates may *directly* influence own academic performance through a peer effect (beyond the indirect effect, via sports participation, that our identification strategy relies on). For example, the number of recruited companymates could affect own grades through choice of academic major or time devoted to extracurriculars (e.g., military responsibilities). To address this—that is, to provide additional support that the exclusion restriction is valid—we test for robustness using a variety of additional controls as proxies for possible peer effect channels. We find no evidence of such issues, thus the instruments are orthogonal to academic outcomes because of company random assignment, conditional on observable pre-USNA characteristics.

While the unique setting of USNA is necessary for causal identification, it differs from traditional Division I university settings in potentially important ways. Institutional Features of the USNA subsection illustrates USNA institutional features in more detail, and the final section of the article discusses possible external validity issues, following the complete presentation of our results.

On average, student-athletes' grades trail nonathletes' by a small margin: 0.17 on a 4-point GPA scale. The gap diminishes to 0.11 grade points upon controlling for background characteristics (via individual fixed effects) and then further to 0.019 upon controlling for time effects. IV estimates suggest that sports participation reduces student-athletes' GPAs by 0.038 grade points, on average, which is 6.2%

of one standard deviation of annual GPA. Our favored specification shows that a small but statistically significant effect exists for recruited athletes. We cannot be certain of the effect's sign, but point estimates are consistently negative across all specifications. The subpopulation that generates this local average treatment effect (LATE) consists of recruits in their senior year, and there is weak evidence that it also includes juniors and relatively high SAT performers. We find no statistically significant effect for unrecruited athletes, but the subpopulation of "complier recruits" is not substantially different from the greater student body, across observable characteristics. Lastly, to contextualize our main findings, we present some descriptive analysis of post-USNA outcomes, such as military tenure and promotion rates, for student-athletes compared to nonathletes. Higher GPA is associated with more military career success in general, but athletics are only associated with greater accomplishments within certain job tracks, such as the Marine Corps. Thus, there is little evidence that the potentially harmful effect of athletics on academics may be mitigated by intangible skills gained "on the field," like leadership, that may help in certain professional tracks.

The article proceeds as follows. The second section briefly discusses previous research on the connection between academics and athletics and illustrates how our data and the USNA environment present an opportunity to build upon these findings. The third section describes the econometric model and conditions necessary for instrument validity, and the fourth section presents estimation results, robustness checks, and findings from alternate models. The fifth section presents descriptive analysis of post-USNA outcomes, and the sixth section concludes.

Background

Athletic Participation and Human Capital Accumulation

There is limited existing research on this topic, and no study has directly addressed causality. Two other papers have examined similar institutional data. Maloney and McCormick (1993) assembled records from 4 academic years of Clemson University, a Division I school with a large athletic program. They first note a performance gap between the GPAs of athletes and nonathletes.[7] To explain the gap, they estimate regressions predicting GPA based on a wide range of students' characteristics, such as gender, race, high school academic standing, and SAT scores, as well as endogenous variables such as college major, college GPA to date, difficulty of course schedule, and athletic participation. After controlling for these observables, the GPA gap vanishes for all but in-season athletes in revenue-generating sports (i.e., men's basketball and football). In the second paper, Robst and Keil (2000) study institutional data from Division III Binghamton University using a similar regression framework. They determine that, opposite to the case of Clemson University, athletic participation is *positively* associated with GPA and graduation rates. Robst and

Keil suggest that their incompatibilities with Maloney and McCormick may stem either from institutional differences (e.g., Division III institutions are forbidden from offering athletic scholarships under NCAA guidelines) or because "athletes may be different from nonathletes in unobserved ways that would increase the likelihood of graduation, regardless of whether they participate in athletics" (Robst & Keil, 2000, p. 557). These papers' conflicting findings point to the importance of addressing such endogeneity issues.

If athletic participation affects achievement, then it follows that subsequent labor force outcomes should also be affected. Previous research has investigated nationally representative data, linking intercollegiate athletic participation to higher graduation rates (Long & Caudill, 1991) and higher earnings after college (Long & Caudill, 1991; Olbrecht, 2009). These studies, however, are descriptive in scope, and it is unclear if college athletics are important because they influence academic outcomes, various aspects of human capital, or if these associations are produced by unobserved traits of athletes.

More recent studies of causal effects have determined that athletic participation in other venues impacts labor market outcomes. High school athletic participation leads to more educational attainment (Barron, Ewing, & Waddell, 2000; Pfeifer & Cornelissen, 2010; Stevenson, 2010) and more labor force participation for women (Stevenson, 2010). Eide and Ronan (2001) use height as an instrument for high school sports participation, estimating effects on educational attainment that differ by gender and race. Athletic participation may also affect other skills. Leadership positions in high school, such as team captaincy, foster higher wages (Kuhn & Weinberger, 2005), and even recreational sports participation in adulthood may aid employability and wages (Lechner, 2009).

Our study contributes to this body of research as the first to directly assess the causal link between college athletics and academic achievement. We examine this link in the setting of USNA, which possesses a number of unique features that aid our efforts but also merit closer inspection.

Data

Our data have been provided by the USNA Office of Institutional Research. The office maintains a multitude of information taken from all students who attended USNA for at least one semester in academic years ending 1991-2009. Table 1 contains summary statistics (across person-years) for the full sample, for the subsample of varsity athletes, and for the subsample of nonathletes.

Precollege characteristics. We observe information for each student before he or she arrived at USNA: gender, math SAT score, verbal SAT score, "standardized" high school class rank (hereafter, "high school standing"),[8] whether recruited for a sport (and which sport, if so), whether previously enlisted in the military, race/ethnicity, and feeder source (if any).[9] As seen in Table 1, our sample is only 15% female—

Table 1. Summary Statistics.

Variable Name	Full Sample Mean	Full Sample Standard Deviation	Athletes Mean	Athletes Standard Deviation	Nonathletes Mean	Nonathletes Standard Deviation
Fixed characteristics						
Female[a]	0.15	0.36	0.21	0.41	0.13	0.34
SAT score (math)	663	64	647	65	669	63
SAT score (verbal)	639	69	620	69	645	68
High school standing[b]	566	122	545	122	574	121
Prior enlisted[a]	0.09	0.29	0.06	0.23	0.10	0.31
Recruited[a]	0.27	0.44	0.61	0.49	0.15	0.35
Race/ethnicity						
Asian[a]	0.04	0.20	0.03	0.17	0.04	0.21
Black[a]	0.06	0.24	0.08	0.28	0.05	0.22
Hispanic[a]	0.08	0.27	0.05	0.21	0.09	0.28
White or other[a]	0.81	0.39	0.84	0.37	0.81	0.40
Feeder source						
NAPS[a]	0.16	0.37	0.21	0.40	0.15	0.36
Foundation school[a]	0.07	0.25	0.07	0.25	0.07	0.25
USNA[a]	0.77	0.42	0.73	0.45	0.78	0.41
Year-specific characteristics						
GPA	2.90	0.61	2.78	0.61	2.95	0.61
Credit hours	34.8	3.7	34.7	3.5	34.9	3.7
Ease of schedule	2.90	0.21	2.86	0.20	2.92	0.21
No. of recruited companymate–classmates	8.67	2.62	8.69	2.63	8.66	2.62
USNA-exit characteristics						
Years of service	8.27	4.82	8.25	4.76	8.29	4.85
Ever attained rank						
O-3[a]	0.86	0.34	0.86	0.34	0.86	0.34
O-4[a]	0.27	0.44	0.27	0.44	0.27	0.44
O-5[a]	0.10	0.30	0.10	0.31	0.10	0.30
Service assignment						
Surface warfare[a]	0.25	0.43	0.27	0.44	0.24	0.42
Nuclear operations[a]	0.15	0.36	0.11	0.31	0.17	0.37
Naval aviation[a]	0.34	0.47	0.30	0.46	0.35	0.48
Marine corps[a]	0.18	0.39	0.19	0.40	0.18	0.38
Other[a]	0.09	0.28	0.13	0.33	0.07	0.26
Observations (person-years)	78,948		21,000		57,948	

Note. USNA = U.S. Naval Academy; GPA = grade point average; NAPS = Naval Academy Preparatory School.

[a]Sample proportion rather than sample mean (i.e., dummy variable). [b]A standardized version of students' high school academic standing, calculated by the Admissions Office. Refer to "Data" subsection for a description of this variable.

although this proportion has increased to nearly 25% in recent years—and students were, on average, strong high school performers. Nine percent of the sample were previously enlisted in the military,[10] and 23% attended a feeder program.

Year-by-year college characteristics. The data contain information for each student during each academic year at USNA: course code (e.g., Econ 101), grade, and number of credits for every course completed (from which we calculate standard GPA on a four point scale); company assignment (the all-important peer group, described in the next subsection); military achievement score (incorporates measures of military performance, conduct, and physical fitness); and whether participated in a varsity sport (and which sport, if so). There are 78,948 person-year observations in total, and roughly one quarter of them are from varsity student-athletes.

College exit characteristics. We observe whether each student graduated. All graduates immediately enter into a commitment of 5 or more years to serve in the U.S. Navy or Marine Corps, and we see their years of service (or current years of experience, if still active), highest military rank achieved (or current rank as of 2012, if still active), and assignment of service community (surface warfare officer, naval aviation, Marine Corps, etc.). Summary statistics for these exit characteristics in Table 1 do not account for timing issues. For instance, graduates who "attained a rank of O-4 or higher" only come from the classes of 2002 and earlier, because officers very rarely attain the O-4 rank without at least 10 years of experience. We describe these outcomes in more detail in our analysis in the fifth section.

Institutional features of the USNA. USNA participates in athletics as a Division I NCAA school. During the sampling period, there have been 31 distinct varsity sports programs, including men's, women's, and co-ed teams. Table 2 displays total participation counts for each varsity sport. Historically, football has the largest contingent of athletes, followed by sprint football, sailing (co-ed), men's track, and women's track. Some sports have much larger contingents of recruits (e.g., swimming, lacrosse) than others (e.g., sprint football, crew), and there is heterogeneity in GPAs across sports but nonrecruited participants tend to earn higher grades.

There are several institutional features of USNA that are directly relevant to our study. We introduce each feature here but elaborate where necessary in the sections that follow. While USNA possesses unique features, described below, that enable causal identification, it differs from traditional Division I universities in potentially important ways. We discuss possible external validity issues in the final section of the article, following the complete presentation of our results.

Random assignment. All students at USNA live in one on-campus dormitory. Single halls within the dormitory house 30 companies of approximately 150 students each, and each contains an even mix of freshmen, sophomores, juniors, and seniors. A

Table 2. Participation Counts and GPA by Varsity Sport.

	Athlete-Years		Average GPA		
Sport Name	Recruits	Nonrecruits	Recruits	Nonrecruits	Note
Men's sports					
Baseball	566	112	2.61	2.79	
Basketball	369	69	2.37	2.60	
Cross country	357	187	2.98	3.01	
Diving/swimming	692	78	2.75	3.10	
Football	2,387	390	2.42	2.64	
Golf	190	72	2.90	3.13	
Gymnastics	26	70		2.81	
Heavyweight crew	326	425	2.87	3.06	
Lacrosse	934	117	2.61	2.74	
Lightweight crew	259	530	2.99	3.15	
Soccer	531	125	2.74	2.80	
Sprint football	309	1,713	2.83	2.82	
Squash	172	183	2.77	3.09	
Tennis	226	63	2.84	2.85	
Track	998	657	2.75	2.82	
Volleyball	24	25			Team present 1991–1993
Water polo	480	49	2.71	2.78	
Wrestling	733	413	2.68	2.79	

(continued)

Table 2. (continued)

	Athlete-Years		Average GPA		
Sport Name	Recruits	Nonrecruits	Recruits	Nonrecruits	Note
Women's sports					
Basketball	290	45	2.59	2.66	
Crew	343	459	2.96	2.89	
Cross country	431	127	3.01	2.99	
Diving/swimming	554	88	2.83	2.81	
Gymnastics	21	11			Team present 1991–1993
Lacrosse	44	76	2.79	3.05	Team present 2007–2009
Soccer	357	50	2.91	2.79	Team present 1994–2009
Tennis	3	8			Team present 2009
Track	860	316	2.86	2.83	
Volleyball	259	65	2.66	2.73	
Co-ed sports					
Fencing	33	62	2.75	2.90	Team present 1991–1993
Rifle/pistol	186	207	2.78	2.87	
Sailing	569	1,748	2.82	2.99	

Note. Some GPA averages are omitted (masked) due to small subsample sizes. GPA = grade point average.

student's company is arguably the most important influence on his or her social experience at USNA. Companymates live, eat, study, drill, and compete together, building unit cohesion, teamwork skills, and morale under arduous and challenging experiences.[11] For many varsity athletes, company interactions complement the social interactions that they already receive within their team.

Upon arrival, every freshman is assigned into a company via a procedure that produces a diverse but randomly allocated mix of students in each company. According to the Admissions Office, students are first randomly spread across companies based on predetermined characteristics: race, gender, home state, prior military service, and attendance at a 1-year Naval Academy preparatory school. After these initial stratifications, administrators randomly assign all remaining students to companies. The key features of the procedure are that students have no control over the outcome—for instance, USNA does not solicit interests, lifestyle details, or roommate preferences as is typical at other universities—and it produces an allocation that is effectively random. A small number (approximately five per year) of assignments are determined by the Academy's administration to avoid placing siblings or relatives in the same company. There have also been so-called shotguns in many years, in which an entire class may be reassigned (following the same stratified assignment procedure) to new companies, or in which class cohorts within companies remain together but are swapped into different companies (via random assignment).[12] Company assignment may also change in exceptional circumstances, such as via "love chits" or "hate chits," which may arise from individual students' personal relationships[13] (again, the reassignment is performed randomly).

The company assignment procedure produces minimal variation in the above pre-USNA traits across companies, but it ensures that one's actual peer group is not self-selected. The mechanism prevents athletes from sorting into residence hall groupings that could offer academic advantages, and it is a key aspect of our identification strategy detailed in the third section. Given the importance of random assignment, we carry out a simple verification test. Similar to a test shown in the peer effects literature by Feld and Zölitz (2014), we regress a freshman's recruitment status R_i on a set of company dummy variables (performing separate regressions for each academic year). Under random assignment, we expect the F-statistics from these regressions to be predominantly statistically insignificant. We observe 19 cohorts of freshmen; thus, we estimate 19 separate regressions. Table 3 shows model F-statistics and corresponding p values for each regression. There is no evidence of systemic groupings of recruits into companies.[14]

Universal scholarships. Students at USNA do not pay tuition and are barred from outside employment. As a result, there is no notion of an "athletic scholarship." In fact, there is nothing requiring recruited athletes to participate in their recruited sport. At other Division I universities, recruited athletes are predominantly on 1-year athletic scholarships.[15] Scholarships may not be renewed from year-to-year due to coaches' updated evaluations of athletic ability, performance, contribution, or even injury or illness. At

Table 3. Randomization Verification.

Academic Year	F-Statistic	p Value
1991	0.81	0.78
1992	0.43	1.00
1993	0.86	0.69
1994	0.86	0.70
1995	0.81	0.78
1996	0.85	0.72
1997	0.78	0.80
1998	1.05	0.39
1999	0.62	0.94
2000	0.87	0.66
2001	0.87	0.66
2002	0.47	0.99
2003	0.12	1.00
2004	0.18	1.00
2005	0.56	0.97
2006	0.56	0.97
2007	0.28	1.00
2008	0.62	0.94
2009	0.80	0.77

Note. Table refers to 19 stratified regressions (by academic year) of recruitment status of freshmen on a set of company dummy variables. For each academic year, the table reports *F*-statistics (and corresponding *p* values) from the test that company dummies' coefficients are jointly equal to zero. We do not reject the hypothesis for any academic year, which is consistent with random assignment of recruits across companies.

USNA, all students, athlete or not, effectively possess full scholarships that are guaranteed for 4 years. Thus, we naturally bypass potentially problematic selection bias that could otherwise stem from athlete attrition due to scholarship loss.

Core course requirements. At USNA, all students must pass or test out of a wide range of core courses in a range of subject areas such as calculus, chemistry, political science, history, ethics, and law. These courses form their entire first-year schedule, with few exceptions.[16] While focusing on academic major requirements, upperclassmen continue to enroll in core courses taken by all, such as electrical engineering, leadership, and navigation, further constraining their scheduling. Thus, students' curricula are relatively homogeneous, conditional on their academic major, so they have limited ability to sort into easier courses. As our goal is to model GPA as a function of observable characteristics, this feature equalizes, to some extent, students' course schedule difficulty. GPA calculated solely from core courses also presents an opportunity to check that our main results are robust to students' (limited amount of) course selection.

4-Year residency limit. All students must complete their degree requirements in 4 academic years, with very few exceptions (none of which are related to athletics).[17] Students must also carry at least 15 credit hours in their academic schedules each semester. According to a 2013 NCAA report(NCAA Research, 2013), nationwide Division I student-athletes graduate at rates comparable to the general student body. Eckard (2010), however, finds that this may vary based on the particular graduation rate metric. At USNA, the experiences of athletes and nonathletes are identical in this dimension.

Physical mission. All USNA students commit to a "physical mission" that may weaken the distinction between athletes and nonathletes. Each student must pass the Physical Readiness Test (PRT)[18] each semester to verify a required fitness level. A larger proportion of the USNA student body competes in time-intensive nonvarsity athletics compared to other universities because USNA hosts several "club sports." These sports teams, although not associated with the NCAA, provide organized athletic opportunities for approximately 400–500 additional students every year. Information provided by club sports administrators suggests that only a few club sports (such as men's and women's rugby and men's ice hockey) entail a similar time commitment to varsity sports. Unfortunately, we only observe club sports participation data during 2000-2009, so we omit it from our analysis to preserve the larger sample size. Additionally, all nonvarsity, nonclub athletes must participate in some type of regular athletic activity, such as an intramural sport. Intramural sports (e.g., soccer or ultimate frisbee) generally practice and compete for no more than 3 hr per week. Although we do not observe records of intramural-level athletic participation, we expect the time commitment to be similar for all participants, because such activity is uniformly required of all nonvarsity, nonclub athletes.

Econometric Model

Baseline specification

Our objective is to estimate the production of educational outcomes for student i in years $t = 1, 2, 3, 4$ of college. We adopt a linear specification which takes the following form,

$$G_{it} = \gamma S_{it} + \delta P_{it} + \lambda E_{it} + \eta_i + \varepsilon_{it}, \tag{1}$$

in which dependent variable G_{it}—academic achievement—is captured by student i's GPA in his or her tth year of college.

Covariates. Our empirical work uses the following explanatory variables for the specification in Equation 1:

- S_{it} is a binary indicator of whether student i participated in varsity athletics in his or her tth year of college.
- P_{it} includes controls for peers' quality, which may be important for validity of the IV strategy described below. We estimate several versions of the model, where P_{it} includes various stratifications and interactions of the average math

and verbal SAT score across student i's companymate peers.[19] This variable is indexed by both i and t because, although its information comes solely from student i's peers' pre-USNA information, i's peer cohort may change exogenously over his or her 4 years at USNA (we elaborate on this in Corrective Measures subsection).

- E_{it} incorporates three types of "environmental factors" at USNA that could change over the sample period and over students' years of attendance: (1) academic year-specific dummy variables (1991-2009, interpreted as annual grade inflation compared to a baseline year); (2) class-specific dummy variables ($t = 1, 2, 3, 4$, interpreted as systemic GPA differences between freshmen, sophomores, juniors, and seniors); (3) E_{it} may also include other time-t (endogenous) variables that are codetermined along with sports participation S_{it}, such as student i's credit load in year t or military performance scores.

Omitted Variables. η_i and ε_{it} represent unobserved factors affecting student i's GPA in period t. η_i represents fixed academic ability. It includes pre-USNA characteristics described in the previous section as well as other unobserved time-invariant traits.

ε_{it} contains time-varying, individual-specific, unobservable factors affecting grades. It is important to note that ε_{it} does *not* contain athletic-related effects on studying effort (e.g., travel for competitions, team parties, athlete-only academic resources), nor does it contain athletic-related personal events that could affect grades (e.g., injuries, fatigue). These events—and potentially other forms of "athletic bias" such as instructor favoritism or discrimination—are natural components of the studying athletics leisure trade-off and thus should be captured by the coefficient for S_{it}. An athlete's decision to quit coincides with the decision to forego these advantages (or disadvantages), thus γ should subsume these sorts of effects.

Problems. Even under this interpretation, simple OLS estimation of Equation 1 poses three main drawbacks vis-à-vis potential endogeneity of S_{it}:

1. Estimates of γ may suffer sample selection bias if a larger (or smaller) proportion of student-athletes depart the institution, compared to nonathletes. If the departing athletes are weaker academically, then we expect an upward bias for estimates of γ (and vice versa).
2. Athletic participation S_{it} may be correlated with fixed academic ability η_i. If athletes tend to have lower unobserved academic ability (perhaps because more academically inclined students steer away from athletics), then estimates of γ would be biased downward (and vice versa).
3. Athletic participation S_{it} may be correlated with individual time-variant factors in ε_{it}. For instance, simultaneity bias could result from a family emergency adversely affecting a student's classroom performance in period t while also inducing the student to cease athletic pursuits in the same period.

Given the assortment of possible problems, we cannot predict the sign of the bias in OLS estimates of γ. Raw averages show that nonathletes slightly outperform student-athletes in the classroom, but we must address these endogeneity issues to ascertain whether sports participation exerts a positive or negative effect at the margin.

Corrective measures

Sample selection bias via attrition (#1 above) does not pose an issue in this setting, which we argue below. Fixed effects estimation handles any time-invariant omitted variables (#2 above). Finally, an IV strategy addresses any further endogeneity issues (#3 above), also developed below.

Attrition. At most Division I schools, varsity athletes who quit may not have the financial means to remain enrolled in school after losing their scholarship. At USNA, however, there is no notion of an "athletic scholarship" because all students receive full tuition, room, board, and a small monthly stipend. Attrition would still pose a problem if athletes tend to depart for nonfinancial reasons at different rates than nonathletes, but empirical evidence suggests otherwise. In our sample, 11.4% of nonathletes and 11.1% of athletes drop out of the institution.[20] More specifically, 4.3%, 5.4%, and 1.7% of nonathletes leave before the start of their sophomore, junior, and senior years, respectively. For athletes, the comparable attrition rates are 3.8%, 5.4%, and 1.7%, respectively.[21] Additionally, using a cross section of students from our panel, we estimate a probit model with graduation as the outcome and a subset of covariates from Equation 1 as controls (results available from the authors by request). The coefficient for varsity athletic participation was statistically significant but tiny; athletes are 0.69% *more* likely to graduate than nonathletes, conditional on demographics, SAT score, high school standing, prior military experience, and graduation year. In summary, empirical evidence reveals a minimal role, if any, for attrition in our study, and consequently we do not model it below.

IVs. We account for students' endogenous athletic participation decisions via an IV strategy. The instrument comes from information on the number of recruited students within a company. As previously mentioned, upon their arrival at USNA, students are randomly assigned into companies that form their primary peer groups in the residence hall and many other aspects of daily life. For IV estimation, we desire a variable that is correlated with S_{it} but otherwise uncorrelated with academic outcomes, conditional on the covariates in Equation 1. We define the instrument Z_{it} to be the number of students within i's company and class year cohort during i's tth year of college (excluding student i himself or herself) who were recruited to play a varsity sport before enrolling at USNA. Z_{it} aims to capture the influence of a student's recruited peers on his or her own sports participation decision.[22] Therefore, for Z_{it} to be a valid instrument, the following relevance condition must hold:

Claim 1: $Cov(Z_{it}, S_{it}|P_{it}, E_{it}, \eta_i) \neq 0$.

The first-stage results (discussed in the next section) support this condition.

It is also crucial that the instrument is derived from a pre-USNA characteristic, athletic recruitment; we cannot use the *current* athletic status of students' companymates as an instrument because, although those peers are randomly allocated, own participation decisions may affect companymates' athletic participation. The number of recruited companymates does not suffer such a simultaneity problem because it is essentially impossible that a USNA student could have influenced the recruitment of one of his or her companymate–classmates.

A potential concern is that recruited companymates may *directly* influence own academic performance (the indirect effect, which works through sports participation, is essential for identification). In other words, it is conceivable that Z_{it} is not actually excludable from the structural equation. Such an effect may primarily act through an academic peer effect channel. For example, a higher number of recruits in 20th company might correspond with a lower proportion of academically focused students in that company. We propose to control for such possible academic peer effects by estimating variants of the IV model that include additional controls as follows:[23]

1. We estimate versions of the model that contain two new covariates of the form $\frac{\sum_{k\neq i} R_k \times \text{SATM}_{kct}}{\sum_{k\neq i} R_k}$ and $\frac{\sum_{k\neq i} R_k \times \text{SATV}_{kct}}{\sum_{k\neq i} R_k}$, representing the average math and verbal SAT scores across each *recruited* student k of i's class cohort within company c. This allows peer effects on grades to exclusively stem from a company's recruited students. (R_i is a binary indicator of whether student i was recruited as a varsity athlete, prior to enrolling at USNA.) These controls are represented in Equation 1 by P_{it}.
2. We estimate versions of the model that contain two new, even more flexible, covariates that permit the peer effects' coefficients to differ by own recruitment status. That is, in this model, we include the two covariates from #1 above as well as their interactions with own recruitment: $R_i \times \frac{\sum_{k\neq i} R_k \times \text{SATM}_{kct}}{\sum_{k\neq i} R_k}$ and $R_i \times \frac{\sum_{k\neq i} R_k \times \text{SATV}_{kct}}{\sum_{k\neq i} R_k}$. As mentioned above, these controls are represented in Equation 1 by P_{it}.

One can imagine other vulnerabilities of the exclusion restriction. For example, the number of recruited classmates–companymates could affect grades through channels other than sports participation and academic peer effects, such as time devoted to physical training, intensity of military studies, choice of academic major, or social interactions (watching sporting events, drinking, smoking, etc.). We propose to proxy for these possible social peer effects by estimating versions of the IV model that contain additional controls for:

1. Military achievement score,[24]
2. Credit hours and the "ease" of each student's course schedule, and[25]
3. Company fixed effects,[26]

These controls are represented in Equation 1 by E_{it}. It remains possible that these covariates are not adequate controls for direct peer effects that could stem from Z_{it}. But, as we will show, the IV estimation is robust to their inclusion, suggesting that such channels are not influential components of the mechanism at hand. In summary, we must assume that the number of recruits within one's company–class cohort is orthogonal to his or her GPA-affecting time-variant unobserved traits, conditional on controls and fixed unobserved factors:

Claim 2: $Cov(Z_{it}, \varepsilon_{it}|P_{it}, E_{it}, \eta_i) = 0$.

As we have argued, this assumption is reasonable because of both the random assignment of students to companies and robustness to a wide variety of observable controls. If the instrument does not adequately capture students' peers' influence on their sports participation decisions, then the first stage would be too weak, but the orthogonality assumption would not be compromised.

The final model utilizes two-stage least square (2SLS) with fixed effects to estimate versions of Equation 1, instrumenting endogenous variable S_{it} with Z_{it}. 2SLS yields consistent estimates of the regression coefficients under the above assumptions.

Interpretation of Causal Estimates of γ. We interpret the corrected estimate of γ as the LATE of athletic participation on grades. Following Angrist, Imbens, and Rubin (1996), the calculated effect is attributable specifically to subpopulations who are affected by changes in the instruments. In other words, γ represents the average effect of participation on GPA for the subset of students whose athletic pursuits respond to changing peer group influences. The LATE does not *directly* apply to students who are "involuntary nonathletes"—perhaps due to a lack of athletic skill—because such students cannot join a varsity team. But if sports participation would hypothetically affect their grades similarly to the observed effect on the treated subpopulation, then estimates of γ *would* extend to the remaining population. We cannot test this counterfactual, but we examine the "complier" subpopulation in the next section, following the main estimation results.

Results

Baseline Estimates

We begin by estimating two baseline specifications of Equation 1 via fixed effects estimation. In all models that follow, standard errors are clustered by class groups within each company (e.g., sophomores in 23rd company in 1995 and juniors in 23rd

Table 4. Main Results.

Dependent Variable: GPA	FE 0	FE 1	IV FE 1	FE 2	IV FE 2
Participated in Sport	−0.109***	−0.0186***	1.112	0.00263	1.523
	(0.00504)	(0.00468)	(1.549)	(0.00604)	(1.917)
× Recruited				−0.0517***	−1.561
				(0.00932)	(1.280)
$H_0 : \gamma_1 = \gamma_2 = 0$ *F*-statistic				23.32	23.53
Academic and class year controls	N	Y	Y	Y	Y
Observations	78,948	78,948	78,948	78,948	78,948

Note. All models account for individual fixed effects. Standard errors (in parentheses) are clustered by company–class groups. Dummy variable controls for academic years and class years are included where specified but not shown in the table. FE = fixed effects; IV FE = Instrumental Variable Fixed Effects.
*p < .05. **p < .01. ***p < .001.

company in 1996 are in the same cluster). In the first column of Table 4, we include the primary regressor of interest S_{it} but no other covariates. Here, S_{it}'s coefficient estimate $\widehat{\gamma} = -0.109$ is the difference between athletes' and nonathletes' yearly GPAs, conditional only on individual fixed effects. Varsity athletes trail nonathletes' by just over one tenth of a letter grade.

Column "FE 1" introduces controls for time effects: class year (e.g., freshman, sophomore) and academic year (e.g., 1995, 1996). Here, much of the previous gap between student-athletes and nonathletes is explained, as it now shrinks to only 0.019 grade points. The baseline estimates are instructive in establishing the basic correlations between grades and sports participation, but conditioning on observables and fixed effects is not sufficient. There are many possible unobserved confounding factors that can jointly influence grades and sports participation.

Corrected Estimates

Column "IV FE 1" of Table 4 provides results from the first stage of fixed effects IV estimation of Equation 1. Here we instrument for sports participation S_{it} with the number of recruited companymate–classmates Z_{it}.[27] Standard error estimates allow for clustering by class groups within each company, as before. Note that we have not yet included peer characteristics P_{it} in the specification; the next subsection shows that main findings are robust to their inclusion.

The instrument's coefficient estimate is small in magnitude, indicating that an additional recruit in one's company–class cohort is associated with a greater likelihood of own sports participation by 0.11 percentage points. It is not statistically significant. Z_{it} does not provide enough power for the first stage, which comes through in the second stage. "IV FE 1" of Table 5 displays second-stage IV estimates of Equation 1. Here we now find no statistically significant sports participation effect on GPA.

Table 5. First-Stage Results.

Dependent Variable	IV FE 1 S_{it}	IV FE 2 S_{it}	IV FE 2 $R_i \times S_{it}$
Number of Recruited Companymate–Classmates	0.00113	−0.00454***	−0.00824***
	(0.00112)	(0.00127)	(0.000821)
× Recruited		0.0211***	0.0317***
		(0.00226)	(0.00146)
H_0: Excluded instruments = 0 F-statistic		43.92	235.75
Academic and class year controls	Y	Y	Y
Observations	78,948	78,948	78,948

Note. All models account for individual fixed effects. Standard errors (in parentheses) are clustered by company–class groups. Dummy variable controls for academic years and class years are included in all models but not shown in the table. FE = fixed effects; IV FE = Instrumental Variable Fixed Effects.
*p < .05. **p < .01. ***p < .001.

To improve the first stage, we posit that recruits may respond to the presence of an additional peer recruit differently than nonrecruits. Many nonrecruits simply may not be athletic enough to participate in a varsity sport, in which case the instrument would be orthogonal to their participation decision. And compared to the nonrecruits who could play a varsity sport, recruits may be more (or less) predisposed to respond to peer stimuli within the company milieu. To incorporate this idea into our econometric model, we rewrite Equation 1, now allowing the coefficient on sports participation to differ by a binary indicator of recruitment status, R_i:

$$G_{it} = \gamma_1 S_{it} + \gamma_2 (R_i \times S_{it}) + \delta P_{it} + \lambda E_{it} + \eta_i + \varepsilon_{it}. \tag{2}$$

We estimate a baseline version of Equation 2 via fixed effects; column "FE 2" of Table 5 contains results. $(R_i \times S_{it})$ is statistically significant at the 0.1% level, but we cannot reject that the coefficient for nonrecruits $\gamma_1 = 0$. The association between athletics and grades appears to stem from recruits.

We now have two endogenous variables in the right-hand side of Equation 2. If Z_{it} is a valid instrument for S_{it}, then $(R_i \times Z_{it})$ is a valid instrument for $(R_i \times S_{it})$. We adopt this instrumentation and estimate Equation 2 using fixed effects IV estimation. The two columns labeled "IV FE 2" of Table 4 show results from the first stage. For recruits, the peer effect on their sports participation is over an order of magnitude larger than in the previous model; a one standard deviation increase in the number of recruited companymate–classmates (2.6 peer recruits) is associated with a 4.3% greater likelihood of own participation. For nonrecruits, the effect is *negative* but very small in magnitude. Coefficients are comparable in the first-stage regression with the interaction term as its dependent variable, and F-tests of the hypotheses that the two excluded instruments are jointly zero are strongly rejected for both regressions. Thus, it appears that recruits are more likely to participate if they are surrounded by athletic peers, while nonrecruits are unaffected or may even shy away from sports when among many recruited companymates. First-stage estimates imply that identification stems primarily from the behavior of recruits in our sample—we explore this point further in Analysis of the Complier Subpopulation subsection.

Table 5 column "IV FE 2" shows second-stage estimates of Equation 2. While neither γ_1 nor γ_2 is individually significant, we reject that they are jointly equal to zero (F-statistics are shown). Therefore, there is a statistically significant effect for recruits but we cannot reject that $\gamma_1 + \gamma_2 = 0$, which means that we are not certain (in a statistically significant sense) of the effect's sign. The point estimate is negative which suggests a small but harmful GPA effect only for recruits, and it is similar in magnitude to the uncorrected estimate. In this specification, our point estimate suggests that athletic participation harms recruited students' GPA by 0.038 grade points, which is 6.2% of one standard deviation of annual GPA. In additional specifications shown below, point estimates of the participation effect for recruits grow larger in magnitude, thus here we may underestimate its size. For nonrecruits, we do not witness a statistically significant effect of athletic participation. We interpret $\widehat{\gamma_1+\gamma_2}$ as a LATE; thus, it remains to be seen whether the estimated sports participation effect applies to the general population of college students.

Before examining the validity of these findings in greater detail, it is instructive to consider an alternate instrumentation that seems intuitive but ultimately leads to biased estimates. The marginal effect of one more recruited student on own athletic participation could increase if the additional peer was recruited *in the same sport.* We can model this behavior by redefining Z_{it} to include two variables: Z_{it}^{same} = the number of companymate – classmates recruited for the same sport as i; Z_{it}^{diff} = the number of companymates – classmates recruited for a different sport from i ($Z_{it}^{\text{diff}} = Z_{it}$ if i was not recruited). Under this instrumentation, we find a *positive but insignificant* effect equal to 0.1 additional grade points due to sports participation (results available upon request). This estimate could be biased, however, because the alternate instrumentation likely violates the exclusion restriction. The interaction specification of Equation 2 is only valid if Z_{it} is excludable from the structural equation. Under the new definition, Z_{it}^{same} by construction includes information on own recruitment status (which is not excludable) because it is only nonzero when student i is recruited. Due to this caveat, we proceed with the ungrouped Z_{it} of "IV FE 2" as our main model to further examine.

Validity of IV Estimates

Validity of the IVs strategy hinges on the excludability of Z_{it} from the second stage. As previously discussed, peer recruits could *directly* affect own academic performance through a classic academic peer effect. That is, if recruits are more likely to participate in sports when more of their companymate–classmates are recruits, this also means that more of their peers are academically weaker, on average.[28] In this case, peer effects such as those documented by Carrell, Fullerton, and West (2009) could bias the IV estimates. Results shown thus far have not accounted for this potential pitfall.

In Table 6, we first check for possible direct athletic peer effects by testing two variants of Equation 2 that include additional control variables (represented by P_{it} in Corrective Measures subsection). Column “IV FE 2” reproduces the main second-stage estimates from Table 5 for easy comparison. Column “IV FE 3” controls for average math and verbal SAT scores across recruited company-mate–classmates, which captures the standard “contextual” pretreatment peer effect from the peer effects literature. Column “IV FE 4” allows coefficients for the average math and verbal SAT scores of recruited companymate–classmates to differ by own recruitment status. Results are qualitatively similar to “IV FE 2,” although point estimates of the participation effect are smaller (0.001 and 0.015 grade points, respectively). Here, peer effects may explain part of the effect seen in the main model. Alternately, effect size differences across models are not statistically significant and thus may stem from noise; there is relatively low within-individual variation in the instrument, which would lead to high variance in the fixed effect IV estimates. Further study of additional specifications below suggests the latter is true.

We are also concerned with possible factors in E_{it} (described in Corrective Measures subsection), representing students’ choices made alongside their athletic participation decision that could affect grades. To check these concerns, column “IV FE 5” is identical to “IV FE 4” but includes additional controls for ease of schedule (Corrective Measures subsection provides the formula for ease of schedule), credit hours, military achievement,[29] and company-level fixed effects.[30] Here, point estimates of recruits’ sports participation effect are jointly significant and larger than before but we still cannot be statistically certain of their negative sign.

As mentioned above, student-athletes might attempt to select traditionally easier electives or coordinate their course choices with their athletic seasons. If such decisions are correlated with Z_{it}, then estimates could be biased. “Schedule stacking” is difficult at USNA compared to other Division I institutions due to the extensive core curriculum, so we do not anticipate a major role for this decision.[31] But if the ease metric of the previous models fails to properly capture such processes, course selection would remain an endogenous decision that could enter the mechanism underlying the athletic effect on grades. Columns “IV FE 6” and “IV FE 7” of Table 6 estimate the corrected model using GPA calculated from only core courses (described in Institutional Features of the USNA subsection) and GPA calculated from only professional courses such as navigation, military law, and leadership for naval officers, respectively. They also include all the other controls in E_{it} and P_{it} utilized in this subsection. Our main finding is preserved; sports participation effects for recruits are jointly statistically significant and point estimates are negative but insignificant with sizes 0.36 and 0.094 grade points, respectively.

Table 6. Additional IV Specifications—Testing Exclusion Restriction via Various Fixed Effects and Other Controls.

Model	IV FE 2	IV FE 3	IV FE 4	IV FE 5	IV FE 6	IV FE 7
Dependent Variable	GPA	GPA	GPA	GPA	GPA (Core)	GPA (Pro)
Participated in Sport	1.523	1.622	1.609	.139	0.779	−.276
	(1.917)	(2.069)	(2.031)	(1.010)	(1.373)	(1.488)
× Recruited	−1.561	−1.623	−1.624	−.500	−1.142	−.370
	(1.280)	(1.376)	(1.365)	(0.678)	(0.928)	(1.001)
Mean SAT Score (Math) of Recruited Companymate–Classmates		0.000953	0.00103	.000329	0.000374	−.000268
		(0.000939)	(0.00115)	(0.000620)	(0.000869)	(0.000913)
× Recruited			−0.000314	−.000120	−0.000782	.000130
			(0.000990)	(0.000523)	(0.000790)	(0.000779)
Mean SAT Score (Verbal) of Recruited Companymate–Classmates		−0.000590	−0.000536	−.0000516	−0.0000485	.000237
		(0.000665)	(0.000709)	(0.000401)	(0.000577)	(0.000586)
× Recruited			−0.000170	−.0000245	−0.000195	−.000142
			(0.000513)	(0.000315)	(0.000497)	(0.000465)
Ease of schedule				.612***	0.142***	.250***
				(0.0252)	(0.0347)	(0.0375)
Credit hours				−.00573***	−0.0112***	−.0108***
				(0.000606)	(0.00105)	(0.000933)
Military score				.330***	0.474***	.647***
				(0.0217)	(0.0286)	(0.0318)
Controls for full company effects	N	N	N	Y	Y	Y
Academic and class year controls	Y	Y	Y	Y	Y	Y
$H_0: \gamma_1 = \gamma_2 = 0$ *F*-statistic	23.53	21.99	21.26	29.29	28.53	25.29
Observations	78,948	78,946	78,946	78,143	77,287	77,968

Note. Results from the main model (IV FE 2) are reproduced above for easy comparison. Sample sizes have changed slightly from model to model because military score is missing for a small number of students, and we were unable to compute GPA (core) and GPA (pro) for those student semesters in which no such courses were taken (fewer than 1,000 observations). All models account for individual fixed effects. Standard errors (in parentheses) are clustered by company–class groups. Dummy variables for academic years and class years are included in all models but not shown in the table. FE = fixed effects; GPA = grade point average; IV FE = Instrumental Variable Fixed Effects.

*$p < .05$. **$p < .01$. ***$p < .001$.

While it remains possible that these additional covariates and alternate GPA specifications are vulnerable to other direct peer effects that could stem from Z_{it}, we have found no evidence of such channels in Table 6. Effect estimates do not vary in sign but vary in size, which is most likely a product of the fixed effects IV estimation strategy.

Analysis of the Complier Subpopulation

As a LATE, the participation effect on GPA stems from students whose sports participation responds to variation in the number of recruits in their companies. First-stage estimates strongly suggest that our findings are driven by recruited students. This finding is logical, as unrecruited students simply may not have the athletic skills to compete at the Division I level. Further, if students tend to group themselves based on homophily, their within-company subgroups may be arranged via their athleticism, among other traits. In other words, an athletic student may be more susceptible to the peer influence of athletic friends, compared to a less athletic student.

To investigate the subset of "compliers," we look more closely at the first stage. Table 7 contains first-stage estimates for specific subsamples of *recruited students only*. The peer effect of the number of recruited companymate–classmates on own sports participation is the largest and statistically significant for seniors. A one standard deviation increase in the number of recruited peers (2.6 peers) is associated with a 1.95% greater likelihood of own participation as a senior.[32] The effect is significant at the 10% level for juniors, but small and insignificant for freshmen and sophomores. There is also evidence at the 10% level that students with SAT scores above 1,300 are driving the first stage. We do not detect any statistical differences in response to the instrument for men versus women or by race/ethnicity. It is important to note that insignificance of the instrument here does not indicate the presence of a weak instrument problem. Lack of significance is likely due to the truncated subsample sizes.

Thus, the subpopulation of compliers contains recruits who are upperclassmen with relatively strong SAT scores. If these subgroups do not represent the general student body, then our main estimate could have limited applicability. Table 8 compares summary statistics for recruits and nonrecruits across the four class years. Overall, recruits have weaker academic qualifications upon their arrival at USNA. These discrepancies persist, as there is a GPA gap across all 4 years of college between recruits and nonrecruits. Recruits are more likely to be women and much less likely to have prior military experience. In other categories, recruits and nonrecruits are comparable. Our compliers tend to be recruits with SAT scores over 1,300, suggesting that they may be closer to nonrecruits in this dimension. By junior and senior year, the gaps in GPA and sports participation between recruits and nonrecruits narrow. Overall, subsample averages suggest that the compliers may not comprise a perfect cross section of the student body, but they do not differ substantially across several observable traits.

Table 7. First Stage for Recruits Only—Stratified by Pre-USNA Observables.

	Class Year				Gender	
Dependent Variable: S_{it}	Freshman	Sophomores	Junior	Senior	Men	Women
No. of recruited companymate–classmates	−.0000125	.00187	.00544	.00750*	.00442	.000764
	(.00272)	(.00279)	(.00314)	(.00318)	(.00300)	(.00505)
Background characteristics' controls	Y	Y	Y	Y	Y (via FE)	Y (via FE)
Academic year controls	Y	Y	Y	Y	Y	Y
Class year controls	N	N	N	N	Y	Y
Observations	5,722	5,440	5,062	4,955	16,614	4,565
		Race/ethnicity			SAT Score	
		White	Minority		SAT > 1,300	SAT < 1,300
No. of recruited companymate–classmates		.00436	−.00433		.00719	.000966
		(.00279)	(.00716)		(.00435)	(.00324)
Background characteristics' controls		Y (via FE)	Y (via FE)		Y (via FE)	Y (via FE)
Academic and class year controls		Y	Y		Y	Y
Class year controls		Y	Y		Y	Y
Observations		18,339	2,840		7,577	13,602

Note. Results are from first-stage regressions estimated only on recruits for each of the indicated subsamples. Standard errors (in parentheses) are clustered by company–class groups. Class year stratifications include controls for gender, race/ethnicity, math and verbal SAT score, high school standing, feeder source, and prior enlisted status, instead of individual fixed effects. All other models account for individual fixed effects. Dummy variables for academic years and class years are included in models where indicated but not shown in the table. FE = fixed effects; USNA = U.S. Naval Academy.
*p < .05. **p < .01. ***p < .001.

Table 8. Sample Means Grouped by Class Year and Recruitment Status.

	Freshman		Sophomores Recruits		Juniors		Seniors	
Variable Name	Recruits	Nonrecruits	Recruits	Nonrecruits	Recruits	Nonrecruits	Recruits	Nonrecruits
Year-specific characteristics								
GPA	2.53	2.80	2.72	2.92	2.85	3.00	2.98	3.11
Credit hours	33.8	34.0	36.4	37.2	35.0	35.8	32.2	32.7
Ease of schedule	2.68	2.73	2.84	2.85	2.97	2.97	3.09	3.11
Participated in sport[a]	0.70	0.17	0.66	0.16	0.56	0.13	0.47	0.10
Fixed characteristics								
Female[a]	0.23	0.14						
SAT score (math)	641	670						
SAT score (verbal)	613	646						
High school standing[b]	526	572						
Prior enlisted[a]	0.02	0.11						
Race/ethnicity								
Asian[a]	0.02	0.05						
Black[a]	0.07	0.06						
Hispanic[a]	0.03	0.10						
White or other[a]	0.87	0.78						
Feeder source								
NAPS[a]	0.21	0.16						
Foundation school[a]	0.07	0.06						
Direct to USNA[a]	0.73	0.78						
Observations	5,722	15,274	5,440	14,780	5,062	14,024	4,955	13,691

Note. Table shows sample means for variables within each category.
[a]Sample proportion rather than sample mean (i.e., dummy variable). [b]A standardized version of students' high school academic standing, calculated by the Admissions Office. Refer to "Data" subsection for a description of this variable.

Identification of students marginally attached to their sports team also provides some clues as to the mechanism through which athletes' grades may be negatively affected. As upperclassmen who are relatively strong performers on the SAT, compliers may enjoy more options upon entering their military service. For instance, competitive service communities such as Nuclear Operations and Naval Aviation require higher academic marks from candidates (see Descriptive Analysis of Post-USNA Outcomes subsection). The transition to military service may gain prominence upon entering junior year because students must sign a "two-for-seven" contract that binds them to 2 more years of school and at least 5 years of military service. Thus, "reluctant athletes" during their first 2 years may be more likely to quit upon entering their junior year when the rise in the importance of future priorities (e.g., cumulative GPA) is not sufficiently offset by the collective gains from participation (like team or individual success or the presence of recruited peers). Although this discussion is speculative, our findings are consistent with the idea that separation from one's sports team affords an opportunity to better focus on academics; the complier subpopulation comprises the very students who we would expect to be receptive to such motivations.

A perhaps larger question is—in the event that the LATE does not extend to the general population of students—is the complier subpopulation of interest to higher education policy makers? We suggest the answer is "yes." Table 8 shows that the treatment effect applies to a group of college students who are "normal" in many observable dimensions. They score in the 600s on their math and verbal SATs, are approximately 13% minorities, and most followed the standard path before enrolling in college (e.g., 75% did not attend a feeder program, 2% were previously enlisted). As with the vast majority of NCAA athletes, USNA student-athletes virtually never enter professional athletics after college. These students—who are on the margin of athletic participation—are those who would be impacted by policy changes to college athletics, such as the small financial incentives for athletic participation currently under review by policy makers.

Descriptive Analysis of Post-USNA Outcomes

If athletics do crowd out grades, then policy makers should consider the relative importance of grades and athletics toward various postcollegiate outcomes. For instance, there may be nonacademic benefits to sports participation, such as leadership or teamwork, that could enhance a career in naval service. In this section, we present descriptive analysis of the post-USNA setting to see whether there is evidence hinting at the importance of such "intangibles." Our data provide three items that track USNA graduates' subsequent careers in the Navy or Marine Corps: years of service in the military (either current or upon retirement from active duty), highest rank attained,[33] and service assignment. There are 28 distinct service assignments, which we group into five major communities: (1) Surface Warfare, officers whose primary mission is focused on the operation of Navy ships; (2) Nuclear Operations, officers whose primary mission is focused on the nuclear propulsion of ships and submarines; (3) Naval Aviation, officers whose primary mission is focused on the operation of Navy aircraft; (4) Marine Corps, all officers commissioned in the Marine Corps; and (5) All others, including various smaller communities, such as Navy Intelligence, Supply Corps, and Medical Corps. In this section, we examine the association of academics and varsity athletic participation with these outcomes.

Summary statistics (Table 1) for years of service and rank attainment outcomes are nearly identical for athletes and nonathletes. Service assignments, however, differ modestly; point estimates show that a larger proportion of athletes select into Surface Warfare, Marine Corps, and "other communities," compared to nonathletes.

To see if any patterns persist, conditional on observables, we estimate simple OLS and probit models. Since we observe military career outcomes only for graduates (students who separate from USNA do not enter military service), we estimate these models using a cross section of individual USNA *graduates*. We also must restrict the sample based on timing constraints. For instance, a model with outcome variable "served in the military for at least 5 years" should use the subsample of graduates from the class of 2006 and earlier because we observe their post-USNA variables only through 2012. Likewise, estimation of a model with outcome variable "attained a rank of O-4 or higher" should only include graduates from 2002 and earlier because it is very rare to attain the O-4 rank without at least 10 years of experience.

It is important to reiterate that we are not changing the identification strategy in this section. Here, we present a separate, simple analysis to examine the conditional correlations between academics, athletics, and military outcomes. We estimate such models with the following functional form:

$$\text{Outcome}_i = \alpha\text{GPA}_i + \beta\text{PreCollege}_i + \gamma\text{Sport}_i + \delta\text{Major}_i + \lambda\text{GradYear}_i + u_i. \quad (3)$$

GPA_i is graduate i's cumulative GPA upon graduation. As before, PreCollege_i includes gender, race/ethnicity, SAT scores, high school standing, feeder source, and prior military experience. Sport_i is a binary variable for whether that graduate i participated in a varsity sport for *at least 2 years* while at USNA.[34] Major_i indicates i's academic major types: engineering, mathematics and natural sciences, or humanities and social sciences. GradYear_i represents a set of dummy variables for graduation year that provide the necessary year-specific intercepts for each outcome. For instance, graduates of the year 1994 have a different unconditional probability of attaining O-5 by the end of the sampling window, compared to graduates of the year 1992.

Table 9 displays results from the estimation of Equation 3. There are five panels, each corresponding to a different variable for Outcome_i: length of military service (in years), military service for at least 5 years (binary), service for at least 10 years (binary), attaining at least O-3 rank (binary), attaining at least O-4 (binary), and attaining at least O-5 (binary). For each outcome variable, columns (1) through (4) include different control variable combinations of pre-USNA characteristics and cumulative GPA upon graduation. The table reports coefficients as marginal effects for the probit models. Coefficients for PreCollege_i, Major_i, and GradYear_i are omitted to conserve space, but their estimates are as expected: On average, women serve 1.5 fewer years than men (model "OLS 1," full set of covariates) and are 16.8% less likely to attain O-4 (model "Probit 4," full set of covariates). Science, Technology, Engineering, and Math (STEM) majors remain in the military longer and receive more promotions than humanities and social science majors. Several other characteristics (SAT score, feeder source, prior military experience) are strong predictors of military service but are not significant in all models.

Without conditioning on pre-USNA characteristics and cumulative GPA upon graduation, varsity athletes leave the military sooner and receive fewer promotions (Table 9, column [1]). This estimate subsumes the negative effect of athletics on GPA estimated above. In column (2), we condition on GPA—which is strongly and positively related to military tenure and promotions—and doing so attenuates the effect of athletic participation. Participation's coefficient remains negative but is statistically insignificant for all outcomes (column [2]). In columns (3) and (4), we additionally condition on pre-USNA characteristics. Sports participation is not statistically significant in any specification, but GPA remains positively related to each outcome except "served 10 or more years." GPA effects are modestly sized (e.g., a 1-point increase in GPA is associated with 4 more months of military service or a 7% greater chance of attaining O-5), but stand in contrast to athletic participation effects. If athletics were to exert first-order effects on military careers—via the accumulation of leadership or physical skills—we would expect them to be visible in column (3). Since column (3) omits cumulative GPA, if there is a first-order (positive) effect of athletic participation on career outcomes, then it is washed out by its second-order (negative) effect on grades estimated in Results section. This could explain athletic participation's statistical insignificance in column (3), but even upon controlling for GPA in column (4), its size and significance do not change. These models suggest that academic effects dominate athletic effects with respect to military career outcomes.

The previous models may fail to capture associations between athletics and military outcomes because the skills practiced "on the field" are more useful in certain branches of service. For example, teamwork could be more vital to the career development of an aviator than a submariner. To test this hypothesis, we estimate versions of Equation 3 that include interactions of five service selection options with GPA and varsity athletic participation:

$$\begin{aligned}\text{Outcome}_i = {} & \alpha\text{GPA}_i + \beta\text{PreCollege}_i + \gamma\text{Sport}_i + \delta\text{Major}_\text{i} + \lambda\text{GradYear}_i \\ & + \psi_1\text{SS}_i + \psi_2(\text{SS}_i \times \text{GPA}_i) + \psi_3(\text{SS}_i \times \text{Sport}_i) + \text{v}_i.\end{aligned} \tag{4}$$

SS_i represents a set of dummy variables for the five service selection categories. Table 10 presents estimation results for Equation 4. For brevity, we only show coefficient estimates of the ψ's; other covariates have similar associations to those described in the previous models. Estimates from interaction models are not as

Table 9. Military Experience and Related Outcomes.

Model	(1)	(2)	(3)	(4)
Control variable sets for each model				
Pre-USNA characteristic controls	N	N	Y	Y
Graduation year controls	Y	Y	Y	Y
Model OLS 1, Dependent variable: Years of military experience ($N = 14{,}854$)				
Varsity participation for 2+ years	−.203*	−.144	.0556	.0557
	(.0877)	(.0800)	(.0706)	(.0697)
Cumulative GPA (upon graduation)		.474***		.326***
		(.0919)		(.0687)
Model Probit 1, Dependent variable: Served 5+ years ($N = 14{,}854$)				
Varsity participation for 2+ years	−.0158	−.00996	.00748	.00754
	(.00875)	(.00853)	(.00799)	(.00800)
Cumulative GPA (upon graduation)		.0440***		.0355***
		(.00919)		(.0107)
Model Probit 2, Dependent variable: Served 10+ years ($N = 10{,}032$)				
Varsity participation for 2+ years	−.0223	−.0200	.00515	.00516
	(.0125)	(.0118)	(.0120)	(.0120)
Cumulative GPA (upon graduation)		.0187**		−.00273
		(.00690)		(.00615)
Model Probit 3, Dependent variable: Attained at least O-3 rank ($N = 16{,}890$)				
Varsity participation for 2+ years	−.00685	−.00226	.00195	.00175
	(.00591)	(.00563)	(.00551)	(.00550)
Cumulative GPA (upon graduation)		.0354***		.0338***
		(.00357)		(.00433)
Model Probit 4, Dependent variable: Attained at least O-4 rank ($N = 10{,}997$)				
Varsity participation for 2+ years	−.0196	−.0111	.00916	.00896
	(.0118)	(.0119)	(.0124)	(.0122)
Cumulative GPA (upon graduation)		.0698***		.0549**
		(.0177)		(.0185)
Model Probit 5, Dependent variable: Attained at least O-5 rank ($N = 5{,}579$)				
Varsity participation for 2+ years	−.0152	−.00224	.0178	.0182
	(.0113)	(.0110)	(.0109)	(.0113)
Cumulative GPA (upon graduation)		.0912***		.0736***
		(.0119)		(.0184)

Note. Reported coefficients for probit models are marginal effects. Pre-USNA characteristic controls include gender, race/ethnicity, SAT, high school standing, feeder source, and prior enlisted status. Standard errors (in parentheses) are clustered by graduation year. The models OLS 1 and Probit 1 are estimated on the subsample who graduated before 2007. Probit 2 uses the subsample who graduated before 2002. Probit 3, Probit 4, and Probit 5 use subsamples who graduated before 2009, 2003, 1997 to account for mandatory minimum tenure requirements for promotion. USNA = U.S. Naval Academy; GPA = grade point average; OLS = ordinary least squares.
*$p < .05$. **$p < .01$. ***$p < .001$.

accurate, likely due to increased dimensionality and collinearity, but we can still infer some relationships. GPA is positive and significant at the 1% level in each service community for at least one outcome. It is most important as a predictor of career success in naval aviation and "other" service assignments. For instance, a 1-point increase in GPA for an aviator is associated with a 12.5% greater likelihood of attaining O-4. Athletic participation is significant at the 1% level in only one category: as a predictor of 10 or more years of service in the Marine Corps (we also find that athletics are correlated with career development in aviation, at the 5% level).

In summary, these models show strong evidence of a positive relationship between academics and military career success but only limited evidence that athletics can be advantageous in certain job tracks. Results for service communities such as surface warfare and nuclear operations reveal no first-order connection between athletics and graduates' military careers, while the IV models of Results section suggest a second-order crowding-out effect of athletics on career success (via its impact on grades). In light of these findings, it is natural to question the value added—to the individual student—of intercollegiate athletics. We should emphasize that the models in this section are not structural in scope. They may be plagued by selection because we do not observe graduates after their exit from the military. Athletes may have gained skills "on the field" in unobserved dimensions that yield higher wages in the private sector, compelling them to depart military service earlier. One could also question the value added—to the Navy—of intercollegiate athletics. While we find very limited evidence of a positive link between athletic participation and post-USNA military outcomes, we should be wary of the notion of "military success" that is captured by our outcome variables. We do not directly observe, for example, whether athletes may become better commanders, earn more commendations, or experience fewer conduct violations than nonathletes. We cannot be certain that these important traits are captured by rank attainment and length of military career, which are coarse measures of success, at best.

Discussion and Conclusion

Discussion of External Validity

While the unique environment of USNA benefits our identification strategy, it may present limitations in terms of applicability to other Division I university settings. For example, as discussed in Institutional Features of the USNA subsection, all nonvarsity athletes at USNA must still partake in organized athletics, via club or intramural sports. These nonvarsity sports programs may simply substitute for more robust menus of extracurricular activities at other universities. But if not, these programs may narrow the gap between athletes and nonathletes, in which case our estimates of the negative effect of athletics on academics may be understated compared to other universities.

Table 10. Military Experience and Related Outcomes With Interaction Terms.

	Int. OLS 1	Int. Probit 1	Int. Probit 2	Int. Probit 3	Int. Probit 4	Int. Probit 5
Dependent Variable	Military Experience	5+ Years	10+ Years	Made O-3	Made O-4	Made O-5
Service selection (ref.: Surface warfare operations)						
Nuclear operations	−0.944	−.491***	.0154	−.0265	.134	.0810
	(0.823)	(.113)	(.121)	(.0725)	(.110)	(.132)
Naval aviation	0.332	−.186**	−.0735	−.0166	−.169	−.0900*
	(0.684)	(.0635)	(.110)	(.0353)	(.117)	(.0412)
Marine corps	0.504	.0209	−.0148	−.0574	.0161	−.302***
	(0.648)	(.0554)	(.123)	(.0680)	(.128)	(.0453)
Other	−4.130**	−.239*	−.413***	.0357	−.344***	−.301***
	(1.231)	(.0991)	(.0610)	(.0213)	(.0852)	(.0663)
GPA (ref.: Surface warfare operations)4	−0.0675	−.0110	−.0508	.0260**	.00437	−.0122
	(0.207)	(.0128)	(.0325)	(.00938)	(.0365)	(.0271)
× Nuclear Operations	0.156	.104***	−.0364	.00915	−.0719	−.0365
	(0.266)	(.0259)	(.0433)	(.0175)	(.0389)	(.0395)
× Naval Aviation	0.521	.104***	.0836	.0164	.125**	.0747***
	(0.298)	(.0192)	(.0432)	(.0140)	(.0441)	(.0191)
× Marine Corps	0.268	.0165	.0376	.0350	.0162	.153**
	(0.279)	(.0234)	(.0495)	(.0185)	(.0566)	(.0493)
× Other	1.549**	.0680**	.209***	−.0205	.175***	.179**
	(0.455)	(.0223)	(.0463)	(.0136)	(.0494)	(.0580)

(continued)

Table 10. (continued)

	Int. OLS 1	Int. Probit 1	Int. Probit 2	Int. Probit 3	Int. Probit 4	Int. Probit 5
Dependent Variable	Military Experience	5+ Years	10+ Years	Made O-3	Made O-4	Made O-5
Participation (ref.: Surface warfare operations)	−0.135	−.00103	−.0310	.000165	−.0155	−.0202
	(0.158)	(.0123)	(.0191)	(.00594)	(.0219)	(.0386)
× Nuclear Operations	0.166	.00953	.0383	−.0142	.00894	.0802
	(0.255)	(.0188)	(.0474)	(.00897)	(.0388)	(.0515)
× Naval Aviation	0.341*	.0394*	.0347	.00723	.0273	.0526
	(0.148)	(.0174)	(.0243)	(.00932)	(.0250)	(.0362)
× Marine Corps	0.442	.00833	.0836**	−.0108	.0627	.0905
	(0.279)	(.0243)	(.0291)	(.0145)	(.0334)	(.0819)
× Other	−0.0617	−.0155	.0223	.0240*	.00302	−.00451
	(0.381)	(.0239)	(.0492)	(.0109)	(.0491)	(.0567)
Observations	14,854	14,850	10,032	16,886	10,997	5,579

Note. Reported coefficients for probit models are marginal effects. Table does not show coefficient estimates for gender, race, SAT, high school standing, feeder source, prior enlisted status, academic major, and graduation year. Standard errors are clustered by graduation year. GPA = grade point average; OLS = ordinary least squares.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Other USNA features may widen the gap between athletes and nonathletes, relative to other universities. As discussed in Institutional Features of the USNA subsection, all USNA students must graduate in 4 years, regardless of their athletic status, and students must remain continuously enrolled with a minimum of 15 credit hours per semester. Combined with the large set of core courses that are mandatory for all students, athletes have limited means to lighten their schedules while in season. At other schools, it is common practice to "redshirt" a year of school,[35] or simply to delay graduation beyond 4 years, while spreading the academic burden over more years of attendance. In these ways, athletic participation at USNA is more time restrictive than at other schools, and our estimates may thus be overstated. On the other hand, the extra structure imposed on USNA student-athletes' daily schedules could *benefit* them, compared to students at traditional universities who lack mandatory evening study periods, strict meal times, required classroom attendance, and so forth.

Nevertheless, the athletic experience at USNA is similar in many ways to other schools. As at other well-funded Division I programs, varsity athletes receive complimentary gear, apparel, medical services, and nutritional supplements that are not available to other students. Although USNA athletes do not receive as significant favorable academic treatment, they do enjoy some perks related to their military responsibilities. When in season, varsity athletes may skip weekly drill practices and parades. They may eat meals during off-hours with their teammates at "team tables," while nonathletes must attend all three meals per day with their companymates. These perks provide time-savings that bring USNA athletes' experiences closer to those at other universities.

Ideally, we would echo Carrell, Hoekstra, and West (2011)—who address similar issues in their study of peer effects on fitness at the U.S. Air Force Academy—to say that our evidence of an athletic effect on academics at USNA suggests that it may also exist in the broader population of universities. But athletic effects differ from peer effects, and given the discussion above, the mixture of issues affecting the external applicability of our findings prevents assignment of a direction to a potential bias.

Our empirical model approximates a randomized control trial in a setting where the treatment group includes students who are "marginally attached" athletes who separate from their sport. Once a treated athlete quits, his or her experience mimics that of nonvarsity students; thus, the control group includes marginally attached athletes who do not quit. Any measurable effect on grades stems from how students behave differently upon separation from their sports team (one can also imagine the opposite experiment, in which students who are marginally *un*attached *join* a varsity team, but this is much rare). As discussed above, many contextual student experiences are naturally held constant across athletes and nonathletes at USNA, so which behaviors might change upon quitting one's team? A student might study more, lose access to athlete-only academic resources, interact differently with peers, or experience newfound favoritism or discrimination from instructors. While we cannot disentangle these effects from one another—our estimates reveal only the net effect of sports participation—Analysis of the Complier Subpopulation subsection presents circumstantial evidence that marginally attached athletes who quit are those more likely to reprioritize their studies. With the caveat that this discussion is speculative, if this is truly the dominant underlying mechanism, then we would expect the same direction of the treatment effect to prevail at other universities, as we can easily imagine that binding time constraints between academics and athletics exist in more traditional settings.

Conclusion

For many students, intercollegiate athletics are an important part of their college experiences. Some previous work has studied the role of sports participation within secondary and higher education systems, but this article is the first attempt to directly estimate the causal impact of athletics on academic achievement in college. At the USNA, student-athletes' grades trail nonathletes' by a small margin, and this gap nearly vanishes upon controlling for student-specific (background) characteristics. Athletic participation is endogenous, however, because of selection (athletes are not randomly selected from within the student body), omitted variables (athletes may possess unobserved traits correlated with their academics), and simultaneity (academic outcomes may affect students' decisions to participate in sports). To identify causal effects, we exploit the Naval Academy's random assignment of each student into a "company" that forms his or her primary peer group. We define an IV for athletic participation as the number of recruited athletes within a student's company–class cohort. We argue that, conditional on observable controls and fixed effects, validity stems from the random assignment, as students cannot self-select into these peer groups, and identification stems from a peer effect compelling athletically inclined students to compete on varsity teams.

IV estimates reveal that sports participation affects recruited student-athletes' GPAs, but we cannot be certain whether the effect is small or as large as 0.37 grade points on average. We are not certain of the effect's sign, but point estimates suggest that the effect for recruits is negative. Additional analysis shows that the "compliers" driving our results are students who were recruited, who are juniors and seniors, and who were relatively strong performers on the SAT. This subset of students appears comparable to the greater student body. Analysis of postgraduation outcomes, such as military experience and promotions, suggests that, while higher GPA predicts more military career success, athletics may only benefit certain career tracks, such as the Marine Corps. Although our findings show no clear benefit of athletic participation to the individual or to the Navy, we must bear in mind that the observable post-USNA outcomes do not reveal the full story. We would be better equipped to analyze this important question with richer data on graduates' post-military careers. It may be insightful to analyze such skills as leadership and teamwork that could be boosted by sports participation. It is also natural to question whether the effect may differ between types of athletes, specific sports, skill levels, or athletic achievements. Does participation in basketball incur the same effect as participation in water polo? Future work could exploit exogenous variation in competitive seasons across sports and years to estimate both heterogeneous effects for different sports as well as in-season versus out-of-season effects on academic achievement. Despite these caveats, our main takeaway is that, for the average student, there are not complementarities, but trade-offs, between academic achievement and athletic participation.

Authors' Note

We are solely responsible for any errors.

Acknowledgment

We are very grateful to Cathy McGuire and Glenn Gottschalk of the USNA Office of Institutional Research for their efforts in compiling the data for us, as well as to Rob Slye for excellent research assistance. We also thank Jonah Rockoff, Ahmed Rahman, three anonymous referees, and participants at a Vassar College Economics Department seminar for their helpful suggestions.

Declaration of Conflicting Interests

The author(s) declared no potential conflict of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. National Collegiate Athletic Association (NCAA) Press Release, March 13, 2007.
2. National Labor Relations Board, Region 13, Case 13-RC-121359.
3. U.S. District Court, N.D. Cal., Case No. 09-CV-3329.
4. NCAA, May 13, 2009. Defining Countable Athletically Related Activities (http://www.ncaa.org/sites/default/files/Charts.pdf).
5. For instance, to be eligible to compete, the NCAA mandates that Division I student-athletes "must achieve 90 percent of the institution's minimum overall grade-point average necessary to graduate (for example, 1.8) by the beginning of year two, 95 percent of the minimum GPA (1.9) by year three and 100 percent (2.0) by year four" (http://www.ncaa.org/remaining-eligible-academics).
6. The assignment process, which is not purely random, is described in detail in Institutional Features of the U.S. Naval Academy (USNA) subsection.
7. The performance gap observed at Clemson is 0.3 grade points, which is nearly twice as large as the average gap in our USNA data.

8. This variable is calculated by the Admissions Office, which provided the following description: “Since not every high school ranks students, some applicants are home schooled, and some schools use different grading scales, this column standardizes all applicants to a common scale reflecting their rank in their high school class.” As per their internal policy, the Admissions Office could not release more information on how this variable is calculated. We do observe “pure” high school class rank (and high school class size) for 85% of our sample. Among this subsample, the size of the correlation between “standardized” high school class rank and “regular” class rank divided by class size is 0.86.
9. Some students attend a 1-year college preparatory program. The largest such feeder is the Naval Academy Preparatory School in Newport, RI, followed by various “foundation schools” that maintain preparatory programs affiliated with USNA. A complete list of the foundation schools may be found online: http://www.usna.com/page.aspx?pid=282
10. There are age restrictions for students. Currently, they must be at least 17 years of age and must not have passed their 23rd birthday on July 1 of their year of admission. Prior to the class of 2001, students could not have passed their 22nd birthday on July 1 of their year of admission.
11. See USNA Viewbook (http://www.usna.edu/Viewbook/life.php).
12. Shotguns occurred for the classes of 1992-1996 following their respective sophomore years. In 1996, USNA downsized the number of companies from 36 to 30. All the midshipmen from six randomly chosen companies were dispersed throughout the rest of the companies in their respective battalion. This process affected the classes of 1997-1999 (the class of 1998 experienced an additional shotgun after their freshman year, prior to the downsizing). The classes of 2000-2005 experienced shotguns following their freshmen years. Shotguns ceased thereafter.
13. “Love chits” result when two students within the same company desire to enter a romantic relationship. USNA regulations require one of them to transfer to another company. “Hate chits” occur when a student requests a transfer to another company based on an event or bias that may have occurred in his or her current company. Based on our data, such chits occur approximately 5–10 times per year.
14. If anything, p values tend to be *larger* than what we would expect under pure random assignment. Thus, although unreported to us by the Admissions Office, recruits in some years may have been (deliberately) evenly spread about the companies (several p values are very close to one). While still consistent with random assignment, low variation in the number of recruits across companies would harm the power of our instrument but would not compromise the exclusion restriction (described in Econometric Model section). Ultimately, we find that there is sufficient variation in the number of recruits across companies to estimate an effect via instrumental variables (IVs) estimation.
15. From the NCAA Division I Manual Section 15.3.5.1.
16. If a student expressed interest in the majoring in a critical language (i.e., Chinese or Arabic), then the student would take courses in that language starting freshman year and, therefore, postpone a few mandatory classes until their sophomore or junior year.

Students may also validate one or more core classes, which would lead them to enroll in higher level classes during their freshman year. For example, a student with sufficient prior experience in calculus could enroll in Calculus III during freshman year while their peers enroll in Calculus I and II.

17. Possible exceptions include conduct or honor offenses, in which case the offending student is typically subject to delayed graduation by one semester. Other exceptions are religious missions, in which students leave the Academy for their mission (up to 2 years) and then assimilate into new graduating classes upon returning, thus delaying graduation. There are fewer than 20 such exceptions per year.
18. The Physical Readiness Test consists of curl ups, push-ups, and a 1.5 mile run (or 500 yard swim). The minimum standards for men are 65 curl ups, 45 push-ups, and a time under 10 min and 30 s on the run (or 11 min and 20 s on the swim). The minimum standards for women are 65 curl ups, 20 push-ups, and a time under 12 min and 40 s on the run (or 11 min and 20 s on the swim).
19. The need for this control is motivated by literature on peer effects in higher education, such as the work of Carrell et al. (2009); it is discussed further in Corrective Measures subsection.
20. Here, "athletes" are students who participated in a varsity sport for at least 1 year.
21. Very few students depart once their junior year has commenced because those who stay beyond sophomore year sign a "two-for-seven" contract, which binds them to a minimum of seven more years of military commitment (two more as a student, five as an officer). Those who sign but later drop out of USNA must repay the cost of their education to the government. This provides a strong incentive to complete the degree.
22. One could also suggest that identification may stem, for instance, from a larger group of recruits building a "culture of athleticism" within their company. In this sense, the *ratio* of a student's recruited peers is also a viable instrumentation for Z_{it}.
23. Here, we follow the work of Carrell et al. (2009) in accounting for academic peer effects.
24. Military achievement, known as Military Quality Point Rating, incorporates measures of military performance, conduct, and physical fitness. Further details are available on the USNA intranet, or upon request.
25. We define a variable Ease_{it} that is based on a formula used by Maloney and McCormick (1993) and Robst and Keil (2000). The following metric captures the "ease" of student i's course schedule by calculating the average grade, weighted by credit hours, received by all *other* students across all courses taken by student i: $\text{Ease}_{it} = \frac{\sum_{j=1}^{N}\left(\text{Credits}_j \times \sum_{k \neq i}^{L} \frac{\text{Grade}_{kj}}{L}\right)}{\sum_{j=1}^{N} \text{Credits}_j}$, where $N =$ the number of courses taken by student i in year t; $L =$ the number of students other than student i taking course k in year t; $\text{Credits}_j =$ the number of credit hours for course j; $\text{Grade}_{kj} =$ the grade earned by student k in course j. This variable is a noisy measure of the true "ease" of student i's course schedule, because it depends on his or her

draws of both classmate quality and instructor "grading difficulty." Since USNA students have very limited ability to self-select into specific course sections, both their classmate and instructor draws should be effectively random.

26. Anecdotal evidence suggests that companies may maintain their own unique "culture" that persists over the years. If so, time-invariant company fixed effects would help account for problematic social peer effect channels.
27. As mentioned in a footnote in Corrective Measures subsection, we can also estimate the IV model using the ratio of peer recruits within a student's company–class cohort. We do not present these results here; coefficient estimates do not significantly change across the main models, but with ratios, the first stage is slightly weaker in the preferred model described below. The number of peer recruits may provide a stronger first stage because—regardless of company size—a student tends to interact with his or her entire company. The marginal effect of an additional athletic peer may affect a student's propensity to participate in a varsity sport more strongly than a marginal increase in the "intensity of company athleticism," which is better captured by the ratio.
28. The reader may wish to skip ahead to Table 7 to verify that recruits tend to be academically weaker.
29. There is a possibility that recruited companymates' peer effects promote a greater (or lesser) devotion to one's military duties, thereby crowding out own grades through a social peer effect.
30. These account for additional channels that may operate through company-persistent effects (e.g., a culture of partying).
31. The reader may skip ahead to Table 7 to see that the "ease of schedule" metric is not substantially different between recruits and nonrecruits. If anything, point estimates suggest that recruits' courses assign, on average, slightly lower grades.
32. The magnitude of 1.95% is not directly comparable to the 3.6% effect size in the main model (Corrected Estimates subsection) because, as mentioned above, the current model is estimated on recruited students only.
33. O-1 is an Ensign in the Navy or a Second Lieutenant in the Marine Corps. O-2 is a Lieutenant (junior grade) in the Navy or a First Lieutenant in the Marine Corps. O-3 is a Lieutenant in the Navy or a Captain in the Marine Corps. O-4 is a Lieutenant Commander in the Navy or a Major in the Marine Corps. O-5 is a Commander in the Navy or a Lieutenant Colonel in the Marine Corps. O-6 is a Captain in the Navy or a Colonel in the Marine Corps. None of the graduates in our sample have been in the armed forces long enough to have been promoted to O-7.
34. Estimation results are qualitatively similar if we redefine $Sport_i$ to indicate participation for "at least 1 year" or "at least 3 years."
35. NCAA guidelines permit student-athletes to "redshirt" up to 2 years, in order to spread their 4 years of athletic eligibility over a longer period. In a "redshirt" year, student-athletes may not compete in their team's competitions (thus their year of athletic eligibility is not consumed), but they may enroll in courses at the university and practice with their team.

References

Angrist, J., Imbens, G., & Rubin, D. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*, 444–455.

Barron, J., Ewing, B., & Waddell, G. (2000). The effects of high school athletic participation on education and labor market outcomes. *Review of Economics and Statistics*, *82*, 409–421.

Carrell, S., Fullerton, R., & West, J. (2009). Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labor Economics*, *27*, 439–464.

Carrell, S., Hoekstra, M., & West, J. (2011). Is poor fitness contagious? Evidence from randomly assigned friends. *Journal of Public Economics*, *95*, 657–663.

Eckard, E. W. (2010). NCAA athlete graduation rates: Less than meets the eye. *Journal of Sport Management*, *24*, 45–58.

Eide, E., & Ronan, N. (2001). Is participation in high school athletics an investment or a consumption good? Evidence from high school and beyond. *Economics of Education Review*, *20*, 431–442.

Feld, J., & Zölitz, U. (2014). *Understanding peer effects: On the nature, estimation and channels of peer effects* (Working Papers in Economics, No. 596). Gothenburg, Sweden: University of Gothenburg.

Gemus, J. (2010). *College achievement and earnings* (Working Paper 2010:1). Uppsala, Sweden: Department of Economics, Uppsala University.

Jones, E., & Jackson, J. (1990). College grades and labor market rewards. *Journal of Human Resources*, *25*, 253–266.

Kuhn, P., & Weinberger, C. (2005). Leadership skills and wages. *Journal of Labor Economics*, *23*, 395–436.

Lechner, M. (2009). Long-run labour market and health effects of individual sports activities. *Journal of Health Economics*, *28*, 839–854.

Long, J., & Caudill, S. (1991). The impact of participation in intercollegiate athletics on income and graduation. *Review of Economics and Statistics*, *73*, 525–531.

Maloney, M., & McCormick, R. (1993). An examination of the role that intercollegiate athletic participation plays in academic achievement athletes' feats in the classroom. *Journal of Human Resources*, *28*, 555–570.

NCAA Research (2013). *Trends in Graduation-Success Rates and Federal Graduation Rates at NCAA Division Institutions*. NCAA Research Staff.

Olbrecht, A. (2009). Do academically deficient scholarship athletes earn higher wages subsequent to graduation? *Economics of Education Review*, *28*, 611–619.

Pfeifer, C., & Cornelissen, T. (2010). The impact of participation in sports on educational attainment—new evidence from Germany. *Economics of Education Review*, *29*, 94–103.

Robst, J., & Keil, J. (2000). The relationship between athletic participation and academic performance: Evidence from NCAA Division III. *Applied Economics*, *32*, 547–558.

Stevenson, B. (2010). Beyond the classroom: Using title IX to measure the return to high school sports. *Review of Economics and Statistics*, *92*, 284–301.

Wise, D. (1975). Academic achievement and job performance. *American Economic Review*, *65*, 350–366.

Author Biographies

Michael A. Insler is an associate professor of Economics at the US Naval Academy. His research focuses on the economics of education, labor, and health. A lifelong sports fan, he is also a "locally competitive" distance runner in the Annapolis, Maryland area.

Jimmy Karam is a retired naval officer who served at the US Naval Academy as the Economics Department associate chair. At the Naval Academy he taught Sports Economics and participated as a mentor and coach on the Varsity Sprint Football and Baseball teams. He currently resides in Poway, CA where he coaches high school football.