Athletic Participation and Educational Outcomes

Brooke Coneeny, Sydney Levy, Ellie Miller, Ally Scheve

December 14, 2021

Econometrics

Professor Bhanot

Abstract

Athletics participation has been a central part to many high school student's educational and social experiences. However, it is unclear if playing sports in high school is actually beneficial to later educational outcomes. Using data from the Education Longitudinal Study collected by the U.S. Department of Education, we investigate the association between athletics participation and educational outcomes such as college attendance and change in math scores. Using a cross-sectional analysis we find that sustained athletics participation is associated with an increased probability of attending college.

Introduction and Background

Everyone wants to put themselves in the best position to succeed in life. Whether that be hiring private coaches, exercising and maintaining a balanced diet, networking, meditating, or staying off of social media, people are constantly partaking in activities that they believe will give them an extra advantage. This is especially true of parents who want to set their children up for success later in life. Should their child join the band or the soccer team? Is attending acting camp or coding camp more beneficial? Should they just sign up for everything? What about hiring a private tutor? These are questions that every concerned parent is asking themselves. These questions are particularly relevant when considering what extracurricular activities, if any, a high school student should participate in. Extracurriculars provide many social and well being benefits (Buckley and Lee, 2018). They are a great way to meet new people and make friends. Many activities such as sports or theatre also provide opportunities for movement and exercise, and they can help students find areas outside of the classroom at which they excel (Winstone et al., 2020, Pérez-Ordás et al., 2019). Furthermore, athletics may level the effects of economic privilege by providing intangible life skills and additional educational opportunities. These are all important considerations, but is there an actual association between participating in extracurricular activities and immediate academic achievement as well as long term collegiate success?

Our research will evaluate this question using data collected through the Educational Longitudinal Study conducted by the US Department of Education. This is a nationally representative sample that was administered to over 15,000 high school students. The data was initially collected in 2002 (the students' sophomore year of highschool), with follow-up surveys conducted in 2004 and 2006. The factors we use to evaluate academic success include the change in a student's math quartile score between their sophomore and senior year of high school and whether or not the student attended college. We then examine if participation in extracurricular activities - athletics and band - has an association to our outcome variables of interest while including demographic controls. We also look to see if continued participation in athletics - remaining on a sports team from sophomore through senior year - has any additional impact on a student's academic performance as compared to students who never participated in athletics, quit after their sophomore year, or only began playing their senior year of high school. Lastly, we determine if athletics has a mediating effect on economic privileges that are associated with a higher probability of attending college.

Our findings suggest that there is a positive relationship between being an athlete and academic achievement. More specifically, sustained engagement in athletics throughout high school appears to be an important factor in college attendance. Athletics participation may facilitate an equalizing effect between those who are economically privileged and those who are not. There are meaningful implications that can be taken from these findings: as a result of the current COVID-19 pandemic, many U.S. schools completely eliminated or severely paired down their extracurricular activities (US News). Now that the world is returning to a sense of normalcy, these activities are starting to come back. The conversation now becomes how to most effectively handle the re-introduction of extracurriculars. Important decisions need to be made about the allocation of the public school systems' limited resources. On an individual level, families must decide which activities to re-enroll their children in. As the world is coming out of the COVID-19 fog, the relationship between extracurricular activities, specifically athletics, and academic performance is as important as ever.

Literature Review

There has been a significant amount of research trying to determine the relationship between participation in extracurricular activities and student academic performance. Specifically, there have been many efforts to estimate the individual academic returns of involvement in these activities. The results from these studies have been mixed. For example, in a 2008 academic performance comparison between athletes and non-athletes in the state of Kansas, Lumpkin and Favor (2008) report that high school athletes earned higher grades, had a higher graduation rate, and dropped out of school less frequently than individuals who did not participate in high school athletics. The positive association between participation in high school athletics and academic performance is supported in the research conducted by Dyer et al. (2017), Wretman (2017), Fox et al. (2010), and Chen et al. (2021). However, in a National Longitudinal Study of Adolescent Health, Rees and Sabia (2010) found only limited evidence that participation in sports was correlated with higher math and English GPA amongst high school students. Rees and Sabia (2010) cite the need to account for other unobservable characteristics possessed by student athletes that could explain academic performance. This claim is supported by research conducted by Eccles et al. (2003), Eccles and Barber (1999), and Eide and Ronan (2001). In addition to unobservable characteristics, an individual's socio-economic status must be taken into account when evaluating the relationship between athletic participation and academic achievement (Jerrim et al., 2020).

Extensive research has also taken place at the collegiate level. In a 2008 paper titled, *In-Season vs. Out-of-Season Academic Performance of College Student-Athletes,* Scott et al. (2008) find that collegiate student athletes across all three NCAA athletic divisions perform better academically when they are out of season for their sport, adding to the evidence that athletic participation does not facilitate improved academic outcomes. Likewise, Aries et al. (2004) found similar results when examining the academic performance among student athletes at highly selective liberal arts colleges. This is a contrast to the results attained by Muñoz-Bullon et al. (2017) which indicated that amongst students at a university in Spain, those who participated in formal athletic activities performed better academically than their peers who did not. Robst and Keil (2000) also found results that indicate student-athletes perform better academically than non-athletes at certain NCAA Division III schools. Thus, evidence remains mixed as to whether collegiate athletes show improved academic outcomes as compared to non-athletes.

However, this inquiry is not exclusive to athletics. In a 2010 study, Knifsend and Graham (2010) examine how the breadth of extracurricular participation influences both a students' academic performance and their sense of belonging at school. They report that engagement across two domains (academic/leadership groups, arts activities, clubs, and sports) leads to the greatest sense of belonging at school amongst 11th and 12th graders, a higher grade point average in 11th grade, and greater academic engagement in 12th grade, relative to those who were more or less involved across domains. There has also been research into how extracurricular involvement can help diminish the gap in academic achievement between individuals of different social identities. In a study conducted using data from the National Educational Longitudinal Study of 2002 (the same data we use), researcher David Morris (2015) reported that less advantaged high school students see greater improvement in academic performance from participation in organized activities than their more advantaged peers. He concludes that participation in organized activities can be a form of resource compensation that helps reduce the achievement gap between students of different socioeconomic classes.

Likewise, in a 2019 study conducted by Billingsley and Hurd (2019), the results indicate that engagement in extracurriculars amongst under-represented college students could be an effective strategy for facilitating academic success. This engagement could counter the negative psychological outcomes of discrimination faced by minority students. Finally, there exists data that suggests that participation in extracurricular activities can be beneficial when transitioning from an academic to professional environment. In a Portuguese experiment, researchers Pinto and Ramalheira (2017) distributed the resumes of business students to various professionals who were asked to rank the employability of the students based on their resumes. Results showed that among individuals with comparable academic performances, those who were involved in extracurricular activities were considered more employable than those who were not.

While these aforementioned reports contribute to literature evaluating the cross-sectional effect of participation in extracurriculars on academic performance, there has been less investigation into the effect of extended participation in extracurricular activities in high school on long term academic success. As a result, our study will contribute to the existing literature in the field by expanding on how high school athletics participation impacts later educational outcomes such as college attendance. One example of this line of research is a study conducted by Mahoney, Cairns, and Farmer (2003). The authors found that consistent involvement in extracurriculars during middle and high school was positively associated with college attendance. An additional instance of researchers examining the impact of extracurricular engagement in high school on college success comes from a study also completed using the National Educational Longitudinal Study. Gardner et al. (2008) found that an individual's likelihood of getting into college was positively associated with an increase in duration and intensity of their extracurricular activity participation. Our research is able to build on this as we look into if there are differing levels of achievement associated with specific extracurriculars. The relationship between participation in extracurriculars, specifically athletics, and academic performance is a highly sought after question. This is evidenced through the numerous studies that have been conducted on the topic, and our research will hope to further contribute to the field.

Data and Methodology

The data used in our research is from the Educational Longitudinal Study by the US Department of Education. The Educational Longitudinal Study is a longitudinal panel study established in 2002 that examines the critical transitions experienced by students as they proceed through high school and into postsecondary education or their careers (US Department of Education 2005). The initial survey was administered to 15,000 high school sophomores in the spring of 2002. The nationally representative sample was constructed by first selecting 750 U.S. schools and then randomly sampling tenth-graders within each school. Catholic and other private schools were sampled at a higher rate to ensure that the sample is large enough to allow for a comparison with public schools. Similarly, Asian students were sampled at a higher rate than White, Black, and Hispanic students so that they could be compared to these groups. Study participants were first surveyed in 2002, which was most participants' sophomore year of high school. Students were then surveyed in 2004, which for most students was their senior year of high school. Finally, students were surveyed in 2006 in order to record post high school outcomes. We use the Educational Longitudinal Survey because it offers unique information on student's later educational successes as well as detailed information about their high school activities, namely sports. Our analyses include respondents who participated in the 2002, 2004, and 2006 surveys.

We employed two measures of academic achievement in this study. Our first dependent variable, *Attend College*, is a binary variable equal to one if the respondent enrolled in a postsecondary institution and 0 otherwise. The second dependent variable, *Change in Math Quartile*, measures the change in math quartile score between the respondents sophomore and senior year of high school. Ideally, we would have also included changes in reading / composite quartiles however this data was not available in the second followup questionnaire (2004). Further, data such as GPA, SAT scores, and the number of advanced placement courses the student was enrolled in were removed from the public dataset to ensure participant privacy. The table below reports the summary statistics for our outcome variables.

	N	Mean	St. Dev.	Min	Max
Attended College	14039	0.750	0.433	0	1
Change in Math Quartile	13394	-0.095	0.616	-3	3

The central independent variable in our study, *Athlete*, records if the respondent participated in high school athletics in the student's sophomore and/or senior year. If the student reported playing at least one sport in a given year, they are considered an athlete for that year in our dataset. *Athlete* is equal to one if a student participated in a sport in either the 2002 or 2004 survey. We also measured athletic participation in four more granular categories: *Athlete Sophomore and Senior Year*, *Athlete Only Sophomore Year*, *Athlete Only Senior Year*, and *Never An Athlete*. We utilized these more specific categories in order to determine whether continued engagement in a sport, quitting a sport, or beginning a sport later in high school had a different impact on educational outcomes.

In addition to our athletic participation measure, we also investigate similar patterns in a different extracurricular activity. *Band Participant* is equal to one if the respondent participated in band either sophomore or senior year of high school. Similarly to athletics participation, we measure band participation in four more granular categories: *Band Sophomore and Senior Year, Band Only Sophomore Year, Band Only Senior Year, and Never In Band*. The table below reports the summary statistics for the key independent variables relating to participation in athletics and band.

	N	Mean	St. Dev.	Min	Max
Athlete	12316	0.681	0.466	0	1
Athlete Sophomore and Senior Year	11661	0.404	0.491	0	1
Athlete Only Sophomore Year	11661	0.184	0.388	0	1
Athlete Only Senior Year	11661	0.075	0.263	0	1
Band Participant	13467	0.314	0.464	0	1
Band Sophmore and Senior Year	15468	0.129	0.336	0	1
Band Only Sophmore Year	14590	0.067	0.249	0	1
Band Only Senior Year	13837	0.055	0.227	0	1

In our cross-sectional specification, we control for individual demographic characteristics. First, we control for gender. *Female* is equal to 1 if the respondent identifies as a female and is equal to 0 if the respondent identifies as male. Then, we controlled for student race. *White* is equal to 1 if the respondent identifies as white (non-Hispanic), and 0 otherwise. *Asian* is equal to 1 if the respondent identifies as Asian, Hawaii/Pacific Islander (non-Hispanic), and 0 otherwise. *Black* is equal to 1 if the respondent identifies as Black or African American, and 0 otherwise. *American Indian* is equal to 1 if the respondent identifies as American Indian / Alaska Native (non-Hispanic), and 0 otherwise. *Hispanic* is equal to 1 if the respondent identifies as Hispanic, and 0 otherwise. Finally we controlled for student age. *Age* represents the number of months before December 1987 that the student was born. We used this classification because the youngest student in the study was born in December 1987 and this way an increase in *Age* by 1 represents a student being 1 month older relative to another student in the sample. The table below represents the summary statistics for our demographic variables.

	N	Mean	St. Dev.	Min	Max
Female	15370	0.502	0.500	0	1
White	15244	0.570	0.495	0	1
Asian	15244	0.096	0.294	0	1
Black	15244	0.133	0.339	0	1
American Indian	15244	0.009	0.092	0	1
Hispanic	15244	0.145	0.353	0	1
Age	15190	23.008	6.777	0	60

Additional controls were included to represent whether a student is a native English speaker, a respondent's family income, whether their parents graduated from college, whether they intended to go to college in their sophomore year of high school, the number of siblings they have, whether or not they held a paid work position their sophomore year of high school, and the amount of time they spent per week watching television or playing video games. *English is Native Language* is equal to 1 if the respondent identifies English as their first language, and 0 otherwise. Parent Income, labeled as *Income above 50*k, is equal to 1 if the respondent's family income is above $50,000, and 0 otherwise. We chose to use $50,000 as a cutoff because this was the median respondent-reported family income. *Parent College Graduate* is equal to 1 if the respondent reports at least one parent as graduating from college, and 0 otherwise. *College Intentions* is equal to 1 if the respondent indicated their sophomore year of high school that they intended to enroll in college, and 0 otherwise. *Number of Siblings* is equal to the number of siblings the respondent lists. Paid work, labeled as *Held Job Soph Year HS*, is equal to 1 if the respondent held a job their sophomore year, and 0 otherwise. Hours spent watching TV or playing video games, labeled as *Hours spent watching TV*, is equal to the total number of hours the respondent reported spending on either activity per week. The table below shows a summary of these control variables.

	N	Mean	St. Dev.	Min	Max
English is Native Langauge	15244	0.830	0.375	0	1
Income above 50k	16197	0.487	0.500	0	1
Parent College Graduate	15321	0.411	0.492	0	1
College Intentions	13794	0.919	0.272	0	1
Number of Siblings	12243	1.458	1.219	0	7
Held Job Soph Year HS	12843	0.377	0.485	0	1
Hours Watching TV per Week	12434	3.563	2.444	0	8

In our central analysis we estimate the individual-level linear probability model[1], regressing *Attend College* on *Athlete* and the demographic control variables. In these regressions we report heteroskedastic robust standard errors. The estimating equation is

$$\textit{Attend College}_i = \beta_0 + \beta_1\, \textit{Athlete}_i + \gamma_i X_i + \epsilon_{i,t}$$

where X_i is a vector of individual demographic controls, β_0, β_1 and γ_i are coefficients to be estimated, and $\epsilon_{i,t}$ is the residual for the individual *i* at time *t*. In order to evaluate more specific hypotheses about the relationship between athletics and college attendance, we employ our more granular measures of participation described above. The estimating equation is

$$\textit{Attend College}_i = \beta_0 + \beta_1\, \textit{Athlete Sophomore and Senior Year}_i + \beta_2\, \textit{Athlete Only Sophomore Year}_i$$
$$+\ \beta_3\, \textit{Athlete Only Senior Year}_i + \gamma_i X_i + \epsilon_{i,j}$$

In this specification *Never An Athlete* is the omitted category. In addition, we estimated an analogous ordinary least squares regressions for which the dependent variable was *Change in Math Quartile*. Finally, we use this general estimating strategy for evaluating the relationship between *Band Participant* (and the more granular versions of band participation) and *Attend College* and *Change in Math Quartile.*

[1] One potential concern with a linear probability model is that it generates predicted probability less than zero or greater than one. Therefore, we also include in the appendix the estimates employing a logit specification. Our results are robust to employing this alternative econometric model.

A central concern of the paper is whether participation in sports, in addition to having an average effect on educational attainment, ameliorates socioeconomic advantages. To investigate this question, we add interaction terms between key measures of socioeconomic advantage and participation to the specifications described above. Specifically, we interact *Income Above 50K* with *Athlete/Band Participant* and *Parent College Graduate* with *Athlete/Band Participant.*

Results

Figure 1, Panel A, reports the cross-sectional OLS estimates for *Attend College* with demographic controls.[2] *Athlete* has a positive and statistically significant coefficient. Being an athlete at any point in high school is associated with a 6.15 percentage point increase in the likelihood of attending college in comparison to non-athletes. In Figure 1, Panel B, *Athlete Sophomore and Senior Year* has a positive and statistically significant coefficient. We find that continued participation in athletics throughout high school is associated with a 8.88 percentage point increase in the likelihood of attending college in comparison to the *Never an Athlete* group. Interestingly, although *Athlete Only Sophomore Year* and *Athlete Only Senior Year* have positive coefficients, neither of them are statistically significant. This suggests that there is potentially a meaningful difference between students who have continued athletic engagement as opposed to those who quit their sport or start playing later in high school.

[2] All estimates shown in coefficient plots are also recorded as tables in the appendix.

Figure 1: Athletic Participation and College Attendance Cross-Sectional Estimates
Note: This figure reports OLS cross-sectional estimates of *Attend College* on *Athlete* (Panel A) and *Attend College* on *Athlete Sophomore and Senior Year*, *Athlete Only Sophomore Year*, and *Athlete Only Senior Year* (Panel B)..

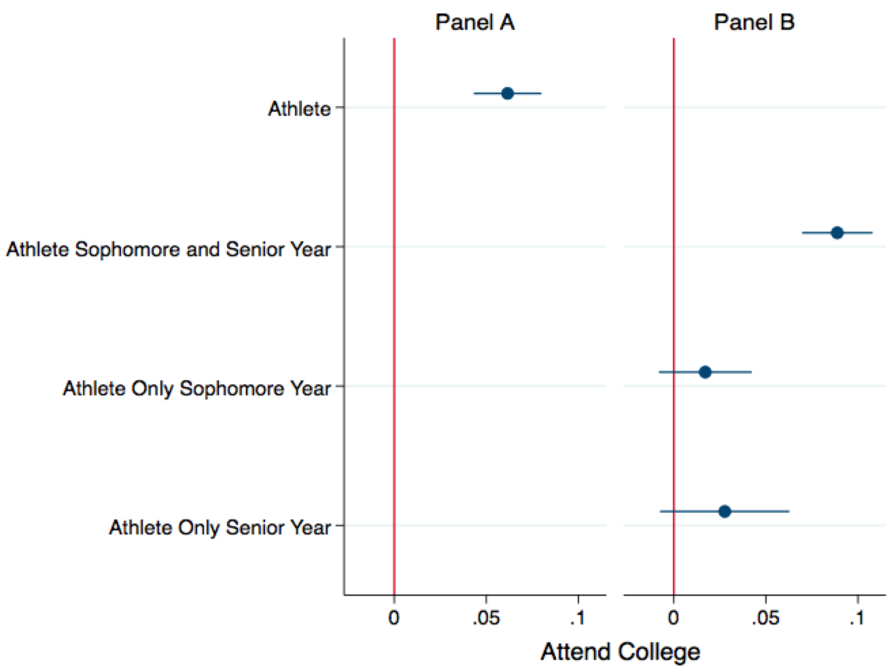
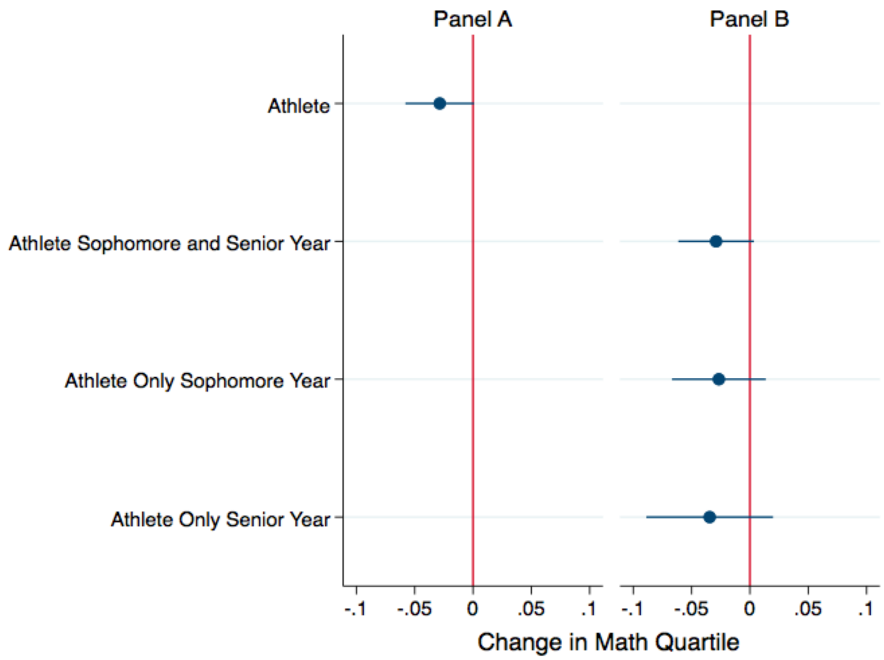


Figure 2, Panel A, reports the cross-sectional OLS estimates for *Change in Math Quartile* with demographic controls. *Athlete* has a negative, but not statistically significant coefficient. In Figure 2, Panel B, *Athlete Sophomore and Senior Year* has a negative coefficient, but again it is not statistically significant. This suggests that there is not a relationship between athletic participation and a change in math quartile scores. However, it should be noted that changing a whole math quartile between sophomore and senior year is not very common with over 69% of respondents having no change between survey waves. A more accurate assessment of this relationship would use raw scores which we unfortunately do not have access to.

Figure 2: Athletic Participation and Change in Math Quartile Cross-Sectional Estimates
Note: This figure reports OLS cross-sectional estimates of *Change in Math Quartile on Athlete* (Panel A) and *Change in Math Quartile* on *Athlete Sophomore and Senior Year*, *Athlete Only Sophomore Year*, and *Athlete Only Senior Year* (Panel B).



Given that our results show that there is some relationship between athletic participation and college attendance, we wanted to explore if this is an outcome specific to athletics or just characteristic of high school extracurriculars. Therefore, we used the same estimating strategy for evaluating the relationship between *Band Participant* and *Attend College* and *Change in Math Quartile*.

Figure 3, Panel A, reports the cross-sectional OLS estimates for *Attend College* with demographic controls. *Band Participant* has a positive but not statistically significant coefficient. In Figure 3, Panel B, *Band Sophomore and Senior Year* has a positive and statistically significant coefficient. We find that being in band throughout high school is associated with a 4.68 percentage point increase in the likelihood of attending college in comparison to the *Never In Band* group. Although *Band Only Sophomore Year* and *Band Only Senior Year* have positive coefficients, neither of them are statistically significant. Similar to the athletics results, this indicates that there is potentially a meaningful difference between students who have continued band engagement and those who do not.

Figure 3: Band Participation and College Attendance Cross-Sectional Estimates
Note: This figure reports OLS cross-sectional estimates of *Attend College* on *Band Participant* (Panel A) and *Attend College* on *Band Sophomore and Senior Year*, *Band Only Sophomore Year*, and *Band Only Senior Year* (Panel B).

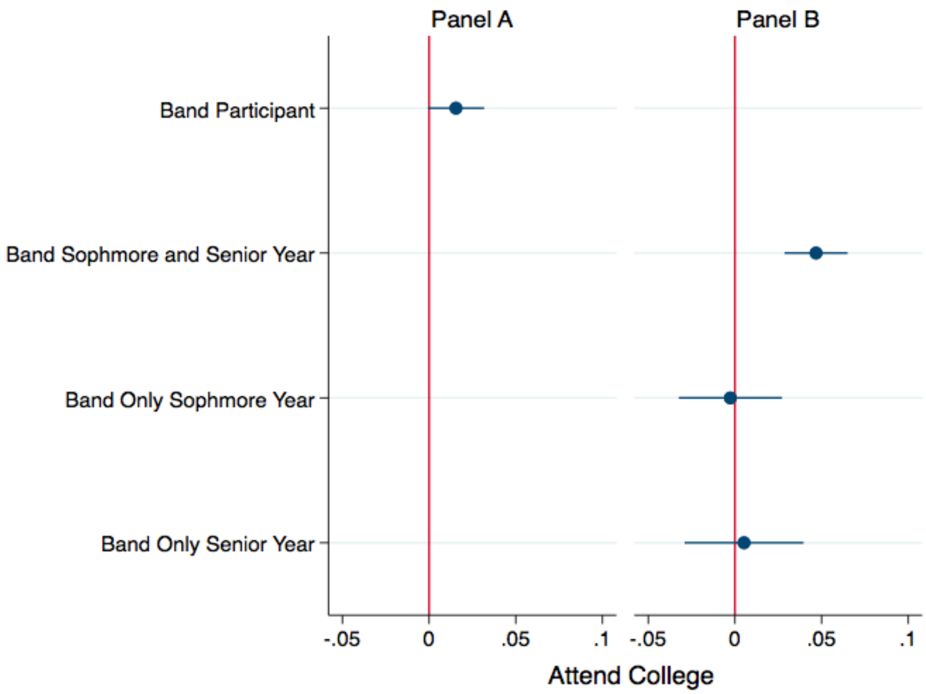
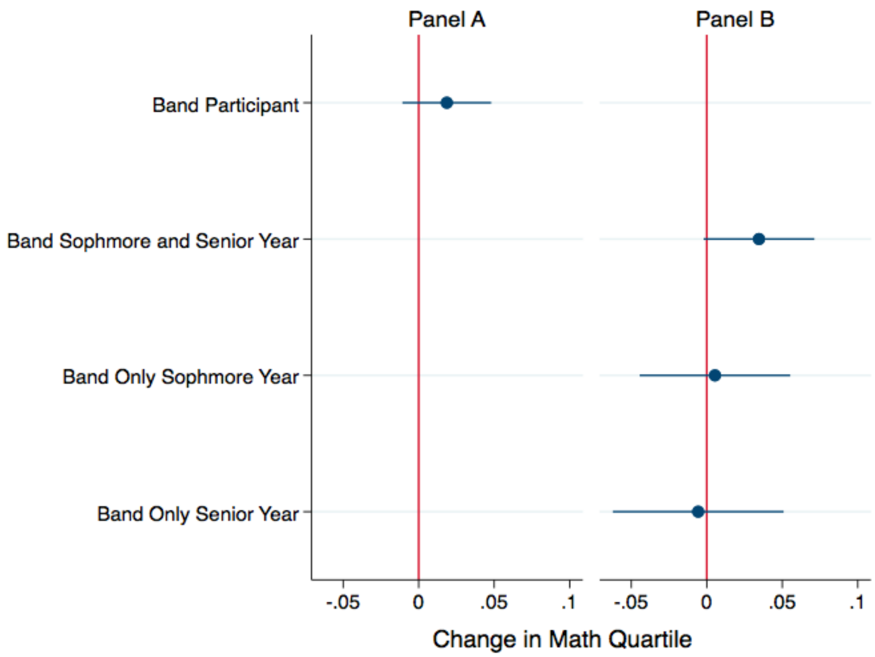


Figure 4, Panel A, reports the cross-sectional OLS estimates for *Change in Math Quartile* with demographic controls. *Band Participant* has a negative but not statistically significant coefficient. In Figure 4, Panel B, *Band Sophomore and Senior Year* has a positive coefficient but again it is not statistically significant. Similar to athletics participation, *Band Participant* is not significantly associated with *Change in Math Quartile*. However, in contrast to *Athlete* the coefficient is positive.

Figure 4: Band Participation and Change in Math Quartile Cross-Sectional Estimates
Note: This figure reports OLS cross-sectional estimates of *Change in Math Quartile* on *Band Participant* (Panel A) and *Change in Math Quartile* on *Band Sophomore and Senior Year*, *Band Only Sophomore Year*, and *Band Only Senior Year* (Panel B).



Overall, these estimates suggest that there may be a stronger association between athletics participation and college attendance compared to band participation. That said, since *Band Sophomore and Senior Year* also shows a positive and significant correlation with *Attend College,* there is still the possibility that there simply exists an association between continued extracurricular engagement and college attendance rather than being specific to sports.

Finally, we wanted to see if these extracurricular activities moderate the effects of economic privilege. The idea is that having wealthy parents or parents who went to college is strongly associated with students attending college. The reasons for this advantage are many but include students having the habits associated with academic success modeled by their parents. It may be that athletic or other extracurricular activities model these same habits and therefore lessen parental advantages. Although our results suggest that potentially any participation in high school extracurricular activities is associated with a greater probability of college attendance, it could be plausible that different kinds of activities are better at leveling the effects of economic privilege. For example, in order to make a high school sports team, a student may need a base level of skill that is only possible through expensive club teams or private lessons. On the other hand, band may welcome beginners and provide students with an instrument, meaning there is no barrier to entry. Alternatively, many sports have low entry costs, wide participation, and talent distributed evenly across socioeconomic groups. Sports participation could be particularly effective in equalizing opportunity.

We evaluate these arguments by adding an interaction between being an athlete and being economically privileged to our specifications above. We estimate this by creating two interaction terms *Income Above 50K*Athlete* and *Parent College Graduate*Athlete*. Table 1 reports coefficient estimates of specifications including these interaction terms. Columns 1 and 2 report the results for the interactions with *Athlete* while Columns 3 and 4 report the results for interactions with *Band Participant.* In Column 1, the interaction between *Income Above 50K* and *Athlete* is equal to -0.051 and is statistically significant while the coefficient estimate on *Income Above 50K* is 0.089 and also significant. This means that the differences in attending college between students with parents who make over $50,000 and those who make less are smaller for athletes. This result is consistent with the hypothesis that participation in sports mitigates economic privilege. Column 2 reports a similar set of results when measuring economic privilege by whether the student had a parent who attended college. The estimated interaction terms for band, however, are not statistically significant and switch signs across specifications in Columns 3 and 4.

To better interpret substantively what the interaction terms means, it is instructive to calculate the marginal effect of a given variable of interest on the probability of attending college. In the specifications in Table 1, Column 1, the marginal effect of having a parent who makes more than $50k for a non-athlete is just the coefficient on the variable *Income Above 50K*. The marginal effect of having a parent who makes more than $50k for an athlete is equal to the sum of the coefficient on the variable *Income Above 50K* and the interaction term. Figures 5 and 6 plot these marginal effects for all four specifications in Table 1.[3] In Figure 5, Panel A we can see that for a non-athlete the marginal effect of having a parent with an income of $50k is about 11.5 percentage points while it is 6.5 percentage point for an athlete. This suggests about half of the advantage is removed through participation in sports. The estimates in Panel B indicate about half of the advantage of having a parent who went to college is eroded for athletes. We do not see the same consistent and significant differences for band participation in the results in Figure 6.

[3] Note that the negative interaction term is also consistent with athlete participation mattering less for privileged students. One could also calculate the marginal effects of being an athlete for privileged and non-privileged students to emphasize this interpretation.

Table 1: Athletic/Band Participation and College Attendance Cross-Sectional Estimates with Interaction Terms
Note: This figure reports OLS cross-sectional estimates of *Attend College* on *Athlete* and *Band Participant* in addition to interaction terms with *Income Above 50K* and *Parent College Graduate*.

	(1)	(2)	(3)	(4)
Athlete	0.0887***	0.0960***		
	(0.0160)	(0.0148)		
Income Above 50K*Athlete	-0.0510**			
	(0.0190)			
Parent College Graduate*Athlete		-0.0771***		
		(0.0177)		
Band Participant			0.0304	0.0122
			(0.0155)	(0.0143)
Income Above 50K*Band Participant			-0.0257	
			(0.0178)	
Parent College Graduate*Band Participant				0.00656
				(0.0164)
Female	0.0280***	0.0285***	0.0127	0.0130
	(0.00844)	(0.00843)	(0.00845)	(0.00846)
Hispanic	0.0362	0.0362	0.0340	0.0330
	(0.0273)	(0.0273)	(0.0268)	(0.0268)
White	0.0764***	0.0768***	0.0836***	0.0833***
	(0.0230)	(0.0230)	(0.0224)	(0.0224)
Black	0.0692*	0.0696*	0.0694**	0.0695**
	(0.0270)	(0.0270)	(0.0265)	(0.0265)
American Indian	-0.0366	-0.0331	-0.0104	-0.0126
	(0.0720)	(0.0721)	(0.0707)	(0.0707)
English is Native Language	-0.0444**	-0.0429**	-0.0457**	-0.0452**
	(0.0157)	(0.0156)	(0.0158)	(0.0158)
Number of Siblings	-0.00872*	-0.00854*	-0.00999**	-0.00992*
	(0.00390)	(0.00389)	(0.00387)	(0.00387)
Parent College Graduate	0.0944***	0.148***	0.105***	0.103***
	(0.00847)	(0.0157)	(0.00857)	(0.0104)
Income Above 50K	0.115***	0.0806***	0.0944***	0.0861***
	(0.0169)	(0.00952)	(0.0115)	(0.00949)
Hours Watching per Week	-0.00943***	-0.00964***	-0.0112***	-0.0112***
	(0.00196)	(0.00196)	(0.00196)	(0.00196)
Held Paid Job as Sophomore	-0.00961	-0.00930	-0.00740	-0.00723
	(0.00846)	(0.00846)	(0.00849)	(0.00849)
Observations	6602	6602	6987	6987
R^2	0.161	0.162	0.167	0.167

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

Figure 5: Marginal Effect of Income Above 50k and Parent College Graduate on College Attendance for Non-Athletes and Athletes

Note: This figure reports the marginal effect of having an income above 50k on college attendance for non-athletes and athletes (Panel A) and reports the marginal effect of having at least one parent graduate college on college attendance for non-athletes and athletes (Panel B).

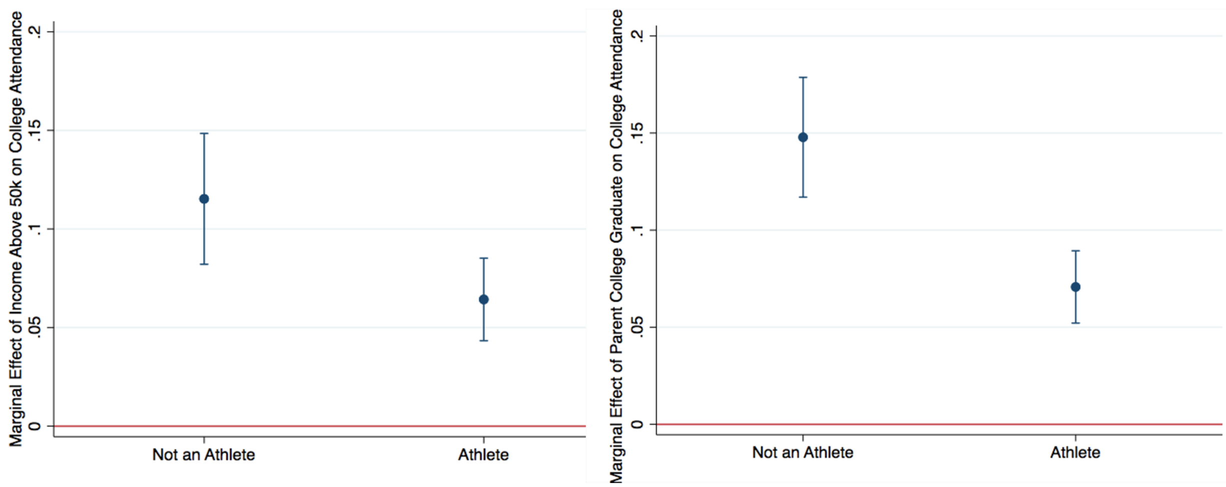
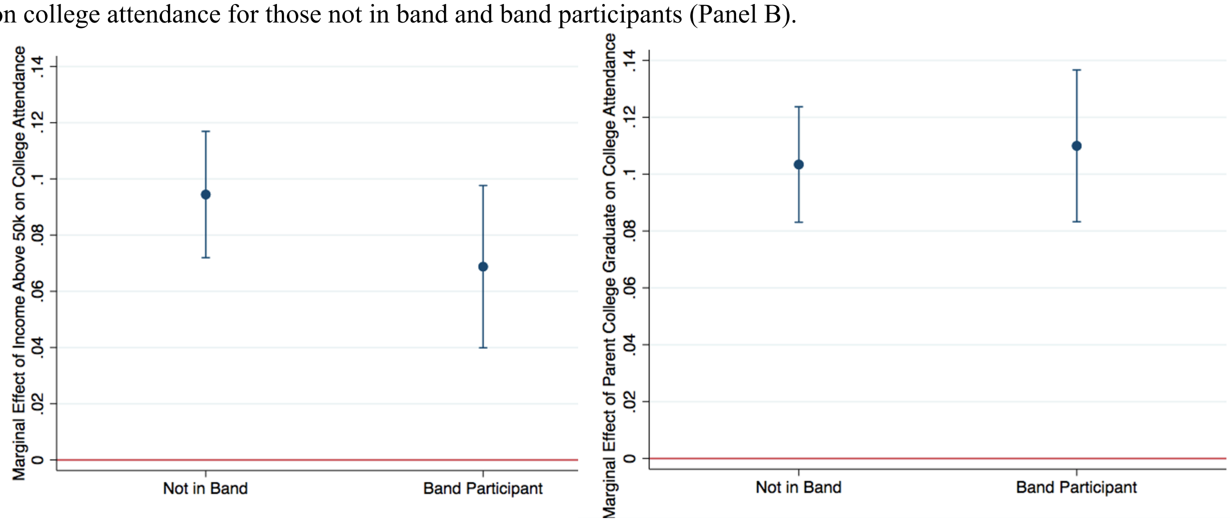


Figure 6: Marginal Effect of Income Above 50k and Parent College Graduate on College Attendance for Non-Band Participants and Band Participants

Note: This figure reports the marginal effect of having an income above 50k on college attendance for those not in band and band participants (Panel A) and reports the marginal effect of having at least one parent graduate college on college attendance for those not in band and band participants (Panel B).



Discussion

High school is an incredibly busy time in an adolescent's life, often made more stressful by the college application process and more generally what to do after high school ends. In the literature there are mixed results regarding the effects of athletics participation on educational outcomes. Some studies, Lumpkin and Favor (2008), Dyer et al. (2017), Wretman (2017), Fox et al. (2010), and Chen et al. (2021) find a positive relationship between participation in athletics and academic performance. Others however are less convinced, Rees and Sabia (2010), Eccles et al. (2003), Eccles and Barber (1999), and Eide and Ronan (2001), or find evidence to the contrary, Scott et al. (2008), and Aries et al. (2004). Considering the initial impact of the COVID-19 pandemic removed many students from their activities and the continued threat of variants such as Delta and Omicron, parents must decide if it is worth it to encourage their children to participate in school sports or other high school extracurriculars.

We find a strong positive association between athletics participation and college attendance. Interestingly, when we break athletics participation down into more granular categories we find the positive effect only remains for students who played sports both Sophomore and Senior year. Those who quit after their Sophomore season or started Senior year did not see the same statistically significant correlation indicating that maintained participation is important to the effect of athletics participation on college attendance. Moreover, we also find in our cross-sectional specifications that this positive relationship also exists for other high school extracurriculars such as band when there is sustained engagement. This suggests it may simply be long-term commitment to an extracurricular activity in general which improves the probability of a student attending college. Our other measure of educational achievement, *Change in Math Quartile*, did not have a significant relationship with either athletic or band participation. However, as noted in the results section a majority of students did not have any change in their math quartile between Sophomore and Senior year making the estimate arguably less informative.

In addition to our findings that support a generally positive association between continued extracurricular engagement and college attendance we find that athletic participation can help moderate the effects of economic privilege. In particular, our marginal effects figures show that about half of the advantage of having parents that make above 50k or having a parent that graduated from college is removed through participation in sports. Potentially sports provide life skills or opportunities that are similar to those given to students when they are economically privileged therefore acting as a moderating force. The same differences in marginal effects were not seen for band participation.

Although our econometric exploration may give a broad idea of the relationship between high school athletics participation and some educational outcomes, namely college attendance, there are a number of limitations to our study. First and foremost, because the data we are working with is cross-sectional and our analysis is correlational, we must consider the possibility of omitted variable bias. One prominent omitted variable we are concerned with is the innate drive / work ethic of each student. Focusing on *Attend College* as our outcome variable and *Athlete* as our independent variable, we believe that innate drive is positively correlated with being an athlete. Athletics often require intense practices, long-term dedication, and a strong mindset so people with higher innate drive are more likely to compete in athletics. Further, we believe that innate drive is positively correlated with college attendance because it requires a strong work ethic to do well in highschool and successfully apply and get accepted into college. Thus, we predict a positive bias by not including a measure for innate drive in our regression of college attendance on being an athlete. Thus, since our estimate for the effect of being an athlete on college attendance is positive and we have positive bias, we are likely overstating the effect of being an athlete on whether a student attends college.

In addition to a student's innate drive, we are also concerned about the quality of the high school as an omitted variable. Although we partially control for school quality by including family income in our regression (higher socioeconomic neighborhoods tend to have higher quality schools), we believe there are factors beyond income that influence the quality of the academics / extracurricular opportunities provided at a given high school. We believe that a better quality / better-funded high school is associated with a higher proportion of students being athletes because the school can provide better facilities and coaches for the students. In addition, a better high school in terms of fundings and academics will send a higher proportion of students to college because students will be better-prepared academically and are offered more college-preparatory support. Thus, since both of these correlations are positive, we expect a positive bias from not including a measure for the quality of the high school in our analysis. Therefore, because the coefficient on athletes is positive in our regression of college attendance on whether a student is an athlete and the presence of positive bias, we are likely overstating the effect of athletic participation on college attendance.

While we have demonstrated a couple relevant omitted variables, we recognize that there could be many more confounding variables that we are unable to include in our analysis due to data limitations. Ideally, we would also want to include measures for the home environment that parents provide to their children. Parents can influence their child's educational outcomes, in ways that are unrelated to income. For example, parents who are below the median income in our study could be reading to their children every night and encouraging extracurricular participation because they want their children to be able to attend college. Further, we are also unable to control for the innate athletic ability that a student has, which is primarily a genetic factor. For example, a student can be a very good athlete and be offered an athletic scholarship to a university which they would not have been able to academically attend without playing their sport. Another factor we would ideally be able to control for is a student's financial ability to attend college. We recognize that the correlational nature of our analysis is prone to omitted variables and thus refrain from making causal claims.

If not limited by our data and time constraints, we would have explored instrumental variable regressions to improve our ability to make causal claims about athletic participation and educational outcomes. Weather could serve as an instrumental variable on athletic participation, especially for outdoor sports, because as the weather gets warmer, athletes are better able to practice more and play additional games. Thus, warm weather is likely positively correlated with athletic participation. Further, we believe that there is no (or very little) correlation between warm weather and college attendance, other than through athletic participation. This means that warm weather, as a potential instrument in our regression analysis, would have instrument relevance and exogeneity. By using warm weather as an instrument, we believe that we would be better able to isolate the causal effect of athletic participation on educational outcomes such as college attendance.

Another method that could be used to better causally estimate the relationship between athletic participation and college attendance could be a regression discontinuity design capitalizing on the arbitrary cut off dates for school enrollment. One could imagine that those who are right above the cutoff and therefore older in their grade are more likely to be athletes because they are bigger and stronger. Therefore, if we had access to a student's athletics participation, exact birthdate, and educational outcomes we could assess if there is a difference in educational outcomes for those just above and below the birthday cut off through athletics participation. This design assumes that there is no education advantage to being older in a given grade cohort. One could imagine violations of this assumption but it is not clear whether they push altogether in a positive or negative direction or what the magnitude of such an effect might be.

References

1. Aries, E., McCarthy, D., Salovey, P., & Banaji, M. R. (2004). A Comparison of Athletes and Non-Athletes at Highly Selective Colleges: Academic Performance and Personal Development. *Research in Higher Education*, *45*(6), 577–602. https://doi.org/10.1023/b:rihe.0000040264.76846.e9
2. Billingsley, J.T., Hurd, N.M. Discrimination, Mental Health and Academic Performance Among Underrepresented College Students: The Role of Extracurricular Activities at Predominantly White Institutions. Soc Psychol Educ 22, 421–446 (2019). https://doi.org/10.1007/s11218-019-09484-8
3. Buckley, P., & Lee, P. (2018). The Impact of Extra-Curricular Activity on the Student Experience. *Active Learning in Higher Education*, *22*(1), 37–48. https://doi.org/10.1177/1469787418808988
4. Chen, S., Li, X., Yan, J., & Ren, Z. (2021). To be a Sportsman? Sport Participation is Associated with Optimal Academic Achievement in a Nationally Representative Sample of High School Students. *Frontiers in Public Health*, *9*. https://doi.org/10.3389/fpubh.2021.730497
5. Gardner, M., Roth, J., & Brooks-Gunn, J. (2008). Adolescents' Participation in Organized Activities and Developmental Success 2 and 8 years After High School: Do Sponsorship, Duration, and Intensity Matter? Developmental Psychology, 44(3), 814–830. https://doi.org/10.1037/0012-1649.44.3.814
6. Dyer, A. M., Kristjansson, A. L., Mann, M. J., Smith, M. L., & Allegrante, J. P. (2017). Sport Participation and Academic Achievement: A Longitudinal Study. American Journal of Health Behavior, 41(2), 179–185. https://doi.org/10.5993/ajhb.41.2.9
7. Eccles, J. S., Barber, B. L., Stone, M., & Hunt, J. (2003). Extracurricular Activities and Adolescent Development. Journal of Social Issues, 59(4), 865–889.
8. Eccles, J. S., & Barber, B. L. (1999). Student Council, Volunteering, Basketball, or Marching Band: What Kind of Extracurricular Involvement Matters? Journal of Adolescent Research, 14(1), 10–43.
9. Eide, E. R., & 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(5), 431–442.
10. Fox, C. K., Barr-Anderson, D., Neumark-Sztainer, D., & Wall, M. (2010). Physical Activity and Sports Team Participation: Associations with Academic Outcomes in Middle School and High School Students. Journal of School Health, 80(1), 31–37. https://doi.org/10.1111/j.1746-1561.2009.00454.x
11. Jerrim, J., Lopez‐Agudo, L. A., & Marcenaro‐Gutierrez, O. D. (2020). Posh but Poor: The Association Between Relative Socio‐economic Status and Children’s Academic Performance. Review of Income and Wealth, 67(2), 334–362. https://doi.org/10.1111/roiw.12476

12. Knifsend, C.A., Graham, S. Too Much of a Good Thing? How Breadth of Extracurricular Participation Relates to School-Related Affect and Academic Outcomes During Adolescence. J Youth Adolescence 41, 379–389 (2012). https://doi.org/10.1007/s10964-011-9737-4
13. Lumpkin, Angela and Favor, Judy (2013) "Comparing the Academic Performance of High School Athletes and Non-Athletes in Kansas 2008-2009," Journal of Applied Sport Management: Vol. 4 : Iss. 1. https://trace.tennessee.edu/jasm/vol4/iss1/31
14. Mahoney, J. L., Cairns, B. D., & Farmer, T. W. (2003). Promoting Interpersonal Competence and Educational Success Through Extracurricular Activity Participation. Journal of Educational Psychology, 95(2), 409–418. https://doi.org/10.1037/0022-0663.95.2.409
15. Moody, J. (2020). How Coronavirus is Upending High School Extracurriculars. US News: World Report. Retrieved December 11, 2021, from https://www.usnews.com/education/best-colleges/articles/how-coronavirus-is-upending-high-school-extracurricular-activities.
16. Morris, D. S. (2016). Extracurricular Activity Participation in High School. American Educational Research Journal, 53(5), 1376–1410. https://doi.org/10.3102/0002831216667579
17. Morris, D. S. (2015). Actively Closing the Gap? Social Class, Organized Activities, and Academic Achievement in High School. Youth & Society, 47(2), 267–290. https://doi.org/10.1177/0044118X12461159
18. Muñoz-Bullón, F., Sanchez-Bueno, M. J., & Vos-Saz, A. (2017). The Influence of Sports Participation on Academic Performance Among Students in Higher Education. Sport Management Review, 20(4), 365–378. https://doi.org/10.1016/j.smr.2016.10.006
19. Pérez-Ordás, R., Aznar Cebamanos, M., Nuviala, R., & Nuviala, A. (2019). Evaluation of Extracurricular Sports Activities as an Educational Element for Sustainable Development in Educational Institutions. *Sustainability*, *11*(12), 3474. https://doi.org/10.3390/su11123474
20. Pinto, L. H., & Ramalheira, D. C. (2017). Perceived Employability of Business Graduates: The Effect of Academic Performance and Extracurricular Activities. Journal of Vocational Behavior, 99, 165–178. https://doi.org/10.1016/j.jvb.2017.01.005
21. Rees, D. I., & Sabia, J. J. (2010). Sports Participation and Academic Performance: Evidence from the National Longitudinal Study of Adolescent Health. Economics of Education Review, 29(5), 751–759. https://doi.org/10.1016/j.econedurev.2010.04.008
22. Robst, J., & Keil, J. (2000). The Relationship Between Athletic Participation and Academic Performance: Evidence from NCAA Division III. Applied Economics, 32(5), 547–558. https://doi.org/10.1080/000368400322453
23. Rombokas, M. (1995). High School Extracurricular Activities & College Grades. ERIC. https://eric.ed.gov/?id=ED391134

24. Scott, B. M., Paskus, T. S., Miranda, M., Petr, T. A., & McArdle, J. J. (2008). In-Season vs. Out-of-Season Academic Performance of College Student-Athletes. Journal of Intercollegiate Sport, 1(2), 202–226. https://doi.org/10.1123/jis.1.2.202
25. Winstone, N., Balloo, K., Gravett, K., Jacobs, D., & Keen, H. (2020). Who Stands to Benefit? Wellbeing, Belonging and Challenges to Equity in Engagement in Extracurricular Activities at University. Active Learning in Higher Education. https://doi.org/10.1177/1469787420908209
26. Wretman, C. J. (2017). School Sports Participation and Academic Achievement in Middle and High School. Journal of the Society for Social Work and Research, 8(3), 399–420. https://doi.org/10.1086/693117

Appendix

Table A1: Athletic Participation and College Attendance Cross-Sectional Estimates

	(1)	(2)	(3)	(4)
Athlete	0.119***	0.0615***		
	(0.00846)	(0.00938)		
Athlete Sophomore and Senior Year			0.155***	0.0888***
			(0.00883)	(0.00975)
Athlete Only Sophomore Year			0.0604***	0.0171
			(0.0118)	(0.0129)
Athlete Only Senior Year			0.0771***	0.0277
			(0.0158)	(0.0179)
Female		0.0281***		0.0336***
		(0.00844)		(0.00852)
Hispanic		0.0344		0.0369
		(0.0273)		(0.0274)
White		0.0749**		0.0760**
		(0.0230)		(0.0231)
Black		0.0681*		0.0750**
		(0.0270)		(0.0272)
American Indian		-0.0361		-0.0365
		(0.0724)		(0.0711)
English is Native Language		-0.0431**		-0.0468**
		(0.0156)		(0.0157)
Number of Siblings		-0.00885*		-0.00957*
		(0.00390)		(0.00390)
Parent College Graduate		0.0941***		0.0909***
		(0.00847)		(0.00848)
Income Above 50K		0.0807***		0.0778***
		(0.00953)		(0.00953)
Hours Watching per Week		-0.00943***		-0.00890***
		(0.00196)		(0.00196)
Held Paid Job as Sophomore		-0.00922		-0.00859
		(0.00846)		(0.00847)
Observations	11055	6602	10509	6555
R^2	0.021	0.160	0.030	0.165

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

Table A2: Athletic Participation and Change in Math Quartile Cross-Sectional Estimates

	(1)	(2)	(3)	(4)
Athlete	-0.0164	-0.0285		
	(0.0116)	(0.0150)		
Athlete Sophomore and Senior Year			-0.00631	-0.0290
			(0.0128)	(0.0165)
Athlete Only Sophomore Year			-0.0311	-0.0266
			(0.0163)	(0.0205)
Athlete Only Senior Year			-0.0375	-0.0345
			(0.0220)	(0.0277)
Female		-0.0377**		-0.0385**
		(0.0143)		(0.0144)
Hispanic		0.0447		0.0488
		(0.0442)		(0.0443)
White		-0.0284		-0.0213
		(0.0383)		(0.0384)
Black		0.00553		0.0150
		(0.0434)		(0.0435)
American Indian		0.127		0.134
		(0.0931)		(0.0946)
English is Native Language		-0.0174		-0.0196
		(0.0265)		(0.0266)
Number of Siblings		0.00706		0.00616
		(0.00599)		(0.00602)
Parent College Graduate		0.0719***		0.0717***
		(0.0149)		(0.0150)
Income Above 50K		0.00601		0.00576
		(0.0158)		(0.0159)
Hours Watching per Week		-0.00737*		-0.00785*
		(0.00317)		(0.00318)
Held Paid Job as Sophomore		-0.0249		-0.0226
		(0.0143)		(0.0144)
Observations	12226	7719	11637	7659
R^2	0.000	0.011	0.000	0.011

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B1: Athletic Participation and College Attendance Cross-Sectional Estimates

	(1)	(2)	(3)	(4)
Band Participant	0.0408***	0.0155		
	(0.00769)	(0.00828)		
Band Sophomore and Senior Year			0.0922***	0.0468***
			(0.00891)	(0.00927)
Band Only Sophomore Year			0.0441**	-0.00263
			(0.0136)	(0.0152)
Band Only Senior Year			0.0232	0.00528
			(0.0159)	(0.0175)
Female		0.0129		0.0156
		(0.00845)		(0.00843)
Hispanic		0.0332		0.0295
		(0.0268)		(0.0268)
White		0.0833***		0.0771***
		(0.0224)		(0.0224)
Black		0.0694**		0.0689**
		(0.0265)		(0.0264)
American Indian		-0.0126		-0.0193
		(0.0707)		(0.0711)
English is Native Language		-0.0452**		-0.0440**
		(0.0158)		(0.0158)
Number of Siblings		-0.00994*		-0.00950*
		(0.00387)		(0.00388)
Parent College Graduate		0.106***		0.104***
		(0.00857)		(0.00852)
Income Above 50K		0.0861***		0.0844***
		(0.00949)		(0.00949)
Hours Watching per Week		-0.0112***		-0.01000***
		(0.00196)		(0.00195)
Held Paid Job as Sophomore		-0.00727		-0.00710
		(0.00849)		(0.00847)
Observations	12010	6987	11641	6890
R^2	0.002	0.167	0.007	0.164

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B2: Band Participation and Change in Math Quartile Cross-Sectional Estimates

	(1)	(2)	(3)	(4)
Band Participant	0.0175	0.0186		
	(0.0117)	(0.0150)		
Band Sophomore and Senior Year			0.0240	0.0346
			(0.0151)	(0.0187)
Band Only Sophomore Year			0.0190	0.00542
			(0.0207)	(0.0254)
Band Only Senior Year			-0.00406	-0.00569
			(0.0225)	(0.0289)
Female		-0.0413**		-0.0429**
		(0.0147)		(0.0147)
Hispanic		0.0409		0.0433
		(0.0450)		(0.0450)
White		-0.0307		-0.0317
		(0.0386)		(0.0386)
Black		-0.00943		-0.0102
		(0.0437)		(0.0438)
American Indian		0.0540		0.0548
		(0.0927)		(0.0927)
English is Native Language		-0.0272		-0.0233
		(0.0277)		(0.0277)
Number of Siblings		0.00492		0.00495
		(0.00619)		(0.00619)
Parent College Graduate		0.0663***		0.0657***
		(0.0153)		(0.0153)
Income Above 50K		0.00666		0.00602
		(0.0162)		(0.0163)
Hours Watching per Week		-0.00705*		-0.00684*
		(0.00323)		(0.00324)
Held Paid Job as Sophomore		-0.0252		-0.0256
		(0.0148)		(0.0148)
Observations	12675	7274	12516	7262
R^2	0.000	0.011	0.000	0.011

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C1: Logit Estimation for Athletic Participation and College Attendance

	(1)	(2)
Athlete	0.562***	
	(0.0799)	
Athlete Sophomore and Senior Year		0.925***
		(0.0970)
Athlete Only Sophomore Year		0.122
		(0.104)
Athlete Only Senior Year		0.218
		(0.154)
Female	0.287***	0.345***
	(0.0827)	(0.0836)
Hispanic	0.315	0.336
	(0.203)	(0.205)
White	0.616***	0.622***
	(0.169)	(0.173)
Black	0.622**	0.674**
	(0.203)	(0.207)
American Indian	-0.104	-0.133
	(0.461)	(0.454)
English is Native Language	-0.402**	-0.437**
	(0.152)	(0.152)
Number of Siblings	-0.0804*	-0.0864*
	(0.0337)	(0.0340)
Parent College Graduate	1.016***	0.990***
	(0.0938)	(0.0942)
Income Above 50K	0.743***	0.721***
	(0.0881)	(0.0889)
Hours Watching per Week	-0.0892***	-0.0868***
	(0.0170)	(0.0172)
Held Paid Job as Sophomore	-0.0998	-0.102
	(0.0820)	(0.0826)
Observations	6602	6555
R^2		

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C2: Logit Estimation for Band Participation and College Attendance

	(1)	(2)
Band Participant	0.162*	
	(0.0812)	
Band Sophomore and Senior Year		0.547***
		(0.117)
Band Only Sophomore Year		-0.0136
		(0.137)
Band Only Senior Year		0.0354
		(0.158)
Female	0.125	0.147
	(0.0773)	(0.0785)
Hispanic	0.276	0.273
	(0.189)	(0.192)
White	0.622***	0.600***
	(0.158)	(0.162)
Black	0.589**	0.609**
	(0.186)	(0.191)
American Indian	0.0433	0.0132
	(0.439)	(0.444)
English is Native Language	-0.389**	-0.390**
	(0.144)	(0.145)
Number of Siblings	-0.0846**	-0.0855**
	(0.0319)	(0.0326)
Parent College Graduate	1.059***	1.071***
	(0.0890)	(0.0906)
Income Above 50K	0.748***	0.748***
	(0.0822)	(0.0839)
Hours Watching per Week	-0.0955***	-0.0867***
	(0.0160)	(0.0162)
Held Paid Job as Sophomore	-0.0737	-0.0724
	(0.0768)	(0.0784)
Observations	6987	6890
R^2		

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$