Do High School Sports Build or Reveal Character?*

Michael R Ransom†

BYU and IZA

Tyler Ransom‡

Oklahoma and IZA

October 21, 2017

Abstract

We examine the extent to which participation in high school athletics has beneficial effects on future education, labor market, and health outcomes. Due to the absence of plausible instruments in observational data, we use recently developed methods that relate selection on observables with selection on unobservables to estimate bounds on the causal effect of athletics participation. We analyze these effects in the US separately for men and women using three different nationally representative longitudinal data sets that each link high school athletics participation with later-life outcomes. We do not find consistent evidence of individual benefits reported in many previous studies—once we have accounted for selection, high school athletes are no more likely to attend college, earn higher wages, or participate in the labor force. However, we do find that men (but not women) who participated in high school athletics are more likely to exercise regularly as adults. Nevertheless, athletes are no less likely to be obese.

JEL codes: I20, J24

Keywords: Human capital, High school sports

*This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (`http://www.cpc.unc.edu/addhealth`). No direct support was received from grant P01-HD31921 for this analysis.

We are grateful to Ivan Kwok for excellent research assistance and to Andrew Beauchamp for sharing helpful code for compiling the Add Health data. We are especially grateful to Jim Moody and Social Science Research Institute (SSRI) at Duke University for generously providing computing resources with which to access the Add Health data. This paper benefited from helpful comments by Patrick Coate, Jeff Denning, Eric Eide, Brian Krauth, and seminar participants at Duke, Carolina Population Center, and WEAI 2017 (San Diego, CA). All errors are our own.

†Contact: Department of Economics, 143 FOB, Provo, UT 84604. ransom@byu.edu. (801) 422-4736

‡Corresponding author. Contact: Department of Economics, 322 CCD1, 308 Cate Center Drive, Norman, OK 73072. ransom@ou.edu. (405) 325-2861.

1 Introduction

Participating in sports is a cultural rite of passage for adolescents in many countries, including the United States. According to the National Federation of State High School Associations (NFHS) in the US, 7.9 million high school students (56%), play some kind of sport. Sports participation has also trended upward over time, and participation in sports organized by high schools has increased steadily over the past 25 years (National Federation of State High School Associations, 2017).

Given widespread participation in sports, it is natural to ask if the benefits outweigh the costs, both to individual athletes and to schools. While potential benefits of sports participation on long-term individual outcomes have been widely publicized (Dick's Sporting Goods, 2017), participating in athletics may be costly for individual students by taking time away from academic pursuits (Coleman, 1961) or increasing injury risk (Fair and Champa, 2017). Moreover, maintaining athletic programs is a non-trivial cost for schools—so much so that athletic programs are being dropped from an increasing number of school districts. It is estimated that 27% of public high schools will have no athletic programs by the year 2020 (Dick's Sporting Goods, 2017; Up 2 Us Sports, 2017). This is a particularly surprising trend in light of the continued growth in the number of students participating.

The primary question amid the debate of whether to maintain funding for high school athletics is whether or not athletic participation benefits students in line with the purposes of schools. That is, does participation enhance human capital of students in ways that will improve their lives, as opposed to simply providing an enjoyable recreational activity? We add our analysis to a large number of previous studies that have used observational data to also investigate this question. The primary empirical approach in existing studies has been to either assume that athletes are randomly assigned, or to use instrumental variables or quasi-experimental policy changes to estimate a plausibly causal effect. We take a different approach by instead asserting that, outside of one-time large-scale policy changes, no plausibly exogenous instruments exist. Instead, we make use of recently developed econometric methods that relate selection on observables with selection on unobservables to bound the causal effects of participation in high school sports (see also Altonji, Elder, and Taber, 2005b; Millimet, Tchernis, and Husain, 2010; Millimet and Tchernis, 2013; Krauth, 2016; Oster, Forthcoming).

The econometric method we utilize in our analysis is developed by Krauth (2016) and allows researchers to empirically test the extent of deviations from exogeneity in a linear model with univariate treatment. Specifically, this method puts bounds on the correlation between the policy variable and the unobservable characteristics relative to the correlation between the policy variable and observable characteristics. We implement the method as a sensitivity analysis to include the case where sports participation is correlated with the error term in the outcome equation.

Athletic participation is strongly positively correlated with a number of outcomes—including high school graduation, college attendance, college graduation, wages, exercise habits, and absence of obesity—but we find that this correlation is almost completely due to selection. For most of the outcomes that we consider, we find that even if the correlation between athletic participation and unobservable characteristics is a small fraction of the correlation between athletic participation and observable characteristics, then there is no effect of sports. Across several different outcomes and different samples, we find no consistent benefit from high school sports. However, in a few cases that we discuss below, we do find statistically significant effects from sports participation that are arguably causal.

We analyze three separate nationally representative longitudinal surveys that link athletic participation in high school with future individual outcomes such as post-secondary education, labor market earnings, health, and propensity to engage in risky behaviors. The three surveys are the National Longitudinal Survey of Youth, 1979 (NLSY79); the National Education Longitudinal Study of 1988 (NELS:88); and the National Longitudinal Study of Adolescent to Adult Health (Add Health). Each of these studies has been used previously by researchers to analyze effects of high school sports, but no study has jointly analyzed all three.[1]

Our primary contributions are three-fold: *(i)* to assess the sensitivity of previous causal claims using recently developed econometric methods; *(ii)* to document the impact of sports participation on health and behavioral outcomes in addition to education and labor market outcomes; and *(iii)* to examine heterogeneity in the effects by gender.

Our generally null results inform the policy debate on high school sports by providing evidence against claims that sports foster skills that improve educational or labor market outcomes. Such skills, often mentioned by proponents of high school athletics, include leadership, teamwork, patience, persistence, and positive health habits (Dick's Sporting Goods, 2017). There are two potential pathways through which this null effect might operate. First, participation in sports requires a minimum level of social or health skill. For sports participation to be causal, it would need to be the case that post-participation skill levels among athletes be even higher than the initial levels of these skills. Second, even if sports raise the level of these skills among participants, it is possible that alternative activities such as non-athletic clubs also foster these skills. That is, sports participation might crowd out other activities that would encourage accumulation of the same or similarly valuable skills.

Our paper proceeds as follows. In the next section, we outline the relevant variables from our various data sources. Section 3 discusses theoretical reasons for why athletics might have an impact on future outcomes, and also discusses identification problems and how our method overcomes them. Section 4 presents our primary empirical results, and Section 5

[1] Ewing (1998), Barron, Ewing, and Waddell (2000), and Ewing (2007) use the NLSY79; Anderson (1998), Anderson (2001), and Leeds, Miller, and Stull (2007) each use the NELS:88; and Gorry (2016) and Rees and Sabia (2010) both use the Add Health.

concludes.

2 Data

Our analysis makes use of three separate nationally representative American data sets that survey youth during their secondary school years with repeated surveys into their adult life. The three studies we use are: *(i)* the National Longitudinal Survey of Youth 1979 (NLSY79); *(ii)* the National Education Longitudinal Study of 1988 (NELS:88); and *(iii)* the National Longitudinal Study of Adolescent to Adult Health (Add Health). Each survey contains slightly different information on sports participation, as well as other contextual variables and outcomes. Below, we summarize the similarities and differences among the three surveys.

2.1 NLSY79

The NLSY79 surveyed 12,686 American youth who were between the ages of 14 and 22 in 1979 and followed respondents annually or biennially for 25 rounds, until 2012. Youth were sampled at the household level, and all interviews were conducted at home. The NLSY79 includes data on the following topics which are relevant to our analysis: *(i)* personal and family background, including intelligence test scores, race, ethnicity, family income, parental education, parental co-residence, and year of birth; *(ii)* high school sports participation; *(iii)* educational attainment, including high school graduation, post-secondary college attendance, and four-year college graduation; *(iv)* labor market outcomes, including full-time employment status and wages; and *(v)* health outcomes such as height and weight, which we use to compute Body Mass Index (BMI), the metric used to diagnose obesity.

The sports participation question in the NLSY79 is asked in the fifth round of survey, when respondents would have been between 19 and 27 years old, and asks respondents to select from a list all high school clubs or extracurricular activities they had participated in. The list of activities includes student government, performing arts, yearbook/newspaper staff, National Honor Society, and "athletics, cheerleading, or pep clubs."

2.2 NELS:88

The NELS:88 was conducted by the United States National Center for Educational Statistics (NCES). The potential sample consists of about 25,000 students from 1,052 randomly selected public and private schools in the United States. Respondents were 8th-grade students in 1988 at the time the survey was initiated. (Each school could contribute up to 26 students to the sample.) The study conducted four additional follow-ups: in 1990 (when most of the cohort was in the 10th grade); in 1992 (12th grade); again in 1994 (two years after most students had left high school); and a final follow-up in 2000 (when most students would have been out of high school for eight years). The survey includes responses from students, parents, teachers, and school administrators, so there are detailed data about parental background, school activities, and school characteristics. The NELS:88 contains information on race, ethnicity, family income, parental education, parental co-residence, and intelligence test scores. We observe post-secondary education and college graduation as educational outcomes. Labor market outcomes include full-time employment status in 2000 and earnings in 1999. Respondents also report exercise and drinking habits in the final round of the survey.

The NELS:88 collects detailed information on sports participation, both at the individual and school levels. At the individual level, respondents select which sports teams they are affiliated with, as well as at which level (intramural, junior varsity, varsity, captain).[2] School administrators also indicate which sports programs (if any) are offered at the school. Furthermore, similar information is collected regarding other extracurricular activities such as performing arts and yearbook/newspaper staff.

2.3 Add Health

The Add Health surveyed a school-level sample of 20,728 students in grades 7-12 in 1995 from 52 middle schools and 80 high schools in the United States. The survey is ongoing and collected four waves as of 2008, with a fifth wave being collected in 2016-17. Wave I of the survey had separate parts that were conducted in school and at home. Later waves were conducted at home. Like the other surveys described previously, the Add Health collects information on personal and family background, sports participation, and later-life outcomes. Personal and family background is less detailed in the Add Health, but includes basic measures such as race, ethnicity, intelligence (measured by the Peabody Picture Vocabulary Test), parental education, and parental co-residence. Like the NELS:88, Add Health also includes measures of the student's school context, such as size of student body, urbanicity, and ethnicity of the student body. Educational outcomes are collected in Waves III and IV and include high school graduation, post-secondary college attendance, and four-year college graduation. As labor market outcomes we measure wages and full-time employment status in Wave IV. The Add Health also contains detailed information on health outcomes. We observe students' height and weight in each wave, from which we compute BMI. We also observe information on alcohol consumption and exercise habits, which are both collected in Wave IV.

High school sports participation was asked of all students in the Wave I in-school survey. Specifically, students were asked to select, from a list of activities, in which activities they

[2] Possible sports in the NELS:88 include baseball/softball, basketball, football, soccer, swim team, other team sports (hockey, volleyball, etc.), individual sports (cross-country, gymnastics, golf, tennis, track, wrestling, etc.) For simplicity and comparability to other surveys, we aggregate all sports into one group.

were "participating or planning to participate." In addition to a list of detailed sports programs, individuals were able to select non-athletic activities such as yearbook, student government, National Honor Society, performing arts, or foreign language/math clubs.[3]

2.4 Defining athletic participation

Due to the slightly different manner in which sports participation was elicited in each survey, we have slightly different definitions of who in each sample is an athlete.

In the NLSY79, we define athletes as those who report having participated in "athletics, cheerleading, or pep clubs" in high school. In contrast to the other surveys, respondents are not able to indicate which specific sports programs they participated in, nor are they able to designate intensity of participation (e.g. varsity versus junior varsity), or indicate that they were members of multiple sports teams. Thus, the wording on this question provides an overly broad definition of athlete. This is particularly problematic for women, who are much more likely to participate in cheerleading or pep club activities that may have much different environments for fostering human capital than extramural competitive sports.

In the NELS:88, we define athletes as those who report having participated at the junior varsity, varsity or team captain level in any of the possible athletic programs, based on responses during the 10th grade and the 12th grade surveys. That is, if a student reports sports participation in either wave at the junior varsity level or higher, we consider the student to be an athlete. This cohort provides the most precise definition of athletic participation.

We define athletes in the Add Health to be those who report participation in any of the athletic programs. Unlike the NELS:88, the Add Health collects this information just once and does not distinguish among various levels of competition. Students who plan to participate in sports programs respond to the survey in the same way as students who are already participating and so are also treated as athletes. Thus, the definition of athletic participation in the Add Health cohort is also overly broad: some who report "planning to participate" will not end up making the team.

As shown at the bottom of Tables 1 and 2, the level sports participation differs significantly between the surveys. For example, for boys, the sports participation rate is about 45 percent in the NLSY79, about 49 percent in the Add Health, but almost 70 percent in the NELS:88. Similarly, for girls, the sport participation rate is much higher in the NELS:88 (49 percent) compared to the NLSY79 (34 percent) and the Add Health (38 percent). However, when compared with the current level of sports participation reported by NFHS (56 percent) the overall NELS:88 sports participation rate (59 percent for both sexes) does not look out of line. One reason for the differences across surveys is due to oversampling of minorities in the

[3]Possible sports in the Add Health include baseball/softball, basketball, football, field hockey, ice hockey, soccer, swimming, tennis, track, volleyball, wrestling, or other. As with the NELS:88, we do not consider heterogeneity in the effect of different sports. We do this primarily to maintain comparability and simplicity in interpreting our results.

NLSY79 and the Add Health. However, it is also possible that the sample attrition among student athletes in the NELS:88 is somewhat lower than other students. Because we use information from the base year survey of 8th-graders, we are unable to include individuals who were added to the sample during succeeding waves to maintain representativeness of the sample.

2.5 Sample selection

This subsection briefly outlines our sample selection criteria. Additional details on sample selection are reported in Appendix Section A.

Our analysis of the NLSY79 focuses on all respondents who are not members of the disadvantaged white or military oversamples, and who were present in their interview at age 25. This leaves us with 4,837 men and 4,926 women. Sample sizes are slightly smaller for labor market outcomes due to sample attrition and selectivity of participation.

In the NELS:88, we restrict ourselves to students who were in school during the base year as well as the 10th grade and 12th grade surveys (even if they were not in the 10th grade or 12th grade at the time), and who participated in the final wave (in 2000). We also restrict our sample to those with parental background information and those who took the cognitive tests in the base year. Our final sample consists of 8,969 individuals—4,227 men and 4,742 women.

In the Add Health, we focus on those who completed the Wave I in-school questionnaire, who were aged 17 or younger in Wave I, and who were not missing certain health measures in Wave I. This leaves us with 11,263 observations—6,113 women and 5,150 men.

2.6 Descriptive statistics

In Table 1 and Table 2 we present basic descriptive statistics of athletes and non-athletes for men and women, respectively. Each table contains a list of outcomes and control variables from each survey, along with the respective sample means for non-athletes, athletes, and the full sample. We denote with an asterisk sample means of athletes that are statistically different from non-athletes at the 5% level. The descriptive statistics provide a high-level understanding of how selection on observables and unobservables might mitigate the effects of sports. We divide the variables into three categories: background characteristics, school characteristics, and outcomes.

Athletes tend to have higher cognitive test scores, be disproportionately white, have parents with higher levels of education, be more likely to co-reside with parents, and come from homes with higher incomes. In short, our basic summary statistics reveal that athletes are strongly positively selected on personal and family background traits.

On the school side, athletes are less likely to be absent from school, more likely to be found in private schools and schools with smaller student bodies, more likely to be found in rural schools, and more likely to attend schools that are more racially segregated. These results hold for both men and women, and are in line with existing literature and theory. Namely, overwhelmingly white, private, and rural schools provide more opportunities for student athletes, for whatever reason. Possible explanations include differences in school funding, or that it is statistically easier to make the team at a school with a smaller student body.

In addition to observing that athletes have different background and school contexts, we also observe that athletes have very different adult outcomes. They attain higher levels of education, measured either by grades completed or degrees attained. Athletes also earn more as adults: about 15 percent higher wages for men, and about 12 percent higher wages for women. Athletes are much more likely to report exercising regularly. Male athletes are neither more nor less likely to be obese as adults, while female athletes are much less likely to be obese. Athletes of both genders report a higher frequency of alcohol abuse as adults.

The results in Tables 1 and 2 are striking in that the different surveys exhibit not only the same sign of sports effects, but also many of the same magnitudes, in spite of the fact that athletic participation is measured quite differently across the three surveys.

3 Human capital theory and identification strategy

With basic descriptive results in hand, we now discuss reasons why high school sports may or may not be beneficial, and also review identification strategies pursued in previous studies and how our approach differs from them.

3.1 Theoretical effects of sports

Sports are thought to be beneficial to youth because they provide a forum to develop important skills whose development otherwise tends to be omitted from traditional education. These skills include teamwork, persistence, patience, time management, and leadership skills (Dick's Sporting Goods, 2017). Furthermore, sports participation may be beneficial by increasing access to higher education through athletic scholarships, keeping troubled youth "off the streets," matching youth with coaches who serve as mentors, and teaching youth proper health, physical fitness, and conditioning habits.

On the other hand, sports participation may be harmful to youth if it distracts too much from academic pursuits (Coleman, 1961), if it causes excess physical injury (Fair and Champa, 2017), or if it encourages youth to spend more time with peers who are less academically inclined or more prone to risky behavior.[4]

[4] Miller et al. (1999) find that male athletes report higher rates of sexual experience than non-athletes. Their analysis, however, is correlational.

While sports participation may have positive effects in some dimensions, it may also have negative effects in other dimensions (Leeds, 2015). A secondary contribution of our paper is to consider a broader set of outcomes to determine if there are any such negative effects.

3.2 Previous identification strategies

In Table 3, we list 21 studies that measure short- or long-run impacts of sports participation on social, health, or economic outcomes. Of these studies, roughly 45% implement some form of instrumental variables strategy, another 45% assume exogeneity of participation (i.e. selection on observables), and the remaining 10% use some other strategy (e.g. quantile techniques, sample selection correction, or hierarchical models such as individual fixed or random effects). Given our descriptive results in the previous section, it seems a far reach to assume that athletic participation is exogenous to personal and family background characteristics, let alone unobservable characteristics like motivation and determination. Below, we discuss in detail the various instruments used in the prior studies and why, with one exception, these are unlikely to be valid.

Height. The most popular instrument in the literature is height (Barron, Ewing, and Waddell, 2000; Eide and Ronan, 2001; Pfeifer and Cornelißen, 2010; Rees and Sabia, 2010; Yeung, 2015). The argument for using height as an instrument is that it is correlated with sports participation, but is assumed to not be correlated with educational or labor market outcomes. In other words, some students are randomly endowed with height (either from birth or via a well-timed growth spurt) and, as a result, are invited to join an athletic team. Their experience on the team then provides them with a set of skills that improve their educational and labor market outcomes, to which they would otherwise not have had access if they had happened to be shorter. The primary problem with using height as an instrument is that it independently affects the outcomes of interest, thus invalidating the excludability condition (Case and Paxson, 2008; Persico, Postlewaite, and Silverman, 2004).

School and peer characteristics. Another set of studies uses characteristics about students' peers' participation decisions or characteristics about the school, such as total enrollment, public/private status, library books per student, teacher-to-pupil ratio, or athletic program offerings to instrument for sports participation (Anderson, 1998, 2001; Barron, Ewing, and Waddell, 2000; Gorry, 2016). The intuition for this approach is that schools with larger student populations do not have more or larger athletic teams, so the opportunity of participating decreases with the size of the school. Alternatively, some schools of similar size my choose to offer many opportunities to participate in sports by sponsoring a larger number of teams or sports. The counterfactual comparison in this setting takes an individual at a small school, or a school with many athletic programs (where it is relatively easy to make the team) and compares his/her outcome with a similar individual at a large school or a school with few athletic programs (where it is more difficult to make the team). Similar to the case of height, this instrument is unlikely to satisfy the excludability condition because there are many differences in the communities surrounding large and small schools which are likely to affect future outcomes. Similarly, schools that offer many athletic programs are likely to differ in the types of academic programs that they offer. In other words, moving the same individual from a large school to a small school also involves moving them from an urban/suburban community to a rural one, or from a resource-rich school to a resource-poor one. For example, the quality of teachers may vary systematically with school size, even conditional on teacher-to-pupil ratio. Furthermore, private schools have been shown to affect future outcomes directly (Altonji, Elder, and Taber, 2005a). It is thus plausible that schools with more athletic programs per student are also better at fostering students' human capital in other ways that are unobservable to researchers.

Family background. Another potential instrument for high school sports participation is family background (Barron, Ewing, and Waddell, 2000). Some students may be interested in and qualify for the team, but cannot participate because of parental income or time constraints. This instrument is unlikely to satisfy the excludability condition, given that parental income is highly correlated with a number of variables that determine educational and labor market outcomes (e.g. parental education, single parenthood, parenting style, time investment, etc.).[5]

Geographic variation in the impact of Title IX regulations. The most convincing analysis to date uses a natural experiment that takes advantage of differences across states in the level of participation among boys prior to the 1972 passing of Title IX of the Educational Amendments to the 1964 Civil Rights Act, more commonly known as Title IX. The program mandated that sports participation among girls match that of boys, so that states with high levels of sports participation among boys were required to offer more sports opportunities for girls. From the standpoint of girls, living in a state that previously had relatively more participation among boys is plausibly random relative to future education or labor supply outcomes. Stevenson (2010) analyzes US Decennial Census microdata samples by comparing outcomes state-by-state of women who came of age just after Title IX with those of women who came of age just before Title IX, and adjusting for pre-Title IX levels of boys' participation and other observable differences. Stevenson concludes that increased sports participation through Title IX increases college attendance and labor force participation for women, though the sizes of the effects are small.

While we might expect the benefits of sports participation to be similar for boys and girls,

[5]See also Aucejo and James (2017), who show that differences in family background are an important determinant of gaps in educational attainment.

it is possible that the benefits differ by gender. This approach cannot measure the effect of sports participation for boys because the policy change did not expand sports opportunities for boys. Furthermore, the quasi-experimental expansion of sports opportunities for girls is a one-time event and these results could be tied to labor market conditions specific to the birth cohorts surrounding the policy change. Finally, the approach taken by Stevenson measures the impact of exposure to sports, not necessarily the impact of sports participation, since individual participation is not observed in Census microdata. Our approach examines heterogeneity in returns to sports by gender, and compares returns across three separate decades using data that identifies participation at the individual level.

3.3 Our econometric strategy

Rather than search for a plausibly exogenous instrument for high school sports participation, we implement a method developed by Krauth (2016) that allows us to directly test the degree of exogeneity in sports participation via sensitivity analysis around the OLS baseline estimate. Similar, but slightly different, methods have been developed by Altonji, Elder, and Taber (2005b); Millimet and Tchernis (2013); and Oster (Forthcoming). Krauth's method is based on the idea that the the researcher can replace the assumption of exogeneity with a weaker assumption that the ratio between selection on unobservables and selection on observables falls within some range. Thus our approach can also function as an indirect test of the validity of instrumental variables approaches used by prior studies on these same data sets.

Our econometric model is as follows:

$$y_i = \alpha d_i + X_i\beta + \varepsilon_i \tag{1}$$

where y_i is an outcome for student i, d_i indicates that i was a student athlete, X_i are observable characteristics of i, and ε_i are unobservable characteristics that determine the outcome, but which are not correlated with observable characteristics. α is a parameter that measures the impact of sports participation on the outcome of interest.

Importantly, we assume that d_i is correlated with ε_i. Typical empirical approaches would address this problem with experimental randomization or leveraging quasi-experimental policy variation. If neither of these approaches is feasible (as in the current setting), then the researcher must either assume exogeneity, or concede that the parameter of interest α is not identified.

Krauth's approach is to define a parameter λ which satisfies

$$\begin{aligned} \text{Corr}(d,\varepsilon) &= \lambda \text{Corr}(d, X\beta), \\ \lambda &= \frac{\text{Corr}(d,\varepsilon)}{\text{Corr}(d, X\beta)}. \end{aligned} \tag{2}$$

λ is referred to as the *relative correlation parameter* because it measures the degree of selection on unobservables relative to selection on observables. The case when $\lambda = 0$ implies that d is exogenously determined, while $\lambda = 1$ is analogous to the methodology of Altonji, Elder, and Taber (2005b). Krauth calls his methodology the *relative correlation restriction*, or RCR.

Krauth suggests two avenues for researchers implementing this methodology:

1. *Partial identification (bounds analysis):* Assume bounds on λ; i.e. assume $\lambda \in \left[\lambda^L, \lambda^H\right]$ and then estimate corresponding α's in the interval $\left[\alpha^L, \alpha^H\right]$.

2. *Sensitivity analysis:* Estimate α by OLS, then find the smallest (absolute) value of λ such that the OLS estimate is statistically zero. We label this parameter λ^*.

We consider both of these avenues in our analysis.

A valid question is, "What are reasonable values of λ^* such that a researcher could claim that the estimate of α represents a causal effect?" A simple baseline, introduced by Altonji, Elder, and Taber (2005b), is $\lambda = 1$, which implies that selection on unobservables is no larger than selection on observables. Altonji et al. (2013) suggest a useful way to think about this issue. They argue that only some of the variables that determine an outcome are available in a data set. This is due, in part, to the fact that the researcher is unaware of *all* of the variables that are relevant. But more importantly, it is costly to conduct a large survey. The types of data sets that we use have been collected to serve multiple purposes, not just to study the outcomes that we are interested in. Thus, the types of questions asked and the information collected are variables that are likely to be useful or interesting for a broad range of topics. As they state, "Because of limits on the number of the factors that we know matter, that we know how to collect, and that we can afford to collect, many ... [important explanatory variables]... are left out" (Altonji et al., 2013). Thus, they argue that it may be reasonable to think of the set of variables that are available to us in a data set as a random subset of all the variables that matter. In this case, we would expect that the correlation between d_i and the observables should be about the same as the correlation between d_i and the unobservables, which are contained in ε_i.

Of course, it is possible that the correlation with unobservables is even higher than the correlation with observables. Suppose that there is some character trait, such as competitiveness, that is not observable to the researcher. If that is a very strong determinant of success in later life, and is also a strong determinant of whether someone chooses to participate in sports, then the assumption that $\lambda = 1$ might be too restrictive—competitiveness is more important than the average observed variable, so $\lambda > 1$ is a more appropriate assumption. On the other hand, one could argue using a similar line of reasoning that $\lambda < 1$ is a more appropriate bound. For this to be the case, one would need to believe that the data collection process happened to disproportionately sample the most important variables. Thus, the average observed variable is more important than the average unobserved variable. While it may be the case that $\lambda > 1$ or $\lambda < 1$, we argue that binding λ within the unit interval is reasonable for the cases we examine here.

Other empirical examples verify $\lambda = 1$ as a reasonable limit. For example, Krauth (2016) tests how sensitive the Project STAR randomized experiment is to noncompliance in the treatment.[6] He estimates $\hat{\lambda}^* \approx 2$, where $\hat{\lambda}^*$ is the value for λ that would be consistent with a null effect of the treatment. Thus, notwithstanding experimental noncompliance, smaller class sizes have causal effects on test scores. On the other hand, Krauth (2016) presents a different example using observational data, which estimates $\hat{\lambda}^* \approx 0.1$, indicating that the measured effect is unlikely to be causal because the estimated effect is not robust to even small deviations from exogeneity. In the next section, we show that this latter case holds true for high school athletics participation across multiple nationally representative data sets for almost all of the outcomes that we examine.

4 Results

We now present and discuss our main empirical findings from the econometric model introduced in equations (1) and (2). For each of the surveys we analyze, we present impacts of high school sports participation on graduating from high school, attending college, graduating from college, full-time employment, wages or income, and exercise habits, obesity, and alcohol abuse. Each outcome is measured at approximately age 25 for each of the three survey cohorts. This corresponds to calendar years 1983-1990 for the NLSY79, calendar year 2000 for the NELS:88 (when respondents were aged 25-27), and calendar year 2008 for the Add Health (when respondents in our sample were aged 26-30). We present OLS estimates as well as estimates using the relative correlation restriction (RCR) method.

4.1 Educational outcomes

Tables 4 and 5 contain estimates of α, the effect of sports participation on educational outcomes, from all three surveys, respectively for men and women. The first two rows of the table report OLS estimates, while the remaining rows report RCR estimates. Consistent with the large body of prior research on the topic, as well as the evidence reported in Tables 1 and 2, when we assume that students are randomly assigned to be athletes, we find economically large and statistically significant effects of sports participation on high school graduation, college attendance, and college graduation.[7] Furthermore, the inclusion

[6] Specifically, some students who were assigned to a particular classroom were moved to a different classroom because of behavioral issues or parental request. For further details, see Krueger (1999) and Krauth (2016).

[7] We do not estimate the effect of sports participation on high school graduation for the NELS:88 sample, since our definition of sports participation requires that the respondent be in school during the 12th grade

of explanatory variables dramatically reduces the size of the estimated effect, by more than half in some cases. The estimated effects are slightly smaller for women, but the general pattern is very similar.

We now discuss the RCR estimates presented in Tables 4 and 5. Directly below the OLS estimates, we report the bounds on the treatment effect of sports participation when $\text{Corr}(\text{Sports}, \varepsilon)$ is no larger than $\text{Corr}(\text{Sports}, X\beta)$, i.e. for the range of between $\lambda = 0$ and $\lambda = 1$. Because of the positive selection on observables for these outcomes, the lower limit of the interval represents the case when $\lambda = 1$, and the upper limit corresponds with the OLS estimate. For example, for the outcome variable, "Graduate College," we estimate a range of -0.153 to 0.056 for the NLS79 sample of men. The lower bound corresponds to the assumption that the selection of unobservables is the same as the selection on observables. The line below gives a conservative confidence interval for our estimates—the lower bound of this confidence interval is the lower bound of the 95 percent confidence interval for the case of $\lambda = 1$.

Below the treatment effect bounds we list auxiliary parameters from the RCR procedure. These auxiliary parameters are meant to serve as helpful diagnostics in characterizing the level of selection on unobservables. Specifically, we report $\hat{\lambda}^{\infty}$, which is the value of λ at which identification breaks down, i.e. the value of λ that yields bounds on α equal to $(-\infty, \infty)$. Second, we report $\hat{\lambda}(0)$, which is the smallest value of λ such that the estimated treatment effect bounds include 0. Finally, we report $\hat{\lambda}^{*}$ a similar statistic as $\hat{\lambda}(0)$, but which instead corresponds to the smallest value of λ such that the 95% confidence interval of the bounds includes zero. Thus, $\hat{\lambda}(0)$ is analogous to a point estimate, while $\hat{\lambda}^{*}$ is analogous to a statistical significance test (e.g. a t-test) on that point estimate. Thus, for the confidence interval surrounding the treatment effect bounds to not include zero, we need $\lambda < \min\left(\hat{\lambda}^{*}, \hat{\lambda}(0), \hat{\lambda}^{\infty}\right)$. As each of the λ parameters are estimated with uncertainty, we also indicate significance levels on their corresponding z-tests.[8]

To further supply intuition for this approach, we include a graphical presentation of the RCR auxiliary parameters for college graduation among men in the NELS:88 in Figure 1, panel (a). The figure plots λ as a function of the treatment effect α. In other words, the figure plots the function defined in Equation (2): since β is implicitly a function of α (and hence ε is also implicitly a function of α), we have that λ as defined in Equation (2) is also a function of α so that one can recover the value of α that is consistent with, e.g. $\lambda = 1$.[9] The shaded region on the y-axis denotes the bounds on λ, which we assume to be $[0, 1]$ as discussed in the previous section. The shaded region on the x-axis denotes to the corresponding bounds on α, i.e. $\left[\alpha^{L}, \alpha^{H}\right]$. λ^{∞} is the horizontal dashed line, which marks the value of λ at which

interview.

[8]We do not report standard errors on the λ parameters so as to avoid overly burdening the reader. These are available from the authors upon request.

[9]Another way of thinking about this is that, in an OLS estimation problem, the parameters β will change if the researcher constrains α to be some value. Thus, β is implicitly a function of α.

identification breaks down (i.e. yields bound on α that are completely uninformative). While this line is the asymptote of the $\lambda(\alpha)$ function, $\lambda(\alpha)$ is not a hyperbola. For this reason, it is possible for $\hat{\lambda}(0) > \hat{\lambda}^{\infty}$. The vertical dashed line in the figure is α^{∞}, i.e. the value of α at which $\mathrm{Corr}(d, X\beta(\alpha)) = 0$, where we again emphasize that β is implicitly a function of α. For positively selected outcomes, an increase in λ^H results in a decrease in α^L. The opposite is true for negatively selected outcomes: increasing λ^H results in an increase in α^H. As a final note, panel (b) of Figure 1 shows that it is possible for $\lambda(0) > \lambda^{\infty}$. This is why $\lambda < \min\left(\hat{\lambda}^{*}, \hat{\lambda}(0), \hat{\lambda}^{\infty}\right)$ is the relevant condition for causality.

The RCR estimates in Tables 4 and 5 illustrate the sensitivity of sports to selection on unobservables. For most of the data sets and outcomes, if correlation of sports participation with the unobservables is even a small fraction of its correlation with the observed variables, the estimated impact of sports participation is nil. The exception is with the NELS:88 data. For the education outcomes we consider in these tables, we estimate that participation in sports leads to an increase in the probability of men attending college, even when $\lambda = 1$, although the estimated effect is small, and not statistically significant.[10] Also, our estimate of $\hat{\lambda}(0)$ for college graduation is also quite large at 0.89. Thus, the correlation between sports participation and college graduation propensity for men in the NELS:88 is robust to a large amount of selection on unobservables, although we would not reject a null hypothesis of a null effect if $\lambda = 1$. On the other hand, for women we find little evidence to support a causal effect of sports participation on the educational outcomes that we examine. For all three of these samples, even if the correlation with the unobservables is only half as much as the correlation with the observables, there is no beneficial effect of sports participation.

4.2 Labor market outcomes

Tables 6 and 7 contain estimates for log wages and full-time employment across all three surveys, respectively for men and women. Overall, we find a very similar pattern to what we observed for the educational outcomes: OLS estimates are large and statistically significant, and the inclusion of explanatory variables reduces this effect. One exception is the full-time work status for the NELS:88 and Add Health male samples, where the raw difference between athletes and non-athletes is essentially zero in each. Another exception is that for women's full-time work status in the NELS:88, where we document a negative selection on observables.

The RCR analysis confirms that even if the correlation between sports participation and the unobservables is only half of its correlation with the observables, we estimate a null effect of sports on wages and employment. The only exception to this is again in the

[10] Of course, if one believes that the true value of λ is less than 1 (perhaps because the most important characteristics determining college attendance are measured in the survey), then one could conclude that sports has a causal impact on men's college attendance.

NELS:88 sample of men, where we estimate that athletes earn about 3.3 percent more than non-athletes if we allow $\lambda = 1$. However, the estimate is imprecise and is not statistically significantly different from zero.

For several of the outcomes we examine in Tables 6 and 7 there is no statistically significant difference between athletes and non-athletes once we have controlled for observables. In these cases, there is no reason to apply the RCR estimation, since there is nothing left for the unobservables to explain. We thus report "N/A" for the corresponding RCR estimates.

Our estimates for women's labor supply in Table 7 includes one noteworthy result. In the NELS:88 the observables are actually negatively correlated with the full-time employment outcome. That is, after controlling for observable characteristics, the estimated effect of athletics participation is actually *higher* than the raw difference. So in this case, the bounds for the estimated effect contain only positive numbers, and we can reject the null hypothesis of no effect. However, a lack of agreement from the other two surveys—in which there is no significant difference in employment after controlling for observables—makes us reluctant to conclude that there is a causal effect of sports on women's adult full-time employment in any direction. We discuss at the end of this section possible explanations for the NELS:88 results being different.

4.3 Health and risky behaviors

Tables 8 and 9 contain estimates for our measures of health and risky behaviors: regular exercise, obesity, and alcohol abuse. We base our definition of regular exercise on the responses to questions about physical fitness/exercise activities during the week prior to the interview (for Add Health) or for a typical week (NELS:88). If the respondent reported participating in exercise activities during three or more days per week, we assigned a value of 1 to the regular exercise variable. The NLS79 does not have a variable to measure exercise activities. To indicate alcohol abuse in the NELS:88 cohort, we use responses to questions about how frequently the respondent participates in binge drinking, defined as drinking five or more alcoholic drinks in a row. Those who reported an episode of binge drinking during the past month were assigned a value of 1 for the alcohol abuse variable. For Add Health, we define alcohol abuse as being drunk 25 times or more in the previous year. The NLSY79 does not measure any kind of drug or alcohol abuse during adulthood. Finally, we also measure obesity in both the NLSY79 and the Add Health. This is derived from reported values for height and weight in these two surveys. We calculate the Body Mass Index (BMI) based on the respondent's height and weight.[11] Those with a BMI in the obese range (i.e. 30 or larger) are indicated as obese.[12]

[11]Height and weight are self-reported at various interviews in the NLSY79. In the Add Health, they are collected in each wave by a medical professional.

[12]This definition of obesity is taken from the Centers for Disease Control and Prevention (CDC). See https://www.cdc.gov/obesity/adult/defining.html for more information.

In the case of regular exercise, both surveys indicate that male athletes are more likely to participate in regular adult exercise. Interestingly, after controlling for observables, the athlete/non-athlete difference does not change much (and actually *increases* in the Add Health sample). This result is graphically depicted in Figure 1 (b) for the NELS:88.

In the Add Health sample, we note that, for all non-education outcomes, the treatment effect bounds are nearly completely uninformative. This is because our estimated $\hat{\lambda}^{\infty}$ parameter is much closer to 1 than in the other surveys. As depicted in panel (b) of Figure 1, the λ curves are flat near λ^{∞}, and the corresponding bounds on α are wide as a result. Regular exercise for women in both of these surveys is consistent with unobservables confounding the impact of sports participation.

Tables 1 and 2 show that female athletes have lower obesity rates than non-athletes, but that male athletes in the NLSY79 are actually more obese than non-athletes, with men in the Add Health showing no correlation between athletics and obesity. Our RCR analysis of the male NLSY79 sample shows that adding additional controls *increases* this effect, similar to the case of regular exercise. In the next section, we discuss possible reasons for conflicting results across each of our surveys. As with regular exercise, female athletes' apparent lower obesity rates are mediated by unobservables.

Alcohol abuse is much higher for athletes in the NELS:88 sample for both men and women, and slightly higher for women in the Add Health, as reported in Tables 1 and 2. For women in the NELS:88, this difference disappears once we control for observable characteristics, but it persists for men. If we fix $\lambda = 1$, we estimate a one-half percentage point higher rate for male athletes, but this is not estimated with enough precision to reject the null of no effect at the 5 percent significance level. Women in the Add Health exhibit negative selection on observables, although the effect is small compared to the effect for men in the NELS:88.

4.4 Discussion

With few exceptions, our results are qualitatively similar across all three surveys. The only cases in which we estimate a treatment effect that is statistically significant for λ well above zero are the cases when there appear to be negative selection on observables. Overall, our results agree with Barron, Ewing, and Waddell (2000) who note that much of the differences in outcomes for athletes reflect differences in ability or preferences for leisure.

We now briefly discuss reasons for why one might expect our findings to be quantitatively different across the three cohorts we study. The two primary reasons for this could be *(i)* the effects are different in the different time periods analyzed by the different surveys, or *(ii)* the estimates in the different surveys reflect differences in how treatment or outcomes are measured.

First, the causal effect of sports participation on later-life outcomes may be different at different times for reasons that are specific to that time period. Each of our three cohorts is separated by roughly ten years. If there is an interaction between sports participation and labor market conditions at a certain point in time, this could explain quantitative differences in our findings, although it seems unlikely that the relative value of skills acquired by athletes would yield much different returns across time.

Second, and more likely, the heterogeneity in our measured effects could be induced by differences across the surveys in how athletes are measured, i.e. differences in treatment intensity, or how outcomes are measured. For example, in the NLSY79, respondents report if they participated in athletics, cheerleading, or pep club. Similarly, in the Add Health, respondents answer the question "Are you participating/Do you plan to participate in the following clubs, organizations and teams? (check all that apply)." Combined with the fact that many respondents are under age 16 when answering this question, it is unclear if athlete status in the Add Health reflects a desire to participate, or actual participation. In contrast, we observe treatment much more clearly in the NELS:88 because we track athlete status for each individual across grades 10-12. It is possible that the RCR auxiliary parameter estimates for $\hat{\lambda}(0)$ and $\hat{\lambda}^*$ tend to be larger in magnitude for this set of data for this reason. On a similar note, obesity is collected differently across the NLSY79 and Add Health surveys, which may contribute to the opposing findings for men to the extent that they systematically misreport their height or weight.

In results not reported, but available from the authors upon request, we examine how treatment intensity affects our results by examining students who report participating in more than one sport (in the Add Health), or who report that athletics was the club they "participated in most actively" (in the NLSY79). The results for this more intense definition of treatment are generally stronger than our baseline measure. For example, $\hat{\lambda}(0)$ tends to be larger for the more intense definition. However, none of our conclusions is changed. That is, $\hat{\lambda}(0)$ does not become large enough to reject that sports have a causal impact on later-life outcomes.

As further exploration of treatment intensity, we estimate RCR bounds on the NELS:88 sample where we redefine treatment to be varsity athletic status or varsity captain status. The results in each of these cases are very similar to the baseline definition, and none of our conclusions change.

We also examine treatment effect heterogeneity in the NELS:88 by estimating separate specifications where we interact African American status with athlete status (Eide and Ronan, 2001).[13] We find very little heterogeneity in this dimension, even in specifications that assume selection operates only on observables. Thus, the RCR results are unchanged.

[13] Gorry (2016) examines heterogeneity in the effect of sports by gender and socioeconomic status. She does not find a statistically significant causal impact of sports on labor market outcomes, but does find a positive causal effect of sports on academic performance, particularly for low-achieving individuals. She uses instrumental variables to identify causality.

Additionally, we conduct a separate analysis where we include only individuals who were 15, 16, or 17 in Wave I of the Add Health so as to reduce measurement error in athletics. Our hypothesis is that younger individuals in the Add Health may be participating in sports during middle school, which is a less competitive environment. This additional analysis reveals similar findings as the multi-sport athlete analysis.

5 Conclusion

We revisit the literature on the long-run effects of high school sports participation on educational attainment, labor market outcomes, and adult health behaviors. Many previous studies have found positive effects in each of these dimensions by either assuming that sports participation is exogenous (conditional on other observable characteristics), or by making use of instrumental variables that are unlikely to be valid.

We analyze three separate nationally representative longitudinal surveys that link participation in high school sports with later-life outcomes: the NLSY79, the NELS:88, and the Add Health. We employ an econometric technique that empirically tests the sensitivity of the selection on observables assumption and find that estimates of the returns to sports participation are highly sensitive to this assumption. Specifically, we find that, for most educational and labor market outcomes, if the correlation between sports participation and unobservables is only a fraction of the correlation between sports and observables, the effect of sports participation cannot be statistically differentiated from zero. Thus, we conclude that a causal effect of sports participation is unlikely, and that most of the findings of the literature that report beneficial impacts represent the effect of selection into sports.[14]

There are two exceptions to this general statement: For men in the NELS:88 cohort, there is weak evidence that sports participation increases college attendance and graduation, although we cannot reject a null effect at the 5 percent significance level if $\lambda = 1$. We also find that female athletes of the NELS:88 cohort are slightly more likely to work full time.

Our analysis of health benefits of high school sports is also quite weak. However, results based on male samples from both the NELS:88 and the Add Health indicate a higher rate of regular adult exercise, and this effect is statistically significant and rather large. This is the only outcome for which we see a statistically significant impact in two different cohorts. Curiously, we also find a small, statistically significant increase in obesity for men in the NLSY79 sample. One possible explanation for this is a side effect of high-intensity weight training for male-only sports such as football.

Our largely null results inform the policy debate on high school sports by providing evidence against claims that sports foster skills that improve educational or labor market outcomes. However, despite having very little human capital value, sports may still have a

[14] An important exception is Stevenson (2010), who convincingly leverages a natural experiment, but whose results are small in magnitude and may not generalize to other contexts.

place in high school as a social or cultural activity. We generally confirm the assertion of sports commentator Heywood Hale Broun, who said, “Sports reveals character, it doesn’t build it” (Phillips, May 12, 1974).

References

Altonji, Joseph G., Timothy Conley, Todd E. Elder, and Christopher R. Taber. 2013. "Methods for Using Selection on Observed Variables to Address Selection on Unobserved Variables." Working paper.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005a. "An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schooling." *Journal of Human Resources* 40 (4):791–821.

———. 2005b. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1):151–184.

Anderson, Deborah J. 1998. *"If You Let Me Play": The Effects of Participation in High School Athletics on Students' Behavior and Economic Success*. Ph.D. thesis, Cornell University.

———. 2001. "'If You Let Me Play': The Effects of Participation in High School Athletics on Students' Educational and Labor Market Success." Working paper, University of New Mexico.

Aucejo, Esteban M. and Jonathan James. 2017. "Catching Up to Girls: Understanding the Gender Imbalance in Educational Attainment Within Race." Working paper, Arizona State University and Cal Poly SLO.

Barron, John M., Bradley T. Ewing, and Glen R. Waddell. 2000. "The Effects of High School Athletic Participation on Education and Labor Market Outcomes." *Review of Economics and Statistics* 82 (3):409–421.

Case, Anne and Christina Paxson. 2008. "Stature and Status: Height, Ability, and Labor Market Outcomes." *Journal of Political Economy* 116 (3):499–532.

Clarke, Phoebe and Ian Ayres. 2014. "The Chastain Effect: Using Title IX to Measure the Causal Effect of Participating in High School Sports on Adult Women's Social Lives." *Journal of Socio-Economics* 48:62–71.

Coleman, James. 1961. *The Adolescent Society*. Glencoe, Ill.: Free Press.

Darling, Nancy, Linda L. Caldwell, and Robert Smith. 2005. "Participation in School-based Extracurricular Activities and Adolescent Adjustment." *Journal of Leisure Research* 37 (1):51.

Dick's Sporting Goods. 2017. "Sports Matter." URL http://www.sportsmatter.com/. [Online; accessed 20-February-2017].

Eccles, Jacquelynne S. and Bonnie L. Barber. 1999. "Student Council, Volunteering, Basketball, or Marching Band: What Kind of Extracurricular Involvement Matters?" *Journal of Adolescent Research* 14 (1):10–43.

Eide, Eric R. and Nick Ronan. 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.

Ewing, Bradley T. 1998. "Athletes and Work." *Economics Letters* 59 (1):113–117.

———. 2007. "The Labor Market Effects of High School Athletic Participation: Evidence From Wage and Fringe Benefit Differentials." *Journal of Sports Economics* 8 (3):255–265.

Fair, Ray and Christopher Champa. 2017. "Estimated Costs of Contact in Men's Collegiate Sports." Discussion Paper 3001, Cowles Foundation.

Gorry, Devon. 2016. "Heterogenous Effects of Sports Participation on Education and Labor Market Outcomes." *Education Economics* 24 (6):622–638.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (1):153–161.

Krauth, Brian. 2016. "Bounding a Linear Causal Effect Using Relative Correlation Restrictions." *Journal of Econometric Methods* 5 (1):117–141.

Krueger, Alan B. 1999. "Experimental Estimates of Education Production Functions." *Quarterly Journal of Economics* 114 (2):497–532.

Leeds, Michael A. 2015. "Youth Sports and the Accumulation of Human Capital." *IZA World of Labor* 129:1–10.

Leeds, Michael A., Cristen Miller, and Judith Stull. 2007. "Interscholastic Athletics and Investment in Human Capital." *Social Science Quarterly* 88 (3):729–744.

Lipscomb, Stephen. 2007. "Secondary School Extracurricular Involvement and Academic Achievement: A Fixed Effects Approach." *Economics of Education Review* 26 (4):463–472.

Miller, Kathleen E., Michael P. Farrell, Donald F. Sabo, Grace M. Barnes, and Merrill J. Melnick. 1999. "Sports, Sexual Behavior, Contraceptive Use, and Pregnancy among Female and Male High School Students: Testing Cultural Resource Theory." *Sociology of Sport Journal* 16 (4):366–387.

Millimet, Daniel L. and Rusty Tchernis. 2013. “Estimation of Treatment Effects Without an Exclusion Restriction: With an Application to the Analysis of the School Breakfast Program.” *Journal of Applied Econometrics* 28 (6):982–1017.

Millimet, Daniel L., Rusty Tchernis, and Muna Husain. 2010. “School Nutrition Programs and the Incidence of Childhood Obesity.” *Journal of Human Resources* 45 (3):640–654.

National Federation of State High School Associations. 2017. “High School Sports Participation Increases for 27th Consecutive Year.” URL http://www.nfhs.org/articles/high-school-sports-participation-increases-for-27th-consecutive-year/. [Online; accessed 20-February-2017].

Oster, Emily. Forthcoming. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business & Economic Statistics* .

Persico, Nicola, Andrew Postlewaite, and Dan Silverman. 2004. “The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height.” *Journal of Political Economy* 112 (5):1019–1053.

Pfeifer, Christian and Thomas Cornelißen. 2010. “The Impact of Participation in Sports on Educational Attainment—New Evidence from Germany.” *Economics of Education Review* 29 (1):94–103.

Phillips, Barbara. May 12, 1974. “He Wows Them at K-Wesleyan.” *The Salina Journal* URL https://www.newspapers.com/image/43981481/?terms=salina%2Bjournal%2Bhe%2Bwows%2Bthem. Accessed October 18, 2017.

Rees, Daniel I. and Joseph J. Sabia. 2010. “Sports Participation and Academic Performance: Evidence from the National Longitudinal Study of Adolescent Health.” *Economics of Education Review* 29 (5):751–759.

Rehberg, Richard A. and Walter E. Schafer. 1968. “Participation in Interscholastic Athletics and College Expectations.” *American Journal of Sociology* 73 (6):732–740.

Sabo, Donald, Merrill J. Melnick, and Beth E. Vanfossen. 1993. “High School Athletic Participation and Postsecondary Educational and Occupational Mobility: A Focus on Race and Gender.” *Sociology of Sport Journal* 10 (1):44–56.

Spreitzer, Elmer and Meredith Pugh. 1973. “Interscholastic Athletics and Educational Expectations.” *Sociology of Education* 46 (2):171–182.

Stevenson, Betsey. 2010. “Beyond the Classroom: Using Title IX to Measure the Return to High School Sports.” *Review of Economics and Statistics* 92 (2):284–301.

Troutman, Kelly P. and Mikaela J. Dufur. 2007. "From High School Jocks to College Grads: Assessing the Long-Term Effects of High School Sport Participation on Females' Educational Attainment." *Youth & Society* 38 (4):443–462.

Up 2 Us Sports. 2017. "Even the Score." URL https://www.up2us.org/solution. [Online; accessed 20-February-2017].

Videon, Tami M. 2002. "Who Plays and Who Benefits: Gender, Interscholastic Athletics, and Academic Outcomes." *Sociological Perspectives* 45 (4):415–444.

Yeung, Ryan. 2015. "Athletics, Athletic Leadership, and Academic Achievement." *Education and Urban Society* 47 (3):361–387.

Figures and Tables

Table 1: Summary stats of outcome and control variables by athlete status, men

	NLSY79			NELS:88			Add Health		
	Non-athlete	Athlete	Overall	Non-athlete	Athlete	Overall	Non-athlete	Athlete	Overall
Cognitive score	-0.35	0.30*	-0.06	65.25	66.60*	66.20	0.09	0.27*	0.18
	(1.05)	(0.97)	(1.06)	(19.02)	(19.09)	(19.07)	(0.98)	(0.91)	(0.95)
Non-cognitive score	8.85	8.47*	8.68	—	—	—	—	—	—
	(2.40)	(2.39)	(2.40)	(—)	(—)	(—)	(—)	(—)	(—)
White	46.23	56.07*	50.70	72.16	73.49	73.10	52.21	56.70*	54.39
Black	30.70	29.16	30.00	7.08	7.91	7.67	18.88	20.43	19.63
Hispanic	23.07	14.76*	19.30	11.67	10.76	11.02	18.12	11.96*	15.13
Other	—	—	—	9.09	7.84	8.21	10.80	10.92	10.85
Mother's years of education	10.23	11.66*	10.89	—	—	—	—	—	—
	(3.48)	(2.94)	(3.32)	(—)	(—)	(—)	(—)	(—)	(—)
Mother HS dropout	—	—	—	—	—	—	14.16	9.60*	11.94
Mother HS grad	—	—	—	—	—	—	24.80	25.51	25.15
Mother Some college	—	—	—	—	—	—	25.56	24.71	25.15
Mother 4-year college grad	—	—	—	20.03	29.83*	26.95	11.48	15.03*	13.20
Mother Advanced degree	—	—	—	—	—	—	6.83	9.68*	8.21
missing Mother's education	8.78	5.23*	7.17	—	—	—	17.18	15.47	16.35
Father's years of education	10.22	11.94*	11.03	—	—	—	—	—	—
	(4.13)	(3.73)	(4.04)	(—)	(—)	(—)	(—)	(—)	(—)
Father HS dropout	—	—	—	—	—	—	12.00	8.56*	10.33
Father HS grad	—	—	—	—	—	—	18.01	19.31	18.64
Father Some college	—	—	—	—	—	—	17.18	20.67*	18.87
Father 4-year college grad	—	—	—	22.93	35.42*	31.75	10.04	12.00*	10.99
Father Advanced degree	—	—	—	—	—	—	7.51	10.52*	8.97
missing Father's education	17.74	11.58*	14.94	—	—	—	35.26	28.95*	32.19
Maternal co-residence	92.41	93.59	92.95	92.28	94.50*	93.85	44.55	89.16*	66.21
missing Maternal co-residence	—	—	—	—	—	—	50.85	2.92*	27.57
Paternal co-residence	—	—	—	77.55	79.66	79.04	36.54	76.61*	56.00
missing Paternal co-residence	—	—	—	—	—	—	50.92	3.56*	27.92
log family income	9.72	10.02*	9.85	—	—	—	—	—	—
	(1.16)	(0.97)	(1.09)	(—)	(—)	(—)	(—)	(—)	(—)
Family income < 20k	—	—	—	22.20	17.33*	18.76	—	—	—
Family income > 50k	—	—	—	26.07	34.25*	31.84	—	—	—
missing family income	18.76	18.66	18.72	—	—	—	—	—	—
Frequently absent	—	—	—	5.39	4.26	4.59	—	—	—
Has handicap	—	—	—	2.57	1.84	2.06	—	—	—
School size decile 1	—	—	—	8.37	14.08*	12.40	10.86	12.83*	11.84
School size decile 2	—	—	—	9.25	13.74*	12.42	10.09	12.63*	11.34
School size decile 3	—	—	—	9.98	11.09	10.76	11.64	12.12	11.88
School size decile 4	—	—	—	12.55	13.77	13.41	11.37	11.48	11.42
School size decile 5	—	—	—	13.68	11.83	12.37	8.10	10.80*	9.43
School size decile 6	—	—	—	16.65	14.01*	14.79	8.84	11.40*	10.10
School size decile 7	—	—	—	13.44	9.92*	10.95	10.67	9.16	9.92
School size decile 8	—	—	—	7.80	5.86*	6.43	13.36	10.84*	12.11
School size decile 9	—	—	—	8.29	5.70*	6.46	5.33	3.56*	4.46
School size decile 10	—	—	—	—	—	—	9.74	5.20*	7.50
Public school	—	—	—	90.27	82.61*	84.86	—	—	—
Catholic school	—	—	—	6.11	8.41*	7.74	—	—	—
Non-Catholic religious school	—	—	—	1.37	2.58*	2.22	—	—	—
Non-religious private school	—	—	—	2.25	6.40*	5.18	—	—	—
Urbanicity tercile 3	—	—	—	27.92	27.18	27.40	30.15	26.10*	28.15
Urbanicity tercile 2	—	—	—	39.10	41.69	40.93	54.84	54.21	54.53
Urbanicity tercile 1	—	—	—	32.98	30.97	31.56	15.01	19.69*	17.32
0% White	—	—	—	—	—	—	8.27	14.11*	11.16
1%-66% White	—	—	—	—	—	—	42.06	32.60*	37.38
67%-93% White	—	—	—	—	—	—	30.03	27.87	28.96
94%-100% White	—	—	—	—	—	—	19.64	25.42*	22.50
Experience	3.41	3.29	3.35	—	—	—	9.11	8.54*	8.84
	(2.01)	(1.91)	(1.97)	(—)	(—)	(—)	(2.96)	(2.98)	(2.99)
Tenure	—	—	—	—	—	—	3.06	3.06	3.06
	(—)	(—)	(—)	(—)	(—)	(—)	(2.99)	(2.87)	(2.93)
Years of education	11.97	13.49*	12.66	—	—	—	13.57	14.32*	13.93
	(2.30)	(2.15)	(2.36)	(—)	(—)	(—)	(2.18)	(2.17)	(2.21)
Graduated high school	78.25	94.57*	85.66	—	—	—	89.50	95.04*	92.19
Attended college	24.78	52.28*	37.27	77.39	86.39*	83.75	54.72	69.85*	62.07
Graduated 4-year college	9.64	25.73*	16.95	26.52	42.29*	37.65	34.94	44.36*	39.99
Log wage	1.80	1.98*	1.88	10.07	10.21*	10.17	2.61	2.76*	2.69
	(0.49)	(0.49)	(0.50)	(0.72)	(0.69)	(0.70)	(0.78)	(0.73)	(0.76)
Employed full-time	56.63	62.74*	59.40	83.91	84.62	84.41	88.04	91.00*	89.48
Exercise regularly	—	—	—	52.62	69.73*	64.70	71.24	78.99*	75.00
Obese	4.90	6.05	5.42	—	—	—	29.54	28.86	29.21
Alcohol abuse	—	—	—	30.99	42.97*	39.45	20.60	22.10	21.33
N	2,345	1,951	4,296	1,243	2,984	4,227	2,649	2,501	5,150

Notes: Standard deviation below continuous variables in parentheses. * indicates significantly different means between athletes and non-athletes at the 5% level. Experience refers to actual experience in the NLSY79, but potential experience in the Add Health. College attendance is higher in the NELS due to the fact that the sample is restricted to students in the already in the 12th grade. For this reason we also don't consider high school graduation as an outcome in the NELS. Work experience is not reliably measured in the NELS.

Table 2: Summary stats of outcome and control variables by athlete status, women

	NLSY79			NELS:88			Add Health		
	Non-athlete	Athlete	Overall	Non-athlete	Athlete	Overall	Non-athlete	Athlete	Overall
Cognitive score	-0.23	0.32*	-0.04	64.15	69.35*	66.72	-0.03	0.19*	0.05
	(0.97)	(0.88)	(0.97)	(18.21)	(18.66)	(18.61)	(0.97)	(0.91)	(0.95)
Non-cognitive score	8.97	8.48*	8.81	—	—	—	—	—	—
	(2.38)	(2.41)	(2.40)	(—)	(—)	(—)	(—)	(—)	(—)
White	44.88	59.72*	49.93	67.27	75.37*	71.28	49.38	58.71*	52.92
Black	32.37	26.42*	30.35	10.77	7.33*	9.07	22.88	20.99	22.17
Hispanic	22.74	13.86*	19.72	14.03	8.99*	11.54	17.48	10.56*	14.85
Other	—	—	—	7.93	8.31	8.12	10.26	9.74	10.06
Mother's years of education	10.23	11.66*	10.73	—	—	—	—	—	—
	(3.39)	(2.81)	(3.27)	(—)	(—)	(—)	(—)	(—)	(—)
Mother HS dropout	—	—	—	—	—	—	16.79	10.09*	14.25
Mother HS grad	—	—	—	—	—	—	24.70	25.95	25.18
Mother Some college	—	—	—	—	—	—	22.86	25.69*	23.93
Mother 4-year college grad	—	—	—	16.20	29.48*	22.78	10.86	14.78*	12.35
Mother Advanced degree	—	—	—	—	—	—	7.54	9.22*	8.18
missing Mother's education	6.96	3.11*	5.65	—	—	—	17.24	14.27*	16.11
Father's years of education	10.24	11.93*	10.85	—	—	—	—	—	—
	(4.01)	(3.69)	(3.98)	(—)	(—)	(—)	(—)	(—)	(—)
Father HS dropout	—	—	—	—	—	—	12.15	7.67*	10.45
Father HS grad	—	—	—	—	—	—	18.96	19.05	18.99
Father Some college	—	—	—	—	—	—	14.87	18.02*	16.06
Father 4-year college grad	—	—	—	20.92	34.60*	27.69	8.78	11.34*	9.75
Father Advanced degree	—	—	—	—	—	—	6.67	10.43*	8.10
missing Father's education	17.76	10.30*	15.22	—	—	—	38.57	33.49*	36.64
Maternal co-residence	91.94	95.01*	92.99	94.86	95.19	95.02	58.56	94.40*	72.16
missing Maternal co-residence	—	—	—	—	—	—	38.02	1.42*	24.13
Paternal co-residence	—	—	—	72.53	79.38*	75.92	44.87	77.24*	57.16
missing Paternal co-residence	—	—	—	—	—	—	38.41	1.68*	24.47
log family income	9.65	9.96*	9.76	—	—	—	—	—	—
	(1.20)	(1.06)	(1.16)	(—)	(—)	(—)	(—)	(—)	(—)
Family income < 20k	—	—	—	26.51	17.09*	21.85	—	—	—
Family income > 50k	—	—	—	24.47	36.90*	30.62	—	—	—
missing family income	21.67	18.85*	20.71	—	—	—	—	—	—
Frequently absent	—	—	—	7.35	4.35*	5.86	—	—	—
Has handicap	—	—	—	1.42	1.53	1.48	—	—	—
School size decile 1	—	—	—	9.73	16.40*	13.03	10.22	15.78*	12.36
School size decile 2	—	—	—	10.23	14.40*	12.29	10.33	14.18*	11.82
School size decile 3	—	—	—	9.85	12.82*	11.32	11.39	11.98	11.62
School size decile 4	—	—	—	11.27	11.16	11.22	10.79	10.30	10.60
School size decile 5	—	—	—	14.74	12.31*	13.54	9.14	11.98*	10.24
School size decile 6	—	—	—	16.24	13.59*	14.93	8.74	10.73*	9.51
School size decile 7	—	—	—	12.53	9.54*	11.05	9.93	8.10*	9.22
School size decile 8	—	—	—	8.06	4.94*	6.52	13.88	7.72*	11.50
School size decile 9	—	—	—	7.35	4.81*	6.09	5.92	4.35*	5.32
School size decile 10	—	—	—	—	—	—	9.66	4.87*	7.81
Public school	—	—	—	90.35	81.93*	86.19	—	—	—
Catholic school	—	—	—	6.14	8.35*	7.23	—	—	—
Non-Catholic religious school	—	—	—	1.71	3.41*	2.55	—	—	—
Non-religious private school	—	—	—	1.80	6.31*	4.03	—	—	—
Urbanicity tercile 3	—	—	—	28.64	26.63	27.65	29.66	30.29	29.91
Urbanicity tercile 2	—	—	—	38.91	41.54	40.22	55.33	50.09*	53.30
Urbanicity tercile 1	—	—	—	32.32	31.49	31.91	15.01	19.62*	16.80
0% White	—	—	—	—	—	—	7.90	14.87*	10.60
1%-66% White	—	—	—	—	—	—	44.78	30.25*	39.15
67%-93% White	—	—	—	—	—	—	28.84	28.18	28.58
94%-100% White	—	—	—	—	—	—	18.48	26.71*	21.67
Experience	2.63	3.11*	2.79	—	—	—	8.63	8.00*	8.39
	(2.03)	(1.81)	(1.97)	(—)	(—)	(—)	(2.98)	(2.91)	(2.97)
Tenure	—	—	—	—	—	—	2.98	2.82*	2.91
	(—)	(—)	(—)	(—)	(—)	(—)	(2.76)	(2.53)	(2.67)
Years of education	12.32	13.59*	12.75	—	—	—	14.12	14.89*	14.41
	(2.28)	(2.01)	(2.27)	(—)	(—)	(—)	(2.22)	(2.20)	(2.24)
Graduated high school	85.08	97.47*	89.30	—	—	—	92.17	96.90*	93.96
Attended college	33.65	55.12*	40.96	84.59	91.22*	87.87	65.78	78.53*	70.62
Graduated 4-year college	11.74	26.23*	16.67	33.71	52.14*	42.85	37.70	49.85*	42.71
Log wage	1.68	1.79*	1.72	9.77	9.90*	9.83	2.39	2.52*	2.44
	(0.47)	(0.46)	(0.47)	(0.78)	(0.83)	(0.80)	(1.00)	(0.94)	(0.98)
Employed full-time	40.27	48.32*	43.01	70.69	73.36*	72.01	78.96	82.67*	80.38
Exercise regularly	—	—	—	53.95	65.29*	59.56	66.90	73.59*	69.44
Obese	6.12	3.82*	5.34	—	—	—	35.69	30.21*	33.60
Alcohol abuse	—	—	—	13.37	18.69*	16.00	8.80	10.49*	9.44
N	2,990	1,544	4,534	2,395	2,347	4,742	3,793	2,320	6,113

Notes: See notes to Table 1.

Table 3: Previous studies of the effects of high school sports on various outcomes

Study	Id. Strategy	Outcomes	Causal Findings
Anderson (1998, 2001)	IV: Peer participation, availability of extracurricular substitutes	Education, wages	Mixed evidence
Barron, Ewing, and Waddell (2000)	IV: School size, private status, health, family background	Education, wages	Little to no effect
Clarke and Ayres (2014)	IV: Cross-state timing differences	Social outcomes: secularism, motherhood, single motherhood	Positive effects
Darling, Caldwell, and Smith (2005)	Selection on observables	Education expectations, GPA	Positive effects
Eccles and Barber (1999)	Selection on observables	Education, risky behaviors	Positive effects
Eide and Ronan (2001)	IV: Height	Education, wages	Positive effect for black men, white women; neg. effect for white men
Ewing (1998)	Selection on observables	Job quality	Positive effects
Ewing (2007)	Selection on observables	Wages, fringe benefits	Positive effects
Gorry (2016)	IV: School size, private status; quantile regression	GPA, HS graduation, employment, earnings, welfare receipt	No effect
Leeds, Miller, and Stull (2007)	Heckman (1979) selection	Time spent studying	Mixed effects
Lipscomb (2007)	Fixed effects	Test scores, education	Positive effects
Miller et al. (1999)	Selection on observables	Sexual behavior	Mixed effects
Pfeifer and Cornelißen (2010)	IV; Height, city size	HS, college graduation	Positive effects
Rehberg and Schafer (1968)	Selection on observables	College expectations	Positive effects
Rees and Sabia (2010)	Fixed effects; IV: Height	GPA, attention to academics, college expectations	No effect
Sabo, Melnick, and Vanfossen (1993)	Selection on observables	College graduation, occupational prestige	Positive effects
Spreitzer and Pugh (1973)	Selection on observables	College expectations	Positive effects
Stevenson (2010)	IV: Cross-state differences in male athletic particip.	College attendance, labor supply	Positive effects
Troutman and Dufur (2007)	Random effects logit	College graduation	Positive effects
Videon (2002)	Selection on observables	GPA, college expectations, absences	Positive effects
Yeung (2015)	IV: height	Test scores	Positive effects

Notes: "Id. Strategy" stands for "identification strategy," and indicates the empirical method used to assert causality. "IV" stands for "instrumental variables."

Table 4: Effect of sports on educational outcomes for men

	Graduate HS		Attend college			Graduate college		
	NLSY79	Add Health	NLSY79	NELS:88	Add Health	NLSY79	NELS:88	Add Health
OLS, no controls	.163***	.055***	.275***	.090***	.149***	.161***	.158***	.096***
	(.010)	(.007)	(.014)	(.014)	(.018)	(.011)	(.016)	(.022)
OLS, full controls	.062***	.028**	.120***	.067***	.078***	.056***	.095***	.079***
	(.010)	(.008)	(.013)	(.013)	(.016)	(.010)	(.013)	(.022)
Bounds, $\lambda \in [0, 1]$	[-.183,.062]	[-.414,.028]	[-.182,.120]	[.005,.067]	[-.511,.078]	[-.153,.056]	[-.012,.095]	[-.481,.079]
	(-.217,.080)	(-.529,.045)	(-.227,.147)	(-.034,.093)	(-.701,.110)	(-.189,.076)	(-.055,.120)	(-.893,.122)
$\hat{\lambda}^{\infty}$	2.75***	1.51***	2.75***	4.53***	1.51***	2.75***	4.51***	1.49***
$\hat{\lambda}(0)$	0.27***	0.18***	0.43***	1.07***	0.29***	0.29***	0.89***	0.55**
$\hat{\lambda}^{*}$	0.18***	0.07	0.31***	0.78**	0.17	0.17***	0.49***	0.21
N	4,296	5,043	4,296	4,227	5,044	4,296	4,196	3,427

Notes: Additional controls include those listed in Tables 1 and 2, as well as birth year dummies. $\hat{\lambda}^{\infty}$ corresponds to the value of λ at which identification breaks down, i.e. $\lambda > \hat{\lambda}^{\infty}$ yields bounds on $\hat{\alpha} \in (-\infty, \infty)$. $\hat{\lambda}(0)$ is the value of λ at which the bounds of $\hat{\alpha}$ include 0. $\hat{\lambda}^{*}$ is the value of λ such that the 95% confidence interval on the bounds of $\hat{\alpha}$ include 0. Standard errors below OLS coefficients in parentheses. 95% confidence interval below each set of bounds in parentheses. ** indicates significance at the 5% level; *** at the 1% level.

Table 5: Effect of sports on educational outcomes for women

	Graduate HS		Attend college			Graduate college		
	NLSY79	Add Health	NLSY79	NELS:88	Add Health	NLSY79	NELS:88	Add Health
OLS, no controls	.124***	.047***	.215***	.066***	.127***	.145***	.184***	.121***
	(.010)	(.007)	(.015)	(.010)	(.016)	(.011)	(.014)	(.022)
OLS, full controls	.044***	.019***	.072***	.035***	.083***	.047***	.086***	.070***
	(.009)	(.007)	(.014)	(.010)	(.015)	(.010)	(.012)	(.018)
Bounds, $\lambda \in [0, 1]$	[-.135,.044]	[-.226,.019]	[-.201,.072]	[-.077,.035]	[-.179,.083]	[-.149,.047]	[-.091,.086]	[-.256,.070]
	(-.161,.057)	(-.291,.032)	(-.245,.099)	(-.113,.053)	(-.311,.113)	(-.184,.070)	(-.136,.110)	(-.433,.104)
$\hat{\lambda}^{\infty}$	3.17***	1.92***	3.17***	3.43***	1.92***	3.17***	3.43***	1.84***
$\hat{\lambda}(0)$	0.38***	0.14***	0.35***	0.35***	0.49***	0.26***	0.51***	0.37***
$\hat{\lambda}^{*}$	0.17***	0.02	0.17***	0.18*	0.27**	0.13*	0.31***	0.19
N	4,534	5,971	4,534	4,742	5,971	4,534	4,705	4,609

Notes: See notes to Table 4.

Table 6: Effect of sports on labor market outcomes for men

	Log wages			Full-time employment		
	NLSY79	NELS:88	Add Health	NLSY79	NELS:88	Add Health
OLS, no controls	.172***	.138***	.090***	.061***	.007	.017
	(.017)	(.025)	(.029)	(.015)	(.012)	(.010)
OLS, full controls	.077***	.128***	.072**	.037***	.013	.015
	(.017)	(.025)	(.030)	(.014)	(.012)	(.013)
Bounds, $\lambda \in [0, 1]$	[-.250,.077]	[.033,.128]	[-.886,.764]	[-.027,.037]	[.013,.059]	[-.541,.568]
	(-.321,.111)	(-.010,.177)	(-1.52,2.46)	(-.086,.064)	(-.011,.118)	(-.723,.746)
$\hat{\lambda}^{\infty}$	2.47***	4.56 ***	1.46***	2.48***	N/A	N/A
$\hat{\lambda}\,(0)$	0.28***	1.23**	0.47	0.65**	N/A	N/A
$\hat{\lambda}^{*}$	0.16*	0.68	0.07	0.18	N/A	N/A
N	3,296	3,864	3,128	4,296	4,227	3,246

Notes: See notes to Table 4. Additionally, the wage and employment models follow a more Mincerian approach by including years of education, dummies for high school and four-year college graduation, and a cubic in work experience. As noted in Table 1, work experience is excluded for the NELS because it is not reliably measured. The bounds on Add Health treatment effects are nearly uninformative because $\hat{\lambda}^{\infty}$ is close to 1, thus making identification tenuous.

Table 7: Effect of sports on labor market outcomes for women

	Log wages			Full-time employment		
	NLSY79	NELS:88	Add Health	NLSY79	NELS:88	Add Health
OLS, no controls	.115***	.130***	.076**	.080***	.027**	.027**
	(.018)	(.024)	(.034)	(.015)	(.013)	(.011)
OLS, full controls	.017	.083***	.056	-.024	.034**	.022
	(.016)	(.024)	(.032)	(.014)	(.013)	(.012)
Bounds, $\lambda \in [0, 1]$	[-.173,.017]	[-.210,.083]	[-.233,.056]	[-.227,-.024]	[.034,.152]	[-.407,.422]
	(-.224,.048)	(-.335,.131)	(-1.01,.120)	(-.274,.003)	(.008,.267)	(-.585,.639)
$\hat{\lambda}^{\infty}$	N/A	3.32***	N/A	N/A	3.43***	N/A
$\hat{\lambda}(0)$	N/A	0.35***	N/A	N/A	-0.92	N/A
$\hat{\lambda}^{*}$	N/A	0.14	N/A	N/A	-0.37	N/A
N	2,935	4,072	3,977	4,604	4,741	4,316

Notes: See notes to Tables 4 and 6.

Table 8: Effect of sports on health and risky behavior outcomes for men

	Regular exercise		Obesity		Alcohol abuse	
	NELS:88	Add Health	NLSY79	Add Health	NELS:88	Add Health
OLS, no controls	.171***	.077***	.011	-.007	.120***	.017
	(.017)	(.016)	(.007)	(.012)	(.016)	(.012)
OLS, full controls	.175***	.091***	.016**	.004	.107***	.017
	(.017)	(.018)	(.007)	(.013)	(.016)	(.013)
Bounds, $\lambda \in [0, 1]$	[.175,.234]	[-.778,.967]	[.016,.082]	[-.895,.913]	[.005,.107]	[-.735,.769]
	(.142,.350)	(-.960,1.14)	(.002,.146)	(-1.09,1.10)	(-.080,.139)	(-.928,.966)
$\hat{\lambda}^{\infty}$	4.55***	1.51***	2.75***	N/A	4.54***	N/A
$\hat{\lambda}\,(0)$	4.76	0.55*	-1.04	N/A	1.04***	N/A
$\hat{\lambda}^{*}$	-1.25	-1.13	-0.63	N/A	0.78**	N/A
N	4,224	5,040	4,296	5,001	4,213	5,020

Notes: See notes to Table 4. The bounds on Add Health treatment effects are nearly uninformative because $\hat{\lambda}^{\infty}$ is close to 1, thus making identification tenuous.

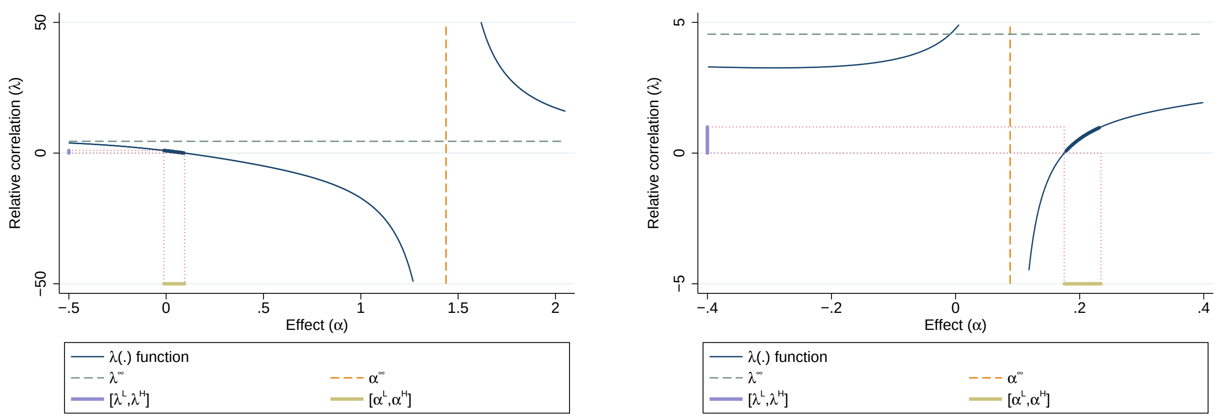
Table 9: Effect of sports on health and risky behavior outcomes for women

	Regular exercise		Obesity		Alcohol abuse	
	NELS:88	Add Health	NLSY79	Add Health	NELS:88	Add Health
OLS, no controls	.113***	.067***	-.023***	-.055**	.053***	.017
	(.014)	(.013)	(.007)	(.014)	(.011)	(.009)
OLS, full controls	.093***	.052***	-.016**	-.042**	.043***	.019**
	(.014)	(.014)	(.007)	(.016)	(.011)	(.009)
Bounds, $\lambda \in [0, 1]$	[-.116,.093]	[-.496,.516]	[-.016,.046]	[-.042,.414]	[-.097,.043]	[-.299,.345]
	(-.194,.122)	(-.629,.754)	(-.029,.089)	(-.073,.625)	(-.173,.065)	(.411,.459)
$\hat{\lambda}^{\infty}$	3.44***	1.92***	3.17***	1.89***	3.44***	1.92***
$\hat{\lambda}(0)$	0.54***	0.27**	0.37*	0.34	0.43***	0.95
$\hat{\lambda}^{*}$	0.29**	0.11	0.08	0.06	0.25	-0.16
N	4,740	5,965	4,534	5,838	4,731	5,960

Notes: See notes to Table 4.

Figure 1: Plots of relative correlation for positively and negatively selected outcomes

(a) College graduation (men, NELS:88) (b) Regular exercise (men, NELS:88)



Notes: Above are plots of λ as a function of the treatment effect α for college graduation (a positively selected outcome) and regular exercise (a negatively selected outcome). The shaded region on the y-axis denotes the bounds on λ, which we assume to be $[0, 1]$. The shaded region on the x-axis denotes to the corresponding bounds on α. λ^{∞} is the horizontal dashed line, which denotes the value of λ at which identification breaks down. The vertical dashed line is α^{∞}, which is the value of α at which $\text{Corr}\,(d, X\beta\,(\alpha)) = 0)$, where $\beta\,(\alpha)$ emphasizes that β is implicitly a function of α. For additional details, see Section 4.1 of the text or Krauth (2016).

For the positively selected outcome, if we widen the bounds on λ in the positive direction, the corresponding interval on α widens in the negative direction. The opposite is true for the negatively selected outcome: widening $[\lambda^L, \lambda^H]$ in the positive direction also widens $[\alpha^L, \alpha^H]$ in the positive direction.

Finally, Panel (b) shows that $\lambda\,(0) > \lambda^{\infty}$ is possible because $\lambda\,(\alpha)$ is not a hyperbola.

A Data appendix

This section presents further details regarding how we select our samples from each of the nationally representative longitudinal surveys that track high school athletic participation and adult outcomes.

A.1 National Longitudinal Survey of Youth 1979 (NLSY79)

The National Longitudinal Survey of Youth, 1979 (NLSY79) surveyed a nationally representative sample of youth and young adults who were aged 14-22 in 1979. Respondents were asked about high school sports participation during the 3rd interview in 1981. Individuals were followed annually until 1994 and biennially thereafter until 2012 for a total of 25 survey rounds. For comparability with the other studies, we focus on individual outcomes at age 25.

A summary of our sample selection and resultant sample sizes is listed in Table A.1.

Table A.1: NLSY79 Sample Selection

Selection criterion	Resultant persons	Resultant person-wave observations
Full NLSY sample	12,686	243,641
Not in disadvantaged white or military oversamples	9,763	214,572
Answered high school athletics question	9,292	209,624
Valid education data	9,292	209,365
Present at age 25	8,830	8,830
Final estimation sample		
Women	4,534	4,534
Men	4,296	4,296

Note: Respondents were asked about high school club participation during the 1984 interview.

A.2 National Education Longitudinal Study of 1988 (NELS:88)

We use data from waves 1 through 4 of the National Education Longitudinal Study of 1988 (NELS:88). The NELS:88 is a nationally representative longitudinal study of adolescents who were in 8th grade in the 1987-88 school year. Respondents were followed four additional times: in 1990 as 10th graders, in 1992 as 12th graders, in 1994 during post-secondary schooling, and in 2000 after the completion of all post-secondary programs. We focus on the base year, 10th grade, 12th grade, and year 2000 interviews.

A summary of our sample selection and resultant sample sizes is listed in Table A.2.

Table A.2: NELS:88 Sample Selection

Selection criterion	Resultant persons	Resultant person-wave observations
Full NELS sample	24,599	100,856
Participated in the base year, 10th grade, 12th grade and final interviews	9,840	9,840
Valid responses for basic demographic variables (race, sex, birth year)	9,505	9,505
Took cognitive tests in base yearbook	9,207	9,207
Parental education and co-residence in base year	9,033	9,033
School enrollment size (10 grade school)	9,032	9,032
Final estimation sample		
Women	4,742	4,742
Men	4,227	4,227

A.3 National Longitudinal Study of Adolescent to Adult Health (Add Health)

The data we use come from waves 1 through 4 of the National Longitudinal Study of Adolescent Health (Add Health). The Add Health is a nationally representative longitudinal study of adolescents who were in grades 7-12 during the 1994-95 school year. Respondents were followed three additional times, most recently in 2008. A fifth wave is in the process of being collected from 2016-18.

We exclude from our analysis those respondents who are missing the Add Health Peabody Picture Vocabulary Test (AH-PVT) score, as well as those who are older than 17 in the first wave. We analyze only respondents who were present in Wave IV, making use of their responses from both Wave I and Wave IV. A summary of our sample selection and resultant sample sizes is listed in Table A.3.

Table A.3: Add Health Sample Selection

Selection criterion	Resultant persons	Resultant person-wave observations
Full Add Health sample	20,728	51,578
Not missing AH-PVT score	19,713	21,762
Age 17 or under in Wave I	14,662	51,535
Present in Waves I and IV	11,263	11,263
Final estimation sample		
Women	6,113	6,113
Men	5,150	5,150

Note: Person-wave observation counts reflect attrition from the survey. AH-PVT score is the Add Health Peabody Vocabulary Test score, which is an abbreviated version of the Peabody Picture Vocabulary Test (PVT), which is designed to measure cognitive verbal skills.