

EXTRACURRICULAR PARTICIPATION, “AT-RISK” STATUS, AND THE HIGH SCHOOL DROPOUT DECISION

Laura M. Crispin

Department of Economics
Saint Joseph's University
Philadelphia, PA 19131
lcrispin@sju.edu

Abstract

I estimate the effect of extracurricular participation on the high school dropout decision with a particular focus on at-risk students. Using a sample of tenth grade students from the National Education Longitudinal Study of 1988, I jointly estimate the dropout and participation decisions (including extracurricular offerings per student), and eligibility requirements as instruments for extracurricular participation. I include an interaction between the participation and at-risk indicators in the dropout equation because past disadvantages may differentially affect at-risk students. I also estimate alternative specifications to identify the effect of participation in different types of activities. Local average treatment effect estimates range from 14 to 20 percentage points, indicating that participants are significantly less likely to drop out of high school than they would have been if unable to participate, with similar estimates for both at-risk and not-at-risk students. These findings are relevant to policy makers and administrators seeking to increase high school graduation rates and improve educational outcomes.

doi:10.1162/EDFP_a_00212

© 2017 Association for Education Finance and Policy

1. INTRODUCTION

The importance of completing high school has been documented throughout numerous studies, yet the dropout rate is relatively high.¹ For instance, the average dropout rate for public high school students in the National Education Longitudinal Study of 1988 (NELS:88) was 9.7 percent, and even higher for at-risk students—15.9 percent.² Extracurricular participation may help to reduce dropout rates, and indeed, summary statistics show that participation in extracurriculars was correlated with substantially lower dropout rates—the dropout rate for participants was 7.1 percent, and for nonparticipants the dropout rate was 21.2 percent.

Though these patterns yield suggestive evidence that extracurricular participation may dramatically reduce dropout rates, the goal of this paper is to provide empirical analyses to determine whether the relationship between participation and the dropout decision is causal. The associations above may be spurious if higher-ability students choose to participate in extracurriculars and also are less likely to drop out. In contrast, if the marginal students choose to participate in order to develop skills outside of the classroom, for example, and are also more likely to drop out, then participation may actually have a more negative effect on the dropout decision than the above summary statistics suggest. Determining the causal effect of participation on the dropout decision is also complicated empirically because extracurricular participation is not the only use of after-school time; students may forgo alternative skill-developing activities (e.g., homework and working) in order to participate in extracurriculars. Thus, to determine the causal effect of participation on the dropout decision, selection as well as investments in other activities must all be taken into account. Furthermore, because policies that may restrict participation (e.g., No Pass/No Play and “Pay-to-Play” laws) may differentially affect at-risk students, it is important to account for this heterogeneity when estimating causal effects.³

To model the relationship between participation, at-risk status, and the dropout decision (by weighing the expected value of completion against the expected value of dropping out) I use a static, binary choice framework in which students choose whether or not to complete high school at the end of the compulsory schooling period. The net expected value of completion is a function of students’ extracurricular participation and at-risk status (among other factors), and their interaction. Using data from the NELS:88, I estimate the probability of dropping out as a function of a binary indicator of students’ tenth grade extracurricular participation. One novel feature of my method is the interaction between the participation and at-risk indicators in the dropout equation, where the at-risk indicator equals one if, as of eighth grade, the student’s family was in the lowest quartile of the socioeconomic status distribution or if the student attended a school with more than 40 percent of students receiving free lunch.

-
1. See, for example, the sheepskin effect literature (e.g., Hungerford and Solon 1987; Jaeger and Page 1996) that documents the labor market returns to a high school diploma; see Cameron and Heckman (1993) for a discussion of the benefits of a diploma relative to a General Equivalency Diploma.
 2. I calculate all figures in this section using a sample of public high school students from the NELS:88. A discussion of the NELS:88 and the sample selection is provided in section 4. I define “at-risk” students as those who enter high school already at a disadvantage and are significantly more likely to have adverse educational outcomes relative to their peers. More specifically, students from low-income families or who attended Title I-eligible primary schools.
 3. These policies and other constraints to participation are further discussed below.

Because participation is endogenously determined and both participation and the dropout decision are binary choices, I estimate a bivariate probit model of the dropout and the extracurricular participation decisions using the number of extracurriculars offered per hundred students in the school and the No Pass/No Play (NPNP)-type rules in the state. These instruments (which to my knowledge I am the first to use), generate exogenous variation in extracurricular opportunities across schools, which are indicators of constraints on participation in lieu of the more commonly used instruments, such as school size and the number of books in the school library (see Barron, Ewing, and Waddell 2000).⁴ The number of activities offered relative to the size of the school is related to the amount of competition that participants face—many activities have a limited number of positions available (e.g., athletics and student council) for which students must compete, and an increase in offerings per student increases the likelihood that a student will be able to participate in at least one activity. NPNP-type rules govern the minimum grade/passing requirements for extracurricular eligibility, which affect whether certain students are able to participate and also affect participation rates at the school level. As extensions of this analysis, I test whether the benefits from participation differ with the type of activity by estimating the bivariate probit model separately for each type of activity (i.e., athletics, arts, academics, and other clubs).

I find that for those affected by the instruments, tenth grade extracurricular participation reduces the likelihood of dropping out by 18.3 percentage points for not-at-risk students, and 17.6 percentage points for at-risk students. When estimating separately by activity type, I find that the instruments are strong predictors of athletics and other club participation (including student government) likely because these are competitive activities with a limited number of positions, thus, the activities for which the instruments are relevant. I find that athletic participants are 16.5 percentage points less likely to drop out than had they not participated, and at-risk participants are 16.6 percentage points less likely to drop out than had they not participated.

In light of these findings, educators and policy makers concerned about students' outcomes should carefully consider laws that restrict participation in extracurriculars (such as NPNP and Pay-To-Play) and may disproportionately constrain at-risk students.⁵ Though designed to incentivize stronger academic performance, these policies may have the unintended consequence of increasing dropout rates via the reduction in participation for at-risk students in particular. Several current surveys suggest that access to and participation in extracurriculars has been reduced specifically for low-income households and minority students. For instance, a report by the C.S. Mott Children's Hospital (2012) studied access to extracurriculars by household income and found that in low-income households, 34 percent of students participated in school sports, whereas in higher-income households, more than 51 percent of students participated.

4. No Pass/No Play (NPNP) refers to laws implemented in Texas and several other states in the mid 1980s that limit extracurricular eligibility to those who have passed all classes in previous years. Other states have minimum grade point average (GPA) requirements. Several articles in the education literature (e.g., Lapchick 1989; Burnett 2000) discuss the NPNP policies and their implications for numerous states.
5. Recently enacted policies have increased funding for disadvantaged schools and introduced open enrollment policies for disadvantaged students in disadvantaged schools (e.g., the No Child Left Behind Act of 2001). Pay-to-Play fees, however, which require students to pay for their own extracurricular activities, rather than being funded through the school, have been in the news recently as having detrimental effects primarily for low-income students (see Simon 2011; Bright 2011).

The study also found that increased cost of school sports is associated with a relatively larger reduction in participation for low-income households. A joint study in 2015 by the National Women's Law Council and the Poverty and Race Research Action Council (Goss Graves et al. 2015) found that in schools with a large share of minorities, students have fewer than half as many opportunities to play sports relative to schools with a large share of white students. This study also found that schools with a large share of minorities also have a relatively larger gender gap in access to athletics—girls only have 67 percent of the opportunities that boys have, relative to 82 percent in primarily white schools (p. 3). Similarly, a 2013 report by the Girl Scouts of America found that girls who planned to go to college were the most likely to participate in extracurriculars, and white girls participated in most activities (sports, performing arts, and academic clubs) at higher rates than black or Hispanic girls (see Schoenberg et al. 2013). Thus, by estimating causal effects of extracurriculars (by providing evidence of the impact of policies to improve extracurricular opportunities), the results in this paper are able to address the above concerns regarding access.

This paper proceeds as follows. In section 2, I discuss the current literature, followed by an exposition of the empirical framework and estimation strategy in section 3. I discuss the data for my analysis, discuss the construction of the instruments, and provide descriptive statistics of key variables in section 4. I provide results and robustness checks in section 5 and conclude with discussion in section 6.

2. RELATED LITERATURE

The extracurricular participation and at-risk literatures have developed nearly independently from one another, though, as indicated in the Introduction, these two subjects are certainly intertwined and should be studied simultaneously.

Extracurricular Participation

There is a small but growing body of work in this area that has primarily focused on high school athletic participation and the effect on later outcomes, such as final educational attainment (e.g., Barron, Ewing, and Waddell 2000; Eide and Ronan 2001; Stevenson 2010) and wages (e.g., Barron, Ewing, and Waddell 2000; Eide and Ronan 2001; Kuhn and Weinberger 2005; Kosteas 2010; Stevenson 2010). Nevertheless, participation in other activities—namely, clubs and vocational activities—may also affect educational and labor market outcomes, though these relationships have received less attention in the literature (see, e.g., McNeal 1995; Lipscomb 2007; Costa 2010; Kosteas 2010).

Of this literature, there are four studies (Eide and Ronan 2001; Lipscomb 2007; Rees and Sabia 2010; Stevenson 2010) that estimate the causal relationship between athletics and contemporaneous outcomes such as test scores and high school completion, and only Lipscomb (2007) provides analysis for different types of activities. Stevenson (2010) uses the variation in sports participation induced by the introduction of Title IX to identify the effect of athletics on several outcomes for girls who attended school during the 1970s. Of her results, the most relevant for my purposes is the finding that a 10 percentage point increase in sports participation rates increases high school completion rates by 0.6 percentage points. Although her identification strategy is similar to mine, her results are restricted to sports participation and her analyses

are at the state level. Therefore, she identifies the effect of average participation rates on an individual's outcomes, whereas I estimate the impact of an individual's participation on his/her own outcome. Eide and Ronan (2001) also study the relationship between sports participation and high school completion. Using data from the High School and Beyond survey, their ordinary least squares results show that sports participation reduces the likelihood of dropping out by 8.6 percentage points for boys and 3.9 percentage points for girls. When instrumenting for sports participation using height at age 16, however, they find statistically insignificant results, indicating suggestive evidence of a relationship between extracurriculars and high school completion. Rees and Sabia (2010) also use students' height as an instrument for sports participation to identify the effect on math and reading GPAs, along with three behavioral indicators. Their instrumental variable (IV) results show no significant relationship between sports participation and GPA, though they find a positive effect on college aspirations. Lipscomb (2007) uses a fixed effects approach to identify the effect of a change in participation status on the change in math and reading test scores. Using a sample of students from the NELS:88, he finds that from eighth to twelfth grades, an extra year of athletic participation increased math and science test scores by 2 percent, and an extra year of club participation increased math test scores by 1 percent, which shows that skills developed through extracurriculars are productive in the classroom.

At-Risk Status

The term "at-risk" refers to students who are likely to suffer adverse educational outcomes. From the at-risk literature, Finn (2006) identifies students as at-risk using three different at-risk indicators, the most exogenous of which is the "status risk," which is equal to one if students are in homes with socioeconomic status below the median and enrolled in schools where the share of students receiving free lunch is above the median.⁶ His paper is one of the few to look at the relationship between extracurricular participation and outcomes for at-risk students, though he does not attempt to identify causal relationships. Using data from NELS:88, he finds that only 25 percent of at-risk students successfully complete high school and the remaining students are either marginal completers or noncompleters. In earlier work, using data from two middle schools, Mahoney and Cairns (1997) also examine the interaction between at-risk status and extracurricular participation and the correlation with early dropout behavior. They find that at-risk participants are associated with relatively lower dropout rates than nonparticipants.⁷

Surprisingly, there seems to be no consensus on how to define at-risk status. For instance, Kagaruki-Kakoti (2005) uses another definition of at-risk status to estimate the relationship between being at-risk and students' outcomes—namely, high school completion. Using the same data as myself and Finn, she regresses a vector of individual and family characteristics as of eighth grade on an indicator for whether the student

6. The second and third indicators are based on "behavioral risk factors" such as skipping class or failing to complete assignments, and "academic risk factors" such as low grades or test scores.

7. Similar work by Covay and Carbonaro (2010) studies elementary school children and the relationship between after-school activities and noncognitive skills and test scores. These authors interact socioeconomic status with extracurricular participation indicators to allow for heterogeneous effects by socioeconomic status, but do not attempt to identify causal effects of participation.

drops out of school in any subsequent period. She defines students as at-risk if they are associated with at least one of the six factors that are significantly and negatively related to high school completion: (1) low parental education, (2) limited English proficiency, (3) having a sibling who has dropped out, (4) being left home alone for three or more hours daily, (5) low family income, or (6) living with a stepparent. Using this definition, she estimates the probability of dropping out of high school, the probability of attaining post-secondary education, and weekly salary as functions of at-risk status, student, family, peer, and neighborhood characteristics. She finds that relative to their peers, those at risk are 3 percentage points more likely to drop out, 3 percentage points less likely to attain additional education, and earn 7.8 log points less per week.

In addition to the definitions discussed above, educational policies provide yet another definition of at-risk status. For example, the No Child Left Behind Act of 2001 defines “at-risk” as “at-risk of academic failure, has a drug or alcohol problem, is pregnant or a parent, has come into contact with the juvenile justice system in the past, is at least one year behind the expected grade level for the age of the individual, has limited English proficiency, is a gang member, has dropped out of school in the past, or has a high absenteeism rate at school” (USDOE 2002, p. 1591). Though this is a more comprehensive definition of at-risk, it tends to include endogenous actions like joining a gang rather than exogenous factors like demographic characteristics. Therefore, this definition is difficult to use in empirical analysis given that these at-risk students have chosen to be at-risk rather than live in an environment that leads to adverse outcomes. Furthermore, this definition relies on ex-post characteristics and actions, so it is not useful for policies that seek to intervene at an earlier stage before the student has dropped out, joined a gang, or become a parent, for instance.

Because of this lack of consensus and concerns about exogeneity of these alternatives, I rely on my own measure of at-risk status (discussed in section 4). I do, however, use both Finn and Kagaruki-Kakoti’s definitions of at-risk status as robustness checks on my findings because of their (plausible) exogeneity from the extracurricular participation decision and their measurement prior to high school.

3. EMPIRICAL FRAMEWORK AND ESTIMATION STRATEGY

Empirical Framework

The primary question of this paper is: Does extracurricular participation have a causal effect on the dropout decision? To address this question, I model a student’s high school dropout decision as a static, binary choice made at the end of the compulsory schooling period.⁸ With the objective of maximizing the present discounted value of lifetime utility, which is a function of leisure and consumption, the student must evaluate the expected value of utility from continuing high school relative to dropping out.⁹

-
8. Modeling only one schooling period instead of a sequential decision-making process is common in the extracurricular participation literature (see, e.g., Barron, Ewing, and Waddell 2000; Eide and Ronan 2001; Kosteaas 2010).
 9. A common framework assumes the student maximizes the present discounted value of lifetime earnings (see, e.g., Willis and Rosen 1979; Catsiapis 1987). However, as I discuss in this section, human capital investments such as extracurricular participation may also have consumption value that makes schooling more enjoyable, and schooling itself may be enjoyable. Therefore, I use a utility maximization framework instead. Others such as Barron, Ewing, and Waddell (2000) and Eide and Ronan (2001) have also used the utility maximization framework to explore the effects of extracurricular participation on students’ educational attainment.

The expected value of continuation increases with both the increase in expected future income from the human capital acquired during the final high school years and the current enjoyment associated with school attendance (e.g., spending time with friends). The expected value of dropping out increases with the value of foregone earnings and foregone leisure (Schultz 1963). These expected values vary across students and depend on the student's skills and abilities, tastes, and expectations, the quality of educational inputs, and the current labor market conditions at the time of the decision (Hill 1979; Willis and Rosen 1979). Thus, the student weighs these alternatives at the end of the period and decides to drop out when the expected value of high school continuation net of the expected value of dropping out is negative.

The student's skills and expected satisfaction from additional schooling, and hence the net expected value of high school continuation, are affected by the student's activities (namely, extracurriculars) during the schooling period, the influences and educational inputs from her family and school, and her individual ability. Therefore, the net expected value of high school continuation for individual i at the end of the compulsory period (i.e., at the end of time t) can be written as a function of activities, influences, inputs, and individual attributes during period t , and the decision rule for dropping out can be defined, respectively, as:

$$EV(HS_{it}) = \beta_1 + \beta_2 EP_{it} + \beta_3 AR_{it} + \beta_4 EP_{it} * AR_{it} + \beta_5 X_{it} + \varepsilon_{it}, \tag{1a}$$

$$DO_{it} = 1 \text{ if } EV(HS_{it}) < 0, \text{ and } 0 \text{ otherwise,} \tag{1b}$$

where the student's extracurricular participation during the school period t , EP_{it} , the student's at-risk status, AR_{it} , and their interaction, are the focus of this study and are discussed below. X_{it} represents the remaining student, family, and school characteristics during the schooling period, the labor market conditions, time allocated to homework and work, and the student's ability. The last term includes any omitted factors that affect the net expected value of high school continuation.

Extracurricular participation may affect the dropout decision for several reasons. First, extracurricular participation is often viewed as a human capital investment because participants may develop interpersonal skills, leadership and teamwork skills, and increase their confidence, self-discipline, and self-esteem (Spreitzer 1994; Barron, Ewing, and Waddell 2000; Eide and Ronan 2001; Kosteas 2010). Therefore, extracurricular participation should increase the amount of skill acquired during the remaining high school years, which increases the net expected value of high school continuation. Skills developed during extracurriculars may also be valuable in the labor market, implying higher foregone earnings and, thus, a reduction in the net expected value of remaining in school. Extracurricular activities can also be viewed as an investment in social capital (Feldman and Matjasko 2005), where participants develop peer networks that may increase the enjoyment of schooling during the final high school years so as to increase the net expected value of remaining in school. Thus, with the exception of extreme cases (e.g., the athletic superstar whose labor market opportunities include beginning a professional career), it is most likely that participation leads to a reduction in the likelihood of dropping out.

Additionally, extracurricular activities may also be a form of current consumption, where participants forgo other activities that may develop skills (i.e., homework and work) in order to enjoy extracurriculars (see, e.g., Eide and Ronan 2001; Kosteas 2010), which would decrease the net expected value of high school continuation for participants. Nevertheless, empirical evidence (provided in table A.2, which is available in a separate online appendix that can be accessed on *Education Finance and Policy's* Web site at www.mitpressjournals.org/doi/suppl/10.1162/EDFP_a_00212, and discussed in section 4) suggests extracurricular participants are no more or less likely to work during high school than nonparticipants, and that participants complete more hours of homework than nonparticipants. This suggests that the opportunity cost of participation is forgone leisure rather than forgone skill. Therefore, it is important to account for other uses of time but it is not necessary to explicitly model these other time uses.

The at-risk indicator is intended to measure the history of disadvantage due to aspects of the student's family and school environment that may result in lower levels of acquired skill, less taste for schooling, or lower educational expectations relative to her peers. The disadvantages are largely socioeconomic, where students from low income families or with less-educated parents often attain fewer years of schooling (Cameron and Heckman 2001). This may be due to parents' relatively lower educational expectations or relatively lower investment in their child (Hill and Stafford 1977). A similar comparison can be drawn with respect to the history of schooling inputs, where students who have received a history of low-quality schooling inputs may acquire fewer skills relative to their peers (Hanushek 1979). Thus, the net expected value of high school continuation may be lower for at-risk students relative to not-at-risk students, with all else equal.¹⁰

I include the interaction between the participation and at-risk indicators to test whether the effects of extracurricular participation on the dropout decision differ by at-risk status. There are several reasons to include this interaction. First, if at-risk students have accumulated relatively fewer skills, then at-risk participants are likely to acquire relatively less skill from extracurriculars, and therefore have a relatively lower net expected value of additional schooling. Second, the consumption benefits from peer networks may have a differential effect on at-risk students, though the direction is ambiguous. This is because on the one hand, at-risk students may have relatively lower taste for schooling, and so may derive relatively less enjoyment from peer networks during the final high school years. On the other hand, the positive peer influences or popularity from extracurriculars could help the student to reengage in the schooling environment.¹¹ The interaction term between the participation and at-risk indicators captures the net effect of these mechanisms.

Although a student's at-risk status is (plausibly) exogenously determined by family and schooling characteristics, students are not randomly assigned to extracurriculars during the schooling period. Extracurricular participation is a choice that students

10. I am interested in the history of disadvantage, rather than the current family income or school quality, because the long-term disadvantages are likely to have a cumulative effect, and thus a stronger impact on the student's educational decision. Cameron and Heckman (2001) find evidence of this with respect to the college enrollment decision, showing that socioeconomic circumstances at the time of the college decision explain relatively little of the racial gap in educational attainment, but that the history of family income has more explanatory power.

11. Mahoney and Cairns (1997) make similar arguments with respect to the "marginal" student.

make after weighting the expected value of participation against the alternative (not participating), which are also functions of the student's ability, skill, tastes and preferences, and opportunities during the schooling period. Therefore, the net expected value of participation can be modeled and the decision rule for participation defined, respectively, as follows:

$$EV(EP_{it}) = \gamma_1 + \gamma_2 Z_{it} + \gamma_3 AR_{it} + \gamma_4 X_{it} + v_{it}, \quad (2a)$$

$$EP_{it} = 1 \text{ if } EV(EP_{it}) \geq 1, \text{ and } 0 \text{ otherwise,} \quad (2b)$$

where the net expected value of participating is a function of the opportunities for participation, Z_{it} , the student's at-risk status, and the vector of student, family, and schooling characteristics. The error term includes any omitted factors that affect the net expected value of participation, and is assumed to be correlated with omitted factors that affect the dropout decision, though the direction of the correlation is indeterminant. For instance, students with higher ability may expect to benefit more so from extracurriculars and from schooling than their peers, and thus omitted ability may generate a negative correlation between the participation and dropout decisions. As a contrasting example, students with lower ability may benefit more than their peers from extracurriculars through the development skills outside the classroom (e.g., leadership or teamwork skills) or from the enjoyment of peer networks. They may also anticipate a lower net expected value of high school continuation, resulting in a positive correlation between the participation and dropout decisions. Because the opportunities to participate (Z_{it}) generate exogenous variation in participation that differs across students, this source of variation can be used to identify the causal effect of participation on the high school dropout decision.

Though I have only discussed extracurricular activities in general, different activities may develop different skills or different amounts of social capital. For instance, academic clubs are likely to increase students' cognitive skills more so than participation in athletics, whereas athletic participation may have a substantial impact on students' leadership or teamwork skills. Thus, the effects of participation on the high school dropout decision may differ by activity. To incorporate this, I extend the basic framework by relaxing the assumption that a student makes a binary choice of whether to participate in extracurriculars, and allow for distinctions between types of activities. For each activity, the student chooses to participate in that activity when the net expected value of participation is positive.¹² To use this framework, I make the necessary assumption that the primary component of the marginal cost of participating in a specific activity is foregone leisure, which implies that students do not substitute one activity for another. Though this is restrictive, I find empirical evidence that students participating in one activity are more likely to participate in other activities, which is similar to the relationship between participation and homework hours and suggests this assumption is not unrealistic.¹³

12. Though a time allocation model of extracurricular participation in various activities would be an interesting extension and a valuable contribution to the literature, data are unavailable regarding students' time allocation to different activities and so I leave this model to future work.

13. These descriptive patterns are available upon request.

Estimation Strategy

As discussed above, the binary decisions to drop out (equation 1b) and participate in extracurriculars (equation 2b) are both functions of inputs and characteristics during the schooling period. Therefore, I estimate the basic model above as a system using a bivariate probit specification. The probability of each decision is a function of educational inputs, as well as student, family, and school characteristics during tenth grade (the “schooling period”). At-risk status is determined prior to entering tenth grade. The joint probability of tenth grade participation (EP_{it}) and the subsequent dropout decision (DO_{it}), conditional on observables, can be written as:

$$Pr(DO_{it} = 1, EP_{it} = 1) = \Phi(\beta_1 + \beta_2 EP_{it} + \beta_3 AR_{it} + \beta_4 EP_{it} * AR_{it} + \beta_5 X_{it}, \gamma_1 + \gamma_2 Z_{it} + \gamma_3 AR_{it} + \gamma_4 X_{it}, \rho), \tag{3}$$

where $\Phi(\cdot)$ is the bivariate normal cumulative distribution function, EP_{it} is an indicator equal to one if the student participated in extracurriculars, AR_{it} is an indicator equal to one if the student is at-risk upon entering tenth grade, the vector X_{it} includes measures of the family and school inputs, proxies for student ability and preferences, such as race, gender, and prior test scores, and proxies for labor market condition such as neighborhood demographics (i.e., median household income and racial composition of the neighborhood). The vector Z_{it} includes two measures of extracurricular opportunities: extracurricular offerings per student and state-level eligibility requirements. The errors are assumed to be distributed as:

$$\begin{pmatrix} \varepsilon_{it} \\ v_{it} \end{pmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\varepsilon^2 & \rho \sigma_\varepsilon \sigma_v \\ \rho \sigma_\varepsilon \sigma_v & \sigma_v^2 \end{bmatrix} \right). \tag{4}$$

The effect of participation on the dropout decision (i.e., β_2) is identified due to exclusion restrictions, where extracurricular opportunities (i.e., the vector Z_{it}) predict participation but are excluded from the dropout equation. The first excluded variable is a measure of the extracurricular offerings per student in the school. This instrument should affect participation because larger schools may offer more clubs and sports than smaller schools, but not in proportion to the differences in enrollment. That is, larger schools may offer relatively more extracurriculars, but there may be even more students competing for a position on the football team or in the student government. In contrast, in smaller schools, all students with an interest or talent in a particular area are likely to be needed in order to field the team. Thus, variation in the offerings per student can be used to predict students’ participation status.¹⁴ The concern for excluding this variable from the dropout equation is that high-quality schools may provide more resources per student (including extracurriculars), and thus the instrument could have a direct effect on the dropout decision. To mitigate this concern, I include a vector of schooling inputs to account for differences in schooling resources. Empirical analysis in section 5 provides further validation of this instrument.

The second instrument is the presence of NPNP laws in the state, requiring a minimum GPA to be met or that students pass a minimum number of courses in

14. This argument is an expanded version of that provided by Barron, Ewing, and Waddell (2000).

the prior semester before they are eligible to participate. These laws should restrict participation but not affect the dropout decision directly because these laws have no bearing on graduation requirements. In addition, I have found no evidence that NPNP laws were passed simultaneously with laws that increased graduation requirements. Therefore, NPNP laws should affect the dropout decision only through the effect on students' extracurricular participation.

Although there is an interaction between the participation and at-risk indicators in the dropout equation, identification in this model is no more complicated than in a standard bivariate probit model if the at-risk indicator is exogenous to the participation decision.¹⁵ This condition can be met by choosing variables that reflect socioeconomic disadvantages that are not a function of the student's current or past participation decision. Therefore, I rely on the families' socioeconomic statuses and a proxy for the schools' Title I funding prior to tenth grade, as neither is likely to depend on the student's participation status.

As an extension, using an identical estimation strategy, I estimate the model in which students choose to participate in different types of activities. This time I replace the indicator for extracurricular participation with an indicator for one of four extracurricular categories: athletics, art-related clubs, academic clubs, and other clubs.¹⁶ Although allowing students to simultaneously choose whether to participate in all activities may be more closely related to the students' actual decision making process, there are only two available instruments, and estimating a model of four endogenous participation decisions (or potentially more, assuming that students may choose to participate in multiple activities during the year) jointly with the dropout decision would not be identified. Therefore, I have chosen to estimate each type of activity and its relationship with the dropout decision using a separate bivariate probit regression.

4. DATA DESCRIPTION

Overview and Sample Selection

In my analysis, I use a sample of students from NELS:88. In the 1988 initial survey round, approximately 30 students from each of 1,050 schools, both public and private, from across the country were selected to participate, yielding an initial sample of 24,600 students.¹⁷ In 1988, students were interviewed in eighth grade, then again in tenth and twelfth grades (in 1990 and 1992, respectively) to gather data about their schooling experiences.¹⁸ In the final two rounds in 1994 and 2000, students were interviewed to

15. See Wooldridge (2010, p. 596) for a discussion of bivariate probit models with interactions. An alternative estimation strategy would be to estimate the bivariate probit model separately for at-risk and not-at-risk students. Using this alternative strategy, I find similar marginal effects and hence my conclusions remain unchanged. Nevertheless, the loss of efficiency is much larger under this alternative, therefore I rely on the estimation strategy specified by equation 3.
16. The categorization scheme that I describe is similar to that of McNeal (1995), who uses the following four categories: athletics, fine arts, academics, and vocational clubs.
17. The sampling frame included only schools with an eighth grade, and schools were chosen using a probability-weighted selection procedure. See Ingels et al. (1994, 1998) and the online data appendix for further details.
18. Though this dataset is somewhat dated, this is the only dataset (to my knowledge) that contains detailed information on extracurriculars at both the student and school levels, along with outcome and contextual data. Therefore, this dataset is ideal for the purposes of this study, and, as indicated in the Introduction, extracurricular opportunity is currently a concern for policy makers, schools, and many advocacy groups alike.

Table 1. Sample Selection Procedures

Reason for Elimination	No. of Individuals	% of Original Sample	No. of Schools	% of Original Sample
Original sample size	24,600		1,450	
Missing or ineligible in 1990 due to:				
Language barrier, physical/mental disability	-40	-0.2%		
Out of scope	-100	-0.4%		
Status unknown	-410	-1.7%		
Completed an abbreviated questionnaire	-820	-3.3%		
Not included in subsample	-6,210	-25.2%		
Missing or ineligible in 1992 due to:				
Language barrier, physical/mental disability	-10	0.0%		
Out of scope	-80	-0.3%		
Status unknown	-230	-0.9%		
Did not respond/refused participation	-560	-2.3%		
Eligible sample	16,160	65.7%	1,450	100%
Attending private school in 1988, 1990, or 1992	-2,950	-12.0%	-370	-26%
Attending magnet, vocational, or Indian reservation school in 1990 or 1992	-2,010	-8.2%	-160	-11%
Traditional public school sample	11,210	45.6%	930	64%
Missing administrator questionnaire in 1988 or 1990	-730	-3.0%	-130	-9%
Missing extracurricular participation in 1990	-650	-2.6%	-10	-1%
Missing math or reading test scores, race, or socioeconomic status in 1988	-350	-1.4%	-20	-1%
Final sample	9,480	38.5%	770	53%

Notes: Sample sizes are rounded to the nearest ten for confidentiality. Numbers may not add up because of rounding.

gather data about their post-secondary experiences. To gather additional information, in 1988 questionnaires were also administered to a parent, two teachers, and the school administrator of individual students. In the 1990 survey round, questionnaires were administered to the student, teachers, and the school administrator, and in the 1992 round, to the student, a parent, one teacher, and the school administrator.¹⁹

In table 1, I show the selection criteria used to create the sample for my analysis, documenting the number of students and high schools eliminated due to each criterion.²⁰ Beginning with the 24,600 students who participated in the 1988 round of the survey, I exclude students with missing tenth grade data due to sample design—those who were ineligible because of physical or mental disabilities, were not located, completed an abbreviated questionnaire, or were not included in the subsample in 1990. This eliminates 7,570 students. Similarly, I eliminate nonrespondents and those who were ineligible or not located in 1992, which excludes a further 870 students.

Of the remaining 16,160 students, I create the desired sample of traditional public school students by eliminating any students attending private schools or nontraditional public schools in 1988, 1990, or 1992.²¹ These criteria eliminate 2,950 and 2,010 students in 370 and 160 schools, respectively.

19. A subsample of students in both the 1990 and 1992 round were chosen to complete the full questionnaire. Further details of the subsampling process are available in the online data appendix.
 20. All sample sizes are rounded to the nearest ten to ensure confidentiality of the respondents.
 21. I have chosen to focus on students attending traditional public high schools because dropout rates are typically higher in traditional public schools, and the educational process, including curriculum and emphasis on extracurriculars, is likely to be quite different across different types of schools. Furthermore, I have no data

From the remaining sample of 11,210 students, I delete individuals and schools with missing administrator data in 1988 and 1990, and individuals missing extracurricular participation data in 1990, excluding 730 and 650 students in 130 and 10 schools, respectively. Finally, I eliminate 350 students missing key data in 1988 (i.e., test scores, socioeconomic status, race).²² The final sample size is 9,480 students in 770 schools.

Independent Variables Used to Create the At-Risk Indicator

To represent the history of disadvantage that may place students at risk for adverse educational outcomes in their future, I create the at-risk indicator using eighth grade socioeconomic status and the share of eighth grade students in the free lunch program at school. These are two variables that are exogenous to the student and proxy for the quantity and quality of educational inputs at home and at school. There are three advantages to this measure over prior measures of at-risk status: first, it is exogenous to the student; second, it is measured prior to high school, which allows policy makers to easily identify students for intervention upon entering high school; and third, these variables are readily available in most educational datasets.

Summary statistics are provided in table 2. Socioeconomic status is an index provided in the NELS:88 that includes parental income, education, and occupation. The index ranges from -2.97 to 2.56 , and those in the lowest quartile have an index of less than -0.59 . The at-risk indicator equals one for students in families in the lowest quartile of the socioeconomic status distribution *or* attending schools with more than 40 percent of students receiving free lunch, and zero otherwise. The lowest quartile of the socioeconomic status distribution coincides roughly with a \$22,000 income for a family of four, qualifying these students for free school lunch.²³ The cutoff point for the share of students receiving free lunch reflects a school's eligibility for Title I funding, which is provided to disadvantaged schools in order to provide more resources and help these schools boost achievement (USDOE 1965, 2010). With this definition, I identify 3,330 students, or 35 percent of the sample, as at-risk.

As part of the robustness results, I also use Finn's (2006) and Kagaruki-Kakoti's (2005) definitions of at-risk status. Recall that Finn includes students who are both below the median in household income *and* in schools above the median in share of students receiving free/reduced price lunch. Though I use the same variables as Finn to create my at-risk indicator, I do not use his exact definition because students need not be at a disadvantage at school *and* at home in order to be at-risk, and further, his cutoff point for socioeconomic status is somewhat high relative to the poverty threshold in 1988, and his cutoff point for free lunch is somewhat low relative to the cutoff used

about the tuition costs of private school. Therefore, I leave the analysis of extracurricular participation and dropout rates across heterogeneous types of schools for future work.

22. Though the main variables of interest—extracurricular participation and dropout status—are from the 1990 and 1992 rounds, respectively, at-risk status is generated using pre-high school data. Therefore, students must be present in the first three rounds and must have administrator data in the second round in order to remain in my sample.
23. Students are eligible for free lunch if their family income is less than 130 percent of the poverty level. Students whose family income is between 130 and 185 percent of the poverty level are eligible for reduced-price lunch (USDA 2011). According to the U.S. Census Bureau (1988), the poverty threshold for a family of four in 1988 was \$12,092.

Table 2. Summary Statistics for Dependent, Independent, and Instrumental Variables by At-Risk Status

	All Students		Not-At-Risk Students		At-Risk Students		Mean Difference (3)-(5)
	Mean (1)	S.D. (2)	Mean (3)	S.D. (4)	Mean (5)	S.D. (6)	
Outcome:							
1 if Dropout; 0 otherwise	0.10		0.06		0.16		-0.10*
Extracurricular participation indicators:							
1 if participated in any extracurricular club/athletics in 10th grade	0.82		0.85		0.77		0.07*
10th grade participation by type							
1 if participated in athletics	0.56		0.61		0.49		0.12*
1 if participated in arts	0.28		0.29		0.26		0.03*
1 if participated in academic clubs	0.34		0.37		0.29		0.08*
1 if participated in other club	0.35		0.36		0.34		0.01
1 if participated in athletics or other clubs in 10th grade	0.69		0.72		0.64		0.08*
Other independent variables:							
At-risk indicator (equals 1 if at-risk)	0.36						
Average hours of work/week							
1 if 0-20 hrs/week	0.40		0.42		0.36		0.06*
1 if 20+ hrs/week	0.21		0.20		0.21		-0.01
Average hours of homework/week							
1 if 0-3 hours/week	0.56		0.53		0.60		-0.07*
1 if 3-10 hours/week	0.16		0.17		0.15		0.01
1 if 10+ hours/week	0.20		0.24		0.15		0.09*
8th grade math test score	36.19	11.67	38.93	11.71	31.25	9.84	7.67*
8th grade reading test score	26.97	8.44	28.72	8.41	23.83	7.54	4.89*
Race/ethnicity:							
1 if Asian	0.03		0.03		0.03		0.01
1 if Hispanic	0.09		0.05		0.16		-0.11*
1 if Black	0.10		0.05		0.19		-0.14*
1 if Native American	0.01		0.01		0.01		-0.01*
1 if Female; 0 if Male	0.50		0.49		0.53		-0.04*
1 if newspaper available in the home	0.72		0.77		0.64		0.13*
1 if newspaper availability unknown	0.02		0.01		0.03		-0.01*
1 if computer available in the home	0.79		0.86		0.68		0.18*
1 if computer availability unknown	0.01		0.01		0.02		-0.01*
1 if 50 or more books available in the home	0.88		0.93		0.80		0.13*
1 if 50 or more books availability unknown	0.02		0.02		0.02		-0.01*
% of teachers in the school with a master's degree or higher	0.51	0.19	0.53	0.19	0.47	0.19	0.05*
Teacher-student ratio	0.07	0.02	0.07	0.02	0.07	0.02	0.00
Lowest teacher salary (in thousands)	20.07	2.89	20.32	2.86	19.61	2.89	0.71*
Highest teacher salary (in thousands)	38.79	7.85	39.52	8.02	37.47	7.37	2.04*
No. of school days	179.37	2.83	179.48	2.80	179.16	2.88	0.33*
No. of classes per day	6.75	0.86	6.78	0.89	6.68	0.80	0.1+
Share of Hispanic 10th graders in the school	0.09	0.20	0.06	0.12	0.15	0.28	-0.09*
Share of black 10th graders in the school	0.11	0.19	0.07	0.12	0.17	0.26	-0.10*
Share of students in the school enrolled in remedial math program	0.08	0.09	0.07	0.07	0.11	0.11	-0.04*
Share of students in the school enrolled in remedial reading program	0.08	0.09	0.07	0.07	0.11	0.11	-0.04*
Median household income in the zip code (in thousands)	31.30	11.91	34.60	12.47	25.38	7.93	9.22*
Share of black residents in the school's zip code	0.08	0.15	0.05	0.09	0.13	0.21	-0.08*
Share of Hispanic residents in the school's zip code	0.07	0.17	0.05	0.11	0.12	0.24	-0.07*
1 if suburban	0.47		0.55		0.32		0.23*
1 if rural	0.37		0.32		0.48		-0.16*

Table 2. Continued.

	All Students		Not-At-Risk Students		At-Risk Students		Mean Difference (3)-(5)
	Mean (1)	S.D. (2)	Mean (3)	S.D. (4)	Mean (5)	S.D. (6)	
Census division							
1 if Mid-Atlantic	0.14		0.15		0.13		0.03
1 if East North Central	0.18		0.19		0.15		0.04*
1 if West North Central	0.10		0.11		0.08		0.03***
1 if South Atlantic	0.16		0.14		0.19		-0.05***
1 if East South Central	0.07		0.05		0.09		-0.04*
1 if West South Central	0.13		0.11		0.17		-0.05*
1 if Mountain	0.08		0.07		0.09		-0.01
1 if Pacific	0.11		0.13		0.09		0.04*
Instrumental variables							
No. of extracurriculars offered per 100 students	1.58	1.25	1.56	1.18	1.61	1.36	-0.05
1 if NPNP laws in the state	0.89		0.87		0.92		-0.05*
8th grade covariates to generate at-risk indicator							
Socioeconomic status	-0.08	0.73	0.26	0.54	-0.71	0.61	
Share of 8th graders in free lunch program	0.23	0.20	0.14	0.11	0.39	0.24	
Summary statistics							
No. of students	9,480		6,160		3,330		
No. of schools	770		590		633		

Notes: Sample sizes are rounded to the nearest ten to maintain confidentiality. Summary statistics are weighted using panel weights.

* $p < 0.05$; *** $p < 0.10$.

to determine Title I funding. When using Finn’s definition, 2,930 students, or 30.9 percent of the sample, are considered at-risk. Following Kagaruki-Kakoti’s definition, students are categorized as at-risk if neither parent graduated from high school, if they live in a single-parent home, if the student has limited English proficiency, if one or more of his siblings has dropped out of school, if the student spends three or more hours home alone after school, or if the family’s annual income is less than \$15,000.²⁴ Using this definition, 3,820 students, or 40.2 percent of the sample, are classified as being at-risk. I provide a comparison of results from alternative definitions in section 5.

Dropout Indicator, Extracurricular Participation, and Control Variables

In all analyses, the dependent variable of interest is an indicator equal to one if the student dropped out of school during the time between the first and the second follow-ups, and equal to zero otherwise. Summary statistics for this variable, all independent variables, and instruments are provided in table 2 for the entire sample, and separately by at-risk status. These statistics show that 9.7 percent of the sample are dropouts, though this is a weighted average of the at-risk dropout rate of 16.0 percent and the not-at-risk dropout rate of 6.2 percent.

24. Summary statistics are provided in the online data appendix.

The main independent variable of interest is the extracurricular participation indicator, and I use two different specifications of this variable. The first is an indicator equal to one if the student participated in at least one of eighteen extracurricular activities during tenth grade and equal to zero otherwise. The second uses indicators for participation in four different types of activities: (1) athletic activities (baseball/softball, football, basketball, soccer, swim team, “another team sport,” “another individual sport,” cheerleading, poms/drill team); (2) art-related clubs (band/orchestra and drama club); (3) academic clubs (honors society, and “any other academic club”); and (4) “other” clubs (student government, yearbook/newspaper, and “all remaining clubs”). I use these categories to create four separate indicators, where each indicator equals one if the student participated in any activities that fall into that category and zero otherwise.²⁵

I control for other after-school time uses in tenth grade by including the self-reported average hours per week that the student spends on homework after school and working (outside of the home).²⁶

I also include variables to represent the students’ ability, skill, and tastes, and school quality and labor market conditions. To control for ability, I condition on the students’ math and reading test scores in eighth grade. The test scores I use are the students’ “IRT (item response theory) estimated number right” math and reading test scores provided by the NELS:88. Further details of these tests and IRT methodology used in the NELS:88 are provided in the online data appendix. As shown in table 2, not-at-risk students scored significantly higher than at-risk students in both math and reading, with a mean difference of 7.67 points and 4.89 points, respectively. These test scores are intended to account for some of the potential selection issues, where students with higher initial ability may choose to participate and also may be less likely to drop out. I include the students’ race/ethnicity and gender to control for potential differences in achievement, expectations, and influences that are specific to these particular groups. Indicators for newspaper, computer, and book availability within the home are also included, intending to capture differences in family inputs during tenth grade so that the at-risk indicator can be interpreted in terms of past disadvantages.²⁷

The school variables that I include are intended to control for differences in resources provided by the school and the amount of time that students have to allocate to after-school activities. To proxy for school quality, I include the share of teachers with a master’s or higher degree, the teacher–student ratio in the school, and the lowest and

25. Although a time-allocation model detailing students’ time spent on various activities would be a valuable contribution to the literature, students were not asked about hours allocated to each activity and so this remains an avenue for future research.

26. Table A.2, in the online appendix, shows that extracurricular participants substitute time away from leisure rather than from skill-producing activities. As shown, the share of participants within each work category is quite stable, revealing the work and extracurricular participation decisions are largely uncorrelated. Also, the share of participants in homework hour categories is increasing, indicating that participants are likely to do more homework than nonparticipants. Anderson (2001) found similar patterns in the participation-homework hours relationship. These patterns suggest that students are not replacing homework time with extracurricular activities, and so modeling these additional time usages is not necessary. However, the positive homework-participation correlation does suggest that homework hours should be controlled for in the regression analysis.

27. In some specifications that I tested, I also included socioeconomic status in tenth grade to capture family inputs and influences around the time of the dropout decision. It is highly correlated with at-risk status, however, and so the effects of being at-risk are much less precisely estimated but do not affect my conclusions.

highest teacher salary in the school. To proxy for the amount of time students spend in school, I include the number of school days per year and the number of classes per day. These variables are from the 1990 administrator questionnaire. Missing data are imputed using school size, urban status, and Census Division, and I include an indicator equal to one if the variable was imputed.²⁸

Finally, peers and neighbors may have an influence over students' demands for various after-school activities and on the net expected value of high school completion. Additionally, neighborhood variables are also proxies for the labor market conditions that students face upon dropping out of high school. Therefore, I include the share of black and Hispanic tenth grade students and the share of students enrolled in remedial math or English programs, the median household income in the school's zip code, and the share of black and Hispanic adults in the school's zip code.²⁹ I also include an indicator for urban status and indicators for the nine Census Divisions. Missing peer and zip code variables are also imputed and I include an indicator in the regression to account for this.³⁰

I use two instrumental variables that represent constraints to participation. The first is the total number of extracurriculars offered per hundred students in the school, where the numerator and denominator are from the administrator's questionnaire.³¹ The main drawback from a data perspective is in how the data were coded. The administrator was asked to reply yes or no to whether certain extracurricular activities were offered by the school, but these extracurriculars were highly aggregated and were not divided by gender. For instance, the administrator was asked whether sports were provided by the school, but information detailing the range of sports and the variation in sports across gender was not gathered. Thus, the measure that I use is limited in scope due to these data constraints.

The second instrument that I use is the presence of NPNP laws in the state. This variable was derived from the state-level rules listed in Lapchick (1989). I have coded this list to reflect the presence of NPNP laws using an indicator variable equal to one if the state had the following types of rules: (1) students must pass all courses in the previous semester, (2) students must have a minimum GPA, (3) students must not fail more than one course in the previous semester, and (4) a combination of minimum GPA along with course-passing requirements.³² Many states had some type

28. Only 10 schools were missing teacher–student ratios but 200 schools were missing the share of teachers with a master's degree or higher. Teacher salary information was missing for 170 schools, and the school year and classes per day were missing for 110 schools.

29. In some specifications, I also included the share of tenth graders in the free lunch program, although it is highly correlated with at-risk status, to control for current schooling conditions that exert an influence on the student's decisions. This exclusion does not affect my conclusions.

30. Only twenty schools were missing these variables.

31. There were twenty-two categories that administrators were asked about. These categories are: honors society, band, choir, orchestra, computer club, drama club, service clubs, math club, science club, history club, foreign language club, other subject-based clubs, student government, newspaper, yearbook, religion-based clubs, debate, interscholastic sports, intramural sports, vocational clubs, cheerleading, and international clubs. For schools with valid data, the minimum number of extracurriculars offered was six. When the count of extracurriculars was missing, this minimum was assigned. I tried other methods to impute missing data, but the change in results was not significant.

32. In earlier work, I used a categorical version of this instrument instead of a binary variable. The results were nearly identical to the main results in this paper, but in some specifications, the categorical version had less statistical significance than the binary version, so I opted to use the binary version instead.

of NPNP laws by 1989, but there were still several states (e.g., Maryland and New York) that did not. These laws are particularly appealing as potential instruments because many states were enacting these laws at or around the late 1980s, so students attending school during this time period would have had little opportunity to relocate to another state in anticipation of these laws. However, given that many states had some form of NPNP law in effect by 1990, there may be relatively little variation for identification.³³

Raw correlations indicate that both instruments are correlated with the participation decision, and in the direction expected. The correlation coefficient between participation and the number of offerings per hundred students is 0.10 (p -value of 0.00), showing that as offerings per student increase, participation increases, and the correlation between participation and the presence of NPNP laws is -0.02 (p -value of 0.07), showing that laws restricting participation are negatively correlated with participation. Further discussion of the instruments, including a discussion of the strength of the instruments in the analysis, are provided in section 5.

Descriptive Statistics

In the final column of table 2, I provide the mean difference between not-at-risk and at-risk students. In general, this comparison shows substantial differences in inputs, characteristics, and decisions of at-risk versus not-at-risk students. At-risk students have a 9.8 percentage point higher dropout rate relative to not-at-risk students, and a 7.4 percentage point lower tenth grade extracurricular participation rate. At-risk students begin high school with relatively lower test scores, and in tenth grade, at-risk students are doing fewer hours of homework and working more hours than their non-at-risk peers. In terms of family, school, and neighborhood characteristics during tenth grade, at-risk students tend to receive fewer educational resources at home and at school, and more frequently attend schools and live in neighborhoods with larger shares of minorities and low-income households.

There is also significant variation in outcomes, characteristics, and inputs by at-risk and extracurricular participation status. With respect to dropout rates, nonparticipants have a dropout rate three times higher than participants (21.2 percent and 7.1 percent, respectively). For not-at-risk students, the dropout rate of 17.5 percent for nonparticipants is more than four times that of participants (4.1 percent).³⁴ For at-risk students, of those who participated, 13.1 percent drop out of high school, while the dropout rate for nonparticipants of 25.8 percent is strikingly high. These statistics suggest a potentially large, negative relationship between extracurricular participation and dropout rates for both at-risk and not-at-risk students.

33. Therefore, as an avenue for future research, I note the timing of NPNP laws as an instrument for the change in participation. This would be similar in nature to the use of Title IX as an instrument to predict changes in female sports participation (Stevenson 2010), though it would apply to both genders and all activities. Additionally, as suggested by an anonymous referee, variation in pay-to-play policies across states would serve as an interesting instrumental variable, though these policies are fairly recent and generally do not coincide with the time period of the NELS:88 data.

34. Supporting calculations are provided in table A.1, which is available in the online appendix.

Table 3. Estimated Effect of Extracurricular Participation on the Dropout Decision

Average Partial Effects	Probit (1)	Bivariate Probit (2)	
Difference in predicted dropout probability by participation status for not-at-risk students	-0.088** (-0.015)	-0.183* (0.088)	
Difference in predicted dropout probability by participation status for at-risk students	-0.072** (0.018)	-0.176* (0.018)	
Difference in predicted dropout probability by at-risk status for participants	0.037*** (0.022)	0.031 (0.022)	
Difference in predicted dropout probability by at-risk status for nonparticipants	0.052** (0.010)	0.038** (0.010)	

Regression Results	Extracurricular Participation Regression		Dropout Regression
1 if student is at-risk	0.169*** (0.100)	-0.170** (0.051)	0.123 (0.105)
1 if participated in any extracurricular club/sport	-0.591** (0.081)		-1.140** (0.333)
1 if participant and at-risk	0.227* (0.114)		0.254* (0.114)
No. of extracurriculars offered per 100 students		0.102** (0.026)	
1 if NPNP laws in the state		0.011 (0.070)	
Pseudo R ²	0.173		
Log pseudo-likelihood	-430281	-1,139,503	
Rho		0.318	
p-value of rho		0.093	

Notes: *N* = 9,480 students in 770 schools. Sample sizes are rounded to the nearest ten to maintain confidentiality. Huber-White standard errors clustered at the school level are shown in parentheses. Panel weights are used in all regressions. All other covariates are included in the regression, and results are shown in Appendix table A.3, which is available in the online appendix.

p* < 0.05; *p* < 0.01; ****p* < 0.10.

5. RESULTS

In table 3, I show the results from regressing tenth grade extracurricular participation, at-risk status, and their interaction on the probability of dropping out (conditional on observables), as well as average partial effects for all combinations of extracurricular participation and at-risk status.³⁵ In the lower panel of the table, the first set of columns shows the probit results of the dropout decision treating extracurricular participation as exogenous, conditional on observables; the second set of columns shows the bivariate probit results from estimating the extracurricular participation and dropout decisions simultaneously.

The probit results in table 3 provide a useful benchmark, showing that extracurricular participation has a statistically significant, negative relationship with the dropout decision that is less negative for at-risk students. The average partial effect of participation for not-at-risk students is -8.8 percentage points, and for at-risk students, -7.2 percentage points. The results are suggestive of a large, negative relationship between extracurricular participation and the likelihood of dropping out. The second two rows

35. The estimated coefficients for all covariates are provided in table A.3 (in the online appendix). All partial effects are calculated holding all other variables at the sample mean.

of average partial effects show that, regardless of participation status, being at-risk increases the likelihood of dropping out, although the estimates are larger for those who are nonparticipants.

Whereas the probit specification treats extracurricular participation as exogenous, the bivariate probit model predicts extracurricular participation using exogenous constraints to participation. The average partial effects are shown in column 2 in the top panel, and the second set of columns in the lower panel of table 3 shows the estimated coefficients from the extracurricular participation and the dropout regressions, respectively.

Focusing first on the exogeneity and strength of the instruments, the coefficients of the instrumental variables indicate that more activities offered per student increase participation, and although NPNP laws have no significant independent effect on the participation decision, the sign of the coefficient indicates that these laws increase participation on average, potentially by reducing the number of eligible students. These instruments are jointly significant at the 5 percent level, with a Wald test F -statistic of 13.20 (p -value: 0.0014).^{36, 37}

Another concern is that the instruments, primarily offerings per student, are not exogenous due to positive selection into schools with more resources. To address this concern, I show summary statistics for the instruments by at-risk status in table 2, which show that there is no clear difference in the mean between groups. If more advantaged students were selecting into schools with better resources or fewer participation constraints, I would expect to observe not-at-risk students attending schools with more extracurriculars per student. I also regressed eighth grade math and reading test scores on the vector of covariates and instruments to find whether, conditional on covariates, there was a significant relationship between the instruments and prior skills. These results showed neither a large nor a statistically significant relationship of the instruments with prior skills, suggesting that students with higher initial skill are not selecting into schools with more offerings per student or states without NPNP laws.³⁸ Though this analysis is not irrefutable evidence of exogeneity, it suggests no clear pattern of sorting or selection.

To address the effect of extracurricular participation, I show the average partial effects of participation by at-risk status, and of being at-risk by participation status from the bivariate probit specification. As shown in the upper panel of table 2, column 2, I find that, all else equal, the average partial effects of participation for not-at-risk and at-risk students, respectively, are now -18.3 percentage points and -17.6 percentage points.

-
36. A Lagrange multiplier test is a more common test for the strength of instruments when using a probit regression for an endogenous regressor. This test is not valid with clustered standard errors, however. Therefore, I rely on a Wald test instead. Unfortunately, there are not more rigorous weak instrument tests in the bivariate probit setting (Nichols 2011).
37. To verify that identification is not solely due to nonlinearity in the bivariate probit model, I also estimate linear probability models using two-stage least squares but *excluding* the interaction between the at-risk and participation indicators. This specification shows similar results to the bivariate probit model, though with less precision and less strength of instruments. I rely on the probit and bivariate probit models because the linear model results in more than 16 percent of the sample with a negative predicted probability of dropping out. Furthermore, the results using the nonlinear probability models are more precisely estimated.
38. The t -statistics of the estimated coefficients of instruments in these specifications range from 0.14 to 0.90, indicating no statistically significant relationship between prior test scores and the instruments, conditional on additional covariates. The tables of regression results are available upon request.

These findings indicate that reducing constraints to extracurricular participation will lead to a substantial reduction in dropout rates for those affected. Additionally, within participation categories, I find that with all else equal, at-risk students remain more likely to drop out. For nonparticipants, the effect of being at-risk is a 3.8 percentage point increase in the likelihood of dropping out.

It is important to note that larger magnitudes are to be expected in the bivariate probit model due to the local average treatment effect (LATE) interpretation of the results using instrumental variables. The instruments reflect constraints to participation, and so only affect those who, due to constraints, are forced to change their participation status. Therefore, it is not appropriate to compare these results to the entire population, but only to interpret these results as LATEs.³⁹ This group of students, however, is the group of policy interest, because these are the students who will be affected by cuts or expansions of extracurricular offerings.

With the LATE interpretation in mind, these findings suggest that there is negative selection into participation for those constrained. The increase in magnitude of the estimates after instrumenting is a common finding in the extracurricular participation literature (e.g., Eide and Ronan 2001; Kosteas 2010; Stevenson 2010). For example, Eide and Ronan (2001) find that the estimates of sports participation on college graduation increase from 9.0 percentage points to 14.4 percentage points when moving from ordinary least squares to IV estimation. The intuition is that without constraints, students at all points of the achievement/ability spectrum are able to participate if they choose. Yet with constraints, those at either end of the spectrum are unlikely to be affected by the instruments (the “always” or “never” participants), whereas marginal students are likely to be negatively affected by the constraints. Specifically, when there are fewer opportunities to participate, or when there are GPA minimums or other grade-related restrictions, this is likely to exclude the average (or just below average) student, leading to larger LATE estimates than probit estimates.

Extensions

Next, to determine whether these effects differ with the type of activity in which the student engages, I divide extracurricular participation into four categories. The bivariate probit results and average partial effects are shown in table 4. As shown in the first set of columns, I find that for athletic participation, the average partial effect is a 16.5 percentage point decline in the likelihood of dropping out for not-at-risk students. The effect is negligibly more negative for at-risk students, for whom participation results in a 16.6 percentage point reduction in the likelihood of dropping out. The findings also show that regardless of participation category, being at-risk increases the likelihood of dropping out, specifically for nonparticipants. Note that the instruments are particularly strong in this specification, with a Wald F -statistic of 30.11 (p -value of 0.001). These results show that the benefits from athletics are important for reducing high school attrition for those facing constraints to participation.

39. The population average includes those who would never participate, regardless of constraints, as well as those who would nearly always participate, and so are unaffected by the instrument. See Angrist, Imbens, and Rubin (1996) for discussion of LATE interpretation of IV results.

Table 4. Estimated Effect of Athletics, Arts, Academic, and Other Club Participation on the Dropout Decision

	Athletic Participation		Arts Participation		Academic Club Participation		Other Club Participation	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average Partial Effects from Bivariate Probit								
Diff. in predicted dropout probability by participation status for not-at-risk students	-0.165** (0.069)							-0.208** (0.032)
Diff. in predicted dropout probability by participation status for at-risk students	-0.166** (0.065)							-0.217** (0.028)
Diff. in predicted dropout probability by at-risk status for participants	0.025*** (0.015)							0.031** (0.009)
Diff. in predicted dropout probability by at-risk status for non-participants	0.023** (0.005)							0.021** (0.004)
Regression Results								
1 if student is at-risk	-0.216** (0.042)	-0.038 (0.048)	-0.148** (0.044)	0.373** (0.071)	0.320** (0.059)	-0.004 (0.045)	0.183** (0.054)	
1 if student participated in athletics		-1.390** (0.285)						
1 if athletic participant and at-risk		0.252** (0.096)						
1 if student participated in art clubs				-0.610 (0.549)				
1 if art club participant and at-risk				-0.279* (0.123)				
1 if student participated in academic clubs					0.947*** (0.496)			
1 if academic club participant and at-risk					0.078 (0.114)			
1 if student participated in other clubs								-1.724** (0.111)
1 if other club participant and at-risk								0.238** (0.084)

Table 4. Continued.

Regression Results	Athletics Reg.	Dropout Reg.	Arts Reg.	Dropout Reg.	Academics Reg.	Dropout Reg.	Other Club Reg.	Dropout Reg.
No. of extracurriculars offered per 100 students	0.099** (0.018)		0.033 (0.023)		0.007 (0.022)		0.069** (0.018)	
1 if NPNP laws in the state	0.046 (0.065)		-0.054 (0.071)		-0.147*** (0.075)		0.091 (0.065)	
Log pseudo-likelihood		-1,473,872		-1,334,033		-1,421,581		-1,436,636
Rho		0.586		0.297		-0.650		0.883
p-value of rho		0.012		0.415		0.039		0.000

Notes: $N = 9,480$ students in 770 schools. Sample sizes are rounded to the nearest ten to maintain confidentiality. Huber-White standard errors clustered at the school level are shown in parentheses. Panel weights are used in all regressions. All other covariates are included in the regression, but results are not shown.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.10$.

For art and academic clubs, shown in the second and third sets of columns, the instruments are not strong predictors for either activity group, and so the causal effects of these activities are not identified.⁴⁰ The weakness of the IVs in these specifications presents an interesting robustness check for the instruments and their relevance. The strongest instrument in nearly every specification is the offerings per student, which is intended to capture the effect of competition within the school on the students' participation. Athletics and other clubs (e.g., student government) are highly competitive given that the number of positions is limited, whereas art clubs (e.g., orchestra) and academic clubs (e.g., math club) tend to have a variable number of participants and are unlikely to lead to such competition. Therefore, the weakness of this instrument in predicting art or academic club participation serves to validate the intuition of this instrument. With respect to NPNP laws, though they are negatively correlated with academic club participation, the instruments are not jointly statistically significant, leading to potentially incorrect signed estimates in the bivariate probit model.

As shown in the fourth set of estimates in table 4, I find similar results for other club participation as I found for athletics. The instruments are jointly significant in the participation equation (Wald *F*-statistic of 16.56 [*p*-value of 0.0003]), and the average partial effects indicate large negative effects of participation on the probability of dropping out. I find that for not-at-risk students, the average partial effect of other club participation is 20.8 percentage points, and for at-risk students, the average partial effect is 21.7 percentage points. This indicates that participation in other clubs reduces the likelihood of dropping out overall, though slightly more for at-risk students.

Because of the strength of the instruments in the athletics and other club specifications in table 4, I use a variation of these specifications where the extracurricular participation indicator equals one if the student participated in athletics or "other" clubs and equals zero otherwise. I provide the estimates from bivariate probit model in table 5. I find that for not-at-risk students, participants are 21.0 percentage points less likely to drop out than had they not participated, and for at-risk students, participants are 19.9 percentage points less likely to drop out than had they not participated. This suggests that the results in table 3 are somewhat understated due to the inclusion of activities for which the instruments are weak. Additionally, note that the average partial effect of at-risk status remains fairly stable across specifications in tables 3, 4, and 5.

Robustness

Next, I estimate several variations of the earlier specifications to test whether the results are sensitive to alternative definitions of at-risk and to test how the conclusions would change had I performed the analysis in a way more comparable to previous studies.

In table 6, I present results from specifications identical to the bivariate probit specification from table 3, but replacing my definition of at-risk with Kagaruki-Kakoti's and Finn's definitions, respectively. As shown, Kagaruki-Kakoti's definition of at-risk yields bivariate probit results that show slightly smaller effects of participation than the results in table 3: Not-at-risk participants are 13.2 percentage points less likely to drop out than had they not participated, and at-risk participants are 17.1 percentage

40. The Wald *F*-statistics of the instruments in the arts and academics specifications, respectively, are 3.31 (*p*-value of 0.19) and 4.01 (*p*-value of 0.13).

Table 5. Estimated Effect of Athletic and Other Club Participation on the Dropout Decision

Average Partial Effects from Bivariate Probit	Athletics and Other Clubs Only	
Difference in predicted dropout probability by participation status for not-at-risk students	-0.210** (0.075)	
Difference in predicted dropout probability by participation status for at-risk students	-0.199** (0.07)	
Difference in predicted dropout probability by at-risk status for participants	0.024 (0.02)	
Difference in predicted dropout probability by at-risk status for non-participants	0.035** (0.006)	
Regression Results	Athletics and Other Club Regression	Dropout Regression
1 if student is at-risk	-0.146** (0.046)	0.100 (0.087)
1 if student participated in athletics or other clubs		-1.400** (0.267)
1 if athletic/other club participant and at-risk		0.324** (0.098)
No. of extracurriculars offered per 100 students	0.117** (0.021)	
1 if NPNP laws in the state	0.055 (0.065)	
Log pseudo-likelihood		-1,384,286
Rho		0.498
<i>p</i> -value of rho	0.005	

Notes: *N* = 9,480 students in 770 schools. Sample sizes are rounded to the nearest ten to maintain confidentiality. Huber-White standard errors clustered at the school level are shown in parentheses. Panel weights are used in all regressions. All other covariates are included in the regression, but results are not shown.

** *p* < 0.01.

points less likely to dropout than had they not participated. The differences in these estimates relative to those in table 3 are likely due to several differences in our at-risk measures. For instance, Kagaruki-Kakoti assumes that students are at-risk if they live with a stepparent, but it is not clear that this is a risk factor that would be associated with a disadvantaged background. Turning to the estimates using Finn’s definition, I find quite different results, although this indicator is created using the same variables as the indicator upon which I rely. In the results shown in column 2, being at-risk is not associated with an increase in the probability of dropping out, and there is no effect of the interaction term between the at-risk and participation indicators on the dropout decision. This suggests that at-risk students are no more or less likely to graduate relative to their peers, and no more or less likely to benefit from extracurriculars, which is inconsistent with the model and with the summary statistics. These results illustrate that cutoffs for defining at-risk status should be consistent with definitions of socioeconomic disadvantage, otherwise there is little detectable difference in at-risk and not-at-risk students.⁴¹

41. I perform a similar analysis by replacing the at-risk indicator with race/ethnicity and include controls for socioeconomic status. I find that the interaction between extracurricular participation and race is not statistically significant, suggesting that it is factors related to socioeconomic status rather than race that are disadvantageous. These results are available upon request.

Table 6. Estimated Effect of Extracurricular Participation on the Dropout Decision Using Alternative Definitions of At-Risk

Average Partial Effects from Bivariate Probit	Using Kagaruki-Kakoti's At-Risk Status		Using Finn's At-Risk Status	
	(1)	(2)	(1)	(2)
Difference in predicted dropout probability by participation status for not-at-risk students	-0.132 (0.082)	-0.136*** (0.080)		
Difference in predicted dropout probability by participation status for at-risk students	-0.178* (0.086)	-0.141*** (0.079)		
Difference in predicted dropout probability by at-risk status for participants	0.079** (0.023)	0.009 (0.025)		
Difference in predicted dropout probability by at-risk status for nonparticipants	0.033** (0.006)	0.004 (0.007)		

Regression Results	Extracurricular Participation		Extracurricular Participation	
	Regression	Dropout Regression	Regression	Dropout Regression
1 if student is at-risk	-0.239** (0.044)	0.330** (0.113)	-0.127* (0.060)	0.040 (0.112)
1 if participated in any extracurricular club/sport		-0.921* (0.365)	0.100** (0.026)	-0.860* (0.348)
1 if participant and at-risk		-0.001 (0.121)	-0.000 (0.070)	0.003 (0.123)
No. of extracurriculars offered per 100 students	0.096** (0.026)			
1 if NPNP laws in the state	0.003 (0.069)			
Log pseudo-likelihood		-1,135,382		-1,145,242
Rho		0.268		0.218
p-value of rho		0.163		0.254

Notes: N = 9,480 students in 770 schools. Sample sizes are rounded to the nearest ten to maintain confidentiality. Huber-White standard errors clustered at the school level are shown in parentheses. Panel weights are used in all regressions. All other covariates are included in the regression, and results are available upon request.

Sources: Kagaruki-Kakoti 2005; Finn 2006.

*p < 0.05; **p < 0.01; ***p < 0.10.

In table 7, I present results from a specification similar to that used in previous studies. In this specification, I remove the at-risk indicator and interaction term, but include a control for current socioeconomic status. I estimate each specification for the tenth grade participation indicator and for the tenth grade athletic participation indicator, because these are the two measures of participation that are common in the literature. In table 7, I first show the results from using the tenth grade participation indicator. The results indicate that participants are 13.4 percentage points less likely to drop out than had they not participated. From this specification, the effect of participation on the dropout decision is much smaller in magnitude (in absolute value) than what I found in table 3. In the final set of columns in table 7, I show the results associated with tenth grade athletic participation, which indicate that the effect of athletic participation is a 14.5 percentage point reduction in the probability of dropping out. These results are much closer in magnitude to what I presented in table 4, although this specification still underestimates the effect of participation for at-risk students. Overall, these findings show that participation has a smaller effect on the dropout decision relative to the results in table 3 and table 4, suggesting that when

Table 7. Estimated Effect of Extracurricular Participation on the Dropout Decision after Removing the At-Risk Indicator

	Extracurricular Participation Regression	Dropout Regression (1)	Dropout Marginal Effects	Extracurricular Participation Regression	Dropout Regression (2)	Dropout Marginal Effects
Regression Results from Bivariate Probit						
1 if participated in any extracurricular club/sport		-0.922** (0.299)	-0.134** (0.046)			
1 if participated in athletics					-0.937** (0.323)	-0.145* (0.062)
No. of extracurriculars offered per 100 students	0.103** (0.026)			0.100** (0.019)		
1 if NPNP laws in the state	-0.011 (0.069)			0.009 (0.064)		
Log pseudo-likelihood		-1,131,759	-		-1,467,206	-
Rho		0.269			0.326	
p-value of rho		0.122			0.136	

Notes: *N* = 9,480 students in 770 schools. Sample sizes are rounded to the nearest ten to maintain confidentiality. Huber-White standard errors clustered at the school level are shown in parentheses. Panel weights are used in all regressions. Marginal effects are calculated at the sample mean. All other covariates, including tenth grade socioeconomic status, are included in the regression, and results are available upon request.

p* < 0.05; *p* < 0.01.

not taking into account the differential effect of participation by at-risk status, the effect of participation on the dropout decision is understated.

6. DISCUSSION AND CONCLUSION

In this study, I distinguish between at-risk and not-at-risk students based on disadvantages in their background prior to entering high school. A priori, there are obvious differences between these two groups in terms of dropout rates, extracurricular participation, and observable characteristics. I rely on the results for which the instruments are strong predictors—the results using overall participation, athletic, and “other club” participation. These results indicate that for those affected by the instruments, their likelihood of dropping out is substantially reduced by extracurricular participation (relative to having been unable to participate), and these findings are robust to a number of variations. In addition, I find that in several instances, extracurricular participation has a comparable effect on the likelihood of dropping out for at-risk and not-at-risk students, specifically with respect to athletics. Overall, these findings suggest that policies to increase access to extracurriculars would indeed result in lower dropout rates, in particular for at-risk students who may face more constraints to participation than their peers. For instance, many recent policies (e.g., No Pass/No Play and Pay-to-Play) tend to restrict participation—more so for disadvantaged groups of students.

How cost-effective would programs be to maintain or expand extracurriculars? Although there are not data available in the NELS:88 regarding expenditures on extracurriculars, current data on fees may be used to address this question. The average Pay-to-Play fees in Ohio (a state that recently enacted Pay-to-Play policies) are estimated at \$140 per student per school year.⁴² Therefore, a back-of-the-envelope calculation

42. See www.springfieldnewssun.com/news/news/local-education/pay-to-play-fees-could-be-banned-in-ohio/nnnMn/.

suggests that an average high school of 1,200 students, with all students participating in one activity, would generate approximately \$168,000 in Pay-to-Play revenue per year—roughly equivalent in cost to hiring three additional teachers and reducing the student–teacher ratio by 0.2 standard deviations.⁴³ Though eliminating Pay-to-Play fees would not induce all students to participate, this policy will have a large impact on those who are induced to participate, and the small reduction in the student–teacher ratio is likely to have a negligible effect on outcomes. Lastly, it is worth noting that in my preliminary work, I found a large correlation between eighth and tenth grade extracurricular participation, suggesting that participation may be persistent and may have persistent effects. This would suggest that promoting and facilitating extracurriculars at an early age may be particularly beneficial as well.⁴⁴

ACKNOWLEDGMENTS

Many thanks to Audrey Light, Bruce Weinberg, Lucia Dunn, Manisha Goel, Dimitrios Nikolaou, Seth Gershenson, participants at the SEA, AEFPP, and MEA conferences and seminar participants at Ohio State University, as well as colleagues and two anonymous referees, for helpful discussions and suggestions. Funding for the completion of this paper was provided by Saint Joseph’s University through its summer research grant program. All errors are my own.

REFERENCES

- Anderson, Deborah J. 2001. If you let me play: The effects of high school athletics on students’ educational and labor market success. Unpublished paper, University of Arizona.
- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91(434):444–455. doi:10.1080/01621459.1996.10476902.
- 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. doi:10.1162/003465300558902.
- Bright, Rebecca. 2011. *Public school, for a price*. Available <http://neatoday.org/2011/06/15/public-school-for-a-price/>. Accessed 12 March 2012.
- Burnett, Michael A. 2000. “One strike and you’re out”—An analysis of no pass/no play policies. *High School Journal* 84(2):1–6.
- Cameron, Stephen V., and James J. Heckman. 1993. The nonequivalence of high school equivalents. *Journal of Labor Economics* 11(1):1–47. doi:10.1086/298316.
- Cameron, Stephen V., and James J. Heckman. 2001. The dynamics of educational attainment for black, Hispanic, and white males. *Journal of Political Economy* 109(3):455–499. doi:10.1086/321014.
- Catsiapis, George. 1987. A model of educational investment decisions. *Review of Economics and Statistics* 69(1):33–41. doi:10.2307/1937898.

43. The average teacher salary in public secondary schools is \$53,520 in 2011–12 (USDOE 2015). In my sample, the average number of teachers is 92.2, with an average student–teacher ratio of 16.19 (standard deviation = 4.34).

44. Therefore, as data become available, in future studies I intend to explore the timing and duration of extracurricular participation.

Costa, Cristiano M. 2010. Extracurricular activities and wage differentials. Unpublished Paper, University of Pennsylvania.

Covay, Elizabeth, and William Carbonaro. 2010. After the bell: Participation in extracurricular activities, classroom behavior, and academic achievement. *Sociology of Education* 83(1):20–45. doi:10.1177/0038040709356565.

C.S. Mott Children's Hospital. 2012. *Pay-to-play sports keeping lower-income kids out of the game*. Available www.mottnpch.org/sites/default/files/documents/051412paytoplayreport.pdf. Accessed 22 June 2016.

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. doi:10.1016/S0272-7757(00)00033-9.

Feldman, Amy F., and Jennifer L. Matjasko. 2005. The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. *Review of Educational Research* 75(2):159–210. doi:10.3102/00346543075002159.

Finn, Jeremy D. 2006. *The adult lives of at-risk students: The roles of attainment and engagement in high school (NCES 2006–328)*. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Graves, Goss, Neena Chaudry Fatima, Katharine Gallagher Robbins, Anne Morrison, Lauren Frohlich, Adaku Onyeka-Crawford, Philip Tegeler, Michael Hilton, and Silva Mathema. 2015. *Finishing last: Girls of color and school sports opportunities*. Washington, DC: National Women's Law Center and Poverty and Race Research Action Council.

Hanushek, Eric A. 1979. Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources* 14(3):351–388. doi:10.2307/145575.

Hill, C. Russell. 1979. Capacities, opportunities and educational investments: The case of the high school dropout. *Review of Economics and Statistics* 61(1):9–20. doi:10.2307/1924826.

Hill, C. Russell, and Frank P. Stafford. 1977. Family background and lifetime earnings. In *The distribution of economic well-being*, edited by F. Thomas Juster, pp. 511–556. Cambridge, MA: Ballinger.

Hungerford, Thomas, and Gary Solon. 1987. Sheepskin effects in the returns to education. *Review of Economics and Statistics* 69(1):175–177. doi:10.2307/1937919.

Ingels, Steven J., Kathryn L. Dowd, John D. Baldridge, James L. Stipe, Virginia H. Bartot, and Martin R. Frankel. 1994. *National education longitudinal study of 1988: Second follow-up: Student component data file user's manual (NCES 94–374)*. Washington, DC: National Center for Education Statistics U.S. Department of Education.

Ingels, Steven J., Leslie A. Scott, John R. Taylor, Jeffrey Owings, and Peggy Quinn. 1998. *National education longitudinal study of 1988 (NELS:88) base year through second follow-up: Final methodology report*, Working Paper No. 98–06. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Jaeger, David A., and Marianne E. Page. 1996. Degrees matter: New evidence on sheepskin effects in the returns to education. *Review of Economics and Statistics* 78(4):733–740. doi:10.2307/2109960.

Kagaruki-Kakoti, Generosa. 2005. An economic analysis of school and labor market outcomes for at-risk youth. Unpublished paper, Georgia State University.

Kosteas, Vasilios D. 2010. High school clubs participation and earnings. Unpublished Paper, Cleveland State University.

Kuhn, Peter, and Catherine Weinberger. 2005. Leadership skills and wages. *Journal of Labor Economics* 23(3):395–436. doi:10.1086/430282.

Lapchick, Richard E. 1989. *Pass to play: Student athletes and academics*. Washington, DC: National Education Association.

Lipscomb, Stephen. 2007. Secondary school extracurricular involvement and academic achievement: A fixed effect approach. *Economics of Education Review* 26(4):463–472. doi:10.1016/j.econedurev.2006.02.006.

Mahoney, Joseph L., and Robert B. Cairns. 1997. Do extracurricular activities protect against early school dropout? *Developmental Psychology* 33(2):241–253. doi:10.1037/0012-1649.33.2.241.

McNeal, Ralph B., Jr. 1995. Extracurricular activities and high school dropouts. *Sociology of Education* 68(1):62–80. doi:10.2307/2112764.

Nichols, Austin. 2011. Causal inference for binary regression. Paper presented at Stata Conference, Chicago, June.

Rees, Daniel, and Joseph Sabia. 2010. Sports participation and academic performance: Evidence from the national longitudinal study of adolescent health. *Economics of Education Review* 29(5):751–759. doi:10.1016/j.econedurev.2010.04.008.

Schoenberg, Judy, Kamla Modi, Kimberlee Salmond, Mark Mather, and Linda Jacobsen. 2013. *The state of girls: Unfinished business, extracurricular and out-of-school activities*. Available www.girlscouts.org/content/dam/girlscouts-gsusa/forms-and-documents/about-girl-scouts/research/sog_full_report.pdf. Accessed 24 August 2016.

Schultz, Theodore W. 1963. *The economic value of education*. New York: Columbia University Press.

Simon, Stephanie. 2011. Public schools charge kids for basics, frills. *The Wall Street Journal*, 25 May.

Spreitzer, E. 1994. Does participation in interscholastic athletics affect adult development? A longitudinal analysis of an 18–24 age cohort. *Youth & Society* 25(3):368–387. doi:10.1177/0044118x94025003004.

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. doi:10.1162/rest.2010.11623.

U.S. Census Bureau. 1988. *Poverty thresholds for 1988 by size of family and number of children*. Available www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html. Accessed 12 June 2012.

U.S. Department of Agriculture (USDA). 2011. *National school lunch program*. Available www.fns.usda.gov/cnd/lunch/AboutLunch/NSLPFactSheet.pdf. Accessed 12 June 2012.

U.S. Department of Education (USDOE). 1965. *Title I of the Elementary and Secondary Education Act of 1965 (20 U.S.C. 6301 et seq.)*. Washington, DC: USDOE.

U.S. Department of Education (USDOE). 2002. *Public law 107–110—Jan. 8, 2002*. Available www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf. Accessed 1 June 2016.

U.S. Department of Education (USDOE). 2010. *Improving basic programs operated by local educational agencies (Title I, Part A)*. Available www2.ed.gov/programs/titleiparta/index.html. Accessed 12 June 2012.

U.S. Department of Education (USDOE). 2015. *Table 211.10: Average salaries for full-time teachers in public and private elementary and secondary schools, by selected characteristics: 2011–12*. Available https://nces.ed.gov/programs/digest/d14/tables/dt14_211.10.asp?current=yes. Accessed 22 June 2016.

Willis, Robert J., and Sherwin Rosen. 1979. Education and self-selection. *Journal of Political Economy* 87 (5):S7–S36. doi:10.1086/260821.

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross-sectional and panel data*. Cambridge, MA: MIT Press.