Spring 5-17-2010

Extracurricular Activities and Wage Differentials

Cristiano M. Costa
costam@sas.upenn.edu

Extracurricular Activities and Wage Differentials

Abstract
This dissertation develops and estimates a structural dynamic model in which individuals decide among participation in extracurricular activities, study time, and employment during high school. The specification allows for returns to participation in extracurricular activities on salaries and unobserved heterogeneity in students' preferences over study time, extracurricular activities, employment during high school and the acquisition of a bachelor's degree. The model is estimated using a selected sample of 1,875 white males from NELS88. The objective is to evaluate quantitatively the consequences of public policies that eliminate the opportunity for students to participate in extracurricular activities and policies that give payments to students if they obtain higher test scores. Results of counterfactual experiments show that policies that eliminate extracurricular activities in high schools decrease students' tests scores, college graduation rates, and future wages, while policies that give payments based on test scores increase scores and the percentage of students who obtain bachelor's degrees.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Economics

First Advisor
Kenneth Wolpin

Second Advisor
Petra Todd

Third Advisor
Flavio Cunha

Keywords
extracurricular activities, structural model, wage differentials, non-cognitive skills

Subject Categories
Economics

EXTRACURRICULAR ACTIVITIES AND WAGE DIFFERENTIALS

Cristiano Machado Costa

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

2010

Supervisor of Dissertation

____

Kenneth I. Wolpin, Professor, Department of Economics

Graduate Group Chairperson

____

Dirk Krueger, Professor, Department of Economics

Dissertation Committee

Kenneth I. Wolpin, Professor, Department of Economics

Petra E. Todd, Professor, Department of Economics

Flávio Cunha, Assistant Professor, Department of Economics

EXTRACURRICULAR ACTIVITIES AND WAGE DIFFERENTIALS

To my lovely wife Luciana.

ACKNOWLEDGMENTS

I thank Kenneth Wolpin for his insightful guidance and support throughout the development of this dissertation, and Petra Todd and Flávio Cunha for valuable comments, suggestions and encouragement. I also would like to thank Áureo de Paula, Roberto Pinheiro, Andrew Postlewaite, David Russo, and Ilton Soares for comments on earlier versions of this research. Seminar participants at FUCAPE Business School, Inter-American Development Bank (IADB) and the Empirical Microeconomics Club at University of Pennsylvania also provided valuable feedback. Financial support from the Ministry of Education of Brazil (CAPES-Fulbright Program) is gratefully acknowledged. I am responsible for all errors.

ABSTRACT

EXTRACURRICULAR ACTIVITIES AND WAGE DIFFERENTIALS

Cristiano M. Costa

Kenneth I. Wolpin

This dissertation develops and estimates a structural dynamic model in which individuals decide among participation in extracurricular activities, study time, and employment during high school. The specification allows for returns to participation in extracurricular activities on salaries and unobserved heterogeneity in students' preferences over study time, extracurricular activities, employment during high school and the acquisition of a bachelor's degree. The model is estimated using a selected sample of 1,875 white males from NELS88. The objective is to evaluate quantitatively the consequences of public policies that eliminate the opportunity for students to participate in extracurricular activities and policies that give payments to students if they obtain higher test scores. Results of counterfactual experiments show that policies that eliminate extracurricular activities in high schools decrease students' tests scores, college graduation rates, and future wages, while policies that give payments based on test scores increase scores and the percentage of students who obtain bachelor's degrees.

Contents

List of Tables

Chapter 1

Introduction

In the United States, more than 70% of high-school students participate in athletics or school clubs on a weekly basis.[1] Participation in extracurricular activities is often linked to higher test scores, educational attainment, and future income.[2] For example, average test scores of students who participate in athletics or clubs are 7.5% higher than scores obtained by students who do not participate. On average, 51% of students who participate in athletics or clubs during the four years of high school obtain a bachelor's degree, while this rate is only 24% among those who do not participate. Frequent participants in extracurricular activities earn about 20% more than nonparticipants eight years after high-school graduation.

One potential explanation for these differences is that participation in extracurricular activities facilitates the development of non-cognitive skills (motivation, leadership, self-esteem, etc.), which are not well measured by academic test scores or

[1] All statistics in this paragraph are based on a sample of 7,131 students from the National Educational Longitudinal Study of 1988 (NELS88) for which all information was available. Please refer to Section 3 or to the Appendix for details about the definitions and construction of variables.

[2] Recent economic literature reports positive effects of participation in extracurricular activities on test scores, educational attainment, and wages during adulthood. See, for example, Lipscomb (2007), Barron, Ewing and Waddell (2000), Eide and Ronan (2001), Persico, Postlewaite and Silverman (2004), and Kuhn and Weiberger (2005).

schooling measures, like years of education. Recent studies have shown that involvement in extracurricular activities is related to higher aspirations for the future, the likelihood of avoiding the development of antisocial behavior problems, and low levels of depression and anxiety (Mahoney, Larson and Eccles, 2005). Studies also report that participation is associated with a greater desire to help others, higher self-esteem and better decision-making skills (Miller, 2003).

At the same time, the literature on the technology of skill formation reports that non-cognitive skills promote the development of cognitive skills (Cunha and Heckman, 2007; Cunha and Heckman, 2008). For example, there is evidence that programs that focus on the acquisition of non-cognitive skills are associated with increased engagement in learning and test scores during adolescence (Miller, 2003). These results point to a complementary effect of the acquisition of non-cognitive skills on outcomes usually related to cognitive skills, like test scores. For example, skills learned in school clubs, like assertiveness or critical thinking, may improve academic performance. Additionally, research shows that non-cognitive skills directly influence schooling decisions and wages during adulthood (Heckman, Stixrud, and Urzua, 2006). Finally, skills like teamwork, leadership, and motivation are frequently listed as very important on employers' descriptions of necessary qualifications for jobs in today's economy (Murnane and Levy, 1996).

If participation in extracurricular activities actually helps the development of non-cognitive skills, the decision to participate in sports and school clubs can be interpreted as an investment. One would expect the student to anticipate future benefits and take them into account when deciding to participate in extracurricular activities. The decision depends not only on expected future outcomes, but also on students' preferences. Spending time on extracurricular activities takes time away from other activities, such as studying, working and leisure. Students face a trade-off between investing in non-cognitive skills (extracurricular activities) and investing in cognitive skills (study time) during high school.

In this context, public policies that emphasize academic achievement (test scores), may ignore the important role of non-cognitive skills in the formation of cognitive skills.[3] These types of policies may induce students to spend more time studying and doing homework and less time participating in extracurricular activities. Also, policies that reduce the availability of sports and clubs during high school may reduce students' opportunities to acquire non-cognitive skills valuable later in life. To capture the result of such policies accurately, hwoever, a model of high-school decision-making has to take into account the fact that students will change their choices once an extracurricular activity is no longer available and may compensate for this lack of opportunity by studying more or acquiring more working experience. Therefore, to estimate the potential consequences of such policies a model must account for the nature of the student's choices during high-school years and its consequences for future outcomes, like college graduation and wages during adulthood.

This article develops and estimates a structural dynamic model in which high-school students make decisions concerning study time, participation in athletics, participation in clubs, and employment during high school. Students face a time constraint, which leads to trade-offs among their choices. They derive utility or disutility from the activities on which they spend time (study, extracurricular activities and work) and derive utility from leisure. After graduation, students decide either to enter

[3] In particular, there is evidence that non-cognitive skills promote the formation of cognitive skills, but, in general, the opposite is not true (Cunha and Heckman, 2007). Additionally, research suggests that non-cognitive skills are easier to acquire at later ages than cognitive skills (Carneiro and Heckman, 2003).

the labor market or to obtain a bachelor's degree and postpone entrance into the labor market for four years. The model allows for unobserved heterogeneity in students' preferences, ability to take exams, and productivity in the labor market. Students' choices during high school affect outcomes throughout their lives, such as standardized test scores, wages during high school and salaries during their adulthoods. Preferences and outcomes are affected by individual shocks that are observed by the students but unobserved by the econometrician. The specification allows for different returns to participation in extracurricular activities on salaries for those holding a bachelor's degree and for those with a high-school diploma. This feature captures the idea that non-cognitive skills, acquired through participation in extracurricular activities, may be more important at certain educational levels than at others.

The economic literature on the effects of participation in extracurricular activities on wages during adulthood is relatively sparse. Persico, Postlewaite and Silverman (2004) investigate the relationship between participation and future wages using a sample of white males from the National Longitudinal Survey of Youth (NLSY). Results from ordinary least squares (OLS) estimates show a wage differential of about 9% in adult wages for those who participate in athletics and 3% for those who participate in clubs, after controlling for ability (AFQT) and schooling measures. Kuhn and Weiberger (2005) use data on white males from the Project TALENT, the National Longitudinal Study of 1972 (NLS-72) and High School and Beyond (HSB) to estimate the effects of leadership positions during high school on wages during adulthood. Results from different OLS estimates show that students who participate in sports and school clubs receive a wage differential between 2% and 10%, while those who were also team captains or club presidents receive additional wage premiums ranging from zero to more than 20%.

Because participation in extracurricular activities is a choice, a simple regression of wages on participation in athletics and clubs may lead to a biased estimator, due to self-selection.[4] To control for endogeneity of participation in athletics, Eide and Ronan (2001) use height as an instrument for athletic participation. After controlling for high-school characteristics, family income, and other observables, the effect of participation in athletic activities on adult wages is not statistically significant for a sample of white males from the HSB survey. Barron, Ewing and Waddell (2000) use high-school enrollment and measures of health, family income, and location as instruments. The authors control for ability (AFQT), years of education and other individual characteristics and report wage premiums of about 4% for males in the NLS-72, but the effect is not statistically significant for students in the NLSY cohort. The authors point out that these results depend on the quality of the instruments for athletic participation that they use.

In order to account for potential endogeneity problems and to assess the consequences of different public policies, the model developed in this article is structurally estimated using a sample of 1,875 white males from the National Educational Longitudinal Study of 1988 (NELS88). This survey is particularly informative because it contains information about students' participation in athletic activities (baseball, basketball, football, soccer, swimming, and other team and individual sports) and in school clubs (band, dramatic or musical, government, newspaper, and service clubs) for more than one period. Students were interviewed when they were sophomores (1990) and again when they were seniors (1992). Subjects also answered questions

[4]Suppose, for example, that students who are more productive in the labor market are also the ones who derive higher utility from participation in sports, due to some unobserved characteristic. Then, a simple OLS regression of wages on participation in sports would overestimate the effects of participation in sports on wages, and lead to biased results.

about their post-secondary education and work experience later in life (in 1994 and in 2000). The is structured in a dynamic discrete-choice model framework and is estimated using the Simulated Maximum Likelihood (SML) method.

Results show that participation in athletics is estimated to increase wages in adulthood up to 6.7% for individuals with a bachelor's degree and up to 4.2% for those with only a high-school diploma. The wage premium for participation in school clubs is estimated to be up to 5.7% for college graduates and 3.9% for those with a high-school diploma only. These results are similar to those found in the previous literature. Because the model is structural, these estimates are already free of potential biases caused by unobserved heterogeneity and self-selection. We can then conclude that the labor market is willing to pay more to workers with these kinds of experience.

The amount of experience accumulated in each activity directly impacts the choice of educational level (bachelor's degree or high-school diploma only), because experiences affect future wages in each educational level differently. Unobserved heterogeneity in students' preferences play an important role on determining who participates in school activities. In particular, estimates of the utility parameters indicate that all activities are costly (parameters are negative) but are more costly for certain groups of students. Those with strong preferences for clubs derive lower utility levels from participating in sports, but their disutility from studying is the lowest among all types. Therefore, this type of student tends to participate more in clubs, and, because the payoffs for such experience are more valuable after obtaining a bachelor's degree, this type of student anticipates that and attends going to college more often than the average student.

Estimates also show that test scores increase with study time and participation in clubs and decrease with participation in sports and work. These results can be interpreted as evidence that skills acquired thriugh clubs help the development of cognitive skills, as measured by test scores, while skills acquired through athletic activities or part-time jobs are detrimental to the acquisition of cognitive skills and result in lower test-score levels.

Having estimated the parameters governing preferences and outcomes, it is possible to evaluate the impact of changes in public policies on choices during high school, the choice of educational level, and wages during adulthood. Four different policies are analyzed. The first three experiments investigate the effects of eliminating clubs, athletic activities, and both types of activities simultaneously. These policy experiments aim to measure the long-term consequences of recent proposed cuts in high-school budgets, which usually have extracurricular activities as prime candidates for cutbacks. Results show that public policies that eliminate expenditures in extracurricular activities may lead to lower test scores, fewer students obtaining bachelor's degrees, and lower wages during adulthood. The fourth experiment intends to estimate the consequences of policies that transfer income to students who perform better in exams, in the same fashion as policies recently implemented in Chicago, Washington, and New York. Results show that a policy that pays 700 dollars to every student who obtains a score above the median of the test distribution increases test score and the number of students obtaining bachelor's degrees.

This dissertation is organized as follows. Chapter 2 outlines the dynamic-decision model and the estimation strategy, Chapter 3 describes the NELS88 sample used in the empirical analysis, and Chapter 4 presents the results of the structural estimation. Chapter 5 provides the results of four policy experiments, and Chapter 6 concludes.

Chapter 2

Model and Estimation Strategy

This Chapter presents a dynamic-choice model in which high-school students make decisions on how to allocate their time among study, work, and extracurricular activities. These choices affect their current utility in each period of high school and lead to trade-offs during those years. The decisions may affect their test score results, and the experience accumulated in each of these activities during high school affects the students' post-secondary schooling decisions and their wages during adulthood.

2.1 Timing

The model has three periods, $t = 1, 2, 3$. The first period corresponds to a student's first two years of high school: freshman and sophomore years. He chooses how to allocate his time among study, work, and extracurricular activities. At the end of the first period, he observes the outcomes of his choices: test scores and wages, and accumulates experience in extracurricular activities and work. In the second period, which corresponds to the last two years of high school, junior and senior years, he again chooses among the same options as in the first period and observes the resulting outcomes. In the third period, the student decides whether to enter the labor market or obtain a college degree. The third-period utility is the sum of all future wages after leaving high school. If the student decides to obtain a college degree, the direct benefit/cost is added to his utility, and his entry into the labor market is delayed by four years.

2.2 The Choice Set

During high-school years ($t = 1, 2$), the choice set consists of four binary decisions: study ($d_s^t \in \{0, 1\}$), participation in athletics ($d_a^t \in \{0, 1\}$), participation in clubs ($d_b^t \in \{0, 1\}$), and work ($d_w^t \in \{0, 1\}$). Upon finishing high school ($t = 3$), the student decides whether to enter the labor market ($d_c = 0$) or to obtain a college degree ($d_c = 1$). In the first and second periods, the decision set allows for sixteen ($16 = 2^4$) mutually exclusive and exhaustive choices. In the third period, there are only two alternatives.

2.3 The State Space

Working and participating in extracurricular activities build skills in the form of cumulative experience. The transition function is deterministic and is simply the sum of the past decisions. For any $k = a, b, w$: $e_k^{t+1} = e_k^t + d_k^t$. The initial experience (e_k^1) is assumed to be equal to zero for all students. For example, participating in athletics and working in the first period, $(d_a^1, d_b^1, d_w^1) = (1, 0, 1)$, and participating in athletics and clubs while not working in the second period, $(d_a^2, d_b^2, d_w^2) = (1, 1, 0)$, will yeld cumulative experience of $(e_a^3, e_b^3, e_w^3) = (2, 1, 1)$. While time spent on working and participating in extracurricular activities accumulates over time, study time does not.

In the first and second periods, individuals receive random utility shocks that shift their utility levels of studying, participating in athletics and clubs, and working.

These shocks should be understood as individual choice-specific characteristics that are related to the utility of joining the activities but are not observed by the econometrician. The four random shocks are known by the students at the time they make their decisions. The individual shocks are drawn from a multivariate normal distribution and are allowed to be correlated within periods, $(u_s^t, u_a^t, u_b^t, u_w^t)' \sim N(\mathbf{0}, \Sigma)$. The diagonal elements are denoted σ_s^2, σ_a^2, σ_b^2 , and σ_w^2, and covariance elements are σ_{sa}, σ_{sb}, σ_{sw}, σ_{ab}, σ_{aw}, and σ_{bw}. The matrix Σ is assumed to be same in periods one and two. In the third period, students also receive a choice-specific shock (u_c), which is assumed to be normally distributed, $u_c \sim N(0, \sigma_c^2)$. Additionally, students face shocks on their wage offers during high school: $\xi^t \sim N(0, \sigma_{\xi^t}^2)$. This shock is assumed to be observed by the students before they decide. Over the first and second periods, students also face a shock (ϵ^t) on their test scores, which is not known to the students before they decide how to allocate their time and is assumed to be normally distributed, $\epsilon^t \sim N(0, 1)$.

The state space for an individual i at time t, $\Omega(i,t)$, describes the accumulated experience ($e_i^t \equiv (e_{ai}^t, e_{bi}^t, e_{wi}^t)'$), current-period utility shocks u_i^t (where $u_i^t \equiv (u_{si}^t, u_{ai}^t, u_{bi}^t, u_{wi}^t)'$ for $t = 1, 2$, and $u_i^3 \equiv (u_{ci}^3)$), wage offer shock ξ_i^t, and the scores in the previous period, g_i^{t-1}. Hence, the state space for individual i at time t is $\Omega(i,t) = \left\{e_i^t, u_i^t, \xi_i^t, g_i^{t-1}\right\}$, and $g_i^0 = \phi$ for all i.

2.4 Outcomes

While in high school, individuals receive wage offers that are assumed be constant and independent of students' skills and choices. The wage offers in the first two periods $(t = 1, 2)$ are modeled as the sum of a constant term and an individual-specific shock:

$$w_i^t = \exp(w^t + \xi_i^t),$$

where w^t is constant across students at time t and ξ_i^t is the individual-specifi component of the wage offer at time t. The shock is known by the students when they decide whether to work.

The test score result for student i in period t is a function of his decisions, $d_i^t = (d_{si}^t, d_{ai}^t, d_{bi}^t, d_{wi}^t)'$. This is the mechanism by which students transform their experiences or skills learned in extracurricular activities, into cognitive skills (test scores). This specification allows for positive and negative direct effects. For example, participation in the newspaper club may increase students' test scores because they learn how to manage time better or it may decrease scores because students get confused with some parts of the material learned in class because they were focused on memorizing lines for a play to be presented in the drama club.

The test score process also depends on a stochastic component (ϵ_i^t) unknown to the student and assumed to be independent over time and across individuals. Scores in the second period are allowed to depend on scores in the first period to account for accumulated knowledge.

Students take exams in the first period and receive a score, g_i^{1*} , which is assumed to follow the process:

$$g_i^{1*} = \rho_{11} d_{si}^1 + \rho_{12} d_{ai}^1 + \rho_{13} d_{bi}^1 + \rho_{14} d_{wi}^1 + \epsilon_i^1,$$

where $\rho_1^\intercal = (\rho_{11}, \rho_{12}, \rho_{13}, \rho_{14})$ is a vector of parameters and ϵ_i^1 is a shock to test scores. What is observed, however, is the final score, g_i^1, which is assumed to be discrete and ranges over integervalues from one to four. Therefore, the process can be modeled as an ordered probit:

$$
\begin{aligned}
P(g_i^1 = 1) &= \Phi\left(\mu_1^1 - \rho_1^\intercal d_i^1\right) \\
P(g_i^1 = 2) &= \Phi\left(\mu_2^1 - \rho_1^\intercal d_i^1\right) - \Phi\left(\mu_1^1 - \rho_1^\intercal d_i^1\right) \\
P(g_i^1 = 3) &= \Phi\left(\mu_3^1 - \rho_1^\intercal d_i^1\right) - \Phi\left(\mu_2^1 - \rho_1^\intercal d_i^1\right) \\
P(g_i^1 = 4) &= 1 - \Phi\left(\mu_3^1 - \rho_1^\intercal d_i^1\right),
\end{aligned}
$$

where μ_1^1, μ_2^1, and μ_3^1 are parameters such that $\mu_1^1 < \mu_2^1 < \mu_3^1$, $d_i^1 = (d_{si}^1, d_{ai}^1, d_{bi}^1, d_{wi}^1)'$, and $\Phi(.)$ is the cumulative normal distribution function. Because the student does not know ϵ_i^1 at the beginning of the period and the shock is normally distributed, he infers that the probabilities of receiving the scores as described above. In the second period, scores are allowed to depend on the first-period scores, and the process is given by:

$$
g_i^{2*} = \rho_{20} g_i^1 + \rho_{21} d_{si}^2 + \rho_{22} d_{ai}^2 + \rho_{23} d_{bi}^2 + \rho_{24} d_{wi}^2 + \epsilon_i^2,
$$

where ϵ_i^2 is also independent and normally distributed, $\epsilon_i^2 \sim N(0,1)$. Probabilities over the second-period scores are also modeled as an ordered probit as described above, but with parameters μ_1^2, μ_2^2, and μ_3^2.

In the final period, students decide between two alternatives: enter the labor market or obtain a college degree. These choices lead to two different wage profiles. The wage profiles for each educational level, high-school diploma (HS) or college degree (C), depend on second period scores, experience in athletics, clubs and work, and age (a). The wages (W^{HS} and W^C) are assumed to be log-normally distributed, and each include random components unknown by the student when finishing high school.

The wage profile faced by individual i at age a is written as follows:

$$W^{HS}(i,a) = \exp(\alpha_0^{HS} + \alpha_1^{HS}(a-18) + \alpha_2^{HS}(a-18)^2 + \alpha_3^{HS}g_i^2 + \alpha_4^{HS}I_{\{e_{ai}^3=1\}}$$
$$+\alpha_5^{HS}I_{\{e_{ai}^3=2\}} + \alpha_6^{HS}I_{\{e_{bi}^3=1\}} + \alpha_7^{HS}I_{\{e_{bi}^3=2\}} + \alpha_8^{HS}I_{\{e_{wi}^3=1\}} + \alpha_9^{HS}I_{\{e_{wi}^3=2\}} + \varepsilon_i^{HS}(a))$$
$$\text{for } a \geq 18$$

and

$$W^{C}(i,a) = \exp(\alpha_0^{C} + \alpha_1^{C}(a-22) + \alpha_2^{C}(a-22)^2 + \alpha_3^{C}g_i^2 + \alpha_4^{C}I_{\{e_{ai}^3=1\}}$$
$$+\alpha_5^{C}I_{\{e_{ai}^3=2\}} + \alpha_6^{C}I_{\{e_{bi}^3=1\}} + \alpha_7^{C}I_{\{e_{bi}^3=2\}} + \alpha_8^{C}I_{\{e_{wi}^3=1\}} + \alpha_9^{C}I_{\{e_{wi}^3=2\}} + \varepsilon_i^{C}(a))$$
$$\text{for } a \geq 22$$

where α^{HS} and α^{C} are vectors of parameters and $I_{\{x=y\}}$ is an indicator function that equals one if the statement in braces is true and zero otherwise. The shocks on wages to high-school graduates ($\varepsilon_i^{HS}(a)$) and to college graduates ($\varepsilon_i^{C}(a)$) are assumed to be independent across individuals and ages and normally distributed: $\varepsilon_i^{HS}(a) \sim N(0, \sigma_{HS}^2)$ and $\varepsilon_i^{C}(a) \sim N(0, \sigma_C^2)$.

2.5 Preferences and Student's Problem

In the first and second periods, students choose how to use their time, $d_i^t = (d_{si}^t, d_{ai}^t, d_{bi}^t, d_{wi}^t)'$. They derive utility/disutility from each these four activities. They also derive utility from consumption (c_i^t) if they work and from leisure (l_i^t), which is measured as the time available after choosing d_i^t. Test scores (g_i^t) do not give immediate utility but affect future outcomes, such as second-period scores and wage profiles after high-school graduation (W^{HS} and W^{C}).

The problem of student i in period t, for $t = 1$ or 2, is:

$$\max_{\{d_i^t\}} \left\{ U_t(c_i^t, l_i^t, d_i^t, \Omega(i,t)) + \beta^2 E V_{t+1}(\Omega(i,t+1) | d_i^t, \Omega(i,t)) \right\}$$

$$s.t. \qquad d_i^t \in \{0,1\}^4$$

$$c_i^t = w_i^t d_{wi}^t$$

$$l_i^t = 4 - d_{si}^t - d_{ai}^t - d_{bi}^t - d_{wi}^t$$

where U_t is the utility in period t, β is the discount factor, and V_{t+1} is the utility in all subsequent periods, which will be defined in the following paragraph.[1] The contemporaneous utility at t is:

$$U_t(c_i^t, l_i^t, d_i^t, \Omega(i,t)) = c_i^t + \psi^t \sqrt{l_i^t} + (\gamma_1^t + u_{si}^t) d_{si}^t + (\gamma_2^t + u_{ai}^t) d_{ai}^t + (\gamma_3^t + u_{bi}^t) d_{bi}^t + (\gamma_4^t + u_{wi}^t) d_{wi}^t$$

where ψ^t is a parameter that represents the preference for leisure in period t, γ^t is a vector of parameters that assign the direct utility of enrolling in an activity in period t, and u_{si}^t, u_{ai}^t, u_{bi}^t, and u_{wi}^t are random shocks that shift the utility/disutility levels. The features of the utility function are easy to interpret. The utility is measured in monetary terms because it is linear in consumption. The preference for leisure is concave, so the marginal utility of leisure is decreasing. The fact that leisure is linear in d_i^t implies that the disutility that comes from decreasing leisure time by one unit is the same for all four activities, and it is transformed into monetary terms by ψ^t. It is also important to mention that the expected continuation value depends not only on the distribution of the future utility and wage offer shocks but also on the distribution of shocks to scores during period t, which is unknown to the student at period t.

In the third period, students have finished high school and choose whether to enter the labor market or to obtain a college degree and enter the labor market four years

[1] The next-period expected value is discounted at the rate β^2 because each of the first two periods is equivalent to two years in the data.

later. If students decide to enter the labor market, their utility values are equal to the expected discounted value of the future wages, $E(V_i^{HS})$. If they decide to obtain a college degree, their utility value will be the utility/disutility of attending college $\left(\gamma^C\right)$ plus the expected discounted value of the future wages, $E(V_i^C)$. This utility is also affected by a shock on preferences (u_{ci}) that is known by the student by the time he has to make his choice. The student's problem in the third period is:

$$V_3(d_{ci}, \Omega(i,3)) = \max_{d_{ci}\in\{0,1\}} \{E(V_i^{HS}),\ \gamma^C + E(V_i^C) + u_{ci}\}$$

where

$$E(V_i^{HS}) = E\left[\sum\nolimits_{a=18}^{65} \beta^{(a-18)} W^{HS}(i,a)\right]$$
$$E(V_i^{C}) = E\left[\sum\nolimits_{a=22}^{65} \beta^{(a-18)} W^{C}(i,a)\right]$$

The utility/disutility of attending college $\left(\gamma^C\right)$ should be understood as the sum of all pecuniary and non-pecuniary costs and benefits from attending a four-year college.

When deciding between the two educational paths, students do not know the realizations of future shocks on wages. Therefore, they make their decisions based on the expected discounted values of both educational wage profiles. They know, however, the wage profile parameters and the distribution of the shocks, and have sufficient information to calculate the expected value of the two options. The wage functions can be rewritten as

$$W^{HS}(i,a) = \exp(\alpha_1^{HS}(a-18) + \alpha_2^{HS}(a-18)^2)\exp(\alpha_0^{HS} + \alpha_3^{HS} g_i^2 + \alpha_4^{HS} I_{\{e_{ai}^3=1\}}$$
$$+\alpha_5^{HS} I_{\{e_{ai}^3=2\}} + \alpha_6^{HS} I_{\{e_{bi}^3=1\}} + \alpha_7^{HS} I_{\{e_{bi}^3=2\}} + \alpha_8^{HS} I_{\{e_{wi}^3=1\}} + \alpha_9^{HS} I_{\{e_{wi}^3=2\}})\exp(\varepsilon_i^{HS}(a))$$
$$\text{for } a \geq 18$$

and

$$W^C(i,a) = \exp(\alpha_1^C(a-22) + \alpha_2^C(a-22)^2)\exp(\alpha_0^C + \alpha_3^C g_i^2 + \alpha_4^C I_{\{e_{ai}^3=1\}}$$
$$+\alpha_5^C I_{\{e_{ai}^3=2\}} + \alpha_6^C I_{\{e_{bi}^3=1\}} + \alpha_7^C I_{\{e_{bi}^3=2\}} + \alpha_8^C I_{\{e_{wi}^3=1\}} + \alpha_9^C I_{\{e_{wi}^3=2\}})\exp(\varepsilon_i^C(a))$$
$$\text{for } a \geq 22.$$

This leads to the following simplification:

$$W^{HS}(i,a) = \varphi^{HS}(a)\exp(\alpha_0^{HS} + \phi^{HS}X_i)\exp(\varepsilon_i^{HS}(a)) \quad \text{for } a \geq 18$$

and

$$W^C(i,a) = \varphi^C(a)\exp(\alpha_0^C + \phi^C X_i)\exp(\varepsilon_i^C(a)) \qquad \text{for } a \geq 22,$$

where

$$\varphi^{HS}(a) = \exp(\alpha_1^{HS}(a-18) + \alpha_2^{HS}(a-18)^2)$$
$$\varphi^C(a) = \exp(\alpha_1^C(a-22) + \alpha_2^C(a-22)^2)$$
$$\phi^{HS} = (\alpha_3^{HS}, \alpha_4^{HS}, \alpha_5^{HS}, \alpha_6^{HS}, \alpha_7^{HS}, \alpha_8^{HS}, \alpha_9^{HS})$$
$$\phi^C = (\alpha_3^C, \alpha_4^C, \alpha_5^C, \alpha_6^C, \alpha_7^C, \alpha_8^C, \alpha_9^C)$$
$$X_i = (g_i^2, I_{\{e_{ai}^3=1\}}, I_{\{e_{ai}^3=2\}}, I_{\{e_{bi}^3=1\}}, I_{\{e_{bi}^3=2\}}, I_{\{e_{wi}^3=1\}}, I_{\{e_{wi}^3=2\}})^\intercal.$$

Hence, it is possible to rewrite the expected discounted value as:

$$E\left[V_i^{HS}\right] = \exp(\alpha_0^{HS} + \phi^{HS}X_i)\textstyle\sum_{a=18}^{65}\beta^{(a-18)}\varphi^{HS}(a)E\left[\exp(\varepsilon_i^{HS}(a))\right]$$
$$E\left[V_i^C\right] = \exp(\alpha_0^C + \phi^C X_i)\textstyle\sum_{a=22}^{65}\beta^{(a-18)}\varphi^C(a)E\left[\exp(\varepsilon_i^C(a))\right]$$

Because the wage errors, $\varepsilon_i^{HS}(a)$ and $\varepsilon_i^C(a)$, are normally distributed, the terms $\exp(\varepsilon_i^{HS}(a))$ and $\exp(\varepsilon_i^C(a))$ are log-normally distributed.

As a result, $E[\exp(\varepsilon_i^{HS}(a))] = \exp(\sigma_{HS}^2/2)$, and $E[\exp(\varepsilon_i^C(a))] = \exp(\sigma_C^2/2)$. Therefore, the expressions can be written as:

$$E\left[V_i^{HS}\right] = \exp(\alpha_0^{HS} + \phi^{HS} X_i + \sigma_{HS}^2/2)\lambda^{HS}$$
$$E\left[V_i^C\right] = \exp(\alpha_0^C + \phi^C X_i + \sigma_C^2/2)\lambda^C$$

where $\lambda^{HS} = \sum_{a=18}^{65} \beta^{(a-18)} \varphi^{HS}(a)$ and $\lambda^C = \sum_{a=22}^{65} \beta^{(a-18)} \varphi^C(a)$. The values of λ^{HS} and λ^C are direct functions of $\beta, \alpha_1^{HS}, \alpha_2^{HS}, \alpha_1^C$ and α_2^C; they are not additional parameters.

2.6 Solution Method

The model is solved by backward induction. In the last period students know the relevant information to make their decisions at each state point $(\Omega(i,3))$. The optimal decision (d_{ci}) maximizes the lifetime value $V_3(d_{ci}, \Omega(i,3))$. Knowing these maximum alternative-specific value functions for every element of the state space at the third period, we can integrate over the utility shock (u_{ci}) to obtain the expected maximum alternative-specific value function $(EV_3(\Omega(i,3))$ for each element in the state space at the beginning of period three.

At the beginning of the second period, the students know the relevant information $(\Omega(i,2)$ and $EV_3(\Omega(i,3))$, at each attainable $\Omega(i,3))$ to make their decisions. In this period, however, before computing the optimal decision, a student needs asses the probability of each test score outcome to form beliefs about the expected value of the utility in the third period, as each decision at the beginning of the second period affects the distribution of scores at the end of the second period. After integrating over the test score shock (ϵ_i^2), students have all the information they need (including $EV_3(\Omega(i,3)|\Omega(i,2), d_i^2)$ for all d_i^2) to make their decisions in the second period at each state point $(\Omega(i,2))$.

In the first period, they calculate the maximum alternative-specific value function for every element of the state space at the second period $V_2(\Omega(i,2))$ and integrate over the test score shock (ϵ_i^1) and second-period shocks (u_i^2) to obtain $EV_2(\Omega(i,2)|\Omega(i,1),d_i^1)$. Given this information, a student chooses the alternative that maximizes his expected utility.

2.7 Estimation Strategy

The model is a three-period multinomial choice problem and can be estimated using the Simulated Maximum Likelihood method. The method consists of obtaining the so-called Emax values for periods two and three by simulating the errors given the set of parameters and then using these Emax vakues to calculate the likelihood. There are, however, two minor complications in the process: (i) the errors on test scores are not known before the decision is made, and (ii) there is no closed form solution for choice probabilities in the first and second periods.

The first complication affects the way the Emax values are calculated. Since test scores are known only at the end of the period, Emax values are calculated in two steps. In the second period, $EV_3(\Omega(i,3))$ is obtained by calculating the utility value at each state point, which includes the scores (g_i^2), via simulation[2], and after that integrating over the four possible final scores ($g_i^2 = 1, 2, 3,$ or 4) to obtain $EV_3(\Omega(i,3)|\Omega(i,2),d_i^2))$. The probability of each score level, however, has a closed form solution. This procedure is also used in the first period.

The second complication (no closed form solution for the choice probabilities)

[2] The Emax values were calculated using 200 draws. For each state point and draw, the maximum utility is calculated. For each state point, the Emax value is simply the average of the maximum utilities over the 200 draws.

requires the use of a simulator for the calculations of the probabilities. The method consists of drawing many sets of shocks for each student, calculating the values of the utility function in each draw, and calculating the probabilities of each possible choice given the observed state point and expected future values.[3]

Since shocks are independent across time, the likelihood function can be written as the product of within-period outcome probabilities:

$$\mathcal{L}(\theta|y,x) = \prod_{i=1}^{I}\prod_{t=1}^{3} P(y(i,t)|\theta, x(i,t))$$

where θ is the vector of parameters, y is the set of all observed decisions/outcomes, and x is the set of all observables (past decisions). $P(y(i,t)|\theta, x(i,t))$ is the probability of observing the decisions and outcomes $y(i,t)$ for the individual i at period t, conditional on the vector of parameters θ and the observation of $x(i,t)$.

The decisions/outcomes are $y(i,1) = [d_i^1, w_i^1, g_i^1]$, $y(i,2) = [d_i^2, w_i^2, g_i^2]$ and $y(i,3) = [d_i^3, W_i]$. Notice that, if the student chooses $d_i^3 = 0$, then we observe $W^{HS}(i, a = 26)$ and if $d_i^3 = 1$, then we observe $W^C(i, a = 26)$[4]. The observables are such that $x(i,3) = [e_i^3, g_i^2]$, $x(i,2) = [e_i^3, g_i^1]$, and $x(i,1) = \emptyset$.

In the last period, the conditional probability is:

$$\begin{aligned} P(y(i,3)|\theta, x(i,3)) &= P(d_i^3, W_i(a=26)|\theta, e_i^3, g_i^2) \\ &= P(d_i^3|\theta, e_i^3, g_i^2) P(W_i(a=26)|\theta, e_i^3, g_i^2, d_i^3) \end{aligned}$$

[3]The method used in this article is a *logit-smoothed AR simulator*, as in Train (2003, Chapter 5.6.2). This simulator was originally suggested by McFadden (1989). Ben-Akiva and Bolduc (1996) call it a "logit-kernel probit," when applied to a probit model. The smooth-simulated probabilities for the first and second periods were calculated based on 200 draws. The smoothing parameter, τ, was set to be equal to 500. Please refer to Eckstein and Wolpin (1999) for an example of use of this simulator.

[4]In NELS88, the dataset used in the estimation, the only infomation available about wages after high school is when students are 26 years old. This issue is discussed in Chapter 3.10.

$$= I_{\{d_i^3=1\}} P(\gamma^C + E\left[V_i^C\right] + u_{ci} > E\left[V_i^{HS}\right] | \theta, e_i^3, g_i^2) P(W_i = W^C(i, a=26) | \theta, e_i^3, g_i^2, d_i^3)$$
$$+ I_{\{d_i^3=0\}} P(E\left[V_i^{HS}\right] > \gamma^C + E\left[V_i^C\right] + u_{ci} | \theta, e_i^3, g_i^2) P(W_i = W^{HS}(i, a=26) | \theta, e_i^3, g_i^2, d_i^3).$$

The probability is calculated by first calculating the values of $E\left[V_i^{HS}\right]$ and $E\left[V_i^C\right]$ for which there are closed-form solutions[5]. After that it is possible to calculate $P(\gamma^C + E\left[V_i^C\right] + u_c i > E\left[V_i^H S\right] | \theta, e_i^3, g_i^2)$ and $P(E\left[V_i^H S\right] > \gamma^C + E\left[V_i^C\right] + u_c i | \theta, e_i^3, g_i^2)$, using the cumulative normal distribution, since $u_c \sim N(0, \sigma_c^2)$. The probabilities $P(W_i = W^C(i, a = 26) | \theta, e_i^3, g_i^2)$ and $P(W_i = W^H S(i, a = 26) | \theta, e_i^3, g_i^2)$ also have a closed-form solution, the density of a standard normal distribution, since $\varepsilon_i^H S(a) \sim N(0, \sigma_H S^2)$ and $\varepsilon_i^C(a) \sim N(0, \sigma_C^2)$.

In the second period the likelihood contribution does not have a closed-form solution. The probability is:

$$\begin{aligned} P(y(i,2)|\theta, x(i,2)) &= P(d_i^2, w_i^2, g_i^2 | \theta, e_i^2, g_i^1) \\ &= P(g_i^2 | \theta, d_i^2, w_i^2, e_i^2, g_i^1) P(d_i^2, w_i^2 | \theta, e_i^2, g_i^1) \\ &= P(g_i^2 | \theta, d_i^2, w_i^2, e_i^2, g_i^1) P(d_i^2 | \theta, w_i^2, e_i^2, g_i^1) P(w_i^2 | \theta, e_i^2, g_i^1). \end{aligned}$$

where $P(g_i^2 | \theta, d_i^2, w_i^2, e_i^2, g_i^1)$ is obtained from an ordered probit, $P(w_i^2 | \theta, e_i^2, g_i^1)$ comes from a standardized normal density and $P(d_i^2 | w_i^2, e_i^2, g_i^1, k)$ is simulated using shocks drawn from a multivariate normal distribution. The simulation, however, has to take into account the fact that the econometrician observes the realization of the shock on wages, $\xi_i^2 = \ln(w_i^2) - w^2$, for those who worked in period two, as it affects the individual utility.

[5] The probability $P(d_i^3, W_i | \theta, e_i^3, g_i^2)$ can be rewritten as $P(d_i^3 | \theta, e_i^3, g_i^2) P(W_i | \theta, e_i^3, g_i^2, d_i^3)$ because the decision to attend college does not depend specifically on the errors on $W_i(i, a = 26)$. Hence, observing $W_i(i, a = 26)$ does not change the probability of $P(d_i^3 | \theta, e_i^3, g_i^2)$. This happens because we can write both $E\left[V_i^C\right]$ and $E\left[V_i^{HS}\right]$ in a closed form solution that is a function of the wage profile parameters, and not of its future errors.

The first-period contribution to the likelihood is:

$$\begin{aligned} P(y(i,1)|\theta, x(i,1))) &= P(d_i^1, w_i^1, g_i^1|\theta) \\ &= P(g_i^1|\theta, d_i^1, w_i^1)P(d_i^1, w_i^1|\theta) \\ &= P(g_i^1|\theta, d_i^1, w_i^1)P(d_i^1|\theta, w_i^1)P(w_i^1|\theta). \end{aligned}$$

and the procedure is the same as in the second period.

It is important to remember that the Emax values for the future periods are required for the calculations of $P(y(i,2)|\theta, x(i,2))$ and $P(y(i,1)|\theta, x(i,1)))$. Hence, the recursive method to maximize $\mathcal{L}(\theta|y,x)$ proceeds as follows: (1) choose the starting values for the parameters and distribution of types, (2) calculate the Emax values using the frequency simulator method, (3) calculate $\mathcal{L}(\theta|y,x)$, and (4) use a maximization method to find a better set of parameters and distribution of types. Repeat the steps from (1) until the likelihood is maximized.

2.8 Unobserved Heterogeneity

To allow for unobserved differences in preferences, test-taking ability, and labor market ability, assume that type 1 individuals are characterized by the following vector of parameters:

$$[\gamma_1^1, \gamma_2^1, \gamma_3^1, \gamma_4^1, \gamma_1^2, \gamma_2^2, \gamma_3^2, \gamma_4^2, \gamma^C, w^1, w^2, \mu_1^1, \mu_2^1, \mu_3^1, \mu_1^2, \mu_2^2, \mu_3^2, \alpha_0^{HS}, \alpha_0^C],$$

which was previously defined. The other types ($k > 1$) are defined by their differences relative to type 1.

For $k > 1$, a student of type k will have parameters

$$
\begin{aligned}
&\gamma_{jk}^{t} = \gamma_{j}^{t} + \chi_{jk} && \text{for } t = 1,2 \text{ and } j = 1,2,3,4 \\
&\gamma_{k}^{C} = \gamma^{C} + \chi_{5k} && \\
&w_{k}^{t} = w^{t} + \chi_{6k} && \text{for } t = 1,2 \\
&\mu_{jk}^{t} = \mu_{j}^{t} + \chi_{7k} && \text{for } t = 1,2 \text{ and } j = 1,2,3 \\
&\alpha_{0k}^{HS} = \alpha_{0}^{HS} + \chi_{8k} && \\
&\alpha_{0k}^{C} = \alpha_{0}^{C} + \chi_{9k}. &&
\end{aligned}
$$

Hence, students may have different levels of direct utility from each decision (χ_{1k}, χ_{2k}, χ_{3k}, χ_{4k}), including the utility/disutility from a college degree (χ_{5k}). They may also differ in their productivity when working during high school (χ_{6k}), in their ability to take test (χ_{7k}), and in productivity in market work determinated by each of the two educational levels during adulthood (χ_{8k}, χ_{9k}).

Although heterogeneity is unobserved by the econometrician, it is assumed that there are K types. Also, it is assumed that each type k is observed with proportion π_k in the sample, such that $\sum_{k=1}^{K} \pi_k = 1$, and $\pi_k > 0$, $\forall k$. In this version of the model the number of types is set to be equal to four, $K = 4$. Therefore, allowing for unobserved heterogeneity adds 27 ($= 9 \times 3$) parameters to the model, as we use type 1 as the base type. Also, the proportions of each type (π_1, π_2, and π_3) add 3 more parameters to the model.

Unobserved heterogeneity leads to a slightly different likelihood function. The likelihood in this case will be a finite mixture of type-specific likelihoods:

$$
\mathcal{L}(\theta|y,x) = \prod_{i=1}^{I} \sum_{k=1}^{K} \pi_k \prod_{t=1}^{3} P(y(i,t)|\theta, x(i,t), k).
$$

The advantages of this method are its simplicity and flexibility. The unobserved heterogeneity is modeled using 27 parameters to capture differences in preferences, ability in test-taking, and market productivity. If one tries to capture these characteristics using observables, like family background variables, it would require 9 times the number of observed variables. The model leads to a flexible distribution of types, as it does not impose a parametric assumption on the distribution of types.

Chapter 3

Data

This Chapter describes the sample used in the empirical estimation and the methods by which variables are created from the original survey. Descriptive statistics are presented, as are the results of OLS regressions. Data restrictions and the identification of the model are explained in the final section of the Chapter.

3.1 Sample

The model is estimated using a selected sample from the National Educational Longitudinal Study of 1988 (NELS88). The study is a nationally representative sample of eighth graders surveyed in 1988. Students were interviewed four more times: in 1990, 1992, 1994, and 2000. They responded to questions on a range of topics including school, work, and extracurricular activities. The NELS88 also collected information about certain outcomes during their adulthood, including the highest educational level obtained and wages.

The final sample used in this article includes all white men[1] who attended 10th

[1] The estimation uses only white individuals to avoid measuring wage differentials that could be correlated with race. Also, the sample excludes women to prevent measuring differentials caused by decisions related to childbearing.

grade in 1990, 12th grade in 1992, studied in the same institution throughout high school and obtained a high-school diploma by 1994. The sample also excludes individuals who studied in schools where sports and clubs were both not available. After dropping individuals with missing information, the final sample contains 1,875 individuals.

3.2 Variables

While NELS88 has the drawback of not having many waves after high-school graduation, it contains detailed information about students' participation in extracurricular activities. Among many other questions concerning their choices, students answered questions about the use of their time, including (i) the number of hours per week usually spent doing homework outside the classroom, (ii) participation in organized athletic activities, (iii) participation in organized school clubs, and (iv) number of hours per week spent working. Using this information, discrete variables were created as follows.

The variable *study* (d_s) is set equal to one if the time spent doing homework out of school is greater than five hours per week, and zero otherwise. The variable participation in *athletics* (d_a) is set equal to one if the student participated in at least one school-organized sport (baseball, softball, basketball, football, soccer, swimming, or others). If the student did not participate in any sport, the variable is set equal to zero. The variable participation in *clubs* (d_b) is set equal to one if the student participated in at least one school-organized club (band/orchestra, theater/musical clubs, student government, newspaper/yearbook, or school service clubs). If he did not participate in any of these clubs, the variable is set to zero. The variable *work* (d_w) is set equal to one if the student was employed and reported working more than ten hours per week during high school and zero otherwise.

During the survey, students took tests in many subjects. The variable *test score* (g) was created using the average of the IRT Theta Score obtained by the student across the following subjects: reading, mathematics, science, and history/citizenship/geography. After taking the average, the variable was discretized using the four quartiles levels within each period. Students who worked while in high school reported hourly wages. The variable *high-school wage* (w) is the weekly rate, computed using the hourly wage and assuming the sutdent worked twenty hours per week.

In the fourth follow-up of the survey, students reported their educational level. The variable *college* (d_c) was set equal to one if the student obtained a bachelor's degree or higher by the time of the fourth follow-up. Otherwise, the variable is equal to zero. In this last follow-up, individuals also reported wages received in their current or most recent job. The variable *wage (W)* is the weekly wage reported by the individual. The variable was later separated in two, W^{HS} and W^{C}, according to the educational level obtained by the student (high-school diploma, HS, or bachelor's degree, C).

3.3 Descriptive Statistics

Table 1 displays averages for the variables study, participation in athletics, participation in clubs, and work, during sophomore and senior years. The most popular among the four choices is athletics. Approximately 70.0% of the students participated in sports while in their sophomore year, and 61.3% participated in the activity during senior year. Participation in clubs was 36.4% during sophomore year and 46.1% during senior year. Almost one-fifth of students worked when sophomores, while about two-fifths decided to work during the senior year. 22.5% of students reported having studied more than five hours per week while sophomores. Students increased their study time, on average, as seniors, when 37.3% of students reported spending more than five hours per week studying outside the classroom.

Cross-participation rates in athletics, clubs and labor market are presented in Table 2. The most common combination is clubs and athletics. While 26.6% of sophomores reported to have participated in both activities, 31.2% of students chose to participate in both activities in the later years of high school. Also, there is an increase in participation in all activities during the last two years of study. On the other hand, the group of students who did not participate in any of these three activities shrinks by about thirty percent. It is important to notice that students are reducing their leisure time from one year to another. The distribution of leisure time chosen in each period is presented in Table 3.

Student participation is persistent during the high-school years, as students who participate in a given activity in the first period are more likely to participate in the same activity in the second period. Table 4 displays the transition rates for athletics, clubs and work. The strong persistence in athletics and clubs (about 78%) is evidence that either the students may have heterogeneous preferences over these two activities or that two years of experience is more valuable than one year of experience.

During high school, the average weekly wage received by those who worked was 96.29 dollars during sophomore year and 116.65 dollars in senior year, an increase of 21.1% (see Table 5). This difference may reflect an increase in the productivity of labor between the ages of sixteen and eighteen. Average test scores were 54.05 and 57.77 for sophomores and seniors, respectively (see Table 6). The increase of about seven percent in the average test score can be interpreted as an accumulation of knowledge between periods[2] and is accompanied by an increase of 14.8 percentage points in the number of students who said they had studied more than five hours per week outside the classroom. Table 7 displays the distribution of scores by activity and period. In both periods, students who studied or participated in clubs were more likely to obtain scores above the median of the distribution. Those who participated in athletics do not obtain scores much different than the average student in the sample. Individuals who worked during high school were more likely to be in the bottom of the test score distribution. This pattern is more evident in the the first period than in the second.

Eight years after high-school graduation, in 2000, 44.7% of the individuals had obtained a bachelor's degree. Table 8 presents the same statistic conditional on having participated for at least one year in each of the extracurricular activities. Those who participated in athletics or clubs obtained a bachelor's degree more often than the average student (49.7% and 54.7%, respectively). Students with working experience were less likely to obtain a bachelor's degree, 33.5%.

Table 9 presents the average weekly wages in 2000 by educational level and conditional on having participated in extracurricular activities. On average, individuals who obtained a bachelor's degree received a wage premium of 26.3% over the wages received by individuals with only a high-school diploma. Additionally, students who had participated in clubs received a wage premium of about 8.6% compared to the wages of those who had not participated. Working during high school, however, seems to harm the wages of those holding a bachelor's degree, leading to below-average wages. Students who had participated in athletics received the highest wage premium among

[2] For details about the IRT Test and the IRT Theta Score please refer to *Second Follow-up: Student Component Data File User's Manual (*Ingels, S. J., Dowd, K. L., Baldridge, J. D., Stipe, J. L., Bartot, V. H.; Frankel, M. R., 1994).

college graduates. The difference between average weekly wages of those who had participated in athletics and those who had not is 25.9%. This difference is of the same magnitude as the college premium (26.3%).

Average wages conditional on participation in athletics and clubs are higher among those who decided to enter the labor market right after high-school graduation. While the difference for those who had participated in athletics is about 4%, this difference is 1.7% for those who had participated in clubs. Finally, working experience seems to be very important for high-school graduates who do not go to college. Wages are 15% higher among those who have at least one year of experience during high school. While working seems to hurt test scores early, working experience seems to pays off for those entering the labor market.

3.4 OLS Regressions

Descriptive statistics suggest significant returns to participation in extracurricular activities as shown in average weekly wages. In order to obtain a more precise measure of the magnitude of the effects of experiences in extracurricular activities on adult wages, the results of regressions of log-wages on experience and selected observed variables is presented in Table 10. Regressing log-wages on test scores and experience (one or two periods) in each activity show that returns to two periods of participation in athletics or working are correlated with about 16% higher wages eight years after finishing high school. This number is about 5% for experience in organized school clubs. After controlling for the post-secondary educational level (dummy variable for bachelor's degree), a measure of self-esteem, family background, and school characteristics, the magnitude of the returns are lower for participation in athletics (about

10%) and zero for participation in school clubs.[3] Work experience, however, still demonstrates strong returns to wages at the age of 26. Those who worked for one period receive on average a wage premium of 8%, while those who worked for two periods obtain an average wage premium of 20%.

The returns to participation in athletics and clubs seem to be different between those who received college degrees and those holding only a high-school diploma. Experience in extracurricular activities is more likely to increase future income for those who received college degrees, while it seems to be less important for those holding only a high-school diploma. At the same time, work experience is more important for those who do not pursue a college degree. Table 11 reports the returns to participation in extracurricular activities when regressions are estimated separately for college graduates and those with a high-school diploma only, after controlling for self-esteem, family background, and school characteristics. Results show that two periods of experience in athletics increases the wages of college graduates by 17% and wages of students with only a high-school diploma by 7%. The same level of experience in the workplace increases wages of college graduates by 19% and wages of those who completed only high school by 22%. Returns to participation in clubs are not statistically different from zero.

Because participation in extracurricular activities and future educational level (to attend college) are choices and students potentially have different preferences, a simple regression of log-wages on participation in athletics and clubs may lead to

[3] Table 10 presents the results of regressing the log-weekly wages in 2000 on the quartiles of test scores, and on dummies for accumulated experience in athletics, clubs, and work. The additional variables included in specifications (2) through (5) are a dummy if the student has a college degree (variable f4hhdg in NELS88), a measure of self-esteem (variable f2locus1), family background (a measure of socio-economic status, number of siblings and a dummy that is equal to one if the mother has a college degree or higher, variables bys32, f2ses and bys34b), and school characteristics (dummies if the school is private, urban, or rural, variables g12ctrl1 and g12urbn3).

a biased estimate due to self-selection. Suppose, for example, that students who are more productive in the labor market are also the ones who derive higher utility from participation in sports, due to some unobserved characteristic. Then, a simple OLS regression of wages on participation in sports would overestimate the effects of participation in sports on wages, leading to biased results. The model and the estimation method presented in Chapter 2 attempt to eliminate the effect of this bias caused by the self-selection problem, and the results are presented in Chapter 4. The NELS88 survey, however, imposes some limitations on the estimation method and the assumptions made to obtain identification are presented in the section below.

3.5 Data Restrictions and Identification Issues

After including the 27 parameters that allow for unobserved heterogeneity and the 3 type-proportion parameters, the structural model has 94 parameters. The total number of estimated parameters is 88. One particular data restriction - not observing a long series of wages during adulthood - implies that five parameters have to be fixed: β, α_1^{HS}, α_2^{HS}, α_1^C, and α_2^C. Parameter β is the discount rate and its value is set equal to 0.97. Parameters α_1^{HS}, α_2^{HS}, α_1^C, and α_2^C are taken from Heckman, Lochner and Todd (2003). The values are: $\alpha_1^{HS} = \alpha_1^C = 0.1301$ and $\alpha_2^{HS} = \alpha_2^C = -0.0023$.[4]

The identification of the model is obtained as follows. The identification of the remaining parameters on the wage profiles $(\alpha^H, \alpha^C, {\sigma_H}^2, \sigma_C^2)$ are obtained by observing variation in the wages received by students in 2000, when they are twenty-six years

[4] It is not possible to estimate these parameters because NELS88 does not contain information about wages during adulthood, except in 2000. The survey ends in 2000 when most students are 26 years old. Even if the survey observed wages in all years before 2000, estimation of these parameters would lead to a very steep wage profile, as most salaries increase fast in the initial years in the labor force. Such feature is not obtained in long-term wage profiles, as observed, for example, in the salaries of individuals in the Decennial Censuses. This is the reason why these parameters are taken from Heckman, Lochner and Todd (2003).

old and by assuming that the shocks on log-wages during adulthood are normally distributed, identifying 18 parameters. Identification of γ^C is obtained by imposing normally distributed errors and observing the educational choice, d_c, for each individual. This identification, however, is not obtained without normalizing the standard deviation of the shock, σ_c, as in a probit model. Therefore, σ_c is normalized to 2,000.

The parameters of the grading process in the first and second periods are obtained from observing variation in the scores in both periods and from imposing an ordered probit specification, identifying 15 parameters. The variance of the test scores in both periods was normalized to one. Imposing log-normality on the distribution of wages during high school and observing the wages for those who worked leads to the identification of the high-school wage function's intercepts and variances, for a total of 4 parameters.

The utility function parameters describing the preferences of the base-type (ψ^1, ψ^2, γ^1's, and γ^2's) in periods one and two are identified by imposing a multivariate normal distribution on the preference shocks and from observing the decisions made by students and their consequent leisure time choices in the first two periods (for a total of 10 parameters). This information also helps to identify the ten parameters in Σ. Because the matrix of shocks is constant between periods one and two, and because sixteen different choices are observed in each period, it is possible to identify all elements in the variance-covariance matrix, since we have sixteen alternatives and only 10 elements in the matrix. The only requirement is that we observe groups of individuals choosing all sixteen alternatives in both periods in the data.

There are 30 parameters left. They are the 27 unobserved heterogeneity parameters and the 3 type-proportion parameters. The unobserved heterogeneity parameters are identified, as are the parameters for the base-type (type 1), by observing variations in students' choices between periods one and two. The distribution of types is identified by imposing the restriction that $\pi_1 < \pi_2 < \pi_3$. This restriction, however, does not have to be imposed during the estimation process, as types can always be relabeled after estimation.

Chapter 4

Estimation Results

This Chapter presents the results of the structural estimation[1]. Tables 12 to 18 report the parameter estimates and standard deviations, and Tables 19 to 27 present the model fit and chi-square statistics for selected moments of the data.

4.1 Parameter Estimates

Table 12 displays the parameters of the wage functions during adulthood for individuals with a bachelor's degree and with only a high-school diploma. High-school scores contribute to wages if the individual earns a bachelor's degree, resulting in an increase of about 1.7% on wages. The same return on scores is not obtained by individuals with a only high-school diploma. Participating in athletics is estimated to increase wages by 3.6% for students with two years of experience and 6.7% for those who participated during four years of high-school, if they finish college. Wage differentials for students holding a high-school diploma are of smaller magnitude, 1.2% and 4.2%, for those who participated for two and four years, respectively. This result implies that experience in athletic extracurricular activities is more important to individuals

[1] The maximum log-likelihood achieved was -15,540.

who finish college than to those who enter the job market right after graduating from high school. Additionally, persistence in athletic activities pays off. Students get a higher increase in wages by participating in sports during four years of high school, instead of only two years, in both wage profiles.

Among college graduates, estimates show that the wage premium for participation in school clubs is 2.7% for the first two years of experience, and 5.7% for a total of four years of experience. Returns to participation in clubs are lower among students holding only a high-school diploma. Two years of experience lead to an extra 1.3% of wages, while four years increase wages by 3.9%. Overall, experience in school clubs is less valuable than experience in athletics. As in athletic activities, skills learned in school clubs are more valuable to college graduates than to students who get only a high-school diploma, and persistence leads to higher returns. Working experience, however, is more valuable if the individual enters the labor market right after high-school graduation. For this group of individuals, two years of experience increases wages by 4.1%, while four years of experience leads to a wage premium of 7.6%. For individuals with a bachelor's degree these numbers are 0.08% and 2.9%, respectively.

Parameters representing unobserved heterogeneity in adult wages can be interpreted as an absolute advantage. Individuals who are more likely to receive higher wages if they obtain a bachelor's degree are also more likely to receive higher wages if they work in jobs requiring only a high-school diploma, when compared to the base-type individual. That type of individual is also more likely to have strong preferences for participation in athletics, as will be discussed later in this section. On the other hand, conditional on test scores and experience levels, students who have strong preferences for participating in clubs are more likely to receive lower wages than individual type 1 in both educational levels.

Tables 13 and 14 show the parameter estimates for wages during high school and test scores. Wages increase during the high-school years, possibly due to an increase in the productivity of labor between periods (which is not in the model). Test scores increase with study time and participation in clubs and decrease with participation in sports and work during the first two years of high school. In the latter two years, studying and participating in clubs still help scores, participating athletic activities still has a negative effect, but working no longer affects scores. Additionally, the quartile of the test score received in the first two years positively affects the quartile of the score obtained in the last two years of high school. These results can be interpreted as evidence that skills acquired in clubs help the development of cognitive skills measured in test scores, while skills acquired in athletic activities or in part-time jobs are detrimental to the acquisition of cognitive skills and result in lower test scores.

The utility function parameters are presented in Tables 15 and 16. Parameters ($\gamma_1, \gamma_2, \gamma_3, \gamma_4$, and γ^C) indicate that all activities are costly (parameters are negative) for all individuals. In the first period, working is the costliest activity, followed by participation in clubs, participation in sports, and studying. In the second period, participating in sports is the costliest activity, followed by working, participation in clubs, and studying. Relative to the type-1 student, students with lower utility costs from participation in athletics (type 2) have higher utility costs from obtaining a bachelor's degree and derive lower utility levels from participation in clubs and working, but they have stronger preferences for studying. Individuals with strong preferences for clubs (type 3) derive lower utility levels from participating in sports, working, and obtaining a bachelor's degree, but, their disutility from studying is the lowest among all types. Finally, when compared to the base type students, individuals with lower disutility from working (type 4) have higher disutility from participating in sports or clubs or from studying. The estimated distribution of types is presented in Table 17, and the variance-covariance matrix parameters are presented in Table 18.

4.2 Model Fit

In general, the model predictions do well when compared to selected sample moments. Table 19 presents the simulated participation rates in activities by schooling year. The model captures the patterns of study choices, participation in extracurricular activities and work, as well as most of the cross-participation rates (Table 20). Based on the chi-square measure of goodness of fit, the results underestimate the number of students simultaneously participating in clubs and working in sophomore year and the proportion of students simultaneously participating in athletics, clubs and working in the senior year. The model is able to fit most of the moments related to the distribution of the number of activities in which students participated by schooling year (Table 21) and the transition rates between activities from sophomore to senior year (Tables 22A, 22B and 22C). The model, however, fails to match the transition rates from clubs to work, work to clubs, and work to athletics. The model also underestimates the persistence rate in participation in clubs.

Simulated results fit the distribution of test scores conditional on participation in extracurricular activities reasonably well (Tables 23A, 23B and 23C). In particular, the model captures the fact that students who study or participate in clubs obtain higher scores than those who participate in sports or work. Statistically, however, the model is not able to capture fully the negative skewness in the distribution of scores of those who studied. The simulated average wages during high school and their standard deviations are presented in Table 24. Average wages match the data, but the model underestimates the standard deviations. The specification does a good job matching the bachelor's degree rate conditional on choices (Table 25). The model, however, underestimates the graduation rate for those who worked during high school.

Tables 26A, 26B and 26C present the distribution of cumulative experience in athletics, clubs, and work for the sample data and for the simulated data by educational level (high-school diploma vs. bachelor's degree) together with its chi-square statistics. At first glance, the model looks to fit well the distribution of experience. But, an analysis of the chi-square statistics reveals that for those with only a high-school diploma the model underestimates the cumulative experience in clubs and in the workplace. Additionally, among individuals holding a bachelor's degree, the model overestimates the level of experience students will have in athletic activities.

Finally, Table 27 shows the average wages in the sample and in the simulation, conditional on participation in activities. The model captures wage levels and the differences in wages reasonably well. For those with a bachelor's degree, the model seems to capture the wage premium for athletes and for those with experience in clubs. However, for this same group of individuals, the simulation delivers a positive wage differential for those who have working experience. This feature is not present in the sample data. For individuals holding only a high-school diploma, the model captures the wage differentials and fits the sample data reasonably well. The model does a poor job, however, generating the larger differences observed in the sample data. In particular, the simulation does not generate the one-hundred dollar wage premium experienced by students who worked during high school and earned only a high-school diploma.

Chapter 5

Counterfactual Experiments

Having estimated the structural model parameters governing preferences and outcomes, it is possible to evaluate the impact of changes in public policies during high-school years on choices during high school, the choice of educational level (bachelor's degree), and wages during adulthood. In each experiment, the model is simulated under the new policy and a new set of decisions and outcomes is obtained. Because structural parameters are primitives in the model, they are invariant to policy changes.[1]

This chapter presents the results of four policy experiments[2]. The first three experiments investigate the effect of the availability of sports and clubs during high school. These policy experiments aim to measure the long-term consequences of proposed cuts in high school budgets, which usually have extracurricular activities as prime candidates for cutbacks. This situation has become more common recently, with school districts facing substantial budget cuts across the country. Budget cuts

[1] It is important to mention that the counterfactual experiments are conducted using parameters estimated using a selected sample of white male individuals. The results obtained in the experiments may not be the same for females or individuals of other ethnic groups, as they may have different preferences or face different choices or distribution of shocks during adolescence.

[2] Each simulated sample has 5,000 students.

have affected small and large districts. The Los Angeles Unified School District, for example, expects to face a twenty-percent loss in its high-school athletic funding in the coming school year.[3] The Inter-Lakes School (at the Inter-Lakes School District, in New Hampshire) is planning to cut $1,150,940 in expenditure, which are "expected to be felt in virtually every school and at every grade level."[4] As a result, the high-school newspaper will be eliminated. These experiments aim to replicate this type of situation.

The fourth experiment intends to estimate the consequences of policies that provide income to students who perform better in exams. Public policies that give prizes to students who obtain better test scores, grades, or other measures of achievement have recently been implemented in schools in Chicago, Washington, and New York. The immediate results in terms of grades and scores, however, are mixed.[5] This experiment aims to measure the long-run consequences for students' income, given the trade-off between study time and time spend in extracurricular activities during high school.

5.1 Counterfactual 1: Eliminating Clubs

In this first counterfactual exercise, students no longer have the option to participate in clubs. The change in the choice set aims to replicate a situation in which schools face a budget cut in which they have to reduce their investment in extracurricular activities: in particular they have to eliminate clubs. In this version of the model, students face fewer options, as participation in clubs is no longer available.

[3] As reported by Sondheimer (2010) for *Los Angeles Times*.

[4] As reported by Tunning (2010) for *Citizen.com*.

[5] See Fryer (2010) for a detailed description of the programs and their results.

Because participation in clubs affects future outcomes, students anticipate the consequences of not having the opportunity to accumulate experience and make different choices while in high school. The results of interest are participation in activities during high-school, test scores, the rate at which students obtain bachelor's degrees, and average wages during adulthood.

Table 28 shows students' participation rates in activities in the original model and under the case in which clubs are no longer available. Students change their choices towards the accumulation of working experience. In both periods there is a decrease in the number of individuals who study or participate in sports, while the number of students working increases. The consequence of eliminating clubs on the distribution of test scores is presented in Tables 29A and 29B. Even conditioning on having studied, test scores decrease, and the distribution becomes more positively skewed after clubs are no longer available. At the same time, the number of students choosing to obtain a bachelor's degree decreases from 43.9% to 28.6% and are as low as 19.7% for those who worked at least one period during high school (Table 30). Finally, wages during adulthood decrease in almost every conditional category of participation during high school, with bigger drops among those who obtain a bachelor's degree (Table 31).

The intuition behind this result is as follows. Because the payoff for obtaining experience in participation in clubs had been higher for those who actually ended up obtaining a bachelor's degree, the option of getting a degree becomes less valuable than entering the labor market right after finishing high school. Anticipating that, students are more likely to work during high school and decrease their participation in athletics. Test scores decrease as a consequence of these choices. Hence, policy makers and school district managers should be aware that public policies that eliminate expenses on school clubs may lead to lower test scores, fewer of students untimately obtaining a bachelor's degree, and lower wages during adulthood. An analysis of the lifetime value lost by the average student reveals that individuals should be willing to pay up to 790 dollars per month during the four years of high school to have the opportunity to participate in school clubs.

5.2 Counterfactual 2: Eliminating Sports

In the second counterfactual exercise, students no longer have the opportunity to participate in athletics. The same set of outcomes (choices, scores, college rates and adulthood wages) is analyzed below.

The results of eliminating the opportunity for students to participate in sports are in the same direction of those in Counterfactual 1. More students choose to work, and fewer students decide to study or participate in clubs (Table 32). Test scores decrease (Tables 33A and 33B), and the bachelor's degree attainment decrease by more than 50% (Table 34). Additionally, at the age of 26, average adult wages decrease more than 5% for those with a bachelor's degree and more than 2% for those with only a high-school diploma (Table 35). This decrease in wages, together with a lower rate of earning bachelor's degrees, decreases the expected discounted value of students' future income. In fact, according to the model predictions, students' would be willing to pay up to $1,233 per month for the opportunity to participate in sports.

Results of Counterfactual 1 and 2 reveal that students change their choices as a result of the lack of opportunities to acquire experience through participation in school clubs and sports. A more dramatic experiment would eliminate extracurricular activities all together. This possibility is investigated in Counterfactual 3.

5.3 Counterfactual 3: Eliminating Sports and Clubs

The third counterfactual aims to analyze the consequences of eliminating both types of extracurricular activities at the same time. The results are presented in Tables 36, 37A, 37B, 38, and 39. The consequences are similar to the ones in Counterfactuals 1 and 2. More students choose to work during high school, test scores decrease, and bachelor's degree attainment decreases to only 15%. Wages for those holding a bachelor's degree are 8% lower, while this decrease is about 2% for individuals holding only a high-school diploma. The final result is a combination of the consequences of Counterfactuals 1 and 2, and the decrease in future income is larger. According to the model, the lifetime value lost by the average student is such that they would pay up to $1,500 per month during the four years of high school to get extracurricular activities back in schools.

5.4 Counterfactual 4: Paying for Better Scores

The fourth counterfactual estimates the effects of a policy that gives $700 dollars per period to students who obtain a score above the median of the test score distribution. As before, the results of interest are participation in activities during high school, test scores, the rate at which students obtain a bachelor's degree, and the average wages during adulthood.

Table 40 reports students' participation rates in activities in the original model and after the introduction of the policy. With the policy, the number of students spending more than 5 hours a week studying at home and participating in clubs increases, while the number of students who work decreases. Participation rates in athletics remain basically unchanged.

The effects of the policy on the distribution of test scores are presented in Tables 41A and 41B. As a consequence of the increases in study and participation in clubs and the decline in the number of students working, test scores increase, and about 55% of the students obtain scores that would be above the median of the original test score distribution. The number of students choosing to obtain a bachelor's degree increases from 43.9% to 50.3% (Table 42). Finally, wages during adulthood increase slightly among those who decide to get a bachelor's degree (Table 43).

The intuition behind this result is as follows. Students anticipate that if on top of studying more they increase their participation in clubs they will be more likely to receive the $700 payment. Because the payoff for obtaining better test scores and experience in participation in clubs is higher for those who actually end up obtaining a bachelor's degree, it increases their lifetime value of obtaining a college degree, increasing the rate of students who obtain a bachelor's degree. The wages by educational level do not increase much conditional on experience, but the average wage for all students increase, because more students now have a bachelor's degree. The average weekly wage increases by about 2%.

Chapter 6

Concluding Remarks

This dissertation develops and estimates a structural dynamic decision model in which individuals decide about participation in extracurricular activities, study time, and employment during high school. The specification allows for returns to participation in extracurricular activities on salaries and unobserved heterogeneity in students' preferences over study time, extracurricular activities, employment during high school and the acquisition of a bachelor's degree. The empirical results suggest that the model is able to replicate most of the patterns in the data well. Results show that participation in extracurricular activities increase wages during adulthood as much as 6.7%. Additionally, results show that while participation in school clubs increases test scores, working or participating in sports in the last two years of high school may harm test scores.

Having estimated the parameters governing preferences and outcomes, four different public policies were analyzed. The first three experiments investigate the effect of eliminating clubs, athletic and both activities simultaneously. These policy experiment aim to measure the long-term consequences of proposed cuts in high-school budgets, which often mark extracurricular activities as prime candidates for cutbacks. The results are the following: public policies that eliminate expenses in extracurricular activities lead to lower test scores, fewer students obtaining a bachelor's degree, and lower wages during adulthood. The fourth experiment intends to estimate the consequences of policies that transfer income to students who perform better on exams. Results show that a policy that pays $700 dollars to every student who obtains a score above the median of the test distribution increases test scores and the number of students obtaining a bachelor's degree.

Table 1: Participation Rates

	Sophomores	Seniors
Study	22.50	37.28
Athletics	69.97	61.28
Clubs	36.42	46.13
Work	19.20	40.69

Table 2: Cross-Participation Rates

	Sophomores	Seniors
Athletics and Clubs	26.61	31.25
Athletics and Work	12.10	21.28
Clubs and Work	6.24	16.10
Athletics, Clubs and Work	3.94	9.65
None	15.41	10.88

Table 3: Number of Activities

	Sophomores	Seniors
One Activity	47.52	39.79
Two Activities	33.12	39.68
Three Activities	3.95	9.65
None	15.41	10.88

Table 4: Transition Rates

	Athletics	Clubs	Work
Athletics	78.13	48.78	37.65
Clubs	65.45	78.18	37.19
Work	57.50	40.83	65.83

Table 5: Average Weekly Wages during High School

	Sophomores	Seniors
Wage	96.29	116.65
Std. Dev	51.15	60.85
N	360	763

Table 6: IRT Theta Scores during High School

	Sophomores	Seniors
Mean	54.05	57.77
Std. Dev.	8.09	8.62
Min.	28.17	31.79
Max	71.73	78.45
Median	54.71	58.55

Table 7: Test Scores Distribution by Activity Participation

Quartile	Sophomores				Seniors			
	Study	Athletics	Clubs	Work	Study	Athletics	Clubs	Work
1^{st}	12.09	23.63	15.67	32.78	18.03	22.63	18.15	30.28
2^{nd}	22.75	25.61	21.38	25.00	21.60	24.98	20.92	26.73
3^{rd}	28.49	26.37	27.82	22.22	28.75	27.85	27.17	23.20
4^{th}	36.49	24.39	35.14	20.00	31.62	24.54	33.76	19.79

Table 8: Bachelor's Degree Attainment by Activity Participation

	%
Athletics	49.79
Clubs	54.73
Work	33.52
All Sample	44.74

Table 9: Average Weekly Wages by Experience and Educational Level

	High School Diploma	Bachelor's Degree
Athletics		
Participated	688.12	886.25
Not	662.00	703.66
Clubs		
Participated	686.66	882.92
Not	675.02	813.02
Work		
Worked	720.83	841.47
Not	626.62	869.01
All		
	680.18	859.26

Table 10: OLS Estimates Ln(Wage) Equation

Covariates	(1)	(2)	(3)	(4)	(5)
Intercept	6.2589	6.2745	6.2910	6.2748	6.2622
	(0.0507)	(0.0504)	(0.0505)	(0.0576)	(0.0596)
Test Scores	0.0184	-0.0167	-0.0220	-0.0276	-0.0289
	(0.0143)	(0.0146)	(0.0147)	(0.0152)	(0.0152)
Exp. Atheltics = 1	0.0604	0.0465	0.0467	0.0399	0.0380
	(0.0428)	(0 0423)	(0.0423)	(0.0423)	(0.0424)
Exp. Athletics = 2	0.1669	0.1167	0.1155	0.1078	0.1015
	(0.0348)	(0.0352)	(0.0353)	(0.0351)	(0.0353)
Exp. Clubs = 1	0.0268	0.0066	0 0031	0.0002	-0.0014
	(0.0352)	(0.0349)	(0.0347)	(0.0345)	(0.0343)
Exp. Clubs = 2	0.0499	0.0212	0.0145	0.0047	0.0011
	(0.0364)	(0.0363)	(0.0362)	(0.0364)	(0 0367)
Exp. Work = 1	0.0430	0.0739	0.0738	0.0792	0.0812
	(0.0319)	(0.0320)	(0.0320)	(0.0320)	(0.0319)
Exp. Work = 2	0.1644	0.1963	0.1935	0.2008	0.2024
	(0.0393)	(0.0390)	(0 0386)	(0.0387)	(0.0388)
Bachelor's Degree		0.2255	0.2163	0.1859	0.1790
		(0.0313)	(0.0316)	(0.0330)	(0.0331)
Self-Esteem			X	X	X
Family Background				X	X
School Characteristics					X
N	1,673	1,673	1,673	1,673	1,673
R^2	0.0229	0.0501	0.0540	0.0592	0.0624
F-Stat. (K,N-K-1)	5.89	11.51	11.03	9.21	7.67

Standard errors robust to heteroskedasticity are in parentheses.

Table 11: OLS Estimates Ln(Wage) Equation by Educational Level

Covariates	(1)	(2)
Intercept	6.3255	6.2871
	(0.0753)	(0.1050)
Test Scores	-0.0424	-0.0087
	(0.0198)	(0.0242)
Exp. Atheltics = 1	-0.0417	0.2170
	(0.0515)	(0.0735)
Exp. Athletics = 2	0.0771	0.1782
	(0.0424)	(0.0636)
Exp. Clubs = 1	-0.0075	0.0054
	(0.0428)	(0.0554)
Exp. Clubs = 2	-0.0141	0.0090
	(0.0533)	(0.0519)
Exp. Work = 1	0.1371	-0.0124
	(0.0410)	(0.0501)
Exp. Work = 2	0.2225	0.1901
	(0.0491)	(0.0647)
Self-Esteem	X	X
Family Background	X	X
School Characteristics	X	X
N	906	767
R^2	0.0428	0.0548
F-Stat. (K,N-K-1)	2.70	3.60

Standard errors robust to heteroskedasticity are in parentheses.

Table 12: Wage Profiles

	W^{HS}	W^{C}
Intercept (α_0)	5.3900	5.9514
	(0.0633)	(0.0831)
Diff. Type 2 - Type 1 (χ_{82}, χ_{92})	0.0026	0.0071
	(0.0110)	(0.057)
Diff. Type 3 - Type 1 (χ_{83}, χ_{93})	-0.0079	-0.0052
	(0.0132)	(0.086)
Diff. Type 4 - Type 1 (χ_{84}, χ_{94})	0.0022	0.0002
	(0.0184)	(0.015)
Scores (α_3)	0.0044	0.0170
	(0.0038)	(0.0025)
Exp. Athletics = 1 (α_4)	0.0124	0.0362
	(0.0341)	(0.0211)
Exp. Athletics = 2 (α_5)	0.0427	0.0675
	(0.0332)	(0.0204)
Exp. Clubs = 1 (α_6)	0.0137	0.0276
	(0.0234)	(0.0145)
Exp. Clubs = 2 (α_7)	0.0399	0.0579
	(0.0460)	(0.0284)
Exp. Work = 1 (α_8)	0.0410	0.0077
	(0.0247)	(0.0152)
Exp. Work = 2 (α_9)	0.0765	0.0288
	(0.0277)	(0.0302)
Std. Dev. (σ^2_{HS}, σ^2_{C})	0.5721	0.5852
	(0.1891)	(0.1841)

Table 13: High-School Wages

	Sophomores		Seniors
Intercept (w^t)	4.4209		4.6412
	(0.3166)		(0.3161)
Diff. Type 2 - Type 1 (χ_{62})		-0.0101	
		(0.1628)	
Diff. Type 3 - Type 1 (χ_{63})		-0.0098	
		(0.1733)	
Diff. Type 4 - Type 1 (χ_{64})		0.0069	
		(0.1736)	
HS Wage Std. Dev ($\sigma^2_{\xi^t}$)	0.4340		0.3905
	(0.2445)		(0.1254)

Table 14: Test Scores

	Sophomores	Seniors
Score in $t-1$ (ρ_0)	-	1.3704
		(0.1502)
Study (ρ_1)	0.4343	0.4943
	(0.0276)	(0.0258)
Part. in Athletics (ρ_2)	-0.0501	-0.4661
	(0.0392)	(0.0400)
Part. in Clubs (ρ_3)	0.4278	0.2300
	(0.0325)	(0.0330)
Working (ρ_4)	-0.1426	-0.1676
	(0.03879)	(0.0337)
Intercept 1 (μ_1)	-0.5144	1.9918
	(0.0585)	(0.5903)
Intercept 2 (μ_2)	0.1783	3.4112
	(0.0585)	(0.5912)
Intercept 3 (μ_3)	0.9088	4.7817
	(0.0586)	(0.5927)
Diff. Type 2 - Type 1 (χ_{72})	0.02757	
	(0.0618)	
Diff. Type 3 - Type 1 (χ_{73})	0.04218	
	(0.0644)	
Diff. Type 4 - Type 1 (χ_{74})	-0.03171	
	(0.0726)	

Table 15: First and Second-Period Utility Function

	Sophomores	Seniors
Leisure (ψ)	46.2670	50.2556
	(24.9559)	(15.3608)
Study Type 1 (γ_1)	-431.8265	-223.3407
	(133.5025)	(132.5548)
Diff. Type 2 - Type 1 (χ_{12})	16.0288	
	(137.7338)	
Diff. Type 3 - Type 1 (χ_{13})	28.1681	
	(145.9075)	
Diff. Type 4 - Type 1 (χ_{14})	-27.8922	
	(160.7079)	
Part. in Athletics Type 1 (γ_2)	-667.9982	-857.2047
	(434.4340)	(440.30)
Diff. Type 2 - Type 1 (χ_{22})	358.1105	
	(397.9374)	
Diff. Type 3 - Type 1 (χ_{23})	-154.6604	
	(422.9212)	
Diff. Type 4 - Type 1 (χ_{24})	55.1001	
	(453.7270)	
Part. in Clubs Type 1 (γ_3)	-1,091.0698	-809.5591
	(477.6131)	(480.3632)
Diff. Type 2 - Type 1 (χ_{32})	-158.0102	
	(469.9277)	
Diff. Type 3 - Type 1 (χ_{33})	278.0013	
	(415.0422)	
Diff. Type 4 - Type 1 (χ_{34})	19.0246	
	(516.6749)	
Working Type 1 (γ_4)	-1,102.1523	-853.2133
	(455.5958)	(459.8158)
Diff. Type 2 - Type 1 (χ_{42})	-161.1918	
	(461.9802)	
Diff. Type 3 - Type 1 (χ_{43})	-123.8164	
	(457.3783)	
Diff. Type 4 - Type 1 (χ_{44})	390.0222	
	(413.6932)	

Table 16: Third-Period Utility Function

	Parameter
Bachelor Degree Type 1 (γ^C)	-16,096.0074
	(2,634.4742)
Diff. Type 2 - Type 1 (χ_{52})	-138.1181
	(155.1608)
Diff. Type 3 - Type 1 (χ_{53})	-97.0964
	(196.1260)
Diff. Type 4 - Type 1 (χ_{54})	2.0128
	(20.2775)

Table 17: Distribution of Types

	%
Type 1 (π_1)	10.0507
Type 2 (π_2)	50.8952
Type 3 (π_3)	25.7620
Type 4 (π_4)	13.2921

Table 18: Vaiance-Covariance Matrix

	Value	Std. Dev.
Std. Dev. Study (σ_s)	431.9505	153.3420
Std. Dev. Athletics (σ_a)	339.8556	142.1868
Std. Dev. Clubs (σ_b)	420.5623	146.2949
Std. Dev. Work (σ_w)	399.7791	207.1738
Covariance Study and Athletics (σ^2_{sa})	-372.270	128.7944
Covariance Study and Clubs (σ^2_{sb})	693.681	240.8815
Covariance Study and Work (σ^2_{sw})	146.671	104.0153
Covariance Athletics and Clubs (σ^2_{ab})	-875.744	299.5988
Covariance Athletics and Work (σ^2_{aw})	917.794	628.8120
Covariance Clubs and Work (σ^2_{bw})	790.945	542.4398

Table 19: Participation Rates

	Sophomores			Seniors		
	Sample	Simulation	χ^2	Sample	Simulation	χ^2
Study	22.50	23.18	0.49	37.28	37.82	0.23
Athletics	69.97	68.58	1.72	61.28	60.60	0.36
Clubs	36.42	35.14	1.32	46.13	45.16	0.71
Work	19.20	20.48	1.98	40.69	41.62	0.67

$\chi^2_{1,0.01} = 6.64$

Table 20: Cross-Participation Rates

	Sophomores			Seniors		
	Sample	Simulation	χ^2	Sample	Simulation	χ^2
Athletics and Clubs	26.61	27.66	1.05	31.25	29.54	2.55
Athletics and Work	12.10	10.98	2.21	21.28	20.18	1.35
Clubs and Work	6.24	4.56	9.04	16.10	13.98	6.23
Athletics, Clubs and Work	3.94	3.28	2.15	9.65	6.98	15.33
None	15.41	15.72	0.13	10.88	9.34	4.58

$\chi^2_{1,0.01} = 6.64$

Table 21: Number of Activities

	Sophomores			Seniors		
	Sample	Simulation	χ^2	Sample	Simulation	χ^2
One Activity	47.52	47.64	0.01	39.79	40.92	0.99
Two Activities	33.12	33.36	0.04	39.68	42.76	7.43
Three Activities	3.95	3.28	2.21	9.65	6.98	15.33
None	15.41	15.72	0.13	10.88	9.34	4.58

$\chi^2_{1,0.01} = 6.64$

Table 22: Transition Rates

(A) Sample

	Athletics	Clubs	Work
Athletics	78.13	48.78	37.65
Clubs	65.45	78.18	37.19
Work	57.50	40.83	65.83

(B) Simulation

	Athletics	Clubs	Work
Athletics	76.06	47.59	35.96
Clubs	65.85	75.13	32.56
Work	47.75	37.40	66.60

(C) Chi-Square

	Athletics	Clubs	Work
Athletics	4.70	1.06	2.28
Clubs	0.13	10.22	17.21
Work	72.94	9.13	0.49

$\chi^2_{1,0.01} = 6.64$

Table 23: Test Scores Distribution by Activity Participation

(A) Sample

Quartile	Sophomores				Seniors			
	Study	Athletics	Clubs	Work	Study	Athletics	Clubs	Work
1^{st}	12.09	23.63	15.67	32.78	18.03	22.63	18.15	30.28
2^{nd}	22.75	25.61	21.38	25.00	21.60	24.98	20.92	26.73
3^{rd}	28.49	26.37	27.82	22.22	28.75	27.85	27.17	23.20
4^{th}	36.49	24.39	35.14	20.00	31.62	24.54	33.76	19.79

(B) Simulation

Quartile	Sophomores				Seniors			
	Study	Athletics	Clubs	Work	Study	Athletics	Clubs	Work
1^{st}	15.88	24.38	17.13	30.76	18.35	25.94	17.36	31.81
2^{nd}	20.28	25.31	22.42	26.27	26.81	23.99	22.14	27.63
3^{rd}	27.96	25.46	26.98	23.44	24.59	24.85	27.06	21.48
4^{th}	35.89	24.85	33.47	19.53	30.25	25.21	33.44	19.08

(C) Chi-Square

	Sophomore	Seniors
Study	27.67	36.07
Athletics	1.26	16.22
Clubs	5.46	2.04
Work	5.00	4.88

$\chi^2_{3,0.01} = 11.35$

Table 24: Average Weekly Wages during High School

		Sample	Simulation
Sophomores	Wage	96.29	96.17
	Std. Dev.	51.15	45.83
Seniors	Wage	116.65	115.80
	Std. Dev.	60.85	48.61

Table 25: Bachelor's Degree Attainment by Activity Participation

	Sample	Simulation	χ^2
Athletics	49.79	51.67	2.65
Clubs	54.73	56.25	1.74
Work	33.52	29.01	17.11
All Sample	44.74	43.98	0.43

$\chi^2_{1,0.01} = 6.64$

Table 26: Cumulative Experience by Activity and Educational Level

(A) Sample

	Only High-School Diploma			Bachelor Degree		
Experience	Athletics	Clubs	Work	Athletics	Clubs	Work
0	30.41	55.69	43.15	14.78	33.85	64.60
1	25.68	23.67	41.22	17.28	28.01	26.46
2	43.92	20.66	15.64	67.94	38.14	8.94

(B) Simulation

	Only High-School Diploma			Bachelor Degree		
Experience	Athletics	Clubs	Work	Athletics	Clubs	Work
0	33.56	57.91	38.59	9.50	31.06	68.03
1	24.35	25.53	43.45	25.51	30.01	23.83
2	42.09	16.57	17.96	64.98	38.93	8.14

(C) Chi-Square

	High-School Diploma	Bachelor Degree
Athletics	8.83	111.28
Clubs	19.58	7.29
Worl	17.75	9.65

$\chi^2_{2,0.01}$ =9.21

Table 27: Average Weekly Wages by Experience and Educational Level

	High School Diploma		Bachelor's Degree	
	Sample	Simulation	Sample	Simulation
Athletics				
Participated	688.12	682.97	886.25	868.10
Not	662.00	680.34	703.66	786.71
Clubs				
Participated	686.66	694.62	882.92	871.68
Not	675.02	672.98	813.02	835.27
Work				
Worked	720.83	704.72	841.47	877.01
Not	626.62	646.08	869.01	852.55
All				
	680.18	682.09	859.26	860.37

Table 28: Participation Rates

	Sophomores		Seniors	
	Full Model	Conterfactual 1	Full Model	Counterfactual 1
Study	23.18	21.28	37.82	36.14
Athletics	68.58	58.72	60.60	51.96
Clubs	35.14	-	45.16	-
Work	20.48	23.88	41.62	48.62

Table 29: Test Scores Distribution by Activity Participation

(A) Full Model

Quartile	Sophomores				Seniors			
	Study	Athletics	Clubs	Work	Study	Athletics	Clubs	Work
1^{st}	15.88	24.38	17.13	30.76	18.35	25.94	17.36	31.81
2^{nd}	20.28	25.31	22.42	26.27	26.81	23.99	22.14	27.63
3^{rd}	27.96	25.46	26.98	23.44	24.59	24.85	27.06	21.48
4^{th}	35.89	24.85	33.47	19.53	30.25	25.21	33.44	19.08

(B) Counterfactual 1

Quartile	Sophomores				Seniors			
	Study	Athletics	Clubs	Work	Study	Athletics	Clubs	Work
1^{st}	20.77	30.59	-	34.67	23.24	33.29	-	35.38
2^{nd}	23.50	26.36	-	25.80	28.83	25.25	-	27.31
3^{rd}	26.32	23.47	-	21.27	24.68	22.98	-	20.61
4^{th}	29.42	19.58	-	18.26	23.24	18.48	-	16.70

Table 30: Bachelor's Degree Attainment by Activity Participation

	Full Model	Counterfactual 1
Athletics	51.67	36.73
Clubs	56.25	-
Work	29.01	19.71
All Sample	43.98	28.64

Table 31: Average Weekly Wages by Experience and Educational Level

	High School Diploma		Bachelor's Degree	
	Full Model	Counterfactual 1	Full Model	Conterfactual 1
Athletics				
Participated	682.97	673.38	868.10	844.09
Not	680.34	673.50	786.71	777.86
Clubs				
Participated	694.62	-	871.68	-
Not	672.98	-	835.27	-
Work				
Worked	704.72	692.69	877.01	855.65
Not	646.08	638.97	852.55	820.34
All				
	682.09	673.43	860.37	834.20

Table 32: Participation Rates

	Sophomores		Seniors	
	Full Model	Conterfactual 2	Full Model	Counterfactual 2
Study	23.18	19.44	37.82	34.70
Athletics	68.58	-	60.60	-
Clubs	35.14	18.30	45.16	26.94
Work	20.48	25.74	41.62	52.90

Table 33: Test Scores Distribution by Activity Participation

(A) Full Model

Quartile	Sophomores				Seniors			
	Study	Athletics	Clubs	Work	Study	Athletics	Clubs	Work
1^{st}	15.88	24.38	17.13	30.76	18.35	25.94	17.36	31.81
2^{nd}	20.28	25.31	22.42	26.27	26.81	23.99	22.14	27.63
3^{rd}	27.96	25.46	26.98	23.44	24.59	24.85	27.06	21.48
4^{th}	35.89	24.85	33.47	19.53	30.25	25.21	33.44	19.08

(B) Counterfactual 2

Quartile	Sophomores				Seniors			
	Study	Athletics	Clubs	Work	Study	Athletics	Clubs	Work
1^{st}	17.08	-	16.72	31.86	17.87	-	15.44	27.52
2^{nd}	22.43	-	23.61	26.34	25.99	-	20.86	26.62
3^{rd}	27.78	-	27.10	22.84	25.07	-	26.06	23.10
4^{th}	32.72	-	32.57	18.96	31.07	-	37.64	22.76

Table 34: Bachelor's Degree Attainment by Activity Participation

	Full Model	Counterfactual 2
Athletics	51.67	-
Clubs	56.25	30.50
Work	29.01	14.61
All Sample	43.98	19.94

Table 35: Average Weekly Wages by Experience and Educational Level

	High School Diploma		Bachelor's Degree	
	Full Model	Counterfactual 2	Full Model	Conterfactual 2
Athletics				
Participated	682.97	-	868.10	-
Not	680.34	-	786.71	-
Clubs				
Participated	694.62	685.18	871.68	830.60
Not	672.98	661.93	835.27	788.37
Work				
Worked	704.72	683.09	877.01	822.88
Not	646.08	640.02	852.55	799.09
All				
	682.09	668.68	860.37	809.97

Table 36: Participation Rates

	Sophomores		Seniors	
	Full Model	Conterfactual 3	Full Model	Counterfactual 3
Study	23.18	19.04	37.82	34.16
Athletics	68.58	-	60.60	-
Clubs	35.14	-	45.16	-
Work	20.48	26.98	41.62	55.42

Table 37: Test Scores Distribution by Activity Participation

(A) Full Model

Quartile	Sophomores				Seniors			
	Study	Athletics	Clubs	Work	Study	Athletics	Clubs	Work
1^{st}	15.88	24.38	17.13	30.76	18.35	25.94	17.36	31.81
2^{nd}	20.28	25.31	22.42	26.27	26.81	23.99	22.14	27.63
3^{rd}	27.96	25.46	26.98	23.44	24.59	24.85	27.06	21.48
4^{th}	35.89	24.85	33.47	19.53	30.25	25.21	33.44	19.08

(B) Counterfactual 3

Quartile	Sophomores				Seniors			
	Study	Athletics	Clubs	Work	Study	Athletics	Clubs	Work
1^{st}	19.54	-	-	33.73	20.20	-	-	29.20
2^{nd}	23.63	-	-	26.54	27.58	-	-	27.79
3^{rd}	25.95	-	-	21.57	25.29	-	-	22.05
4^{th}	30.88	-	-	18.16	26.93	-	-	20.971484

Table 38: Bachelor's Degree Attainment by Activity Participation

	Full Model	Counterfactual 3
Athletics	51.67	-
Clubs	56.25	-
Work	29.01	11.22
All Sample	43.98	14.84

Table 39: Average Weekly Wages by Experience and Educational Level

	High School Diploma		Bachelor's Degree	
	Full Model	Counterfactual 3	Full Model	Conterfactual 3
Athletics				
Participated	682.97	-	868.10	-
Not	680.34	-	786.71	-
Clubs				
Participated	694.62	-	871.68	-
Not	672.98	-	835.27	-
Work				
Worked	704.72	678.83	877.01	798.01
Not	646.08	634.35	852.55	784.27
All				
	682.09	664.74	860.37	791.08

Table 40: Participation Rates

	Sophomores		Seniors	
	Full Model	Conterfactual 4	Full Model	Counterfactual 4
Study	23.18	39.62	37.82	43.16
Athletics	68.58	68.38	60.60	60.90
Clubs	35.14	53.38	45.16	53.46
Work	20.48	15.14	41.62	37.96

Table 41: Test Scores Distribution by Activity Participation

(A) Full Model

Quartile	Sophomores				Seniors			
	Study	Athletics	Clubs	Work	Study	Athletics	Clubs	Work
1^{st}	15.88	24.38	17.13	30.76	18.35	25.94	17.36	31.81
2^{nd}	20.28	25.31	22.42	26.27	26.81	23.99	22.14	27.63
3^{rd}	27.96	25.46	26.98	23.44	24.59	24.85	27.06	21.48
4^{th}	35.89	24.85	33.47	19.53	30.25	25.21	33.44	19.08

(B) Counterfactual 4

Quartile	Sophomores				Seniors			
	Study	Athletics	Clubs	Work	Study	Athletics	Clubs	Work
1^{st}	14.74	20.42	15.85	26.68	14.50	20.72	13.95	27.92
2^{nd}	18.98	22.70	21.13	26.16	25.35	21.67	20.50	24.82
3^{rd}	27.51	27.00	27.58	23.91	27.29	27.42	29.22	23.92
4^{th}	38.77	29.89	35.44	23.25	32.85	30.18	36.33	23.34

Table 42: Bachelor's Degree Attainment by Activity Participation

	Full Model	Counterfactual 4
Athletics	51.67	57.97
Clubs	56.25	60.04
Work	29.01	33.31
All Sample	43.98	50.28

Table 43: Average Weekly Wages by Experience and Educational Level

	High School Diploma		Bachelor's Degree	
	Full Model	Counterfactual 4	Full Model	Conterfactual 4
Athletics				
Participated	682.97	681.45	868.10	873.14
Not	680.34	680.02	786.71	816.28
Clubs				
Participated	694.62	689.22	871.68	878.80
Not	672.98	671.50	835.27	826.80
Work				
Worked	704.72	699.07	877.01	907.62
Not	646.08	656.51	852.55	852.48
All				
	682.09	680.98	860.37	868.14

Bibliography

[1] Barron J. M., Ewing, B. T., and Waddell, G. R. The Effects of High School Athletic Participation on Education and Labor Market Outcomes, *The Review of Economics and Statistics*, Vol. 82, No. 3. pp. 409-421, 2000.

[2] Ben-Avika, M., and D. Bolduc. Multinomial probit with a logit kernel and a general parametric specification of the covariance structure, Working Paper, Dept. of Civil Engineering, MIT, 1996.

[3] Carneiro, Pedro, and Heckman, James. Human Capital Policy, *NBER Working Paper Series*, No. 9495, 2003.

[4] Cunha, Flavio, and Heckman, James. The Technology of Skill Formation, *American Economic Review*, Vol. 97-2, pp. 31-47, 2007.

[5] Cunha, Flavio, and Heckman, James. Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation, *Journal of Human Resources*. Vol. 43, No. 4, 2008.

[6] Curtin, T.R., Ingels, S.J., Wu, S., and Heuer, R. National Education Longitudinal Study of 1988: Base-Year to Fourth Follow-up Data File User's Manual (NCES 2002-323). Washington, DC: U.S. Department of Education, National Center for Education Statistics, 2002.

[7] Eckstein, Zvi, and Wolpin, Kenneth. Why Youths Drop Out of High School: The Impact of Preferences, Opportunities and Abilities, *Econometrica*, Vol. 67, No. 66, pp. 1295-1339, 1999.

[8] Eide, E.R. and Ronan, N. Is participation in high school athletics an investment or a consumption good? Evidence from high school and beyond. *Economics of Education Review,* Vol. 20, pp. 431–442, 2001.

[9] Fryer Jr., R. Financial Incentives and Student Achievement: Evidence from Randomized Trials, *Working Paper*, Dept. of Economics, Harvard University, 2010.

[10] Ingels, S. J., Dowd, K. L., Baldridge, J. D., Stipe, J. L., Bartot, V. H.; Frankel, M. R. (1994). Second Follow-Up: Student Component Data File User's Manual (NCES 94-374). Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, 1994.

[11] Heckman, J., Lochner, L. and Todd, P. Fifty Years of Mincer Earnings Regressions, IZA Discussion Paper, No. 775, 2003.

[12] Heckman, James, Stixrud, Jora, and Urzua, Sergio. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior, *Journal of Labor Economics*, Vol. 24, No. 3, 2006.

[13] Kuhn, P. and Weinberger, C. Leadership Skills and Wages, *Journal of Labor Economics*, Vol. 23, No. 3, pp. 395-436, 2005.

[14] Lipscomb, S. Secondary School Extracurricular Involvment and Academic Achievement: A Fixed Effects Approach, *Economics of Education*, Vol. 26, pp. 463-472, 2007.

[15] McFadden, D. A method of simulated moments for estimation of discrete response models without numerical integration, *Econometrica*, Vol. 57, pp. 995-1026, 1989.

[16] Mahoney, Joseph L. (Ed); Larson, Reed W. (Ed); Eccles, Jacquelynne S. (Ed) Organized activities as contexts of development: Extracurricular activities, after-school and community programs. Mahwah, NJ: Lawrence Erlbaum Associates Publishers, 2005.

[17] Miller, B. M. Afterschool Programs and Educational Success (Critical Hours: Executive Summary). Quincy, MA: Nellie Mae Education Foundation, 2003.

[18] Murnane, R.J. and Levy, F. Teaching the new basic skills: Principles for educating children to thrive in a changing economy. New York, NY: The Free Press, 1996.

[19] Persico, N., Postlewaite A. and Silverman, D. The Effect of Adolescent Experience on labor Market Outcomes: The Case of Height, *Jornal of Political Economy*, Vol. 112, No. 5. pp. 1019-1053, 2004.

[20] Sondheimer, E. Los Angeles City Section program faces $1.4-million budget cut, *Los Angeles Times*, January 25, 2010. Link: http://articles.latimes.com/2010/jan/25/sports/la-sp-sportscuts-citysection26-2010jan26.

[21] Taber, Christopher. The Rising College Premium in the Eighties: Return to College or Return to Unobserved Ability?, *Review of Economic Studies*, Vol. 68, pp. 665-691, 2001.

[22] Train, K. Discrete Choice Methods with Simulation. New york, NY: Cambridge University Press, 2003.

[23] Tunning, R. Budget cuts recommended to school board, *The Citizen*, January 19, 2010. Link: http://www.citizen.com/apps/pbcs.dll/article?AID=/20100119/GJNEWS02/701199933/-1/citnews.

[24] U.S. Dept. of Education, National Center for Education Statistics. NATIONAL EDUCATION LONGITUDINAL STUDY:BASE YEAR THROUGH FOURTH FOLLOW-UP, 1988-2000 [Computer file]. ICPSR version. Washington, DC: U.S. Dept. of Education, National Center for Education Statistics [producer], 2002. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2004.

Appendix

The sample used in the paper is originally from the *National Educational Longitudinal Study: Base Year Through Fourth Follow-up, 1988-2000* (NELS88). The original data has 12,144 students who were interviewed in 1988 (base-year), 1990 (first follow-up), 1992 (second follow-up), 1994 (third follow-up), and 2000 (fourth follow-up). The inital sample consists of 3,071 white men (f4race=4 and f4sex=1), who had finished high school by the time of the third follow-up (f4univ1=1126), and stayed in the same high school over the years (f2f1scfl=1). During the construction of variables, individuals with missing values were dropped from the sample as explained below for each variable.

Study

The variables *study time* (d_s^1, d_s^2) are equal to one if the time spend doing homework out of school is greater than five hours per week (f1s36a2>3 & f1s36a2<8) in the first follow-up, and if the students spends more than five hours per week doing homework (f2s25f2>3 & f2s25f2<9) in the second follow-up, respectively. Otherwise, the variables are equal to zero. If the value was missing, the student responded multiple values, or the student skipped the questions, the observation was dropped.

Participation in Athletics

The variables *participation in athletics* (d_a^1, d_a^2) are equal to one if the student participated in baseball/softball (f1s41aa>2 & f1s41aa<7), basketball (f1s41ab>2 & f1s41ab<7), football (f1s41ac>2 & f1s41ac<7), soccer (f1s41ad>2 & f1s41ad<7), swimming (f1s41ae>2 & f1s41ae<7), other1 (f1s41af>2 & f1s41af<7), or other2 (f1s41ag>2 & f1s41ag<7) in the first follow-up, and if he participated in team1 (f2s30aa>2 & f2s30aa<6), ind1 (f2s30ab>2 & f2s30ab<6), team2 (f2s30bj>2 & f2s30bj <5), or ind2 (f2s30bk>2 & f2s30bk<5) in the second follow-up, respectively. If he did not participated in any sport, the variables are equal to zero. If there was missing values or multiple values for all questions, the student skipped all questions, or the school does not offer any athletic activity, the observation was dropped.

Participation in Clubs

The variables *participation in clubs* (d_b^1, d_b^2) are equal to one if the student participated in the band or orchestra (f1s41ba>2 & f1s41ba<5), theater or musical clubs (f1s41bb>2 & f1s41bb<5), student government (f1s41bc>2 & f1s41bc<5), newspaper or yearbook (f1s41be> 2 & f1s41be<5), or school service clubs (f1s41bf>2 & f1s41bf<5) in the first follow-up, and if he participated in the band or orchestra (f2s30ba>2 & f2s30ba<5), theater or musical clubs (f2s30bb>2 & f2s30bb<5), student government (f2s30bc>2 & f2s30bc<5), newspaper or yearbook (f2s30be>2 & f2s30be<5), or school service clubs (f2s30bf>2 & f2s30bf<5) in the second follow-up, respectively. If he did not participated in any of the above clubs, the variable is equal to zero. If there was missing values or multiple values for all questions, the student skipped all questions, or the school does not offer any club activity, the observation was dropped.

Work

The variables *work* (d_w^1, d_w^2) are equal to one if the student was currently employed (f1s84=5) and reported to work more than ten hours a week (f1s85>1 & f1s85<6) by the time of the first follow-up, and was currently employed (f2s86a=2) and working more than ten hours per week (f2s88>2 & f2s88<10) by the time of the second follow-up, respectively. Otherwise, the variables are equal to zero. If the value was missing, the student responded multiple values, or the student skipped the questions, the observation was dropped.

College

The variable *college* (d_c) is equal to one if the student obtained a bachelor degree or higher (f4hhdg>3 & f4hhdg<7) by the time of the fourth follow-up. Otherwise, the variable is equal to zero. If the value was missing the observation was dropped. All the cases of students who skipped the question were actually cases of students who never attended a post secondary institution (f4attpse=2), so their variable *college* was set equal to zero as it was a legitimated skip.

Test Scores

The variables *test scores* (g^1, g^2) are based on the average IRT Theta Score obtained by the student on reading (f12xrth), mathematics (f12xmth), science (f12xsth) and history, citizenship, and geography (f12xhth) IRT Tests during the first follow-up, and on reading (f22xrth), mathematics (f22xmth), science (f22xsth) and history, citizenship, and geography (f22xhth) IRT Tests during the second follow-up, respectively. After taking the averages of the tests, the variable was discretized using the four quartiles levels within each period. If the score was missing or the student did not take one of the tests, the observation was dropped. For details about the IRT Test and the IRT Theta Score please refer to *NELS88 Second Follow-up: Student Component Data File User's Manual (1994)*.

Wages during High School

The variables *wages* (w^1, w^2) are non-missing and positive if the student was currently employed (f1s84=5) and had reported a hourly wage rate (f1s88) in the first follow-up, and was currently employed (f2s86a=2) and had reported a hourly wage rate (f2s91) by the time of the second follow-up, respectively. The wage rates in the original survey are categorical. A numerical variable was created using the average of each category. If the student responded that he worked, but the wage rate was missing, the student responded multiple values, refused to answer the question, or the student skipped the question, the observation was dropped. The final variable, weekly wage rate, is the hourly rate multiplied by 20.

Wages during Adulthood (2000)

The variable *wage in 2000* (W) was created using the earning period (f4bratp) and wage rate (f4brate) reported by the student by the time of the fourth follow-up (2000). The wages were annualized whenever the earning period reported was shorter than a year (for example, monthly, weekly, etc.). If the individual reported an hourly rate, the variable was adjusted by the number of hours worked per week (f4bjhpw) reported by the student, as long as this number was not greater than 50 hours a week. If the information about the wage rate was missing the observation was dropped.