WHO BENEFITS MOST FROM COLLEGE?

EVIDENCE FOR NEGATIVE SELECTION IN HETEROGENEOUS ECONOMIC RETURNS TO HIGHER EDUCATION

JENNIE E. BRAND
University of California-Los Angeles

YU XIE
University of Michigan

Forthcoming in American Sociological Review, April 2010

Direct all correspondence to Jennie E. Brand, Department of Sociology, University of California-Los Angeles, 264 Haines Hall, Los Angeles, CA 90095-1551 (brand@soc.ucla.edu). Financial support for this research was provided by the National Institutes of Health, Grant 1 R21 NR010856-01. Versions of this article were presented at the Institute for Research on Poverty at the University of Wisconsin-Madison, the Department of Sociology at Yale University, the Department of Sociology at the University of California-Los Angeles, Carolina Population Center at the University of North Carolina-Chapel Hill, the Center for Poverty and Inequality at Stanford University, the 2008 Annual Meeting of the Population Association of America, and the 2008 Annual Meeting of the American Sociological Association. We thank Sara Goldrick-Rab, Eric Grodsky, Robert Hauser, James Heckman, Michael Hout, Robert Mare, Stephen Morgan, Kimberlee Shauman, Jeff Smith, Hung-Hua Tien, Florencia Torche, Donald Treiman, and several anonymous reviewers from *ASR* for helpful comments and suggestions. We are grateful to Ben Jann for his STATA programming assistance. The ideas expressed herein are those of the authors.

ABSTRACT

In this article, we consider how the economic return to a college education varies across members of the U.S. population. Based on principles of comparative advantage, scholars commonly presume that positive selection is at work, that is, individuals who are most likely to select into college also benefit most from college. Net of observed economic and noneconomic factors influencing college attendance, we conjecture that individuals who are *least* likely to obtain a college education benefit the most from college. We call this theory the *negative selection hypothesis.* To adjudicate between the two hypotheses, we study the effects of completing college on earnings by propensity score strata using an innovative hierarchical linear model with data from the National Longitudinal Survey of Youth 1979 and the Wisconsin Longitudinal Study. For both cohorts, for both men and women, and for every observed stage of the life course, we find evidence suggesting negative selection. Results from auxiliary analyses lend further support to the negative selection hypothesis.

WHO BENEFITS MOST FROM COLLEGE? EVIDENCE FOR NEGATIVE SELECTION IN HETEROGENEOUS ECONOMIC RETURNS TO HIGHER EDUCATION

Educational expansion is one of the most apparent, enduring, and consequential features of modern society. Considering the significant educational expansion in the United States during the twentieth century, particularly at the postsecondary level, Fischer and Hout (2006:247) conclude that "the division between the less- and more-educated grew and emerged as a powerful determiner of life chances and lifestyles." In 2007, the U.S. Census Bureau reported that college graduates earned about $55,000, on average, compared with less than $30,000 for individuals who had only a high school diploma. Social scientists have long been interested in questions about access to and the impact of higher education (e.g., Blau and Duncan 1967; Hout 1988; Hout and DiPrete 2006). Scholars have asked: (1) What family and individual attributes are associated with the attainment of higher education? and (2) What are the causal effects of higher education on subsequent socioeconomic outcomes?

In the rational-behavioral model, common in the economics literature, the questions posed above are intrinsically intertwined: individuals make decisions about whether to pursue higher education on the basis of cost-benefit analyses. People choose higher education only if it increases their lifetime earnings expectations (Becker 1964; Card 1995, 2001; Heckman and Honoré 1990; Manski 1990; Mincer 1974; Willis and Rosen 1979). In other words, barring imperfect information, constraints on borrowing funding, or uncertainty, individuals choose to attend college according to expected economic returns; people attain college educations only if the economic returns outweigh the costs. Although this utility maximization paradigm can accommodate noneconomic factors in principle, scholars partial to this approach seldom consider such factors in studying higher education.[1] If economic factors are the main determinants of college attainment, it follows that individuals who are most likely to attend college would also benefit most from college (Carneiro, Hansen, and Heckman 2003; Carneiro, Heckman, and Vytlacil 2001, 2007; Heckman, Urzua, and Vytlacil 2006; Willis and Rosen 1979). We call this thesis the *positive selection hypothesis.*

The sociological literature usually treats the two research questions posed above separately, due to the recognition that higher education is an achieved status subject to the influences of numerous factors (Boudon 1974; Bourdieu 1977; Bowles and Gintis 1976; Coleman 1988; DiMaggio 1982; Jencks et al. 1972; Lucas 2001; MacLeod 1989; Mare 1981; Morgan 2005; Sewell, Haller, and Ohlendorf 1970). That is, a key theme of this literature is that college-going behavior is governed not only by rational choice but also by cultural and social norms and circumstances (Coleman 1988). As such, mechanisms influencing college attainment may differ by social background. For some individuals from socially advantaged backgrounds, college is a culturally expected outcome. For this group, college is less exclusively and intentionally linked to economic gain than it is for people in less advantaged groups, for whom a college education is a novelty that may well demand economic justification (Beattie 2002; Boudon 1974; Smith and Powell 1990). In addition, less-educated workers' earnings prospects are bleak, particularly if they come from disadvantaged backgrounds. By contrast, people from advantaged backgrounds have a high likelihood of attending college and relatively high earnings prospects. Once we partial out observed covariates that help predict college education, it is possible that, due to differential selection mechanisms and earnings prospects, individuals who are *least* likely to obtain a college education benefit most from college. We call this conjecture the *negative selection hypothesis.*

To adjudicate between the positive and negative selection hypotheses, we conduct an empirical study analyzing data from two large U.S. longitudinal surveys: the National Longitudinal Survey of Youth 1979 cohort and the Wisconsin Longitudinal Study 157 cohort. Using these two data sources enables us to curb relative strengths and weaknesses across the datasets, such as quality of available control variables, and to establish robustness of our results. They also enable us to examine possible differences in returns over the life course across cohorts. Because individuals invest in higher education with the expectation of obtaining economic benefits over the lifetime (Mincer 1974), it is important to consider variation in returns to higher education over the life course.

We use a three-step methodological approach. First, we invoke an ignorability assumption that, after we control for a rich set of observed covariates, there are no additional confounders between individuals who do and do not complete college. Under the ignorability assumption, we summarize in estimated propensity scores systematic differences in covariates between college and non-college goers (Rosenbaum and Rubin 1983, 1984; Rubin 1997). Second, we estimate the effects of college completion on earnings by propensity score strata and examine patterns of effects using a hierarchical linear model (Xie and Wu 2005). This innovative key step allows us to find either a positive or a negative pattern between the effects of a college education and the likelihood of obtaining a college education. Third, we revisit the ignorability assumption and conduct auxiliary analyses that aid our interpretation of the results. In a sensitivity analysis, we omit several key covariates to explore the consequences of violating the ignorability assumption. We conduct all analyses separately for men and women.

THEORETICAL AND METHODOLOGICAL ISSUES IN ASSESSING COLLEGE RETURNS

Two Sources of Selection Bias

As is well-known in the causal inference literature, but seldom acknowledged in empirical sociological research, there are two types of selection bias in observational data (Morgan and Winship 2007). The first type is due to heterogeneity in preexisting conditions, or attributes that are associated with both the treatment condition and the outcome. In the case of economic returns to higher education, attributes such as mental ability and work habits may be positively associated with the likelihood of attaining higher education and higher earnings. The second type of selection bias is due to heterogeneity in treatment effects, that is, systematic differences between individuals who do and do not attain a college education in the causal effect of a college education on earnings. Economic returns to higher education should vary across members of a society (Card 1999) because it is implausible to assume that different members of a population respond identically to treatment college education. In this research, we depart from the population homogeneity assumption and focus on group-level variability by aggregating individuals according to their estimated likelihood of completing college.[2] Based on observed attributes, we ask whether individuals who are more likely to attain college educations receive higher or lower returns to college education relative to individuals who are less likely to attain college educations. This approach allows us to explore the potential association between the two sources of population heterogeneity.

To illustrate the first source of bias, let us begin with a standard model in which the effect of a college education is homogeneous. For the *i*th person, the following regression function decomposes the observed logged earnings into the sum of three additive parts (in addition to a constant term)—the treatment effect of college education, a linear combination of covariates, and the residual:

$$y_i = \alpha + \delta d_i + \boldsymbol{\beta}'\boldsymbol{X}_i + U_i, \qquad (1)$$

where y is the natural logarithm of earnings, d is a dummy representing whether the respondent completes college (1 if yes; 0 otherwise), $\boldsymbol{X}$ is a vector of earnings determinants that may also influence the probability of completing college, and U is the residual unexplained by the baseline model. The parameters $\boldsymbol{\beta}$ are regression coefficients measuring the changes in log earnings associated with changes in the earnings determinants $\boldsymbol{X}$, which typically include several measures of family socioeconomic status, geographic residence, academic achievement, and, in many studies, some measure of mental ability. The exponential transformation of the regression coefficient δ represents the multiplicative increase in earnings associated with the receipt of a college degree, *ceteris paribus.* In Equation 1, δ is assumed to be an unknown constant parameter, invariant across all members of the population. Extensions to this standard model may incorporate additional explanatory factors for earnings, such as institutional selectivity, academic major, and academic performance (Brewer, Eide, and Ehrenberg 1999; Dale and Krueger 2002; Thomas 2003; Thomas and Zhang 2005; Zhang 2005).

If homogeneity is true, the main threat to causal inference is that an ordinary least squares (OLS) regression of y on d, even controlling for $\boldsymbol{X}$, is subject to the first source of selection bias due to a non-zero correlation between U and d (Griliches 1977). Under the homogeneity assumption, the conventional wisdom is that OLS estimates of the economic return to schooling are upwardly biased (Griliches 1977; Hauser and Daymount 1977), because factors such as unobserved ability and work ethics should positively affect both education and earnings. The actual direction of the bias, however, has not been empirically settled. For instance, Ashenfelter and Krueger (1994) contend that OLS estimates of the effect of education on earnings are downwardly biased because such estimates are often below instrumental variable (IV) estimates of returns to schooling. If we relax the unrealistic homogeneity assumption, however, there is no simple answer as to whether OLS estimates are biased upward or downward: the OLS estimate is essentially a weighted average of heterogeneous effects, some of which are necessarily higher, while others are lower, than the population average (Angrist and Krueger 1999; Morgan and Winship 2007). Under this more realistic conceptualization of underlying heterogeneity in the returns to education, individuals differ not only in background attributes but also in the economic benefits they reap from a college education.

To systematically study heterogeneous treatment effects of higher education on earnings, we adopt a simple approach using rich covariates and invoking ignorability, at least provisionally. This approach allows us to find empirical patterns of treatment-effect heterogeneity as a function of observed covariates. A common way of studying heterogeneous treatment effects by observed covariates is to examine the interaction between education and specific factors that influence wages and the probability of attaining a college education, such as race or gender (Barrow and Rouse 2005; Perna 2005; Welch 1973), or parents' education or occupation (Altonji and Dunn 1996; Hauser 1973; Olneck 1979). When comparing returns to college between individuals who complete college and those who do not, however, the most meaningful interaction is between college education and the propensity to complete college (Heckman et al. 2006).[3] We aggregate heterogeneous college effects to propensity score group-level mean effects and directly observe trends in effects (Xie and Wu 2005).[4]

The ignorability assumption (also called “unconfoundedness” or “selection on observables”) states that potential outcomes are uncorrelated with treatment status, conditional on observed covariates. The assumption can never be verified, and indeed should not be taken as true in practice for observational data; its plausibility depends on the availability of observed covariates that differ between college graduates and non-college graduates and also influence earnings. It is reasonable to suspect that models that do not control for cognitive ability, for instance, do not satisfy ignorability. Still, measurement of meaningful confounders renders ignorability tentatively more plausible, although not necessarily true. While we do not think the ignorability assumption is true, analyses under this assumption are the most the data can tell us without additional unverifiable assumptions.[5] Using our strategy, we focus on group differences by the propensity to complete college and adjudicate between two potential patterns in observed heterogeneous effects of college completion on earnings: positive selection (individuals most likely to benefit from college are most likely to complete college) versus negative selection (individuals most likely to benefit from college are *least* likely to complete college).

Positive versus Negative Selection

In economics, human capital theory is an influential explanation for educational acquisition (Becker 1964; Mincer 1974). The core idea of the theory is that a gradation in earnings by education level reflects returns to individuals’ rational investment in education. If λ represents the present value of the lifetime economic return to college, and c the cost of college, attending college produces a net gain if $\lambda > c$, with the benefit thus defined as $\pi = \lambda - c$. The association between the returns to college and the decision to attend college is at the core of more recent literature that links variation in returns to education to heterogeneous schooling behavior. Premised on principles of self-selection and comparative advantage, the thesis is that the most “college worthy” individuals, in the sense of having the highest returns to college, are the most likely to select into college (Averett and Burton 1996; Carneiro, Hansen, and Heckman2003; Carneiro, Heckman, and Vytlacil 2001, 2007; Roy 1951; Willis and Rosen 1979). These individuals are also in a better position to cover the economic costs of a college education, particularly at high-cost institutions (Zhang 2005). According to this literature, positive selection should occur because individuals who stand to benefit the most from a college education are most likely to select into college.

The positive selection thesis is widely, albeit not universally, accepted in economics. In our view, it is more a theoretic argument than a proposition that can readily be subject to empirical tests. In economics, empirical research on choice relies heavily on the revealed preference framework (e.g., Manski and Wise 1983; Train 2003). Applied to our research question here, the framework essentially states that a researcher can infer that $\lambda > c$, at least in expectation, if a person is observed to complete college, and $\lambda \leq c$ otherwise. Willis and Rosen (1979) use this strategy in their classic study that applies Roy’s (1951) model to the college education question, with the difference in expected utility between college and high school education determining the likelihood of attending college. More recently, Carneiro, Heckman, and Vytlacil (2007) also report evidence they interpret as positive self-selection (i.e., individuals with the greatest expected returns are the most likely to attend college).

Sociologists, too, recognize heterogeneity in returns to college. Raftery and Hout (1993:57), for example, state that it “seems likely that the perceived benefit of education varies among individuals” as a function of individual attributes. Like economists, sociologists infer that the choice of attending college can result from a cost-benefit analysis (Boudon 1974; Breen and Goldthorpe 1997; Raftery and Hout 1993); sociologists, however, emphasize that the costs and benefits are not purely economic. For instance, in terms of costs, sociologists have considered heterogeneity in terms of both the financial burden and the family pressure stemming from deviating from class-based cultural norms (Boudon 1974; Raftery and Hout 1993).

In contrast to the strictly economic cost-benefit model of college attendance, much research indicates that multiple actors and factors influence college attendance. Beginning with the Blau-Duncan model, sociologists have recognized the significance of numerous family background factors for educational attainment, such as parents' education and occupation, family structure (McLanahan, and Sandefur 1994), and sibship size (Blake 1981).[6] The "Wisconsin model" of status attainment further specifies the concrete processes by which family background affects educational attainment: family socioeconomic status and measured ability affect occupational and educational aspirations, as does encouragement from parents and significant others (Hauser, Tsai, and Sewell 1983; Sewell et al. 1970; Sewell, Haller, and Portes 1969; Sewell and Hauser 1975).[7] Coleman (1988), too, offers insight into how family background factors influence children's attainment via the concept of social capital, that is, social relationships consisting of expectations, information channels, and social norms. Encouragement, expectations, information, and norms differ by family background, generating differential mechanisms of selection into college (Morgan 2005).

In addition, sociologists have developed a neo-Marxist conflict perspective that helps explain differences in educational attainment by social background. For instance, cultural capital scholars stress the importance of family background for educational attainment, emphasizing general cultural background, knowledge, disposition, and skills that children acquire from their parents. Sociologists further argue that schools systematically reward the cultural capital of the advantaged classes and devalue that of the lower classes (Bourdieu 1977; DiMaggio 1982;

Lareau 2003). Social reproduction theorists elaborate on this theme, maintaining that primary and secondary schools train advantaged students to take their positions at the top of the socioeconomic order (e.g., by pursuing postsecondary schooling), while conditioning the poor to accept their lower status in the class structure (Bowles and Gintis 1976; MacLeod 1989). In summary, this literature suggests that high social background individuals are likely to go to college even in the absence of a rational economic cost-benefit analysis, whereas low social background individuals must overcome considerable odds to attend college.

<FIGURE 1>

Past research in social stratification provides a compelling theoretical and empirical basis for postulating variation in the effects of education on earnings by social background. This research shows that the direct relationship between social origin and destination (both measured by occupational status) is much weaker for college graduates than for workers without college degrees (Hout 1984, 1988). Figure 1 depicts this empirical pattern. If we change the perspective and examine returns to schooling (Goldthorpe and Jackson 2008) (i.e., the difference in destination between college-educated and less-educated workers) as a function of social origin, this interaction pattern yields a smaller difference by college education for individuals of high social origin (δ_2) than for individuals of low social origin (δ_1). In other words, individuals with relatively disadvantaged social backgrounds, or those with the lowest probability of completing college, benefit the *most* from completing college. This pattern results from the particularly poor labor market prospects for workers with low levels of education combined with low levels of other forms of human, social, or cultural capital.[8] This collective theoretical and empirical tradition leads to our negative selection hypothesis.

We are not the first sociologists to discuss possible patterns of negative selection (Brand and Halaby 2006; Bryk, Lee, and Holland 1993; DiPrete and Engelhardt 2004; Hoffer, Greeley, and Coleman 1985; Morgan 2001; Tsai and Xie 2008). Studies show, for example, that high school environment has a stronger effect on marginal college attendees than on more advantaged students (Bryk et al. 1993; Hoffer et al. 1985). The economics literature also provides direct empirical evidence in support of negative selection in higher education. An economic study reports that a randomly chosen person might expect to receive a 9 percent increase in wages due to college education, while those actually selecting into college receive about a 4 percent increase (Heckman, Tobias, and Vytlacil 2001). Additionally, studies that use compulsory schooling laws, differences in the accessibility of schools, or similar features as instrumental variables find larger economic returns than do OLS estimates (Card 2001). This suggests larger returns to education for individuals on the margin of school continuation.[9]

Behavioral Model

We specify the behavioral model for college education as the following: let d^* represent the potential likelihood that the ith person completes college, and d_i the observed outcome (1 if yes; 0 otherwise). It is customary to relate the two through a threshold measurement model:

$$d_i = 1 \text{ if } d_i^* > 0; \qquad (2)$$

$$d_i = 0 \text{ otherwise.}$$

We further specify that college attainment is determined by a weighted average of an economic component π_i , a noneconomic component η_i, and a residual ε_i:

$$d_i^* = w_i\pi_i + (1 - w_i)\eta_i + \varepsilon_i, \qquad (3)$$

where ε_i is assumed to be independent of π_i, η_i, and w_i, with $0 \le w_i \le 1$. A key insight from the sociological literature is that the relative weight w_i given to the economic component may decrease with the noneconomic determinant η_i (i.e., a negative correlation between the two in the population). We further assume π_i to be a linear function of observed covariates (λ_1'X) plus an unobserved component μ_i, and η_i to be a linear function of observed covariates (λ_2'X). We can rewrite Equation 3 as the following:

$$d_i^* = w_i\lambda_1'X_i + (1-w_i)\lambda_2'X_i + w_i\mu_i + \varepsilon_i. \qquad (4)$$

The likelihood of completing college is high when d_i* is large. Writing out the model of Equations 1 through 4 makes it easier to appreciate the key difference between the economic and noneconomic factors influencing college attainment. In the traditional Roy-type college behavioral model, $w_i = 1$, and μ_i drives the college education decision, conditional on $\boldsymbol{X}$ (Willis and Rosen 1979). In most sociological literature, familial, personal, and institutional characteristics dominate (i.e., w_i is much smaller than 1), so that the observed covariates $\boldsymbol{X}$ primarily determine the decision rule, with the self-selection component given the secondary role or sometimes ignored (i.e., $w_i\mu_i = 0$).

Equation 4 cannot be estimated because it is unidentified. As a research strategy, we invoke the ignorability assumption and thus assume away the unobserved self-selection component (μ_i) as a first step in the data analysis. We further simplify the equation into a misspecified but estimable reduced-form propensity score model:

$$d_i^* = \boldsymbol{\lambda}'\boldsymbol{X}_i + \nu_i. \qquad (5)$$

How does the misspecification of Equation 5 affect our ability to make inferences regarding propensity score-specific causal effects of college education on earnings? The sociological literature suggests that because w should be negatively correlated with the observed propensity score, the extent of misspecification caused by omitting μ declines with the observed propensity score; that is, the decision to go to college among children from high-status families is dictated less by rational choice and self-selection than it is among children from low-status families. When a person who is not expected to go to college based on observed characteristics does go to college, there are strong factors involved, one of which may be the economic incentive.

STATISTICAL MODELS

To fix ideas, we adopt the potential outcome approach to causal inference. The potential outcome approach has early roots in experimental designs (Neyman 1923) and economic theory (Roy 1951) and has been extended and formalized for observational studies in statistics (e.g., Holland 1986; Rosenbaum and Rubin 1983, 1984; Rubin 1974), economics (e.g., Heckman 2005; Manski 1995), and sociology (e.g., Morgan and Winship 2007; Sobel 2000; Winship and Morgan 1999). The approach makes explicit the issues that concern the identification and estimation of causal effects. Let y be logged earnings, and again let d be a variable scored 1 for an individual who completes college and 0 otherwise. We ask what individual i's earnings would be if he or she were to receive the treatment (i.e., complete college), compared with not receiving the treatment (i.e., not complete college). As only one of the two earnings values, y_i^1 or y_i^0, is actually observed, causal inference is impossible at the individual level; it always requires statistical analysis at the group level on the basis of some homogeneity assumption (Holland 1986).

To infer causality with observational data, it is necessary to introduce unverifiable assumptions. In this research, we first introduce the ignorability assumption:

$$E(y^0 \mid \boldsymbol{X}, d = 1) = E(y^0 \mid \boldsymbol{X}, d = 0) \qquad (6a)$$

and

$$E(y^1 \mid \boldsymbol{X}, d = 0) = E(y^1 \mid \boldsymbol{X}, d = 1). \qquad (6b)$$

Equation 6a assumes that the average earnings of college-educated workers, had they not completed college, would be the same as the average earnings of non-college-educated workers, conditional on observed covariates. Likewise, Equation 6b assumes that the average earnings of non-college-educated workers, had they completed college, would be the same as the average earnings of college-educated workers, conditional on observed covariates.

Models for Heterogeneous Treatment Effects

When treatment effects are heterogeneous, there can be two types of selection bias, as we discuss above: pretreatment heterogeneity bias and treatment-effect heterogeneity bias. Both types of bias can threaten the validity of causal inference with observational data. Estimators such as fixed-effects and the difference-in-differences attempt to eliminate pretreatment heterogeneity bias but not treatment effect heterogeneity bias (Angrist and Krueger 1999).

If we allow the coefficient of treatment in Equation 1 to be heterogeneous, we can, at least theoretically, write out the two types of heterogeneous components. Equation 1 becomes the following:

$$y_i = \alpha_i + \delta_i d_i + \boldsymbol{\beta}'\boldsymbol{X}_i + U_i. \qquad (7)$$

In this specification, α_i represents pretreatment heterogeneity, while δ_i represents treatment-effect heterogeneity.[10] If there is pretreatment heterogeneity bias, correlation $\rho(\alpha, d) \neq 0$. If there is treatment-effect heterogeneity bias, correlation $\rho(\delta, d) \neq 0$ (Heckman et al. 2006; Winship and Morgan 1999). The individual-level heterogeneity model is not identifiable, as α_i and δ_i cannot be separated from U_i without further constraints; we invoke the ignorability assumption. In practice, conditioning on $\boldsymbol{X}$, which is typically multidimensional, proves difficult due to the "curse of dimensionality"; we cannot often find treated and untreated units with identical values on $\boldsymbol{X}$ if $\boldsymbol{X}$ is of a high dimension. However, Rosenbaum and Rubin (1983, 1984) show that, given the ignorability assumption, it is sufficient to condition on the propensity score as a function of $\boldsymbol{X}$. The propensity score is defined as the probability of assignment to the treatment group (college completion) given covariates $\boldsymbol{X}$:

$$P = p(d_i = 1 \mid \mathbf{X}) \qquad (8)$$

In this study, we evaluate heterogeneity in treatment effects by decomposing δ in Equation 7 into a nonparametric function of the propensity score and use a hierarchical linear model to reveal a pattern of returns.[11] Based on observed family and personal attributes, we can divide a group into subpopulations with similar predicted propensity scores to complete education. We then assess whether population heterogeneity in the propensity to complete college is associated with heterogeneity in returns to college. Specifically, we ask if the estimated effect of college is positively or negatively associated with the estimated propensity to complete college.

Our analytic strategy proceeds in three steps: (1) We estimate binary logistic regressions predicting the probability of completing college and derive propensity scores for each individual in the sample. We group respondents into strata of estimated propensity scores to balance the distributions of the covariates between college graduates and non-college graduates ($p < .001$). (2) In level 1, we estimate the treatment effects specific to balanced propensity score strata using ordinary least squares regression. (3) In level 2, we examine the heterogeneous results by propensity score strata and summarize the trend in the variation of effects using a hierarchical linear model (Xie and Wu 2005). Our approach is similar to propensity score matching, as respondents' observed differences are characterized by propensity scores. The two methods differ in how comparisons are constructed. In a typical propensity score matching analysis, comparison by treatment status is made on an individual basis and averaged over the population or a subpopulation. In our approach, comparison by treatment status is constructed for a relatively homogeneous group based on propensity scores and examined across different groups of similar propensity scores through a hierarchical linear model.[12]

DATA, MEASURES, AND DESCRIPTIVE STATISTICS

Data Description

To examine heterogeneous treatment effects of education on earnings, we use two large panel datasets containing extensive information about respondents' social backgrounds, abilities, and schooling experiences: the National Longitudinal Survey of Youth 1979 (NLSY)[13] and the Wisconsin Longitudinal Study (WLS).[14] Both samples are cohort-based. Single-cohort longitudinal surveys are advantageous in controlling for the potential confounding effect of cohort with experience. The NLSY is a nationally representative sample of 12,686 respondents who were 14 to 22 years old when first surveyed in 1979. These individuals were interviewed annually through 1994 and biennially thereafter. We restrict our sample to respondents who were 14 to 17 years old at the baseline survey in 1979 (N = 5,581), had not graduated from high school at the time the Armed Services Vocational Aptitude Battery (ASVAB) tests were administered (N = 3,885), had completed at least the 12th grade as of 1990 (N = 3,034), and do not have any missing data on the set of covariates used in our analysis (N = 2,474). We set these sample restrictions to examine a cohort with little age variation, to ensure that all measures we use are precollege, and to compare college graduates with respondents who completed at least a high school education. We evaluate effects of completing college on earnings for respondents ages 29 to 32 (in 1994), 33 to 36 (in 1998), and 37 to 40 (in 2002), that is, from early- to mid-career years.

The WLS is a regional panel study based on a random sample of 10,317 men and women who graduated from Wisconsin high schools in 1957. Research shows that for processes of socioeconomic attainment, patterns found in the WLS mirror those found in national probability samples (Sheridan 2001). We restrict our sample to respondents who do not have any missing data on the set of covariates used in our analysis (N = 7,905).[15] Replication of the analysis through these data sources for two different cohorts allows us to check the robustness of the core findings. The two data sources are also complementary in their relative strengths and weaknesses. While the NLSY offers national representation, the WLS contains a much larger sample of relatively homogeneous respondents with many well-measured precollege covariates, including a notably reliable measure of cognitive ability.

Variable Measurement

Table 1 lists the precollege variables we use to construct propensity score strata for our two data sources. Most of these measures figure prominently in sociological studies of educational and occupational attainment, and their measurement is straightforward. There are, however, a few differences across data sources in the measurement of these variables. Parents' income is measured as total net family income in 1979 dollars in the NLSY; the WLS uses parents' income in 1957 dollars. "Residence/proximity to college or university" indicates whether a respondent lived in an SMSA in 1979 in the NLSY, and whether a respondent's high school was within 15 miles of a college or university in the WLS. College-prep indicates whether a student was enrolled in a college-preparatory curriculum in the NLSY and had completed the requirements for UW-Madison in the WLS. The measurement of mental ability also differs across the data sources. In 1980, 94 percent of the NLSY respondents were administered the ASVAB, a battery of 10 intelligence tests measuring knowledge and skill in areas such as mathematics and language. We first residualize each of the ASVAB tests on age at the time of the test separately by race and gender, with the residuals standardized to have mean zero and variance one. We then combine the items (with equal loadings that sum to one) into a composite scale (Cronbach's α = .92) (Cawley et al. 1997). In the WLS, we use the 1957 Henmon-Nelson Test of Mental Ability scores. We use hourly wages as the outcome variable in the logarithm form.[16] In the NLSY, our outcome is logged hourly wages and salary for respondents' late 20s through early 40s (in 1994, 1998, and 2002). In the WLS, our outcome is logged yearly earnings at age 35 (in 1975) and logged hourly wages at age 53 (in 1993).[17] We add a small positive constant ($.50) before taking the logs. Unemployed workers are eliminated.[18]

<TABLE 1>

Descriptive Statistics

A higher probability of attaining a college degree is among the most important causal mechanisms for realizing the advantage associated with high socioeconomic origins, a key finding of Blau and Duncan's (1967) classic study. As Table 1 shows, college graduates, compared with individuals who did not graduate from college, are more likely to come from families with high income, highly educated parents, intact family structure, and few siblings. High levels of secondary school academic success, cognitive ability, and encouragement from teachers and parents to attend college, as well as friends who plan to attend college, are also predictive of college education. These statistics suggest that many noneconomic factors figure prominently in youths' educational attainment. Finally, for multifaceted reasons (Kao and Thompson 2003), the likelihood of completing college varies by race and Hispanic origin, with whites and Asians being more likely than blacks and Hispanics to complete college.

MAIN ANALYSIS AND FINDINGS

College Returns under the Assumption of Homogeneity

Table 2 provides the estimated effects of college completion on earnings, separately by sex, through regression analyses under the homogenous effect assumption, controlling for the full set of covariates described above.[19] For NLSY employed men, college completion yields a highly significant positive effect on logged hourly wages that steadily increases over time, from a 20 percent advantage in men's late 20s to early 30s to a 51 percent advantage in their late 30s to early 40s. This is consistent with the human capital model. Given the known increasing temporal trend in returns to college, it is not surprising that the effect of college completion is smaller in magnitude in the earlier WLS cohort. Still, results for WLS men indicate significant and increasing returns associated with a college degree over the life course.

<TABLE 2>

Results for NLSY employed women reveal a large, significant effect of college completion in their late 20s to early 30s, a smaller effect in their mid-30s relative to their early 30s, and then a comparatively larger effect in their late 30s to early 40s.[20] Differences in life course patterns between men and women may reflect the influences of traditional gender roles in the family and corresponding intermittent labor force attachment among women relative to men, particularly during childbearing years (Becker 1991; Bianchi 1995; Mincer and Polachek 1974). That is, women's life course pattern of effects may reflect the selection of some women out of the labor force or a lower additive return to college during childbearing years. In the WLS, the effect of a college degree also declines over the life course for women, but the effect in their mid-30s is larger for WLS respondents than for those in the NLSY. We speculate that sample selection may explain this somewhat peculiar finding. Women's labor force participation was much lower for the WLS cohort (57 percent of WLS women were employed at age 35, compared with 76 percent of NLSY women in their mid-30s); it was thus more selective with respect to earnings than for the NLSY cohort.

Generating Propensity Score Strata

Our next objective is to examine the heterogeneous effects of college completion by propensity score strata. We estimate binary logistic regressions predicting the odds of completing college by the covariates described in Table 1 for each data source, separately by sex, and derive estimated propensity scores for each individual (Becker and Ichino 2002). Table A1 in the Appendix reports results for the logistic regressions. We then generate balanced propensity score strata; balancing is satisfied when within each interval of the propensity score the average propensity score and the means of each covariate do not significantly differ between college and non-college graduates. We restrict the balancing algorithm to the region of common support, that is, to regions of propensity scores in which both treated and control units are observed.[21] To demonstrate the balance achieved within each stratum, we present covariate means by propensity score strata for NLSY men in Table 3.[22] Table 3 also elucidates the characteristics of a typical individual within each stratum. For instance, a characteristic person in stratum 1 has parents who are high school drop-outs, three siblings, low ability, friends who do not plan to go to college, and is enrolled in a nonacademic track. By contrast, a characteristic person in stratum 5 has parents with some college, one sibling, high ability, friends who plan to go to college, and is enrolled in an academic track. Table 4 provides the number of cases in each stratum, separately by college attainment, gender, and data source. As expected, the frequency distributions for college- and non-college-educated individuals run in opposite directions. In the case of college-educated individuals, the frequency count increases with the propensity score, whereas for non-college-educated individuals, the count decreases with the propensity score. Still, we achieve overlap within each stratum: for each propensity score stratum there are individuals with $d = 1$ and other individuals with $d = 0$.[23]

<TABLE 3>

<TABLE 4>

Heterogeneous College Returns

Figures 2 through 5 present the main results of our study. We first estimate treatment effects specific to propensity score strata and then detect the pattern of effects by propensity score with a hierarchical linear model (HLM). Points in Figures 2 through 5 represent estimates of stratum-specific effects of college completion on logged earnings. The linear plots and reported level-2 slopes in the figures are based on the HLMs (i.e., level-2 variance-weighted least squares models estimated by level-1 college effects specific to propensity score strata regressed on propensity stratum rank). All point estimates and associated *t* values corresponding to Figures 2 through 5 are provided in the Appendix, Table A2 for the NLSY and Table A3 for the WLS.

Figure 2 depicts results for NLSY men's college effects on earnings at ages 29 to 32 (in 1994), 33 to 36 (in 1998), and 37 to 40 (in 2002). The downward linear slopes illustrate the declining trend in effects with propensity stratum rank at every observed time period. For instance, for men in their late 20s to early 30s, a unit change in stratum rank is associated with a 5 percent reduction in the treatment effect, such that the predicted effect of college completion on earnings in stratum 1 is about 30 percent, while the predicted effect in stratum 5 is about 10 percent. This means, for example, that an individual with parents who are high school drop-outs, and who himself has low measured ability, benefits more from completing college, on the magnitude of an estimated 20 percent, than would an individual whose parents went to college and who himself has high measured ability. We also find evidence suggesting a declining trend in college effects on earnings for men in their mid- and late-30s and early-40s. In the late-30s to early-40s, for instance, the 5 percent reduction in treatment effect per stratum rank again results in an estimated 20 percent difference between the lowest and highest strata, or between the least and most advantaged college goers. The level-2 slopes thus offer support for the negative selection hypothesis at each observed stage of the life course. As expected, college completion is associated with an increasing economic return over the life course, and this is true across propensity score strata.

<FIGURE 2>

The results for NLSY women, shown in Figure 3, are similar to those for men in suggesting negative selection at each observed stage of the life course. For instance, for women in their late-30s to early-40s, a unit change in stratum rank is associated with a 4 percent reduction in the treatment effect. The predicted effect of college completion on earnings is about 40 percent for stratum 1 women with disadvantaged socioeconomic backgrounds, versus about 25 percent for stratum 5 women with advantaged socioeconomic backgrounds. In contrast to men, however, we observe an oscillating return to college over the life course among women, as we discussed for results under the assumption of homogeneity. Again, differences in life course earnings returns between men and women likely reflect intermittent labor force attachment among women during childbearing years; these family processes could affect women differently by propensity score strata.

<FIGURE 3>

Figures 4 and 5 present results for WLS men and women, respectively.[24] Figure 4 depicts results for WLS men's earnings at age 35 (in 1975) and age 53 (in 1993). Level-2 slopes in the WLS indicate a less than 1 percent reduction per stratum rank at age 35 (the flattest of the level-2 slopes we observe), and a 2 percent reduction per stratum rank at age 53. There are nine strata in the WLS, versus five in the NLSY; a 2 percent reduction per stratum rank at age 53 thus means an 18 percent decrease in the earnings return to college education for the highest stratum relative to the lowest stratum. Although we observe generally lower returns to education in the earlier cohort, the results from the WLS are consistent with those from the NLSY in lending support for negative selection. At every observed stage in the life course, the level-2 slopes reveal that the benefit to completing college is greatest among men *least* likely to complete college (Figures 2 and 4).

<FIGURE 4>

WLS women display a much steeper downward slope at age 35 compared with NLSY women in their mid-30s, the result of a very high return among women in stratum 1. Women from disadvantaged social backgrounds who obtained college degrees in the early 1960s were likely particularly selective, generating an unusually high return. Such women may have been less likely to assume traditional family roles due to a lower likelihood they were married to men with economic resources sufficient for role-specialization within the family relative to their advantaged peers (Hill and Stafford 1974).

<FIGURE 5>

Figures 2 through 5 demonstrate a systematic selection mechanism at work: when individuals with a low propensity of completing college (i.e., individuals from the most disadvantaged social origins and with the lowest ability and achievement) actually complete college, they benefit the most from doing so. Tables A2 and A3 in the Appendix show that the wage gap between the treatment and control groups (level-1 coefficient) is statistically significant within several, but by no means every, propensity score stratum. This facet of our findings is consistent with the negative selection hypothesis: 8 out of the 10 estimated effects in stratum 1 are statistically significant, while only 3 out of the 10 estimated effects in the final stratum are significant. Still, there are few statistically significant level-2 slope coefficients.

Furthermore, level-2 slope coefficients are based on very few data points. There is no "population" of propensity score strata as true level-2 units of analysis. We fit the HLM model to provide an overall one-degree-of-freedom summary of the direction of the pattern of effect heterogeneity as a function of propensity score; the direction is negative in every case. At the least, we can say that the selection patterns in Figures 2 through 5 are clearly not positive. The evidence for the negative selection hypothesis is only suggestive for each case we present. Nevertheless, the core finding of a negative pattern holds true for two different data sources with varying quality of measures, at every observed life course stage for different cohorts, and for both men and women—a total of 10 distinct negative level-2 slopes. Moreover, we found the same pattern for another cohort-based U.S. longitudinal dataset (Brand and Xie 2007).[25] Sill, our results should be taken as descriptive and suggestive, not definitive.

AUXILIARY ANALYSIS

Given the evidence suggesting negative selection, we now consider the question of causal mechanisms. It is plausible, indeed likely, that multiple mechanisms account for the pattern we observe. We first test the idea of differential selection mechanisms by propensity score strata with a measure of the value of a college education among high school seniors in the WLS. "Value of college" is a weighted average of the scores in response to a series of 18 statements regarding perceptions of the value of going to college (Amer 1964). The most highly weighted items include: "I would rather start earning money quickly, and learn on the job"; "learning on a job is more practical than most school learning"; and "going to college would be a waste of time for me." (See WLS Memo 129 for further details on the items and the variable construction.) In Table 5, we examine values of college by propensity score strata and education among WLS men. We find a large differential between college graduates and non-college graduates in low propensity strata: the former values college more than the latter does, while the gap gradually decreases across propensity score strata to almost no difference among those in the highest propensity stratum. The atypically high value that disadvantaged youth who attend college actually place on college contrasts markedly with the uniformly high value (i.e., undifferentiated by actual college completion status) that advantaged youth place on college. This result suggests, as we hypothesized, that mechanisms leading to college attainment differ by social background.

<TABLE 5>

Because the value of college variable encompasses both economic and noneconomic incentives, we use an additional variable to determine whether low propensity college attendees are more economically driven than high propensity college attendees, for whom college attendance is a cultural expectation. Since field of study affects earnings (Thomas and Zhang 2005), we examine stratum-specific college majors for college-educated men in the WLS. Table 6 shows that while low propensity students are more likely to concentrate in business and education—majors that yield immediate economic return—high propensity students are more likely to major in the sciences and humanities, subjects that require strong academic interests and are less likely to be motivated by immediate economic rewards.[26]

<TABLE 6>

Our results suggesting negative selection raise a question: Why do some prior studies suggest empirical support for positive selection? We suspect that one explanation lies in the choice, or availability, of covariates in the analyses. Empirical support for positive selection is sometimes based on models that omit key variables such as ability, high school academic performance, and parents' and teachers' encouragement. Omitting these important confounders may introduce a distortion to the observed pattern of selection from negative to positive. To test this possibility, we now act as if we do not have access to the full set of covariates at our disposal and restrict covariates in the WLS to a set comparable to that used by Carneiro, Heckman, and Vytlacil (2001).[27, 28] We omit ability and academics, social-psychological variables, and religion from our models, highly significant factors in our expanded model reported in the Appendix, Table A1. As expected, we find large mean differences in these covariates between college- and non-college-educated respondents within each propensity score stratum.

Figure 6 shows the revised results for WLS men at age 35, omitting the aforementioned variables. When we restrict models to a more limited set of covariates, we find evidence for positive selection. This figure should be compared with Figure 4, as the analysis is parallel for the same sample, the difference lying in the specification of covariates for the propensity model. Omitting these variables not only changes the overall size of the college effect (as in the case when effects are assumed to be homogeneous), but it also changes the direction of association between propensity of treatment and treatment effects. With the full set of covariates at our disposal, we observe a modest pattern of negative selection; when we trim covariates to a more limited set, we observe positive selection. We cannot attribute the change in direction to the omission of one single covariate; a detailed decomposition analysis modeled after Xie and Shauman (1998) shows that no single covariate can be held accountable.[29] Rather, the omission of the whole set of covariates (ability and academic achievement, aspirations, and encouragement) induces the observed change from a positive to a negative slope. This change is attributable to the fact that these additional covariates exert greater power explaining college completion in higher propensity score strata than in lower propensity strata. This is consistent with our earlier discussion of Equations 3 and 4, where we state that noneconomic factors should play a greater role in predicting college education for persons likely to complete college than for their peers who are less likely to complete college.

<FIGURE 6>

DISCUSSION AND CONCLUSIONS

Heterogeneity in response to a common treatment is a norm, not an exception. Individuals differ not only in background attributes but in how they respond to a particular treatment. An important task of sociological research is to summarize systematic patterns in population variability, a longstanding demographic tradition that Xie (2007) attributes to Otis Dudley Duncan. In this article, we consider population heterogeneity in returns to schooling, examining the effects of completing college by propensity score strata in a hierarchical linear model. We first estimate effects of college for groups based on the likelihood of a college education; we then examine systematic heterogeneity in those group-specific effects. Our analysis depicts whether patterns of population heterogeneity reflect positive or negative selection, that is, whether economic benefits of college are greater among persons most or least likely to complete college. Our evidence suggests negative selection: individuals most likely to benefit from a college education are the *least* likely to obtain one. This finding holds for both men and women, for every observed stage over the life course, and for two different cohorts.

Empirical patterns are generally consistent with our hypothesized interaction effect depicted in Figure 1. The increasing demand for educated workers alongside the decreasing demand for less-educated workers has resulted in an increase in the earnings differential between educated and less-educated workers (Farley 1996). We find, however, that this differential is especially large among individuals with a low propensity for completing college. Therefore, a principal reason for low propensity college-educated workers' relatively large economic return is that their social position is marked by substantial disadvantage. In the absence of a college degree, low propensity men and women have limited human, cultural, and social capital and hence particularly limited labor market prospects. By contrast, in the absence of a college degree, individuals from more advantaged social backgrounds can still rely on their superior resources and abilities. The negative selection pattern does not emerge because low propensity college goers earn more wages than do high propensity college goers; they do not. Rather, the pattern emerges because low propensity non-college goers earn so little.

We realize that using the propensity score to identify heterogeneous treatment effects has limitations, as it relies on the ignorability assumption and overlooks heterogeneity due to unobserved variables. The plausibility of the ignorability assumption is specific to each research setting, depending on the richness of the observed covariates. We invoke the ignorability assumption because we wish to know what the observed data alone can reveal, knowing that we have a set of rich covariates. However, there are always lingering unobservable causal factors or mechanisms. Nevertheless, there are several benefits in focusing on observable heterogeneity in treatment effects. Although treatment-effect heterogeneity is potentially observable, it is seldom studied in empirical sociological research. With a focus on observable heterogeneity, we uncover an important finding: the most disadvantaged individuals with respect to observed social background, achievement, and ability are the most likely to benefit from a college education. Auxiliary analyses lend further support to the negative selection hypothesis: individuals from disadvantaged social backgrounds who attend college may use education as a means for economic mobility, while those from advantaged social backgrounds, for whom college is a cultural norm, may be less purposively driven by an economic rationale. Moreover, we find that empirical support in prior research for positive selection may be a result of missing certain key variables.

We have several plans for future research. First, we limit our focus here to the earnings gap between individuals who complete college and those who complete only high school, which allows us to easily borrow from the methodological literature on causal inference. Although there is a well-documented difference between the two groups in the labor market (Grubb 1993; Kane and Rouse 1995; Mare 1981), it is clearly a simplification to treat education as a dichotomous treatment. In future studies, we will measure higher education more precisely by amount, quality, and major. Second, in this study, we note some interesting differences between men and women patterned by propensity score strata across the life course, which we conjecture reflects differences across strata in women's labor market intermittency. Future research will analyze the heterogeneous effects of higher education on women's family formation patterns in greater depth.

The widespread belief in the socioeconomic return to higher education has prompted policy efforts that expand educational opportunities for all Americans. While many policymakers implicitly assume homogeneity in the return to schooling, potential heterogeneity in returns is receiving more attention as many countries are experiencing rapid expansion in college enrollment. This has led some to question the relative costs and benefits of higher education for those who were not previously receiving it. Yet, in the presence of heterogeneous treatment effects, no simple summary statement can be invoked regarding the benefit of completing college, either for individuals already receiving higher education or for those likely to benefit from educational expansion. The average benefit depends on the composition at any given time of the group of students who complete college. One interpretation of our results is that a college education may be particularly beneficial among groups targeted by educational expansion efforts—that is, individuals who are otherwise unlikely to attend college based on their observed characteristics.

Due to our simplifying methodological assumption, the above conclusion is only tentative. The very pattern of heterogeneous treatment effects of college education on earnings by the propensity to complete college suggests an unobserved selection mechanism at work: individuals from disadvantaged social backgrounds, for whom college is not a culturally expected outcome, overcome considerable odds to attend college and may be uniquely driven by the economic rationale. Thus, if educational expansion results in a larger number of college goers who are otherwise unlikely to attend college, unobserved selectivity due to economic motivation may go down. This could equalize the unobserved selectivity across the spectrum of the propensity to complete college and change the overall negative pattern we observed in this study to a flat one. Hence, an alternative interpretation of our results is that the observed pattern of negative selection is due to differential selectivity, with persons of low propensity to complete college more selective than persons of high propensity. While we cannot adjudicate between these two alternative interpretations, we have produced an important empirical finding: individuals who are less likely to obtain a college education benefit more from college.

REFERENCES

Altonji, Joseph G. and Thomas A. Dunn. 1996. "The Effects of Family Characteristics on the Return to Education." *The Review of Economics and Statistics* 78:692–704.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effect Using Instrumental Variables." *Journal of the American Statistical Association* 91(434):444–55.

Angrist, Joshua D. and Alan B. Krueger. 1999. "Empirical Strategies in Labor Economics." Pp. 1277–1366 in *Handbook of Labor Economics,* Vol. 3A, edited by O. Ashenfelter and D. Card. Amsterdam: Elsevier.

Armer, J. Michael. 1964. *Community and School Environments and College Plans of Public High School Seniors.* PhD dissertation, University of Wisconsin, Madison, WI.

Ashenfelter, Orley and Alan B. Krueger. 1994. "Estimates of the Economic Return to Schooling from a New Sample of Twins." *American Economic Review* 84:1157–73.

Ashenfelter, Orley and Cecilia Rouse. 1998. "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins." *Quarterly Journal of Economics* 113:253–84.

Averett, Susan L. and Mark L. Burton. 1996. "College Attendance and the College Wage Premium: Differences by Gender." *Economics of Education Review* 15:37–49.

Barrow, Lisa and Cecilia E. Rouse. 2005. "Do Returns to Schooling Differ by Race and Ethnicity?" *American Economic Review* 95:83–87.

Beattie, Irene. 2002. "Are all Adolescent Econometricians Created Equal? Racial, Class, and Gender Differences in College Enrollment." *Sociology of Education* 75:19–43.

Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* New York: Columbia University Press.

———. 1991. *A Treatise on the Family.* Cambridge, MA: Harvard University Press.

Becker, Sascha and Andrea Ichino. 2002. "Estimation of Average Treatment Effects Based on Propensity Scores." *Stata Journal* 2:358–77.

Bianchi, Suzanne. 1995. "Changing Economic Roles of Women and Men." Pp. 107–154 in *State of the Union: America in the 1990s; Volume One: Economic Trends,* edited by R. Farley. New York: Russell Sage Foundation.

Bjorklund, Anders and Robert Moffitt. 1987. "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models." *Review of Economics and Statistics* 69:42–49.

Blake, Judith. 1981. "Family Size and the Quality of Children." *Demography* 18:421–42.

Blau, Peter and Otis Dudley Duncan. 1967. *The American Occupational Structure.* New York: John Wiley &Sons.

Boudon, Raymond. 1974. *Education, Opportunity and Social Inequality.* New York: John Wiley & Sons.

Bourdieu, Pierre. 1977. "Cultural Reproduction and Social Reproduction." Pp. 487-510 in *Power and Ideology in Education,* edited by J. Karabel and A. H. Halsey. New York: Oxford University Press.

Bowles, Samuel and Herbert Gintis. 1976. *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life.* New York: Basic Books.

Brand, Jennie E. and Charles N. Halaby. 2006. "Regression and Matching Estimates of the Effects of Elite College Attendance on Educational and Career Achievement." *Social Science Research* 35:749–70.

Brand, Jennie E. and Yu Xie. 2007. "Who Benefits Most From College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." California Center for Population Research, University of California-Los Angeles CCPR 033-07.

Breen, Richard and John H. Goldthorpe. 1997. "Explaining Educational Differentials: Towards a Formal Rational Action Theory." *Rationality and Society* 9:275–305.

Brewer, Dominic, Eric Eide, and Ronald Ehrenberg. 1999. "Does it Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings." *Journal of Human Resources* 34(1):104–123.

Bryk, Anthony, Valerie Lee, and Peter Holland. 1993. *Catholic Schools and the Common Good.* Cambridge, MA: Harvard University Press.

Cameron, Stephen V. and James J. Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy* 109:455–99.

Card, David. 1995. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." Pp. 201–222 in *Aspects of Labour Market Behavior: Essays in Honour of John Vanderkamp,* edited by L. Christofides, E. Kenneth Grant, and R. Swidinsky. Toronto, Canada: University of Toronto Press.

———. 1999. "The Causal Effect of Education on Earnings." Pp. 1801–1863 in *Handbook of Labor Economics,* Vol. 3A, edited by O. Ashenfelter and D. Card. Amsterdam: Elsevier.

———. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica* 69:1127–60.

Carneiro, Pedro, Karsten Hansen, and James J. Heckman. 2003. "Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice." *International Economic Review* 44:361–422.

Carneiro, Pedro and James J. Heckman. 2002. "The Evidence on Credit Constraints in Post-Secondary Schooling." *Economic Journal* 112:705–34.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil. 2001. "Estimating the Return to Education When it Varies Among Individuals." Unpublished manuscript. Department of Economics, University College London.

———. 2007. "Estimating Marginal and Average Returns to Education." Unpublished manuscript. Department of Economics, University College London.

Cawley, John, Karen Conneely, James Heckman, and Edward Vytlacil. 1997. "Cognitive Ability, Wages, and Meritocracy." Pp. 179-192 in *Intelligence, Genes, and Success: Scientists Respond to the Bell Curve,* edited by B. Devlin, S. E. Feinberg, D. Resnick, and K. Roeder. New York: Springer.

Coleman, James A. 1988. "Social Capital and the Creation of Human Capital." *American Journal of Sociology* 94:S95–S120.

Dale, Stacy B. and Alan B. Krueger. 2002. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *Quarterly Journal of Economics* 117(4):1491–1527.

DiMaggio, Paul. 1982. "Cultural Capital and School Success: The Impact of Status Culture Participation on the Grades of U.S. High School Students." *American Sociological Review* 47:189–201.

DiPrete, Thomas A. and Henriette Engelhardt. 2004. “Estimating Causal Effects with Matching Methods in the Presence and Absence of Bias Cancellation.” *Sociological Methods and Research* 32:501–528.

Farley, Reynolds. 1996. *The New American Reality: Who We Are, How We Got Here, Where We Are Going.* New York: Russell Sage Foundation.

Fischer, Claude S. and Michael Hout. 2006. *Century of Difference: How America Changed in the Last One Hundred Years.* New York: Russell Sage Foundation.

Goldthorpe, John and Michelle Jackson. 2008. “Education-Based Meritocracy: The Barriers to its Realization.” Pp. 93–117 in *Social Class: How Does it Work?* edited by A. Lareau and D. Conley. New York: Russell Sage Foundation.

Griliches, Zvi. 1977. “Estimating the Returns to Schooling: Some Econometric Problems.” *Econometrica* 45:1–22.

Grubb, W. Norton. 1993. “The Varied Economic Returns to Post-Secondary Education: New Evidence from the Class of 1972.” *The Journal of Human Resources* 28:365–82.

Hauser, Robert M. 1973. “Socioeconomic Background and Differential Returns to Education.” Pp. 129-45 in *Does College Matter? Some Evidence on the Impacts of Higher Education,* edited by L. C. Solomon and P. J. Taubman. New York: Academic Press.

———. 2004. “Progress in Schooling.” Pp. 271–318 in *Social Inequality,* edited by K. Neckerman. New York: Russell Sage Foundation.

Hauser, Robert M. and Thomas N. Daymount. 1977. “Schooling, Ability, and Earnings: Cross-Sectional Findings 8 to 14 Years after High School Graduation.” *Sociology of Education* 50:182–206.

Hauser, Robert M., Shu-Ling Tsai, and William H. Sewell. 1983. “A Model of Stratification with Response Error in Social and Psychological Variables.” *Sociology of Education* 56:20–46.

Heckman, James J. 1978. “Dummy Endogenous Variables in a Simultaneous Equation System.” *Econometrica* 46(4):931–59.

———. 2005. “The Scientific Model of Causality.” *Sociological Methodology* 35:1–98.

———. 2007. “Schools, Skills and Synapses.” Lecture at Peking University, China. Retrieved 08-29-07 (http://jenni.uchicago.edu/papers/pku_2007/).

Heckman, James J. and Bo E. Honoré. 1990. “The Empirical Content of the Roy Model.” *Econometrica* 58:1121–49.

Heckman, James J., Justin L. Tobias, and Edward Vytlacil. 2001. “Four Parameters of Interest in the Evaluation of Social Programs.” *Southern Economic Journal* 68:210–223.

Heckman, James, Sergio Urzua, and Edward Vytlacil. 2006. “Understanding Instrumental Variables in Models with Essential Heterogeneity.” *The Review of Economics and Statistics* 88:389–432.

Hill, C. Russell and Frank P. Stafford. 1974. “Allocation of Time to Preschool Children and Educational Opportunity.” *Journal of Human Resources* 9:323–41.

Hoffer, Thomas, Andrew Greeley, and James Coleman. 1985. “Achievement Growth in Public and Catholic Schools.” *Sociology of Education* 58(2):74–97.

Holland, Paul W. 1986. “Statistics and Causal Inference” (with discussion). *Journal of American Statistical Association* 81:945–70.

Hout, Michael. 1984. “Status, Autonomy, and Training in Occupational Mobility.” *American Journal of Sociology* 89:379–409.

———. 1988. “More Universalism, Less Structural Mobility: The American Occupational Structure in the 1980s.” *American Journal of Sociology* 93:1358–1400.

Hout, Michael and Thomas A. DiPrete. 2006. “What We Have Learned: RC28’s Contributions to Knowledge about Social Stratification.” *Research in Social Stratification and Mobility* 24:1–20.

Imbens, Guido W. and Joshua Angrist. 1994. “Identification and Estimation of Local Average Treatment Effects.” *Econometrica* 62:467–76.

Jann, Ben, Jennie E. Brand, and Yu Xie. 2008. –hte– Stata module to perform heterogeneous treatment effect analysis, available from (http://ideas.repec.org/).

Jencks, Christopher, Marshall Smith, Henry Acland, Mary Jo Bane, David Cohen, Herbert Gintis, Barbara Heyns, and Stephan Michelson. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America.* New York: Basic Books.

Kane, Thomas and Cecilia Rouse. 1995. “Labor-Market Returns to Two- and Four-Year College.” *American Economic Review* 85(3):600–614.

Kao, Grace and Jennifer S. Thompson. 2003. “Racial and Ethnic Stratification in Educational Achievement and Attainment.” *Annual Review of Sociology* 29:417–42.

Lareau, Annette. 2003. *Unequal Childhoods.* Princeton, NJ: Princeton University Press.

Lucas, Samuel. 2001. “Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects.” *American Journal of Sociology* 106:1642–90.

MacLeod, Jay. 1989. *Ain’t No Makin’ It: Aspirations and Attainment in a Low-Income Neighborhood.* Boulder, CO: Westview Press.

Manski, Charles F. 1990. “Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?” Pp. 43-60 in *Studies of Supply and Demand in Higher Education,* edited by C. T. Clotfetter and M. Rothschild. Chicago, IL: University of Chicago Press.

———. 1995. *Identification Problems in the Social Sciences.* Boston, MA: Harvard University Press.

Manski, Charles F. and David A. Wise. 1983. *College Choice in America.* Cambridge, MA: Harvard University Press.

Mare, Robert D. 1981. “Change and Stability in Educational Stratification.” *American Sociological Review* 46:72–87.

McLanahan, Sara and Gary Sandefur. 1994. *Growing up with a Single Parent: What Hurts, What Helps.* Cambridge, MA: Harvard University Press.

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings.* New York: Columbia University Press.

Mincer, Jacob and Solomon Polachek. 1974. “Family Investments in Human Capital: Earnings of Women.” *Journal of Political Economy* 82:S76–S108.

Morgan, Stephen. 2001. “Counterfactuals, Causal Effect Heterogeneity, and the Catholic School Effect on Learning.” *Sociology of Education* 74(4):341–74.

———. 2005. *On the Edge of Commitment: Educational Attainment and Race in the United States.* Stanford, CA: Stanford University Press.

Morgan, Stephen L. and Jennifer J. Todd. 2008. “A Diagnostic Routine for the Detection of Consequential Heterogeneity of Causal Effects.” *Sociological Methodology* 38:231–81.

Morgan, Stephen and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research.* Cambridge, UK: Cambridge University Press.

Neyman, J. 1923. "On the Application of Probability Theory to Agricultural Experiments: Essay on Principles, Section 9." *Statistical Science* 5(4):465–80.

Olneck, Michael. 1979. "The Effects of Education." Pp. 159-90 in *Who Gets Ahead?* edited by C. Jencks. New York: Basic Books.

Orfield, Gary, S. E. Eaton, and E. R. Jones. 1996. *Dismantling Desegregation: The Quiet Reversal of Brown v. Board of Education.* New York: New Press.

Perna, Laura. 2005. "The Benefits of Higher Education: Sex, Racial/Ethnic, and Socioeconomic Group Differences." *Review of Higher Education* 29(1):23–52.

Raftery, Adrian and Michael Hout. 1993. "Maximally Maintained Inequality: Expansion, Reform, and Opportunity in Irish Education, 1921–75." *Sociology of Education* 66:41–62.

Rosenbaum, James. 2001. *Beyond College for All*. New York: Russell Sage Foundation Press.

Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70:41–55.

———. 1984. "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *Journal of the American Statistical Association* 79:516–24.

Rouse, Cecilia. 2004. "Low-Income Students and College Attendance: An Exploration of Income Expectations." *Social Science Quarterly* 85(5):1299–1317.

Roy, Andrew D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Paper* 3:135–46.

Rubin, Donald B. 1974. “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.” *Journal of Educational Psychology* 66:688–701.

———. 1997. “Estimating Causal Effects from Large Data Sets Using Propensity Scores.” *Annals of Internal Medicine* 127:757–63.

Sewell, William, Archibald O. Haller, and G. W. Ohlendorf. 1970. “Educational and Early Occupational Status Attainment Process: Replication and Revision.” *American Sociological Review* 35:1014–1027.

Sewell, William H., Archibald O. Haller, and Alejandro Portes. 1969. “The Educational and Early Occupational Attainment Process.” *American Sociological Review* 34:82–92.

Sewell, William H. and Robert M. Hauser. 1975. *Education, Occupation, and Earnings: Achievement in the Early Career.* New York: Academic Press.

Sheridan, Jennifer. 2001. *Occupational Attainment across the Life Course: Sources of Stability and Change in Three Occupational Characteristics.* PhD Dissertation, Department of Sociology, University of Wisconsin, Madison, WI.

Smith, Herbert L. and Brian Powell. 1990. “Great Expectations: Variations in Income Expectations among College Seniors.” *Sociology of Education* 63:194–207.

Sobel, Michael E. 2000. “Causal Inference in the Social Sciences.” *Journal of the American Statistical Association* 95:647–51.

Thomas, Scott. 2003. “Longer-Term Economic Effects of College Selectivity and Control.” *Research in Higher Education* 44(3):263–99.

Thomas, Scott and Liang Zhang. 2005. “Post-baccalaureate Wage Growth within Four Years of Graduation: The Effects of College Quality and College Major.” *Research in Higher Education* 46(4):437–59.

Train, Kenneth. 2003. *Discrete Choice Methods with Simulation.* Cambridge, UK: Cambridge University Press.

Tsai, Shu-Ling and Yu Xie. 2008. "Changes in Earnings Returns to Higher Education in Taiwan since the 1990s." *Population Review* 47:1–20.

U.S. Census Bureau. 2007. "Earnings Gap Highlighted by Census Bureau Data on Educational Attainment." *U.S. Census Bureau News,* March 15, 2007, Press Release.

Welch, Finis. 1973. "Black-White Differences in Returns to Schooling." *American Economic Review* 63:893–907.

Willis, Robert J. and Sherwin Rosen. 1979. "Education and Self-Selection." *Journal of Political Economy* 87:S7–36.

Winship, Christopher and Stephen L. Morgan. 1999. "The Estimation of Causal Effects from Observational Data." *Annual Review of Sociology* 25:659–707.

Xie, Yu. 2007. "Otis Dudley Duncan's Legacy: The Demographic Approach to Quantitative Reasoning in Social Science." *Research in Social Stratification and Mobility* 25(2):141-56.

Xie, Yu and Kimberlee A. Shauman. 1998. "Sex Differences in Research Productivity Revisited: New Evidence about an Old Puzzle." *American Sociological Review* 63:847–70.

Xie, Yu and Xiaogang Wu. 2005. "Market Premium, Social Process, and Statisticism." *American Sociological Review* 70:865–70.

Zhang, Liang. 2005. *Does Quality Matter? Benefits of Attending a High-Cost, Prestigious College.* New York: Rutledge.

FOOTNOTES

[1] There are noteworthy exceptions to this statement. For instance, Heckman (2007) emphasizes the importance of socioemotional skills—such as motivation, sociability, self-esteem, and health—for educational attainment.

[2] Unfortunately, individual-level variation in returns to higher education cannot be estimated (Holland 1986).

[3] A full interaction approach would also quickly exhaust degrees of freedom.

[4] With appropriate weights (Morgan and Todd 2008), we can obtain average treatment effects for specific populations, such as the average treatment effect, the treatment effect on the treated, and the treatment effect on the untreated (Brand and Halaby 2006). Other approaches to studying heterogeneous treatment effects include the switching regression model (Heckman 1978; Roy 1951; Willis and Rosen 1979), local average treatment effects (Angrist, Imbens, and Rubin 1996; Angrist and Krueger 1999; Imbens and Angrist 1994), and marginal treatment effects (Bjorklund and Moffitt 1987; Heckman et al. 2006).

[5] For example, studies that use switching regression models invoke a parametric assumption and require a strong theory. Studies that use instrumental variables (in the case of local average and marginal treatment effects) face the difficulty of finding a meaningful IV that affects treatment assignment directly but affects the outcome only indirectly through treatment.

[6] Economists, too, recognize the important role of family background (Ashenfelter and Rouse 1998), youthful expectations (Rouse 2004), ability (Cameron and Heckman 2001), and socioemotional skills (Heckman 2007) for educational attainment.

[7] Neighborhood and school characteristics, influenced by economic and racial segregation, also affect youths' academic achievement (MacLeod 1989; Orfield, Eaton, and Jones 1996).

[8] Still, a number of well-paid jobs do not require a college education, and some scholars question how many college graduates the U.S. labor force actually needs (Rosenbaum 2001).

[9] In other words, IV estimates can be interpreted as local average treatment effects, effects that pertain to units induced by the instrument. Instrumental variables estimates might also exceed OLS estimates if the instruments are correlated with ability (see Carneiro and Heckman 2002).

[10] Note that α in Equation 1 can be viewed as heterogeneous, that is, α_i, as it cannot be separated from the error term U_i. We write out the heterogeneous intercept explicitly in Equation 7.

[11] Rosenbaum and Rubin (1984) used propensity score strata, although they did not look for the variation of treatment effects as a function of the propensity score.

[12] To facilitate implementation of our method, we developed a new Stata module "hte" (Jann, Brand, and Xie 2008), which is available for public use.

[13] The NLSY79 was sponsored by the Bureau of Labor Statistics of the U.S. Department of Labor. The survey was conducted under contract with the Center for Human Resource Research and the National Opinion Research Center. Additional funding was provided by the National Institute of Child Health and Human Development and the National Institute on Drug Abuse.

[14] Since 1991, the WLS has been supported principally by the National Institute on Aging (AG-9775 and AG-21079), with additional support from the Vilas Estate Trust, National Science Foundation, Spencer Foundation, and Graduate School of the University of Wisconsin-Madison.

[15] Final analysis samples for the NLSY and WLS are generally more advantaged than full samples, although the differences are small.

[16] The log form measures proportional earnings differences rather than raw dollar differences. In high propensity score strata, there are likely larger raw differences than log differences in earnings between college and non-college completers than in low score strata. Although we

could invoke substantive arguments to favor proportional differences, we do not find this distinction to be a significant issue. We use the log form for comparability to prior studies and ease in interpretation.

[17] In the WLS, there is no hourly wage measure in 1975 comparable to 1993. We also analyzed earnings in respondents' early 20s for the NLSY and early 60s for the WLS. However, given higher levels of unemployment (and differential unemployment across strata) as a result of early career transitions in the NLSY and retirement transitions in the WLS, we limit our focus to results pertaining to prime working ages.

[18] We also ran all models where unemployed workers were maintained and assigned zero earnings. The substantive conclusions are analogous to those presented here. There is little difference in labor force participation rates across strata during prime earnings years for men. College-educated women, however, are more likely to exit the labor force during childbearing years in high than in low propensity score strata.

[19] Estimates for the control variables are available from the authors upon request.

[20] In the NLSY, we adjust for an indicator of marriage and the presence of children at age 25.

[21] We exclude 238 men and 74 women in the NLSY, and 16 men and 590 women in the WLS, who do not meet this requirement.

[22] To reduce the number of tables, we show results only for NLSY men. Results, available from the authors upon request, are comparable for NLSY women and for the WLS.

[23] The numbers of non-college graduates for NLSY and WLS men and women in our final stratum, and of college graduates in our first stratum for WLS women, are very small. This pattern is not surprising, as these individuals complete or do not complete college against the

expected odds. We thus collapsed strata in these instances and adjusted for the estimated propensity score.

[24] We expand the range of the y-axis for WLS results to accommodate the smaller (for men) and larger (for women) impact of college education relative to the NLSY.

[25] We used data from the National Longitudinal Study of the High School Class of 1972 (NLS) and examined effects of college completion on earnings by propensity score rank in 1986 (age 32). These data represent a third cohort, positioned between the WLS and NLSY cohorts. Results from the NLS lend further support for the theory of negative selection (Brand and Xie 2007). Due to space constraints, we do not present these findings.

[26] In Brand and Xie (2007), we explore this idea more explicitly by examining a ratio of the importance of monetary to nonmonetary factors in selecting a career across propensity score strata for college-educated men in the National Longitudinal Study of the High School Class of 1972. We found that men in low propensity score strata are more likely than men in high propensity strata to state that monetary factors are more important than nonmonetary factors. While all women state that nonmonetary factors are more important than monetary ones, low propensity women are more likely to state that monetary factors are important, yielding a propensity score pattern of results comparable to that of men.

[27] We use WLS data for this analysis because of the rich set of covariates at our disposal. We restrict analysis to men's earnings at age 35 for comparison with prior studies.

[28] Carneiro, Heckman, and Vytlacil (2001) do not accept the ignorability assumption. Their approach therefore differs from ours, offering another possible reason for the difference in results. We do not explore this possibility here.

[29] Results are available from the authors upon request.

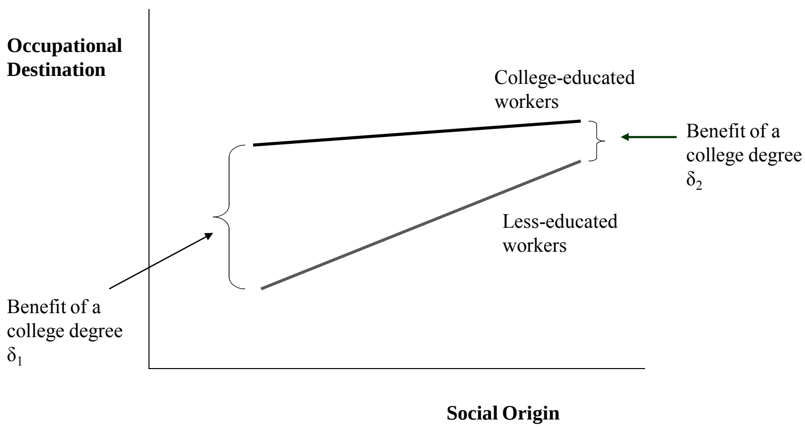

Figure 1. Hypothetical Model: Origin, Education, and Destination

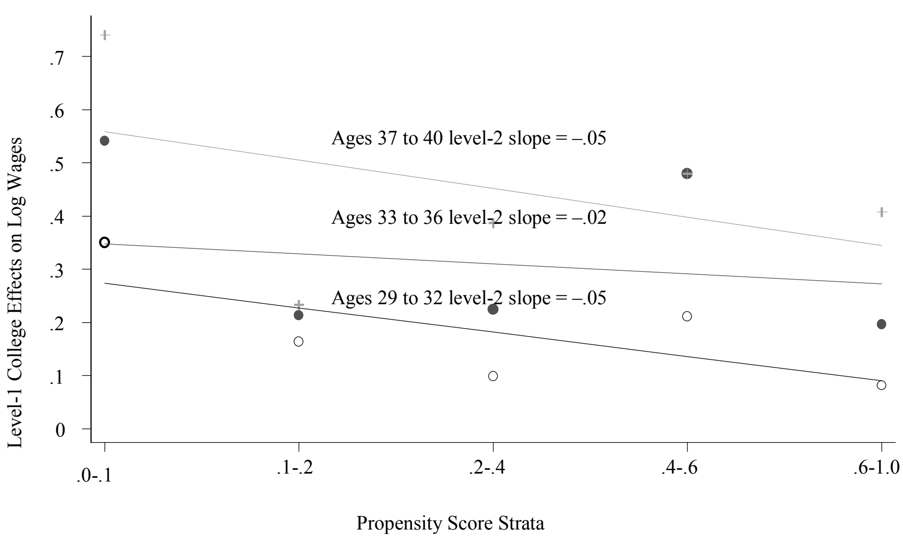

Figure 2. HLM of Economic Returns to College; NLSY Men

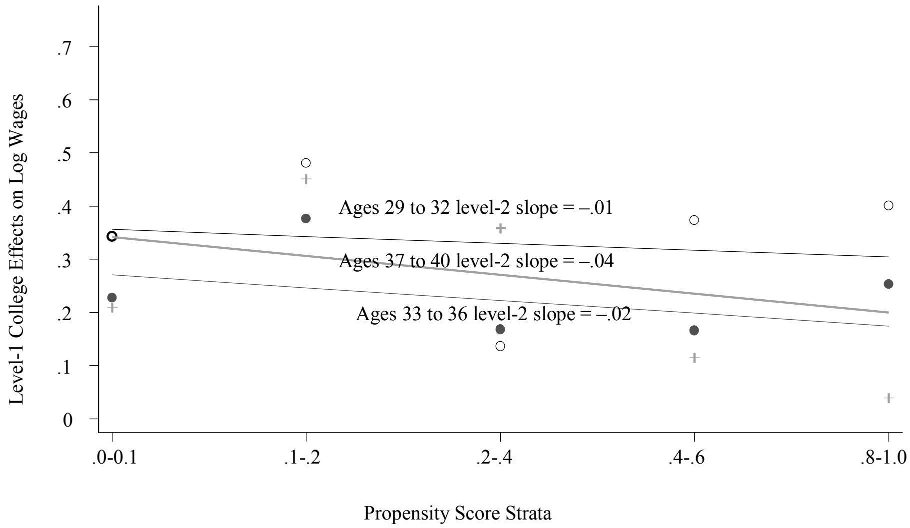

Figure 3. HLM of Economic Returns to College; NLSY Women

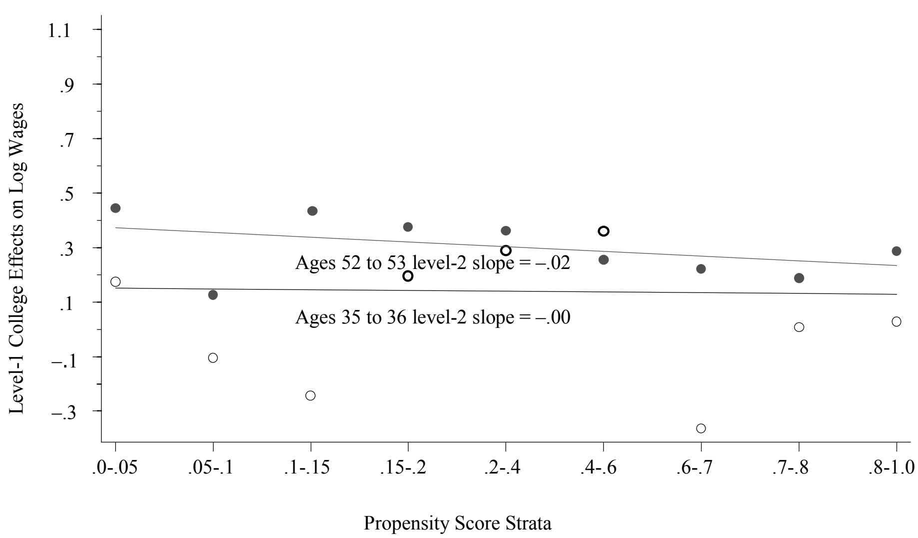

Figure 4. HLM of Economic Returns to College; WLS Men

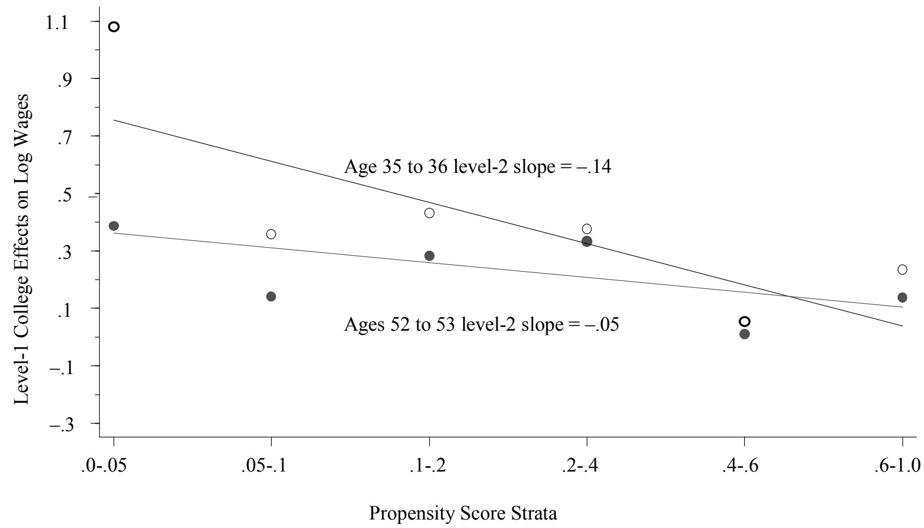

Figure 5. HLM Model of Economic Returns to College; WLS Women

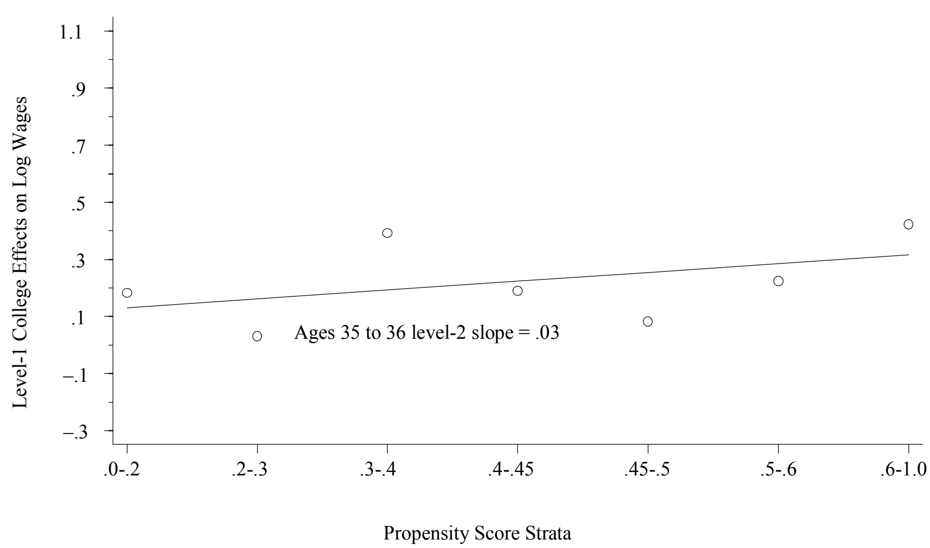

Figure 6. HLM of Economic Returns to College; WLS Men Small Covariate Set

Table 1. Descriptive Statistics of Precollege Covariates

	NLSY Means				WLS Means			
	Men (N = 1,265)		Women (N = 1,209)		Men (N = 3,690)		Women (N = 4,215)	
Variables	Non-College Graduate	College Graduate	Non-College Graduate	College Graduate	Non-College Graduate	College Graduate	Non-College Graduate	College Graduate
Race								
Black	.18	.07	.15	.07				
Hispanic	.07	.03	.07	.03				
Social Background								
Parents' income	17870	26538	18174	25991	5605	8123	5622	9262
Mother's education	11.26	13.32	11.18	13.37	10.15	11.56	9.94	12.02
Father's education	11.23	14.39	11.16	14.14	9.10	11.37	9.21	11.79
Intact family (0–1)	.72	.83	.67	.85	.90	.92	.90	.92
Number of siblings	3.29	2.34	3.40	2.45	3.45	2.61	3.51	2.40
Rural residence (0–1)	.25	.19	.24	.21	.22	.12	.20	.16
Urban res. / prox. to college	.77	.78	.75	.80	.42	.50	.50	.53
Jewish (0–1)	.00	.03	.00	.04	.00	.02	.00	.03
Ability and Academics								
Class rank					35.76	65.49	53.78	79.51
Mental ability (IQ)	–.09	.69	–.04	.64	97.03	111.75	98.67	112.00
College-prep (0–1)	.23	.59	.23	.49	.54	.91	.46	.89
Social-Psychological								
Teachers' encouragement					.35	.75	.36	.77
Parents' encouragement					.47	.91	.39	.90
Friends' college plans	.42	.79	.48	.81	.22	.66	.30	.76
Weighted Sample Proportion	.76	.24	.77	.23	.69	.31	.82	.18

Notes: Parents' income is measured as total net family income in 1979 dollars in the NLSY and in 1957 dollars in the WLS. Urban residency / proximity to college indicates whether a respondent lived in an SMSA in the NLSY and whether a respondent's high school was within 15 miles of a college or university in the WLS. Mental ability is measured with a scale of standardized residuals of the ASVAB in the NLSY and with the Henmon-Nelson IQ test in the WLS. College prep indicates whether a student was enrolled in a college-preparatory curriculum in the NLSY or whether a student completed the requirements for UW-Madison in the WLS.

Table 2. Effects of College Completion on Log Wages under the Assumption of Homogeneity

	Men	Women
NLSY		
1994 Wages	.180***	.276***
(ages 29 to 32)	(.047)	(.051)
1998 Wages	.296***	.188***
(ages 33 to 36)	(.054)	(.052)
2002 Wages	.410***	.216**
(ages 37 to 40)	(.069)	(.075)
WLS		
1975 Earnings	.124	.380**
(age 35)	(.067)	(.113)
1993 Wages	.302***	.225***
(age 53)	(.034)	(.038)

Note: Numbers in parentheses are standard errors. Treatment effects are conditional on the set of covariates for each data source described in Table 1. NLSY estimates further condition on age at baseline. NLSY estimates for women also condition on an indicator for married with children at age 25. All outcome variables are current hourly wages, except for WLS 1975 earnings, which are current yearly earnings. Unemployed workers are omitted.
* $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed tests).

Table 3. Mean Covariate Values by Propensity Score Strata: NLSY Men

	Propensity Score Strata									
	(.0–.1)		(.1–.2)		(.2–.4)		(.4–.6)		(.6–1.0)	
Variables	Non-Coll. Grad.	Coll. Grad.	Non-Coll. Grad.	Coll. Grad.	Non-Coll. Grad.	Coll. Grad.	Non-Coll. Grad.	Coll. Grad.	Non-Coll. Grad.	Coll. Grad.
Black	.37	.25	.21	.40	.23	.23	.19	.12	.15	.07
Hispanic	.18	.30	.12	.08	.12	.09	.12	.09	.07	.05
Parents' income	13381	12253	17614	18482	19324	18422	23062	21348	23469	34702
Mother's edu.	10.31	10.05	11.67	12.16	11.98	12.21	12.71	12.54	13.67	14.79
Father's edu.	10.17	9.95	11.79	10.72	12.08	12.53	13.33	13.97	15.11	16.30
Intact family	.63	.55	.63	.80	.74	.67	.85	.80	.85	.91
Num. of siblings	3.84	4.05	3.04	3.04	2.64	2.47	2.88	2.46	2.04	2.17
Rural resident	.21	.30	.26	.20	.21	.21	.19	.12	.11	.20
Availability coll.	.76	.70	.80	.84	.75	.77	.73	.77	.81	.78
Jewish	.00	.00	.00	.00	.01	.00	.00	.02	.04	.08
Mental ability	–.14	–.01	.31	.48	.62	.57	.79	.76	.90	1.05
College track	.17	.16	.32	.37	.41	.52	.57	.55	.83	.73
Friends' plans	.35	.55	.61	.52	.66	.74	.90	.85	.93	.93

Table 4. Frequency Counts per Propensity Score Stratum

NLSY						WLS					
Men			Women			Men			Women		
P-Score	$d = 0$	$d = 1$	P-Score	$d = 0$	$d = 1$	P-Score	$d = 0$	$d = 1$	P-Score	$d = 0$	$d = 1$
(.00, .10)	454	20	(.00, .05)	573	12	(.00, .05)	931	28	(.00, .05)	1367	27
(.10, .20)	135	25	(.05, .10)	181	17	(.05, .10)	418	33	(.05, .10)	441	38
(.20, .40)	130	43	(.10, .20)	156	28	(.10, .15)	255	25	(.10, .20)	367	67
(.40, .60)	52	65	(.20, .40)	147	47	(.15, .20)	155	45	(.20, .40)	391	172
(.60, 1.00)	27	76	(.40, .60)	37	48	(.20, .40)	386	149	(.40, .60)	204	185
			(.60, 1.00)	19	55	(.40, .60)	208	200	(.60, 1.00)	101	265
						(.60, .70)	72	122			
						(.70, .80)	46	173			
						(.80, 1.00)	48	380			

Table 5. Mean "Value of College" by Propensity Score Strata: WLS Men

	(.0–.05)	(.05–.1)	(.1–.15)	(.15–.2)	(.2–.4)	(.4–.6)	(.6–.7)	(.7–.8)	(.8–1.0)
Non-college grad.	45.12	54.55	63.29	66.09	74.82	82.52	84.90	88.73	85.36
College grad.	58.86	70.59	81.00	82.77	85.26	86.79	89.13	88.48	88.92

Note: The variable "value of college" is a weighted average of the scores in response to a series of 18 statements regarding perceptions of the value of going to college.

Table 6. Proportion of College Majors for College-Educated Men by Propensity Score Strata: WLS Men

	Propensity Score Strata								
College Major	(.0–.05)	(.05–.1)	(.1–.15)	(.15–.2)	(.2–.4)	(.4–.6)	(.6–.7)	(.7–.8)	(.8–1.0)
Physical science	.00	.06	.04	.02	.03	.05	.05	.04	.05
Math	.00	.06	.04	.02	.06	.09	.08	.04	.05
Biological science	.11	.03	.04	.02	.09	.09	.11	.07	.12
Engineering	.04	.06	.13	.12	.06	.14	.13	.23	.22
Pre-professional	.00	.00	.00	.00	.00	.01	.01	.01	.02
Computer science	.04	.00	.04	.00	.01	.02	.01	.01	.01
Business	.19	.27	.17	.19	.16	.15	.10	.11	.10
Social science	.15	.15	.25	.17	.18	.19	.10	.22	.21
Humanities	.04	.03	.00	.10	.13	.08	.13	.11	.10
Art and music	.11	.09	.04	.07	.04	.05	.05	.01	.05
Education	.22	.18	.21	.14	.15	.08	.07	.06	.05
Communications	.04	.03	.00	.02	.06	.01	.01	.04	.01
Agriculture	.04	.00	.00	.02	.01	.01	.02	.04	.01
Other	.04	.03	.04	.10	.02	.03	.03	.04	.02
Number	27	33	24	42	145	196	120	171	375

Table A1. Logit Models Predicting College Completion for the Generation of Estimated Propensity Scores

	NLSY		WLS	
Variables	Men	Women	Men	Women
Black	–.651*	–.208		
	(.268)	(.256)		
Hispanic	–.792*	–1.100**		
	(.335)	(.326)		
Parents' income	.000	.000	.000***	.000*
	(.000)	(.000)	(.000)	(.000)
Parents' income^2	.000	.000	.000*	.000
	(.000)	(.000)	(.000)	(.000)
Mother's education	–.363*	–.407**	–.334***	–.200
	(.162)	(.137)	(.092)	(.109)
Mother's education^2	.020**	.021***	.017***	.018*
	(.007)	(.006)	(.004)	(.005)
Father's education	.124**	.072*	.088***	.055**
	(.037)	(.036)	(.017)	(.017)
Intact family	.281	.507*	.073	.025
	(.236)	(.230)	(.169)	(.185)
Number of siblings	–.089	–.123*	–.018	–.067**
	(.051)	(.050)	(.021)	(.025)
Rural residence	.066	–.098	–.099	.667***
	(.251)	(.225)	(.143)	(.150)
Availability of college	–.419	–.065	–.092	.092
	(.239)	(.228)	(.099)	(.106)
Jewish	1.999		1.280*	1.387**
	(1.032)		(.609)	(.409)
Class rank			.028***	.027***
			(.002)	(.003)
Mental ability	1.984***	1.514***	.023***	.018***
	(.298)	(.239)	(.004)	(.005)
Mental ability^2	–.386	–.305		
	(.218)	(.176)		
College track	.603**	.514*	.618***	.693***
	(.196)	(.199)	(.132)	(.003)
Teachers' enc.			.438***	.565***
			(.107)	(.117)
Parents' enc.			.996***	1.323***
			(.130)	(.146)
Friends' plans	.992***	.645**	.946***	.723***
	(.209)	(.204)	(.101)	(.116)
LR Chi-Sq.	445.23	304.11	1765.50	1429.10
Prob. > Chi-Sq.	.00	.00	.00	.00
Sample size	1,265	1,203	3,690	4,215

Note: Numbers in parentheses are standard errors.
* $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed tests).

Table A2. Effects of College Completion on Log Wages by Propensity Score Strata: NLSY

	Level-1 Slopes					Level-2 Slopes
	Stratum 1	Stratum 2	Stratum 3	Stratum 4	Stratum 5	
Men						
1994 Earnings	.351**	.164	.099	.212*	.082	–.046
(ages 29 to 32)	(.123)	(.121)	(.107)	(.104)	(.082)	(.038)
1998 Earnings	.542***	.214	.225*	.479***	.197	–.019
(ages 33 to 36)	(.150)	(.108)	(.099)	(.120)	(.164)	(.046)
2002 Earnings	.740***	.234	.387**	.479*	.408*	–.053
(ages 37 to 40)	(.163)	(.191)	(.133)	(.188)	(.178)	(.055)
Women						
1994 Earnings	.343**	.480***	.137	.373	.401*	–.013
(ages 29 to 32)	(.119)	(.109)	(.107)	(.142)	(.168)	(.044)
1998 Earnings	.228*	.376**	.169	.167	.253	–.024
(ages 33 to 36)	(.103)	(.121)	(.114)	(.140)	(.265)	(.047)
2002 Earnings	.210	.452*	.359**	.116	.039	–.035
(ages 37 to 40)	(.160)	(.176)	(.134)	(.193)	(.368)	(.068)

Note: Numbers in parentheses are standard errors.
* $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed tests).

Table A3. Effects of College Completion on Log Wages by Propensity Score Strata: WLS

	Level-1 Slopes									
	Stratum 1	Stratum 2	Stratum 3	Stratum 4	Stratum 5	Stratum 6	Stratum 7	Stratum 8	Stratum 9	Level-2 Slopes
Men										
1975 Earnings	.175	–.104	–.243	.195	.288*	.360*	–.364	.007	.029	–.003
(age 35)	(.268)	(.269)	(.277)	(.225)	(.132)	(.146)	(.244)	(.261)	(.249)	(.033)
1993 Earnings	.444***	.126	.428*	.374**	.360***	.255***	.221	.188	.285*	–.017
(age 52)	(.118)	(.133)	(.168)	(.112)	(.080)	(.073)	(.127)	(.151)	(.132)	(.017)
Women										
1975 Earnings	1.079*	.358	.433	.376	.053	.235				–.143
(age 35)	(.462)	(.403)	(.321)	(.208)	(.223)	(.329)				(.087)
1993 Earnings	.385**	.141	.283**	.333***	.011	.138				–.052
(age 52)	(.147)	(.137)	(.106)	(.076)	(.087)	(.103)				(.029)
WLS Men, Small Set of Covariates										
1975 Earnings	.181	.031	.392**	.189	.082	.224	.423			.031
(age 35)	(.117)	(.099)	(.119)	(.204)	(.202)	(.219)	(.246)			(.032)

Note: Numbers in parentheses are standard errors.
* $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed tests).