The Consequences of Academic Match between Students and Colleges

Eleanor Wiske Dillon

Jeffrey Andrew Smith

Abstract: We consider the effects of student ability, college quality, and the interaction between the two on academic outcomes and earnings using data on two cohorts of college enrollees. Student ability and college quality strongly improve degree completion and earnings for all students. We find evidence of meaningful complementarity between student ability and college quality in degree completion at four years and long-term earnings, but not in degree completion at six years or STEM degree completion. This complementarity implies some tradeoff between equity and efficiency for policies that move lower ability students to higher quality colleges.

Eleanor Dillon is Visiting Assistant Professor at Amherst College. Jeffrey Smith is Paul T. Heyne Distinguished Chair in Economics and Richard Meese Chair in Applied Econometrics at the University of Wisconsin-Madison and an affiliate of the NBER, IZA, CESifo and HCEO. We thank Dan Black, Hoyt Bleakley, John Bound, David Deming, Sue Dynarski, Jose Galdo, Josh Goodman, Audrey Light, Lois Miller, Peter Mueser, Sarah Turner, Ophira Vishkin, and Martin West for helpful comments, along with seminar participants at Aarhus University, the 2014 Bergen-Stavanger Workshop on Labour Markets, Families and Children, the 2013 CESifo education group meetings, Bristol, Chicago (family economics), Cleveland Fed, Colorado, Cornell PAM, Dartmouth, Guelph, HEC, Institute for Fiscal Studies, MDRC, Michigan CIERS, Mannheim, National University of Singapore, Ohio State, the 2017 Ottawa-Carlton Graduate School in Economics Launch Conference, Penn GSE, Seton Hall Conference on College Match, Stanford CEPA, Toronto, UIC, Washington, Washington University in St. Louis, and Wilfred Laurier, and students in the Winter 2017 versions of Economics 622 and Public Policy 713 at the

doi:10.3368/jhr.55.3.0818-9702R1

This open access article is distributed under the terms of the CC-BY- NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0) and is freely available online at: http://jhr.uwpress.org

Downloaded from by guest on October 15, 2026. Copyright 2019

University of Michigan, particularly Ellen Stuart.

The online appendix can be found at http://jhr.uwpress.org

Disclosure statement: This research was supported by NSF SES 0915467.

JEL Codes: I23, I26, J3

Downloaded from by guest on October 15, 2026. Copyright 2019

I. Introduction

How students of varying ability sort into colleges of varying qualities has captured the attention not only of academic researchers studying higher education but also of the policy literature, the popular press, and the blogosphere. Until the last decade, the literature focused almost exclusively on outcomes for relatively low ability students at high-quality colleges, particularly students admitted under racial and ethnic preference policies at selective colleges. More recently, high ability students at relatively low-quality colleges have moved into the spotlight via the widely read studies by Bowen, Chingos, and McPherson (2009) and Roderick et al. (2008). Despite this recent focus on student-college sorting, the literature offers few credible estimates of the main object of interest: the interaction effects of student ability and college quality.

We ask whether the effects of attending a higher-quality college varies by student ability. Many studies conclude that increased college resources improve average student outcomes. These main effects of college quality should inform individual students' preferences over colleges, but cannot speak to the efficiency of various approaches to sorting students into colleges. In the presence of differential effects, resorting students via policy, even when respecting existing capacity constraints, has the potential to produce gains or losses in both efficiency and equity. In contrast, if the effects of college quality do not vary by student ability, then resorting can yield only equity gains or losses. The applied theory literature on college sorting, such as Rothschild and White (1995) and Sallee, Resch, and Courant (2008), posit complementarities between student ability and college quality; casual discussion about "fit" often presumes them. Under this assumption, positive assortative matching maximizes output. We look for evidence of these complementarities or of alternative relationships. Knowledge of the heterogeneous effects (if any) of college quality has clear value to students and parents

Downloaded from by guest on October 15, 2026. Copyright 2019

making decisions about college enrollment, and to researchers and policymakers concerned with the design, operation, and effects of state university systems with diversified quality portfolios.

Our analysis contributes to the small but growing literature on the effects of student-college sorting in a number of important ways. First, much of the literature frames the discussion of student-college sorting in terms of match, with relatively low ability students at relatively high quality colleges labeled "overmatched," and relatively high ability students at relatively low quality colleges labeled "undermatched." While we retain these practical labels for the two types of deviations from full assortative matching, we clarify the conceptual distinction between the main effects of college quality and student ability and their interaction when interpreting outcomes for these groups of students.[1]

Second, we present estimates from the 1979 and 1997 cohorts of the National Longitudinal Survey of Youth (hereinafter NLSY-79 and NLSY-97). By considering two different datasets, we can examine the stability of our estimates between cohorts of college students separated by over two decades, during which post-secondary education in the United States changed in important ways. We code our outcome and conditioning variables and design our analyses in the same way for the two datasets in order to make our analysis across cohorts as compelling as possible. Our analysis also implicitly replicates (in a broad sense) and extends the earlier analyses of the college quality main effect in the NLSY-79 presented in Black, Daniel, and Smith (2005).

Third, we examine a variety of outcome measures. With a couple of important recent exceptions discussed in more detail below, the earlier literature focuses primarily on degree completion. Bowen and Bok's (1998) early finding of no apparent impact on degree completion for overmatched students suggested to us that these students might find other ways to deal with

Downloaded from by guest on October 15, 2026. Copyright 2019

better-prepared colleagues and a high-pressure environment. For example, they might follow the increasingly common path of increased time to degree, as highlighted in Bound, Lovenheim, and Turner (2010). Or they might follow scholarship athletes at some colleges in taking easy courses, as suggested in journalistic exposés such as Steeg et al. (2008) and Thompson (2008). Or they might transfer to another school. Our examination of transfers, highlighted by Arcidiacono and Lovenheim (2016) as an understudied outcome, as well as of earnings in the years immediately following college enrollment, tells us more about the mechanisms through which college quality and ability affect educational and labor market outcomes. Our analysis of earnings up to 11 years after initial enrollment quantifies the medium-term labor market effects of college quality, ability and their interaction. For the NLSY-79 cohort, we present estimates for earnings up to 30 years following college start, the first longer-term estimates of college quality effects that vary with student ability.

Fourth, following Black and Smith (2006) and our earlier analysis of the determinants of college quality choice in Dillon and Smith (2017), we use composite indices as our measures of student ability and college quality. We expect these measures to embody substantially less measurement error than the single measures (for example, the student's own SAT score and the average SAT score of the entering class) commonly used in the literature and so to provide more accurate estimates.

Finally, we make an explicit case for our "selection on observed variables" identification strategy, which we view as credible in our context. Relative to the small set of existing studies of college sorting that use this identification strategy, we have a richer and more compelling set of relevant conditioning variables. Moreover, the literature provides strong evidence of the importance of factors likely conditionally unrelated to outcomes in driving college choice; these

Downloaded from by guest on October 15, 2026. Copyright 2019

factors provide exogenous variation in college quality. Unlike earlier papers, we show that our estimates stabilize as we add marginal sets of conditioning variables.

We find substantial deviations from full positive assortative matching in both cohorts, with both high ability students at low quality colleges and low ability students at high quality colleges. These deviations modestly decline from the NLSY-79 to the NLSY-97. Our examination of the effects of ability, college quality, and their interaction reveals substantively strong and statistically significant main effects of college quality and student ability on degree completion and earnings. These marginal effects of college quality are always positive, even for relatively low ability students, indicating that resorting policies like affirmative action likely pass a private benefit-cost test, on average, for the affected students. College quality matters relatively more, compared to ability, in the later cohort.

We find evidence of a causal effect of the interaction of quality and ability, but only for certain outcomes. We find clear evidence of complementarity between student ability and college quality in time to degree in both cohorts: abler students benefit relatively more from college quality for this outcome. We also find clear evidence of complementarity for long-run earnings outcomes in the NLSY-79 cohort; this pattern shows up in the point estimates for the later cohort as well. These patterns indicate an equity-efficiency tradeoff associated with policies that increase the enrollment of (relatively) less able students at high quality colleges.

Looking at mechanisms, we find some evidence that transfers tend to reduce gaps between student ability and college quality: undermatched students have a higher conditional probability of transferring up, though overmatched students do not have a similarly higher conditional probability of transferring down. Unlike Arcidiacono, Aucejo, and Hotz (2016), we

Downloaded from by guest on October 15, 2026. Copyright 2019

find no interaction effects related to STEM (= Science, Technology, Engineering and Math) degree completion.

II. Literature

Consideration of how the effects of college quality vary with student ability represents a natural extension of the broader literature on the causal effect of college quality on student outcomes.[2] If students benefit from greater college resources, then we expect academic undermatch to be mechanically costly for students, even without any role for student-college interactions, because it implies attending a lower-quality college. Overmatch is likewise mechanically beneficial. We review the subset of the literature on college quality that explicitly considers heterogeneous effects of college quality that vary with student ability.[3]

A few papers devote their full attention to the ways in which students with particular characteristics sort into colleges of different qualities. Our earlier work, Dillon and Smith (2017), uses the same NLSY-97 data we employ here and finds important roles for financial constraints in explaining college quality choices, as well as the in-state public college options available to the student. We show that more informed students and parents, proxied by variables such as parental education and the fraction of high school peers attending four-year colleges, act as though they believe that the main effect of college quality dominates any negative effects of academic overmatch. Smith, Pender, and Howell (2013) conduct a similar analysis using different data sets and a different definition of match. Reassuringly, they obtain similar findings. Lincove and Cortes (2019) use administrative data from Texas and find, among other interesting patterns, an important role for "social matching," which they define as attending a college with a high share of students in one's own racial or ethnic group. Hoxby and Avery (2012) and Hoxby

Downloaded from by guest on October 15, 2026. Copyright 2019

and Turner (2015) emphasize the role of information about college choices in driving college quality decisions for high achieving students from disadvantaged backgrounds while Griffith and Rothstein (2009) highlight the role of geographic distance for all students.

The studies most similar to ours examine heterogeneous college quality effects using "selection on observed variables" identification strategies to deal with non-random selection of students into colleges of varying qualities.[4] None of these studies find meaningful interaction effects, though several identify strong main effects of both student ability and college quality.[5] Mattern, Shaw, and Kobrin (2010) use data from a large number of colleges and a relatively limited set of observed characteristics, with student and college quality both measured using SAT scores and discretized into quartiles. They study how the effect of college quality on first year college GPA and persistence into the same college in the second year varies with student ability. The analysis in Chingos (2012) resembles our own in imposing capacity constraints but employs different data (the National Educational Longitudinal Study), cruder measures of student ability and college quality, a linear specification in college quality and student ability, and a less compelling set of conditioning variables. Black, Daniel, and Smith (2005) look at the simple interaction of student ability and college quality in the context of a parametric linear model of log wages applied to the NLSY-79 data. We compare our results to theirs in Section 6.7. Finally, Bowen, Chingos, and McPherson (2009) examine heterogeneous effects of college quality on college completion using impressive administrative data from various state university systems. They use relatively crude selectivity categories for universities, a modest conditioning variable set, and high school GPA and/or ACT/SAT scores as their measure of student ability.[6]

Light and Strayer (2000) also look at how the effect of college quality varies with student ability using the NLSY-79, but employ an empirical approach that differs substantially from

Downloaded from by guest on October 15, 2026. Copyright 2019

ours. They consider two sequential choices. The first choice, which they model as a multinomial probit, consists of either not attending four-year college (which combines entering the labor force and attending a two-year college) or going to college in one of four ordered quality quartiles. The second choice, which they model as a probit, concerns college completion. To address the potential for non-random selection on unobserved variables, they allow correlated errors between the two choices. Identification comes from conditioning on observed variables, from some (arguably implausible) exclusion restrictions, from some restrictions on the coefficients on the interactions of student ability and college quality, and from restrictions on the covariance matrix of the errors in the college choice model. Their estimates reveal substantively important heterogeneous effects that imply worse (average) outcomes for some lower ability students at higher quality colleges; we say more about why their qualitative findings differ from ours in Section V.G.

Another genre of studies focuses primarily on academic overmatch. Bowen and Bok (1998) find strong positive college quality effects on degree completion and earnings among black students attending the selective schools included in the "College and Beyond" data, which likely rule out negative net effects from overmatch within this group. Arcidiacono, Aucejo, and Spenner (2012) study Duke University, where the average African-American student starts out somewhat less prepared academically than other students, presumably due to affirmative action, but somewhat more likely to express a desire to major in the natural sciences, engineering, or economics. The authors find that black students at Duke differentially migrate away from these majors toward majors in the humanities or other social sciences and show that this pattern results almost entirely from their differential preparation. Put into our conceptual framework, their findings suggest that relatively overmatched black students at Duke adapt by changing majors

Downloaded from by guest on October 15, 2026. Copyright 2019

within an institutional context where essentially all students finish their degree. Arcidiacono, Aucejo, and Hotz (2016) continue this line of work by examining students at different University of California (UC) campuses entering school between 1995 and 1997, years prior to the ban on affirmative action in admissions in that state. The authors provide compelling evidence that overmatched minority students at UC schools who intend a STEM major have lower probabilities of completing any degree at a UC school and of graduating in STEM.[7]

A final group of papers uses "natural" experiments or discontinuities to isolate the experiences of students on the margin of admission to a higher quality college. Hoekstra (2009) began this strand of the literature by using a discontinuity in student SAT scores (conditional on high school GPA) to examine the effects of admission to, and enrollment in, a state flagship university. He finds large positive effects of flagship acceptance and attendance on earnings 10-15 years after high school completion for men but surprisingly small effects for women. Kurlaender and Grodsky (2013) exploit an unusual event in 2004 in which the admissions offices of UC schools initially offered their marginal acceptances deferred admission due to budget issues but were later able to offer them regular admission. They find that marginal students accumulate fewer credits compared to similar students at lower-ranked UC schools, but have higher graduation probabilities. Thus, they find evidence of overmatch effects on the "intensive" course-taking margin but the college quality effect dominates for degree completion. The estimates in this group of papers correspond to quite narrowly defined populations of marginal students. Moreover, these findings shed only indirect light on match. The nature and extent of overmatch for students at a college quality margin depends on the particular definition of match employed as well as the quality of the student's next best alternative and the homogeneity of

Downloaded from by guest on October 15, 2026. Copyright 2019

student ability within each college. Still, at the very least, the evidence from these papers stands at odds with large, negative consequences of what we might call local overmatch.[8]

Overall, we view the literature as providing strong evidence of causal effects of college quality and student ability on academic and labor market outcomes. In contrast, most but not all of the literature finds little in the way of interaction effects, other than on intermediate outcomes such as transfer and major choice.

III. Data

A. NLSY

We use the NLSY-79 data, which includes Americans who were ages 14 to 22 on January 1, 1979, and the NLSY-97 data, which includes Americans who were 12 to 16 years old as of December 31, 1996. In both cohorts, participants were interviewed annually starting in 1979 and 1997, respectively, and continuing through their college years. They have been interviewed biannually since 1994 and 2011, respectively. We include the representative samples from each survey along with the supplemental samples of blacks and Hispanics.[9] Most respondents in the NLSY-79 sample graduated from high school and made their college choice between 1975 and 1983, while the NLSY-97 sample did the same between 1998 and 2002. We focus on students who enroll in a four-year college by age 21 (39% of high school graduates and GED holders in the NLSY-79 sample and 36% in the NLSY-97 sample).

One of the strengths of the NLSY data for both cohorts lies in the rich set of individual and family covariates it provides.[10] Using the restricted access geocode data provides additional information on the identities of colleges attended and allows the use of contextual information based on the respondent's residential location. The following sections describe our ability and

Downloaded from by guest on October 15, 2026. Copyright 2019

college quality measures, as well as our outcome variables; Appendix Tables A1, A2, along with Online Appendix Tables A1 and A2 describe the construction of our analysis sample and summarize our conditioning variables.

B. Ability

We follow Dillon and Smith (2017) in designing our measures of student ability and college quality for the NLSY-97 sample and construct comparable measures for the earlier NLSY-79 cohort. Our measures of student ability draw on the Armed Services Vocational Aptitude Battery (ASVAB). In the 1997 cohort, 86% of respondents who started at a four-year college completed the test; the corresponding number for the 1979 cohort is 93%. We use the method developed by Altonji, Bharadwaj, and Lange (2012) to construct comparable measures of eight of the ASVAB test components common to the two cohorts, adjusting for the transition between 1979 and 1997 from pen-and-paper to computer adaptive testing and for the varying ages at which the respondents took the test. We do not use the scores on the purely vocational components.

We then construct the first two principal components of these eight section scores. Our primary measure of ability, which we call ASVAB1, equals each respondent's percentile of the first principal component within the sample distribution of college-bound respondents in their NLSY cohort.[11] As shown in Online Appendix Table A3, the first principal component explains 68% and 66% of the total variance in test scores across the eight sections for the 1979 and 1997 cohorts, respectively. In both cohorts, the first component places the highest weight on academic subjects such as arithmetic reasoning and paragraph comprehension. Not surprisingly giving the loadings, the correlation between ASVAB1 and the respondent's SAT (or rescaled ACT) score equals 0.79 in NLSY-79 and 0.80 in NLSY-97.

Downloaded from by guest on October 15, 2026. Copyright 2019

The second component of the ASVAB scores, which we call ASVAB2, explains a further 10-11% of the variance. As in Cawley, Heckman, and Vytlacil (2001), who perform a similar analysis using the NLSY-79 data, the second component places the most weight on the two timed sections of the test: numerical operations and coding speed. We include ASVAB2 as an additional control variable in our multivariate analyses. To capture further dimensions of ability we also include high school GPA and SAT scores along with multiple proxies for non-cognitive or socio-emotional skills.[12]

C. College Quality

We construct a one-dimensional index of college quality by combining measures related to selectivity and college resources. The available data limit us to using measures of inputs as proxies for quality, but we note that Hoxby (2015)'s value-added estimates correlate with one important component of our index, namely college selectivity. In particular, our index combines the mean SAT or ACT score of entering students, the percent of applicants rejected, the average salary of all faculty engaged in instruction, and the undergraduate faculty-student ratio. We combine data from the U.S. Department of Education's Integrated Post-Secondary Education Data System (IPEDS) and U.S. News and World Report, using data from 1992 for the NLSY-79 and data from 2008 for the NLSY-97.[13]

Following Black and Smith (2004, 2006) and Black, Daniel and Smith (2005) we estimate the principal components across these four measures of quality. We use the eigenvector of the first principal component (reported in Online Appendix Table A4) to calculate a weighted average of the proxies available for each college, then calculate percentiles of this average across colleges.[14] We interpret our index as an estimate of latent college quality, which we view as continuous and one-dimensional. Combining multiple proxies into a single index measures latent

Downloaded from by guest on October 15, 2026. Copyright 2019

quality with less error than a single proxy. Our index reveals remarkable stability in college quality between our two cohorts; weighted by full-time undergraduates the correlation equals 0.86.[15] Our measure does not capture differences in the quality that different students experience within the same university due to, for example, quality differences across fields of study or participation in honors programs. Our index also speaks only indirectly to absolute differences in college quality. In practice, the four individual quality proxies underlying our index increase modestly but steadily with the index for the bottom 90% of four-year colleges but more steeply for the top 10% of colleges. Figure 1 documents this pattern for expenditures per student, a measure we do not include in our index but which strongly correlates with its components. This very general scaling issue with latent indices, emphasized in this literature by Bastedo and Flaster (2014), also applies to the two other most common proxies for college quality in the literature: the mean SAT score of the entering class and the Barron's selectivity categories.

We analyze the quality of the first four-year college a student attends rather than the last, as in Black, Daniel, and Smith (2005) and some other studies. Our concern with how students initially sort into colleges motivates this choice, as treating the quality of the first college as the choice variable allows us to treat subsequent transfer and completion choices, some of which may result from initial college choices, as intermediate outcomes on the way to earnings effects.

D. Sorting Among Colleges by Student Ability

To assess the degree of sorting across colleges by student ability we consider the joint distributions of the student ability and college quality measures just described. As we weight the quality percentile by student body size, a college in the n^{th} quality percentile is the college that a student in the n^{th} ability percentile would attend under positive assortative academic matching.[16] We label substantial deviations from this type of sorting as overmatch and undermatch. One

Downloaded from by guest on October 15, 2026. Copyright 2019

appealing feature of our measures is the possibility of achieving full assortative matching without violating institutional enrollment constraints. The measures employed in other studies in the literature, such as Roderick et al. (2008), Bowen, Chingos, and McPherson (2009), and Smith, Pender, and Howell (2013) lack this feature.[17]

Table 1 gives the joint distributions of student ability and college quality for the 1979 and 1997 cohorts, with both variables discretized into quartiles. In both cohorts, students differentially concentrate along the diagonal indicating positive assortative matching. The four diagonal cells account for 34.3% of students in 1979 and 37.1% in 1997, rather than the 25% implied by random sorting. The three upper right cells, corresponding to low ability students at high quality colleges, account for 11.3% of students in 1979 and 10.2% in 1997, while the three lower left cells, corresponding to high ability students at low quality colleges, account for 14.9% in 1979 and 13.2% in 1997.[18] Thus, we find substantial departures from full assortative matching in both cohorts. Viewed longitudinally, our data (perhaps surprisingly given the recent policy focus on match) reveal only a small, though meaningful, increase in the correlation between student ability and college quality.[19]

E. Outcomes

We examine five educational outcomes: graduation within four or six years of starting, obtaining a STEM degree, and transfer to a higher or lower quality college. The NLSY-79 survey did not begin asking questions on the specific college attended until even the younger sample respondents were several years into college, making it difficult to follow transfer behavior in the earlier cohort.[20] We therefore calculate the transfer outcomes only for the NLSY-97 cohort. We define graduation as completing a four-year degree at any college. We define STEM degree completion based on the last reported major(s) prior to graduation and code majors as STEM or

Downloaded from by guest on October 15, 2026. Copyright 2019

non-STEM using the (uncontroversial) system in Arcidiacono, Aucejo, and Hotz (2016).[21] Some restless students transfer more than once; we code our transfer variable based on the first observed transfer and only count transfers that involve a change of at least five percentiles (up or down) in our college quality index. Transfers from any four-year college to any two-year college always count as a transfer down.

On the labor market side, in the spirit of the program evaluation literature, we examine the level of real (2010$) earnings (rather than the log) in all years from the start of college, without conditioning on degree completion.[22] We look relative to the start of college rather than the end because we want to capture the opportunity cost of college, because college quality may (heterogeneously) affect the probability of working while in school, and because we want to capture effects on time to degree. We average earnings in two-year intervals, using observed earnings for one year when the value for the other equals zero or missing. We omit two-year intervals if the respondent did not report non-zero earnings in either year. This pooling reduces variance at minimal cost to sample size and temporal fineness, as nearly everyone in our sample of four-year college attendees works in almost every year. Comparisons with the information on job spells suggest that a non-trivial fraction of the zeros represent measurement error.

Table 2 summarizes these outcomes for our sample. The 1997 cohort has a higher graduation rate, consistent with the pattern documented in Archibald, Feldman, and McHenry (2015) that U.S. graduation rates reached a nadir for students starting college in the mid-1980s and have recovered since then. The probability of graduating with a STEM degree has fallen a bit, from 15% to 13%. Consistent with the somewhat earlier cohort studied by Goldrick-Rab (2006) and with the Texans in Andrews, Li, and Lovenheim (2014), we find a great deal of

Downloaded from by guest on October 15, 2026. Copyright 2019

transfer behavior among the NLSY-97 students; 27% transfer at least once. Earnings increase both over the life cycle within cohorts and between the 1979 and 1997 cohorts.

IV. Econometric Framework

To determine whether the data provide evidence of important interactions between ability and college quality, we want to look flexibly at the conditional relationship between these two variables and the outcomes of interest. Several econometric frameworks comport with this goal. This section describes two: our preferred estimator based on a flexible polynomial approximation and an alternative estimator that uses indicators for bins of the discretized joint distribution of ability and college quality.

For binary outcomes, we estimate probit models. In our preferred specification, we estimate the conditional probability function as:

$$(1) \qquad \Pr(Y_i = 1 | A_i, Q_i, X_i) = \Phi(\beta_0 + \beta_P(A_i, Q_i) + \beta_X X_i)$$

In Equation 1, Y denotes the binary outcome of interest, A denotes student ability, Q denotes college quality, $\beta_p(A_i, Q_i)$ denotes a flexible polynomial of ability and quality, and X denotes a vector of other conditioning variables. For earnings, we estimate a parametric linear regression model using the same specification by ordinary least squares. In both cases,

$$(2) \qquad \beta_P(A_i, Q_i) = \beta_1 A_i + \beta_2 A_i^2 + \beta_3 Q_i + \beta_4 Q_i^2 + \beta_5 A_i Q_i + \beta_6 A_i^2 Q_i + \beta_7 A_i Q_i^2$$

We chose this specification after a rigorous round of statistical testing.[23] The polynomial in ability and quality becomes non-parametric once we promise to include additional higher-order terms as our sample size increases. Equation 1 then becomes a partially linear model in which we non-parametrically estimate the effects of ability and quality while conditioning parametrically on the other variables.

Downloaded from by guest on October 15, 2026. Copyright 2019

Polynomial approximations sometimes mislead, especially around the edges of the data. As a sensitivity check, we implement a different semi-parametric framework that includes indicators for combinations of college quality quartile and student ability quartile. We include indicators for 15 of the 16 possible combinations, with ability and quality both in the lowest quartile serving as the omitted category. This approach avoids the oft-observed instability of higher order polynomials away from the center of the data but cannot capture any within-quartile variation. In practice, the two estimators tell the same substantive story; see Online Appendix Table A5 for the estimates from the second approach for a subset of our outcomes.

The NLSY surveys include several measures of respondents' cognitive skills. For ease of interpretation we interact only ASVAB1 with college quality. We therefore want to concentrate the effects of any common component of ability in this variable. To accomplish this, we orthogonalize the SAT score and GPA variables against ASVAB1 prior to including them in the multivariate analyses.[24]

V. Identification

This section considers the case for interpreting our estimates as causal. We argue that we have a sufficiently rich conditioning set that the remaining variation in college quality that serves to identify our effects is uncorrelated with the error term in the outcome equation. Our identification strategy captures causal effects if two conditions hold. First, we need the observed covariates included in our model to capture, either directly or as proxies, all the factors that affect both the college quality choice and the outcomes we study. Second, in order to avoid identification via functional form, we need some conditionally exogenous variation in college quality choices. Put differently, we need instrumental variables to exist, even though we do not

Downloaded from by guest on October 15, 2026. Copyright 2019

observe them, as they produce the conditional variation in college quality we implicitly use in our estimation.

We divide our conditioning variables into four sets, each of which proxy for one broad factor affecting educational choices: pre-college skill, student demographic and family characteristics, neighborhood characteristics as of the first survey, and other social factors. We list these variables and describe their construction in detail in Appendix Table A2. Our preferred specification includes the first three sets of covariates; the fourth set provides a test, described below, for our identification strategy. We never condition on whether the student remains enrolled in college each year or whether they have completed a degree, which we view as intermediate outcomes.

We make the case that our conditioning set suffices to solve the problem of non-random selection into colleges in two ways. First, we can think about whether our conditioning set contains those things (or compelling proxies for those things) that existing theory and empirical evidence deem important. Much recent literature, e.g. Heckman and Kautz (2012), emphasizes the importance of non-cognitive skills for educational and labor market outcomes. The broader literature, including our own earlier study, illustrates the need to condition on family resources, both intellectual and financial. More money makes many things about college easier, including longer time-to-degree, more frequent visits home, and not having to work during school, and so affects outcomes; it also surely affects the college quality choice. Parental education will correlate with their knowledge of the college choice process and of how to succeed at college in both the institutional and academic senses. Parental education also likely correlates with taste for education and otherwise unobserved features of the student's childhood environment that affect both outcomes and college choice. Becker and Lewis (1973) highlight a quality-quantity tradeoff

Downloaded from by guest on October 15, 2026. Copyright 2019

for parents, so number of siblings may reflect both resources and preferences. We expect that our county education variable will both help with measurement error in the direct parental resource variables and proxy for primary and secondary school quality as well as peer pressure and expectations.

The second way to think about our covariate set asks whether the marginal covariates make any difference to the estimates. In Heckman and Navarro's (2004) framework there exist multiple unobserved factors on which we need to condition. As we increase the number of proxy variables in our conditioning set, the amount of selection bias in our estimates should decrease to zero, so long as we keep adding proxies for all factors. Turning this around, if we observe that the estimates stabilize as we increase the richness of the conditioning set, this suggests we are doing a good job of proxying for the unobserved factors, unless there exists an additional unobserved factor uncorrelated with all of our covariates. Oster (2017) cautions that a finding of coefficient stability means little if the newly added variables do not capture any conditional variation in the dependent variable. We perform such analyses by adding sets of related variables to the conditioning set in an order that reflects our prior about their importance for solving the problem of non-random selection into colleges of different qualities.

The literature suggests that plenty of exogenous variation exists in college quality choices conditional on our observed covariates. First, differences in state college quality mix, admission policies, and pricing strategies provide plausibly exogenous variation in the budget sets facing students and their parents. Niu and Tienda (2010) and Daugherty, Martorell, and McFarlin (2014) estimate large effects of guaranteed admission to in-state public colleges through the Texas Top 10% rule on college choices for eligible students. Cohodes and Goodman (2014) estimate similarly large effects of in-state-specific scholarships in Massachusetts, lowering the

Downloaded from by guest on October 15, 2026. Copyright 2019

average quality of college attended for eligible students who chose in-state colleges over more selective outside options. Second, distances to colleges of various qualities provide variation in the costs of attendance as in Card (1995) and Currie and Moretti (2003). Third, what normally represents a sad feature of this literature, namely the consistent finding that many students, parents, and high school guidance counselors have little idea about how to choose a college, provides support for our identification strategy. Hoxby and Avery (2012) and Hoxby and Turner (2013) show the difference a small amount of reliable information can make for many students. Similarly, the literature provides many examples of small behavioral economics tricks having non-trivial effects on college choices. Pallais (2015) finds that you can change college choices by changing the number of colleges to which students can send their ACT scores for free and Bettinger et al. (2012) find that having H & R Block help with the federal financial aid form can have real effects on college-going. Scott-Clayton (2012) reviews the literature showing that students and parents often know very little about the likely costs and benefits of college. Finally, both the descriptive and ethnographic literature, such as Roderick et al. (2008), and the quantitative literature on sorting, such as Lincove and Cortes (2019), suggest that many students choose among colleges for reasons unrelated to academic quality such as the football team or the presence of high school friends. While the value students place on these non-academic traits may well *unconditionally* correlate with outcomes, we expect that variation among students in the nature and extent of the trade-off between academic and non-academic aspects of colleges that they face to produce useful, and random, *conditional* variation in college quality choice.

Two main issues motivate concerns about selection bias in estimates of the college quality main effect: First, students, their parents, and college admissions officers may have access to information on student ability that we, the researchers, do not. To the extent that those

Downloaded from by guest on October 15, 2026. Copyright 2019

unobserved factors affect admissions, we would expect an upward bias in the estimated effect of college quality because it proxies in part for higher unobserved student ability or ambition (and we might expect this bias primarily at the upper end of the college quality distribution, where "holistic" rather than rule-based admissions dominate). Second, we might worry about measurement error in college quality, as in Black and Smith (2006). Though our use of a quality index based on multiple proxies addresses this issue some measurement error surely remains, which we expect will push the estimated effect toward zero. Of course, we have no basis for arguing that these two biases cancel out in practice.

Now think about the interaction of college quality and student ability. If we overstate the effect of a high quality college for all students, then overmatched students will look better than they should relative to other students of the same ability. Similarly, undermatched students will look relatively worse than they should. Thus, upward bias in the estimated effect of college quality should lead us to understate the effects of overmatch and to overstate the effects of undermatch, and so potentially to overstate the degree of complementarity between student ability and college quality. Measurement error in ability and/or in college quality, in contrast, should attenuate our estimates of the effects of both overmatch and undermatch; indeed, Griliches and Ringstad (1970) highlight the particularly pernicious effects of measurement error in non-linear contexts such as interactions.

VI. Effects of College Quality and Ability on College Outcomes and Earnings

A. Graduation Rates

Table 3 presents our estimates of Equation 1 for degree completion within four and six years for both cohorts. The first three rows of estimates report the mean marginal effect of ability

Downloaded from by guest on October 15, 2026. Copyright 2019

percentile at different points in the college quality distribution, constructed from our estimates of the flexible polynomial of ability and quality percentiles. The second three rows report the mean marginal effect of college quality at different points in the ability distribution. We scale both *A* and *Q* to lie in [0, 1].

Our first key finding consists of substantively meaningful and statistically significant main effects of both college quality and student ability on graduation within six years of starting college. For example, for a student in the NLSY-97 cohort attending a college at the 25th percentile of the quality distribution each 10 percentile increase in a student's ability increases her (conditional) probability of graduating within six years by 3.10 (0.1 x 100 x 0.310) percentage points. Along the same lines, for a college starter in the NLSY-97 cohort of median ability, increasing the quality of the first college attended by 10 percentiles increases the probability of graduating within six years by 3.60 percentage points. The magnitudes of our estimates fit comfortably within the existing literature. While strong main effects of student ability and college quality emerge for both cohorts, the relative importance of college quality increases noticeably for the later cohort.

If ability and quality have only independent effects, then we would expect a uniform effect of college quality across students of different ability levels. Alternatively, the effect of quality could vary with student ability. For example, college quality might increase degree completion probabilities more for students lower in the ability distribution. Our second key finding is that the effect of college quality on graduation varies very little with student ability. Likewise, the effect of student ability is quite steady at different points in the college quality distribution. Figure 2 plots the average predicted six-year graduation probability at each percentile of college quality. It shows that at the 25th, 50th and 75th percentiles of the ability

Downloaded from by guest on October 15, 2026. Copyright 2019

distribution the probability of graduating within six years increases almost linearly in college quality percentile for the NLSY-79 cohort. The NLSY-97 cohort shows some evidence of complementarity: the probability of graduating within six years increases more slowly with college quality above the 60th percentile of colleges for students in the 25th percentile of ability.

We can quantify the evidence for heterogenous effects in our college completion results in two ways. First, because our model nests a model with only main effects of college quality and ability, we can test the restriction that all coefficients on the interactions of ability and college quality jointly equal zero. The p-values from these tests appear in the third row in the bottom panel of Table 3. The p-values of 0.554 for the NLSY-97 cohort and 0.363 for the NLSY-79 cohort indicate that the restrictions implicit in the main-effects-only model cause little trouble for the data. Alternative statistical tests consider the null of equal average derivatives with respect to student ability and with respect to college quality, at the 25th, 50th and 75th percentiles of each. P-values for these tests appear in the last two rows of Table 3; they comport with our interpretation based on the magnitudes above.

Second, we can look to Tables 4A and 4B, which compare the observed completion rate with the completion rate implied by our model in a counterfactual world of full positive assortative matching. We obtain this value by predicting degree completion for every observation with their college quality percentile recoded to equal their ability percentile. Based on our model, we find that degree completion rises less than one percentage point, moving from 59.9% to 60.4% for the younger cohort and from 49.7% to 50.3% for the older one. The negative effect of moving lower ability students away from high quality colleges to their matched quality level almost entirely cancels out the positive effect of moving higher ability students out of low

Downloaded from by guest on October 15, 2026. Copyright 2019

quality colleges. Chingos (2012) performs a similar calculation and also finds virtually no effect of resorting students.

The net effect of moving to full assortative matching masks large improvements in outcomes from moving some students to higher-quality colleges. The last columns of Tables 4A and 4B present a second counterfactual in which we ignore capacity constraints (and general equilibrium considerations) and assume that all students attend a college in the 90th percentile of college quality. Our model predicts that moving all students to a high-quality college would increase degree attainment within six years by 9.1 percentage points to 58.8% for the older cohort and by 10.6 percentage points to 70.5% for the younger cohort. This increase might seem smaller than expected, but student characteristics matter as well, and differ strongly between students presently at the 90th percentile of college quality and those further down the distribution.

Now consider the results for graduation in four years, a standard that remains normative but has become increasingly aspirational for many students, as documented in e.g. Bound, Lovenheim, and Turner (2010). The average derivative estimates for completion in four years resemble those for completion in six years in sign, all positive, but differ in showing a clear and substantively meaningful pattern of complementarity between student ability and college quality, one that gets stronger for the NLSY-97 cohort. We can reject the nulls of equal mean derivatives for the 1997 cohort. In parallel (and substantively similar) results from linear probability models presented in Online Appendix Table A6, we strongly reject the null of zero coefficients on the interaction terms and the nulls of equal average derivatives with respect to ability and college quality for the 1997 cohort (and come closer to rejection of these nulls in the 1979 cohort). Overall, we find serious evidence of complementarities between ability and quality for on-time degree completion. In addition, the difference between the two cohorts in the relative importance

Downloaded from by guest on October 15, 2026. Copyright 2019

of student ability and college quality, with college quality playing a smaller role for the older cohort, remains striking.

B. Intermediate Educational Outcomes

To shed light on the mechanisms underlying our findings on completion rates, we consider the effects of student ability and college quality on some intermediate college outcomes. Students might react to large gaps between their ability and the quality of their college by changing their major, as in Arcidiacono, Aucejo, and Hotz (2016), who argue that some overmatched students switch from STEM majors to other, less challenging majors. A change of major could delay graduation, thereby lowering four-year, but not six-year, completion rates. Table 5 presents our estimates for STEM degree completion. For both cohorts, we find substantively and statistically significant effects of student ability for all levels of college quality, with only modest differences between the effects at the 25th, 50th, and 75th percentiles. We find substantively small effects, not statistically different from zero, of college quality at all levels of student ability. The p-values for the null of zero interaction equal 0.791 and 0.708 for the NLSY-97 and NLSY-79 cohorts. Our estimates predict that full positive assortative matching would change STEM degree completion by less than 0.6 percentage points in either cohort.

Students may react to learning they have made a poor initial choice by transferring to another school. Again, this mid-course adjustment could delay graduation beyond four years. We find some evidence consistent with ability-quality complementarity when looking at transfer behavior in the NLSY-97 cohort. The third column of estimates in Table 5 corresponds to Equation 1 with transfer up as the dependent variable, while the fourth column of estimates corresponds to transfer down.[25] Increasing a student's ability percentile by 10 percentage points raises the probability that she will transfer to a higher quality college by 1.4 percentage points if

Downloaded from by guest on October 15, 2026. Copyright 2019

she starts at a 25th percentile college. In contrast, student ability has virtually no effect on the probability of transferring to a higher quality college if the student starts at a 75th percentile college. The second three rows show an expected pattern: increasing the quality of the first college a student attends lowers the probability that she will transfer to an even higher quality college, with a larger effect for students higher in the ability distribution. The pattern of derivatives with respect to ability reflects students preferentially transferring toward assortative matching, while the pattern of derivatives with respect to quality is partly mechanical.

We see the reverse patterns when considering transfers to lower quality colleges. More able students transfer to a lower quality college less often, though only modestly and imprecisely so. Increasing the quality of the first college attended raises the probability that students will transfer down; these effects differ statistically from zero but not much by the level of student ability. Taken together, these transfer results provide some support for ability-quality complementarity along with a strong (and again partly mechanical) main effect of college quality.

We cannot reject the null of only ability and quality main effects on transfer behavior: As shown at the bottom of Table 5, the p-values equal 0.384 and 0.328 for transferring to a higher and lower quality college, respectively. In Table 4B, we predict that full positive assortative matching would modestly decrease transfers to a higher quality college and modestly increase transfers to lower quality colleges. This counterfactual sorting mechanism substantially decreases the transfer probability for students who strongly undermatch at their first college, but such students constitute only a small fraction of the total. Since transfers often delay graduation, these moves have real costs for students (and often for the taxpayer) in terms of more time in school and less time in the labor force.

Downloaded from by guest on October 15, 2026. Copyright 2019

C. Earnings

Tables 6A and 6B present our estimates for the effects of ability and college quality on average annual earnings during and after college. In years 2-3, both college quality and student ability have generally negative effects on annual average earnings in both cohorts. In the NLSY-79, 2-3 years after starting college a student at the 50th percentile of ability earns $208 less per year for each 10 percentile point increase in the quality of first college attended. Students at lower quality colleges are more likely to have left without a degree and begun working full time 2-3 years after starting college. Higher quality colleges may also require greater effort to keep up with course work, limiting the time students have to work while still in college. Finally, near the top of the college quality distribution, marginal increases in college quality may give students access to more financial aid and reduce their need to work during college. The negative relationship between ability and earnings likely reflects a similar short-run tradeoff between current earnings and investment in skill accumulation, as well as access to more merit-based financial aid.

At 10-11 years after students begin college these patterns have completely reversed: both college quality and student ability strongly raise average annual earnings. For a student of median ability in the NLSY-97 cohort, each 10 percentile point increase in the quality of the first college is associated with an additional $1,480 of annual earnings. In keeping with the completion rate estimates, we find much larger ability effects in the NLSY-79 cohort and smaller college quality effects. For example, at the median of college quality, a 10 percentile point increase in student ability increases earnings at 10-11 years by $915 in the NLSY-79, compared to just $417 for the later cohort. While our average derivative estimates have large standard errors, we find persuasive evidence that college quality increases future earnings throughout the ability distribution.[26]

Downloaded from by guest on October 15, 2026. Copyright 2019

As with degree completion in four years, the estimates for earnings 10-11 years after college start suggest a substantively important complementarity between college quality and student ability, particularly for the NLSY-97 cohort. The average derivative of earnings with respect to student ability has a much larger value, around $127 per percentile point for the NLSY-97, for students at the 75th percentile of college quality than for those at the 25th percentile or at the median. Similarly, the average derivative of earnings with respect to college quality increases with student ability, from about $104 per percentile point at the 25th percentile of ability to about $186 per percentile point at the 75th percentile of ability. Still, we cannot reject the nulls of no interaction effects (as well as other nulls involving much larger interaction effects) or of equal derivatives.

Table 4 shows that moving to full positive assortative matching in the NLSY-97 cohort would increase mean earnings by about $1,328 10-11 years after beginning college. The corresponding change for the NLSY-79 cohort is $694. We do not emphasize these point estimates given how far this scenario projects outside the data, and given the likely importance of equilibrium effects of uncertain direction and magnitude (including the fact that resorting the students would change the quality of all of the colleges as we measure it). The data provide no evidence of any harmful individual effects for low ability students who attend higher quality colleges. At the same time, our estimates suggest that policies that place some students with lower ability at top colleges do impose some efficiency costs due to the complementarity between student ability and college quality.

The NLSY-79 cohort, now into its fifties, allows us to examine earnings outcomes for several decades after college start. These results appear in the last columns of Table 6A and in Figure 3. The data provide large, positive, and generally statistically significant estimates of the

Downloaded from by guest on October 15, 2026. Copyright 2019

main effects of student ability and college quality at all durations from 10-31 years after college start. Even at long durations we can clearly rule out negative average effects of overmatch. We can also rule out standard models in which college quality simply signals student ability, as employer learning would surely have overwhelmed college quality effects over the horizons we consider.[27]

In general, the average derivatives get larger as the time elapsed from college start increases, sometimes quite substantially so. The average derivative with respect to quality for students at the 75th percentile of ability increases from $724 for a 10 percentile-point increase in quality at 10-11 years to $2,527 at 20-21 years to $4,048 at 30-31 years. As our data embody only one cohort, we have no way of separating these increases into components due to age and period effects. Additionally, a pattern consistent with complementarity between student ability and college quality, which was fairly weak in the estimates of earnings 10-11 years after college start for the NLSY-79 cohort, appears quite strongly in the longer-term follow-up estimates. As shown in the bottom row of the table, for earnings 20-21 and 30-31 years out we can reject the null of equal average derivatives with respect to college quality at different levels of student ability. Finally, we remind the reader of our imprecise estimates, and of the gentle decline in the sample size as individuals gradually attrit from the panel.[28]

D. Subgroups

We consider subgroups defined by sex and by parental education, where we partition the latter into "low" and "high" subgroups based on whether at least one parent attended college. We interpret parental education as a proxy for several things, including tastes for college (and college quality) and family resources. In the NLSY-79 cohort nearly half of college entrants have parents with no more than a high school education, but by the NLSY-97 cohort only a quarter of

Downloaded from by guest on October 15, 2026. Copyright 2019

college entrants have parents with no college education. We lack the sample size to usefully examine finer categories. Similarly, though of great substantive interest, we lack the sample sizes to present meaningful subgroup estimates for black and Hispanic students.

Tables 7 and 8 report the effects of student ability and college quality on earnings 10-11 years after starting college separately by subgroups.[29] To limit the demands on our relatively small sample, we estimate these effects by interacting the ability-quality polynomials with subgroup indicators and continuing to estimate pooled coefficients on the other covariates. The main finding from the pooled estimates, positive effects of both student ability and college quality at all levels, generally holds for all sub-groups, with more volatility in point estimates and predictably larger standard errors. In both cohorts, the pattern of complementarity between student ability and college quality is most apparent for the children of more educated parents, though the average effect of student ability is larger for the children of less educated parents. We mostly find larger effects of ability for female students than for male students, much larger in the NLSY-79 cohort. However, the two main differences between the two cohorts hold for both men and women: student ability plays a relatively larger role in determining earnings in the earlier NLSY-79 cohort and the younger cohort displays more evidence of complementarity between ability and college quality in degree completion (the latter in results not shown).

E. Identification

We now consider some evidence regarding our identification strategy. Tables 9 and 10 present estimates based on increasingly rich sets of conditioning variables for our two most important outcomes: degree attainment within six years and earnings in years 10-11 after starting college. The lower rows of each table indicate the set of included conditioning variables; the categories

Downloaded from by guest on October 15, 2026. Copyright 2019

correspond to those in Appendix Table A2. The estimates in column (4) of each table correspond to those in Tables 3 and 6.

Overall, the tables reveal a substantial amount of movement in the coefficients when moving from column (1) to column (2) by adding additional measures of ability and socio-emotional skills, and when moving from column (2) to column (3), which corresponds to adding demographics and family characteristics.[30] We see somewhat less movement (how much less varies across outcomes and across derivatives) when we add neighborhood characteristics in column (4). Finally, and in parallel to the similar analysis in the Black, Daniel, and Smith (2005) study of college quality, moving from column (4) to column (5) changes the estimates very little. With each transition, including the last, the r-squared values meaningfully increase. These findings support a causal interpretation of our estimates or, at least, suggest that any remaining biases would not overturn our qualitative conclusions.

F. Comparing the NLSY-79 and NLSY-97 Results

Our big picture stories apply to both cohorts: large amounts of overmatch and undermatch in the unconditional joint distribution of student ability and college quality, substantively and statistically significant positive main effects of student ability and college quality for college completion and earnings in the medium and long terms, some evidence of heterogeneous college quality effects in time to degree, and little or no evidence of heterogeneous effects for degree completion or for STEM degree completion. This stability surprised us somewhat. In this section, we briefly remark on two specific differences in the results: (1) the modest but not trivial shift towards positive assortative matching between cohorts; and (2) the relatively smaller role of college quality in determining outcomes for the NLSY-79.

Downloaded from by guest on October 15, 2026. Copyright 2019

As we noted in Section 3.4, the changing sorting patterns between the cohorts comports with some other evidence in the literature. Following Hoxby (2009), we suspect that it results from ongoing reductions in the cost that students (especially high ability students) face in obtaining information about admissions criteria, real (as opposed to posted) prices, and optimal strategy. Reductions in transportation and communication costs likely also play a role.

One reason why college quality may have smaller measured effects on outcomes in the NLSY-79 cohort is that colleges varied less in our input-based measures of quality for this earlier cohort. As shown in Figure 1, each percentile increase in our college quality index corresponds to a somewhat smaller change in expenditure per student in 1992 than in 2008. We may also suffer from attenuation bias due to measurement error in matching students to their first college attended. As noted in Section 3.3, the NLSY-79 did not ask students the name of the college(s) they attended until 1984, part way through most students' college enrollment. We use the first reported college, which we suspect functions as an excellent proxy for first college attended, but memory lapses and transfers could yield some mis-measurement.

G. Comparisons with Other Studies

Two earlier published papers, Light and Strayer (2000) [hereinafter LS] and Black, Daniel and Smith (2005) [hereinafter BDS] estimate college quality effects interacted with ability using the NLSY-79 data. The estimates from the LS probit model of degree attainment (they do not look at earnings) that appear in the two right-most columns of their Table 8 correspond most closely to our own.[31] These estimates assume, implicitly, selection on observed variables; i.e. they frame them as a sensitivity analysis in which they shut down their apparatus for dealing with selection on unobserved variables. Unlike us, they find that ability does not always increase degree attainment across all college quality quartiles, nor does college quality monotonically increase

Downloaded from by guest on October 15, 2026. Copyright 2019

degree completion across all ability quartiles.[32] In the latter case their estimates support the view that overmatch could make some students worse off. Several differences between the LS setup and our own strike us as potential candidates to account for the difference in findings: (1) they treat transfers as dropouts; (2) they restrict some of the interactions between ability quartile and college quality quartile to have zero coefficients on a priori grounds; (3) they condition on variables that we think plausibly endogenous, namely living at home, receipt of financial aid, and actual tuition paid; and (4) their remaining covariate set represents (in essence) a modest subset of our own, which raises the possibility of residual selection bias not present in our analysis.[33]

The analysis in BDS, not surprisingly given the authorial overlap, differs less from our own. Qualitatively, we reach similar conclusions. While they examine interactions only for their log wage outcome, and not for degree completion, they find strong main effects of both student ability and college quality for both degree attainment and log wages. Their Appendix Table 7 presents estimates from a parametric linear model with hourly wages as the dependent variable and a rich covariate set similar to our own (other than in its inclusion of years of schooling), along with main effects in ability and college quality and interactions between college quality and their versions of ASVAB1 and ASVAB2. They offer separate estimates for men and women; for both groups and both interactions they obtain estimated coefficients near zero and far from statistical significance.[34]

As far as we know, just three other studies consider the persistence of the earnings effects of college quality over a very long interval after college start; none of these studies consider how the effects vary with student ability. Turner (2002) considers earnings of men in the Panel Study of Income Dynamics (PSID) who completed a BA by 1975. She finds large and growing effects of college quality from 1975 to 1992 and provides suggestive evidence that the increases

Downloaded from by guest on October 15, 2026. Copyright 2019

primarily represent period effects rather than lifecycle effects. The average SAT score of entering students proxies for college quality. As discussed in her note 13, the lack of a compelling proxy for student ability in the PSID hampers a causal interpretation of her estimates, which rely on a "selection on observed variables" identification strategy. Figure 1 of BDS (2005) shows impacts on log wages for the NLSY-79 for calendar years 1987 to 1998, about 15 years after starting college. They find persistent, stable, and substantively and statistically meaningful effects of college quality for both men and women using a selection on observed variables identification strategy (and conditioning on years of schooling).

Dale and Krueger (2014) link social security earnings data to two cohorts of the "College and Beyond" dataset that includes students entering a non-random sample of relatively high quality colleges in the fall of 1976 and 1989. They present estimates of log earnings impacts through 2007, or 31 years after college enrollment for the older cohort and 18 years after for the younger one.[35] Dale and Krueger (2014) present estimates using two identification strategies; the first assumes "selection on observed variables" in the context of a covariate set that includes standardized test scores as well as a rich (but not as rich as ours) array of other relevant variables. Their second identification strategy attempts to deal with any remaining selection on unobserved variables by conditioning on the average SAT score of the schools to which each student applied; they call this their "self-revelation" model.[36] The first identification strategy yields persistent and sizeable effects of college quality on later earnings for all groups; in marked contrast, the "self-revelation model" estimates reveal such impacts only for black and Hispanic students and those from disadvantaged family backgrounds.

Finally, in a broad sense our results coincide with the descriptive analysis presented in Chetty et al. (2017), who find using US income tax data linked across generations that within

Downloaded from by guest on October 15, 2026. Copyright 2019

college quality tiers, average child income varies only very modestly with parental income. This pattern conflicts with strong negative effects of overmatch as students with relatively low income parents within a quality tier will also have relatively lower average ability as we define it.

VII. Summary and Conclusions

This paper examines the effects of college quality and student ability on academic and labor market outcomes for two cohorts of college goers using the NLSY-79 and NLSY-97 datasets. We adopt a "selection on observed variables" identification strategy in both cases and do our best to ensure comparability in coding and conditioning. In both cohorts, we find strong evidence that college quality and student ability increase the probability of degree completion and later earnings. The relative importance of college quality increases in the later cohort; this finding parallels that of Castex and Dechter (2014), who document a similar change in the relative importance of ability and years of schooling as determinants of wages in the two NLSY cohorts.

For college students and their families, our most salient conclusion is that increasing college quality increases graduation rates and earnings at all points in the ability distribution.[37] At the margin, all students will benefit in expectation from attending higher quality colleges. In Dillon and Smith (2017) we find that well-informed and well-resourced students seek to attend higher quality colleges, even if they will be overmatched at these institutions. Our current work validates this unconditional pursuit of college quality. Policies targeted at increasing the representation of certain groups of students at high-quality colleges will (on average) benefit the targeted students, but may do little to improve overall outcomes if the total number of seats at high-quality colleges remains unchanged.

The simple and compelling applied theory models in Rothschild and White (1995) and

Downloaded from by guest on October 15, 2026. Copyright 2019

Sallee, Resch, and Courant (2008) posit complementarities between student ability and college quality, which can justify the observed long-term increase in positive assortative matching described in e.g. Hoxby (2009). We find modest but substantively important support for these theories. We can reject uniform effects of college quality across the ability distribution for some but not all of the outcomes we examine. The effects of college quality vary with student ability in the production of graduation in four years (but not six), particularly for the NLSY-97 cohort, for transfers, and for long-term (but not immediately post-college) earnings. The interaction effects we find do not overwhelm the uniformly positive main effects. These results suggest modest efficiency gains from better sorting the strongest students into the top colleges, and some efficiency costs to policies that weaken this sorting.

Less prepared students appear to adjust to the demands of higher quality colleges by slowing their studies, leading to smaller gains in four-year graduation rates than their higher-ability classmates but similar gains in six-year graduation rates. In contrast to some other papers, we do not find similar evidence of adjustment by shifting out of STEM majors. One interpretation of our findings on earnings is that the networking and recruiting benefits of attending a higher-quality college benefit all students in their first job, but higher-ability students build more successfully on these early gains as they move through their careers, perhaps because of greater skill acquisition in college.

We conclude with five caveats. First, we interpret our estimates in partial rather than general equilibrium terms; as such, they apply primarily to moving around small numbers of students. Second, we pay for the plausibility of the conditional independence assumption with modest sample sizes. Particularly in the context of high variance outcomes such as earnings, some of the patterns we find show up more clearly in the estimates than in the statistical tests.

Downloaded from by guest on October 15, 2026. Copyright 2019

Third, measurement error remains a concern in multiple senses. While using multiple proxies for student ability and college quality reduces measurement error, it does not eliminate it. Less trivially, we know that individual students at larger colleges experience very different parts of what their institutions have to offer; for example, faculty research and teaching quality may differ across departments. Thus, even if our quality measure does well at capturing the average quality of a college, it may embody substantial measurement error at the student level at which our analysis operates.

Fourth, we consider only undergraduates at four-year colleges. Our results may not generalize to contexts, such as law schools, that provide students with fewer dimensions on which to respond to an environment that proves too challenging or not challenging enough. In law school, for example, students cannot easily change majors or take fewer courses. For this reason, overmatch might have very different overall effects in these contexts than in ours.

Finally, this paper only examines interactions between one characteristic of students, namely their academic ability, and college quality. As noted in Smith (2008), sorting based on other student and college characteristics represent an important omission from most of the literature. Perhaps the most obvious concerns matching on social class or socio-economic status, or what an economist might prefer to call (at the cost of losing some nuance in interpretation) family resources. Recent scholarly books such as Armstrong and Hamilton's (2012) *Paying for the Party* and Radford's (2013) *Top Student, Top School?* highlight this dimension of sorting, as does Tom Wolfe (2004) in his novel of college life entitled *I Am Charlotte Simmons*. Because other student characteristics like social class correlate with academic ability, they represent a potentially confounding treatment in our context.

Downloaded from by guest on October 15, 2026. Copyright 2019

References

Alon, Sigal, and Marta Tienda. 2005. "Assessing the `Mismatch' Hypothesis: Differences in College Graduation Rates by Institutional Selectivity" *Sociology of Education* 78(4): 294-315.

Altonji, Joseph, Prashant Bharadwaj, and Fabian Lange. 2012. "Changes in the Characteristics of American Youth: Implications for Adult Outcomes." *Journal of Labor Economics* 30(4): 783-828.

Andrews, Rodney, Jing Li and Michael Lovenheim. 2014. "Heterogeneous Paths through College: Detailed Patterns and Relationships with Graduation and Earnings." *Economics of Education Review* 42: 93-108.

Archibald, Robert, David Feldman, and Peter McHenry. 2015. "A Quality-Preserving Increase in Four-Year College Attendance." *Journal of Human Capital* 9(3): 265-97.

Arcidiacono, Peter, Esteban Aucejo and V. Joseph Hotz. 2016. "University Differences in the Graduation of Minorities in STEM Fields: Evidence from California." *American Economic Review* 106(3): 525-62.

Arcidiacono, Peter, Esteban Aucejo, Andrew Hussey, and Kenneth Spenner. 2013. "Racial Segregation Patterns in Selective Universities." *Journal of Law and Economics* 56(4): 1039-60.

Arcidiacono, Peter, Esteban Aucejo and Ken Spenner. 2012. "What Happens After Enrollment? An Analysis of the Time Path of Racial Differences in GPA and Major Choice." *IZA Journal of Labor Economics*, 1:5.

Arcidiacono, Peter, and Michael Lovenheim. 2016. "Affirmative Action and the Quality-Fit Trade-Off." *Journal of Economic Literature* 54(1): 3-51.

Armstrong, Elizabeth, and Laura Hamilton. 2013. *Paying for the Party: How College Maintains Inequality.* Cambridge, MA: Harvard University Press.

Aucejo, Esteban, and Jonathan James. 2019. "Catching Up to Girls: Understanding the Gender Imbalance in Educational Attainment Within Race." *Journal of Applied Econometrics* 34(4): 502-25.

Bastedo, Michael, and Allyson Flaster. 2014. "Conceptual and Methodological Problems in Research on College Undermatch." *Educational Researcher* 42(3): 92-99.

Becker, Gary, and H. Gregg Lewis. 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy* 81(2, part 2): S279-S288.

Bettinger, Eric, Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. 2012. "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment." *Quarterly Journal of Economics* 127(3): 1-38.

Black, Dan, Kermit Daniel, and Jeffrey Smith. 2005. "College Quality and Wages in the United States." *German Economic Review* 6(3): 415-43.

Black, Dan, and Jeffrey Smith. 2004. "How Robust is the Evidence on the Effects of College Quality? Evidence from Matching." *Journal of Econometrics* 121(1): 99-124.

Downloaded from by guest on October 15, 2026. Copyright 2019

Black, Dan, and Jeffrey Smith. 2006. "Evaluating the Returns to College Quality with Multiple Proxies for Quality." *Journal of Labor Economics* 24(3): 701-28.

Bound, John, Michael Lovenheim and Sarah Turner. 2010. "Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources." *American Economic Journal: Applied Economics* 2(3): 129-57.

Bowen, William, and Derek Bok. 1998. *The Shape of the River: Long-Term Consequences of Considering Race in College and University Admissions*. Princeton, NJ: Princeton University Press.

Bowen, William, Matthew Chingos, and Michael McPherson. 2009. *Crossing the Finish Line: Completing College at America's Public Universities*. Princeton, NJ: Princeton University Press.

Cadena, Brian, and Ben Keys. 2015. "Human Capital and the Lifetime Costs of Impatience," *American Economic Journal: Economic Policy*. 7(3): 126-53.

Card, David. 1995. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." In *Aspects of Labour Market Behavior: Essays in Honor of John Vanderkamp,* ed. Louis Christophides, E. Kenneth Grant and Robert Swidinsky, 201-22. Toronto: University of Toronto Press.

Castex, Gonzalo, and Evgenia Kogan Dechter. 2014 "The Changing Roles of Education and Ability in Wage Determination." *Journal of Labor Economics* 32(4): 685-710.

Cawley, John, James Heckman, and Edward Vytlacil. 2001. "Three Observations on Wages and Measured Cognitive Ability." *Labor Economics* 8(4): 419-42.

Chetty, Raj, John Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2017. "Mobility Report Cards: The Role of College in Intergenerational Mobility." NBER Working Paper No. 23618.

Chingos, Matthew. 2012. "Graduation Rates at America's Universities: What We Know and What We Need to Know." In *Getting to Graduation: The Completion Agenda in Higher Education*, ed. Andrew Kelly and Mark Schneider, 48-70. Baltimore, MD: Johns Hopkins University Press,

Cohodes, Sarah, and Joshua Goodman. 2014. "Merit Aid, College Quality, and College Completion: Massachusetts' Adams Scholarship as an In-kind Subsidy." *American Economic Journal: Applied Economics* 6(4): 251-85.

Cunha, Flavio. 2009. "Crossing the Finish Line: A Review Article." *Journal of Human Capital* 3(4): 354-78.

Cunha, Jesse, and Trey Miller. 2014. "Measuring Value-Added in Higher Education: Possibilities and Limitations in the Use of Administrative Data." *Economics of Education Review* 42: 64-77.

Currie, Janet, and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." *Quarterly Journal of Economics* 118(4): 1495-1532.

Dale, Stacy Berg, and Alan Krueger. 2002. "Estimating the Payoff of Attending a More Selective College: An Application of Selection on Observables and Unobservables." *Quarterly Journal of Economics* 107(4): 1491-1527.

Downloaded from by guest on October 15, 2026. Copyright 2019

Dale, Stacy Berg, and Alan Krueger. 2014. "Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data." *Journal of Human Resources* 49(2): 323-58.

Daugherty, Lindsay, Paco Martorell, and Isaac McFarlin. 2014. "Percent Plans, Automatic Admissions, and College Outcomes." *IZA Journal of Labor Economics* 3:10.

Dillon, Eleanor, and Jeffrey Smith. 2017. "Determinants of the Match between Student Ability and College Quality." *Journal of Labor Economics* 35(1): 45-66.

Goldrick-Rab, Sara. 2006. "Following Their Every Move: An Investigation of Social-Class Differences in College Pathways." *Sociology of Education* 79(1): 61-79.

Goodman, Joshua, Michael Hurwitz, Jonathan Smith. 2017. "Access to Four-Year Public Colleges and Degree Completion." *Journal of Labor Economics* 35(3): 825-67.

Griffith, Amanda, and Donna Rothstein. 2009. "Can't Get There From Here: The Decision to Apply to a Selective College." *Economics of Education Review* 28: 620-28.

Griliches, Zvi, and Vidar Ringstad. 1970. "Error-in-the-Variables Bias in Nonlinear Contexts." *Econometrica* 38(2): 368-70.

Heckman, James, and Tim Kautz. 2012. "Hard Evidence on Soft Skills." *Labour Economics* 19(1): 451-64.

Heckman, James, and Salvador Navarro. 2004. "Using Matching, Instrumental Variables and Control Functions to Estimate Economic Choice Models." *Review of Economics and Statistics* 86(1): 30-57.

Herrnstein, Richard, and Charles Murray. 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. New York: NY: Free Press.

Hershbein, Brad. 2013. "Worker Signals among New College Graduates: The Role of Selectivity and GPA," W.E. Upjohn Institute for Employment Research Working Paper 13-190.

Hoekstra, Mark. 2009. "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach." *Review of Economics and Statistics* 91(4): 717-24.

House, Emily. 2017. Finding the Best Fit: Exploring Postsecondary Undermatch in Tennessee." Unpublished doctoral dissertation, University of Michigan.
https://deepblue.lib.umich.edu/bitstream/handle/2027.42/137094/ehouse_1.pdf
Accessed July 23, 2019.

Hoxby, Caroline. 2009. "The Changing Selectivity of American Colleges." *Journal of Economic Perspectives* 23(4): 95-118.

Hoxby, Caroline. 2015. "Computing the Value-Added of American Post-Secondary Institutions," Unpublished manuscript, Stanford University.
https://www.irs.gov/pub/irs-soi/15rpcompvalueaddpostsecondary.pdf
Accessed July 23, 2019

Downloaded from by guest on October 15, 2026. Copyright 2019

Hoxby, Caroline and Christopher Avery. 2012. "The Missing `One-Offs': The Hidden Supply of High-Achieving, Low Income Students." NBER Working Paper No. 18586.

Hoxby, Caroline and Sarah Turner. 2013. "Expanding College Opportunities for High-Achieving, Low Income Students." SIEPR Discussion Paper 12-014.

Hoxby, Caroline, and Sarah Turner. 2015. "What High-Achieving Low-Income Students Know about College." *American Economic Review* 105(5): 514-17.

Kurlaender, Michal, and Eric Grodsky. 2013. "Mismatch and the Paternalistic Justification for Selective College Admissions." *Sociology of Education* 86(4): 294-310.

Light, Audrey, and Wayne Strayer. 2000. "Determinants of College Completion: School Quality or Student Ability?" *Journal of Human Resources* 35(2): 299-332.

Light, Audrey, and Wayne Strayer. 2004. "Who Receives the College Wage Premium? Assessing the Labor Market Returns to Degrees and College Transfer Patterns." *Journal of Human Resources* 39(3): 746-73.

Lincove, Jane, and Kalena Cortes. 2019. "Match of Mismatch? Automatic Admissions and College Preferences of Low- and High-Income Students." *Educational Evaluation and Policy Analysis* 41(1): 98-123.

Long, Mark. 2008. "College Quality and Early Adult Outcomes" *Economics of Education Review* 27: 588-602.

Long, Mark. 2010. "Changes in the Returns to Education and College Quality." *Economics of Education Review* 29: 338-47.

Loury, Linda Datcher, and David Garman. 1995. "College Selectivity and Earnings." *Journal of Labor Economics* 13(2): 289-308.

Mattern, Krista, Emily Shaw, and Jennifer Kobrin. 2010. "Academic Fit: Is the Right School the Best School or is the Best School the Right School?" *Journal of Advanced Academia* 21(3): 368-91.

Neal, Derek, and William Johnson. 1996. "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy* 104(5): 869-95.

Niu, Sunny Xinchun, and Marta Tienda. 2010. "The Impact of the Texas Top Ten Percent Law on College Enrollment: A Regression Discontinuity Approach." *Journal of Policy Analysis and Management* 29(1): 84-110.

Oster, Emily. 2017. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business and Economic Statistics*, forthcoming.

Pallais, Amanda. 2015. "Small Differences that Matter: Mistakes in Applying to College." *Journal of Labor Economics* 33(2): 493-520.

Radford, Alexandria. 2013. *Top Student, Top School? How Social Class Shapes Where Valedictorians Go to College*. Chicago: University of Chicago Press.

Downloaded from by guest on October 15, 2026. Copyright 2019

Roderick, Melissa, Jenny Nagaoka, Vanessa Coca, and Eliza Moeller with Karen Roddie, Jamiliyah Gilliam, and Desmond Patton. 2008. *From High School to the Future: Potholes on the Road to the Future*. Chicago, IL: Consortium on Chicago School Research.

Rodriguez, Awilda. 2015. "Tradeoffs and Limitations: Understanding the Estimation of College Undermatch." *Research in Higher Education* 56(6): 566-94.

Rothschild, Michael, and Lawrence White. 1995. "The Analytics of the Pricing of Higher Education and Other Services in Which the Customers are Inputs." *Journal of Political Economy* 103(3): 573-86.

Sallee, James, Alexandra Resch, and Paul Courant. 2008. "On the Optimal Allocation of Students and Resources in a System of Higher Education." *The B.E. Journal of Economic Analysis and Policy: Advances* 8(1): Article 11.

Sander, Richard, and Stuart Taylor. 2012. *Mismatch: How Affirmative Action Hurts Students It's Intended to Help, and Why Universities Won't Admit It*. New York, NY: Basic Books.

Scott-Clayton, Judith. 2012. "Information Constraints and Financial Aid Policy," NBER Working Paper No. 17811.

Segal, Carmit. 2012. "Working When No One is Watching: Motivation, Test Scores, and Economic Success." *Management Science* 58(8): 1438-57.

Smith, Jeffrey. 2008. "Heterogeneity and Higher Education." In *Succeeding in College: What it Means and How to Make it Happen*, ed. Michael McPherson and Morton Schapiro, 131-144. New York, NY: College Board.

Smith, Jonathan, Matea Pender, and Jessica Howell. 2013. "The Full Extent of Student-College Academic Undermatch." *Economics of Education Review* 23: 247-61.

Steeg, Jill, Jodi Upton, Patrick Bohn, and Steve Berkowitz. 2008. "College Athletes [sic] Studies Guided Toward `Major in Eligibility'." USA Today, November 19 http://usatoday30.usatoday.com/sports/college/2008-11-18-majors-cover_N.htm Accessed July 23, 2019.

Thompson, Leisa. 2008. "University of Michigan Athletes Steered to Professor." Ann Arbor News, March 6. https://www.mlive.com/wolverines/academics/stories/2008/03/athletes_steered_to_prof.html Accessed July 23, 2019.

Turner, Sarah. 2002. "Changes in the Return to College Quality," Unpublished manuscript, University of Virginia.

Wolfe, Tom. 2004. *I Am Charlotte Simmons*. New York, NY: Farrar, Straus, and Giroux.

Zimmerman, Seth. 2014. "The Returns to College Admission for Academically Marginal Students," *Journal of Labor Economics* 32(4): 711-54.

Downloaded from by guest on October 15, 2026. Copyright 2019

Table 1A: **Joint distribution of college quality and ability—NLSY-79**

Ability Quartiles	1st Quartile (lowest)	2nd Quartile	3rd Quartile	4th Quartile (highest)	Total
1st Quartile	10.6	7.0	4.7	2.7	
(lowest)	(42.6)	(27.9)	(18.7)	(10.8)	(100.0)
	[37.6]	[27.9]	[17.2]	[13.8]	[25.0]
2nd Quartile	7.7	7.0	6.4	3.9	
	(30.8)	(28.0)	(25.7)	(15.5)	(100.0)
	[27.3]	[28.0]	[23.7]	[19.9]	[25.1]
3rd Quartile	5.9	6.0	8.4	4.7	
	(23.6)	(24.1)	(33.5)	(18.8)	(100.0)
	[20.8]	[24.0]	[30.8]	[23.9]	[24.9]
4th Quartile	4.0	5.0	7.6	8.3	
(highest)	(16.1)	(20.0)	(30.6)	(33.2)	(100.0)
	[14.3]	[20.0]	[28.2]	[42.4]	[25.0]
	(28.3)	(25.0)	(27.1)	(19.6)	
Total	[100.0]	[100.0]	[100.0]	[100.0]	[2,497.0]

Table 1B: **Joint distribution of college quality and ability—NLSY-97**

Ability Quartiles	1st Quartile (lowest)	2nd Quartile	3rd Quartile	4th Quartile (highest)	Total
1st Quartile	12.0	7.0	4.3	1.8	
(lowest)	(48.0)	(27.8)	(17.0)	(7.2)	(100.0)
	[44.0]	[25.7]	[18.2]	[8.1]	[25.0]
2nd Quartile	7.4	8.0	5.5	4.1	
	(29.6)	(31.9)	(21.9)	(16.6)	(100.0)
	[27.0]	[29.4]	[23.4]	[18.7]	[25.0]
3rd Quartile	4.9	6.9	7.1	6.1	
	(19.6)	(27.7)	(28.2)	(24.5)	(100.0)
	[18.0]	[25.6]	[30.1]	[27.6]	[25.0]
4th Quartile	3.0	5.3	6.6	10.1	
(highest)	(12.0)	(21.0)	(26.6)	(40.4)	(100.0)
	[11.0]	[19.4]	[28.3]	[45.5]	[25.0]
	(27.3)	(27.1)	(23.4)	(22.2)	
Total	[100.0]	[100.0]	[100.0]	[100.0]	[2,071.0]

Notes: Each cell contains the overall percentage, (the row percentage), and [the column percentage]. College quality is measured by the 4-factor index. Ability is measured by the first principal component of the ASVAB scores. Percentages are weighted as described in the text. A Pearson's chi-squared test rejects equality of the ability-quality distributions across cohorts (p-value = 0.078).

Downloaded from by guest on October 15, 2026. Copyright 2019

Table 2: Summary of College Outcomes

	NLSY-79	NLSY-97
4-year college starters	2,497	2,071
Graduate within 6 years	50%	60%
In 5 years	42%	53%
In 4 years or less	26%	34%
Complete STEM degree	15%	13%
Transfer		27%
Transfer to a higher quality college		7%
Transfer to a lower quality college		16%
Labor force participation 2-3 years after entering	95%	92%
Men	94%	89%
Women	95%	94%
Labor force participation 10-11 years after entering	95%	92%
Men	96%	93%
Women	94%	91%
Average earnings 2-3 years after entering	$8,371	$9,382
Men	$9,556	$10,842
Women	$7,228	$8,246
Average earnings 10-11 years after entering	$39,022	$42,828
Men	$45,362	$48,830
Women	$32,627	$37,970

Notes: All percentages are of all four-year college starters. Two-year average earnings are calculated among those who worked in at least one of the target years. Weighted as described in the text.

Downloaded from by guest on October 15, 2026. Copyright 2019

Table 3: Effect of College Quality and Ability on Degree Attainment

		NLSY 79		NLSY 97	
		Within 6 years	Within 4 years	Within 6 years	Within 4 years
∂Outcome/∂A	Q = p25	0.487*	0.265*	0.310*	0.141*
		(0.055)	(0.056)	(0.058)	(0.061)
	Q = p50	0.495*	0.310*	0.264*	0.252*
		(0.055)	(0.053)	(0.060)	(0.059)
	Q = p75	0.467*	0.346*	0.313*	0.347*
		(0.057)	(0.048)	(0.064)	(0.058)
∂Outcome/∂Q	A = p25	0.223*	0.027	0.330*	0.241*
		(0.058)	(0.048)	(0.060)	(0.067)
	A = p50	0.273*	0.102	0.360*	0.323*
		(0.062)	(0.057)	(0.056)	(0.056)
	A = p75	0.210*	0.141*	0.308*	0.416*
		(0.064)	(0.058)	(0.065)	(0.051)
Observations		2,441	2,441	2,071	2,071
R-squared		0.129	0.096	0.201	0.180
Pr(interaction=0)		0.363	0.510	0.554	0.360
Pr(∂Outcome/∂A equal)		0.806	0.451	0.440	0.033
Pr(∂Outcome/∂Q equal)		0.304	0.207	0.537	0.068

Notes: Mean marginal effects are calculated from the coefficients of a polynomial of ability and college quality as described in the text. Standard errors in parentheses. * indicates that the estimated effect is statistically significant at 5%. The final three rows present the p-values from Wald tests that the coefficients on the interaction terms of the ability-quality polynomial are jointly equal to zero, that the mean marginal effects of ability are equal across the three percentiles of college quality, and that the mean marginal effects of college quality are constant across the three percentiles of ability.

Downloaded from by guest on October 15, 2026. Copyright 2019

Table 4A: Counterfactuals from Re-assigning Students to Colleges, NLSY-79

	Actual outcome	If all students attend a matched college	If all students attend a 90th pctile college
Graduate within 6 years	49.7	50.3#	58.8#
	(0.15)	(0.19)	(0.36)
Graduate within 4 years	26.0	27.4#	29.5#
	(0.13)	(0.18)	(0.35)
Major in STEM	15.5	16.0#	14.4#
	(0.12)	(0.17)	(0.25)
Earnings 2-3 years after starting college	8,383	8,354	8,190#
	(25)	(29)	(89)
Earnings 10-11 years after starting college	39,014	39,708#	40,135#
	(86)	(116)	(220)
Earnings 20-21 years after starting college	64,330	67,169#	69,768#
	(182)	(279)	(563)
Earnings 30-31 years after starting college	79,201	84,639#	88,589#
	(295)	(406)	(907)

Table 4B: Counterfactuals from Re-assigning Students to Colleges, NLSY-97

	Actual outcome	If all students attend a matched college	If all students attend a 90th pctile college
Graduate within 6 years	59.9	60.4#	70.5#
	(0.18)	(0.21)	(0.45)
Graduate within 4 years	33.8	36.1#	46.9#
	(0.16)	(0.20)	(0.46)
Major in STEM	13.3	13.4	14.1#
	(0.12)	(0.15)	(0.30)
Transfer to a higher quality college	7.6	6.9#	0.9#
	(0.09)	(0.11)	(0.05)
Transfer to a lower quality college	16.7	17.0#	17.5#
	(0.12)	(0.16)	(0.39)
Earnings 2-3 years after starting college	9,401	8,896#	10,218#
	(31)	(36)	(111)
Earnings 10-11 years after starting college	42,730	44,058#	49,516#
	(115)	(161)	(291)

Notes: Counterfactual outcomes are calculated using the estimates reported in Tables 3, 5, and 6. Annual earnings are in 2010 dollars. Bootstrapped standard errors from 200 draws in parentheses. # indicates that the counterfactual predictions are statistically different from the observed outcomes with 5% confidence.

Downloaded from by guest on October 15, 2026. Copyright 2019

Table 5: Effect of College Quality and Ability on Intermediate Outcomes

		NLSY 79	NLSY 97		
		STEM degree	STEM degree	Transfer up	Transfer down
∂Outcome/∂A	Q = p25	0.384*	0.250*	0.139*	-0.074
		(0.059)	(0.061)	(0.060)	(0.039)
	Q = p50	0.390*	0.236*	0.050	-0.025
		(0.052)	(0.048)	(0.033)	(0.055)
	Q = p75	0.361*	0.237*	0.014	-0.064
		(0.047)	(0.040)	(0.011)	(0.064)
∂Outcome/∂Q	A = p25	0.014	0.036	-0.132*	0.090
		(0.018)	(0.031)	(0.031)	(0.047)
	A = p50	0.063	0.046	-0.165*	0.079*
		(0.036)	(0.041)	(0.041)	(0.039)
	A = p75	0.079	0.037	-0.262*	0.108*
		(0.052)	(0.052)	(0.061)	(0.028)
Observations		2,396	2,067	2,044	2,044
R-squared		0.145	0.114	0.111	0.039
Pr(interaction=0)		0.708	0.791	0.384	0.328
Pr(∂Outcome/∂A equal)		0.688	0.960	0.098	0.363
Pr(∂Outcome/∂Q equal)		0.239	0.920	0.106	0.512

Notes: Mean marginal effects are calculated from the coefficients of a polynomial of ability and college quality as described in the text. Standard errors in parentheses. * indicates that the estimated effect is statistically significant at 5%.

Downloaded from by guest on October 15, 2026. Copyright 2019

Table 6A: Effect of College Quality and Ability on Earnings, NLSY-79

		Year 2-3	Year 10-11	Year 20-21	Year 30-31
∂Earnings/∂A	Q = p25	-3,118*	8,674*	28,372*	30,299*
		(1,160)	(3,379)	(7,675)	(11,487)
	Q = p50	-3,365*	9,150*	30,619*	40,513*
		(1,276)	(3,562)	(8,030)	(12,278)
	Q = p75	-2,858*	12,827*	42,637*	62,602*
		(1,341)	(3,588)	(8,052)	(12,302)
∂Earnings/∂Q	A = p25	-1,797	4,183	12,945	12,242
		(1,227)	(2,839)	(6,691)	(9,718)
	A = p50	-2,075	5,752	27,397*	29,541*
		(1,141)	(3,057)	(8,098)	(11,896)
	A = p75	-1,759	7,243*	25,270*	40,482*
		(1,039)	(3,145)	(7,808)	(11,222)
Observations		2,094	2,228	1,792	1,593
R-squared		0.093	0.141	0.186	0.199
Pr(interaction=0)		0.816	0.710	0.035	0.083
Pr(∂Earnings/∂A equal)		0.842	0.502	0.219	0.077
Pr(∂Earnings/∂Q equal)		0.871	0.712	0.015	0.044

Table 6B: Effect of College Quality and Ability on Earnings, NLSY-97

		Year 2-3	Year 10-11
∂Earnings/∂A	Q = p25	-2,093	3,581
		(1,189)	(3,891)
	Q = p50	-3,936*	4,173
		(1,231)	(4,283)
	Q = p75	-7,145*	12,667*
		(1,745)	(4,458)
∂Earnings/∂Q	A = p25	767	10,393*
		(1,597)	(4,089)
	A = p50	-2,649*	14,820*
		(1,336)	(4,065)
	A = p75	-4,033*	18,634*
		(1,394)	(3,986)
Observations		1,855	1,732
R-squared		0.138	0.136
Pr(interaction=0)		0.079	0.281
Pr(∂Earnings/∂A equal)		0.034	0.149
Pr(∂Earnings/∂Q equal)		0.038	0.289

Notes: Mean marginal effects are calculated from the coefficients of a polynomial of ability and college quality as described in the text. Standard errors in parentheses. * indicates that the estimated effect is statistically significant at 5%. Annual earnings are in 2010 dollars.

Downloaded from by guest on October 15, 2026. Copyright 2019

Table 7: Earnings 10-11 Years after Starting College, by Parents' Education

		NLSY 79		NLSY 97	
		Parents H.S. only	Parents some college	Parents H.S. only	Parents some college
∂Earnings/∂A	Q = p25	17,427*	2,151	9,498	3,252
		(5,336)	(4,055)	(11,359)	(4,154)
	Q = p50	15,743*	5,495	15,805	3,303
		(5,190)	(4,665)	(11,458)	(4,678)
	Q = p75	15,417*	12,815*	17,449	12,039
		(4,985)	(5,016)	(11,324)	(5,053)
∂Earnings/∂Q	A = p25	7,902*	-76	12,274	9,519
		(3,765)	(4,165)	(7,077)	(4,915)
	A = p50	7,331	4,952	7,600	14,741
		(4,641)	(3,987)	(7,419)	(4,764)
	A = p75	5,442	9,236*	19,246	17,464
		(5,201)	(4,071)	(10,480)	(4,506)
Observations		2,199	2,199	1,712	1,712
R-squared		0.145	0.145	0.139	0.139
Pr(interaction=0)		0.969	0.527	0.612	0.424
Pr(∂Earnings/∂A equal)		0.881	0.180	0.688	0.211
Pr(∂Outcome/∂Q equal)		0.921	0.246	0.406	0.402
N, subgroup		1,076	1,123	421	1,291

Notes: Mean marginal effects are calculated from the coefficients of a polynomial of ability and college quality as described in the text. Standard errors in parentheses. * indicates that the estimated effect is statistically significant at 5%. Annual earnings are in 2010 dollars. Students are divided into those who grew up with at least one parent who had some college education or who had no parents with more than a high school education.

Downloaded from by guest on October 15, 2026. Copyright 2019

Table 8: Earnings 10-11 Years after Starting College, by Sex

		NLSY 79		NLSY 97	
		Women	Men	Women	Men
∂Earnings/∂A	Q = p25	10,060*	7,077	4,944	2,120
		(4,276)	(4,663)	(4,438)	(6,369)
	Q = p50	9,231*	8,108	7,547	294
		(4,242)	(5,276)	(4,887)	(6,953)
	Q = p75	10,874*	13,926*	12,711	11,724
		(4,106)	(5,368)	(5,105)	(7,340)
∂Earnings/∂Q	A = p25	5,856	2,755	9,788	11,904
		(3,161)	(4,751)	(5,075)	(6,401)
	A = p50	6,542	4,977	13,967	16,481
		(3,525)	(5,013)	(4,924)	(6,650)
	A = p75	5,821	7,973	17,338	20,086
		(3,994)	(4,536)	(4,702)	(6,694)
Observations		2,228	2,228	1,732	1,732
R-squared		0.142	0.142	0.138	0.138
Pr(interaction=0)		0.942	0.892	0.674	0.768
Pr(∂Earnings/∂A equal)		0.847	0.461	0.478	0.229
Pr(∂Earnings/∂Q equal)		0.933	0.688	0.464	0.650
N, subgroup		1,171	1,057	966	766

Notes: Mean marginal effects are calculated from the coefficients of a polynomial of ability and college quality as described in the text. Standard errors in parentheses. * indicates that the estimated effect is statistically significant at 5%. Annual earnings are in 2010 dollars.

Downloaded from by guest on October 15, 2026. Copyright 2019

Table 9A: Covariate Set Comparisons for Graduating within 6 Years, NLSY-79

		(1)	(2)	(3)	(4)	(5)
∂Outcome/∂A	Q = p25	0.494*	0.514*	0.476*	0.487*	0.486*
		(0.049)	(0.051)	(0.054)	(0.055)	(0.055)
	Q = p50	0.508*	0.527*	0.489*	0.495*	0.490*
		(0.049)	(0.051)	(0.055)	(0.055)	(0.055)
	Q = p75	0.528*	0.532*	0.486*	0.467*	0.458*
		(0.048)	(0.049)	(0.055)	(0.057)	(0.057)
∂Outcome/∂Q	A = p25	0.224*	0.236*	0.222*	0.223*	0.233*
		(0.056)	(0.054)	(0.055)	(0.058)	(0.058)
	A = p50	0.299*	0.316*	0.292*	0.273*	0.280*
		(0.062)	(0.060)	(0.061)	(0.062)	(0.062)
	A = p75	0.248*	0.255*	0.235*	0.210*	0.210*
		(0.062)	(0.061)	(0.062)	(0.064)	(0.064)
Additional ability measures			Yes	Yes	Yes	Yes
Demographics and family				Yes	Yes	Yes
Neighborhood					Yes	Yes
Additional covariates						Yes
Observations		2,497	2,497	2,497	2,441	2,441
R-squared		0.095	0.113	0.123	0.129	0.135
Pr(interaction=0)		0.502	0.355	0.363	0.363	0.333
Pr(∂Outcome/∂A equal)		0.862	0.944	0.940	0.806	0.783
Pr(∂Outcome/∂Q equal)		0.195	0.124	0.187	0.304	0.280

Table 9B: Covariate Set Comparisons for Graduating within 6 Years, NLSY-97

		(1)	(2)	(3)	(4)	(5)
∂Outcome/∂A	Q = p25	0.330*	0.358*	0.306*	0.310*	0.288*
		(0.060)	(0.058)	(0.059)	(0.058)	(0.059)
	Q = p50	0.264*	0.316*	0.260*	0.264*	0.255*
		(0.061)	(0.058)	(0.060)	(0.060)	(0.060)
	Q = p75	0.341*	0.362*	0.304*	0.313*	0.295*
		(0.066)	(0.063)	(0.064)	(0.064)	(0.064)
∂Outcome/∂Q	A = p25	0.479*	0.393*	0.355*	0.330*	0.332*
		(0.057)	(0.059)	(0.058)	(0.060)	(0.059)
	A = p50	0.508*	0.420*	0.380*	0.360*	0.361*
		(0.055)	(0.056)	(0.055)	(0.056)	(0.056)
	A = p75	0.436*	0.371*	0.330*	0.308*	0.319*
		(0.067)	(0.064)	(0.064)	(0.065)	(0.065)
Additional ability measures			Yes	Yes	Yes	Yes
Demographics and family				Yes	Yes	Yes
Neighborhood					Yes	Yes
Additional covariates						Yes
Observations		2,071	2,071	2,071	2,071	2,071
R-squared		0.127	0.173	0.195	0.201	0.212
Pr(interaction=0)		0.225	0.586	0.606	0.554	0.704
Pr(∂Outcome/∂A equal)		0.237	0.536	0.491	0.440	0.618
Pr(∂Outcome/∂Q equal)		0.411	0.615	0.591	0.537	0.608

Notes: Mean marginal effects are calculated from the coefficients of a polynomial of ability and college quality. Standard errors in parentheses. * indicates that the estimated effect is statistically significant at 5%. Annual earnings are in 2010 dollars. Covariate categories defined in Appendix Table A2.

Downloaded from by guest on October 15, 2026. Copyright 2019

Table 10A: Covariate Set Comparisons for Earnings 10-11 Years after Start, NLSY-79

		(1)	(2)	(3)	(4)	(5)
∂Earnings/∂A	Q = p25	12,376*	10,613*	8,369*	8,674*	8,385*
		(2,981)	(3,169)	(3,322)	(3,379)	(3,434)
	Q = p50	13,391*	11,495*	8,913*	9,150*	8,979*
		(3,259)	(3,461)	(3,547)	(3,562)	(3,573)
	Q = p75	17,932*	15,856*	13,338*	12,827*	12,916*
		(3,165)	(3,378)	(3,515)	(3,588)	(3,569)
∂ Earnings/∂Q	A = p25	5,942*	6,583*	6,641*	4,183	3,602
		(2,749)	(2,750)	(2,683)	(2,839)	(2,798)
	A = p50	7,526*	8,096*	8,500*	5,752	5,516
		(3,064)	(3,007)	(2,931)	(3,057)	(3,013)
	A = p75	10,336*	10,677*	10,296*	7,243*	6,990*
		(3,100)	(3,049)	(3,008)	(3,145)	(3,155)
Additional ability measures			Yes	Yes	Yes	Yes
Demographics and family				Yes	Yes	Yes
Neighborhood					Yes	Yes
Additional covariates						Yes
Observations		2,278	2,278	2,278	2,228	2,228
R-squared		0.057	0.068	0.134	0.141	0.145
Pr(interaction=0)		0.510	0.553	0.533	0.710	0.648
Pr(∂Earnings/∂A equal)		0.321	0.357	0.336	0.502	0.448
Pr(∂Earnings/∂Q equal)		0.531	0.579	0.599	0.712	0.637

Table 10B: Covariate Set Comparisons for Earnings 10-11 Years after Start, NLSY-97

		(1)	(2)	(3)	(4)	(5)
∂Earnings/∂A	Q = p25	3,306	3,769	2,871	3,581	3,732
		(3,620)	(3,631)	(3,835)	(3,891)	(3,927)
	Q = p50	4,451	5,087	3,369	4,173	3,929
		(4,111)	(4,043)	(4,234)	(4,283)	(4,356)
	Q = p75	14,481*	15,225*	12,085*	12,667*	12,731*
		(4,400)	(4,481)	(4,453)	(4,458)	(4,478)
∂ Earnings/∂Q	A = p25	14,158*	12,877*	12,334*	10,393*	9,764*
		(4,123)	(4,205)	(4,032)	(4,089)	(4,106)
	A = p50	20,075*	19,064*	16,831*	14,820*	14,284*
		(4,149)	(4,134)	(4,005)	(4,065)	(4,110)
	A = p75	24,461*	23,489*	20,671*	18,634*	17,868*
		(3,928)	(3,969)	(3,912)	(3,986)	(3,994)
Additional ability measures			Yes	Yes	Yes	Yes
Demographics and family				Yes	Yes	Yes
Neighborhood					Yes	Yes
Additional covariates						Yes
Observations		1,732	1,732	1,732	1,732	1,732
R-squared		0.075	0.079	0.130	0.136	0.141
Pr(interaction=0)		0.176	0.172	0.262	0.281	0.259
Pr(∂Earnings/∂A equal)		0.085	0.082	0.137	0.149	0.134
Pr(∂Earnings/∂Q equal)		0.155	0.145	0.283	0.289	0.291

Notes: Mean marginal effects are calculated from the coefficients of a polynomial of ability and college quality. Standard errors in parentheses. * indicates that the estimated effect is statistically significant at 5%. Annual earnings are in 2010 dollars. Covariate categories defined in Appendix Table A2.

Downloaded from by guest on October 15, 2026. Copyright 2019

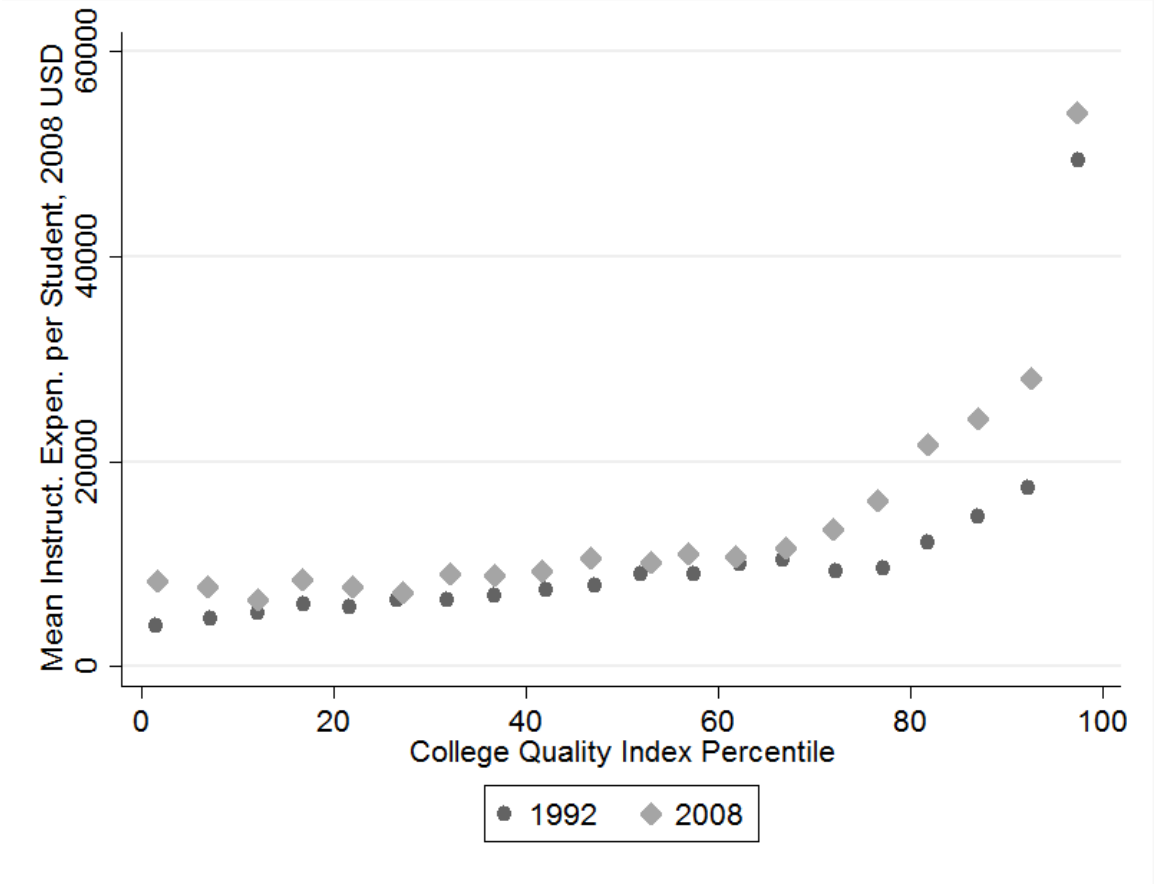

Figure 1: Expenditures per Student as a Function of the College Quality Index

Notes: Instructional expenditure per student from iPEDS and U.S. News and World Report. College quality indices calculated as described in the text.

Downloaded from by guest on October 15, 2026. Copyright 2019

Downloaded from by guest on October 15, 2026. Copyright 2019

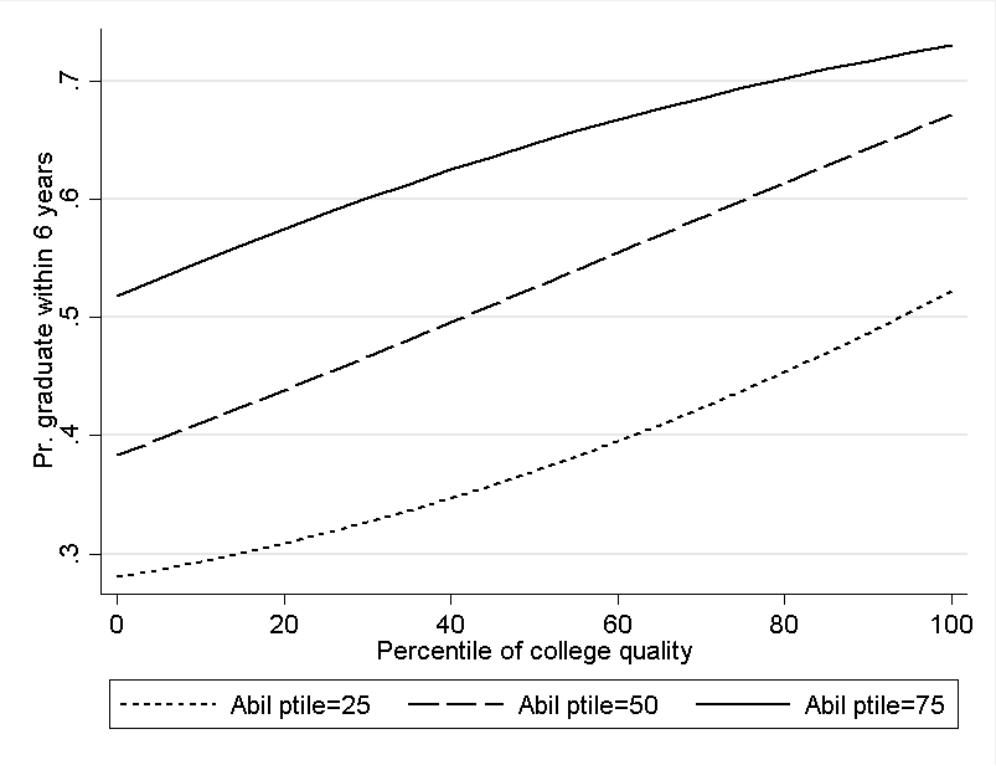

Figure 2a: Effect of College Quality on Graduating within 6 Years--NLSY-79 Cohort

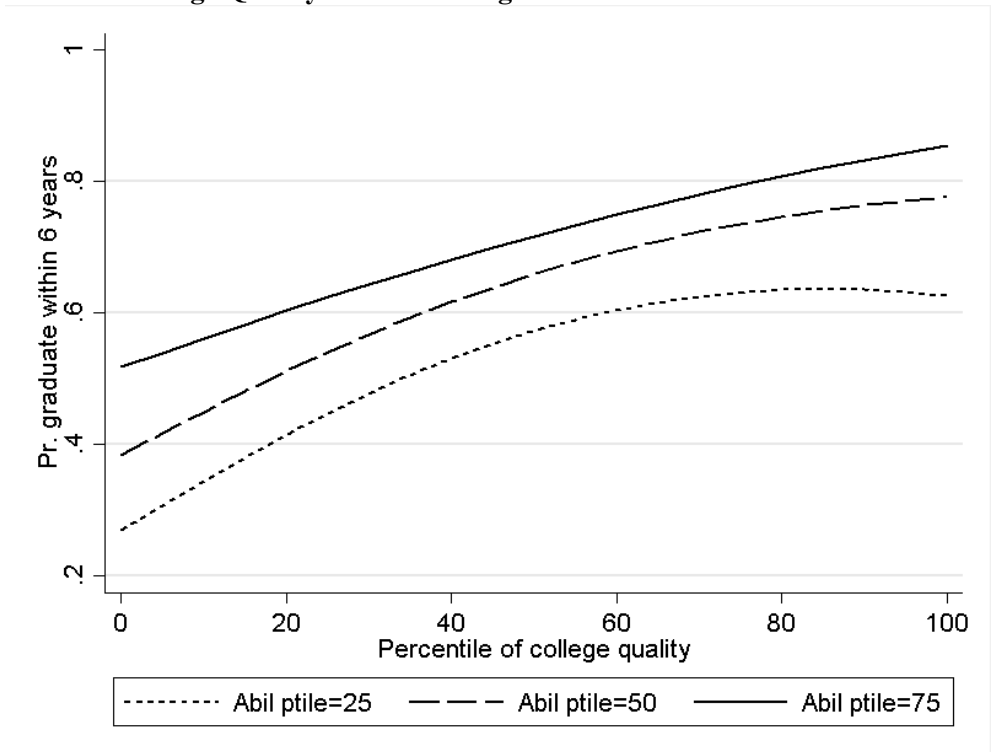

Figure 2b: Effect of College Quality on Graduating within 6 Years--NLSY-97 Cohort

Notes: Projected from estimates summarized in Table 3.

Downloaded from by guest on October 15, 2026. Copyright 2019

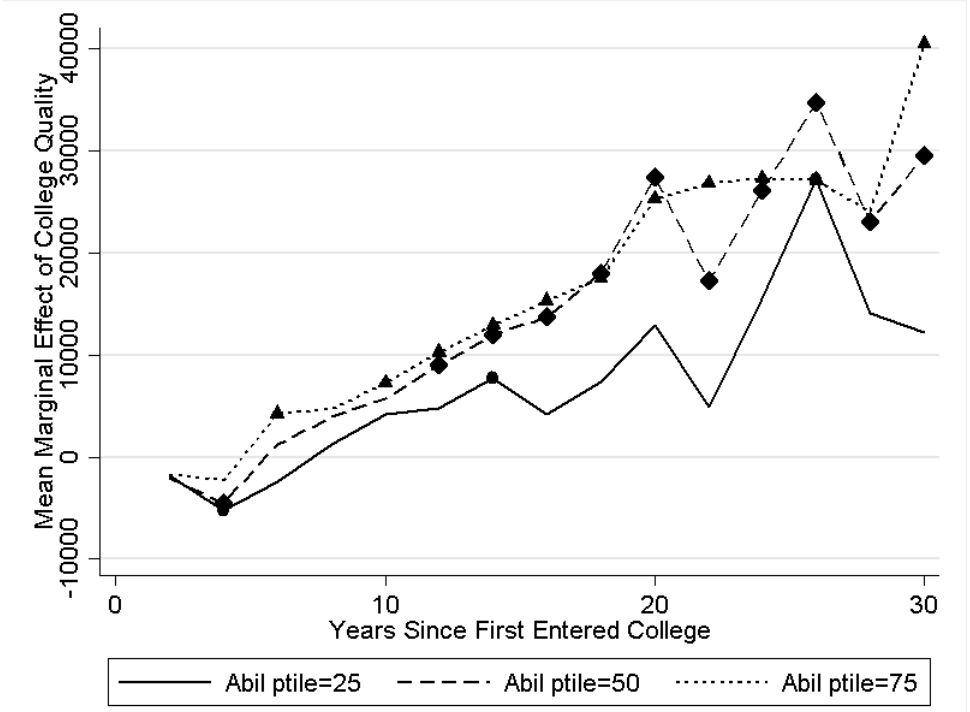

Figure 3a: Long-term Earnings Estimates for NLSY-79--Effects of College Quality

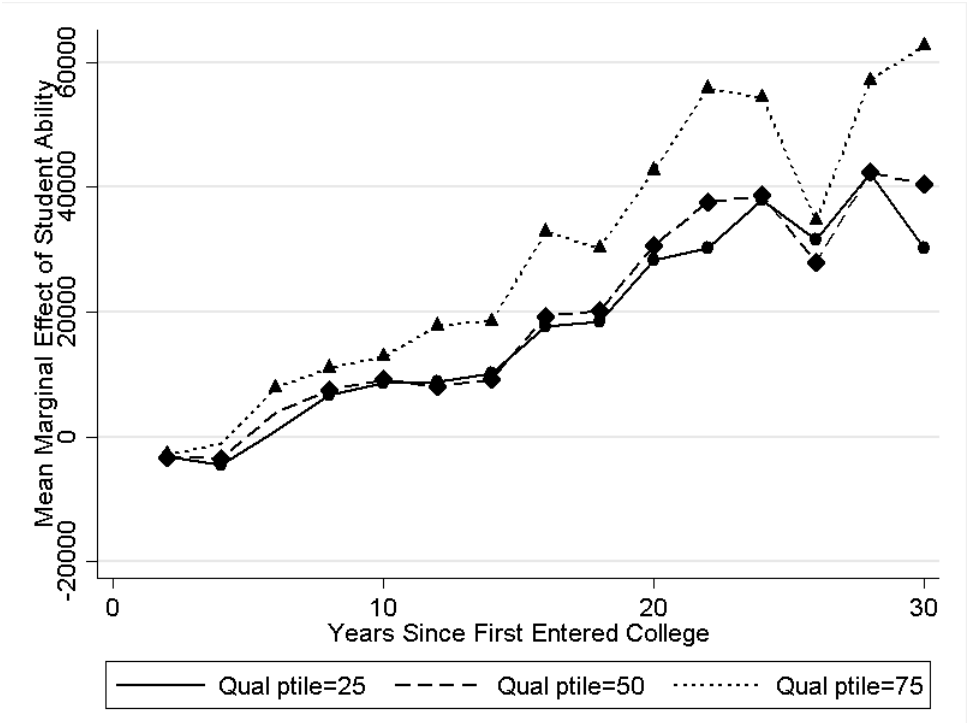

Figure 3b: Long-term Earnings Estimates for NLSY-79--Effects of Student Ability

Notes: Mean marginal effects are calculated from the coefficients of a polynomial of ability and college quality as described in the text. Markers indicate that the estimated effect is statistically significant at 5%. A subset of these estimates is presented in Table 6A.

APPENDIX

Appendix Table A1: Sample Composition

	NLSY 1979 Cohort	NLSY 1997 Cohort
Total Observations	9,369	8,984
Graduated HS or received GED	8,012	7,845
Started at a 4-year college by age 21	3,157	2,831
Interviewed at least 5 years after starting	3,157	2,696
Of eligible 4-year college starters:		
Missing 4-year college quality index	445	94
Has quality, but missing ASVAB	215	531
Analysis sample	*2,497*	*2,071*

Downloaded from by guest on October 15, 2026. Copyright 2019

Appendix Table A2: Description of Independent Variables

College quality	The college quality index is the first principal component of four college quality proxies as described in Online Appendix Table A4. We calculate percentiles of this index across all 4-year-degree granting colleges in IPEDS, weighting each institution by full-time equivalent undergraduate enrollment.
ASVAB 1 percentile	Percentile over 4-year college starters in each cohort of the first principal components of the ASVAB test scores, as described in Online Appendix Tables A3.
Additional ability measures	
ASVAB 2 percentile	Percentile over 4-year college starters in each cohort of the second principal components of the ASVAB test scores, as described in Appendix Tables A3.
High school GPA, percentile	Percentile over 4-year college starters in each cohort of cumulative high school grade point average
SAT or ACT percentile	Percentile over 4-year college starters in each cohort of reported SAT score (or ACT score converted to SAT scale)
Any bad behaviors	Indicator that respondent reported any of the following petty anti-social behaviors by 1980 in NLSY 79 or by 8th grade in NLSY-97: ever suspended from school, ever intentionally destroyed or damaged someone else's property, and ever stolen something worth $50 or less
Interviewer rated uncooperative	Indicator that the NLSY interviewer rated the respondent as somewhat uncooperative in any of the first 3 interviews. In the NLSY 79 this corresponds to a classification of impatient, where the other options are friendly, cooperative, and hostile and friendly is the modal response. In the NLSY 97, this designation corresponds to a score of 3-8 on a scale of 1=hostile to 10=very cooperative, where 10 is the modal response.
Had sex before age 15	Indicator that respondent reported having sex before the age of 15
Demographics and family	
Male	Indicator that the respondent identifies as male
Race and ethnicity	Indicators for white, black, non-white Hispanic, or other non-white (last category in NLSY 97 only). White is omitted group.

Family income quartiles	Indicators for quartile (calculated within the weighted NLSY sample) of total household income in 1979 or 1997. In the NLSY97, this information is taken from the 1997 parent survey where available or from the youth survey (98.6% from parent survey). The NLSY79 did not give parents a separate survey. First quartile is omitted group.
Siblings	Number of siblings reported by the respondent in the NLSY 79 or children age 18 and under living at the respondent's address in 1997 for the NLSY-97.
Parental education	Indicators for the highest educational attainment of either of the respondent's resident parents (or only parent in single parent households) as reported in the first survey waves for each cohort. We include at most one resident mother and father figure using the following prioritization: biological, adopted, step, or foster. High school diploma is the omitted category.
Neighborhood	
Regional indicators	Indicators for region of the U.S. (Northeast, South, Midwest, West) where the respondent lived in 1979 or 1997. Midwest is the omitted category.
Rural	Indicator that the respondent did not live within a Metropolitan Statistical Area (MSA) in 1979 or 1997.
% Adults w/college deg. in county	The share of the over-25 population that has a 4-year college degree in the county where the respondent lived in the first year of each survey, from the 1972 and 1994 County and City Databooks, respectively.
Additional covariates	
Overweight/Obese	Indicators that the respondent was overweight or obese (using BMI and CDC definitions) in 1979 or 1997.
Religious observance per year	How many times per year the respondent attended religious services in 1979 or 1997 (entered in regressions as indicators for each range of values offered in the survey).
Count of enriching resources	Count of educational resources the respondent said he or she had regular access to at home in the 1979/1997 surveys. In the NLSY-79, these resources are: a magazine subscription, a newspaper subscription, and a library card. In the NLSY-97 these resources are: a computer, a dictionary, and a quiet place to study.
Had contact with biological mother/father	Indicator that respondent had ever lived with each biological parent for at least three months by the age of 18 (NLSY-79) or had any contact with each biological parent by 1997 (NLSY 97).

Downloaded from by guest on October 15, 2026. Copyright 2019

[1] Kurlaender and Grodsky (2013) provide similar clarification in the sociology literature.

[2] Recent studies that examine the "main effect" of college quality include Black and Smith (2004), Bowen, Chingos, and McPherson (2009), Cohodes and Goodman (2014), Cunha and Miller (2014), Dale and Krueger (2002, 2014),

Hoekstra (2009), Hoxby (2015), Long (2008, 2010), and Zimmerman (2014). All agree on a positive causal effect for at least some groups.

[3] We focus here on papers that address academic match at the undergraduate level and that use U.S. data. See Sander and Taylor (2012) for a survey of the related, tendentious, literature on academic match in law school.

[4] Alon and Tienda (2005) examine academic match using the High School and Beyond and National Educational Longitudinal Study of 1988 (NELS:88) datasets. Unfortunately, they look only for effects of selectivity (their proxy for college quality) conditional on ability rather than for effects of the interaction of selectivity and ability.

[5] The exception is Loury and Garman (1995), who find substantively important match effects on degree completion (including negative effects of college quality for black students) and post-college earnings in their study that uses the National Longitudinal Study of the High School Class of 1972. Their earnings estimates condition both on a much less rich set of background variables and on several intermediate outcomes – college GPA, major, years of college – and so correspond to a very different estimand than our own. Their completion estimates do not have the issue of conditioning on intermediate outcomes and so remain a puzzle. A replication in light of the subsequent literature would add value.

[6] See in particular their Figures 10.5a, 10.5b, 11.1, 11.2 and 11.3. They use the terminology of match somewhat differently than we do; in particular, they sometimes refer to what we call the main effect of quality as a match effect when it applies to overmatched or undermatched students. Cunha (2009) critiques this study.

[7] The fine Arcidiacono et al. (2013) paper looking at friendship networks in college also sheds some light on potential mechanisms.

[8] In other recent RD papers, Zimmerman (2014) and Goodman, Hurwitz, and Smith (2017) find substantively important effects of college quality (and/or type) on labor market outcomes toward the other end of the college quality spectrum, namely the margin between low quality four-year colleges and two-year colleges.

[9] In the NLSY-79 we omit the military and low-income white samples because both were dropped from the survey before most respondents had completed college. We use custom probability of inclusion weights, constructed by the NLSY, to combine the sampling groups in each survey, and to control for differing response rates by age, sex, and race-ethnicity.

[10] The NLSY datasets also feature impressively high response rates, over 80 percent of the initial respondents in most survey rounds. See https://www.nlsinfo.org/content/cohorts/nlsy79/intro-to-the-sample/retention-reasons-

Downloaded from by guest on October 15, 2026. Copyright 2019

noninterview and https://www.nlsinfo.org/content/cohorts/nlsy97/intro-to-the-sample/retention-reasons-non-interview.

[11] The ASVAB test is not a straightforward measure of "innate" ability because it includes the influences and training that the student has experienced up to the point she takes the test. See Neal and Johnson (1996) for a more thorough discussion of what the ASVAB test measures. We do not mind if the ASVAB also measures intrinsic motivation, as argued by Segal (2012). More broadly, we use the term "ability" quite agnostically to mean the set of skills, innate or acquired, that students possess around the time of the college choice.

[12] We follow Aucejo and James (2019) and include an index of petty anti-social behaviors before age 14 and early sexual activity. Following Cadena and Keys (2015), we also include an indicator of whether the NLSY interviewer rated the respondent as somewhat uncooperative in any of the first three rounds of interviews.

[13] U.S. News and IPEDS collect many of the same statistics and often report identical values. Combining data from the two sources gives us the most complete sample of colleges. We measure college quality somewhat later than the years when each cohort entered college. Many of the component measures first become available in IPEDS in 1992 and 2008 is the earliest year for which we could obtain recent U.S. News data. In both cases, the stability of the underlying proxies over time assuages any concerns about the modest temporal distance.

[14] We weight the percentiles by full-time-equivalent undergraduates. In each reference year, our sample includes all colleges that identify as offering four-year degrees in IPEDS that year, have at least one first-degree-seeking undergraduate enrolled, and report at least two of our quality proxies. We exclude specialty institutions such as nursing colleges and theological seminaries (Carnegie codes 51-60).

[15] Our estimates also exhibit face validity. For example, for the NLSY-97 cohort the University of Michigan gets a 93, Michigan State a 74, Wayne State a 36 and Eastern Michigan a 28.

[16] Our measure of student body size is full-time equivalent undergraduates.

[17] Our measure also differs from that in Chetty et al. (2017). Rodriguez (2015) and House (2017) analyze various measures of academic match and find that different measures provide qualitatively different pictures of the nature and extent of deviations from assortative matching.

[18] Black and Smith (2004, Table 4) and Light and Strayer (2000, Table 3) present alternative estimates of the joint distribution for the NLSY-79 cohort that tell the same basic story. See also Mattern, Shaw, and Kobrin (2010, Table 1).

Downloaded from by guest on October 15, 2026. Copyright 2019

[19] Statistically, a chi-squared test rejects the null of a common joint distribution in the two cohorts. Substantively, our results comport with Hoxby (2009, Figure 1) and Chapter 1 of Herrnstein and Murray (1994), which show that much, but not all, of the large increase in stratification by ability among colleges had played out by the time the NLSY-79 cohort entered college. Smith, Pender, and Howell (2013) reach a quite different conclusion from ours, arguing that undermatch decreased dramatically between the two cohorts they consider. The different timing of their two cohorts, their (quite) different definition of undermatch, and their inclusion of two-year schools stymie a detailed accounting of the differences between their finding and ours.

[20] Light and Strayer (2004) document the nature and extent of transfers in the NLSY-79 and consider their relationship to wages. In keeping with the limited nature of the available data, they distinguish between two-year and four-year colleges but not more finely by quality.

[21] See Table A-1 of their online appendix.

[22] The NLSY datasets offer two different earnings measures: a CPS-like measure based on a question about total earnings the previous year and a constructed variable that builds on information about wages, hours, and weeks on individual job spells. We use the CPS-like measure for both cohorts.

[23] We compared many more and less parsimonious specifications, with and without additional covariates. For several outcomes, these tests do not reject the exclusion of *all* ability-quality interaction terms. We include the most parsimonious specification that still allows for non-linear interaction effects and report tests of the joint significance of these interaction terms in our results. Our thorough search leads us to think that other paths to functional form flexibility, such as substituting the log of earnings for the level or considering the levels of our ability and quality measures rather than their percentiles, would do little to alter our qualitative findings.

[24] We experimented with two other ways of using the SAT and GPA variables: one set of analyses omitted them while the other combined them with the ASVAB components to create a broader ability index. Neither strategy affects our qualitative conclusions. We do not use the broader index as our primary ability measure because of the large number of observations with missing information on SAT and/or GPA.

[25] Replacing the five percentile-point cutoff for a transfer up or down with a zero cutoff or a 10 percentile-point cutoff yields qualitatively similar findings.

[26] Our estimates for 6-7 years after college start (not shown) lie in between the 2-3 and 10-11 year estimates.

[27] See Hershbein (2013) for subtler signaling theories of college quality.

Downloaded from by guest on October 15, 2026. Copyright 2019

[28] All of the qualitative findings in Table 6A related to long-term earnings impacts for the NLSY-79 cohort persist if we restrict ourselves to a balanced panel.

[29] Completion rate estimates by subgroup tell broadly the same story as the pooled completion estimates, with more volatility in individual point estimates due to smaller sample sizes.

[30] The estimated effects of ability and college quality on graduation rates in the NLSY-79 cohort display less sensitivity to the conditioning set than the other outcomes we consider. We lack a good explanation for this pattern.

[31] The corresponding completion probabilities, which we find easier to interpret, appear in their Table 12. Because the model underlying their Table 12 assumes independent errors, the distribution of unobserved variables does not depend on the choice of college and college quality in the first period. Thus, we interpret the three rows for the "overall sample" for each ability quartile as three independent simulations of the same parameter values.

[32] Though substantively different, our estimates do not quite differ statistically from theirs. For example, in their Table 12 students in AFQT quartile 1, roughly our ASVAB quartile 1, suffer a reduction in college completion probabilities of around 0.07 from moving from their first to second quartiles of college quality. The corresponding estimate in our alternative specification using quartile indicators presented in Online Appendix Table A5 is an increase of 0.025 with a standard error of 0.049. The comparison is complicated by the fact that they do not present standard errors on their predictions and the fact that the two estimates, which rely on the same data, presumably have a non-zero covariance.

[33] Other differences seem to us as a priori less likely to account for a large portion of the difference, e.g. (1) LS measure college quality differently than we do; (2) LS measure student ability a bit differently than we do, relying on the Armed Forces Qualifying Test (AFQT) score, a weighted average of four ASVAB component scores; and (3) their sample differs somewhat from ours, as indicated by their sample size of 2,754 compared to ours of 2,441.

[34] We also used our data to replicate their specification and then marched, one change at a time, from their setup to our setup. When we did what they did, we got estimates that look very much like what they got. Key differences result from using the first college rather than the last college attended, which reduces the estimated effect of college quality somewhat, and from including the county conditioning variables, which also reduce the estimated effect of college quality.

[35] The earnings measure is the median value of log annual earnings over five-year intervals, excluding individuals with low enough values to suggest only marginal labor market attachment.

[36] This identification strategy has issues of its own; see Hoxby (2009) for discussion.

Downloaded from by guest on October 15, 2026. Copyright 2019

[37] With the possible statistically insignificant and empirically irrelevant exception of students in the 25th percentile of ability attending colleges in the top 10% of the quality distribution.

Downloaded from by guest on October 15, 2026. Copyright 2019