The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

Working Papers
www.cesifo.org/wp

The Consequences of Academic Match between Students and Colleges

Eleanor Wiske Dillon
Jeffrey Andrew Smith

CESIFO WORKING PAPER NO. 6344
CATEGORY 5: ECONOMICS OF EDUCATION
FEBRUARY 2017

An electronic version of the paper may be downloaded
- *from the SSRN website:* *www.SSRN.com*
- *from the RePEc website:* *www.RePEc.org*
- *from the CESifo website:* *www.CESifo-group.org/wp*

ISSN 2364-1428

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 6344*

The Consequences of Academic Match between Students and Colleges

Abstract

We consider the effects of student ability, college quality, and the interaction between the two on academic outcomes and future earnings using data on two cohorts of college enrollees drawn from the NLSY-79 and the NLSY-97. We find that student sorting has increased modestly between cohorts, and that student ability and college quality strongly improve degree completion and earnings. These patterns imply that, on average, students benefit from "overmatch" of the sort generated by affirmative action in admissions. We find little evidence of match effects on degree completion at eight years or on STEM degree completion, but suggestive evidence of some complementarity between student ability and college quality in degree completion at four years and long-term earnings. Such complementarity implies a tradeoff between equity and efficiency for policies that move lower ability students to higher quality colleges.

JEL-Codes: J310, I240.

Keywords: college mismatch, college quality.

Eleanor Wiske Dillon
Arizona State University
USA - Tempe, AZ 85287-9801
ewdillon@asu.edu

Jeffrey Andrew Smith
University of Michigan
USA - 48109-1220 Ann Arbor MI
econjeff@umich.edu

Version of January 28, 2017
We thank Dan Black, Hoyt Bleakley, John Bound, David Deming, Sue Dynarski, Josh Goodman, 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, Familes and Children, the 2013 CESifo education group meetings, Bristol, Chicago (family economics), Cleveland Fed, Cornell PAM, Dartmouth, MDRC, Michigan CIERS, National University of Singapore, Penn GSE, Seton Hall Conference on College Match, Stanford CEPA, and Washington University in St. Louis. This research was supported by NSF SES 0915467. Any foolishness the careful reader might uncover is our own damn fault. Please be sure and tell us about it.

1. Introduction

How students of varying ability sort into colleges of varying qualities has captured the attention not only of economists (and other academic researchers) studying higher education but also of the policy literature, the popular press and the blogosphere. Much of the literature frames the discussion in terms of the match between student ability and college quality, with relatively low ability students at relatively high quality colleges labeled "overmatched" and relatively high ability students at relatively low quality colleges labeled "undermatched". Until the last decade or so, the literature focused almost exclusively on overmatch, particularly overmatch induced by racial and ethnic preference policies at selective colleges. More recently, undermatch has moved into the spotlight as a result of the widely-read studies by Bowen, Chingos and McPherson (2009) and Roderick et al. (2008).

Despite the current ubiquity of the match conversation, we lack credible estimates of the effects of student-college match. The importance of both student ability and college quality for educational and labor market outcomes is well established. We ask whether college quality has different effects for students of different abilities. For example, an overmatched student might flounder and drop out or might rise to the challenge of a more demanding program. In the presence of such 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. Knowledge of the effects (if any) of academic match has clear value to students and parents making decisions about college enrollment, and to researchers and policymakers concerned with the design, operation and effects of state university systems with diversified college quality portfolios. While we follow the literature in using the "match" terminology, we emphasize that we use the term without ex ante normative intent. Instead, we use it to frame an empirical question about the form of the higher education production function.

Our analysis contributes to the small but growing literature on the effects of academic match at the college level in a number of important ways. First, we clarify the conceptual distinction between the main effects of college quality and student ability and the match effects that may or may not result from their interaction.[1]

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

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. As a natural byproduct of our analysis we replicate (in a broad sense) and extend the earlier analyses of the college quality main effect in the NLSY-79 presented in Black et al. (2005). We code our outcome and conditioning variables and design our analyses in the same way for the two datasets so as to make our analysis across cohorts as compelling as possible.

Third, we examine a variety of outcome measures. With a couple of important recent exceptions discussed in greater 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 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 et al. (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 Ann Arbor News (2008). Or they might transfer to another school that represents a better match. Our examination of transfers, highlighted by Arcidiacono and Lovenheim (2016) as an understudied outcome in this literature, 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 match. For the NLSY-79 cohort, we present estimates for earnings up to 30 years following college start, the first longer-term estimates of match effects.

Fourth, following Black and Smith (2006) and our earlier analysis of the determinants of academic match 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 (e.g. 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 match that use this identification strategy, we have a richer and more compelling set of relevant conditioning variables. The literature provides strong evidence of the importance of factors likely conditionally unrelated to outcomes in driving college choice; these factors provide exogenous variation in college quality. Unlike earlier papers, we show that our estimates stabilize as we add marginal sets of conditioning variables.

To preview our results, we find substantial amounts of both overmatch and undermatch in both cohorts, with a modest decline from the NLSY-79 to the NLSY-97. Our examination of the effects of ability, college quality and their interaction on college completion reveals substantively strong and statistically significant main effects of college quality and student ability on degree completion and earnings. These effects appear even for relatively low ability students, indicating that e.g. affirmative action likely passes a private benefit-cost test, on average, for the affected students.

In contrast, we find very little evidence of a causal effect of the interaction of quality and ability for these outcomes, with two exceptions: First, we find clear evidence of complementarity between student ability and college quality in time to degree: more able students benefit relatively more than less able students from college quality for this outcome. Second, we find suggestive evidence of complementarity for long-run earnings outcomes. These patterns appear for both cohorts of the NLSY, with some variation in timing for the earnings effects, and suggest 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 strong evidence of match-related transfer behavior: overmatched students have a higher conditional probability of transferring down and undermatched students have a higher conditional probability of transferring up. In contrast, and unlike Arcidiacono et al. (2016), we find no match effects related to STEM degree completion.

We structure the remainder of the paper as follows: Section 2 briefly reviews the literature on academic match. Section 3 describes our excellent data, with particular attention to the construction of our student ability and college quality measures and to the outcomes we consider. Section 4 lays outs our econometric framework. Section 5 considers identification and makes the case for a causal interpretation of our estimates. Section 6 presents our main findings on the consequences of academic match and compares results between the two cohorts and between our findings and key earlier studies. Finally, Section 7 offers our conclusions.

2. Literature

We view the study of student-college match as an 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-resource college. Overmatch is likewise mechanically beneficial. We now review the very modest subset of the now quite large literature on college quality that explicitly considers match. [3]

A few papers devote their full attention to the determinants of academic match. 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 patterns of match, 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 et al. (2013) conduct a similar analysis using different data sets and a different definition of match. Reassuringly, they obtain similar findings. Lincove and Cortes (2016) 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) emphasize the role of information about college choices in driving match 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 academic match using "selection on observed variables" identification strategies to deal with non-random selection of students into colleges of

[2] Recent studies that examine the "main effect" of college quality include Black and Smith (2004), Bowen et al. (2009), Dale and Krueger (2002, 2014), Hoekstra (2009), and Long (2008, 2010). 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.

varying qualities.[4] None of these studies find meaningful match effects, though several identify strong main effects of both student ability and college quality. [5] Mattern et al. (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 the effect of academic match on first year college GPA and persistence into the same college in the second year. 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 et al. (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 discuss their results in comparison to our own in Section 6.7. Finally, Bowen et al. (2009) examine match effects 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) look at academic match using the NLSY-79, but employ an empirical approach that differs substantially from 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 binary 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 (not super plausible) exclusion restrictions,

[4] Alon and Tienda (2005) examine academic match using the High School and Beyond data and the National Educational Longitudinal Study of 1988 (NELS:88). 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, which renders their results difficult to interpret in terms of match effects.

[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 mismatch effect when it applies to mismatched students.

from some restrictions on the coefficients on the interactions of student ability and college quality, and from restrictions on the covariance matrix of the college choice model errors. Their estimates reveal substantively important match effects; we say more about why their qualitative findings differ from our own in Section 6.7.

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 et al. (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 say he wishes 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 undermatched black students at Duke adapt by changing majors within an institutional context where essentially all students finish their degree. Arcidiacono et al. (2016) continue this line of work by examining students at different University of California (UC) campuses entering school in the years 1995 to 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 (= Science, Technology, Engineering and Math) major have lower probabilities of completing any degree at a UC school and of graduating in STEM. We compare our estimates to theirs in Section 6.2.[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. Kurlaender and Grodsky (2013) exploit an unusual event in 2004 in which the admissions offices of UC schools initially offered their marginal

[7] The Arcidiacono, Aucejo, Husser and Spenner (2013) paper looking at mismatch and friendship networks in college also sheds some light on potential mechanisms.

acceptances deferred admission due to a budget crisis but later repented and offered 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 academic match 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 evidence on match. The nature and extent of mismatch (as conceived of in the broader literature) 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 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 academic match effects, other than on intermediate outcomes such as transfer and major choice.

3. Data

3.1. NLSY

We use the younger part of the NLSY-79 data, which includes Americans who were ages 14 to 18 on January 1, 1979, and the full NLSY-97 data, which includes Americans who were 12 to 16 years old as of December 31, 1996.[9] 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 both the representative samples from each survey along with the supplemental samples of blacks and Hispanics.[10] Most respondents in the NLSY-79 sample graduated high school and made their

[8] In other recent RD papers, Zimmerman (2014) and Goodman et al. (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] We drop the oldest NLSY-79 respondents because many of these individuals were already living independently as of the first interview (so we lack information about parents' income) and because they took the ASVAB, our main measure of ability, after starting college.

[10] 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

college choice between 1979 and 1983, while the NLSY-97 sample did the same between 1998 and 2002. We focus on students who start at a four-year college by age 21 (39% of high school graduates and GED holders in the NLSY-79 sample and 37% 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.[11] 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 college quality measures, as well as our outcome variables; Appendix Tables 1 and 2 describe the construction of our analysis sample and our conditioning variables.

3.2. 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 Forces 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 et al. (2012) to construct comparable measures of eight ASVAB test components common to the two surveys, 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: auto and shop information and electronics information.

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.[12] As shown in Appendix Table 3, the first principal component explains 68% and

---

NLSY, to combine the sampling groups in each survey, and to control for differing response rates by age, gender, and race-ethnicity groups.

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

[12] 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.

66% of the total variance in test scores across the eight sections for the 1997 and 1979 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 ACT score equals 0.79 in NLSY-79 and 0.80 in NLSY-97.

The second component of the ASVAB scores, which we call ASVAB2, explains a further 10-11% of the variance. As in Cawley et al. (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[13] along with multiple proxies for non-cognitive or socio-emotional skills, listed in the appendix.[14]

3.3. 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 see Hoxby (2016) for an ambitious attempt to measure colleges' value-added). 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,[15] using data from 1992 for the NLSY-79 and data from 2008 for the NLSY-97.[16]

---

[13] Because we only interact ASVAB1 with college quality we 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. We experimented with two other ways of using the SAT and GPA variables: one set of analyses simply omitted them while the other set of analyses combined them with ASVAB1 to create a "super" ability index. Neither strategy affects our qualitative conclusions. We do not use the super index as our primary ability measure because of the large number of observations with missing information on SAT and/or GPA.

[14] We follow Aucejo (2012) 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.

[15] U.S. News and IPEDS collect many of the same statistics and for the same college in the same year the numbers are often identical. U.S. News has average SAT or ACT scores for the students at a number of schools that do not report test scores to IPEDS. However, U.S. News focuses on selective schools. Combining data from the two sources gives us the most complete sample of colleges. We use U.S. News data to fill in each college quality measure when these statistics are missing from IPEDS.

[16] While most NLSY-97 respondents started college between 1999 and 2002, 2008 is the earliest year for which we could obtain recent U.S. News data. In regard to NLSY-79, many of the component measures first become available

Following Black and Smith (2004, 2006) and Black et al. (2005) we use the first principal component across these four measures of quality as our quality index.[17] 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 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.[18] Our measure does not capture differences in the quality that different students experience within the same university due to e.g. 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 measures 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. 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 et al. (2005) and some other studies. Our concern with academic match 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 mismatch, as intermediate outcomes on the way to earnings effects. Appendix Table 5 summarizes student characteristics by college quality quartile; we relegate it to the appendix because it contains no surprises relative to the research on the determinants of match surveyed earlier

3.4 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. We calculate the college's quality percentile across all four-year institutions in the United States included in the IPEDS. We weight the quality percentile by student body size, so a college in the nth quality percentile is the college that a student in the nth ability percentile would attend under perfect

in IPEDS in 1992. In both cases, the stability of the underlying proxies over time assuages any concerns about the modest temporal distance.

[17] See the Appendix Table 4 for the details of the principal components analysis.

[18] Our estimates also exhibit face validity; choosing one corner of one state at random, for the NLSY-97 cohort Michigan gets a 93, Michigan State a 74, Wayne State a 36 and Eastern Michigan a 28.

assortative academic matching of students and colleges.[19] Academic mismatch occurs when students deviate substantially from this type of matching. One appealing feature of our measure is the possibility of achieving perfect assortative matching without violating institutional enrollment constraints. The academic match measures employed in other important studies in the literature, such as Roderick et al. (2008), Bowen et al. (2009), and Smith et al. (2013) lack this feature.[20]

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, which corresponds to academic match. The four diagonal cells account for 34.1% of students in 1979 and 37.8% 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 11.4% of students in 1979 and 9.7% in 1997, while the three lower left cells, corresponding to high ability students at low quality colleges, account for 14.1% in 1979 and 13.4% in 1997.[21] Thus, we find quite substantial departures from academic match in both cohorts. Viewed longitudinally, our data (perhaps surprisingly given the recent policy focus on match) reveal only a small, though meaningful, increase in assortative matching.[22]

3.5. Outcomes

We examine five educational outcomes: graduation within four or eight years of starting, obtaining a Science, Technology, Engineering and Mathematics (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. 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

---

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

[20] Rodriguez (2015) and House (2017) analyze the various measures of mismatch and find that the measure matters for the amount of mismatch measured.

[21] 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 et al. (2010, Table 1).

[22] Statistically, a chi-squared test strongly 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 et al. (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 mismatch, and their inclusion of two-year schools stymie a detailed accounting of the differences between their finding and ours.

last reported major(s) prior to graduation and code majors as STEM or non-STEM using the (uncontroversial) system in Arcidiacono et al. (2016).[23] Some restless students transfer more than once; we code our transfer variable based on the first observed transfer and only count transfers that change college quality by at least five percentiles (up or down) on 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.[24] We look relative to the start of college rather than the end because we want to capture the opportunity cost of college, because academic match may 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 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. Furthermore, comparisons with the information on job spells suggest that a non-trivial fraction of the zeros represent measurement error.

Table 2 presents summary statistics for these outcomes for our sample. The 1997 cohort has a higher eight-year graduation rate (65% versus 55%) but a lower four-year graduation rate (24% versus 32%) reflecting the increased time to degree over recent decades documented in Bound et al. (2010). 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 et al. (2014), we find a great deal of transfer behavior among the NLSY-97 students, about 26% transfer at least once. Earnings increase both over the life cycle within cohorts and between the 1979 and 1997 cohorts.

4. Econometric framework

[23] See Table A-1 of their on-line appendix.

[24] 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.

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) $$\Pr(Y_i = 1 \mid A_i, Q_i, X_i) = \Phi(\beta_0 + \beta_p(A_i, Q_i) + \beta_X X_i)$$

In (1), Y denotes the binary outcome of interest, A denotes student ability, Q denotes college quality, $\beta_p\left(A_i, Q_i\right)$ 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) $$\beta_P(A_i, Q_i) = \beta_1 A_i + \beta_2 A_i^2 + \beta_3 Q_i + \beta_4 Q_i^2 + \beta_5 A_i Q_i + \beta_6 A_i^2 Q_i + \beta_7 A_i Q_i^2$$

We chose this specification after a fairly rigorous round of statistical testing.[25] The polynomial in ability and quality becomes non-parametric once we promise to include higher-order terms (but not too quickly!) on those happy occasions when our sample size increases. This approach therefore represents a partially linear model in which we non-parametrically estimate the effects of ability and quality while conditioning parametrically on the other variables.

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-

[25] We conducted a series of specification tests, with and without additional covariates, starting with higher-order terms in ability, quality, and their interactions and gradually moving towards more parsimonious specifications using the NLSY-97 data. For most 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 between ability and quality and report tests of the joint significance of these interaction terms in our results. We use this same specification for the NLSY-79 cohort data to make the two analyses symmetric.

quartile mismatch. In practice, the two estimators tell the same substantive story; see Appendix Table 5 for the estimates from the second approach for a subset of our outcomes.

5. 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. To accomplish this, we need two things. 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 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 at the time of high school completion, and other social factors. We list these variables and describe their construction in detail in Appendix Table 2. Our preferred specification includes the first three sets of covariates; the fourth set provides a test of sorts, described below, for our identification strategy. Note that we do not ever 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, 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 for parents, so number of children may reflect both resources and preferences. We expect that our census tract income and education variables 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. More generally, many of our conditioning variables affect both college quality choice and outcomes in the broader literature, including our earlier study, Dillon and Smith (2017).

The second way to think about our covariate sets asks whether the marginal covariates make any difference to the estimates. In the framework of Heckman and Navarro (2004), 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 our covariates. Oster (2013) cautions that this line of reasoning related to coefficient stability means little if the newly added variables do not capture any (conditional) variation in the dependent variable. We perform such analyses below by adding sets of 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 and pricing strategies provide plausibly exogenous variation in the budget sets facing students and their parents. 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 or no idea about how to choose a college, provides aid and comfort 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, the descriptive and ethnographic literature, such as Roderick et al. (2008) indicates that many students explicitly choose among colleges for reasons unlikely to be strongly related to outcomes, such as the football team or the presence of high school friends.

What can we say about any remaining selection bias? Putting aside match for the moment, two worries usually arise about 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 unobserved bits 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 (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. 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.

6. Effects of college quality and ability of college outcomes and earnings

6.1 Graduation rates

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

Our first key finding consists of substantively meaningful and statistically significant main effects of both college quality and student ability on graduation within eight 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 eight years by 3.05 (10 x 0.305) percentage points. Along the same lines, for a college starter in the NLSY-97 cohort of median ability, increasing the quality of first college attended by 10 percentiles increases the probability of graduating within five years by 3.69 percentage points. These strong main effects have remained broadly similar over the 20 years between the NLSY-79 and NLSY-97 cohorts, although ability plays a slightly larger role for the older cohort and college quality a slightly smaller one.

If ability and quality have only independent effects, then we would expect a uniform effect of college quality across students of different ability levels. In contrast, in a world of substantively important academic match effects, the effect of quality should 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 3 plots the average derivative of college quality. It shows that, at the 25th, 50th and 75th percentiles of the ability distribution, the probability of graduating within eight years increases almost linearly in college quality for the NLSY-79 cohort. The NLSY-97 cohort shows some evidence of a weak overmatching effect: the probability of graduating within 8 years does not increase with college quality above the 60th percentile of colleges for students in the 25th percentile of ability. However, as shown in Table 3, this concavity is not statistically significant. Overall, the patterns predicted by the mismatch hypothesis simply do not appear strongly in our data.

We can quantify the lack of evidence for mismatch 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 chi-squared statistics and p-value from this test appear in the bottom two rows of Table 3. The p-values of 0.224 for the NLSY-97 cohort and 0.303 for the NLSY-79 cohort indicate that the restrictions implicit in the main-effects-only model cause little trouble for the data in either cohort.

Second, we can look to Tables 6A and 6B, which compare the observed completion rate with the completion rate implied by our model in a counterfactual world of perfect matching. We obtain this value by predicting degree completion for every observation with their college quality percentile recoded to match their ability percentile. Based on our model, we find that degree completion rises less than one percentage point if we eliminate academic mismatch, moving from 64.1% to 64.7% for the younger cohort and from 55.4% to 56.0% 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 quality colleges. Chingos (2012) performs a similar calculation and also finds virtually no effect of resorting students.

This net effect of eliminating academic mismatch masks large improvements in outcomes from moving some students to higher-quality colleges. The last columns of Tables 6A and 6B 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 by about eight percentage points in both cohorts, to 72.5% for the younger cohort and 63.5% for the older 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 Bound et al. (2010). The average derivative estimates for completion in four years resemble those for completion in eight years in sign, all positive, but differ in showing a clear pattern of complementarity between student ability and college quality, particularly for the NLSY-97 cohort. In both cohorts, student ability has a statistically significant effect on the probability of completion in four years at all colleges, but student ability has a substantively larger effect at high quality colleges than at low quality colleges. This pattern comports with a version of academic match in which student ability complements college quality in the production of academic outcomes. The most striking difference between the two cohorts remains the reversal in the relative importance of student ability and college quality, with college quality playing a smaller role for the older cohort.

We do not want to overemphasize this suggestion of complementarity, given that we still cannot reject the null of no interactions in either cohort and given that the substantive implications for resorting remain small. Nonetheless, we note the clear and substantively meaningful changes in the average derivatives relative to the eight-year completion case. Less able students receive somewhat less benefit from higher quality colleges when considering four-year graduation rates, but the same benefit as more prepared students when considering eight-year graduation rates.

6.2 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, for example, react to the experience of mismatch by changing their major, as in Ariciacono et al. (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 eight-year, completion rates. Table 4 presents our estimates of 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 25^{th}, 50^{th}, and 75^{th} 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 0.783 and 0.397 for the NLSY-97 and NLSY-79 cohorts. Our estimates predict that resorting to perfect academic match would change STEM degree completion by less than 0.5 percentage points in either cohort.

To compare our results to those in Arcidiacono et al. (2016), we calculated the quality percentile for each of the UC schools they consider. For the NLSY-97 cohort, the relevant comparison to their sample, our estimates imply that moving from UC-Santa Cruz (UCSC) to UC-Berkeley would increase the probability of graduating with a STEM degree by 0.98, 1.65, and 0.77 for students at the 25th, 50th, and 75th percentiles of ability, respectively.[26] In contrast, in their Table 6 they find that overall, initial STEM majors would *decrease* their probability of obtaining a STEM degree by 2.3 percentage points by moving from UCSC to UC-Berkeley. Their overall result conceals substantial heterogeneity in the form of large decreases for less prepared students and large increases for more prepared students. One could reconcile the two sets of findings by arguing that, for this outcome, match matters more in California than elsewhere or more for students who initially plan on completing a STEM major, but we have no evidence for such a claim.[27]

Students may react to learning they have made a poor initial match by transferring to another school. Again, this mid-course adjustment could delay graduation beyond four years. We find evidence consistent with match effects when looking at transfer behavior in the NLSY-97 cohort. The first column of estimates in Table 4 corresponds to model (1) with transfer up as the dependent variable, while the second column of estimates corresponds to transfer down.[28] Increasing a student's ability percentile by 10 percentage points raises the probability that she will transfer to a higher quality college by 1.6 percentage points if 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 first pattern reflects students preferentially transferring to better matches, while the second 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 and the effect of ability on transferring down is particularly strong at relatively high quality colleges. In general, increasing the quality of the first college attended raises the probability that students will eventually transfer down, particularly for students farther down the ability distribution. Taken together, these transfer

[26] UCSC to UC-Berkeley is a move of 35 = (96-61) percentile points: 35 x 0.028 = 0.98, 35 x 0.047 = 1.65, and 35 x 0.022 = 0.77.

[27] We do not have good data on pre-college major preferences in our samples, and too few observations to run estimates separately for California.

[28] 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.

results provide some support for the mismatch hypothesis and also support for 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 4, the p-values equal 0.278 and 0.296 for transferring to a higher and lower quality college, respectively. In Table 6A, we predict that eliminating initial mismatch would modestly decrease transfers to a higher quality college and modestly increase transfers to lower quality colleges. Eliminating initial mismatch substantially decreases the transfer probability for students who presently severely mismatch in their initial college choice 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.

6.3 Earnings

Table 5A presents our estimates for the effects of ability and college quality on average annual earnings for the NLSY-97 cohort 2-3, 6-7, and 10-11 years after students start college. In years 2-3, both college quality and student ability have negative effects on annual average earnings; for example, 2-3 years after starting college a student at the 50th percentile of ability earns \$303 less per year for each 10 percentile point increase in the quality of first college attended. Students at less selective colleges are more likely to have dropped out of college 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, each 10 percentile point increase in the quality of the first college is associated with an additional \$1,552 of annual earnings. While our average derivative estimates carry with them frustratingly large standard errors, we find strong evidence that college quality increases future earnings throughout the ability distribution. The estimates for years 6-7 after college start, perhaps not surprisingly, lie in between those for 2-3 years and for 10-11 years.

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. The average derivative of earnings with respect to student ability has a much larger value, around $110 per percentile point, 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 $113 per percentile point at the 25th percentile of ability to about $185 per percentile point at the 75th percentile of ability. At the same time, we cannot reject the null of no interaction effects (as well as other nulls involving much larger interaction effects).

Table 6A shows that re-sorting the students in our data so as to eliminate mismatch in the NLSY-97 cohort would increase mean earnings by about $1,350 10-11 years after beginning college. While we do not put much stock in the particular number 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 clearly do not shout out harmful effects of college quality, on average, for low ability students. 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 average derivative estimates for the NLSY-79 cohort, presented in Table 5B, resemble those for the NLSY-97 cohort in their sign and, for the most part, their magnitudes and statistical significance. At 2-3 years after college start, student ability and college quality both have statistically significant negative effects on annual earnings. By 10-11 years after starting college, both student ability and college quality have a positive relationship with earnings. Two key differences emerge. First, 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 $107, compared to just $29 for the later cohort. Second, both 2-3 and 10-11 years after college start, we find far less variation in the effects of student ability across the college quality distribution, and vice versa, and the variation follows the opposite pattern. In this earlier cohort, the effects of ability decrease (in absolute value) with college quality and the small college quality effects decrease with ability. Again, as in the NLSY-97 cohort, we cannot reject the null of zero coefficients on the interaction terms; unlike the later cohort, we predict that resorting students so that all attend their college match would change mean earnings by less than $200.

The NLSY-79 cohort, now into its fifties, allows us to examine earnings outcomes for several decades after college start. These results appear in Table 7 and Figure 3. We highlight four features of these estimates. First, the data provide large, positive, and generally statistically significant estimates of the effect of student ability and college quality at all durations from 10-31 years after college start. Thus even at long durations we can clearly rule out negative average effects of overmatch. We can also rule out simple versions of the "college quality is just a signal" argument, as employer learning would surely have overwhelmed college quality effects over the horizons we consider.[29] Second, 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 about $2,900 at 10-11 years (or about $290 for a 10 percentile point increase in quality) to about $25,000 at 20-21 years to about $50,100 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. Third, a pattern consistent with complementarity between student ability and college quality, which was absent from 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 for this older cohort. 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.[30] The second two patterns we describe appear strongly in the point estimates but largely lack statistical significance at traditional levels.

6.4 Subgroups

We consider subgroups defined by sex and by parental education, where we partition the latter into "low" and "high" subgroups based whether or not at least one parent attended college. We think of parental education as a proxy for several things, including tastes for college (and college quality) as well as 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

[29] See Hershbein (2013) for more subtle signaling theories of college quality.
[30] All of the qualitative findings in Table 7 persist if we restrict ourselves to a rectangular sample.

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 African-Americans and Hispanics.

Tables 8 and 9 report the effects of student ability and college quality on earnings 10-11 years after starting college separately by subgroups.[31] 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, including two statistically insignificant negative estimated effects, and larger standard errors due to the smaller sample sizes. In the NLSY-97 sample, the pattern of complementarity between student ability and college quality is most apparent for the children of more educated parents. For the children of less-educated parents in both cohorts, ability and college quality appear to behave more like substitutes in determining early-career earnings. In both cohorts, we 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.

6.5 Identification

As promised in Section 5, we now consider some evidence regarding our identification strategy. Tables 10A-11B present estimates based on increasingly rich sets of conditioning variables for our two most important outcomes: degree attainment within eight 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 correspond to those in Appendix Table 2. The estimates in column (4) of each table correspond to those in Tables 3 and 5.

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. 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 et al. (2005) study of

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

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.

6.6 Comparing the NLSY-79 and NLSY-97 results

Our big picture stories apply to both cohorts: large amounts of overmatch and undermatch appear in the unconditional 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 match effects in time to degree, and no substantive or statistical evidence of match effects for degree completion or for STEM degree completion. This stability surprised us somewhat. In this section we briefly remark on three specific differences in the results: (1) a modest but not trivial reduction in mismatch between cohorts; (2) the relatively smaller role of college quality in determining outcomes for the NLSY-79; and (3) the timing of the suggestive evidence of complementarity for earnings, which appears at 10-11 years for NLSY-97 but only later for NLSY-79.

As we noted in Section 3.4, the small decrease in mismatch between the cohorts comports with some other evidence in the literature. We suspect that it results from ongoing reductions in the cost that students (especially high ability students) face in obtaining information about admissions criteria and 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 is an excellent proxy for first college attended, but memory lapses combined with the occasional transfer representing a large change in college quality could introduce some mis-measurement.

We do not have a single compelling explanation for the delayed onset of match effects in earnings for the NLSY-79 cohort. The differences between the two cohorts could result from effects of student ability and college quality that vary with student characteristics paired with changes in the composition of four-year college starters over time (even if the effects for each group remain constant); for example, college students in the NLSY-97 cohort are more female and have more educated parents on average. The subgroup estimates presented in Section 6.4 partially support this hypothesis: ability and college quality exhibit more complementarity for the children of more educated parents, at least in the younger cohort, but we did not find strongly different patterns for male and female students.

Changes in the demands placed on workers, driven by changes in production technology, might also generate some of the differences between the two cohorts. For example, rapid technological change could decrease the demand for specific learned skills and increase the rewards for adapting problem solving techniques to new situations as in Autor et al., (2003). If developing these flexible problem-solving skills requires both strong starting student preparation and intense college resources, then earnings in later calendar years, including 10 years after starting college for the NLSY-97, but only 20 or 30 years after starting college for the NLSY-79, would exhibit more complementarity between ability and college quality.

6.7 Comparisons with earlier 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.[32] 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 degree completion across all ability quartiles. In the latter case their estimates comport with the predictions of standard theories of academic mismatch. Several differences between the LS setup

[32] 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.

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 drastically from our own. Qualitatively, we reach similar conclusions. While they examine mismatch 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 match effects. 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 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 Black et al. (2005) shows impacts on log wages for

[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 2754 compared to ours of 2111.

[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 census tract conditioning variables, which also reduce the estimated effect of college quality.

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 both a "selection on observed variables" identification strategy that includes test scores and a strategy that 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, which they call their "self-revelation" model.[36] The first identification strategy finds 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.

7. 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 at both four and eight years after college start as well as later earnings. At the same time, we obtain only very modest evidence that the effects of college quality vary with student ability as suggested by various theories of match, except for the intermediate outcomes of transfer and on-time degree completion. We find some evidence of complementarity between student ability and college quality in the production of graduation in four years (but not eight) and in later earnings, particularly for the NLSY-97 cohort. To our surprise, and in contrast to some other papers in the literature, we do not find strong match effects on STEM degree completion (or, indeed, much in

[35] The earnings measure is actually 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 e.g. Hoxby (2009) for discussion.

the way of college quality main effects for this outcome). Combined with our finding that students of all abilities benefit on average from college quality these results suggest potential equity gains from moving less able students to higher quality colleges at some efficiency cost.

Viewed from a different angle, the strong complementarities between student ability and college quality implicit in the simple and compelling applied theory models in Rothschild and White (1995) and Sallee, Resch and Courant (2008), complementarities that justify the observed long-term increase in college match described in e.g. Hoxby (2009), leave only a modest evidentiary trail in our data. Thus, our findings suggest paying more attention to alternative models such as those outlined in Cunha (2009).

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. 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 undergraduate mismatch. 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, mismatch, particularly overmatch, might have very different overall effects in these contexts than in ours.

Finally, this paper considers only academic match. As noted in Smith (2008), other types of mismatch between students and their undergraduate institutions represent an important omission from most of the literature (all of it inside economics and much of it outside as well). Perhaps the most obvious concerns mismatch in terms of 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 form of mismatch, as does Tom Wolfe (2004) in his novel of college life entitled *I Am Charlotte Simmons*. Because mismatch on social class will likely correlate with academic mismatch, it represents a potentially confounding treatment in our context.

References

Alon, Sigal and Marta Tienda (2005): "Assessing the `Mismatch' Hypothesis: Differences in College Graduation Rates by Institutional Selectivity," *Sociology of Education*, 78, pp. 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), pp. 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, pp. 93-108.

Ann Arbor News (2008): "Academics and Athletics: A Four-Day Ann Arbor News Series on the University of Michigan." http://www.mlive.com/wolverines/academics/stories/

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): 525-562.

Arcidiacono, Peter, Esteban Aucejo, Andrew Hussey, and Kenneth Spenner (2013): "Racial Segregation Patterns in Selective Universities," *Journal of Law and Economics*, 56(4), pp. 1039-1060.

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), pp. 3-51.

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

Aucejo, Esteban (2012): "Explaining Cross-Racial Differences in the Educational Gender Gap," Unpublished manuscript, London School of Economics.

Autor, David, Frank Levy, and Richard Murnane (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, 118 (4), pp. 1279-1333.

Bastedo, Michael and Allyson Flaster (2014): "Conceptual and Methodological Problems in Research on College Undermatch," *Educational Researcher*, 42(3), pp. 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), pp. 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), pp. 1-38.

Black, Dan, Kermit Daniel, and Jeffrey Smith (2005): "College Quality and Wages in the United States," *German Economic Review*, 6(3), pp. 415-443.

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), pp. 99-124.

Black, Dan and Jeffrey Smith (2006): "Evaluating the Returns to College Quality with Multiple Proxies for Quality," *Journal of Labor Economics*, 24(3), pp. 701-728.

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), pp. 129-157.

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), pp. 126-153.

Card, David (1995): "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." In L. Christophides, E. Grant and R. Swidinsky, eds., *Aspects of Labour Market Behavior: Essays in Honor of John Vanderkamp*. Toronto: University of Toronto Press, pp. 201-222.

Cawley, John, James Heckman, and Edward Vytlacil (2001): "Three Observations on Wages and Measured Cognitive Ability," *Labor Economics*, 8(4), pp. 419-442.

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

Cunha, Flavio (2009): "Crossing the Finish Line: A Review Article," *Journal of Human Capital*, 3(4), pp. 354-378.

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), pp. 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*, 2002, 107(4), pp. 1491-1527.

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), pp. 323-358.

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

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

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

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, pp. 620-628.

Griliches, Zvi and Vidar Ringstad (1970): "Error-in-the-Variables Bias in Nonlinear Contexts," *Econometrica*, 38(2), pp. 368-370.

Heckman, James and Tim Kautz (2012): "Hard Evidence on Soft Skills," *Labour Economics* 19(1), pp. 451-464.

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), pp. 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," Unpublished manuscript, 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), pp. 717-724.

House, Emily (2017): "Finding the Fit: Exploring Postsecondary Undermatch in Tennessee." Unpublished doctoral dissertation, University of Michigan.

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

Hoxby, Caroline (2016): “Computing the Value-Added of American Post-Secondary Institutions,” Unpublished manuscript, Stanford University.

Hoxby, Caroline and Sarah Turner (2013): “Expanding College Opportunities for High-Achieving, Low Income Students,” SIEPR Discussion Paper 12-014.

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

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), pp. 299-332.

Lincove, Jane and Kalena Cortes (2016): “Match of Mismatch? Automatic Admissions and College Preferences of Low- and High-Income Students,” NBER Working Paper No. 22559.

Long, Mark (2008): “College Quality and Early Adult Outcomes,” *Economics of Education Review*, 27, pp. 588-602.

Long, Mark (2010): “Changes in the Returns to Education and College Quality,” *Economics of Education Review*, 29, pp. 338-347.

Loury, Linda Datcher and David Garman. 1995. “College Selectivity and Earnings,” *Journal of Labor Economics*, 13(2), pp. 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), pp. 368-391.

Neal, Derek and William Johnson (1996): “The Role of Premarket Factors in Black-White Wage Differences,” *Journal of Political Economy*, 104(5), pp. 869-895.

Oster, Emily (2013): “Unobservable Selection and Coefficient Stability: Theory and Validation,” NBER Working Paper No. 19054.

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

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

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, pp. 566-594.

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), pp. 573-586.

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: 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), pp. 1438-1457.

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

Smith, Jonathan, Matea Pender, and Jessica Howell (2013): "The Full Extent of Student-College Academic Undermatch," *Economics of Education Review*, 23, pp. 247-261.

Steeg, Jill, Jodi Upton, Patrick Bohn, and Steve Berkowitz (2008): "College Athletes [sic] Studies Guided Toward `Major in Eligibility'." USA Today, Updated 11/19/08. http://usatoday30.usatoday.com/sports/college/2008-11-18-majors-cover_N.htm

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

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

Zimmerman, Seth (2014): "The Returns to College Admission for Academically Marginal Students," *Journal of Labor Economics*, 32(4), pp. 711-754.

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

	Quality Quartiles				
Ability Quartiles	1st Quartile (lowest)	2nd Quartile	3rd Quartile	4th Quartile (highest)	Total
1st Quartile (lowest)	10.5 (41.9) [37.4]	6.9 (27.6) [26.3]	5.1 (20.3) [19.4]	2.6 (10.3) [13.1]	(100.0) [25.0]
2nd Quartile	8.2 (32.6) [29.1]	7.5 (30.1) [28.7]	5.6 (22.2) [21.3]	3.7 (15.0) [19.2]	(100.0) [25.0]
3rd Quartile	5.6 (22.3) [20.0]	7.1 (28.4) [27.1]	7.6 (30.4) [29.2]	4.7 (18.8) [24.1]	(100.0) [25.1]
4th Quartile (highest)	3.8 (15.2) [13.5]	4.7 (18.9) [17.9]	7.9 (31.6) [30.2]	8.5 (34.2) [43.6]	(100.0) [24.9]
Total	(28.2) [100.0]	(23.4) [100.0]	(25.9) [100.0]	(22.0) [100.0]	N=1,623

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

	Quality Quartiles				
Ability Quartiles	1st Quartile (lowest)	2nd Quartile	3rd Quartile	4th Quartile (highest)	Total
1st Quartile (lowest)	12.5 (49.8) [44.2]	6.6 (26.6) [24.7]	4.2 (16.7) [18.0]	1.7 (7.0) [8.0]	(100.0) [25.0]
2nd Quartile	7.6 (30.5) [27.1]	8.1 (32.4) [30.2]	5.3 (21.1) [22.7]	4.0 (16.0) [18.4]	(100.0) [25.0]
3rd Quartile	5.1 (20.3) [18.0]	6.8 (27.1) [25.2]	7.2 (28.7) [30.8]	6.0 (23.9) [27.5]	(100.0) [25.0]
4th Quartile (highest)	3.0 (12.1) [10.7]	5.3 (21.3) [19.8]	6.6 (26.4) [28.5]	10.0 (40.1) [46.1]	(100.0) [25.0]
Total	(28.2) [100.0]	(26.9) [100.0]	(23.2) [100.0]	(21.7) [100.0]	N=2,111

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.

Table 2: Summary of College Outcomes

	NLSY-97	NLSY-79
4-year college starters	2,240	1,623
Graduate within 8 years	64%	55%
In 4 years or less	31%	30%
In 5 years	20%	14%
In 6-8 years	13%	11%
Complete STEM degree	13%	15%
Transfer	26%	
Transfer to a higher quality college	7%	
Transfer to a lower quality college	16%	
Transfer to a similar or unknown college	3%	
Did any work 2-3 years after entering	91%	95%
Men	89%	94%
Women	93%	96%
Did any work 10-11 years after entering	92%	96%
Men	93%	97%
Women	92%	95%
Average earnings 2-3 years after entering	$9,343	$7,481
Men	$10,680	$8,287
Women	$8,266	$6,720
Average earnings 10-11 years after entering	$42,942	$40,338
Men	$48,799	$46,348
Women	$38,017	$34,376

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.

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

		NLSY-97 Cohort		NLSY-79 Cohort	
		Within 8 years	Within 4 years	Within 8 years	Within 4 years
$\partial Outcome / \partial A$	Q = p25	0.305*	0.119*	0.464*	0.307*
		(0.056)	(0.060)	(0.066)	(0.081)
	Q = p50	0.222*	0.237*	0.428*	0.395*
		(0.058)	(0.055)	(0.070)	(0.067)
	Q = p75	0.274*	0.359*	0.429*	0.415*
		(0.061)	(0.054)	(0.068)	(0.062)
$\partial Outcome / \partial Q$	A = p25	0.325*	0.220*	0.229*	0.064
		(0.060)	(0.068)	(0.069)	(0.063)
	A = p50	0.369*	0.329*	0.297*	0.162*
		(0.055)	(0.055)	(0.073)	(0.071)
	A = p75	0.282*	0.430*	0.182*	0.197*
		(0.064)	(0.048)	(0.077)	(0.071)
Observations		2,111	2,111	1,623	1,618
R-squared		0.209	0.188	0.143	0.125
Test: interaction=0, Chi2/F		4.372	4.489	3.818	3.436
Pr(interaction=0)		0.224	0.213	0.282	0.329

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%.

Table 4: Effect of College Quality and Ability on Intermediate Outcomes

		NLSY-97 Cohort			NLSY-79
		Transfer up	Transfer down	STEM degree	STEM degree
$\partial Outcome / \partial A$	Q = p25	0.160*	-0.083*	0.291*	0.419*
		(0.064)	(0.038)	(0.069)	(0.085)
	Q = p50	0.061	-0.031	0.253*	0.349*
		(0.038)	(0.053)	(0.051)	(0.061)
	Q = p75	0.017	-0.055	0.244*	0.318*
		(0.012)	(0.061)	(0.042)	(0.055)
$\partial Outcome / \partial Q$	A = p25	-0.119*	0.091	0.028	0.022
		(0.029)	(0.047)	(0.032)	(0.020)
	A = p50	-0.151*	0.079*	0.047	0.077
		(0.039)	(0.038)	(0.042)	(0.047)
	A = p75	-0.258*	0.110*	0.022	0.045
		(0.062)	(0.028)	(0.055)	(0.070)
Observations		2,082	2,082	2,107	1,583
R-squared		0.108	0.037	0.105	0.156
Test: interaction=0, Chi2		3.847	3.694	1.076	2.964
Pr(interaction=0)		0.278	0.296	0.783	0.397

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%.

Table 5A: Effect of College Quality and Ability on Earnings, NLSY-97

		Year 2-3	Year 6-7	Year 10-11
$\partial Earnings / \partial A$	Q = p25	-2,411	-1,880	3,885
		(1,273)	(2,252)	(3,925)
	Q = p50	-4,117*	392	2,851
		(1,219)	(2,724)	(4,308)
	Q = p75	-7,525*	958	11,035*
		(1,632)	(3,141)	(4,298)
$\partial Earnings / \partial Q$	A = p25	751	7,980*	11,254*
		(1,602)	(2,817)	(3,982)
	A = p50	-3,025*	9,387*	15,517*
		(1,337)	(2,524)	(4,027)
	A = p75	-4,386*	10,815*	18,459*
		(1,393)	(2,438)	(4,040)
Observations		1,895	1,942	1,744
R-squared		0.147	0.078	0.137
Test: interaction=0, F		2.579	0.446	1.306
Pr(interaction=0)		0.052	0.720	0.271

Table 5B: Effect of College Quality and Ability on Earnings, NLSY-79

		Year 2-3	Year 6-7	Year 10-11
$\partial Earnings / \partial A$	Q = p25	-3,716*	3,488	14,570*
		(1,284)	(2,967)	(4,538)
	Q = p50	-3,200*	4,276	10,756*
		(1,224)	(2,989)	(4,627)
	Q = p75	-2,081	6,559*	10,090*
		(1,106)	(2,680)	(4,073)
$\partial Earnings / \partial Q$	A = p25	-3,656*	1,544	7,207*
		(1,178)	(2,443)	(3,385)
	A = p50	-4,280*	3,248	4,717
		(1,320)	(2,956)	(4,202)
	A = p75	-1,862	4,909	2,933
		(1,185)	(2,822)	(4,216)
Observations		1,346	1,327	1,323
R-squared		0.100	0.099	0.166
Test: interaction=0, F		2.864	0.527	0.596
Pr(interaction=0)		0.036	0.663	0.618

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.

Table 6A: Counterfactuals from Re-assigning Students to Colleges, NLSY-97

	Actual outcome	If all students attended a matched college	If all students attended a 90th pctile college
Graduate within 8 years	64.1	64.7	72.5
Graduate within 4 years	31.0	33.7	44.3
Major in STEM	14.3	14.2	14.8
Transfer to a higher quality college	7.5	6.8	0.8
Transfer to a lower quality college	16.6	17.0	17.4
Earnings 6-7 years after starting college	28,694	29,174	33,149
Earnings 10-11 years after starting college	42,620	43,979	49,500

Table 6B: Counterfactuals from Re-assigning Students to Colleges, NLSY-79

	Actual outcome	If all students attended a matched college	If all students attended a 90th pctile college
Graduate within 8 years	55.4	56.0	63.5
Graduate within 4 years	30.5	32.0	35.3
Major in STEM	15.0	15.4	14.5
Earnings 6-7 years after starting college	26,190	26,739	26,800
Earnings 10-11 years after starting college	40,338	40,166	42,014

Counterfactual outcomes are calculated using the estimates reported in Tables 3-5. Annual earnings are in 2010 dollars.

Table 7: Long-run Effect of College Quality and Ability on Earnings, NLSY-79

		Year 10-11	Year 20-21	Year 30-31
$\partial Earnings / \partial A$	Q = p25	14,570*	33,122*	26,280
		(4,538)	(12,444)	(14,926)
	Q = p50	10,756*	22,258*	36,348*
		(4,627)	(10,998)	(14,807)
	Q = p75	10,090*	30,114*	46,433*
		(4,073)	(9,798)	(14,317)
$\partial Earnings / \partial Q$	A = p25	7,207*	27,735*	28,100*
		(3,385)	(10,244)	(13,381)
	A = p50	4,717	38,144*	44,636*
		(4,202)	(12,412)	(16,150)
	A = p75	2,933	25,003*	50,145*
		(4,216)	(12,326)	(14,724)
Observations		1,323	1,034	895
R-squared		0.166	0.192	0.202
Test: interaction=0, F		0.596	2.497	1.156
Pr(interaction=0)		0.618	0.058	0.326

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.

Table 8: Effects on Earnings 10-11 Years after Starting College, by Parents' Education

		NLSY-97 Cohort		NLSY-79	
		H.S. only	Some college	H.S. only	Some college
$\partial Outcome / \partial A$	Q = p25	3,924	2,449	15,189*	10,014
		(6,973)	(4,467)	(5,206)	(7,027)
	Q = p50	611	1,398	14,086*	8,206
		(8,568)	(4,789)	(5,660)	(6,806)
	Q = p75	-3,409	9,445*	12,237*	8,586
		(8,862)	(4,806)	(5,532)	(6,141)
$\partial Outcome / \partial Q$	A = p25	20,223*	13,174*	13,209*	3,322
		(6,888)	(4,973)	(3,674)	(5,736)
	A = p50	10,467	18,465*	14,517*	56
		(8,232)	(4,728)	(6,073)	(5,791)
	A = p75	11,707	20,272*	10,561	2,015
		(8,427)	(4,660)	(6,138)	(6,319)
Observations		426	1,297	655	646
R-squared		0.251	0.126	0.272	0.141
Test: interaction=0, F		0.645	1.080	0.432	0.528
Pr(interaction=0)		0.587	0.356	0.730	0.663

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.

Table 9: Effects on Earnings 10-11 Years after Starting College, by Sex

		NLSY-97 Cohort		NLSY-79	
		Men	Women	Men	Women
$\partial Outcome / \partial A$	Q = p25	2,740	5,773	11,718	17,490*
		(6,967)	(4,613)	(6,666)	(6,381)
	Q = p50	-443	6,165	6,354	15,738*
		(7,105)	(5,117)	(6,904)	(5,885)
	Q = p75	9,322	10,747*	7,048	15,279*
		(7,408)	(4,906)	(6,443)	(5,078)
$\partial Outcome / \partial Q$	A = p25	12,779	9,571*	4,330	7,592
		(6,798)	(4,716)	(5,897)	(4,231)
	A = p50	17,369*	11,923*	-462	7,695
		(6,831)	(4,806)	(6,587)	(5,244)
	A = p75	20,068*	14,315*	490	5,679
		(7,329)	(4,788)	(6,117)	(5,995)
Observations		770	974	642	681
R-squared		0.109	0.147	0.151	0.123
Test: interaction=0, F		0.779	0.337	1.090	0.059
Pr(interaction=0)		0.506	0.799	0.353	0.981

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.

Table 10A: Covariate Set Comparisons for Graduating within 8 Years, NLSY-97

		(1)	(2)	(3)	(4)	(5)
∂Outcome / ∂A	Q = p25	0.332*	0.351*	0.299*	0.305*	0.281*
		(0.061)	(0.057)	(0.057)	(0.056)	(0.056)
	Q = p50	0.217*	0.265*	0.218*	0.222*	0.213*
		(0.060)	(0.057)	(0.059)	(0.058)	(0.058)
	Q = p75	0.294*	0.315*	0.265*	0.274*	0.253*
		(0.063)	(0.060)	(0.061)	(0.061)	(0.061)
∂Outcome / ∂Q	A = p25	0.493*	0.403*	0.354*	0.325*	0.327*
		(0.057)	(0.059)	(0.058)	(0.060)	(0.058)
	A = p50	0.539*	0.445*	0.393*	0.369*	0.366*
		(0.054)	(0.054)	(0.053)	(0.055)	(0.054)
	A = p75	0.425*	0.356*	0.307*	0.282*	0.287*
		(0.067)	(0.063)	(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		2,111	2,111	2,111	2,111	2,111
R-squared		0.128	0.178	0.204	0.209	0.221
Test: interaction=0, Chi2		7.037	4.073	4.047	4.372	3.255
Pr(interaction=0)		0.071	0.254	0.256	0.224	0.354

Table 10B: Covariate Set Comparisons for Graduating within 8 Years, NLSY-79

		(1)	(2)	(3)	(4)	(5)
∂Outcome / ∂A	Q = p25	0.501*	0.521*	0.466*	0.464*	0.457*
		(0.060)	(0.060)	(0.066)	(0.066)	(0.066)
	Q = p50	0.453*	0.470*	0.430*	0.428*	0.411*
		(0.062)	(0.063)	(0.069)	(0.070)	(0.070)
	Q = p75	0.494*	0.485*	0.435*	0.429*	0.407*
		(0.058)	(0.060)	(0.068)	(0.068)	(0.069)
∂Outcome / ∂Q	A = p25	0.248*	0.261*	0.239*	0.229*	0.255*
		(0.066)	(0.065)	(0.066)	(0.069)	(0.068)
	A = p50	0.338*	0.340*	0.312*	0.297*	0.310*
		(0.072)	(0.071)	(0.070)	(0.073)	(0.071)
	A = p75	0.235*	0.215*	0.198*	0.182*	0.186*
		(0.075)	(0.074)	(0.075)	(0.077)	(0.076)
Additional ability measures			Yes	Yes	Yes	Yes
Demographics and family				Yes	Yes	Yes
Neighborhood					Yes	Yes
Additional covariates						Yes
Observations		1,623	1,623	1,623	1,623	1,623
R-squared		0.098	0.120	0.138	0.143	0.162
Test: interaction=0, Chi2		4.259	4.310	4.003	3.818	3.909
Pr(interaction=0)		0.235	0.230	0.261	0.282	0.271

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%. Covariate sets are described in Appendix Table 2.

Table 11A: Effects on Earnings 10-11 Years after Start, NLSY-97

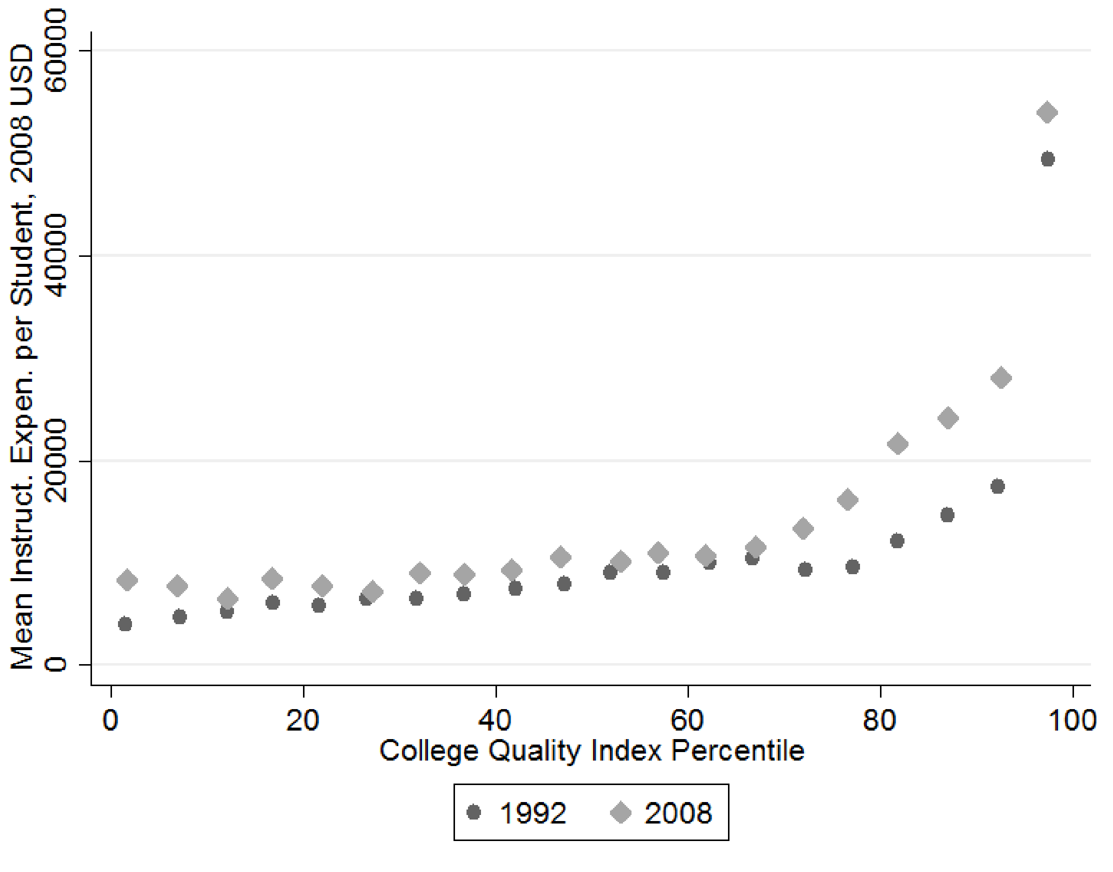
		(1)	(2)	(3)	(4)	(5)
$\partial Earnings / \partial A$	Q = p25	3,598	3,967	3,306	3,885	4,220
		(3,638)	(3,667)	(3,866)	(3,925)	(3,959)
	Q = p50	3,083	3,695	2,218	2,851	2,682
		(4,130)	(4,065)	(4,259)	(4,308)	(4,386)
	Q = p75	12,980*	13,754*	10,608*	11,035*	11,056*
		(4,188)	(4,259)	(4,287)	(4,298)	(4,324)
$\partial Earnings / \partial Q$	A = p25	14,919*	13,525*	13,011*	11,254*	10,766*
		(4,014)	(4,081)	(3,923)	(3,982)	(4,008)
	A = p50	20,700*	19,616*	17,370*	15,517*	15,084*
		(4,109)	(4,095)	(3,960)	(4,027)	(4,070)
	A = p75	24,327*	23,325*	20,370*	18,459*	17,677*
		(3,948)	(4,001)	(3,959)	(4,040)	(4,041)
Additional ability measures			Yes	Yes	Yes	Yes
Demographics and family				Yes	Yes	Yes
Neighborhood					Yes	Yes
Additional covariates						Yes
Observations		1,744	1,744	1,744	1,744	1,744
R-squared		0.076	0.079	0.130	0.137	0.141
Test: interaction=0, F		1.715	1.740	1.369	1.306	1.370
Pr(interaction=0)		0.162	0.157	0.250	0.271	0.250

Table 11B: Effects on Earnings 10-11 Years after Start, NLSY-79

		(1)	(2)	(3)	(4)	(5)
$\partial Earnings / \partial A$	Q = p25	20,773*	17,430*	13,623*	14,570*	14,227*
		(3,949)	(4,284)	(4,483)	(4,538)	(4,669)
	Q = p50	15,970*	12,388*	9,367*	10,756*	10,168*
		(4,111)	(4,490)	(4,632)	(4,627)	(4,749)
	Q = p75	16,211*	11,929*	9,016*	10,090*	9,752*
		(3,626)	(3,984)	(4,136)	(4,073)	(4,281)
$\partial Earnings / \partial Q$	A = p25	10,098*	11,421*	10,349*	7,207*	7,005*
		(3,455)	(3,399)	(3,286)	(3,385)	(3,378)
	A = p50	8,054	9,174*	7,943	4,717	4,658
		(4,328)	(4,213)	(4,078)	(4,202)	(4,117)
	A = p75	5,855	6,204	5,993	2,933	2,765
		(4,234)	(4,108)	(4,077)	(4,216)	(4,233)
Additional ability measures			Yes	Yes	Yes	Yes
Demographics and family				Yes	Yes	Yes
Neighborhood					Yes	Yes
Additional covariates						Yes
Observations		1,323	1,323	1,323	1,323	1,323
R-squared		0.077	0.096	0.157	0.166	0.180
Test: interaction=0, F		0.709	0.780	0.680	0.596	0.611
Pr(interaction=0)		0.547	0.505	0.564	0.618	0.608

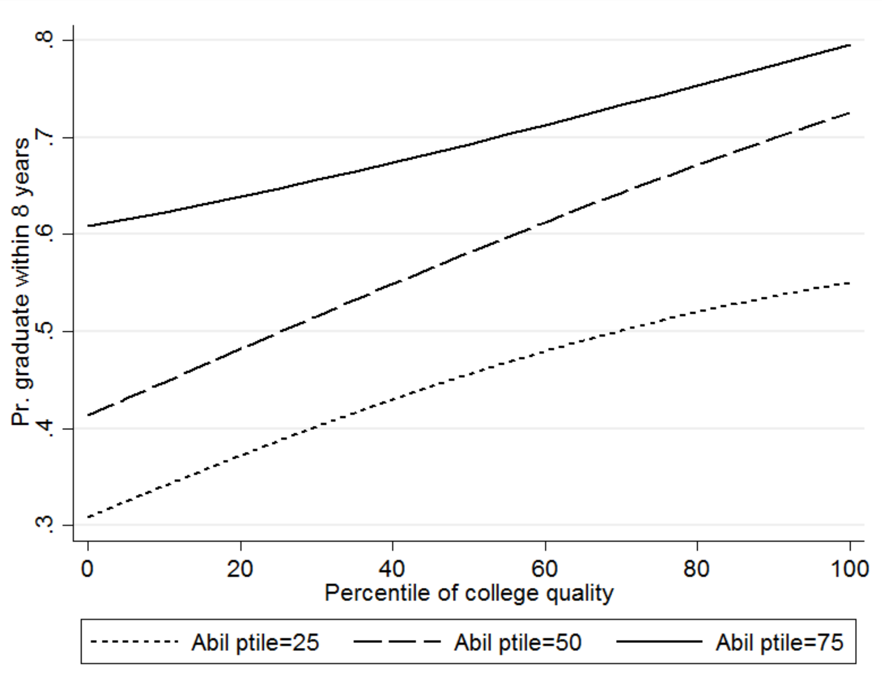
Mean marginal effects are calculated from the coefficients of a polynomial of ability and college quality as described in the text. Standard errors in parentheses. * indicates that the estimated effect is statistically significant at 5%. Covariate sets are described in Appendix Table 2. Annual earnings are in 2010 dollars.

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

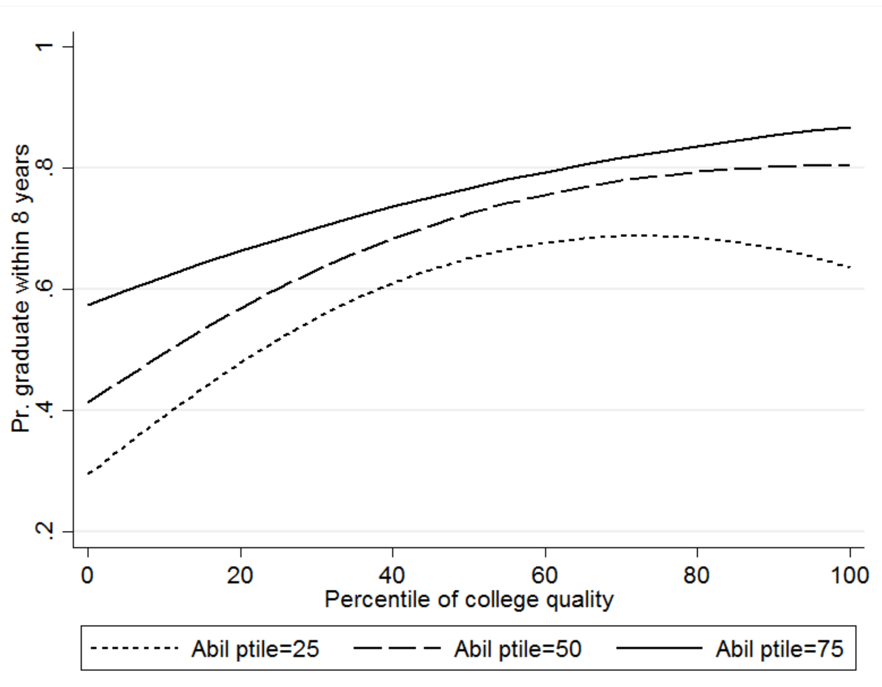


Instructional expenditure per student from iPEDS and U.S. News and World Report. College quality indices calculated as described in the text.

Figure 2: Effect of College Quality on Graduating within 8 Years
a. NLSY-79 Cohort



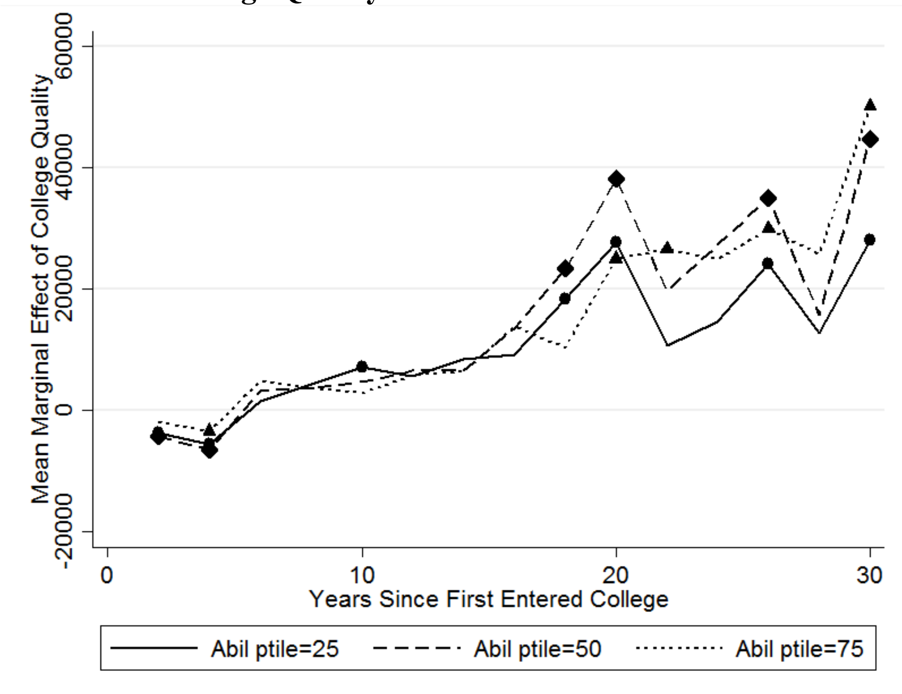
b. NLSY-97 Cohort



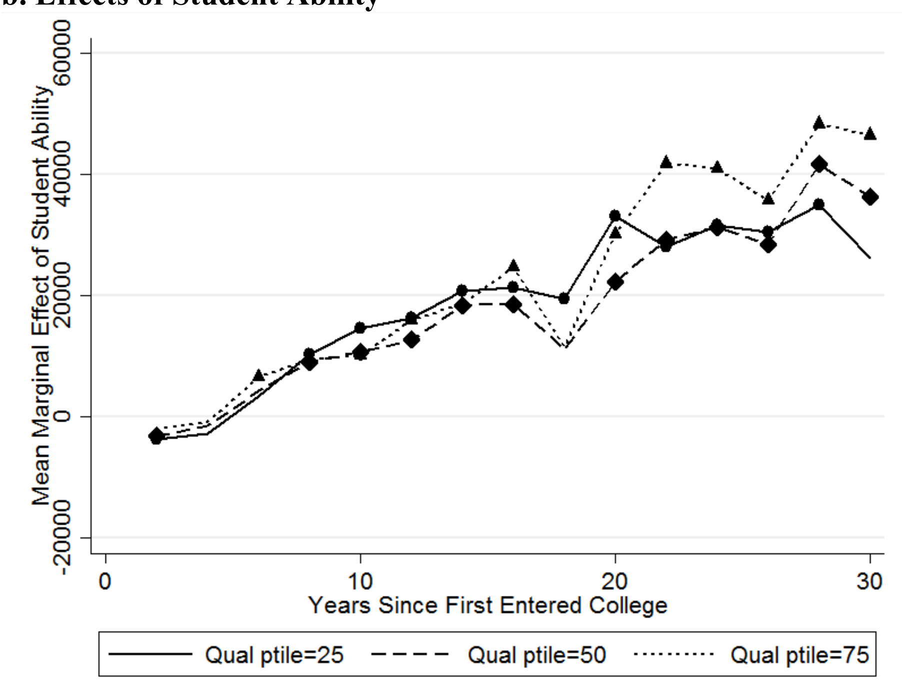
Projected from estimates summarized in Table 3.

Figure 3: Long-term Earnings Estimates for 79

a. Effects of College Quality



b. Effects of Student Ability



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 are presented in Table 7.

Consequences of Mismatch Appendix Tables
January 3, 2017

Appendix Table 1: Sample

	NLSY 1979 Cohort	NLSY 1997 Cohort
Total Observations	5,970	8,984
Graduated HS or received GED	5,103	7,845
Started at a 4-year college by age 21	2,007	2,891
Interviewed at least 5 years after starting	2,007	2,782
Of eligible 4-year college starters:		
Missing 4-year college quality index	*258*	*101*
Has quality, but missing ability	*126*	*570*
Analysis sample	**1,623**	**2,111**

The NLSY97 cohort is restricted to students who first started a four-year college before age 21 to enable us to measure earnings 10 years after college for all starters as of the most recent survey wave (2013/14). The NLSY79 cohort is restricted in the same way for symmetry. We further restrict the NLSY79 sample to respondents who were 18 or younger in 1979, to ensure at least one pre-college interview. In both waves, we also restrict the sample to individuals who give at least one interview five or more years after starting college to minimize ambiguous graduation outcomes from respondents who were still in college as of their last interview. In both waves, we include both the cross-sectional and oversamples, appropriately weighted. However, we cannot include the military oversample and the poor non-black non-Hispanic oversamples from the NLSY79 cohort, because these groups were dropped from the sample in 1984 and 1990 respectively, too early to observe graduation outcomes and post-college earnings.

Appendix Table 2A: Principal Components of the 8 Test Sections of the ASVAB, 1979 Cohort

	1st Component	2nd Component	Unexplained variance
Eigenvalue	5.46	0.84	
Total variance explained	0.68	0.10	
Eigenvectors:			
General Science	0.37	-0.28	0.19
Arithmetic Reasoning	0.37	-0.17	0.21
Word Knowledge	0.38	-0.08	0.20
Paragraph Comprehension	0.37	0.05	0.26
Numerical Operations	0.33	0.51	0.18
Coding Speed	0.30	0.63	0.16
Mathematics Knowledge	0.37	-0.10	0.25
Mechanical Comprehension	0.32	-0.47	0.25

Note: Following Altonji, Bharadwaj, and Lange (2011), we adjust scores on each test component the age of the respondent when they took the test by calculating age-specific percentiles and then assigning each student the score that corresponds to their percentile for 16 year olds.

Appendix Table 2B: Principal Components of the 8 Test Sections of the ASVAB, 1997 Cohort

	1st Component	2nd Component	Unexplained variance
Eigenvalue	5.25	0.88	
Total variance explained	0.66	0.11	
Eigenvectors:			
General Science	0.37	-0.32	0.18
Arithmetic Reasoning	0.38	-0.05	0.23
Word Knowledge	0.37	-0.23	0.23
Paragraph Comprehension	0.38	-0.09	0.24
Numerical Operations	0.31	0.55	0.23
Coding Speed	0.28	0.62	0.25
Mathematics Knowledge	0.38	0.09	0.23
Mechanical Comprehension	0.34	-0.37	0.28

Note: Following Altonji, Bharadwaj, and Lange (2011), we convert the scores on the computer-adaptive test taken by the NLSY-97 cohort to their pen-and-paper score equivalents to allow direct comparison with the NLSY-79 scores. We also adjust scores on each test component the age of the respondent when they took the test by calculating age-specific percentiles and then assigning each student the score that corresponds to their percentile for 16 year olds.

Appendix Table 3: Principal Components of the College Quality Indices

	1992		2008	
	1st Component	Unexplained variance	1st Component	Unexplained variance
Eigenvalue	2.46		2.56	
Variance explained	61%		64%	
Eigenvectors:				
Mean SAT	0.56	56%	0.55	23%
Rejection rate	0.47	47%	0.46	47%
Faculty/Student ratio	0.45	45%	0.47	43%
Avg. faculty salaries	0.51	51%	0.52	31%

Calculated from the set of four-year colleges in IPEDS in each year that report all four college quality proxies (1,157 institutions in 1992, 1,346 in 2008). In both years, quality measures that are missing in IPEDS are filled in where possible using data from U.S. News and World Report.

Appendix Table 4A: **Starting Student Characteristics, NLSY-79**

		Initial College Choice			
	Attendees	1, lowest	2	3	4, highest
N	1,623	526	403	399	295
ASVAB 1 percentile	49%	38%	46%	55%	63%
ASVAB 2 percentile	49%	53%	49%	48%	47%
HS GPA, percentile	47%	41%	45%	48%	56%
SAT or ACT percentile	48%	34%	45%	52%	66%
Bad behavior index	0.4	0.4	0.4	0.4	0.4
Rated uncooperative	4%	5%	4%	4%	2%
Had sex before age 15	15%	19%	14%	12%	12%
Male	48%	48%	49%	48%	48%
White	4%	5%	2%	4%	5%
Black	12%	19%	10%	9%	9%
Hispanic	84%	77%	88%	88%	86%
Family income quartile 1	13%	19%	12%	10%	9%
Family income quartile 2	19%	26%	19%	11%	19%
Family income quartile 3	30%	28%	34%	32%	23%
Family income quartile 4	38%	26%	36%	46%	49%
Siblings	2.8	3.0	2.7	2.8	2.5
No parent completed HS	8%	14%	6%	6%	5%
At least one parent grad. HS	34%	39%	39%	30%	25%
One parent has some college	17%	18%	16%	19%	12%
One parent has degree	42%	29%	39%	46%	57%
Northeast region	23%	8%	18%	28%	45%
South region	33%	36%	39%	35%	18%
Midwest region	31%	45%	27%	25%	25%
West region	13%	11%	16%	12%	12%
Rural	20%	25%	21%	18%	14%
% Adults w/college in county	11%	10%	10%	12%	12%
Overweight	7%	11%	7%	6%	4%
Obese	2%	2%	2%	2%	1%
Religious observance, per yr.	36.3	42.0	33.5	36.9	31.0
Enriching home index	2.6	2.3	2.6	2.6	2.7
Contact with bio. mother	98%	98%	99%	98%	97%
Contact with bio. father	95%	95%	97%	95%	95%

This table describes the characteristics of students at each college quality quartile. For example, the "Male" row shows the percent of students attending each college type who are male. All results are weighted as described in the text. Ability percentiles are among 4-year college starters, with the ASVAB measures adjusted by age when taking the test. All characteristics are measured as of the respondents last year of high school, except as noted in Appendix Table 5.

Appendix Table 4B: **Starting Student Characteristics, NLSY-97**

	Attendees	Initial College Choice 1, lowest	2	3	4, highest
N	2,111	628	576	471	436
ASVAB 1 percentile	49%	35%	47%	55%	66%
ASVAB 2 percentile	49%	49%	50%	48%	52%
HS GPA, percentile	49%	38%	47%	53%	62%
SAT or ACT percentile	48%	31%	43%	53%	67%
Bad behavior index	0.4	0.5	0.4	0.4	0.4
Rated uncooperative	38%	42%	36%	37%	34%
Had sex before age 15	9%	15%	8%	8%	4%
Male	45%	46%	42%	43%	48%
White	78%	78%	77%	80%	78%
Black	12%	15%	15%	11%	7%
Hispanic	3%	4%	3%	3%	3%
Other non-white	6%	3%	5%	6%	12%
Family income quartile 1	12%	19%	11%	7%	9%
Family income quartile 2	21%	27%	20%	19%	15%
Family income quartile 3	29%	30%	31%	29%	27%
Family income quartile 4	38%	25%	39%	44%	49%
Siblings	2.2	2.3	2.2	2.1	2.2
No parent completed HS	3%	5%	3%	2%	2%
At least one parent grad. HS	19%	27%	22%	16%	10%
One parent has some college	27%	30%	26%	25%	25%
One parent has degree	51%	38%	49%	57%	64%
Northeast region	20%	15%	16%	21%	32%
South region	31%	33%	36%	31%	22%
Midwest region	32%	36%	29%	32%	30%
West region	17%	17%	19%	16%	16%
Rural	19%	30%	15%	20%	10%
% Adults w/college in county	21%	18%	20%	22%	23%
Overweight	11%	13%	11%	10%	8%
Obese	8%	10%	7%	8%	4%
Religious observance, per yr.	15.2	17.4	17.2	13.0	12.0
Enriching home index	2.1	1.9	2.1	2.3	2.3
Contact with bio. mother	99%	100%	99%	98%	99%
Contact with bio. father	97%	96%	97%	97%	97%

This table describes the characteristics of students at each college quality quartile. For example, the “Male” row shows the percent of students attending each college type who are male. All results are weighted as described in the text. Ability percentiles are among 4-year college starters, with the ASVAB measures adjusted by age when taking the test. All characteristics are measured as of the respondents last year of high school, except as noted in Appendix Table 5.

Appendix Table 5: Effect on College Completion and Earnings, Quartile Dummies

	NLSY-97		NLSY-79	
	Graduate within 8 years	Earnings, 10-11 years	Graduate within 8 years	Earnings, 10-11 years
ASVAB q1, Quality q2	0.132*	6,573*	-0.017	2,572
	(0.042)	(3,172)	(0.058)	(2,475)
ASVAB q1, Quality q3	0.146*	94	0.039	9,892*
	(0.051)	(2,904)	(0.065)	(4,165)
ASVAB q1, Quality q4	0.164*	11,147*	0.109	2,826
	(0.071)	(5,101)	(0.076)	(2,830)
ASVAB q2, Quality q1	0.129*	8,424*	0.020	5,467
	(0.041)	(2,534)	(0.058)	(3,068)
ASVAB q2, Quality q2	0.203*	8,002*	0.146*	5,764
	(0.042)	(3,040)	(0.061)	(3,106)
ASVAB q2, Quality q3	0.262*	8,203*	0.189*	13,812*
	(0.047)	(3,436)	(0.066)	(4,092)
ASVAB q2, Quality q4	0.418*	11,426*	0.256*	3,673
	(0.061)	(3,673)	(0.076)	(4,853)
ASVAB q3, Quality q1	0.102*	3,650	0.141*	4,448
	(0.048)	(3,662)	(0.067)	(2,946)
ASVAB q3, Quality q2	0.283*	3,357	0.224*	12,250*
	(0.044)	(2,746)	(0.064)	(3,110)
ASVAB q3, Quality q3	0.279*	9,163*	0.167*	8,582*
	(0.044)	(3,313)	(0.063)	(3,483)
ASVAB q3, Quality q4	0.369*	17,787*	0.340*	12,941*
	(0.048)	(3,850)	(0.075)	(4,481)
ASVAB q4, Quality q1	0.239*	4,026	0.338*	15,608*
	(0.058)	(3,131)	(0.082)	(4,285)
ASVAB q4, Quality q2	0.318*	9,223*	0.351*	11,390*
	(0.049)	(3,257)	(0.076)	(4,907)
ASVAB q4, Quality q3	0.324*	15,146*	0.364*	13,988*
	(0.047)	(3,794)	(0.065)	(3,390)
ASVAB q4, Quality q4	0.423*	22,247*	0.406*	15,928*
	(0.043)	(3,723)	(0.065)	(4,210)
Observations	2,111	1,744	1,623	1,323
R-squared	0.207	0.138	0.143	0.174
Test: constant slope, Chi2/F	5.959	1.849	7.125	1.271
Pr(constant slope)	0.744	0.055	0.624	0.248

Test is whether the differences in the effects of adjacent quality quartiles is constant across ability quartiles (aka A2Q2-A2Q1 = A1Q2-A1Q1)

Appendix Table 6: Description of Independent Variables

Measure of college quality

Described in Appendix Table 5

Measures of student abilities

Percentile over 4-year college starters in each cohort of the NLSY of the first (ASVAB1) and second (ASVAB2) principal components of the ASVAB test scores, as described in Appendix Tables 4A and 4B.

High school GPA from respondent's high school transcript in the NLSY97 and self-reported in the NLSY79, standardized to a 4-point scale weighted by Carnegie credits. GPA is orthogonalized against ASVAB1 and then the percentile is calculated within our [weighted] sample of college-goers in the same way as the ASVAB percentile.

Combined math and verbal SAT scores (max 1600) or the composite score on the ACT converted to the SAT scale from the respondent's high school transcript NLSY97 and self-report in NLSY79. SAT is orthogonalized against ASVAB1 and then the percentile is calculated within our [weighted] sample of college-goers in the same way as the ASVAB percentile.

Bad behavior index: count of petty anti-social behaviors by 1980 in NLSY 79 or by 8th grade in NLSY97: ever suspended from school, ever intentionally destroyed or damaged someone else's property, and ever stolen something worth $50 or less.

Indicator that respondent has sex before the age of 15.

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.

Family and demographic characteristics

Sex

Race and ethnicity: indicators for white, black, non-white Hispanic, or other non-white (last category in NLSY 97 only)

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.

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 NLSY97.

Highest educational attainment of either of the respondent's resident parents (or only parent in single parent households) as reported in the fall before the respondent finished high school (or earlier if that year is unavailable). We include at most one resident mother and father figure using the following prioritization: biological, adopted, step, or foster.

Neighborhood characteristics

Region of the U.S. where the respondent lived (Northeast, South, Midwest, or West) in 1979 for NLSY79 or in last year of high school for NLSY97.

Indicator that the respondent did not live within a Metropolitan Statistical Area (MSA) in 1979 for NLSY79 or in last year of high school for NLSY97.

Log median income (from 1990 census) in the census tract where the respondent lived in last year of high school.

The share of the over-25 population that has a 4-year college degree in the county where the respondent lived in 1979 (NLSY79) or in the last year of high school (NLSY97) from the 1972 and 1994 County and City Databooks, respectively.

Additional covariates included only in Tables 7 and 8

Indicators that the respondent was overweight or obese (using BMI and CDC definitions) in the last year of high school, or the closest earlier year available.

How many times per year the respondent attended religious services in the last year of high school, or the closest earlier year available (entered in regressions as indicators for each range of values offered in the survey).

Enrichment index: count of how many educational resources the respondent said they had regular access to at home in the 1979/1997 surveys. In the NLSY79, these resources are: a magazine subscription, a newspaper subscription, and a library card. In the NLSY97 these resources are: a computer, a dictionary, and a quiet place to study.

Indicator that respondent had ever lived with each biological parent for at least three months by the age of 18 (NLSY79) or had any contact with each biological parent by the 1997 (NLSY 97).