CLEF Canadian Labour Economics Forum

WORKING PAPER SERIES

The Consequences of Academic Match between Students and Colleges

Eleanor Dillon (Arizona State University)
Jeffrey Smith (University of Michigan)

Conference Volume CEA-2016

ABSTRACT

The Consequences of Academic Match between Students and Colleges*

We consider the effects of student ability, college quality, and the interaction between the two on academic outcomes and future earnings. Both ability and college quality strongly improve outcomes and earnings. We find little evidence to support the “mismatch” hypothesis that college quality and ability interact in substantively important ways. All students benefit from attending higher quality colleges. Our estimates imply that resorting students to eliminate mismatch, without changing the capacity of any colleges, would raise expected graduation rates by only 0.6 percentage points and mean earnings by $400 per year. The substantial gains for students who move to higher quality colleges under this reshuffling roughly cancel out the losses of students who move down.

JEL Classification: I21, J31

Keywords: college quality, mismatch

Corresponding author:

Jeffrey Smith
Department of Economics
University of Michigan
238 Lorch Hall
611 Tappan Street
Ann Arbor, MI 48104
USA
E-mail: econjeff@umich.edu

* 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 the 2014 Bergen-Stavanger Workshop on Labour Markets, Familes and Children, the 2013 CESifo education group meetings, Dartmouth, MDRC, Michigan CIERS, National University of Singapore, Penn GSE, Stanford CEPA, Cornell PAM, and Washington University in St. Louis. 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 labelled "overmatched" and relatively high ability students at relatively low quality colleges labelled "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 perhaps they might rise to the challenge and do better than they otherwise would have done. In the presence of such differential effects, resorting students via policy, even when respecting existing capacity constraints, has the potential to produce gains 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.

As its primary substantive contribution, this paper applies a “selection on observed variables” identification strategy to the data from the U.S. National Longitudinal Survey of Youth 1997 Cohort to provide estimates of academic match effects. These data provide a nationally representative sample of the most recent cohort for which sufficient time has elapsed since the completion of high school to allow for a serious analysis of the effects of match on both post-secondary outcomes and initial labor market outcomes. Though always somewhat heroic, the NLSY-97 contains a vast enough array of relevant conditioning variables to make our selection-on-observed-variables assumption at least moderately compelling.

As our second major contribution to the academic match literature, we examine a wide variety of outcome measures other than simply degree completion. With a couple of important recent exceptions discussed in greater detail below, the earlier literature focuses primarily on degree completion as the outcome of interest. Bowen and Bok’s (1998) finding of no apparent impact on degree completion for the overmatched students in the “College and Beyond” data 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 and completing easy majors, as suggested in journalistic exposes 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, as well as graduate school attendance and earnings in the years immediately following 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 a decade after initial enrollment quantifies the early labor market effects of college quality, ability and match.

Given our conceptual and empirical framing of academic match in terms of college quality interacted with student ability, as a natural byproduct of our analysis, and as our third major contribution, we replicate (in a broad sense), update and extend the earlier analyses of the college quality main effect in Black, et al. (2005), but using the NLSY-97 cohort. This allows us to compare estimates of the impact of college quality for the NLSY-79 and NLSY-97 cohorts. Our fourth contribution lies in our measures of student ability and college quality which, as we discuss in detail in Section 3, embody less measurement error than those used in other studies, as well as having other virtues relative to the measures commonly adopted in the literature.

To preview our results, we find substantial amounts of both overmatch and undermatch in the NLSY-97 cohort. Dillon and Smith (2013) concur with the literature in arguing that this mismatch results largely from the choices of students rather than the choices of college admissions offices. Our examination of the effects of ability, college quality and their interaction on college completion reveals strong main effects of college quality and ability, which comports with almost all of the existing literature. In contrast, we find little evidence of a casual effect of the interaction of quality and ability, which is to say that we find no substantively important effects of academic match. College quality raises both completion probabilities and post-college earnings for weak students, strong students, and those in between. Only our analysis of transfer behavior reveals patterns partially, but not fully, consistent with students correcting for initial mismatch.

We structure the remainder of the paper as follows: Section 2 reviews the literature on the college quality main effect and on academic match. Section 3 describes the many wonders of our data, with particular attention to the construction of our student ability and college quality measures and to the outcomes we consider. Section 4 presents our econometric framework. Section 5 displays and interprets our findings. Section 6 considers identification and makes the case for a causal interpretation of our estimates. Finally, Section 7 offers our conclusions.

2. Literature

We frame the literature on academic match as a subset of the literature that examines the effect of college quality on academic, labor market and other outcomes. In particular, the match literature considers the extent to which the effect of college quality varies according to the ability of the student. Our brief survey here organizes the literature by identification strategy and focuses in detail on the most recent studies and the ones that, in our view, illustrate the key issues involved. Though the literature has become international in recent years, we limit ourselves to the U.S. literature.[1] We also restrict ourselves to studies of academic match at the undergraduate level, putting to the side the tendentious literature on law school quality (see Sander and Taylor (2012) and the references therein) as well as that on business school quality. Black, et al. (2005) provides links to the earlier college quality literature. We are unaware of a good recent survey though one is surely warranted.

Recent studies relying on a selection on observed variables identification strategy to look at the college quality main effect include Black and Smith (2004), Black, et al. (2005), Black and Smith (2006), Long (2010), and Coate (2015). Bowen and Bok (1998), Alon and Tienda (2005), Bowen, et al. (2009) and Chingos (2011) examine college quality and academic match. Turner

[1] To get a flavor of the international literature, see e.g. Milla (2012) and Betts et al. (2013) for Canada, Hussain, et al. (2009) and Chevalier (2014) for the United Kingdom and Bordón and Braga (2015) for Chile.

(2002), Black, et al. (2005) and Dale and Krueger (2011) provide evidence on the persistence of college quality effects estimated under selection on observed variables. The conditioning sets available in these studies range from modest to impressively rich, with all but the studies using the PSID containing at least one standardized test score to measure student ability.

Another set of papers relies on instrumental variables strategies, sometimes embedded in structural models, though the institutional context does not offer up the most inspiring of instruments. These papers include Light and Strayer (2000), Arcidiacono (2004), and Long (2008).

Hoekstra (2009) looks at the effect of attending a flagship university using a regression discontinuity design. His study finds large positive effects on labor market outcomes, but the setup complicates their interpretation. First, as is well known the impacts strictly apply only to students at the margin of admission. In this sense, they inform about academic match, as they implicitly compare the weakest students admitted to the flagship with the strongest students not admitted to the flagship. However, the context provides not a sharp discontinuity but rather a fuzzy one, meaning that the effect properly applies only to the "compliers" at the discontinuity, those students whose enrollment in the flagship depends on crossing the admissions threshold. Hoekstra's positive finding represents strong evidence against the importance of mismatch effects for these students, though statistical discrimination issues further complicate the interpretation. If employers, at least initially, rely primarily on college attended as a proxy for ability, then the short run impact at the discontinuity will overstate both the longer-run impact and the impact for enrollees away from the discontinuity.[2] Zimmerman (2014) also finds substantively important effects on labor market outcomes in an RD towards the other end of the college quality spectrum, namely the admissions threshold of a lesser four-year public university

[2] See Hershbein (2013) for a thoughtful discussion of college quality and signaling.

in Florida. Goodman, et al. (2015) performs a similar analysis using administrative data for multiple states and obtain similar findings.

Dale and Krueger (2002, 2011) adopt a pair of provocative and original identification strategies that try to get around the problem that students and their parents, as well as the college admissions officers who read their applications, have information that the researcher does not by making use of the partial revelation of that information in students' application choices and the resulting college acceptance decisions. One strategy, which they call their self-revelation model, conditions on the average SAT score of colleges to which the student applied as well as the number of colleges to which they apply. The second strategy compares students accepted into (roughly) the same sets of colleges who make different choices regarding where to attend. We have concerns about both strategies that parallel those in Hoxby (2009).

Recently, Arcidiacono, et al. (2013) use the variation induced by California's Proposition 209, which attempts to ban racial preferences in university admissions, to study academic match effects on minority graduation in STEM fields, while Arcidiacono, et al. (2012) study the same phenomenon using administrative data from Duke University. These analyses mirror our emphasis on the use of intermediate outcomes to account for the absence of mismatch effects on completion.

Our main takeaways from the literature: First, a near consensus that college quality improves both academic and labor market outcomes. Second, academic match may affect some collegiate outcomes, such as major choice, but any negative effects on college completion appear small, if indeed they exist at all.

3. Data

3.1. NLSY

The NLSY-97 data include Americans born between 1980 and 1984. The first interview was in 1997 with follow-up interviews each year since. We include both the representative sample and the over-samples of blacks and Hispanics.[3] Most respondents graduated high school and made their college choice between 1999 and 2002. We focus on students who start at a 4-year college by 2006 (37% of high school graduates and GED holders in this sample). One of the strengths of the NLSY-97 data lies in the rich set of individual and family covariates it provides. 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. We describe the construction of our ability and college quality measures here; the appendix details the construction of the analysis sample and the definition and construction of our conditioning variables.

3.2. Ability

Our measures of student ability draw on the Armed Forces Vocational Aptitude Battery (ASVAB). The ASVAB is designed for applicants to the U.S. military. NLSY-97 respondents were invited to take it in 1997 as part of a norming exercise and were paid $75 for their time. 84% of respondents who started at a 4-year college completed the test. The ASVAB has twelve components, covering both the sorts of skills measured by the SAT and ACT such as algebra and reading comprehension, other skills such as electronics knowledge and spatial reasoning, and

[3] We use probability of inclusion (in the overall NLSY-97 sample) weights, constructed by the NLSY, to combine the two samples, and also to control for differing sampling and response rates in different regions of the U.S. and by age, gender, and race-ethnicity groups.

two timed sections wherein respondents complete as many short questions as possible in a fixed time period.[4]

When survey participants took the ASVAB, they ranged in age from 12 to 18. We adjust the scores for age at testing and then take the first principal component of the 12 section scores. Our primary measure of ability, which we call ASVAB1, is each respondent's percentile of this first principal component within the sample distribution of college-bound NLSY-97 respondents. As shown in the appendix, the first principal component explains 60% of the total variance in test scores across the 12 sections. The first component places the highest weight on academic subjects: arithmetic, word knowledge, and paragraph comprehension. Not surprisingly giving the loadings, the correlation between ASVAB1 and the respondent's SAT or ACT score equals 0.81.

The second component, which we call ASVAB2, explains a further 11% of the variance. It places the most weight on the two timed sections of the test: numerical operations and coding speed. Cawley, et al. (2001) calculate the principal components of the ASVAB scores in the earlier NLSY-79 cohort and estimate a similar loading pattern over the first two components. They find that the first two principal components of the ASVAB scores both predict later earnings in the NLSY-79 sample. We use ASVAB1 as our primary measure of student ability, but include ASVAB2 as an additional variable in our multivariate analyses.

The ASVAB offers a richer measure of ability than SAT or ACT scores because its 12 sections cover a greater range of material. The SAT and ACT are also taken as part of the college application process and shared with colleges. Students with sufficient resources may invest in coaching and take them multiple times in the hope of improving their performance. In contrast,

[4] 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 otherwise, that students possess around the time of the college choice.

the ASVAB provides a cleaner pre-college measure of ability. We can therefore capture some of the college mismatch generated because colleges and students have incomplete information. On the flip side, the ASVAB may capture more variation in effort because there was nothing riding on this test for the NLSY participants.

While we prefer our ASVAB-based ability measure to the SAT or ACT scores commonly relied on in the literature, it still does not capture all the abilities that make for a strong college student. Even if it did attempt to measure all relevant abilities, the score from a single ASVAB test would be an imperfect measure of ability because some students will perform above or below their usual level on any given day. To capture additional dimensions of ability we also include high school GPA and SAT scores along with multiple proxies for non-cognitive or socio-emotional skills; we describe these in more detail below and list them in the appendix. To avoid colinearity concerns with our primary cognitive skill measure we orthogonalize these additional ability measures against ASVAB1 when including them in the multivariate analyses.

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. In particular, our index combines the mean SAT score (or ACT score converted to the SAT scale) 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,[5] both from 2008.[6]

[5] 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

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.[7] Many colleges report two or three of our measures of college quality but not all four. Rather than drop these colleges from our analyses we modify the standard principal component analyses to adjust for partial non-response. Our college quality index is the weighted average across our four normalized quality measures, where the weights come from a principal component analysis within colleges who report all four quality measures. When colleges lack all four measures, we use the available measures and rescale the weights to sum to one.

Like Black and Smith (2006), we interpret our index as an estimate of latent college quality, which we view as continuous and one-dimensional. Within this framework, combining multiple proxies for college quality into a single index measures latent quality with less error than using a single proxy or the categorical quality ratings (e.g. from Barron's) used in much of the literature. Our index corresponds well to a priori notions of relative quality. For example, taking one corner of one state at random, the University of Michigan lies at the 93rd percentile, Michigan State at the 74th, Wayne State at the 36th, and Eastern Michigan at the 28th.

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. Thus, for example, the fact that the difference between Michigan and MSU equals 19 percentiles and that between Wayne State and EMU equals eight does not mean that the first difference

---

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.

[6] While most NLSY-97 respondents started college between 1999 and 2002, 2008 is the earliest year for which we could obtain U.S. News data and the first year that IPEDS reported faculty-student ratios focused only on undergraduates. The other components of our college quality measure remain quite stable between 2000 and 2008, so we feel the improved data available in 2008 outweigh the measurement error from observing college quality later.

[7] See the appendix for the details of the principal components analysis.

exceeds the second when mapped into the latent quality variable. In practice, the four individual quality measures underlying our index increase quite steadily across percentiles of our index for the bottom 90% of 4-year colleges, but increase more with each percentile rise in our index for the top 10% of colleges. 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 measures of college quality in the literature, namely the mean SAT score of the entering class, when interpreted as a proxy for latent quality, and the Barron's categories.

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, weighted by student body size.[8] Because we weight the quality percentile by student body size, a college in the n^{th} quality percentile is the college that a student in the n^{th} ability percentile would attend under perfect assortative academic matching of students and colleges. We consider students academically mismatched when they deviate substantially from this type of matching.

Table 1 gives the joint distribution of student ability and college quality. Students differentially concentrate along the diagonal, which corresponds to academic match, but there are also many mismatched students. The three upper right cells, corresponding to low ability students at high quality colleges, account for 10.7% of the sample, while the three lower left cells, corresponding to high ability students at low quality colleges, account for 12.5%. A comparison of Table 1 to Black and Smith (2004) and Light and Strayer (2000) reveals (perhaps

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

surprisingly given the recent policy focus on mismatch) no dramatic changes in the joint distribution between the NLSY-79 and NLSY-97 cohorts.[9]

One appealing feature of our measure is the possibility of achieving perfect assortative matching without violating institutional enrollment constraints. Other important studies in the literature, such as Roderick, et al. (2008), Bowen, et al. (2009), and Smith, et al. (2013) group students into test score bins and colleges into quality bins, usually using Barron's or another categorical quality measure. For each student test score bin, they then determine the highest quality bin with a high probability of admission. Students in the highest bin get labeled well-matched, with undermatch then defined by the distance (measured in bins) between the bin of the college the student actually enrolled in and the well-matched bin. For every student to be well-matched by their measures would require a sizeable expansion in the enrollment capacity of higher quality colleges. Undermatch resulting from limited slots at high quality colleges merits study (keeping in mind the general equilibrium issues that result from college quality depending in part on the average ability of student peers), but we focus on deviations from perfect sorting in the current system, without the added complication of relaxing capacity constraints.[10]

Table 2 presents descriptive statistics for students by the quality of the first 4-year college they attend. Throughout our analyses we consider the quality of the first 4-year college a student attends as the "treatment" under study; we thus view later transfers as intermediate outcomes. As expected, the students attending higher quality colleges have higher average ability by any of our measures. As in Dillon and Smith (2013) students with wealthier and more educated parents are

[9] Smith, et al. (2013) reach a different conclusion using other data sets and a different academic match measure, as does Long (2010). A thorough study of patterns of college sorting over time using consistent definitions of ability and quality would be a useful addition to the literature.

[10] House (2014) illustrates a broader problem in this literature by using Tennessee data to show the stark differences between the nature and quantity of academic match resulting from the various definitions in the literature when applied to a common dataset.

more likely to attend higher quality colleges, as are students from neighborhoods with higher median incomes and more educated adults.

3.5. Outcomes

We examine four educational outcomes: graduation within five years, transfer to a higher quality college, transfer to a lower quality college, and enrollment in graduate programs. We also consider earnings in all years from the start of college.

Table 3 presents summary statistics for these outcomes for our sample. Among our 2,406 4-year college starters, 62 percent have graduated by the most recent survey wave in 2011, but only 27 percent did so in four years or less.[11] Thus, our data contain substantial variation in time-to-degree. Among those not observed to graduate, the majority left school without a four-year degree, though small fractions remain in school at a two-year or four-year institution or left the data before the final year. As Bound et al. (2010) point out, time-to-degree has increased over the past 30 years and now represents an important implicit source of variation in the direct and opportunity costs of obtaining a degree.

Table 4 breaks down outcomes by college quality and student ability. The data clearly show that, unconditionally, completion probabilities increase in college quality. Degree completion probabilities generally increase in ability, although this pattern is less steep than the college quality gradient. Our multivariate analysis below shows that these patterns hold conditionally as well. 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 our students. As with time-to-degree this represents a change from earlier cohorts; for example, Light and Strayer (2000) did not find enough transfers to bother with in their analysis

[11] We omit the few respondents who first entered college after 2006 in our analyses so that all college starters we consider have now had at least 5 years to complete their degree.

of the earlier NLSY-79 cohort. Transfers down the college quality distribution (which includes transfers to any two-year college) occur substantially more often than transfers up and occur often even among students who begin in the lower half of the 4-year college quality distribution. Transfers up decline strongly with initial quality and show hints of match adjustment, with high ability students more likely to transfer up.

Finally, Table 4 considers earnings in various years after college enrollment. Throughout, we consider the level of real (1997) annual earnings, including zeroes, without conditioning on whether individuals are still enrolled in college or have completed their degree. 8-9 years after college, when almost all students have completed their studies, earnings increase in both college quality and student ability. 2-3 years after the start of college, when students who have not dropped out are still completing their degree, the pattern is reversed; earnings decrease in both ability and college quality. Earnings 2-3 years after starting college reflect both different probabilities of dropping out and beginning full-time work and different patterns of working while in college.

4. Econometric framework

To determine whether the data provide evidence of important interactions between ability and college quality, we want to look flexibly at the relationship between these two variables and the outcomes of interest. We can think of several alternative econometric frameworks that would allow us to do so. This section describes two: our preferred estimator based on a flexible polynomial approximation and an alternative estimator that uses indicators for bins of the 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:

$$\Pr(Y_i = 1 | A_i, Q_i, X_i) = \Phi(\beta_0 + \beta_p(A_i, Q_i) + \beta_X X_i) \tag{1}$$

In (1), Y denotes the binary outcome of interest, A denotes student ability, Q denotes college quality, $\beta_p\left(A_i, Q_i\right)$ is a polynomial of ability and quality, and X denotes a vector of other covariates. For earnings, we estimate a parametric linear regression model using the same specification of the independent variables, which is given by

$$\beta_p\left(A_i, Q_i\right) = \beta_{A1}A_i + \beta_{A2}A_i^2 + \beta_{A3}A_i^3 + \beta_{Q1}Q_i + \beta_{Q2}Q_i^2 + \beta_{Q3}Q_i^3 + \beta_{AQ1}A_iQ_i + \beta_{AQ2}A_i^2Q_i + \beta_{AQ3}A_iQ_i^2 \tag{1}$$

We chose this specification after a fairly rigorous round of statistical testing.[12] We can view this approach as a partially linear model in which we non-parametrically estimate the effects of ability and quality while conditioning parametrically on the other variables. The polynomial in ability and quality becomes non-parametric once we promise to include higher-order moments (but not too quickly!) on those happy occasions when our sample size increases.

Higher-order 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-

[12] We conducted a series of specification tests, with and without additional covariates, starting with higher-order moments of ability, quality, and their interactions and gradually moving towards more parsimonious specifications. 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.

quartile mismatch. In practice, the two estimators tell the same substantive story; see the Appendix for the estimates from the second approach.

5. Results

5.1 Degree Attainment

Table 5A presents our estimates of equation (1) for our binary education outcomes. The first five rows of the table 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 next five rows report the mean marginal effect of college quality at different points in the ability distribution. We present average derivatives (or finite differences in the case of binary covariates) of the conditional probability, which we regard as more interpretable than the probit coefficient estimates. For example, the value of 0.269 in the first column means that each 10 percentile increase in a student's ability increases her probability of graduating within 5 years by 2.69 percentage points if she is attending a college at the 25th percentile of the quality distribution. Holding student ability and other covariates constant, increasing the quality of the initial college in which a student enrolls also leads to a higher probability of obtaining a degree within five years. For a college starter of median ability, increasing the quality of first college attended by 10 percentiles increases the probability of graduating within five years by 3.02 percentage points. Thus, the college quality main effect found in (almost all of) the literature persists in our data.

Our second important finding emerges when we consider how the effect of college quality varies across the distribution of student ability and vice versa. If ability and quality have only independent effects, then the estimated effect of college quality should be uniform across students of different ability levels. In contrast, in a world of substantively important academic match effects, the effect of quality should vary across students. For example, college quality might increase degree completion probabilities more for students higher in the ability distribution. Empirically, we find virtually identical effects of college quality on graduation probability all along the distribution of student ability. Likewise, the effect of student ability is quite steady at different points in the college quality distribution. Figure 1 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 5 years is monotonically increasing in college quality, with a roughly constant difference between students of different ability levels. Put differently, 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 p-value from this test appears in the last row of Table 5A, and the corresponding chi-squared statistic appears in the penultimate row. We obtain a p-value of 0.77, indicating that the restrictions required for the main-effects-only model cause very little trouble for the data.

Second, we can look to Table 6, which compares 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 mismatch, moving from 48.1% to 48.6%. Chingos (2012) conducts a simpler version of this calculation with linear effects, fewer covariates and a different data set and also finds virtually no effect of resorting students. This small effect occurs because the positive effect of moving higher ability students at low quality colleges to their matched quality level is almost entirely cancelled out by the negative effect of moving lower ability students away from high quality colleges to their matched quality level.

This small net effect of eliminating academic mismatch masks the large improvements in outcomes from moving some students to higher-quality colleges. The last column of Table 6 presents 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 from 48.1% to 59.2%. This increase might seem smaller than expected, but student characteristics matter as well, and differ strongly between students at the 90th percentile and those further down the distribution.

Our estimates of the effect of student ability and college quality on the probability of enrolling in graduate school, presented in the second column of Table 5A, closely mirror our results on graduation probability and so we do not discuss them in detail here.

5.2 Transfer Behavior

We do find some evidence consistent with mismatch when looking at transfer behavior. The third column of estimates in Table 5A corresponds to model (1) with transfer up as the dependent variable, while the fourth column of estimates corresponds to transfer down. For this table, we consider only transfers that result in a change in college quality of five percentile points in one direction or the other. We have repeated the analysis using a zero cutoff (i.e. even moving by one percentile point counts) and a 10 percentile point cutoff and we obtain qualitatively similar findings.

Consider first our results for transferring to a higher quality college. Increasing a student's ability percentile by 10 percentage points raises the probability that she will transfer to a higher quality college by 2.4 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 five 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. Both patterns are consistent with a hypothesis that students preferentially transfer to better-matched colleges; more able students are more likely to transfer up the college quality distribution, particularly if they begin at a lower quality college.

We see the reverse pattern when considering transfers to lower quality colleges. For this outcome we consider transferring from any 4-year college to a 2-year college as transferring down. More able students are less likely to transfer to a lower quality college 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 results provide some support for the mismatch hypothesis and also support for a strong (and partly mechanical) main effect of college quality.

We cannot reject the null of only ability and quality main effects on transfer behavior, but this restriction does not fit the data as easily as it does for completion rates. As shown at the bottom of Table 5A, the p-values equal 0.19 and 0.33 for transferring to a higher and lower quality respectively. In Table 6, we find that eliminating mismatch would generate modest decreases in both types of transfers. The share of students transferring to a higher quality college would fall from 6.8 percent to 6.3 percent while the share transferring to a lower quality college would fall from 14.0 to 13.4 percent. Eliminating mismatch substantially decreases the transfer probability for severely mismatched students, but as shown in Table 1 these mismatched students are a small fraction of all college students. Since transfers often delay graduation, these moves are costly for the student and often the taxpayer in terms of more time spent in school and less time spent in the labor force.

5.3 Earnings

Table 5B presents our estimates of the effects of ability and college quality on earnings. These estimates use the same set of conditioning variables as our analyses of academic outcomes and again consist of average derivatives. Importantly, we do not condition on whether the student remains enrolled in college each year or whether they have completed a degree, which we view as intermediate outcomes. In the spirit of the program evaluation literature we consider the level of real earnings rather than the logs and include zero earnings observations for respondents who do not work. We respond to the high level of earnings volatility in these early years by trading a bit of detail for some precision and so consider averages of earnings in pairs of years relative to initial college enrollment.

The first column of results in Table 5B shows the effects of ability and college quality on average annual earnings in the year students start college and the year afterwards while the second column presents estimates for earnings 2-3 years after starting. In all four of these early years both college quality and 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 \$430 less per year for each 10 percentile point increase in the quality of first college attended. We suspect the negative relationship between college quality and earnings reflects three factors. First, students at less selective colleges are less likely complete their degree and are therefore more likely to have dropped out and begun working full time 2-3 years after starting college. Secondly, higher quality colleges may 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. Higher-ability students are less likely to drop out of college early, may be more inclined to focus exclusively on school work rather than working while in college, and may be eligible for more merit-based financial aid.

By 8-9 years after students begin college these patterns reverse: both college quality and ability 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 $1,400 of annual earnings. While the mean marginal effect of college quality is a little unstable over the ability distribution (and with large standard errors), increasing college quality has a positive effect on future earnings throughout the ability distribution. For this pair of years only, we reject (barely) the null of no interaction terms at the five percent level. We do not emphasize this statistical finding both because it disappears in years 10-11 and due to the modest substantive magnitudes at issue (about which we say more below). Because we are studying a recent cohort of college entrants our sample size falls starting nine years after the start of college. The youngest members of the sample graduated high school on time in 2002 and could therefore be at most nine years past the start of college in 2011, our last survey year. The last column of Table 5B suggests that both ability and quality continue to have increasingly positive effects on earnings 10-11 years after starting college, but the small number of observations makes it difficult to draw strong conclusions.

As with degree completion, we do not find much evidence that supports a mismatch hypothesis. If overmatching is harmful to students we might expect increases in college quality to lower earnings for low-ability students, either because overmatch pushes them into easier and less lucrative majors or because they become overwhelmed by higher expectations and fail to complete their degree. If undermatch is a concern we should see that improving college quality has a larger positive effect for higher-ability students. Neither pattern emerges in our data. Figure 3 plots the derivative of earnings 8-9 years after starting college with respect to college quality at several points in the ability distribution. Here we see that more able students experience somewhat larger gains from college quality in the top half of the college quality distribution. Nonetheless, the derivative of earnings with respect to quality is positive for all students at almost all points in the quality distribution. Earnings decline very slightly over the bottom 20% of colleges for high-ability students, but only a very small number of students in the extreme corner of the ability-quality distribution drive this pattern. In any case, negative returns to college quality for high-ability students at low-quality colleges are the opposite of what the mismatch hypothesis would suggest.

Table 6 shows that re-sorting the students in our data so as to eliminate mismatch would increase mean earnings by about $400 8-9 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 (or even talk in their inside voice about) the substantive importance of mismatch for later earnings.

6. Identification

This section considers the case for interpreting our estimates as causal. As noted above, we adopt a "selection on observed variables" strategy in this paper. More formally, we assume 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 there to exist variables not included in our model that vary college quality choices in ways unrelated to the unobserved component of the outcomes. 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.

Our preferred specification includes the following variables: the additional cognitive skill measures already discussed, several proxies for socio-emotional skills,[13] sex, race / ethnicity indicators, number of other children in the household, indicators for parental wealth quartiles, indicators for parental education categories, indicators for census region, log median income at the census tract level, percent of adults with a four-year college degree at the census tract level, log of average two-year in-state tuition, log of average four-year in-state tuition, indicators for

[13] We follow Aucejo (2012) and include indicators that the respondent was ever held back a grade, was ever suspended from school, has ever stolen something worth less than $50, or has ever intentionally destroyed property, all by 8th grade. We also include an indicator of whether the respondent had sex before the age of 15. Finally, following Cadena and Keys (2015) we include an indicator of whether the NLSY interviewer rated the respondent as somewhat uncooperative in any of the first three rounds of interviews.

having a matched public four-year college (in the same state) and matched private four-year college (in any state) within 50 miles, and an indicator for rural residence. All of these variables affect academic match in our earlier study of the determinants of mismatch, Dillon and Smith (2013). In most cases, we also have clear predictions regarding how these variables affect the outcomes we study. For example, parental education affects college completion independently of its effect on college quality choice if more educated parents can provide better advice on how to succeed in college and/or if parental education proxies for the otherwise unobserved component of pre-college academic preparation (i.e. it is correlated with the measurement error in student ability).

The two indicators for a well-matched college nearby likely represent the least obvious inclusions for most readers. Indeed, in the spirit of various papers using distance to college as an instrument, such as Card (1995) and Currie and Moretti (2003), we might think of these as instruments rather than as conditioning variables. We include them as conditioning variables because we worry that they correlate with living at home, which likely has its own treatment effect on the outcomes we study, particularly completion. At the same time, we do not condition directly on living at home because it represents an intermediate outcome.

Should we expect our conditioning set to suffice for selection on observed variables to hold, at least approximately? We can make this case in two ways. First, we can think about what we know from existing theory and empirical evidence regarding what we should condition on, and then ask whether our conditioning set contains those things, or at least compelling proxies for those things. We clearly want 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 on. More money may also affect the college quality choice, particularly for students without a good public college nearby. 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. More educated parents may also have a stronger taste for education, and so may push their children harder to finish. As in Becker and Lewis (1973), parents face a quality-quantity tradeoff. As such, number of other children may reflect both resources and preferences. We include direct measures of parental resources, education and number of children. We also expect that our census tract income and education variables will both help with measurement error in the parental resource variables and proxy for primary and secondary school quality as well as peer pressure and expectations. We also clearly want to condition on student ability, and we do so flexibly.

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), we might imagine that there exist multiple unobserved factors that we need to condition on to solve the problem of non-random selection into colleges of different qualities. We can then think of our conditioning variables as proxies for those factors. In a world with just one unobserved factor, as we increase the number of proxies in our conditioning set, the amount of selection bias in our estimates should decrease to zero. The same holds with two unobserved factors so long as we keep adding proxies for both. Turning this around, if we observe that the estimates stabilize as we increase the richness of the conditioning set, this provides evidence that we are doing a good job of proxying for the unobserved factors, unless there exists an additional unobserved factor uncorrelated with our covariates.

Tables 7 and 8 present estimates based on increasingly rich sets of conditioning variables for graduating in five years and for earnings in years 8-9 after starting college, respectively. 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 Table 5. The estimates in column (6) further exploit the richness of the NLSY-97 data by including variables related to family income, respondent obesity, family religiosity, educational experiences outside of formal schooling, various aspects of the home environment, and contact with biological parents. We think of these additional variables as further proxies for otherwise unmeasured differences in academic preparation, family resources, and family enthusiasm for education. Overall, the tables reveal a substantial amount of movement in the coefficients when moving from column (1) to column (2), which corresponds to adding additional measures of ability and socio-emotional skills, when moving from column (2) to column (3), which corresponds to adding demographics and family characteristics, and when moving from column (3) to column (4), which corresponds to adding neighborhood characteristics and local college options. In contrast, and in parallel to the similar analysis in the Black, et al. (2005) study of college quality, moving from column (4) to column (5) changes the estimates very little. Oster (2013) cautions that our line of reasoning related to coefficient stability means little if the newly added variables do not capture any (conditional) variation in the dependent variable. The r-squared values address this concern, and show that even the final set of added covariates capture a respectable amount of additional variation. Thus, we view our analysis in Tables 7 and 8 as providing further evidence in support of a causal interpretation.

As noted above, in addition to having the correct covariates in our model, we need the correct kind of variation (and, ideally, lots of it) conditional on those covariates. In our view, the literature suggests that plenty of exogenous variation exists in college quality choices conditional on our observed covariates. First of all, differences in state college quality mix and pricing strategies provide plausibly exogenous variation in the budget sets facing students and their parents. Second, in this context, 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. For example, the literature on "one-offs" (the occasional strong student at a weak high schools), such as Hoxby and Avery (2012) and Hoxby and Turner (2013), shows 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. Here we have in mind, for instance, the finding in Pallais (2015) that you can change college choices by changing the number of colleges to which students can send their ACT scores for free or the finding in Bettinger et al. (2009) 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, as well as casual observation, indicates that many students explicitly choose among colleges, at least at the margin, for reasons unlikely to be strongly related to outcomes, such as because of the football team or because their best friend from high school is going there.

What, if anything, can we say about the nature of any remaining bias due to selection on unobserved variables? Putting aside match for the moment and thinking just about the college quality main effect, two worries usually arise. First, we might expect students, their parents, and college admissions officers have access to information on student ability that we, the researchers, do not. To the extent that those unobserved bits affect admissions in the expected way, with better unobserved bits leading to admission to higher quality colleges conditional on the observed bits, and worse unobserved bits the reverse, we would expect an upward bias in the estimated effect of college quality because it proxies in part for unobserved student ability. 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 should help with this issue relative to the common strategy of using only the average test scores of the entering class, some measurement error likely remains, and we would expect it to push the estimated coefficient 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.

7. Summary and conclusions

This paper examines the effects of college quality and student ability on academic and labor market outcomes. We use the rich data from the NLSY-97 and adopt a "selection on observed variables" identification strategy. We find strong evidence that college quality increases earnings and the probability of degree completion, and similarly strong evidence that student ability increases both as well. At the same time, we do not find much evidence of the interactive effects of college quality and student ability predicted by the mismatch hypothesis except in the case of transfer behavior. Our findings imply the absence of substantively meaningful efficiency gains from resorting students to reduce mismatch as we measure it.

We conclude with two caveats. First, our results consider only mismatch at the undergraduate level. Our results may not generalize to other 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, the student cannot change majors, or easily reduce their course load. For this reason, mismatch, particularly overmatch, might have very different overall effects in that context than in ours.

Second, 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.

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/

Arcidiacono, Peter (2004): "Ability Sorting and the Returns to College Major," *Journal of Econometrics*, 121, pp. 343-375.

Arcidiacono, Peter, Esteban Aucejo and V. Joseph Hotz (2013): "University Differences in the Graduation of Minorities in STEM Fields: Evidence from California," Unpublished manuscript, Duke University.

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.

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.

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.

Betts, Julian, Christopher Ferrall, and Ross Finnie (2013): "The Role of University Characteristics in Determining Post-Graduation Outcomes: Panel Evidence from Three Canadian Cohorts," *Canadian Public Policy*, 39, pp. 81-106.

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.

Bordón, Paola and Breno Braga (2015): "Employer Learning, Statistical Discrimination and University," Unpublished manscript, the Urban Institute.

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*, forthcoming.

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.

Chevalier, Arnaud (2014): "Does Higher Education Quality Matter in the UK," *Research in Labor Economics*, 40, pp. 257-292.

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.

Coate, Patrick (2015): "Returns to Quality and Location of College – Evidence from the PSID." Unpublished Manuscript, University of Michigan.

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 (2011): "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 (2013): "The Determinants of Mismatch between Student Ability and College Quality," NBER Working Paper No. 19286.

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 (2015): "College Access, Initial College Choice and Degree Completion," NBER Working Paper No. 20996.

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

Hershbein, Brad (2013): "Worker Signals among New College Graduates: The Role of Selectivity and GPA," Unpublished manuscript, W.E. Upjohn Institute for Employment Research.

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 (2014): "Exploring Undermatch in Tennessee: An Analysis Using Multiple Definitions of `Undermatch'" Unpublished manuscript, University of Michigan.

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

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.

Hussain, Iftikhar, Sandra McNally, and Shqiponja Telhaj (2009): "University Quality and Graduate Wages in the UK," IZA Discussion Paper No. 4043.

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

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.

Milla, Joniada (2012): "University Quality and Labor Market Outcomes of Canadian Youth," Unpublished manuscript, University of Guelph.

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 (2014): "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.

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 1: **Joint distribution of college quality and ability—NLSY-97, four-year starters**

	College Quality Quartiles				
Ability Quartiles	1st Quartile (lowest)	2nd Quartile	3rd Quartile	4th Quartile (highest)	Total
1st Quartile (lowest)	10.6 (45.3) [40.8]	6.8 (28.9) [28.3]	4.0 (17.0) [15.4]	2.1 (8.8) [8.4]	(100.0) (N=711)
2nd Quartile	7.1 (28.3) [27.3]	6.5 (25.8) [27.0]	6.9 (27.6) [26.8]	4.6 (18.4) [18.8]	(100.0) (N=602)
3rd Quartile	5.2 (20.5) [20.2]	6.4 (25.1) [26.8]	7.4 (29.1) [28.9]	6.5 (25.4) [26.5]	(100.0) (N=557)
4th Quartile (highest)	3.0 (11.6) [11.6]	4.3 (16.3) [17.8]	7.4 (28.5) [28.9]	11.4 (43.6) [46.4]	(100.0) (N=536)
Total	[100.0] [N=679]	[100.0] [N=579]	[100.0] [N=599]	[100.0] [N=549]	100.0 N=2,406

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. Observation counts are unweighted.

Table 2: **Average characteristics of students by initial college choice, four-year starters**

	College	College quality quartile			
	Attendees	1, lowest	2	3	4, highest
N	2,406	679	579	599	549
ASVAB 1 percentile	51%	37%	45%	55%	66%
ASVAB 2 percentile	50%	44%	49%	50%	58%
High school GPA, percentile	51%	45%	51%	52%	56%
SAT or ACT percentile	50%	35%	47%	51%	62%
Ever suspended by grade 8	8%	13%	8%	8%	5%
Ever held back by grade 8	3%	6%	3%	3%	1%
Ever stole something <$50	29%	31%	26%	33%	24%
Ever intentionally damaged property	24%	26%	24%	24%	19%
Had sex before age 15	9%	14%	9%	8%	5%
Interviewer rated uncooperative	38%	43%	35%	37%	35%
Male	45%	45%	41%	46%	48%
White	79%	76%	80%	80%	80%
Black	12%	17%	15%	10%	5%
Hispanic	3%	4%	2%	4%	3%
Other (not white)	6%	3%	3%	7%	11%
Wealth quartile 1 (lowest)	11%	18%	9%	8%	9%
Wealth quartile 2	19%	27%	21%	14%	13%
Wealth quartile 3	28%	28%	29%	33%	20%
Wealth quartile 4 (highest)	42%	28%	41%	45%	58%
Household members age 18 or under	2.2	2.3	2.3	2.1	2.2
No parent completed high school	3%	6%	2%	2%	2%
At least one parent grad. high sch.	19%	27%	24%	14%	9%
At least one parent has some college	26%	30%	26%	24%	22%
At least one parent completed college	53%	38%	47%	59%	67%
Northeast region	21%	13%	15%	23%	33%
South region	31%	33%	40%	27%	24%
Midwest region	31%	37%	29%	30%	28%
West region	17%	17%	17%	20%	15%
Rural	18%	30%	17%	16%	10%
Median income in census tract	$36,854	$32,858	$35,869	$38,148	$40,679
% Adults w/college deg. in tract	21%	18%	20%	23%	23%
Avg. 2-year in-state tuition	$3,147	$3,032	$3,075	$3,113	$3,373
Avg. 4-year in-state tuition	$1,513	$1,368	$1,508	$1,502	$1,681
Matched public 4-year in 50 mi	65%	60%	63%	65%	74%
Matched private 4-year in 50 mi	77%	68%	76%	79%	84%

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. In-state tuition is measured in the year each student graduated from high school, deflated to 1997 dollars.

Table 3: Summary of College Outcomes

	Respondents	Percent
Of 4-year college starters	2,406	
Graduate in 4 years or less	642	27%
Graduate in 5 years	434	18%
Graduate in 6 or more years	398	17%
Leave school without a BA	837	35%
Still in 4-year college as of last interview	13	1%
Still in 2-year college as of last interview	3	0%
Not in most recent survey wave	79	3%
Of 4-year college starters		
Transfer to a higher quality college	168	7%
Transfer to a lower quality college	346	14%
Transfer to a similar or unknown college	71	3%
Never transfer	1,821	76%

Categories in each panel of the table are mutually exclusive. Respondents who left college while still participating in the survey are counted as graduated or left without BA even if they did not respond to the most recent survey wave. Respondents who were in school as of their last interview and have not participated in the most recent waves of the survey are counted as out of survey.

Table 4: Summary of outcomes by ability and college quality quartile

	Count	Graduate within 5 years	Transfer to a higher quality college	Transfer to a lower quality college	Earnings 2-3 years after starting college	Earnings 8-9 years after starting college
Quality 1, ability 1	324	17%	10%	10%	$10,291	$19,436
Quality 1, ability 2	180	27%	13%	12%	$10,673	$21,697
Quality 1, ability 3	114	34%	16%	7%	$10,729	$23,291
Quality 1, ability 4	61	41%	19%	8%	$9,699	$30,218
Quality 2, ability 1	206	33%	6%	22%	$8,807	$24,037
Quality 2, ability 2	151	46%	6%	16%	$8,023	$25,864
Quality 2, ability 3	135	52%	8%	13%	$6,941	$26,222
Quality 2, ability 4	87	42%	18%	14%	$10,114	$28,294
Quality 3, ability 1	121	43%	1%	18%	$8,430	$24,929
Quality 3, ability 2	165	50%	5%	19%	$7,224	$26,133
Quality 3, ability 3	163	59%	2%	16%	$6,966	$29,071
Quality 3, ability 4	150	59%	5%	19%	$7,132	$30,805
Quality 4, ability 1	60	61%	1%	22%	$8,515	$26,827
Quality 4, ability 2	106	66%	3%	13%	$7,103	$31,625
Quality 4, ability 3	145	71%	2%	12%	$6,259	$35,657
Quality 4, ability 4	238	71%	1%	9%	$5,027	$36,380

Count is the unweighted number of students in each category. Probabilities and earnings are weighted as described in the text. Transferring to a lower quality college includes transferring to a two-year college. Earnings are average annual earnings, deflated to 1997 dollars, over a two-year period.

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

		Graduate within 5 years	Enroll in graduate school	Transfer to higher quality college	Transfer to lower quality college
dOutcome/dA	Q = p10	0.254	0.300	0.210	-0.074
		(0.111)	(0.119)	(0.133)	(0.043)
	Q = p25	0.269	0.342	0.238	-0.069
		(0.089)	(0.094)	(0.106)	(0.064)
	Q = p50	0.274	0.337	0.109	-0.063
		(0.087)	(0.086)	(0.055)	(0.083)
	Q = p75	0.277	0.305	0.015	-0.151
		(0.087)	(0.083)	(0.015)	(0.085)
	Q = p90	0.275	0.278	-0.004	-0.263
		(0.117)	(0.116)	(0.014)	(0.126)
dOutcome/dQ	A = p10	0.360	0.313	-0.113	0.287
		(0.111)	(0.116)	(0.042)	(0.119)
	A = p25	0.329	0.309	-0.112	0.199
		(0.082)	(0.086)	(0.041)	(0.073)
	A = p50	0.302	0.287	-0.129	0.129
		(0.079)	(0.074)	(0.046)	(0.045)
	A = p75	0.326	0.284	-0.177	0.105
		(0.074)	(0.070)	(0.055)	(0.032)
	A = p90	0.361	0.293	-0.233	0.100
		(0.091)	(0.095)	(0.084)	(0.031)
2nd ASVAB factor		0.089	0.024	0.065	-0.037
		(0.038)	(0.035)	(0.022)	(0.029)
HS GPA		0.290	0.110	-0.013	-0.093
		(0.043)	(0.041)	(0.023)	(0.033)
SAT		0.110	0.066	0.007	-0.086
		(0.047)	(0.044)	(0.024)	(0.036)
Observations		2,405	2,405	2,361	2,388
R-squared		0.185	0.108	0.134	0.054
Test: interaction=0, Chi2		1.127	1.480	4.713	3.453
Pr(interaction=0)		0.771	0.687	0.194	0.327

The first ten rows report the mean marginal effect of ASVAB percentile at different points in the college quality distribution and the mean marginal effect of college quality percentile at different points in the student ASVAB distribution. These marginal effects are calculated from the coefficients of a polynomial of ability and college quality as described in the text, conditioning on the set of covariates described in Appendix Table 2.

Table 5B: Effect of College Quality and Ability on Earnings

		Year 0-1	Year 2-3	Year 4-5	Year 6-7	Year 8-9	Year 10-11
dEarnings / dA	Q = p10	-0.006	1.350	-1.641	7.770	10.897	24.611
		(1.373)	(2.096)	(2.701)	(4.650)	(6.761)	(10.366)
	Q = p25	-0.961	0.476	-0.265	6.059	4.528	15.112
		(1.127)	(1.648)	(2.351)	(3.627)	(4.758)	(7.744)
	Q = p50	-2.191	-1.159	0.353	4.792	1.264	8.343
		(1.225)	(1.640)	(2.458)	(3.411)	(4.061)	(7.117)
	Q = p75	-2.830	-2.733	-1.024	5.536	5.802	10.808
		(1.258)	(1.633)	(2.392)	(3.505)	(4.026)	(6.918)
	Q = p90	-3.031	-3.973	-3.168	7.052	12.455	17.437
		(1.549)	(2.185)	(3.173)	(4.775)	(5.668)	(9.729)
dEarnings / dQ	A = p10	0.642	-1.273	4.961	4.413	0.461	7.246
		(1.541)	(2.349)	(3.288)	(4.444)	(5.501)	(8.869)
	A = p25	-0.316	-2.542	3.718	8.763	8.785	16.491
		(1.278)	(1.922)	(2.611)	(3.394)	(4.116)	(6.470)
	A = p50	-1.539	-4.308	2.759	11.139	14.006	19.220
		(1.329)	(1.922)	(2.551)	(3.389)	(4.016)	(6.310)
	A = p75	-2.175	-5.755	2.813	8.400	10.390	11.542
		(1.212)	(1.757)	(2.329)	(3.291)	(4.366)	(6.927)
	A = p90	-2.328	-6.487	3.285	4.506	4.255	3.537
		(1.328)	(1.968)	(2.694)	(4.299)	(6.262)	(9.504)
2nd ASVAB factor		-0.382	-0.985	-0.126	2.510	3.736	1.382
		(0.450)	(0.685)	(1.049)	(1.592)	(1.948)	(3.318)
HS GPA		-2.131	-2.874	-1.247	2.328	2.620	5.632
		(0.527)	(0.753)	(1.143)	(1.718)	(2.234)	(3.727)
SAT		-0.482	-0.634	-0.497	0.648	-0.532	-3.013
		(0.493)	(0.734)	(1.341)	(1.965)	(2.591)	(4.168)
Observations		2,332	2,289	2,228	2,141	1,872	958
R-squared		0.127	0.132	0.045	0.085	0.142	0.181
Test: interaction=0, F		1.640	1.889	0.464	1.085	2.607	1.453
Pr(interaction=0)		0.178	0.129	0.707	0.354	0.050	0.226

The dependent variable is average annual earnings, in thousands of 1997 dollars, averaged over 2-year bands beginning with the first year of college. The first ten rows report the mean marginal effect of ASVAB percentile at different points in the college quality distribution and the mean marginal effect of college quality percentile at different points in the student ASVAB distribution. These marginal effects are calculated from the coefficients of a polynomial of ability and college quality as described in the text, conditioning on the set of covariates described in Appendix Table 2.

Table 6: Counterfactual Outcomes from Re-assigning Students to Colleges

	Actual outcome	If all students attended a matched college	If all students attended a 90th percentile college
Graduate within 5 years	48.1	48.6	59.2
Enroll in a graduate program	24.9	24.4	34.3
Transfer to a higher quality college	6.8	6.3	1.2
Transfer to a lower quality college	14.0	13.4	17.1
Earnings 2-3 years after starting college	$8,041	$7,673	$7,577
Earnings 8-9 years after starting college	$27,937	$28,346	$30,934

This table presents the share of respondents who achieve each outcome in the data and the predicted share of respondents achieving each outcome in the counterfactual case of no mismatch (all students attend a college in the quality percentile that matches their ability percentile) or if all students attended a very high-quality college. Predictions are made using the coefficients in Tables 5A and 5B.

Table 7: Effect of College Quality and Ability on Graduating within 5 years

		(1)	(2)	(3)	(4)	(5)
dOutcome/dA	Q = p10	0.235	0.304	0.272	0.254	0.242
		(0.107)	(0.109)	(0.113)	(0.111)	(0.112)
	Q = p25	0.237	0.329	0.277	0.269	0.264
		(0.087)	(0.086)	(0.089)	(0.089)	(0.089)
	Q = p50	0.224	0.333	0.262	0.274	0.273
		(0.087)	(0.083)	(0.088)	(0.087)	(0.087)
	Q = p75	0.253	0.331	0.252	0.277	0.269
		(0.087)	(0.084)	(0.087)	(0.087)	(0.086)
	Q = p90	0.290	0.319	0.241	0.275	0.256
		(0.122)	(0.118)	(0.117)	(0.117)	(0.117)
dOutcome/dQ	A = p10	0.551	0.457	0.440	0.360	0.358
		(0.112)	(0.111)	(0.104)	(0.111)	(0.108)
	A = p25	0.573	0.439	0.409	0.329	0.321
		(0.079)	(0.080)	(0.078)	(0.082)	(0.081)
	A = p50	0.568	0.411	0.372	0.302	0.290
		(0.072)	(0.076)	(0.075)	(0.079)	(0.078)
	A = p75	0.571	0.424	0.377	0.326	0.315
		(0.072)	(0.072)	(0.072)	(0.074)	(0.074)
	A = p90	0.581	0.449	0.396	0.361	0.353
		(0.093)	(0.091)	(0.092)	(0.091)	(0.091)
Additional ability measures			Yes	Yes	Yes	Yes
Socioemotional skills			Yes	Yes	Yes	Yes
Demographics				Yes	Yes	Yes
Family characteristics				Yes	Yes	Yes
Neighborhood characteristics					Yes	Yes
Local college options					Yes	Yes
Additional covariates						Yes
Observations		2,406	2,405	2,405	2,405	2,405
R-squared		0.098	0.153	0.169	0.185	0.191

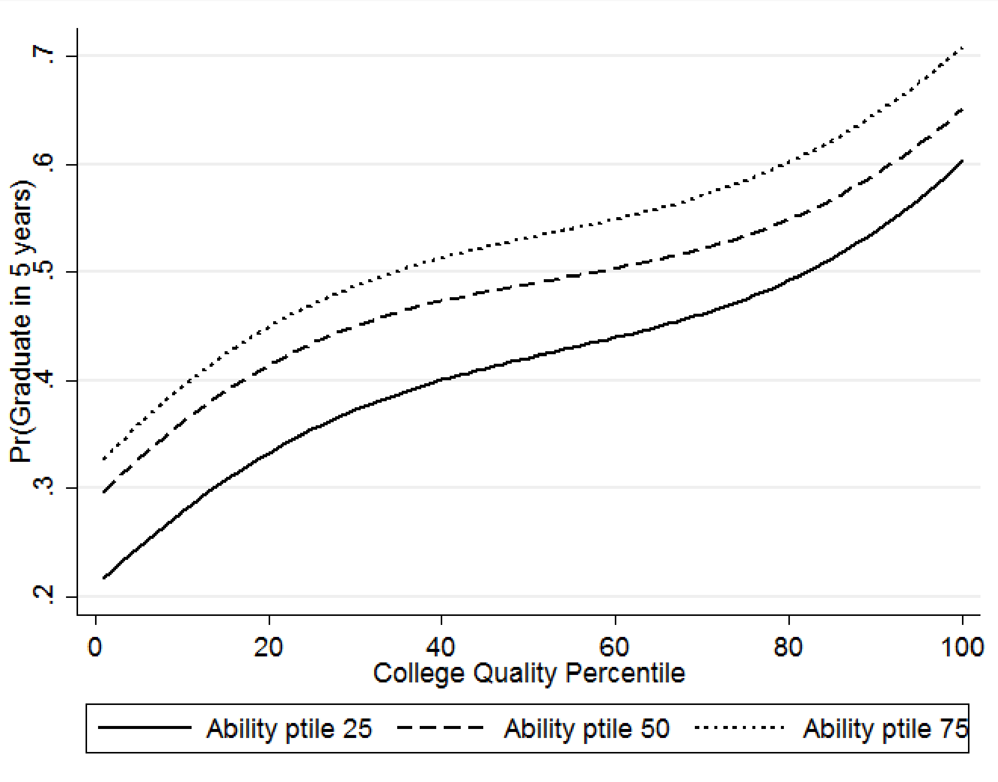
The first ten rows report the mean marginal effect of ASVAB percentile at different points in the college quality distribution and the mean marginal effect of college quality percentile at different points in the student ASVAB distribution. These marginal effects are calculated from the coefficients of a polynomial of ability and college quality as described in the text, conditioning on the indicated subsets of the covariates described in Appendix Table 2.

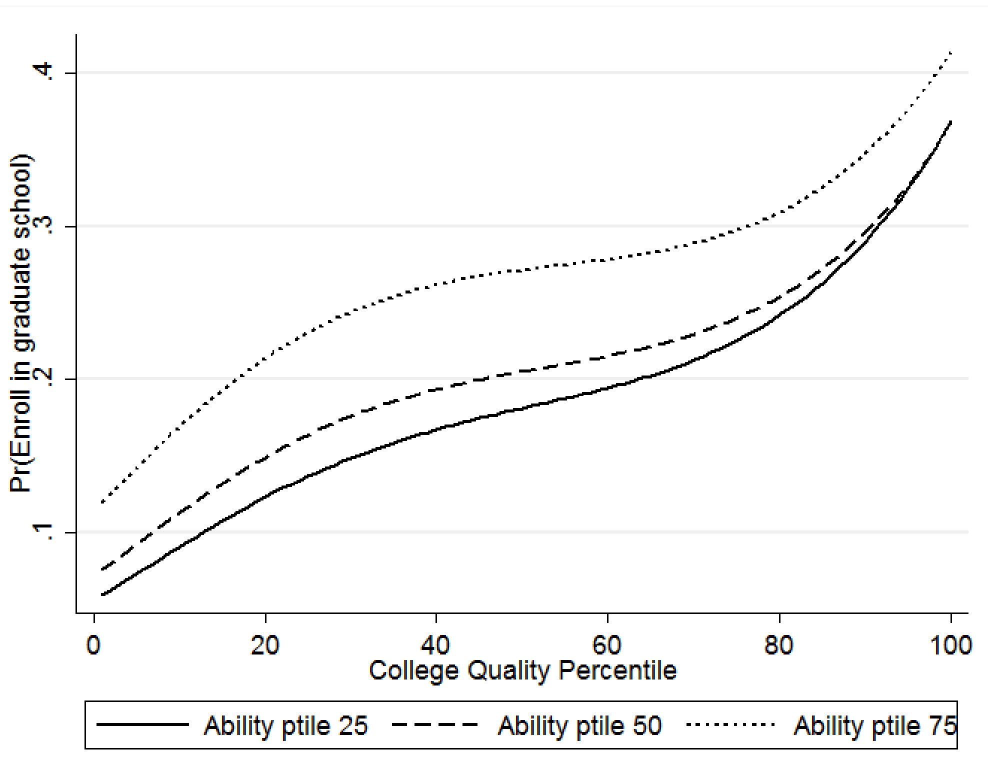
Table 8: Effect of College Quality and Ability on Earnings 8-9 Years after starting college

		(1)	(2)	(3)	(4)	(5)
dEarnings / dA	Q = p10	16.525	16.579	12.398	10.897	11.415
		(6.669)	(6.734)	(6.869)	(6.761)	(6.729)
	Q = p25	9.658	9.744	4.940	4.528	4.849
		(4.543)	(4.637)	(4.769)	(4.758)	(4.798)
	Q = p50	6.038	6.101	0.574	1.264	1.344
		(3.787)	(3.892)	(4.022)	(4.061)	(4.128)
	Q = p75	10.723	10.675	4.742	5.802	5.730
		(3.797)	(3.864)	(3.969)	(4.026)	(4.034)
	Q = p90	17.709	17.549	11.506	12.455	12.326
		(5.644)	(5.686)	(5.637)	(5.668)	(5.629)
dEarnings / dQ	A = p10	4.559	4.128	3.076	0.461	-0.340
		(5.668)	(5.666)	(5.418)	(5.501)	(5.517)
	A = p25	13.410	13.106	11.308	8.785	8.006
		(4.225)	(4.192)	(4.031)	(4.116)	(4.133)
	A = p50	18.860	18.578	15.975	14.006	13.103
		(4.080)	(4.025)	(3.936)	(4.016)	(4.022)
	A = p75	14.833	14.392	11.502	10.390	9.229
		(4.394)	(4.344)	(4.316)	(4.366)	(4.346)
	A = p90	8.147	7.521	4.636	4.255	2.856
		(6.333)	(6.308)	(6.345)	(6.262)	(6.210)
Additional ability measures			Yes	Yes	Yes	Yes
Socioemotional skills			Yes	Yes	Yes	Yes
Demographics				Yes	Yes	Yes
Family characteristics				Yes	Yes	Yes
Neighborhood characteristics					Yes	Yes
Local college options					Yes	Yes
Additional covariates						Yes
Observations		1,873	1,872	1,872	1,872	1,872
R-squared		0.074	0.078	0.126	0.142	0.153

The dependent variable is average annual earnings, in thousands of 1997 dollars, averaged over 2-year bands beginning with the first year of college. The first ten rows report the mean marginal effect of ASVAB percentile at different points in the college quality distribution and the mean marginal effect of college quality percentile at different points in the student ASVAB distribution. These marginal effects are calculated from the coefficients of a polynomial of ability and college quality as described in the text, conditioning on the indicated subsets of the covariates described in Appendix Table 2.

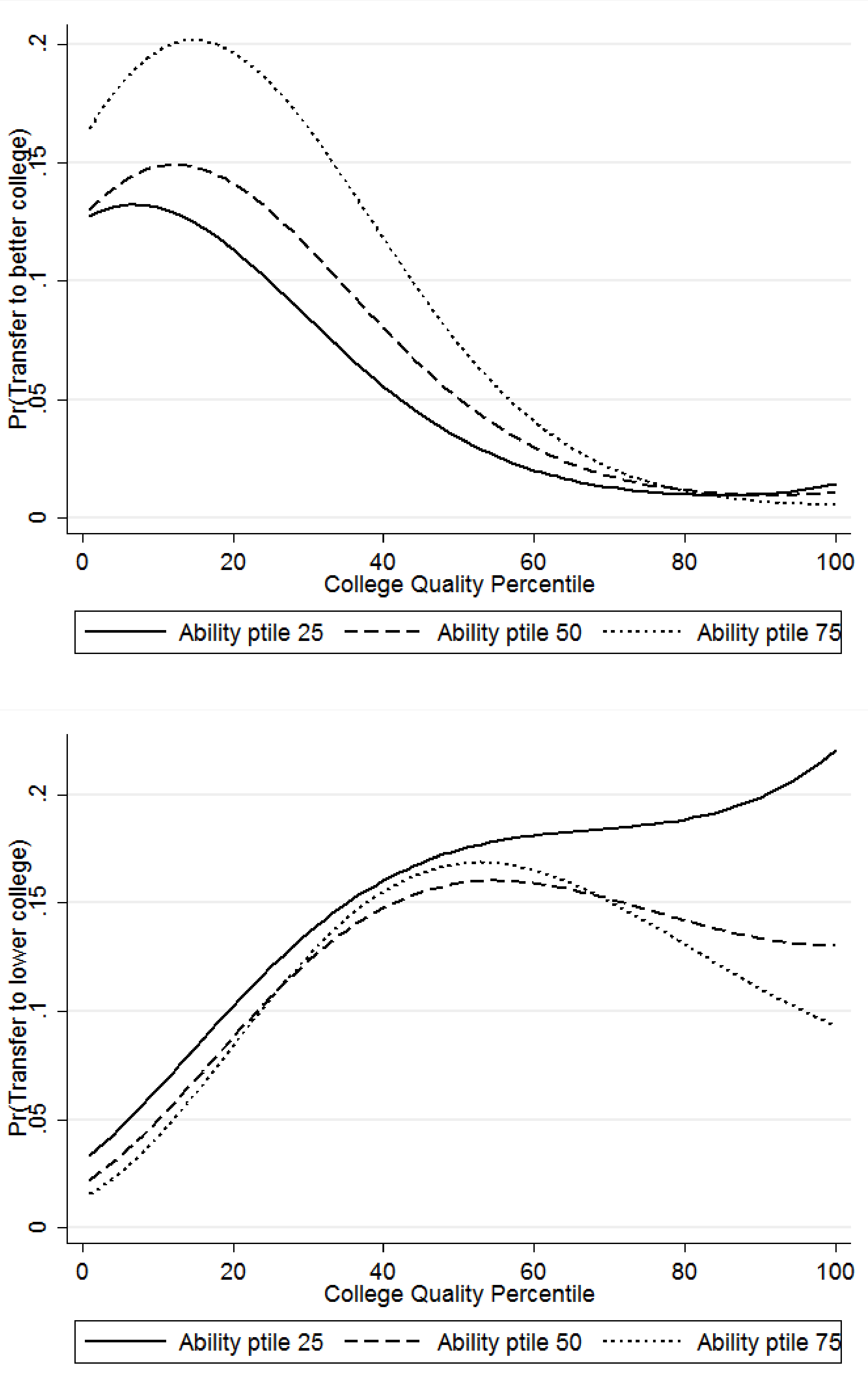
Figure 1: Effect of Student Ability and College Quality on College Outcomes





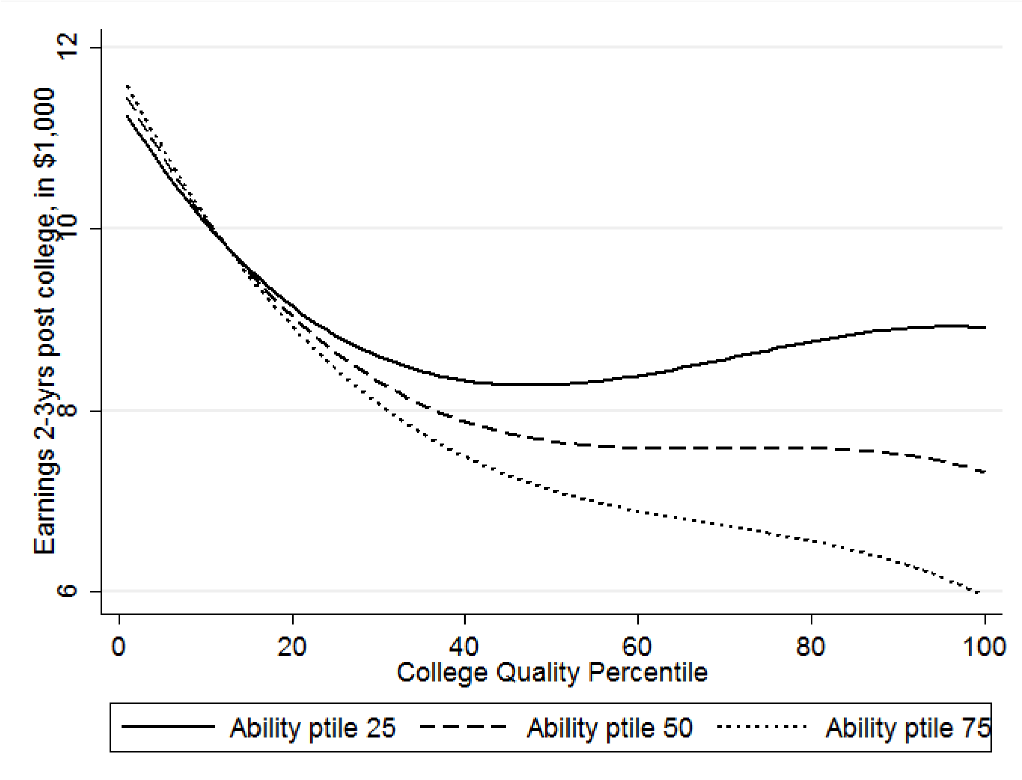
Projected from estimates in Table 5A.

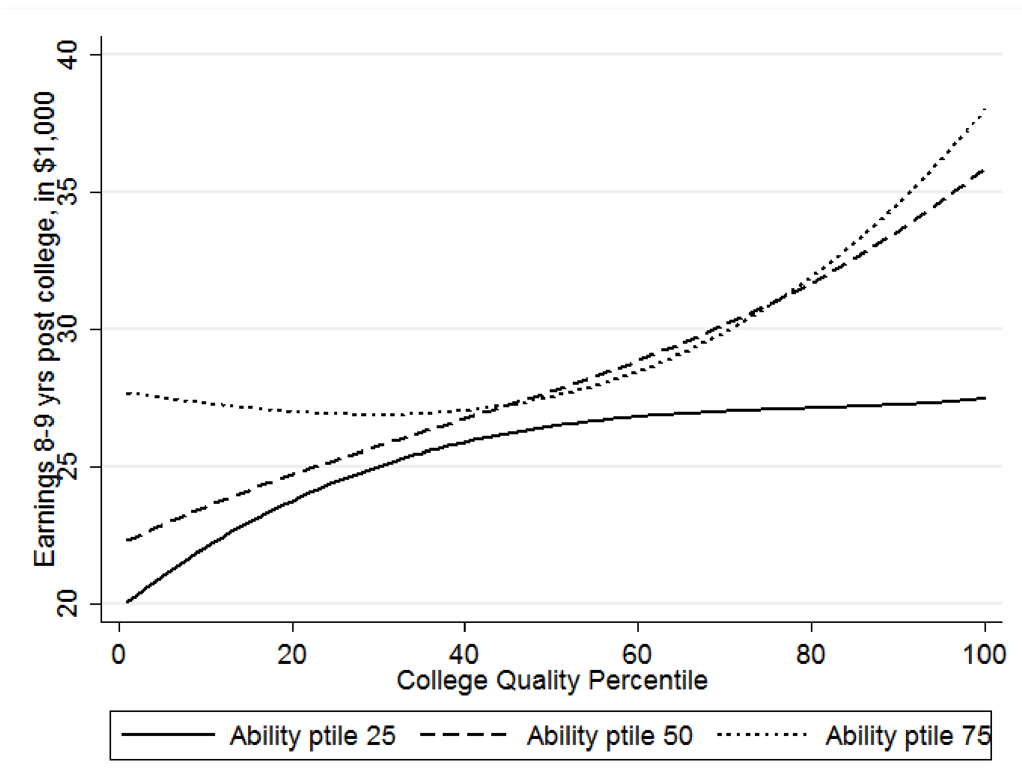
Figure 2: Effect of Student Ability and College Quality on College Transfers



Projected from estimates in Table 5A.

Figure 3: Effect of Student Ability and College Quality on Earnings





Projected from estimates in Table 5B. Earnings in thousands of 1997 dollars.

Appendix Table 1: Sample

Total Observations	8,984
Graduated HS	6,722
Did not graduate HS but got GED	1,123
Started at a 4-year college*	2,915
Starting college qualities	
Of quality quartile 1	819
Of quality quartile 2	682
Of quality quartile 3	714
Of quality quartile 4	641
Missing quality	*59*
Has quality, but missing ability	*450*
Analysis sample	2,406

* The 4-year starters include 53 respondents who got a GED and 4 respondents with no recorded high school graduation date or GED.

College quality is for the first college attended. Of the 59 respondents who started at a 4-year school for whom we do not have a quality index, 29 are missing quality because we could not identify the college and 30 are missing quality because the school was not in IPEDS or did not have enough information to construct a quality measure.

Appendix Table 2: Description of Independent Variables

Measures of student ability

Percentile over 4-year college starters in the NLSY-97 of the first (ASVAB1) and second (ASVAB2) principal components of the 12 sections of the ASVAB test, taken by NLSY-97 respondents in 1997.

High school GPA from respondent’s high school transcript, 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. 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.

Socio-emotional skills

Indicator that the respondent was ever held back a grade in 1^{st}-8^{th} grade, orthogononalized against ASVAB1

Indicator that respondent was ever suspended from school in 1^{st}-8^{th} grade, orthogononalized against ASVAB1

Indicator that respondent said they had ever stolen something worth $50 or less by min($8^{th}$ grade or age 15) , orthogononalized against ASVAB1

Indicator that respondent said they had ever intentionally destroyed or damaged someone else’s property by min(8^{th} grade or age 15) , orthogononalized against ASVAB1

Indicator that respondent has sex before the age of 15, orthogononalized against ASVAB1

Indicator that the NLSY interviewer rated the respondent as somewhat uncooperative (a score of 3-8 on a scale of 1=hostile to 10=very cooperative. 10 is the modal response) in any of the first 3 interviews, , orthogononalized against ASVAB1

Demographic characteristics

Sex

Race and ethnicity: indicators for white, black, non-white Hispanic, or other non-white

Family characteristics

Number of children age 18 and under living at the respondent’s address in 1997 (including the respondent)

Quartile (calculated within the weighted NLSY sample) of Total 1997 net worth for the household where the respondent lived in 1997. Taken from the parent survey where available or from the youth survey (98.6% from parent survey).

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 in last year of high school (Northeast, South, Midwest, or West)

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 (from 1990 census) in the census tract where the respondent lived during his last year of high school.

Indicator that the respondent did not live within a Metropolitan Statistical Area (MSA) in the fall before she finished high school.

Local college options

Average in-state tuition, by year, for public 4-year and 2-year schools is from the State of Washington Higher Education Coordinating Board. “In-state” tuition for District of Columbia residents is calculated as max(national average in-state tuition, national average out-of-state tuition - $10,000) in accordance with DC Tuition Assistance Grant Program. For each respondent, in-state tuition is the in-state tuition in the fall before he finished high school in the state where he lived that fall. All tuition is CPI-deflated to 1997 dollars.

Whether the student has a well-matched (defined as a college whose weighted quality percentile is within 20 percentage points of the student's ASVAB ability percentile) public or private college within 50 miles of the respondent's home in his last year of high school. Distance is measured from the center of the respondent's census tract. The measure for public colleges includes only colleges in the same state. The measure for private college does not impose this restriction.

Additional covariates included only in Tables 7 and 8

Indicators for quintile of family income in 1996. Quintile cutoffs are for families in 1996 as calculated by the Census Bureau.

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

Number of days per week that respondent's family did something religious (before end of high school)

Indicator that respondent answered yes to "In a typical week, did you spend any time taking extra classes or lessons for example, music, dance, or foreign language lessons?"

Sum of indicators that respondent had regular access at home to a computer, a dictionary, and a quiet place to study from 1997 survey

Indicator that respondent had ever had contact with her biological mother by the 1997 survey

Indicator that respondent had ever had contact with her biological father by the 1997 survey

Appendix Table 3: Principal Components of the 12 Test Sections of the ASVAB

	1st Component	2nd Component	Unexplained variance
Eigenvalue	7.18	1.36	
Total variance explained	59.8%	11.3%	
Eigenvectors:			
General Science	0.326	-0.114	21.9%
Arithmetic Reasoning	0.325	0.117	22.2%
Word Knowledge	0.322	-0.038	25.4%
Paragraph Comprehension	0.320	0.114	24.8%
Mathematics Knowledge	0.318	0.239	19.7%
Mechanical Comprehension	0.310	-0.162	27.4%
Electronics Information	0.304	-0.228	26.8%
Assembling Objects	0.273	0.107	45.1%
Shop Information	0.245	-0.462	27.9%
Numerical Operations	0.240	0.444	31.8%
Auto Information	0.225	-0.456	35.6%
Coding Speed	0.223	0.441	37.8%

Note: scores on each test component are adjusted for the age of the respondent when they took the test by regressing the score on age dummies and using the residuals for the principal components analysis. The first two principal components combined explain 71.1% of the total variance of the 12 test section scores.

Appendix Table 4: Principal Components of the College Quality Indices

	1st Component	Unexplained variance
Eigenvalue	2.10	
Total variance explained	52%	
Eigenvectors:		
Mean SAT	0.587	28%
Rejection rate	0.478	52%
Faculty/Student ratio	0.361	73%
Average faculty salaries	0.544	38%

Calculated from the 1,491 4-year colleges in IPEDS in 2008 with all four college quality proxies.

Appendix Table 5: Effect of College Quality and Ability on College Outcomes, quartile dummies

	Graduate within 5 yrs	Enroll in grad sch	Transfer up	Transfer down
ASVAB q1, Quality q2	0.096	0.074	-0.044	0.123
	(0.046)	(0.045)	(0.023)	(0.033)
ASVAB q1, Quality q3	0.165	0.192	-0.146	0.104
	(0.057)	(0.050)	(0.036)	(0.040)
ASVAB q1, Quality q4	0.270	0.137	-0.178	0.170
	(0.069)	(0.065)	(0.048)	(0.050)
ASVAB q2, Quality q1	0.095	0.077	0.014	0.027
	(0.049)	(0.048)	(0.020)	(0.037)
ASVAB q2, Quality q2	0.190	0.117	-0.056	0.080
	(0.050)	(0.049)	(0.024)	(0.038)
ASVAB q2, Quality q3	0.212	0.133	-0.074	0.118
	(0.048)	(0.046)	(0.024)	(0.036)
ASVAB q2, Quality q4	0.298	0.221	-0.106	0.094
	(0.055)	(0.052)	(0.034)	(0.044)
ASVAB q3, Quality q1	0.176	0.098	0.043	-0.034
	(0.056)	(0.055)	(0.023)	(0.047)
ASVAB q3, Quality q2	0.276	0.173	-0.018	0.054
	(0.050)	(0.049)	(0.024)	(0.039)
ASVAB q3, Quality q3	0.296	0.207	-0.103	0.097
	(0.047)	(0.046)	(0.030)	(0.037)
ASVAB q3, Quality q4	0.344	0.299	-0.126	0.068
	(0.050)	(0.046)	(0.032)	(0.039)
ASVAB q4, Quality q1	0.199	0.154	0.037	-0.013
	(0.064)	(0.066)	(0.027)	(0.057)
ASVAB q4, Quality q2	0.193	0.212	0.042	0.067
	(0.058)	(0.056)	(0.024)	(0.045)
ASVAB q4, Quality q3	0.327	0.274	-0.053	0.114
	(0.049)	(0.047)	(0.028)	(0.037)
ASVAB q4, Quality q4	0.365	0.326	-0.140	0.040
	(0.045)	(0.042)	(0.031)	(0.037)
Observations	2,405	2,405	2,361	2,388
R-squared	0.180	0.102	0.135	0.051

This table reports mean marginal effects. Estimates are weighted as described in the text and also include the set of covariates described in Appendix Table 2.

Appendix Table 6: Effect of College Quality and Ability on Earnings, quartile dummies

	Year 0-1	Year 2-3	Year 4-5	Year 6-7	Year 8-9	Year 10-11
ASVAB q1, Quality q2	0.522	-0.652	-1.047	2.554	2.224	3.821
	(0.623)	(0.916)	(1.196)	(1.649)	(1.962)	(3.736)
ASVAB q1, Quality q3	0.186	-0.768	-0.266	2.375	3.516	7.330
	(0.803)	(1.051)	(1.572)	(1.843)	(2.397)	(3.057)
ASVAB q1, Quality q4	0.049	-0.076	2.324	4.016	2.762	6.584
	(0.953)	(1.637)	(2.292)	(3.182)	(3.227)	(6.051)
ASVAB q2, Quality q1	1.107	0.332	-0.237	2.430	1.400	1.534
	(0.610)	(0.901)	(1.028)	(1.436)	(1.747)	(3.179)
ASVAB q2, Quality q2	-0.418	-1.285	-0.799	2.462	4.233	5.281
	(0.507)	(0.852)	(1.083)	(1.475)	(1.939)	(3.240)
ASVAB q2, Quality q3	-0.939	-1.682	-0.271	3.813	3.303	5.064
	(0.515)	(0.857)	(1.160)	(1.689)	(1.999)	(2.947)
ASVAB q2, Quality q4	0.147	-1.173	1.666	7.068	8.109	13.111
	(0.614)	(1.050)	(1.500)	(2.160)	(2.373)	(3.615)
ASVAB q3, Quality q1	0.656	0.087	-0.255	0.138	2.000	3.039
	(0.745)	(1.193)	(1.332)	(1.658)	(2.128)	(3.903)
ASVAB q3, Quality q2	-0.886	-2.846	-2.148	2.745	3.622	7.717
	(0.523)	(0.845)	(1.134)	(1.653)	(2.101)	(3.659)
ASVAB q3, Quality q3	-0.967	-2.177	-1.400	3.642	6.092	6.591
	(0.519)	(0.855)	(1.115)	(1.911)	(2.151)	(3.523)
ASVAB q3, Quality q4	-0.430	-2.173	0.222	6.000	10.515	19.507
	(0.627)	(0.897)	(1.340)	(2.115)	(2.582)	(4.588)
ASVAB q4, Quality q1	0.513	-0.279	-1.400	4.320	6.461	14.287
	(0.767)	(1.180)	(1.609)	(2.451)	(3.433)	(5.371)
ASVAB q4, Quality q2	-0.092	0.087	-0.767	3.110	4.797	4.669
	(0.642)	(1.088)	(1.405)	(1.994)	(2.864)	(3.747)
ASVAB q4, Quality q3	-1.005	-2.362	0.681	4.548	6.352	15.312
	(0.511)	(0.856)	(1.297)	(1.842)	(2.321)	(4.134)
ASVAB q4, Quality q4	-1.719	-3.754	-0.247	5.894	10.118	15.732
	(0.469)	(0.791)	(1.181)	(1.779)	(2.274)	(3.970)
Observations	2,332	2,289	2,228	2,141	1,872	958
R-squared	0.130	0.132	0.045	0.083	0.136	0.183

This table reports mean marginal effects. Estimates are weighted as described in the text and also include the set of covariates described in Appendix Table 2.