

Relevance of Education to Occupation: A New Empirical Approach Based on College Courses*

Aboozar Hadavand¹ and Pooya Almasi²

¹Corresponding Author: Bloomberg School of Public Health, Johns Hopkins University, hadavand@jhu.edu

²Department of Economics, Georgetown University, pa402@georgetown.edu

We introduce a new approach to measuring the match between education and occupation by using the number of college courses related to one's occupation. Previous studies have only considered the match between college "major" and occupation. Using NLSY97 data, we find that those who major in STEM and business are employed in occupations that are better matched to their field of study. Furthermore, the degree to which individuals are matched to their occupation remains the same over time except for business majors. We find that taking courses in college that are relevant to one's occupation is significantly associated with higher wages. Performing well in these courses is associated with an even higher wage premium. A student's wage increases, on average, by 1.6–2.9 percent for each matched course. This effect increases to 2.3–3.8 percent when we use a grade weighted measure of match.

JEL: J24, I26, J62

Keywords: Returns to Education; Skill Match; Horizontal Mismatch

I. Introduction

Economic returns to education depend on a number of factors including years, quality, and type of schooling. Variability of returns with respect to years of schooling has been a focal point of research since the early 1970s, yet fewer studies have factored in the relevance of the knowledge gained in college to occupations taken up later in life.

This paper contributes to the literature on knowledge-occupation matching and its effect on incomes. Following the terminology used in Nordin, Persson and Rooth (2010), we refer to this type of matching as "horizontal" matching distinguishing it from "vertical" matching, which looks at the match between years of schooling obtained and the years of schooling required for an occupation.

* We would like to thank Wim Vijverberg, John Robst, Judith Scott-Clayton, Randy Reback, Basit Zafar, John Winters, Homa Zarghamee, Dan Thomas, Ben Zweig, the two anonymous reviewers, and participants at LERA and EEA for their insightful comments. The data and the code used in this paper may be found here: <https://github.com/ahdvnd/Papers/tree/master/Mismatching>.

Where previous studies (Robst, 2007a; Nordin, Persson and Rooth, 2010) have defined horizontal matching as a discrete match between college major and occupation (a graduate with a degree in economics is matched if working as an economist, but not if working as a software engineer), we construct a non-binary measure of match using course-level data. In our analysis, a software engineer’s course match quality (CMQ) increases for every college course he or she has completed that is related to software engineering. Our first finding is that the CMQ varies significantly across majors. Not surprisingly, those who major in STEM and business are employed in occupations that are better matched to their field of study. The least “matched” individuals are those in humanities. Second, the degree to which individuals are matched to their occupation remains the same over time except for business majors.

We find a positive and significant relationship between CMQ and income for both men and women. Every additional relevant course increases wages by about 1.6–2.9 percent when we control for other individual characteristics. When we weight CMQ by the grades received in each course, the association is 20–40 percent stronger. In other words, more important than course enrollment is the performance in each course. It is not entirely clear, however, if these links are causal.

Measuring horizontal match at the course level rather than at the major level can provide a more nuanced understanding of the returns to college. If more granular measures of the college experience are shown to explain variation in wages, then this would provide evidence against the notion that returns to college are principally a matter of signaling. Furthermore, a focus on course-level horizontal match may also offer valuable insights into how colleges could structure degree requirements to improve occupational outcomes for their students. Generally speaking, educational mismatch is a crucial part of education economics because persistent mismatch can amplify the burdens of student debt, increase recruitment and training costs for employers, and result in structural unemployment and lower aggregate productivity.

Section II of this paper surveys the literature on horizontal matching and summarizes recent findings. Section III introduces our sources of data and explains how we measure CMQ. This section also shows how CMQ varies across our population and subgroups of the population. Section IV summarizes our findings and presents our analysis of the effect of CMQ on wages, and Section V concludes the paper and discusses the limitations of our study.

II. Background

Research examining post-secondary educational content and occupation matching began with three “pioneering” papers by Robst (2007a,b, 2008) in which he observed that individuals who attain the same years of schooling could, nevertheless, experience different returns to education based on the relevance of skills and knowledge acquired in school. In the first of these papers, Robst (2007a)

describes a match as a situation where a person's most recent field of study is related to their current occupation. Using self-reported data from the 1993 National Survey of College Graduates (NSCG), he reports that a fifth of his sample is completely mismatched and a quarter of his sample is partially mismatched. Furthermore, these "partially and completely mismatched men and women earn less than workers who report working in [their] field of study." We deviate from Robst's analyses in that, instead of using a self-reported measure of match, we are constructing our CMQ score using college transcripts and occupations reported in the National Longitudinal Survey of Youth (NLYS97).

In a study conducted in Sweden, Nordin, Perssona and Rooth (2010) use data collected from all individuals aged 28-39 and find that a horizontal mismatch results in a wage penalty of approximately 22 percent for men and 8 percent for women. Given the more vocational nature of Sweden's higher education system, it would be expected that the mismatch penalties in Sweden are higher than they are in the United States.¹

One shortcoming of the approaches taken in Nordin, Perssona and Rooth (2010) and similar studies, such as Boudarbat and Chernoff (2010) in Canada, and Abel and Deitz (2015) in the United States, is the use of non-continuous measures of match. In other words, an individual's college degree is either matched to their occupation or it is not, or, in some cases, perfectly matched, weakly matched or mismatched. Returning to our example, an individual with a degree in economics who works as a software engineer is more likely to be considered mismatched even if a bulk of their college coursework is utilized in their job. In response to this criticism, (Guvenen et al., 2015) use NLSY79 data to create a mismatch metric based on workers ability and their occupation and finds that eliminating the mismatch would raise the average wage of workers by 11 percent. The essence of their paper, however, is not the match between the knowledge acquired in college and occupation but instead the match between "ability" and occupation since they match occupation to math and verbal skills as measured by the ASVAB test and social skills as measured through a behavioral test.

Two other papers have attempted to measure the match between college education and occupation through a non-binary measure. Freeman and Hirsch (2008) use the U.S. Department of Labor's Occupational Information Network (O*NET) to define the knowledge content of each college major and their relevance to the knowledge required in each occupation. Similarly, Yakusheva (2010) and Gelblum (2014) use O*Net data to calculate a matching index reflecting the relevance of one's college major to their occupation. Both papers find that being well-matched is associated with higher wages.

Although these papers are closer to our study, the match is still defined between college major and occupation. The shortcoming of this approach is that the title

¹In Sweden, higher vocational education (HVE) is an alternative to academic track of higher education and is post-secondary education that combines theoretical and practical studies in cooperation with employers and industry.

of one’s college major is only a rough measure of the type of knowledge a person acquires in college. Consider another illustrative example in which a person who, despite having graduated in English, has taken numerous courses in economics. Based on the methods used by Freeman and Hirsch (2008) and Yakusheva (2010), this person is considered mismatched if s/he chooses to work as an economist. We argue that a better measure can be constructed by considering college coursework and by taking into account the grades received in each completed college course.

III. Data and Methodology

We use data from the National Longitudinal Survey of Youth’s 1997 cohort (NLSY97). NLSY97 is a longitudinal study that includes demographic, educational, and occupation related information for 8,984 individuals. To measure and control for ability, we use the Armed Forces Vocational Aptitude Battery (ASVAB) test.² In addition to these shared characteristics, NLSY97 includes a special Post-Secondary Transcript Study (PSTRAN), which collected available post-secondary transcript data for NLSY97 respondents who indicated some degree of post-secondary undergraduate education during Rounds 1-15 of the NLSY97 annual surveys. Transcripts were obtained for 3,818 individuals.³ Unfortunately, the NLSY97 data does not include any information on graduate-level coursework for those who pursued graduate degrees.

A. Determining the relevance of coursework to occupations

To match coursework to occupations, we use a crosswalk jointly developed by the National Center for Education Statistics (NCES) and the Bureau of Labor Statistics (BLS). This crosswalk links occupations with instructional programs that are considered “preparation directly for entry into and performance” in the associated job.⁴ Because the crosswalk uses taxonomies for instructional programs and occupations that differ from the NLSY data, some additional work was needed to pair the two sets of taxonomies.⁵

²Respondents in NLSY97 were asked to take this test in 1997, and 78% of the respondents took the test.

³We should note that of the individuals for whom transcript data was obtained, relevant data may be missing for some. During the PSTRAN study, multiple transcript requests were sent out if individuals had attended more than one undergraduate institution. In some cases, not all transcripts were obtained for reasons of non-response or refusal by the institution to release the transcripts.

⁴See <https://nces.ed.gov/ipeds/cipcode/resources.aspx?y=55> for the Guidelines for using the CIP to SOC Crosswalk.

⁵Courses in the NLSY PSTRAN Study are coded using the NCES’s 2010 College Course Map (CCM) and occupations in NLSY97 are coded using 2002 Census Occupation Codes. The crosswalk we use lists instructional programs and occupations using the Classification of Instructional Program (CIP) and Standard Occupational Classification Codes (SOC) respectively. To translate CCM to CIP we relied on the substantial overlap between the two taxonomies. CCM consists of a slightly longer set of codes, but both the CCM and CIP use a six-digit coding system where the first two digits indicate one of roughly 48 general areas of study. The first two digits of both taxonomies are identical, and for our purposes, we found that dropping the last four digits from both the CCM and CIP was the best solution

We then used the crosswalk to scan the college transcripts of individuals in the PSTRAN study counting the number of courses in their transcript that are relevant to their occupation. The CMQ metric we developed is simply the number of relevant courses listed in the student’s transcript.⁶

Table A1 in Appendix A shows the summary statistics for all variables used in this study for the year 2015 including the CMQ. Although CMQ’s range from zero to 119 across all individuals and all years, a closer investigation reveals that only a few individuals have CMQ scores above 100. When we drop these outliers, we see that, across our sample, the maximum CMQ is equal to 26 and the average CMQ is equal to 7.24.

In our analysis, CMQs for each person change over time due to changes in occupation. We can, therefore, create a longitudinal dataset containing Course Match Quality scores for each person for every year for which we have data on occupations. Our longitudinal data includes years 2003–2011, 2013, and 2015. In each year, we only include those who have an occupation and are included in the PSTRAN study. Figure 1 shows the density plot of CMQs across our sample for the years 2005, 2010, and 2015. We calculated CMQs across all years to observe the progression of matching over time. Note that for a significant share of individuals in our sample, CMQs are small. However, matching increases over time from 2005 to 2010 but remains almost the same between 2010 to 2015.

Mismatching between educational content and occupation can occur for many reasons. Robst (2008) argues that mismatching is a result of both supply-side and demand-side factors. Among the supply-side factors are changes in social norms, changes in interest, lack of information about the return to one’s major, prospects of better job conditions, educational outcomes at the end of college, and the lag between skill acquisition during college years and the demand for skills at the time of graduation. Mismatching can also occur due to general labor market conditions, discrimination, and other supply-side factors. Although we find little difference between the CMQs of men and women, course match quality for whites is, on average, two courses higher than blacks and Hispanics.⁷ The lower matching score for blacks and Hispanics could be due to factors such as poorer labor market conditions, less access to resources and information that guide the choice of which courses to enroll in, and economic considerations at the time of selecting majors.

for our study. The 2002 Census Codes were linked to the 2010 SOC Occupation Codes by combining several existing crosswalks. All of the crosswalks that we use in this paper can be found here: <https://github.com/ahdvnd/Papers/tree/master/Mismatching>.

⁶We initially considered calculating the number of courses matched as a share of all courses required for an occupation as our matching score. However, this method would lead to the same matching score for a person who has taken 1 out of 2 required courses and a person who has taken 20 out of 40 required courses.

⁷Robst (2007b) finds that women were somewhat more likely to report that they are matched to their occupation. 58 percent of women versus 53 percent of men in his sample reported that they are matched to their jobs. He argues, while men are more likely to be mismatched due to career enhancing opportunities like promotions and changes in career interests, women are more likely to report working conditions and working constraints as their reasons for being mismatched.

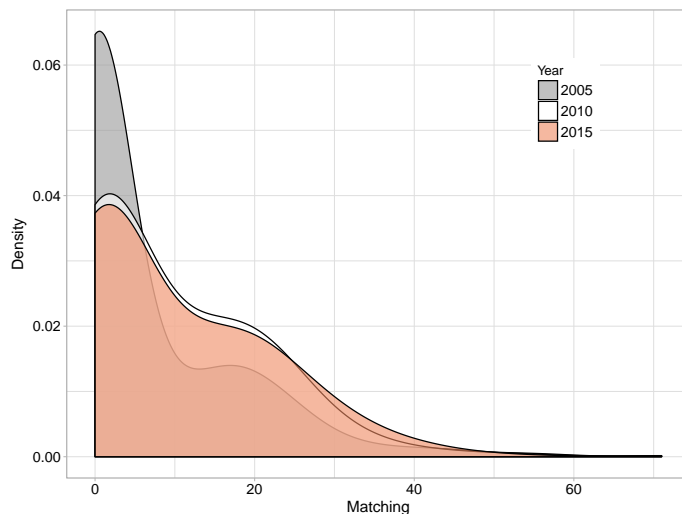


Figure 1. : Distribution of Course Match Quality (CMQ) for years 2005, 2010, and 2015. The CMQ metric is defined as the total number of occupation-relevant courses a student completed while enrolled in college. The CMQs are calculated only for those who have a bachelor’s degree in that year (and not more).

Figure 2 provides the distribution of the CMQS across different college major categories in the last year we have data available, 2015, which is relatively similar to previous years. STEM (science, technology, engineering, and math) and business graduates are, on average, more matched to their occupation. On the other hand, the most “mismatched” individuals are the ones who majored in social sciences, humanities, and the arts. Other researchers have found similar results. Abel and Deitz (2015) find that mismatching is the highest among those who major in the liberal arts. Robst (2007a) finds that mismatching is more prevalent among those who major in English and foreign languages, social sciences, and liberal arts. He argues that education-relevance mismatch is more prevalent among workers with degrees that provide general skills and less commonplace among graduates of majors that are more occupation-specific.

The distributions in Figure 1 are based on CMQs that do not take into account graduation dates. A better way of looking at how matching changes over an individual’s lifetime is to group individuals based on years since their college graduation. This is shown in Figure 3. Over time, those who majored in business and professional and technical fields became better matched to their occupations, while for other major groups CMQs remain more or less constant. We cannot find any explanation for the trend in business and other professional and technical graduates. The large variation in arts is due to the small number of individuals in our sample who majored in the arts, which resulted in high standard deviations.

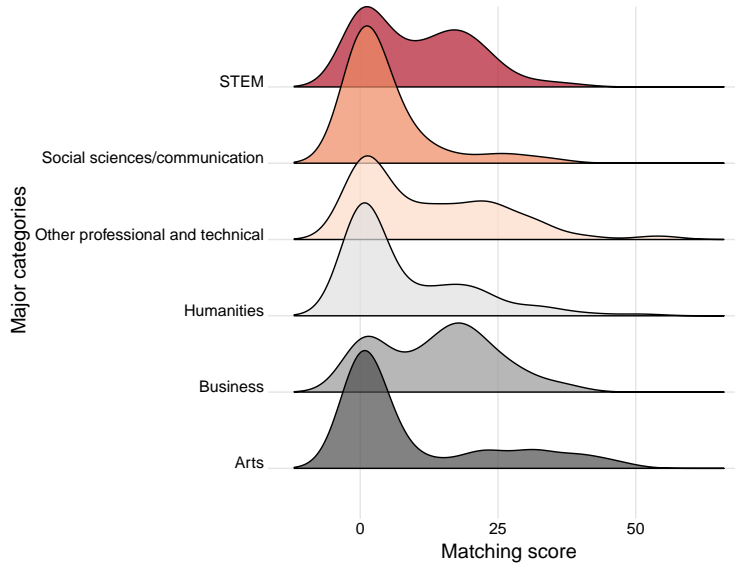


Figure 2. : Distribution of CMQs across different college major groups in 2015. CMQs were only calculated for those who have a bachelor’s degree in that year (and not more). College majors are categorized based on Table A5 in Appendix A.

B. Basic Model

Different forms of Mincer’s human capital model have been used to study the relationship between human capital and labor market earnings. The basic Mincer equation indicates the relationship between formal years of schooling and labor market earnings:

$$(1) \quad \log(Y_i) = \alpha + \beta S_i + \gamma_1 X_i + \gamma_2 X_i^2 + \phi' Z_i + u_i$$

where Y_i is labor market earnings, S_i is the number of years of schooling, X_i represents years of experience, and β is the average premium that comes with one more year of schooling conditional on controlling for individual-specific characteristics Z_i . If we assume that schooling increases the marginal product of labor and that we have a competitive market, then our estimate of β should be positive. The Mincer equation has been altered and other variables have been added to provide a more precise measure of productivity and the wage equation. Including years of schooling alone does not fully account for the applicability of what is learned in school to occupation.

In our empirical model, we try to fix this problem by adding the CMQ metric

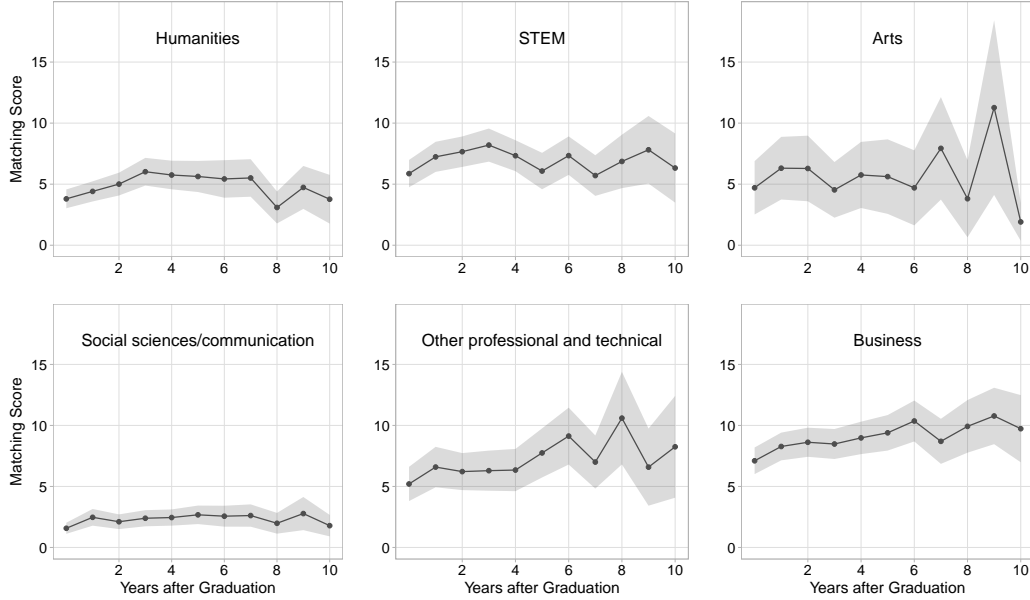


Figure 3. : Trends in the CMQ since the time of graduation from college. College majors are categorized based on Table A5 in Appendix A.

defined above to the wage equation. Since everyone in our sample has a college degree and no degree higher than a bachelor's degree, our primary interest is the effect of our horizontal matching score on the log of wages rather than the amount of schooling itself. The wage equation is then given as follows

$$(2) \quad \log(Y_i) = \alpha + \beta \text{ Matching Score}_i + \gamma_1 X_i + \phi' Z_i + u_i$$

where Y_i is the log of wages for individual i , X_i is the individual experience or years of occupation, and Z_i is a set of individual-specific controls.⁸

IV. Empirical Results

We regress the log of annual wages on CMQ, demographic variables (such as gender, race, and marital status), and other socioeconomic characteristics such as parental education and employment status. For each year, NLSY97 reports an individual's annual wages in the previous years. As a result of this and the fact that employment data are reported biennially after 2011, we were not able to use the annual wages reported in NLSY97 and had to instead calculate annual wages

⁸There is no need for inclusion of a nonlinear term for experience in our model since there are only small variations in ages across our sample.

based on hourly wages, the number of hours worked per week, and the number of weeks worked per year.⁹ Some of the variables in the NLSY97 data used in our analysis have many missing values. As a result, the samples used in the regressions are significantly smaller than the NLSY97 population with a college degree and above: 477 in 2005, 873 in 2010, and 778 in 2015 in our sample as opposed to more than 1,700 in the NLSY97 population with bachelor's degree and above. It is important to check whether our samples are similar to the population. Table A2 in Appendix A compares the two samples in 2015. Although the CMQs in our sample are higher on average by 1.3 points, regarding other covariates the sample used in our analysis seems to be a good representative of the broader NLSY97 sample.

Table 1 provides the estimation results of the basic regression using the CMQ and demographic variables (columns 1–3). The regression results are shown for years 2005, 2010, and 2015 for the sake of space. As can be seen, the coefficient on the CMQ is positive and statistically significant in all three regressions. Every additional relevant course, on average, increases annual wages by 1.8–2.5 percent depending on the year.¹⁰

If other unobserved individual characteristics are correlated with the number of relevant courses, their omission will bias our results. One such characteristic that is critical is individual-specific ability. Heckman and Vytlačil (2001) suggest that ability and schooling are not separable even if ability is observable. Robst (2007a) claims that neglecting ability in the analysis might cause a sorting bias that is due to the fact that individuals with higher ability are more likely to be better matched. Neglecting ability would cause an overestimation of the coefficient on CMQ. We follow the literature by Blackburn and Neumark (1993); Cameron and Heckman (2001); Yakusheva (2010), and others and use the ASVAB score described earlier to control for individual ability. In columns 4–6 of Table 1 we incorporate the ASVAB score as well as other individual and household controls including marital status, the number of children, and a variable for parental education.¹¹ Interestingly, in this new setup, the coefficient for the CMQ does not change significantly. An extra relevant course is still associated with 1.6–2.9 percent increase in annual wages. In our initial estimates, we observe a significant negative gender gap of around 16–23 percent, a dramatic adverse effect on labor market wages.¹² Columns 1–3 of Table 2 suggest that the quadratic term for the

⁹For instance, in 2013, annual wages are reported for 2012 whereas occupations are reported for 2013.

¹⁰We further performed two modifications. We dropped all observations that were in graduate school in the year used in the regression and although we lost some observations the results did not change. We also controlled for the total number of courses reported in the PSTRAN study for each student to see if it is likely that being highly matched is correlated with having taken more courses. The correlation between the two variables was quite low (0.05 in 2015) and including the total number of courses in the regression changed the estimates only by a little. The results are not provided here but are available upon request.

¹¹We use mother's education due to the higher number of observations for this variable compared to father's education.

¹²We also added a variable controlling for on-the-job training. The logic for adding this variable was that on-the-job training might compensate for being mismatched. In the NLSY97 data, individuals were

Table 1—: Regression of the log of annual wages on the matching score in different years

	2005	2010	2015	2005	2010	2015
	(1)	(2)	(3)	(4)	(5)	(6)
Matching Score	0.025*** (0.005)	0.018*** (0.003)	0.019*** (0.003)	0.029*** (0.005)	0.016*** (0.003)	0.019*** (0.003)
Female	-0.164** (0.074)	-0.215*** (0.051)	-0.232*** (0.052)	-0.177** (0.074)	-0.227*** (0.055)	-0.232*** (0.063)
Experience	0.024 (0.023)	0.054*** (0.014)	0.023** (0.009)	0.038 (0.026)	0.051*** (0.015)	0.022** (0.010)
Full-time	1.200*** (0.086)	1.485*** (0.081)	1.337*** (0.093)	1.184*** (0.081)	1.518*** (0.090)	1.311*** (0.107)
Black	-0.100 (0.103)	-0.055 (0.072)	-0.109* (0.064)	0.051 (0.102)	0.105 (0.083)	-0.034 (0.089)
Hispanic	-0.027 (0.115)	0.050 (0.081)	-0.104 (0.065)	0.031 (0.130)	0.091 (0.104)	-0.047 (0.088)
Constant	8.932*** (0.121)	8.641*** (0.156)	9.229*** (0.131)	8.631*** (0.291)	8.680*** (0.258)	9.139*** (0.220)
Other Controls	<i>No</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Observations	477	873	778	390	730	632
Adjusted R ²	0.409	0.484	0.440	0.469	0.497	0.442

Note: The dependent variable is log of annual wages. The CMQ is defined as the total number of relevant courses to a person's occupation. The CMQs are only for those who have a bachelor's degree in that year (and not more). Experience is in years. Full-time is a binary variable indicating whether a person works full-time or not. All individuals in the sample work at a job either part-time or full-time since an occupation title is required to be able to calculate the CMQ. The base category for race is white. Other covariates used in regressions 4–6 are a binary indicator for whether a person has an associate degree, ASVAB score, mother's education, marital status, number of children in the household, and a binary indicator for whether the person lives in rural or urban areas. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are robust.

CMQ is not statistically significant, therefore, there is no quadratic nonlinearity between CMQ and the log of wages.

We limit our sample to students who obtained a bachelor’s degree but did not pursue any further post-secondary education beyond a bachelor’s degree. We, however, add a variable controlling for whether the student had an associate degree prior to the bachelor’s degree. We find the effect of having an associate degree to be small and insignificant. Excluding the sample to only those with a bachelor’s degree and not more means that we are considering the effects of taking relevant coursework conditional on having attended college, and hence, our results say nothing of the effects that college goers have over those who lack a post-secondary degree. Individuals who obtained graduate degrees were left out of our sample as well because our transcript data only covers an individual’s undergraduate studies. As a result, we have no information about the relevance of graduate studies to one’s occupation in our data. Table A3 in Appendix A shows the result for when we relax this condition and include those with more than a bachelor’s degree and control for the level of education. The coefficients on the CMQ change only slightly.

The bilateral relationship between the log of wages and CMQ scores for men and women reveals a steeper positive-sloping line for women than men in 2015 as shown in Figure 4.

To investigate whether gender plays a role in the effect of matching on wages, we use a dummy variable representing gender along with interaction of this variable with the CMQ. The results of the estimation are provided in Table 2 columns 4–6. We find that the effect of the CMQ is larger for women, i.e., higher rewards (penalties) for being matched (mismatched) for women than men. However, the coefficient is not statistically significant in any of the regression models 4–6.

The important but statistically insignificant interaction terms suggest that the model might be miss-specified. For example, the effect of marital status or child-bearing is known to increase male earnings and reduce female earnings. To fully analyze whether CMQs are higher for women than for men we consider a gender-stratified model in which we run a regression separately for men and women. The results of the regression are provided in Table A4 in Appendix A. Similar to before, we find that the effect is larger for women than men by 10 percentage points in 2005 and 3 percentage points in 2015, however, since the confidence intervals of the coefficients for men and women overlap, they are not statistically significantly different at the 95% confidence level.

A. Controlling for College Majors and Occupations

We further question whether the identification we get for the high CMQs is driven by either a persons’ field of study or their occupation. In other words, we

asked whether they had ever received a training certificate or vocational license. We found that the effect was insignificant and adding the variable did not change the coefficient on the matching score.

Table 2—: Regression of the log of annual wages on the matching score for men and women in different years

	2005	2010	2015	2005	2010	2015
	(1)	(2)	(3)	(4)	(5)	(6)
Matching Score	0.045** (0.019)	0.045*** (0.014)	0.027** (0.012)	0.023*** (0.008)	0.014*** (0.005)	0.016*** (0.004)
Female	-0.175** (0.074)	-0.230*** (0.055)	-0.235*** (0.063)	-0.225** (0.096)	-0.250*** (0.079)	-0.268*** (0.088)
Matching Score ²	-0.001 (0.001)	-0.001 (0.001)	-0.0004 (0.001)			
Female x Matching Score				0.010 (0.010)	0.003 (0.006)	0.005 (0.006)
Constant	8.615*** (0.293)	8.665*** (0.257)	9.130*** (0.220)	8.644*** (0.293)	8.694*** (0.263)	9.157*** (0.220)
Other Controls	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Observations	390	730	632	390	730	632
Adjusted R ²	0.469	0.496	0.442	0.469	0.500	0.442

Note: The dependent variable is log of annual wages. The matching score is defined as the total number of relevant courses to a person's occupation. The matching scores are only for those who have a bachelor's degree in that year (and not more). Other covariates used are experience in years, a binary indicator for whether a person works full-time, race, a binary indicator for whether a person has an associate degree, ASVAB score, mother's education, marital status, number of children in the household, and a binary indicator for whether the person lives in rural or urban areas. All individuals in the sample work at a job either part-time or full-time since an occupation title is required to be able to calculate the matching score. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are robust.

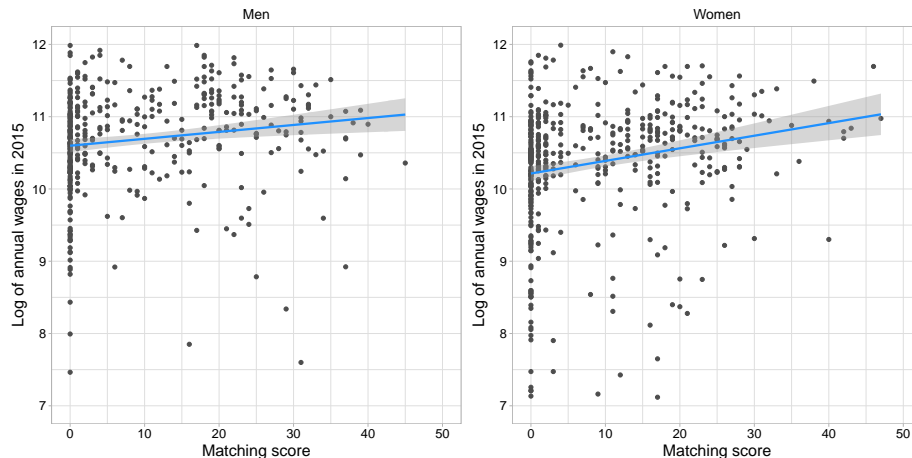


Figure 4. : Log annual wages versus CMQ for Men and Women in 2015. The CMQe is defined as the total number of relevant courses to a person's occupation. The matching scores are only for those who have a bachelor's degree in that year (and not more). All individuals in the sample work at a job either part-time or full-time since an occupation title is required to be able to calculate the CMQ.

want to test the hypothesis that matching based on coursework will still be a significant variable after controlling for college major or occupation. We define six categories of majors for individuals: arts, humanities, social sciences and communications, STEM (science, technology, engineering, and mathematics), business and management, and professional studies. Table A5 in Appendix A summarizes different majors that fall under each of these category. Similarly, we have 14 occupation categories shown in Table A6 in Appendix A. The adjusted- R^2 of the model with only major controls increases when we add the CMQ, which indicates that the CMQ can further explain the variations in wages.¹³

Furthermore, we add students' GPA as an indicator of how well the material is learned to the right-hand side of the regression. The regression results controlling for college majors are provided in Table 3 columns 1–3. Controlling for majors, we observe that the CMQ is still significant and positive. Regardless of a person's major, an additional matched course is associated with a 2.0 percent increase in annual wages in 2015.¹⁴ In the second regression, we control for the occupation categories mentioned above. The regression results are shown in Table 3 columns 4–6. The effect of the matching score on wages is 1.3 percent in 2015.

In summary, our results tend to be robust to fields of study and occupation

¹³adjusted- R^2 increase by 4 points in 2005, by 1 point in 2010, and by 3 points in 2015.

¹⁴Another question is whether the effect of matching varies by major. In order to do this, we interacted major dummies with the CMQ. The result is not shown here since we found no statistically significant interaction coefficient but are available upon request.

types, which indicates that although the number of *required* courses for each occupation may vary depending on the college major or the occupation, the effect of matching does not change when we control for major and occupation categories.

Table 3—: Regression of the log of annual wages on the matching score with major and occupation controls in different years

	2005	2010	2015	2005	2010	2015
	(1)	(2)	(3)	(4)	(5)	(6)
Matching Score	0.028*** (0.006)	0.014*** (0.004)	0.020*** (0.004)	0.024*** (0.006)	0.016*** (0.004)	0.013*** (0.004)
Major Controls	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>No</i>	<i>No</i>	<i>No</i>
Occupation Controls	<i>No</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Other Controls	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Observations	351	649	547	346	670	632
Adjusted R ²	0.476	0.507	0.457	0.506	0.501	0.464

Note: The dependent variable is log of annual wages. The matching score is defined as the total number of relevant courses to a person’s occupation. The matching scores are only for those who have a bachelor’s degree in that year (and not more). Other covariates used are experience in years, a binary indicator for whether a person works full-time, race, a binary indicator for whether a person has an associate degree, ASVAB score, mother’s education, marital status, number of children in the household, and a binary indicator for whether the person lives in rural or urban areas. All individuals in the sample work at a job either part-time or full-time since an occupation title is required to be able to calculate the matching score. We control for GPA in regression 1–3. Major and occupation categories are summarized in Tables A5 and A6 in Appendix A. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are robust.

B. Grade-Based Matching Score

We recognize that merely taking courses does not necessarily reflect an individual’s ability to possess the skills and knowledge required in their jobs. There is a battery of research investigating the relationship between academic achievement and earnings. Wise (1975), Jones and Jackson (1990), and James et al. (1989) are among those who find positive effects for academic achievement such as GPA on wages. However, to our knowledge, no literature investigates academic performance in the context of mismatching.

To address this, we calculate a modified CMQ that is weighted by the grades an individual obtained in each of their relevant courses. We argue that this grade-weighted CMQ, which is highly correlated with the basic CMQ (correlation is 0.94), is a better measure of skill relevance and is likely a better indicator of future wages. Merely enrolling in a course does not sufficiently capture knowledge gained through education. Our grade-weighted CMQ is created in the same way as our basic CMQ, except that we weigh the binary indicator of enrollment in each course by the grade received in that course with grades scaled between 0 and

1. Table 4 includes the regression results for our grade-weighted matching score.

Table 4—: Regression of the log of annual wages on the grade-weighted matching score in different years

	2005	2010	2015	2005	2010	2015
	(1)	(2)	(3)	(4)	(5)	(6)
Grade matching Score	0.034*** (0.007)	0.026*** (0.004)	0.025*** (0.004)	0.038*** (0.007)	0.023*** (0.004)	0.023*** (0.005)
Female	-0.183** (0.075)	-0.209*** (0.051)	-0.222*** (0.052)	-0.198*** (0.075)	-0.219*** (0.056)	-0.218*** (0.063)
Experience	0.021 (0.023)	0.053*** (0.014)	0.025*** (0.009)	0.029 (0.026)	0.050*** (0.015)	0.022** (0.010)
Full-time	1.194*** (0.087)	1.461*** (0.079)	1.341*** (0.093)	1.171*** (0.083)	1.482*** (0.086)	1.331*** (0.108)
Black	-0.049 (0.104)	-0.063 (0.071)	-0.142** (0.063)	0.102 (0.098)	0.113 (0.083)	-0.074 (0.088)
Hispanic	0.026 (0.113)	0.043 (0.080)	-0.123* (0.065)	0.082 (0.127)	0.096 (0.103)	-0.073 (0.087)
Constant	8.964*** (0.124)	8.687*** (0.155)	9.234*** (0.130)	8.679*** (0.290)	8.750*** (0.258)	9.152*** (0.220)
Other Controls	<i>No</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Observations	470	868	777	384	717	625
Adjusted R ²	0.397	0.477	0.436	0.454	0.488	0.438

Note: The dependent variable is log of annual wages. The grade-weighted matching score is defined as the total number of relevant courses to a person's occupation weighted by the grade received in each course. The grade-weighted matching scores are only for those who have a bachelor's degree in that year (and not more). Experience is in years. Full-time is a binary variable indicating whether a person works full-time or not. All individuals in the sample work at a job either part-time or full-time since an occupation title is required to be able to calculate the matching score. The base category for race is white. Other covariates used in regressions 4–6 are a binary indicator for whether a person has an associate degree, ASVAB score, mother's education, marital status, number of children in the household, and a binary indicator for whether the person lives in rural or urban areas. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are robust.

Columns 1–3 present our basic regression incorporating the grade weighted CMQ and columns 4–6 present the regression results with more covariates. Similar to our earlier results, the grade-weighted CMQ is highly significant and positive. The same set of variables are used in models in Table 1 and Table 4, therefore, they are comparable. Comparison of the coefficients reveals that the coefficients on the grade-weighted CMQ are between 0.4–0.9 percentage points larger than the coefficients on the basic CMQ. In other words, performance in a class is even more important than course enrollment when it comes to their impact on future earnings.

C. Panel Analysis of the Effect of the Matching Score on Wages

So far we have examined the effect of matching on wages separately in each year. Guvenen et al. (2015) find that skill mismatching in occupations held early in life

will have substantial effects on future earnings. A more comprehensive analysis of mismatching should investigate earnings over an individual’s lifetime. This is especially important since our results may partially reflect penalties associated with labor market dynamics in the wake of the 2007 financial crisis. Furlanetto and Groshenny (2016) and Zago (2015) find that skill mismatch has increased since the Great Recession.

More importantly, by using panel data, we can control for individual heterogeneity and seasonality, and mitigate the omitted variable bias problem. We use a fixed effects model for our panel analysis to estimate the impact of CMQ on wages. We consider a fixed-effects model mainly because of the plausible correlation between unobserved individual heterogeneity and our regressors. After performing a Hausman test (Hausman, 1978), we find that the fixed effect model is more suitable.

Table 5 presents the regression results of the fixed effects estimator in the sample. Specification (1) is the pooled OLS model controlling for some background characteristics such as mother’s education, ASVAB score, and other covariates (see Table 5 footnote). In specification (2) we use a fixed-effects model to control for individual fixed effects, virtually eliminating all the time-invariant variables including the unobservable variables that might correlate with the CMQ. In the third specification, we control only for time fixed effects and in specification (4) we control for both individual and time fixed effects. In all specifications, the effect of matching on wages remain almost the same, for every 10 matched courses taken during one’s post-secondary education, wages will increase by 8–9 percent. The last specification is the effect of the grade-weighted CMQ on wages which, similar to previous results, is slightly larger than the basic CMQ’s effect.

V. Concluding Remarks

Persistent skill mismatch is not only costly for employers and society as a whole and may also impact workers’ wages throughout their lifetime (Guvenen et al., 2015). Research on the extent of skill match and its impact on wages is scant, and the few studies that have focused on the link between skill match and wages have tended to focus on the match between one’s college major and occupational titles. However, a match based on major may not fully reflect the degree to which the skills and knowledge acquired in school are aligned with the skills and knowledge used in one’s occupation. In this paper, we take advantage of college transcript (PSTRAN) data for students in the National Longitudinal Survey of Youth 1997 (NLSY97) and calculate a Course Match Quality score between one’s college coursework and the skills required in one’s occupation. The advantage of our approach is in creating a non-binary and more content-oriented skill-job matching metric.

If skill mismatch penalizes wages and if it is commonplace across those with lower access to resources such as parental guidance and school counseling, skill mismatch can work as an unequalizing factor affecting the wage distribution.

Table 5—: Panel regression

	(1)	(2)	(3)	(4)	(5)
Matching Score	0.008*** (0.001)	0.009*** (0.002)	0.008*** (0.001)	0.009*** (0.002)	
Grade matching Score					0.011*** (0.003)
Female	-0.174*** (0.032)		-0.154*** (0.033)		
Experience	0.051*** (0.005)	0.085*** (0.006)	0.030*** (0.007)	0.224*** (0.047)	0.224*** (0.047)
Full-time	1.445*** (0.040)	1.138*** (0.035)	1.450*** (0.040)	1.116*** (0.035)	1.116*** (0.036)
Constant	8.653*** (0.110)				
Other Controls	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Individual FE	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>Yes</i>
Time FE	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Observations	5,854	6,959	5,854	6,959	6,942
Adjusted R ²	0.478	0.216	0.459	0.169	0.166

Note: Data includes years 2006–2011, 2013, and 2015. There is no data available for years 2012 and 2014. The dependent variable is the log of annual wages. The matching score is defined as the total number of relevant courses to a person’s occupation. The matching scores are only for those who have a bachelor’s degree in that year (and not more). Other covariates used are a binary indicator for whether a person has an associate degree, ASVAB score, mother’s education, marital status, number of children in the household, and a binary indicator for whether the person lives in rural or urban areas. All individuals in the sample work at a job either part-time or full-time since an occupation title is required to be able to calculate the matching score. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are robust.

At the macroeconomic level, skill mismatch contributes to structural unemployment and under-utilization of the workforce and lowers aggregate productivity. Although on-the-job training and apprenticeships can improve the mismatch between skill and occupation, major/coursework guidance and career counseling to young people and their families in both high school and college levels can perhaps help students make better-informed decisions for their education and career paths. Other policy recommendations include: encouraging coursework and skills that are relevant to a broader set of occupations that can foster mobility within occupations and improve matching, and designing curricula that are more responsive to rapid changes in the skills needed in the economy.

There are, however, some caveats in our analysis. A limitation of our study is the omission of college dropouts. As Weiss (1995) argues, it may be the case that college dropouts are more affected by coursework than those with a degree. As a result, the effect of relevant coursework on earnings may be underestimated.

Second, due to the restricted nature of the NLSY97 data, we were unable to identify the colleges and universities attended by individuals in our sample and, therefore, were not able to control for the quality of education (for example, by controlling for rankings of the colleges individual students attended). Although family background may be a good proxy for this, we think the inclusion of quality may alter our findings. One reason for this bias is the potential correlation between mismatching and school ranking. For instance, a graduate of a more prestigious school may be more likely to be mismatched since they have the advantage of choosing a broader range of occupations. Conversely, a graduate of a less prestigious college may be forced to accept an occupation unrelated to her major due to fewer labor market opportunities.

Lastly, the endogeneity of our matching variable can lead to biased estimations. There are unobserved variables such as employee training, match quality, or other such similar considerations. We controlled for employer learning by adding a binary variable indicating whether the person has had any on-the-job training but, we didn't find any statistically significant effects. However, we believe this binary indicator does not sufficiently reflect on-the-job training. To remedy some of this endogeneity, we considered a fixed effects model in which all the omitted variables that are time-invariant including any unobserved time-invariant variables that are correlated with the matching variable are eliminated from the model. Although, we observe a drop in the effect of matching on wages in the panel analysis compared to the cross-sectional one, the effect after controlling for individual and time fixed effects remains highly statistically significant and economically meaningful.

REFERENCES

- Abel, Jaison R, and Richard Deitz. 2015. "Agglomeration and job matching among college graduates." *Regional Science and Urban Economics*, 51: 14–24.

- Blackburn, McKinley L, and David Neumark.** 1993. "Omitted-ability bias and the increase in the return to schooling." *Journal of labor economics*, 11(3): 521–544.
- Boudarbat, Brahim, and Victor Chernoff.** 2010. "The determinants of education-job match among Canadian university graduates."
- Cameron, Stephen V, and James J Heckman.** 2001. "The dynamics of educational attainment for black, hispanic, and white males." *Journal of political Economy*, 109(3): 455–499.
- Freeman, James A, and Barry T Hirsch.** 2008. "College majors and the knowledge content of jobs." *Economics of Education Review*, 27(5): 517–535.
- Furlanetto, Francesco, and Nicolas Groshenny.** 2016. "Mismatch Shocks and Unemployment During the Great Recession." *Journal of Applied Econometrics*, 31(7): 1197–1214. jae.2498.
- Gelblum, Madeleine.** 2014. "The Early Impact of Postsecondary Career and Technical Education: Do Workers Earn More in Occupations Related to Their College Major? A CAPSEE Working Paper." *Center for Analysis of Postsecondary Education and Employment*.
- Güvenen, Fatih, Burhanettin Kuruscu, Satoshi Tanaka, and David Wiczer.** 2015. "Multidimensional skill mismatch." National Bureau of Economic Research.
- Hausman, Jerry A.** 1978. "Specification tests in econometrics." *Econometrica: Journal of the econometric society*, 1251–1271.
- Heckman, James, and Edward Vytlacil.** 2001. "Identifying the role of cognitive ability in explaining the level of and change in the return to schooling." *Review of Economics and Statistics*, 83(1): 1–12.
- James, Estelle, Nabeel Alsalam, Joseph C. Conaty, and Duc-Le To.** 1989. "College Quality and Future Earnings: Where Should You Send Your Child to College?" *The American Economic Review*, 79(2): 247–252.
- Jones, Ethel B., and John D. Jackson.** 1990. "College Grades and Labor Market Rewards." *The Journal of Human Resources*, 25(2): 253–266.
- Nordin, Martin, Inga Persson, and Dan-Olof Rooth.** 2010. "Education-occupation mismatch: Is there an income penalty?" *Economics of Education Review*, 29(6): 1047–1059.
- Robst, John.** 2007a. "Education and job match: The relatedness of college major and work." *Economics of Education Review*, 26(4): 397–407.

- Robst, John.** 2007b. “Education, College Major, and Job Match: Gender Differences in Reasons for Mismatch.” *Education Economics*, 15(2): 159–175.
- Robst, John.** 2008. “Overeducation and college major: Expanding the definition of mismatch between schooling and jobs.” *The Manchester School*, 76(4): 349–368.
- Weiss, Andrew.** 1995. “Human capital vs. signalling explanations of wages.” *The Journal of Economic Perspectives*, 9(4): 133–154.
- Wise, David A.** 1975. “Academic Achievement and Job Performance.” *The American Economic Review*, 65(3): 350–366.
- Yakusheva, Olga.** 2010. “Return to college education revisited: Is relevance relevant?” *Economics of Education Review*, 29(6): 1125–1142.
- Zago, Riccardo.** 2015. “Skill Mismatch, Job Polarization and the Great Recession.” IZA Conference.

APPENDIX A: ADDITIONAL TABLES

Table A1—: Variable Definitions and Descriptive Statistics in year 2015.

Variable	Definition	Mean	St.Dev	Min	Max
Matching Score	Total no. of relevant courses to a person's occupation	7.24	8.47	0	26
Grade Matching Score	Total no. of relevant courses to a person's occupation weighted by grade in those courses	4.38	5.70	0	19
Total Number of Courses	Total number of college courses taken by the student at the undergraduate level	48.51	17.82	2	130
Log Wages	Log of annual wages and salaries	10.46	0.96	3.29	12.43
Full Time	Works more than 40 wks per yr and at least 35 hours per wk	0.78	0.42	0	1
Experience	Number of years worked at an employee-type job since age 20	10.76	2.65	1.27	15.96
Female		0.55	0.50	0	1
Race	Individual's race categories				
White		0.65	0.48	0	1
Black		0.19	0.39	0	1
Hispanic		0.16	0.37	0	1
Married		0.55	0.50	0	1
Number of Children	Number of household members under the age of 18	0.99	1.13	0	5
Urban	Residence in an urban area	0.86	0.35	0	1
Associate	Whether individual received an Associate's degree before getting his/her Bachelor's	0.16	0.37	0	1
Currently in Grad School	Whether individual is enrolled in a graduate program	0.03	0.18	0	1
ASVAB Score	Combined ASVAB math and verbal score percentage	65.93	23.62	3	100
Mother's Highest Grade	Highest grade completed by individual's biological mother, range is from 1 for first grade to 20 for eight years of postsecondary education or more	13.73	2.86	2	20
Major	Major categories of individual's bachelor's degree				
Arts		0.023	0.15	0	1
Business		0.098	0.30	0	1
Humanities		0.690	0.46	0	1
Professionals		0.100	0.30	0	1
Social Sci. & Communications		0.031	0.17	0	1
STEM		0.053	0.22	0	1

Note: Total number of observations for the variable matching is 819. For other variables the number of observations can be less due to the existence of missing values. The data only includes those who hold a bachelor's degree in 2015 and are included in the PSTRAN transcript study of NLSY97. Major and occupation categories are summarized in Tables A5 and A6 in Appendix A.

Table A2—: Comparison of the sample used in the analysis and the NLSY97 population with college degree and above in 2015

Panel A: Sample used in the study					
Statistic	N	Mean	St. Dev.	Min	Max
Annual earnings	778	48,142	34,130	26	249,997
Female	778	0.56	0.50	0	1
Experience	778	10.76	2.65	1.27	15.96
Full-time	778	0.78	0.42	0	1
Black	778	0.18	0.39	0	1
Hispanic	778	0.16	0.37	0	1
Associate Degree	778	0.17	0.37	0	1
Panel B: The NSLY97 population with bachelor's degree and above					
Statistic	N	Mean	St. Dev.	Min	Max
Annual earnings	1,710	50,845	43,503	0	486,542
Female	1,798	0.57	0.50	0	1
Experience	1,686	10.46	2.82	0.00	16.04
Full-time	1,787	0.71	0.45	0	1
Black	1,798	0.19	0.39	0	1
Hispanic	1,798	0.14	0.35	0	1
Associate Degree	1,798	0.13	0.34	0	1

Note: Panel (A) includes those in the NLSY97 study with bachelor's degree (and not above) in 2015 who participated in the transcript (PSTRAN) study. Those who identified "other" as their race are dropped from the sample. Outliers in the matching score and log of annual wages are removed from the sample. Panel (B) includes those in the NLSY97 study with bachelor's degree and above. Those who identified "other" as their race are dropped. Data in both years are from 2015. Experience is in years and full-time is a binary variable indicating whether the person works full-time. Associate is a binary variable indicating whether a person has an associate degree. The base category for race is white.

Table A3—: Regression of the log of annual wages on the matching score in different years for those with bachelor's degree and beyond

	2005	2010	2015	2005	2010	2015
	(1)	(2)	(3)	(4)	(5)	(6)
Matching Score	0.024*** (0.005)	0.018*** (0.003)	0.015*** (0.002)	0.028*** (0.005)	0.016*** (0.003)	0.015*** (0.003)
Female	-0.178* (0.072)	-0.199*** (0.046)	-0.192*** (0.043)	-0.198*** (0.072)	-0.193*** (0.048)	-0.193*** (0.050)
Experience	0.023 (0.022)	0.050*** (0.012)	0.022** (0.009)	0.037 (0.026)	0.049*** (0.013)	0.021** (0.010)
Full-time	1.224*** (0.085)	1.431*** (0.073)	1.313*** (0.072)	1.219*** (0.079)	1.462*** (0.080)	1.308*** (0.083)
Black	-0.068 (0.101)	-0.039 (0.063)	-0.151*** (0.058)	0.090 (0.099)	0.098 (0.072)	-0.090 (0.078)
Hispanic	-0.003 (0.114)	0.047 (0.073)	-0.084 (0.055)	0.079 (0.128)	0.099 (0.091)	-0.027 (0.073)
Mater's Degree	0.206 (0.132)	0.223*** (0.056)	0.288*** (0.045)	0.218 (0.149)	0.205*** (0.061)	0.252*** (0.051)
Ph.D. Degree		0.961*** (0.333)	0.289*** (0.106)		0.992** (0.498)	0.260** (0.129)
Professional Degree	0.579*** (0.833)	0.732*** (0.186)	0.613*** (0.150)	0.445 (0.796)	0.859*** (0.108)	0.578*** (0.169)
Constant	8.927*** (0.119)	8.705*** (0.138)	9.257*** (0.122)	8.556*** (0.291)	8.641*** (0.218)	9.159*** (0.204)
Other Controls	<i>No</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Observations	498	1,093	1,135	409	921	935
Adjusted R ²	0.417	0.471	0.446	0.476	0.499	0.441

Note: The dependent variable is log of annual wages. The matching score is defined as the total number of relevant courses to a person's occupation. The matching scores are for those who have a bachelor's degree in that year or above in that year. Experience is in years. Full-time is a binary variable indicating whether a person works full-time or not. All individuals in the sample work at a job either part-time or full-time since an occupation title is required to be able to calculate the matching score. The base category for race is white. Other covariates used in regressions 4-6 are a binary indicator for whether a person has an associate degree, ASVAB score, mother's education, marital status, number of children in the household, and a binary indicator for whether the person lives in rural or urban areas. The base category for education is those with bachelor's degree. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are robust.

Table A4—: Regression of the log of annual wages on the matching score in different years for stratified samples of men and women

	Women			Men		
	2005 (1)	2010 (2)	2015 (3)	2005 (4)	2010 (5)	2015 (6)
Matching Score	0.033*** (0.006)	0.015*** (0.004)	0.020*** (0.004)	0.023*** (0.008)	0.015*** (0.005)	0.017*** (0.004)
Constant	8.390*** (0.426)	8.716*** (0.299)	8.722*** (0.294)	8.658*** (0.442)	8.328*** (0.436)	9.309*** (0.324)
Other Controls	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>
Observations	244	422	356	146	308	276
Adjusted R ²	0.439	0.515	0.475	0.506	0.432	0.307

Note: The dependent variable is log of annual wages. The matching score is defined as the total number of relevant courses to a person's occupation. The matching scores are only for those who have a bachelor's degree in that year (and not more). Other covariates used are experience in years, a binary indicator for whether a person works full-time, race, a binary indicator for whether a person has an associate degree, ASVAB score, mother's education, marital status, number of children in the household, and a binary indicator for whether the person lives in rural or urban areas. All individuals in the sample work at a job either part-time or full-time since an occupation title is required to be able to calculate the matching score. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are robust.

Table A5—: Definition of major categories

Major Categories	Majors
Arts	Visual and Performing Arts
Humanities	Foreign Languages, Literature, and Linguistics; Family and Consumer Sciences/Human Sciences; English Language and Literature/Letters; Liberal Arts and Sciences, General Studies and Humanities; Library Science; Multi/Interdisciplinary Studies; Philosophy and Religious Studies; Theology and Religious Vocations; History
Social Sciences and Communications	Area, Ethnic, Cultural, Gender, and Group Studies; Communication, Journalism, and Related Programs; Psychology; Social Sciences
Science, Technology, Engineering, and Mathematics	Agriculture, Agriculture Operations, and Related Services; Natural Resources and Conservation; Communications Technologies/Technicians and Support Services; Computer and Information Sciences and Support Services; Engineering; Engineering Technologies and Engineering-Related Fields; Biological and Biomedical Sciences; Mathematics and Statistics; Military Technologies and Applied Sciences; Science and Technologies/Technicians; Architecture and Related Services
Business and Management	Business, Management, Marketing, and Related Support Services
Professionals	Personal and Culinary Services; Education; Legal Professions and Studies; Parks, Recreation and Leisure Studies; Basic Skills and Developmental/Remedial Education; Citizenship Activities; Health-Related Knowledge and Skills; Interpersonal and Social Skills; Leisure and Recreational Activities; Personal Awareness and Self-Improvement; Physical Sciences; Homeland Security, Law Enforcement, Firefighting and Related Protective Services; Public Administration and Social Service Professions; Construction Trades; Mechanic and Repair Technologies/Technicians; Precision Production; Transportation and Materials Moving; Health Professions and Related Programs; High School/Secondary Diplomas and Certificates; Other majors

Note: Classifications by authors and are based on College Course Map.

Table A6—: Definition of occupation categories

Occupation Categories	Occupations
Management/Business	Management occupations; Business and financial operations occupations
Computer and mathematical occupations	Computer scientists and systems analysts; Network systems and data analysts; Mathematicians; Statisticians
Architecture and engineering occupations	Architects, except naval; Engineers; Engineering technicians, except drafters
Life, physical, and social science occupations	Agricultural and food scientists; Biological scientists; Medical scientists; Astronomers and physicists; Chemists and materials scientists; Environmental scientists and geoscientists; Physical scientists; Economists; Psychologists; Sociologists
Community and social services occupations	Counselors; Social workers; Religious workers
Legal occupations	Lawyers; Judges, magistrates, and other judicial workers; Paralegals and legal assistants
Education, training, and library occupation	Teachers; Archivists, curators, and museum technicians; Librarians; Library technicians; Teacher assistants
Arts, design, entertainment, sports, and media occupations	Artists and related workers; Actors; Producers and directors; Athletes, coaches, umpires, and related workers; Dancers and choreographers; News analysts, reporters and correspondents; Writers and authors; Miscellaneous media and communication workers
Healthcare practitioner and technical occupations	Dentists; Dietitians and nutritionists; Pharmacists; Physicians and surgeons; Physical therapists; Health diagnosing and treating practitioners
Service occupations	Healthcare support occupations; Protective service occupations; Food preparation and serving related occupations; Building and grounds cleaning and maintenance occupations; Personal care and service occupations
Sales and office occupations	Sales and related occupations; Office and administrative support occupations
Natural resources, construction, and maintenance occupations	Farming, fishing, and forestry occupations; Construction and extraction occupations; Installation, maintenance, and repair occupations
Production and transportation occupations	Production occupations; Transportation and material moving occupations
Armed Forces	

Note: Classifications based on Census 2002 Occupational categories, with Standard Occupational Classification (2000 SOC).