Uncertainty, Learning, and the Unemployment-Education Gap Over the Life Cycle*

Jie Duan†
Zhejiang University

Paul Jackson‡
National University of Singapore

July 2025

Abstract

We propose that college graduates enter the labor market with less uncertainty regarding which career they are most productive in, and study how this characteristic contributes to the unemployment-education gap. We document several novel facts to support our hypothesis. Notably, college graduates predict their occupation more accurately than those without a college degree. We then develop and calibrate a life cycle search model featuring differences in uncertainty by education and learning about one's best career fit. Our quantitative analysis suggests large disparities in uncertainty by education, and that such differences can explain a sizeable portion of the unemployment-education gap.

JEL Classification: E24; J24; J62; J64
Keywords: Unemployment; Education; Learning; Life Cycle

*We thank Ismail Baydur, Carlos Carrillo-Tudela, Chaoran Chen, Kevin Devereux, David Dorn, Ying Feng, Siddharth George, Aspen Gorry, Benjamin Lester, Francesco Lippi, Chen Liu, Simon Mongey, Xincheng Qiu, Michelle Rendall, Serene Tan, Shu Lin Wee, Donghai Zhang, Lichen Zhang, Shengxing Zhang, and audiences at the Asian and European meetings of the Econometric Society, Bank of Canada, Georgian Economic Association, Monash Macro and Money Workshop, NUS, NTU, OzMac (Sydney), PHBS, SUFE, SWUFE, UIBE, WAMS (Perth), and Xiamen U. for their insightful comments. The first version of this paper was presented under the title "On the Unemployment-Education Gap." Errors are our own.

†School of Economics, Zhejiang University. West Quad, 866 Yuhangtang Road, Hangzhou, 310058, China. *Email*: duanjie@u.nus.edu.

‡Department of Economics, National University of Singapore. Blk AS2, #04-22, 1 Arts Link, Singapore 117570. *Email*: ecspgj@nus.edu.sg.

1 Introduction

The unemployment rate among those with at least a bachelor's degree in the US is 2.7%, while it is nearly 7% for those without a degree.[1] Moreover, most of the unemployment-education gap is driven by higher separation rates among non-college workers (Elsby et al., 2010). While these facts are well documented, little quantitative research has been done to explain them. This is surprising, as the unemployment rate is one of the most paid attention to measures of labor market performance, unemployment spells are associated with sharp and persistent declines in earnings (Guvenen et al., 2021; Jarosch, 2023), and heterogeneity in separation rates play an important role in explaining variation in lifetime earnings (Ozkan et al., 2023). Additionally, identifying the sources of the unemployment-education gap has the potential to deepen our understanding of the differences between workers with and without a college degree, why their labor market outcomes are vastly different, and inform policies which aim to reduce unemployment. Therefore, this paper's objectives are to (i) propose and provide empirical support for a novel mechanism to explain the unemployment-education gap and (ii) evaluate its quantitative role within a search model of unemployment.

Our hypothesis is that college graduates enter the labor market with a clearer understanding of which career is their best fit (i.e., the career they are most productive in).[2] As such, they (i) enter the labor market having narrowed down the set of careers that are potentially their best fit and (ii) can quickly decipher whether a career is their best fit or not. We refer to these differences between college and non-college workers as the *uncertainty channel*. The connection between the uncertainty channel and the unemployment-education gap is straightforward. If college workers begin their career with fewer potential best fits, then they are less likely to learn they are not in their best fit and subsequently separate from their job, thereby becoming unemployed. Additionally, a faster learning speed allows college workers to find their best fit, which they are less likely to separate from, earlier in their work-life.

We provide empirical support for the uncertainty channel. Our most direct evidence comes from the National Longitudinal Survey of Youth 1979 (NLSY79), where we document that college graduates form more accurate expectations about their future occupation. In our preferred measure of forecast errors, the cosine similarity in skill and task requirements between occupations (Gathmann and Schönberg, 2010; Baley et al., 2022),

[1]Unemployment rates are from the Current Population Survey between 1976-2019. See Section 2.

[2]Broadly speaking, a career is a set of occupations which share a similar composition of skill requirements. Section 2.3.1 provides a precise definition. Throughout, we use "true calling", "best fit" and "good fit" interchangeably. The terminology follows Gervais et al. (2016).

forecast errors are 32% smaller among college graduates.

Further, we compile a set of evidence from the NLSY79 and Current Population Survey (CPS) which indirectly support the uncertainty channel. There are two main supporting facts. First, the unemployment-education gap narrows over the life cycle. Intuitively, as non-college workers begin with higher uncertainty, they experience more separations early on and gradually catch up to college workers as they sample careers, experience fewer separations, and exhibit lower unemployment rates. This is consistent with how separation, unemployment, and career mobility rates behave by age and educational attainment in the data. Second, prior work experience is associated with a lower separation rate. Moreover, this correlation is stronger for non-college workers. While the former has been documented (Topel and Ward, 1992), the latter is a new fact that is aligned with the uncertainty channel: if non-college workers rely more on work experience to learn their best fit, then prior experience should be associated with a more pronounced decline in separations for workers without a college degree.

Having established empirical support for the uncertainty channel, we proceed to develop a life cycle search model of unemployment. Workers are assigned one best fit, where they are most productive, out of a set of careers. Their best fit is initially unknown and workers sample careers to learn their true calling as in Gervais et al. (2016). If the worker learns they are not in their best fit, they can destroy the match in favor of becoming unemployed and sampling a different career. Separations and unemployment decrease with age as older workers are more likely to have found their best fit.

Workers are heterogeneous in their educational attainment (college and non-college). There are three exogenous differences by education. We assume (and find in the calibration) that college workers (i) are more productive in their best fit, (ii) enter the labor market with fewer careers that are potentially their best fit, and (iii) learn their career fit at a higher rate. The second and third differences encompass the uncertainty channel and contribute to the unemployment-education gap, as these differences allow college workers to have fewer separations while they are searching for their best fit and to find their true calling earlier in their career. Non-college workers take longer to find their best fit. However, as they sample careers and find their best fit, the gap in separation and unemployment rates by education narrows.

We calibrate the model by matching a set of moments from the CPS and NLSY79. The uncertainty channel is pinned down by matching the number of careers worked in by education and the shape of the life cycle separation rate profiles. The calibrated model indicates large differences in uncertainty by education. For example, college (non-college) workers enter the labor market with three (eight) careers that are a potential best fit. We validate the model by showing it matches a set of untargeted moments well, such as an unemployment-education gap that narrows over the life cycle and a stronger correlation between experience and separations among non-college workers.

To decompose the unemployment-education gap, we shut down sources of the gap until all that is left is the uncertainty channel. At that point, 24.26% of the gap remains, which defines the lower bound on the uncertainty channel's contribution. Moreover, the decomposition places an upper bound on the uncertainty channel's contribution at 87.85%. The difference between the lower and upper bound stems from a higher (lower) separation rate when a worker knows that they are not (are) in their best fit.[3] This feature of the model contributes to the unemployment-education gap because non-college workers are less likely to be in their best fit. Thus, they are hit with exogenous separation shocks more frequently than college workers.

The decomposition also reveals that the uncertainty channel is an important source of the model-generated gaps in wages and lifetime earnings between non-college and college workers. We find that eliminating exogenous differences in labor productivity at the best fit closes 46.6% (39.9%) of the gap in wages (lifetime earnings) and that the vast majority of the remaining gap is attributable to the uncertainty channel, as non-college workers are more likely to be in a bad career fit, which reduces their effective labor productivity, wages, employment rate, and earnings.

Finally, we use the model to quantify the effects of several education policies. The one we find to be most effective is implementing an education system that reduces uncertainty among high school graduates through specialized/vocational training akin to that received in associate's degree programs. We quantify this policy by re-calibrating the uncertainty channel parameters to separately match career sampling and separation rates for those with a high school diploma only and associate's degree holders. Our findings show that reducing the number of potential best fits for high school graduates from nine to five (the estimated career set size among associate's degree holders) reduces their unemployment rate by 2.5 percentage points, increases average labor productivity (wages) by 5.7% (7.4%), and raises lifetime earnings by 10.2%.

Our paper is related to the literature on the unemployment-education gap. Cairó and Cajner (2018) and Sengul (2017) document that most of the gap is driven by separation rates and develop models where it is more costly to match with a college worker.[4] In

[3] We interpret the difference in separation rates as the outcome of an interaction between match-specific productivity shocks and workers producing less output outside their best fit.

[4] Several studies have documented higher unemployment rates among college graduates than non-college workers outside of the US (Feng et al., 2024; Coskun, 2024; Girsberger and Meango, 2025) and focus on differences in productivity and search frictions across education groups. Our framework allows

Cairó and Cajner (2018), those are training costs while in Sengul (2017), they are screening costs. The additional costs lead to the formation of higher match-specific productivity and lower separations in matches with college workers. While both papers make important contributions, they do not address the unemployment-education gap over the life cycle or why separations are, especially for non-college workers, decreasing in prior experience.[5] We propose and provide empirical support for the uncertainty channel as an alternative, and perhaps complementary, explanation for the unemployment-education gap.[6] We show that the uncertainty channel can not only explain a decent share of the unemployment-education gap, but it is consistent with the evolution of the gap over the life cycle and the relationship between prior experience, separations, and educational attainment. Apart from these empirical and quantitative implications, the uncertainty channel implies a novel difference in workers by educational attainment. Namely, that these two groups of workers differ in how much knowledge they have about their best fit in the labor market, whereas existing work focuses on mechanisms that are, at their core, driven by exogenous differences in labor productivity.[7] We argue that viewing non-college and college workers through this additional layer of heterogeneity opens up new insights into what drives their labor market outcomes and that it may be worthwhile for education policy to reduce pre-labor market career fit uncertainty.

The uncertainty channel is closely related to the literature studying the life cycle implications of learning about one's comparative advantage in the labor market. Papageorgiou (2014) and Gorry et al. (2019) show that learning about occupational fit can explain several life cycle wage and occupational mobility patterns, but do not emphasize separations, unemployment, or differences in uncertainty by educational attainment.[8] Gervais et al. (2016) develop a model that can generate declining separation, occupational mobility, and unemployment life cycle profiles. However, their paper does not study these patterns by educational attainment. We propose that college graduates face less uncertainty over

for differences in labor productivity but does not emphasize differences in search frictions as this margin directly impacts job finding rates which, as we show in Figure 3, do not vary systematically by education and contribute much to the unemployment-education gap in the US.

[5] In models which generate endogenous separations only through variation in match-specific productivity, the expected duration of a match formed with an unemployed worker is independent of the worker's prior experience as nothing about the worker's prior experience is transferrable across matches.

[6] The uncertainty channel and training as in Cairó and Cajner (2018) may complement each other, as firms may be more willing to train workers whom are more likely to be in their true calling.

[7] While our model allows for such differences in productivity by education, most of the gap in average labor productivity by education in our quantitative analysis is driven by non-college workers being more likely to be in a bad career fit.

[8] Neal (1999) develops a model that can replicate a declining complex transition rate over the life cycle, but does not focus on unemployment. Wee (2013) shows that recessions can disrupt the process of learning about one's ability, thereby generating scarring effects of graduating in a recession.

their best fit, provide empirical support for this hypothesis, and show by incorporating Gervais et al. (2016)'s formalization of uncertainty and learning one's best occupational fit into a life cycle search model with heterogenous education, that differences in uncertainty by education can account for a sizeable portion of the unemployment-education gap. Moreover, our emphasis on education leads to new quantitative insights on the role of pre-labor market uncertainty in generating differences in wages and lifetime earnings by education and the effect of education policy on inequality.

Finally, this paper is related to a growing literature which studies life cycle labor market flows. Menzio et al. (2016) and Cajner et al. (2025) generate separation profiles that decrease over the life cycle in environments where older workers are more likely to have formed a match with high match-specific productivity.[9] Gorry (2016) and Esteban-Pretel and Fujimoto (2014) develop models where experienced workers can reject matches with a low productivity. Both models generate decreasing job finding, separation, and unemployment rate profiles over the life cycle. Our contribution to this literature is to study life cycle separations and unemployment by educational attainment. Further, we emphasize the uncertainty channel, rather than learning about match-specific productivity. Section 4.5 relates our findings to the class of models which focus on the formation of match-specific productivity as a driving force of separation rates over the life cycle.

The rest of the paper is organized as follows. Section 2 presents our empirical analysis. Section 3 develops the model. Section 4 carries out the quantitative analysis. Section 5 concludes. Supplementary material from the online appendix is referenced throughout.

2 Empirical Analysis

This section presents the empirical analysis which supports the uncertainty channel. Section 2.1 shows that college graduates form more accurate forecasts of their future occupation. Section 2.2 presents the unemployment-education gap over the life cycle and shows that differences in separations account for most of the gap. Sections 2.3 and 2.4 discuss additional facts. Section 2.5 summarizes the evidence and transitions to the theory.

To begin, we introduce the data sources used throughout our analysis. First is the monthly Current Population Survey (CPS) files covering 1976-2019, which are downloaded from IPUMS (Flood et al., 2022) and described in Appendix A.1. Second is the Occupation Information Network (O*NET), which measures occupational attributes. Third

[9] Chéron et al. (2013) emphasize the effect of retirement on flows over the life cycle while Créchet et al. (2024) analyze how differences in flows by age and gender can explain differences in unemployment rates across European countries.

is the National Longitudinal Survey of Youth (1979), which tracks the lives of 12,686 individuals born between 1957 and 1964. As the NLSY79 is a panel encompassing respondents' entire careers, it allows us to document several patterns that are not feasible in the CPS.[10] Appendix A.2 details our panel of 4,823 male respondents.[11] Throughout, we map occupations to the occ1990dd classification following Autor and Dorn (2013).

2.1 Expected Occupation

This section measures the accuracy of workers' expectations of their future occupation. To do so, we leverage the NLSY79 where respondents were asked, during their initial interview, what kind of work they would like to be doing when they are 35 years old and in 5 years. Among the 4,823 respondents in our sample, 2,565 (1,620) listed an expected occupation at age 35 (in 5 years) and had a realized occupation at that point in time. Among those, 604 (129) obtained a BA or above.[12] While this is a relatively large drop in sample size, Appendix A.9 shows that individuals that we can and cannot compare their expected and realized occupations have similar observable characteristics. Further, we can increase the sample size from 2,565 to 3,260 by comparing the expected occupation at 35 years old to all occupations worked at between 30 and 40 years old, instead of only at age 35. Table 2 below shows that this does not impact our results.

We find stark differences in the anticipated occupation by education. In general, college respondents expect to be working in occupations with higher and more dispersed skill requirements.[13] For example, the three most common expected occupations at age 35 for college (non-college) respondents are managers, lawyers/judges, and physicians (managers, mechanic/repairer, and truck/delivery driver).

For a first pass at measuring forecast errors, we compute the fraction of individuals with the same expected and realized occupation codes at age 35.[14] Table 1 shows that 7.66% (17.38%) of non-college (college) individuals had the same expected and realized occ1990dd occupation code. As the occ1990dd codes are granular (there are 220 unique realized occ1990dd codes in our sample), we map them to broader first- and second-level occupation categories of which there are 6 and 17, respectively, following Dorn (2009). We see that more individuals are employed in their expected occupation with the broader categories. Further, there is a clear difference by education: college graduates are ap-

[10] We use the CPS for its large sample size and because NLSY79 results could be driven by a cohort effect.

[11] We also restrict to males in the CPS. Our findings are not impacted by this restriction.

[12] As respondents were between 15-22 years old during the initial interview, an individual is labeled as "college" within this section if they eventually obtained a BA or above.

[13] See Appendix A.9.

[14] We select the most frequently observed occupation at age 35.

Table 1: Comparison of Expected and Realized Occupations

Occupation Code Level	Non-College	College
occ1990dd	7.66	17.38
Second-level	15.82	37.48
First-level	29.76	60.48

Notes: The occ1990dd codes are mapped to first- and second-level categories following Dorn (2009). See Appendix Table A16 for a list of the first- and second-level occupation categories.

proximately twice as likely to end up employed in their expected occupation than those without a college degree.

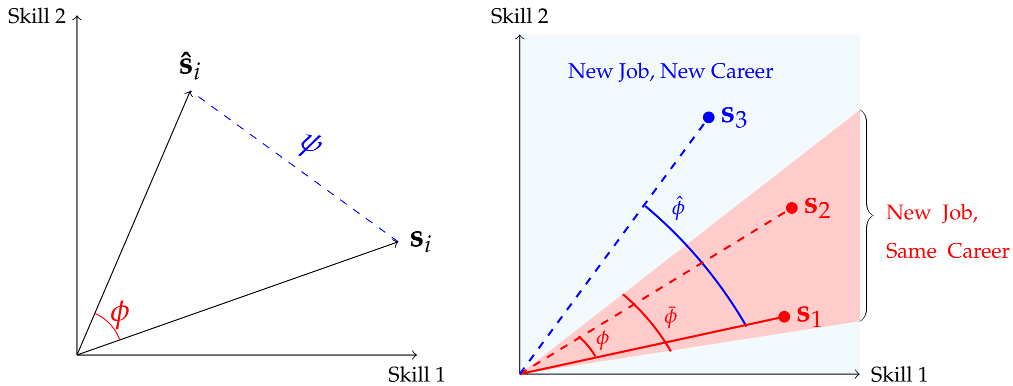
While Table 1 shows in a simple manner that college graduates form more accurate forecasts of their future occupation, it paints an incomplete picture because this comparison not speak to how different the expected and realized occupations are in terms of their skill and task requirements. Moreover, it does not illuminate whether the forecast errors are driven by ending up in an occupation with a different composition or magnitude of skill/task requirements. Therefore, we complement the results presented in Table 1 by computing the distance in skill and task requirements between the realized and expected occupation.[15] To do so, we first measure the verbal, math, and social skill requirements for each occupation as in Guvenen et al. (2020). To capture lower-order skills, we measure an occupation's routine and manual task intensity (Autor and Dorn, 2013). This produces a five-dimensional vector summarizing the skill requirements and task intensity for each occupation.[16] Second, we compute two measures of distance between the vector of requirements for individual i's realized occupation, $\mathbf{s}_i$, and predicted occupation, $\hat{\mathbf{s}}_i$. The first is the angular distance ϕ: $\mathbb{R}^5 \times \mathbb{R}^5 \to [0, \pi/2]$, and is given by:

$$\phi(\mathbf{s}_i, \hat{\mathbf{s}}_i) = \cos^{-1}\left(\frac{\mathbf{s}_i \cdot \hat{\mathbf{s}}_i'}{||\mathbf{s}_i|| \, ||\hat{\mathbf{s}}_i||}\right). \tag{1}$$

Figure 1(a) illustrates the angular distance in the case of two skills. Notably, the angular distance captures the difference in the composition of skill requirements. The second measure is the Euclidean distance, $\psi(\mathbf{s}_i, \hat{\mathbf{s}}_i) = \sqrt{\sum_k (s_{i,k} - \hat{s}_{i,k})^2}$, where $s_{i,k}$ ($\hat{s}_{i,k}$) denotes worker i's realized (expected) occupation's requirement in attribute k. The Euclidean distance accounts for differences in both the composition and magnitude of skill requirements. For an example of the distance measures, Appendix Figure A2 compares dentists

[15] We compute the average requirements across the jobs worked while 35 years old, between 30-40 years old (inclusive), and 5 years from their initial interview.

[16] See Appendix A.1.2 for more details on the measurement of skill and task requirements.



(a) Distance Measurements

(b) Career Transitions

Figure 1: Distance Measurements and Career Transitions. *Notes*: Panel (a) depicts the angular, ϕ, and Euclidean, ψ, distance. Panel (b) demonstrates job changes within and across careers.

to eight occupations and shows how the different skill mix in the occupations relative to dentists maps to an angular and Euclidean distance.

Table 2 reinforces that college workers form more accurate forecasts. From Panel A, the average Euclidean (angular) distance for their occupation at 35 years old is 26% (32%) smaller for college graduates. Panels B and C show similar differences if we (i) compare all realized occupations between ages 30 and 40 to the expected occupation at 35 years old and (ii) the forecast error in 5 years. The third row within each panel shows that 65-77% of the Euclidean distance is attributable to the composition of skill requirements.[17] This suggests that workers have more uncertainty about which composition of skill requirements they are best suited for. This is why we focus on *career* sampling in our model, where a career is a group of occupations with a similar composition of skill requirements. We precisely define a career and measure career mobility in Section 2.3.1.

We conclude this section by assessing the impact of completing more years of schooling on the forecast errors. Before doing so, there are a few caveats to make explicit. First, as occupation expectations were only measured in the initial wave of the NLSY79, we cannot observe any changes in one's expected occupation as they finish more schooling. Second, Appendix Figure A19 shows that the gap in age 35 forecast errors is present at each age at which respondent's expectations were measured. Given that there is a sizeable gap in forecast errors even among 15-18 year olds who had not enrolled in college yet, we do not claim that attending college has a causal effect on an individual's knowledge of their best fit in the labor market. What we present below is suggestive evidence

[17]From the Law of cosines, the fraction of the Euclidean distance that is attributable to differences in the angle, ϕ, is $2\|\mathbf{s}_i\|\|\hat{\mathbf{s}}_i\|(1-\cos(\phi))/\psi^2$. See Appendix A.9.1.

Table 2: Angular and Euclidean Distances by Education

	Non-College	College
Panel A: Occupation at Age 35		
Angular Distance	29.83	20.31
Euclidean Distance	0.77	0.57
% of Euclidean Driven by Angle	73.90	77.04
Observations	1,961	604
Panel B: Occupations Between Ages 30 and 40		
Angular Distance	27.48	19.30
Euclidean Distance	0.73	0.55
% of Euclidean Driven by Angle	69.79	73.33
Observations	2,450	810
Panel C: Expected Occupation in 5 Years		
Angular Distance	25.84	20.09
Euclidean Distance	0.66	0.56
% of Euclidean Driven by Angle	64.85	70.16
Observations	1,491	129

Notes: Angular distance is measured in degrees. Panel A compares an individual's expected occupation at age 35 with their realized occupations at age 35. Panel B compares the expected occupation at age 35 to the (weighted) average of skill requirements across all realized occupations between ages 30 and 40 (inclusive). Panel C compares an individual's expected occupation in 5 years with the realized occupation 5 years after their initial interview. A paired sampled t-test indicates that the forecast error of non-college workers is statistically larger than that of college workers, with the null hypothesis ($H_0 : diff < 0$) being rejected at the 1% significance level. The third row within each panel is the proportion of the Euclidean distance attributable to the angular distance. Data are from the NLSY79.

that individuals who complete more schooling form more accurate expectations about their future occupation and leave a full analysis of the causal impact of attending college on occupational forecast errors to future work.

With those caveats in mind, we estimate the following regression:

$$\text{FCE}_i = \beta_0 + \beta_1 \text{Years}_i + \Gamma X_i + \epsilon_i, \tag{2}$$

where FCE is individual i's forecast error based on their expected and realized occupation at 35 years old, Years is the years of schooling completed, and X contains average ability following Guvenen et al. (2020), race, whether individual i was ever married, had a child, and average family income.

Table 3 presents the estimates of β_1 across the two measures forecast errors (Euclidean and angular) and years of schooling (overall and college). We can see that more years of completed schooling are associated with lower forecast errors. The findings suggest

Table 3: Years of Schooling and Forecast Errors

	(1) Euclidean	(2) Angular	(3) Euclidean	(4) Angular
Education years	-0.01306***	-0.70416***		
College years			-0.03486***	-1.42526***
Observations	2560	2560	1136	1136
R^2	0.056	0.079	0.111	0.119

Notes: The dependent variable in columns 1 and 3 (2 and 4) is the Euclidean (angular) distance between an individual's expected and realized occupation at 35 years old. Education years is the number of years of schooling completed. College years is the number of years of college completed. Columns (3) and (4) include respondents who have completed at least one year of college. All specifications include the full vector of individual controls, X, listed below equation (2). Levels of statistical significance are denoted by ***$(p < 0.01)$. Data are from the NLSY79.

that education, and particularly college, may provide individuals a more accurate understanding of their best occupational fit in the labor market. We revisit this in Section 4.4, where we use the model to quantitatively evaluate the effect of increased educational attainment on unemployment, wages, and lifetime earnings.

2.2 Unemployment-Education Gap

This section presents several facts related to the unemployment-education gap that we argue are consistent with the uncertainty channel.

Figure 2 shows the unemployment rate by age and education, using CPS data. The solid lines show that the unemployment rate for college graduates is lower than those without a college degree and that the unemployment-education gap narrows over the life cycle.[18] Next, Figure 3 presents the job finding and separation probabilities by age and educational attainment.[19] There are several takeaways. First, separations decline with age for each education group. Second, college workers consistently exhibit a lower separation probability. Third, the gap in separation probabilities also narrows over the life cycle.[20]

[18] Appendix A.10 shows that individuals with an associate's degree and college dropouts fall in-between those with no college experience and graduates with a BA or above in our main outcomes of interest.

[19] We correct for time aggregation bias as in Shimer (2012). We also compute the job finding and separation rates as in Shimer (2005) and Elsby et al. (2009). This gives the same conclusions presented in this section. See Appendix A.4.1.

[20] Appendix A.4.1 shows this pattern emerges in both voluntary and involuntary separations. Later, in Section 4.2, we discuss how our model does capture some distinguishing features of voluntary and involuntary separations observed in the data. Moreover, recent work has documented that a small proportion

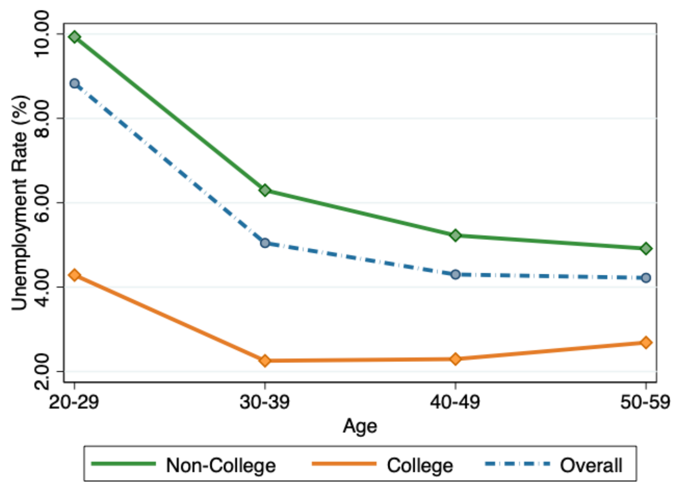


Figure 2: Unemployment-Education Gap over the Life Cycle. *Note*: Unemployment life cycle profiles computed using CPS data between 1976 and 2019.

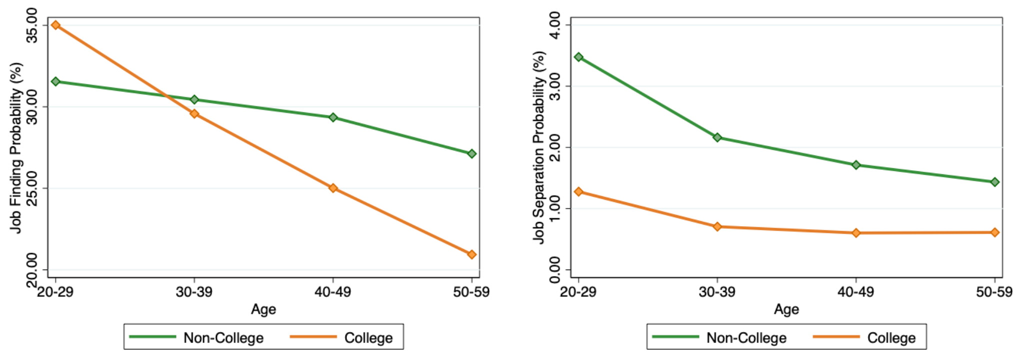
Figure 3 suggests that the unemployment-education gap is primarily driven by differences in separations, as the job finding probability is lower among college graduates throughout most of the life cycle. Applying a decomposition as in Pissarides (2009) shows that differences in separations explains at least 70% of the unemployment-education gap at each age bin.[21] It is for this reason we propose a mechanism that is tightly linked to the separation margin. Intuitively, if non-college (college) workers enter a match with more (less) uncertainty whether they are well-suited for that career, they are more (less) likely to learn it is a bad match and separate from it. To support the connection between the uncertainty channel and the separation margin, we find that the average separation probability among non-college (college) workers in the NLSY79 who were employed in their anticipated occupation at 35 years old is 34% (32%) lower than those who do not.[22]

It is important to reemphasize that the differences in unemployment and separations by educational attainment are widest early in workers' careers. Our hypothesis is consistent with this for the simple reason that college workers, having entered the labor market with less uncertainty, begin their careers with lower separations and hence, a lower un-

of workers frequently transition between employment and unemployment and can account for a disproportionate amount of aggregate unemployment (Hall and Kudlyak, 2022; Gregory et al., 2025). Appendix A.7 demonstrates that the narrowing unemployment-education gap over the life cycle is not driven by "unemployable" non-college workers who exhibit an abnormally large number of separations.

[21] Appendix A.4.2 provides a description of the decomposition, as well as results with alternative transition probabilities and rates.

[22] We also find that, within each education group, the difference in separation rates by forecast error is widest early in workers' careers. See Appendix A.3.1. This supports the notion that workers with more uncertainty over their best fit exhibit higher separations, especially early in their career. Over time, they find their true calling and experience fewer separations.



(a) Job Finding Probability (b) Separation Probability

Figure 3: Job Finding and Separation Probabilities over the Life Cycle. *Note*: All series are computed using CPS between 1976 and 2019 and are corrected for time aggregation bias.

employment rate. Non-college workers enter with more uncertainty and experience more separations. As their career advances, they learn about their best fit, separate from their jobs less frequently, and the unemployment-education gap narrows. This will be formalized in Section 3 and quantified in Section 4.

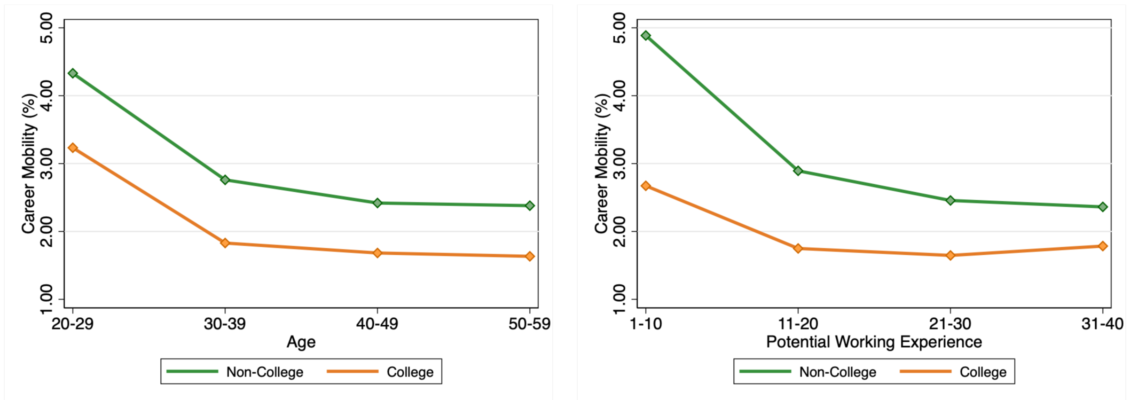
2.3 Supporting Evidence

Section 2.1 presented our most direct evidence for the uncertainty channel and Section 2.2 related the uncertainty channel to the unemployment-education gap. This section presents additional, indirect, evidence for our hypothesis.

2.3.1 Career Mobility

We begin by comparing career mobility rates by age and education. The motivation for doing so is the following: if non-college workers enter the labor market with more uncertainty about their best fit, then they should switch careers at a higher rate, particularly early in their career, as they sample careers and gradually transition to their best fit.

Following Baley et al. (2022), we define a career transition as an occupation switch where the angular distance between the current and previous job exceeds a threshold, $\bar{\phi}$. The threshold is chosen so that the average correlation in occupational requirements is zero in career switches. We find $\bar{\phi} = 21.3$. Intuitively, a career switch occurs when the worker transitions between occupations with very different compositions of skill requirements. This is depicted in Figure 1(b). If the worker switches from occupation 1 to 2, the angle between the skill requirements $\mathbf{s}_1$ and $\mathbf{s}_2$ is $\phi < \bar{\phi}$. So, the worker is moving to a new job within the same career. If the worker transitions between occupation 1 and 3,



(a) Age

(b) Potential Experience

Figure 4: Career Mobility. *Notes*: Career mobility rates are computed using CPS data between 1994-2019 and after applying the Moscarini and Thomsson (2007) correction for measurement error in occupational mobility in the CPS.

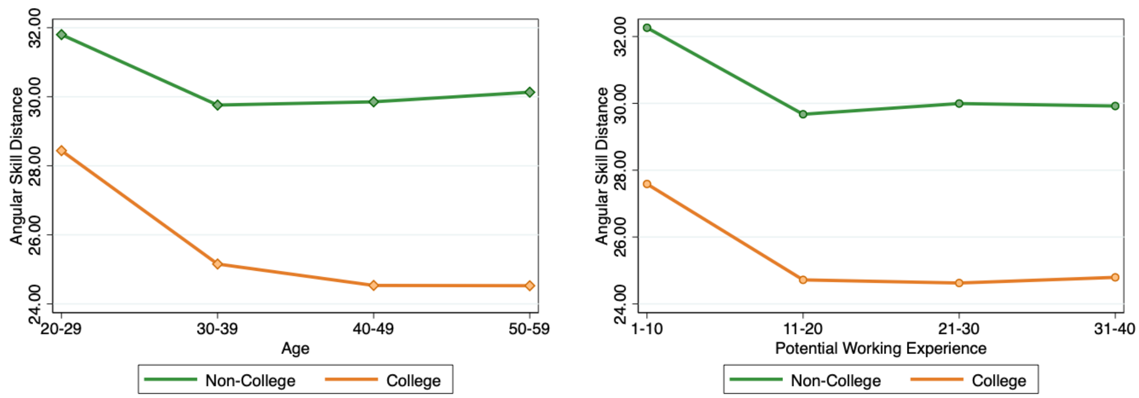
the distance is $\hat{\phi} > \bar{\phi}$. In this case, the composition of skill requirements are sufficiently different, leading to a career switch. Appendix Figure A2 shows eight occupations and whether each is in the same career as dentists based on the angular distances.

Figure 4 presents the career mobility rates. There are two patterns to highlight. First, career mobility is decreasing in age. Second, non-college workers change careers at a higher rate.[23] As education impacts the timing of entry to the labor market, we also show the career mobility rates by years of potential experience, where we assume non-college (college) workers enter the labor market at the age of 18 (22). Figure 4(b) illustrates that while the overall patterns are unchanged, the gap in career mobility rates in the early stages of workers' careers is even larger than when we compare by age, and narrows over the course of workers' careers. Following the intuition at the beginning of this section, these patterns are consistent with the uncertainty channel.

2.3.2 Occupational Distance

Next, we examine another implication of our hypothesis: college graduates should transition between similar occupations whereas those without a college degree make more significant changes when switching occupations. The intuition is, given their lower uncertainty, if college workers learn that their current job is not their best fit, it is still more likely they are in a decent match and that a better match will have fairly similar characteristics to their current job. To evaluate this in the data, we use the CPS to compare skill

[23]We find similar patterns when considering "complex" switches, or a concurrent change in employer, occupation, and industry (Neal, 1999). Results are available upon request.



(a) Age

(b) Potential Experience

Figure 5: Angular Distance in Occupation Transitions. *Note*: The series are computed using CPS data between 1994-2019 and after applying the Moscarini and Thomsson (2007) correction for measurement error in occupational mobility in the CPS.

and task requirements in occupation switches.

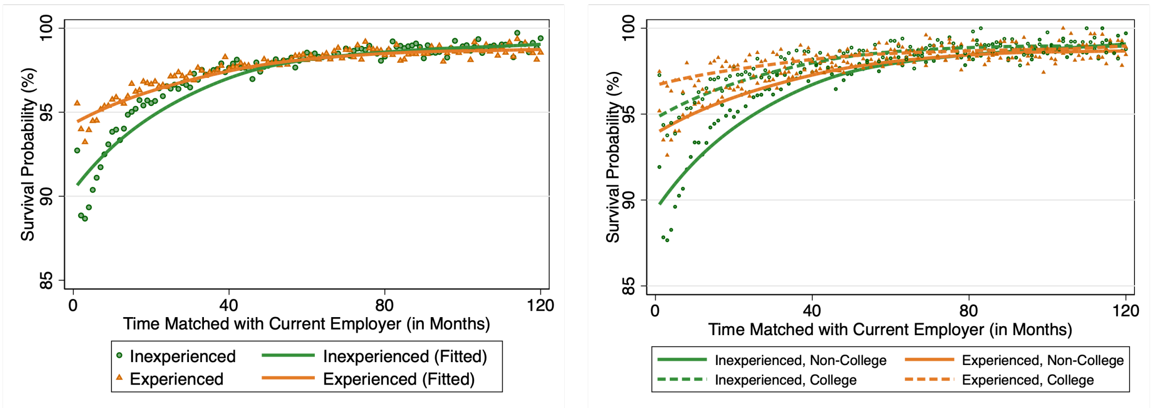
Figure 5 shows that the average angular distance in occupational switches is lower for college graduates across the life cycle. Therefore, not only do college graduates switch careers at a lower rate, but when they do switch occupations, they tend to transition into occupations with a (relatively) similar composition of skill requirements.

2.3.3 Experience and Match Duration

An important feature of our hypothesis is that workers learn their best fit by working, and that they can transfer what they have learned about their best fit between matches. An immediate corollary to this is that the expected duration of a match between a worker and firm is increasing in the worker's prior experience at the time the match is formed. With this in mind, we use the NLSY79 to explore whether the empirical relationship between prior experience and the survival probability of a match is consistent with this intuition.

As a first step, we place workers into two groups based on their level of accumulated experience at the beginning of a match. The first group, experienced workers, consists of those who enter the new match with more than 76 months of work experience, where 76 months is the median months of experience at the formation of new matches in our sample. The second group, inexperienced workers, are those who begin a match with no more than 76 months of experience. The survival probability is simply the fraction of matches that survive between months t and $t+1$.

Figure 6 presents the match survival probability as a function of match tenure and prior experience. As seen in Figure 6(a), experienced workers exhibit a higher survival



(a) All Workers

(b) By Education

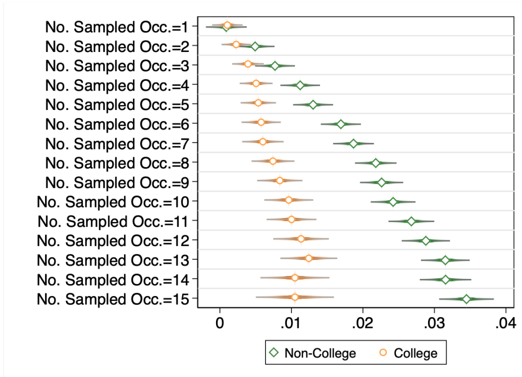
Figure 6: Prior Experience and Match Survival. *Notes*: Panel (a) shows the survival probability of a match over match tenure for experienced and inexperienced workers. Panel (b) further disaggregates by the worker's educational attainment. Data are from the NLSY79.

probability for the first 2–3 years of the match. Figure 6(b) shows that the association between prior experience and the survival probability is stronger among less-educated workers. This can be seen by noting the larger gap in the survival probability between inexperienced and experienced workers for workers with less than a college degree than those with a college degree. This is consistent with the uncertainty channel as non-college workers rely more on experience to find their best fit. Appendix A.6 shows that these findings are robust to excluding matches formed through a job-to-job transition, allowing prior experience to be measured in months rather than two categories, and controlling for observable characteristics such as age.

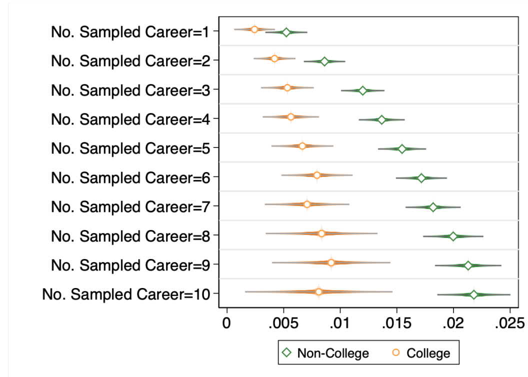
An additional aspect of our hypothesis is that workers learn not just from work experience, but particularly from sampling occupations and careers. Therefore, we estimate the relationship between the survival probability and the number of occupations or careers the worker had formerly worked in when the match was formed. To do so, we separately estimate the following specification on non-college and college workers in the NLSY79:

$$y_{it} = \beta_0 + \sum_{j=1}^{J} \beta_j \mathbb{I}\{\text{NumSam} = j\}_{it} + \gamma \text{Tenure}_{it} + \delta \text{Exp}_{it} + \Phi_i + \epsilon_{it}, \tag{3}$$

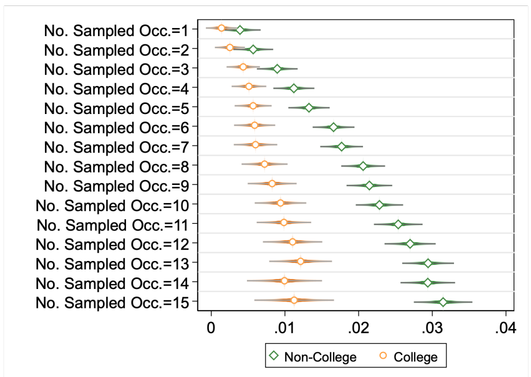
where y_{it} is equal to one (zero) if individual i is employed in month t and employed in the same employer/occupation/career match in month $t+1$ (unemployed in month $t+1$), $\mathbb{I}\{\text{NumSam} = j\}$ is an indicator for the number of occupations or careers individual i had worked in at the time their current match employer/occupation/career match was formed, Tenure_{it} is the employer/occupation/career match tenure, Exp_{it} is total work



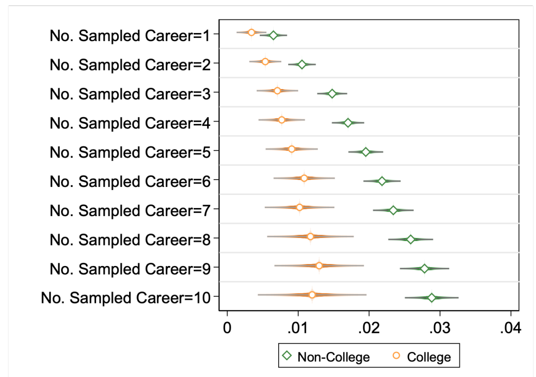
(a) Employer Match: Sampled Occupations



(b) Employer Match: Sampled Careers



(c) Occupation Match: Sampled Occupations



(d) Career Match: Sampled Careers

Figure 7: Job Sampling and Match Survival. *Notes*: Panel (a) ((b)) displays β_j coefficients for sampled occupations (careers) from estimating equation (3) for employer matches. Panel (c) ((d)) displays the coefficients for sampled occupations (careers) when the dependent variable, y_{it}, captures whether a worker remains employed in the same occupation (career), and Tenure$_{it}$ is the worker's accumulated tenure in their current occupation (career). All figures display 99% confidence intervals. Regressions use the NLSY79 sample and full regression output is provided in Appendix A.8.

experience, and Φ_i is an individual fixed effect. The coefficients, β_j for $j = 1, 2, \ldots, J$ with $J = 15$ ($J = 10$), capture the association between the j^{th} occupation (career) sampled and the survival probability, relative to a worker who is forming their first match.

Figures 7(a)-7(b) display the β_j coefficients for employer matches. The results show that, especially for non-college workers, the survival probability increases with the number of sampled occupations and careers. Figures 7(c) and 7(d) show that the same patterns emerge when we consider the survival probability of a worker's occupation and career match. These findings support our hypothesis that non-college workers learn more about their best fit by sampling occupations and careers than college workers do.

2.4 Additional Evidence and Robustness

This section lists additional evidence that complements the analysis in Sections 2.1-2.3. First, college graduates exhibit lower rates of occupational mobility (Appendix A.3.2), consistent with the patterns observed in career mobility. Second, college graduates experience lower skill mismatch throughout the life cycle (Appendix A.3.3), as their lower uncertainty leads to better-suited matches. Third, college graduates work in occupations with more dispersed skill requirements (Appendix A.3.4), which is consistent with the notion that workers with less uncertainty about their skills may be more willing to work in jobs with relatively high requirements in a subset of skills. Fourth, college graduates experience fewer employer, occupation, and career switches (Appendix A.3.5). Fifth, the unemployment-education gap is evident across distinct undergraduate majors, which shows that the gap is not overwhelmingly driven by majors associated with providing specific skills and thereby "locking" graduates into a particular field (Appendix A.3.6).

Further, the CPS patterns can be replicated in the NLSY79. See Appendices A.5.1-A.5.4. Finally, the correlation between educational attainment and our outcomes of interest are robust to controlling for standard observable characteristics. See Appendix A.4.3 for the CPS and A.5.5 for NLSY79 analyses, respectively.

2.5 Summary and Transition

This section has (i) presented a combination of new and previously documented facts and (ii) outlined how each supports the uncertainty channel.[24] As mentioned in Section 1, a leading alternative theory of the unemployment-education gap is that college workers have higher match-specific productivity. These models are consistent with separation rates that (i) are lower among college graduates and (ii) decrease over the life cycle as shown in Figure 3(b). However, there is no particular role for learning and occupations/career sampling in the match-specific productivity theory. As such, these models do not speak to the other patterns shown in this section that are related to occupations and careers. We revisit and elaborate more on the relation between our findings and the match-specific productivity theory in Section 4.5.

Our remaining primary objective is to quantify the uncertainty channel's contribution to the unemployment-education gap. To do so, we develop a search model where workers are heterogeneous in their education and best career fit. Following the evidence on

[24] To the best of our knowledge, the new facts reported in Section 2 are occupational forecast errors (Tables 1-3), life cycle patterns in unemployment and flows by education (Figures 2 and 3), distance in occupational switches (Figure 5), and the relationship between separations and prior experience by education (Figures 6(b) and 7).

forecast errors in Section 2.1, workers do not know their best fit. Further, and based on the evidence presented in Section 2.3.3, workers sample careers to learn their suitability in each. A match may be destroyed upon learning the worker is not in their true calling. Underlying these ingredients are exogenous differences by education in labor productivity, the number of careers that are a potential best fit, and the speed at which workers learn their fit in a career. Section 3.3 outlines how each difference contributes to the unemployment-education gap and Section 4 quantifies their contributions. Finally, we embed these ingredients within a competitive search model with bilaterally efficient contracts (Menzio and Shi, 2011), which allows for a rich amount of heterogeneity among workers in a tractable environment.

3 Model

This section develops a life cycle directed search model. Section 3.1 introduces the environment. Section 3.2 characterizes the equilibrium and Section 3.3 details the sources of the unemployment-education gap.

3.1 Environment

Time is discrete and indexed by $t = 0, 1, \ldots, \infty$. At $t = 0$, there is a unit measure of workers and a large measure of firms. All agents are risk neutral and discount the future according to the discount factor $\beta \in (0, 1)$.

Workers are heterogeneous in four dimensions. The first is age, $a \in \{y, o\}$, for young and old, respectively. Second is educational attainment, $e \in \{0, 1\}$ where $e = 0$ ($e = 1$) is non-college (college). Education is fixed and observable. Third, each worker is best suited for one career, c^*, which we refer to as their best fit (Gervais et al., 2016). For workers with education e, $c^* \in \mathbb{C}_e$ where $\mathbb{C}_e \subset \mathbb{Z}_+$ and $2 < N_1 \equiv |\mathbb{C}_1| < |\mathbb{C}_0| \equiv N_0$. In words, there are fewer careers that are potentially a best fit for college workers. Fourth is a worker's history, i, which denotes one plus the number of careers that the worker has learned is not their true calling. Initially, a worker's true calling is unknown to both the worker and firms. Once the worker learns about their fit in a career, it becomes public information.

The labor market is organized in a continuum of submarkets indexed by $\omega = (a, e, i, s, x)$. In submarket ω, firms search for workers with age a, education e, history i, the worker's status in career i: $s \in \{un, b, g\}$ (unsure, bad, or good fit), and offer workers contracts with lifetime discounted utility x.

Each period is divided into five stages: learning, separation, search, production, and demographics. We proceed to fill in the details of each stage.

In stage 1, employed workers with characteristics (i, e) who are unsure about their current career learn about their fit with probability $\phi_e \in [0, 1]$.[25] Workers who learn that their current career is their true calling become type $i = N_e$. Those who learn that their current career is not their true calling become type $\max\{i + 1, N_e\}$ workers and update their beliefs over the careers they have not sampled according to Bayes rule. A type i worker who has learned that $i - 1$ careers are not their best fit believes that the i^{th} career is their best fit with probability p_{ie}, where

$$p_{ie} = \frac{1}{N_e - (i-1)}. \tag{4}$$

In stage 2, a match with a type (i, e) worker and status s is destroyed with probability $\delta \in [\delta^s, 1]$ where $\delta^{un}_{ie} = p_{ie}\delta^g + (1 - p_{ie})\delta^b$ and $\delta^g < \delta^b$.[26] The destruction probability is specified by the employment contract, and the lower bound represents exogenous separations. A worker who loses their job must wait one period before looking for another.

Next, in stage 3, firms choose which submarket, if any, to post a vacancy in. The vacancy posting cost in submarkets with age a workers is κ_a.[27] Workers choose which submarket to search in. Old workers who look for a new career incur a switching cost ζ. The decision to leave a career is irreversible. There is no search on the job.[28]

[25]We interpret the learning probability as a reduced form representation of a signal extraction problem where, with some probability, the observed match output is perfectly informative of the worker's career fit and with a complementary probability is completely uninformative.

[26]The separation probabilities, δ^b and δ^g, could be indexed by education, e, to capture differences in separation risk in occupations typically worked in by college and non-college workers. In our quantitative analysis in Section 4, δ^b (δ^g) is primarily pinned down by matching the separation rate of young (older) non-college (college) workers. Therefore, our quantitative decomposition of the unemployment-education gap does account for such differences in underlying separation risk across occupations by education, as we study how much lowering δ^b to δ^g closes the unemployment education gap. In other words, this portion of the decomposition could be interpreted as asking how much of the unemployment-education gap closes if we were to assign the same underlying separation risk in occupations typically worked in by college graduates to non-college workers.

[27]Our main intention is not to explain the life cycle job finding profiles shown in Figure 3(a), as these patterns do not contribute much to the unemployment-education gap (see Appendix A.4.2). However, we include age-specific posting costs to ensure the model is broadly in line with these patterns.

[28]In our NLSY79 sample, the probability a worker changes careers through unemployment, i.e. experiences a career switch through an "EUE" transition, is 37%. The probability of switching careers during consecutive months of employment is 1.44%. As workers are much more likely to switch careers through unemployment than employment, we abstract from including on the job search in the model due to the added complexity incorporating it would bring. Moreover, including on the job search is unlikely to alter our quantitative findings, as the model with on the job search would be parameterized to match the frequency of career switches through unemployment and thus, would still give a tight connection between the uncertainty channel and the separation margin.

Let $v(\omega)$ and $u(\omega)$ denote the measure of vacancies and unemployed workers, respectively, searching in submarket ω. The number of matches is given by the CRS matching function $F(u(\omega), v(\omega))$. Define $\theta = v/u$ as tightness in submarket ω. A worker finds a job with probability $f(\theta(\omega)) = F/u(\omega)$ where $f\colon \mathbb{R}_+ \to [0,1]$ is twice continuously differentiable, strictly increasing, and strictly concave. Firms fill their vacancy with probability $q(\theta(\omega)) = F/v(\omega)$ where $q\colon \mathbb{R}_+ \to [0,1]$ is twice continuously differentiable, strictly decreasing, and strictly convex.

In the production stage, stage 4, unemployed workers produce z units of output. Employed workers in their true calling produce y_e, whereas those in a bad fit produce $y_e - \alpha$ units of output where $y_1 > y_0$ and $y_0 - \alpha > z$. The output in unsure matches is $y_{ie}^{un} = p_{ie} y_e + (1 - p_{ie})(y_e - \alpha)$.

At the beginning of stage 5, a fraction λ_o of young agents become old and a fraction λ_d of old agents die. To maintain a constant population, a measure $\mu = \frac{\lambda_o \lambda_d}{\lambda_o + \lambda_d}$ of workers enter the economy as young and unemployed. A fraction π_0 $(1 - \pi_0)$ enter the economy with education $e = 0$ $(e = 1)$. New entrants have their true calling assigned by nature, where the probability any single career is their true calling is $1/N_e$.

Finally, the contract space is complete, giving rise to bilaterally efficient employment contracts. Therefore, employment contracts offered by firms will maximize the joint value of the match (Menzio and Shi, 2011).

3.2 Equilibrium

We begin with the value functions for unemployed workers and the value of a match, which are measured from the beginning of the production stage.

Let $U_{a,e,i}$ denote the value of a worker with age a, education e, and history i searching for their i^{th} career in which they have an unknown fit if $i < N_e$ or a good fit if $i = N_e$. Consider workers who are old at the beginning of the production stage. The worker produces z units of output and survives between periods with probability $1 - \lambda_d$. In the subsequent search stage, they search in submarket ω and find a job with probability $f(\theta(\omega))$. If they find a job, they earn the continuation value of the employment contract, x. If they don't find a job, they earn the value of unemployment, $U_{o,e,i}$. It follows that $U_{o,e,i}$ satisfies

$$U_{o,e,i} = z + \beta(1 - \lambda_d)\{U_{o,e,i} + R(x, U_{o,e,i})\}, \tag{5}$$

$$R(x, U) = \max_{(\theta, x)} f(\theta)(x - U). \tag{6}$$

Now let $\overline{U}_{a,e,i}$ denote the value of an unemployed worker with characteristics (a,e) who knows that career i is a bad fit. The worker decides in the subsequent search stage whether to look for a new career or not. If they leave their current career, they incur the switching cost ζ and search in a submarket for type $i+1$ workers, as they know that their previous i careers are not their best fit. It follows that, for old workers, we have:

$$\overline{U}_{o,e,i} = z + \beta(1-\lambda_d)\{l^*_{o,e,i}[U_{o,e,i+1} - \zeta + R(x, U_{o,e,i+1})] + (1 - l^*_{o,e,i})[\overline{U}_{o,e,i} + R(x, \overline{U}_{o,e,i})]\}, \tag{7}$$

where $l^*_{o,e,i}$ denotes the worker's choice to leave their career and is given by

$$l^*_{o,e,i} = \begin{cases} 1 & \text{if } U_{o,e,i+1} - \zeta + R(x, U_{o,e,i+1}) \geq \overline{U}_{o,e,i} + R(x, \overline{U}_{o,e,i}), \\ 0 & \text{if } U_{o,e,i+1} - \zeta + R(x, U_{o,e,i+1}) < \overline{U}_{o,e,i} + R(x, \overline{U}_{o,e,i}). \end{cases} \tag{8}$$

Consider a young unemployed worker. The difference relative to old workers is the worker becomes old between periods with probability λ_o. Hence, $U_{y,e,i}$ and $\overline{U}_{y,e,i}$ satisfy:

$$U_{y,e,i} = z + \beta\{\lambda_o[U_{o,e,i} + R(x, U_{o,e,i})] + (1-\lambda_o)[U_{y,e,i} + R(x, U_{y,e,i})]\}, \tag{9}$$

$$\begin{aligned} \overline{U}_{y,e,i} = z + \beta\{&\lambda_o\left[l^*_{o,e,i}[U_{o,e,i+1} - \zeta + R(x, U_{o,e,i+1})] + (1 - l^*_{o,e,i})[\overline{U}_{o,e,i} + R(x, \overline{U}_{o,e,i})]\right] \\ &+ (1-\lambda_o)\left[l^*_{y,e,i}[U_{y,e,i+1} + R(x, U_{y,e,i+1})] + (1 - l^*_{y,e,i})[\overline{U}_{y,e,i} + R(x, \overline{U}_{y,e,i})]\right]\}. \end{aligned} \tag{10}$$

We now proceed to value of a match, or the sum of the worker's utility and firm's profits, which is sufficient to characterize the entry of firms and separations as the contracts offered by firms maximize the joint surplus of the match. Starting with an old worker who is employed in a bad match, the match output is $y_e - \alpha$. In the subsequent separation stage, the job is destroyed with probability δ^b, in which case the worker receives the value of unemployment and the firm receives the value of a vacancy (zero).[29] If the match is not destroyed, the continuation value is given by the value of the match. It follows that the value of a bad match with an old worker with education e and history i, $\overline{V}_{o,e,i}$, satisfies

$$\overline{V}_{o,e,i} = y_e - \alpha + \beta(1-\lambda_d)\{\delta^b\overline{U}_{o,e,i} + (1-\delta^b)\overline{V}_{o,e,i}\}. \tag{11}$$

As for old workers with education e and history i who are employed in an unknown or a good fit, the match produces $p_{ie}y_e + (1-p_{ie})(y_e - \alpha)$ units of output. The worker learns about their suitability for their career in the learning stage with probability ϕ_e.

[29] Bad matches with old workers that were not destroyed in the previous separation stage will not be destroyed endogenously as nothing about a bad match changes between periods.

Conditional on learning about their fit, they are in their true calling with probability p_{ie} and the worker's type becomes $i = N_e$. With probability $1 - p_{ie}$, the worker learns they are in a bad fit. In this case, the worker and firm enter the separation stage and choose whether to destroy the match or not. The value of the match, $V_{o,e,i}$, satisfies

$$V_{o,e,i} = p_{ie}y_e + (1-p_{ie})(y_e - \alpha) + \beta(1-\lambda_d)\{\phi_e[p_{ie}(\delta^g U_{o,e,N_e} + (1-\delta^g)V_{o,e,N_e}) + \\ (1-p_{ie})(d^*_{o,e,i}\overline{U}_{o,e,i} + (1-d^*_{o,e,i})\overline{V}_{o,e,i})] + (1-\phi_e)[\delta^{un}_{ie}U_{o,e,i} + (1-\delta^{un}_{ie})V_{o,e,i}]\}, \quad (12)$$

where $d^*_{o,e,i}$, the separation probability after learning the match is a bad fit, is given by

$$d^*_{o,e,i} = \begin{cases} \delta^b & \text{if } \overline{U}_{o,e,i} < \overline{V}_{o,e,i}, \\ 1 & \text{if } \overline{U}_{o,e,i} \geq \overline{V}_{o,e,i}. \end{cases} \quad (13)$$

For young workers, the value of a bad match satisfies:

$$\overline{V}_{y,e,i} = y_e - \alpha + \beta\{(1-\lambda_o)[\delta^b\overline{U}_{y,e,i} + (1-\delta^b)\overline{V}_{y,e,i}] + \lambda_o[\delta^b\overline{U}_{o,e,i} + (1-\delta^b)\overline{V}_{o,e,i}]\}. \quad (14)$$

Finally, we have the value of a young worker in a match with an unsure or good fit, which follows a similar intuition as with old workers:

$$V_{y,e,i} = p_{ie}y_e + (1-p_{ie})(y_e - \alpha) + \beta\sum_a \chi_a\{\phi_e[p_{ie}(\delta^g U_{a,e,N_e} + (1-\delta^g)V_{a,e,N_e}) + \\ (1-p_{ie})(d^*_{a,e,i}\overline{U}_{a,e,i} + (1-d^*_{a,e,i})\overline{V}_{a,e,i})] + (1-\phi_e)[\delta^{un}_{ie}U_{a,e,i} + (1-\delta^{un}_{ie})V_{a,e,i}]\}, \quad (15)$$

where $\chi_a = 1 - \lambda_o$ if $a = y$ and $\chi_a = \lambda_o$ if $a = o$.

The firm's cost to post a vacancy in a submarket with age a workers is κ_a. The expected benefit to posting a vacancy in submarket $\omega = (a, e, i, s, x)$ is $q(\theta(\omega))[V_{a,e,i} - x]$ if $s \in \{un, g\}$ and $q(\theta(\omega))[\overline{V}_{a,e,i} - x]$ if $s = b$. In submarkets visited by a positive amount of workers, tightness is consistent with firms' incentives to create vacancies if

$$\kappa_a \geq \begin{cases} q(\theta)[V_{a,e,i} - x] & \text{for } i \in \{1, 2, \ldots, N_e\} \text{ and } s \in \{un, g\}, \\ q(\theta)[\overline{V}_{a,e,i} - x] & \text{for } i \in \{1, 2, \ldots, N_e - 1\} \text{ and } s = b, \end{cases} \quad (16)$$

and $\theta \geq 0$ with complementary slackness.

Definition 1. A stationary recursive equilibrium consists of a tightness function $\theta(\omega)$, value and policy functions for unemployed workers, $U_{a,e,i}$, $\overline{U}_{a,e,i}$, and $\omega^*_{a,e,i}$, $l^*_{a,e,i}$, joint value and policy functions, $V_{a,e,i}$, $\overline{V}_{a,e,i}$, $d^*_{a,e,i}$, and a distribution of workers that satisfies the following conditions. First, $\theta(\omega)$ satisfies (16) for all ω. Second, the value and policy functions of unemployed workers satisfy equations (5)-(10). Third, the joint value and associated policy functions for a match satisfy equations (11)-(15). Finally, the distribution of workers satisfies the laws of motion specified in Appendix B.1.

As established by Menzio and Shi (2011) for directed search models with free entry and bilateral efficiency, a recursive equilibrium exists and is block-recursive (BRE). As workers self-select into submarkets based on their observable characteristics, firms know they will only meet one type of worker in each submarket. Hence, tightness in each submarket is independent of the distribution of workers across age, educational attainment, history i, and the worker's status in their current career.

3.3 The Unemployment-Education Gap

This section details two broad sources of the unemployment-education gap. The first is the differences in fundamentals by education: labor productivity, y_e, number of potential careers, N_e, and learning speed, ϕ_e. Second is the difference in separations by match status ($\delta^b > \delta^g$). We now proceed to discuss the contribution of each to the unemployment-education gap. First, higher labor productivity ($y_1 > y_0$) leads to a higher match value with college workers, inducing more firms to post vacancies, and for college workers to exit unemployment with a higher probability.

Fewer potential careers, $N_0 < N_1$, has several effects. The first is $p_{i1} > p_{i0}$. Therefore, college workers produce more output in unsure matches, $y^{un}_{i1} > y^{un}_{i0}$, which leads to a higher job finding probability. Also, college workers experience fewer separations as they are more likely to be in their true calling. Moreover, from (4), $\partial p_{i1}/\partial i > \partial p_{i0}/\partial i$. So, ruling out a career has a larger impact on the probability the worker's next career is their true calling for college workers. This enables college workers to find their true calling, where they experience fewer separations, earlier in their career. The difference in learning speeds, $\phi_1 > \phi_0$, has a similar effect as it enables college workers to swiftly decipher a good fit from a bad fit, and quickly find their true calling.

The final implication of $N_0 < N_1$ is that college workers experience fewer separations when their status is unsure. This can be seen by noting

$$\delta^{un}_{i1} = p_{i1}\delta^g + (1 - p_{i1})\delta^b < p_{i0}\delta^g + (1 - p_{i0})\delta^b = \delta^{un}_{i0}, \tag{17}$$

as $p_{i1} > p_{i0}$. From (17), $N_0 < N_1$ generates differences in δ^{un}_{ie} if $\delta^b \neq \delta^g$. In particular, we have assumed $\delta^b > \delta^g$, which is why differences in separation probabilities by status contribute to the unemployment-education gap. It is important to note, however, that $\delta^b > \delta^g$ contributes to the unemployment-education gap because $p_{i1} > p_{i0}$.

What, then, drives the gap between δ^b and δ^g? We interpret $\delta^b > \delta^g$ as the manifestation of underlying match-specific productivity shocks. In a model where match output is made up of a common and idiosyncratic component, matches with a higher common productivity are less likely to be destroyed (e.g., Mortensen and Pissarides (1994)). In our model, $\delta^b > \delta^g$ as workers produce less output in a bad career fit.

Finally, the role of the uncertainty channel and its interaction with $\delta^b > \delta^g$ is more prominent early in a worker's career. As workers age and sample more careers, they are more likely to have found their true calling, experience fewer separations, and are less likely to be unemployed. This is especially true for non-college workers, as they face more uncertainty upon entering the labor market.

4 Quantitative Analysis

This section presents our calibration strategy, model validation, decomposition of the unemployment-education gap, education policy insights, and compares the implications of our model to those centered around match-specific productivity shocks.

4.1 Calibration

A unit of time is one month. The matching function is $F(u,v) = \frac{uv}{(u^\iota + v^\iota)^{1/\iota}}$. There are 18 parameters. The discount factor is $\beta = (0.97)^{1/12}$, and the probabilities of becoming old and dying $\lambda_o = \lambda_d = 1/(12 \times 20)$ so workers expect to spend 20 years in each age. The fraction of college workers is $\pi = 0.30$.[30] The economy is normalized by setting $z = 1$.

The remaining 13 parameters are calibrated via simulated method of moments to match 15 moments. The first moment is $z/[\text{average labor productivity}] = 0.4$ (Shimer, 2005). The second and third are the job finding probabilities for non-college (31.55%) and college workers (35.02%) aged 20 to 29. We also target the average number of unique careers worked by non-college (2.83) and college workers (2.00).[31] Next are the job finding probabilities for all workers between (i) 20 and 39 years old (31.12%) and (ii) 40 and 59 years old (27.61%). The remaining eight moments are the separation probabilities, by education and age bin, displayed in Figure 3(b).

[30] The fraction of 25-30 years old with at least a bachelor's degree between 1992-2017 in the CPS is 30%.
[31] Appendix C.1 details how we count unique careers in the data.

Table 4: Identification of N_e

N_0	5	6	7	8	9
# of Careers, Non-college	2.200	2.398	2.550	2.689	2.800
N_1	1	3	5	7	9
# of Careers, College	1	1.940	2.815	3.636	4.401

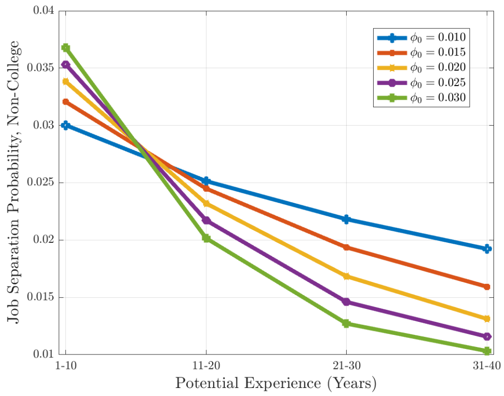
Note: The number of careers is computed by simulating the model economy and counting the number of careers each worker had worked in by the time they exit the labor market. All other parameter values are held constant at their calibrated values shown in Table 6.

While the targeted moments are affected by more than one parameter, one can view $\{\delta^g, \delta^b\}$ as targeting the separation probability for college workers in the last age group and non-college workers in the first age group. This is because the matches of college (non-college) workers are primarily composed of good (bad) matches in the last (first) age group. Next, κ_y (κ_o) targets the job finding probability of young (old) workers, as it affects firm entry for all young (old) workers. Next, $\{y_0, y_1\}$ target the job finding probability at 20-29 years old, by education, as higher output is associated with more vacancies.

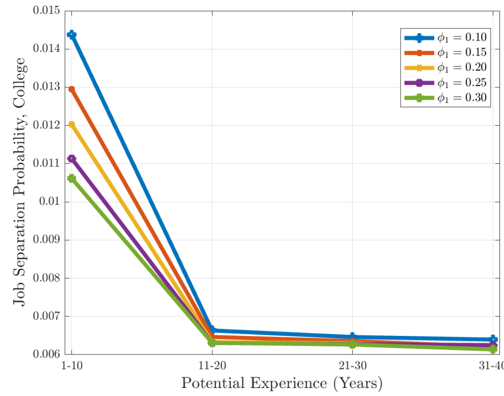
As for the parameters governing the uncertainty channel, $\{N_0, N_1\}$ targets the average number of unique careers worked by education. With a larger set of potential careers, workers expect to undergo more career experimentation before eventually settling into their best fit. This is demonstrated in Table 4.

We then use the "convexity" of the separation profile to pin down the probabilities of learning, $\{\phi_0, \phi_1\}$. As ϕ_e increases, workers learn about their fit at a higher rate. Once they realize that the current match is bad, they may endogenously separate from that match, leading to higher separations earlier in their career. However, with more learning occurring early in the worker's career, workers settle into their best fit and exhibit fewer separations later in their career. Therefore, ϕ_e influences the convexity of the separation profile, as it impacts how many separations workers experience early on and how quickly they can find their best fit (and therefore how rapidly separations decline over the life cycle). This is demonstrated in Figure 8.

Next, α, the output loss in a bad match, targets the $z/[\text{labor productivity}]$ ratio as it impacts the average output produced across all matches. The two remaining parameters, $\{\zeta, \iota\}$, fine-tune the model fit's to the 15 moments. The switching cost for old workers, ζ, improves the model's fit of the separation profile in the later half of workers' careers, as it influences how many old workers will stay in a bad match and be subject to a higher job destruction probability. Finally, ι impacts the responsiveness of job finding probabilities to changes in tightness and improves the fit of moments related to job finding.



(a) ϕ_0, Non-College



(b) ϕ_1, College

Figure 8: Identification of ϕ_e. *Notes*: The separation profiles are computed by simulating the model economy and computing the separation probability within each bin of potential experience. All other parameter values are held constant at their calibrated values shown in Table 6. The separation probabilities are corrected for time aggregation bias.

Denoting $\tilde{m}$ (m) as the vector of 15 model generated (empirical) moments, the vector of 13 parameters, $\hat{\vartheta}$, is given by

$$\hat{\vartheta} = \arg\min \left(\tilde{m} - m\right)' W \left(\tilde{m} - m\right), \tag{18}$$

where $W = I/m^2$ and I is the identity matrix. From (18), $\hat{\vartheta}$ minimizes the sum of squared percentage deviations between the model and data and does not place more weight on moments which are larger in magnitude. Table 5 shows that the model matches the targeted moments well.

Table 6 displays the parameter values. The calibration ascribes the gradual decline in the separation profile of non-college workers to a slower learning speed than college workers as $\phi_1 = 0.159$ and $\phi_0 = 0.019$. Next, we find $N_1 = 3$ and $N_0 = 8$, indicating that non-college workers enter the labor market with nearly three times as many careers that are potentially their best fit. It is important to emphasize that our results do not imply that college workers can work in fewer careers than non-college workers. Rather, they should be interpreted as college workers enter the labor market having narrowed down which careers are potentially their best fit. Thus, the calibration implies a large gap in uncertainty by education, enabling college workers to experience fewer separations and settle into their best fit at an earlier career stage.

Finally, $y_1/y_0 = 1.134$, indicating that there are large differences in productivity in the best fit. However, the model generates a larger gap in average labor productivity by education of 30.09%, as college workers are more likely to be in their best fit.

Table 5: Model and Data Comparison

Moment	Target	Model
JFP, 20-29, Non-College	0.316	0.316
JFP, 20-29, College	0.350	0.349
# Careers, Non-College	2.830	2.692
# Careers, College	2.000	1.938
JFP, 20-39	0.311	0.313
JFP, 40-59	0.276	0.276
SP by age bin, non-college ($\times 10^{-2}$)	[3.48, 2.16, 1.71, 1.43]	[3.37, 2.34, 1.71, 1.34]
SP by age bin, college ($\times 10^{-2}$)	[1.28, 0.71, 0.60, 0.61]	[1.27, 0.65, 0.63, 0.63]
z/[Labor Productivity]	0.400	0.400

Notes: Moments are computed by simulating the model economy. JFP (SP) stands for job finding probability (separation probability). The four age bins are: $[20-29, 30-39, 40-49, 50-59]$. Labor productivity is the average output across all matches.

4.2 Model Validation

Table 7 compares the model and data along some untargeted moments. The first two rows show that the model generates a life cycle unemployment pattern that closely tracks the data. This occurs even though we do not target the complete job finding profiles because, as shown in the third row, the unemployment-education gap (U-E gap for brevity, henceforth) is primarily driven by differences in separation probabilities.

As discussed in Section 2.3.3, the empirical relationship between prior experience and expected duration of a match is consistent with the uncertainty channel. The fourth row of Table 7 presents the estimated coefficient from regressing prior experience (in months) on the survival probability of the match in both the NLSY79 and simulated data.[32] The model captures this association well. Moreover, the fifth row shows that, just as in the data, the association between prior experience and match survival is significantly lower for college workers. This lends support to our hypothesis that the weaker association between prior experience and match duration for college workers is driven by the uncertainty channel. The sixth and seventh rows show that the model generates a realistic learning trajectory for each education group. In particular, college workers quickly settle into a good career fit and experience fewer unique careers, while non-college workers have sampled more careers at each stage. The eighth row shows that the model generates an average elasticity of job finding probabilities with respect to market tightness that is within an empirically

[32]The regression specification is detailed in Appendix A.6 and the untargeted moments are presented in Table A9, Panel B, Column (4).

Table 6: Parameter Values

	Definition	Value		Definition	Value
β	Discount factor	0.997	α	Penalty, bad fit	0.618
λ_o	Pr. of becoming old	0.004	y_0	Prod. of non-college	2.641
λ_d	Pr. of becoming retired	0.004	y_1	Prod. of college	2.996
π	Pr. endowed with $e = 1$	0.300	ζ	Switching cost	164
z	Utility while unemployed	1.000	N_0	# of careers, non-college	8
δ^g	Sep. pr., good fit	0.006	N_1	# of careers, college	3
δ^b	Sep. pr., bad fit	0.028	κ_y	Vacancy cost, young	1.023
ϕ_0	Learning pr., non-college	0.019	κ_o	Vacancy cost, old	3.202
ϕ_1	Learning pr., college	0.159	ι	Matching parameter	0.683

Notes: "Pr." is short for probability and "sep" is short for separation. The first five parameters in the left column are assigned while the remaining thirteen are estimated via simulated method of moments.

supported range of 0.5 to 0.7 (Petrongolo and Pissarides, 2001).

We now turn our attention to the last row of Table 7 and, in doing so, analyze wages for the first time. We have not discussed wages to this point as there are many wage contracts that can deliver the lifetime utility to the worker prescribed by the bilaterally efficient contract (Menzio and Shi, 2010). To study the model's implications for wages and lifetime earnings, we assume that wages are determined through Nash bargaining with constant renegotiation.[33] The last row of Table 7 shows that the model generates a college wage premium that is within the estimated range in our NLSY79 sample.[34]

We conclude this section by discussing the model's implications for involuntary and voluntary separations. While the difference between these types of separations, especially theoretically, can be difficult to interpret, let us label an exogenous (endogenous) separation in the model as involuntary (voluntary). Under our baseline calibration, the involuntary separation probability for non-college (college) workers is 1.84 (0.67). The voluntary separation probability for non-college (college) workers is 0.66 (0.22). While these do not exactly match the empirical rates shown in Appendix A.4.1, they are consistent in three ways. First, the separation probabilities for non-college are higher than college for both types of separations. Second, the gap in separation probabilities between non-college and college workers is higher for involuntary than voluntary separations. These first two features of the model are driven by non-college workers having more uncertainty over their career fit and, as a result, being hit with exogenous separation shocks at a higher rate while they are in the unsure state ($\delta^{un}_{i0} > \delta^{un}_{i1}$). Non-college workers expe-

[33] See Appendix B.2 for the derivation of wages under this assumption.

[34] See Appendix C.2 for details on the empirical estimates of the college wage premium.

Table 7: Model Validation

Untargeted Moments	Data	Model
Urate by age bin, non-college (%)	[9.9, 6.3, 5.2, 4.9]	[9.7, 7.5, 6.0, 5.0]
Urate by age bin, college (%)	[4.3, 2.3, 2.3, 2.7]	[4.1, 1.8, 1.8, 1.8]
Frac. of U-E gap explained by SP	1.213	0.778
β(PriorExp)	5×10^{-5}	5×10^{-5}
β(PriorExp $\times$ College)	-4×10^{-5}	-2×10^{-5}
# of careers by age bin, non-college	[2.6, 2.8, 2.8, 2.8]	[2.2, 2.6, 2.7, 2.7]
# of careers by age bin, college	[1.9, 2.0, 2.0, 2.0]	[1.9, 1.9, 1.9, 1.9]
Elasticity of JFP with respect to θ	0.5 - 0.7	0.572
College wage premium (%)	27.2-37.9	34.3

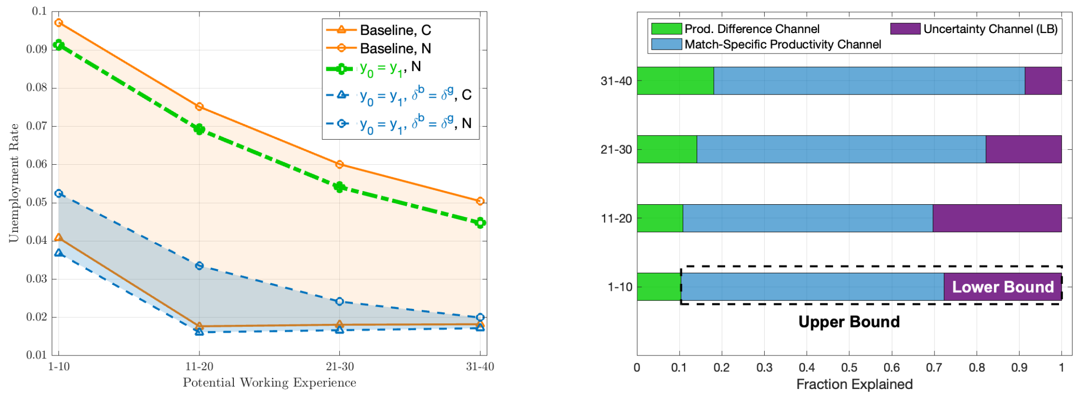
Notes: "Urate" refers to the unemployment rate, "frac." is fraction, "SP" is separation probability, and "JFP" is job finding probability. The four age bins are: $[20-29, 30-39, 40-49, 50-59]$. Rows four and five contain regression coefficients from estimating the regression detailed in Appendix A.6 on both the NLSY79 sample and simulated data. The last row presents estimates of the college wage premium in our NLSY79 sample from estimating the regression detailed in Appendix C.2 and the ratio of average wages of college to non-college workers in the simulated data.

rience voluntary separations at a higher rate for a similar reason, as they are more likely to learn they are in a bad career fit and subsequently separate. However, their voluntary separation rate is dampened by a low learning probability (ϕ_0), which limits the number of opportunities they have to experience a voluntary separation. Third, the model captures that involuntary separations occur at a higher rate than voluntary separations. The model's involuntary separation rate is 2.78 (3.04) times higher than the voluntary separation rate for non-college (college) workers, whereas this ratio typically falls between 3 and 6 at different stages of the life cycle for each education group in the data.

4.3 Decomposing the U-E Gap

There are three sources of the unemployment-education gap in the model: (i) labor productivity, (ii) the uncertainty channel, and (iii) the differences in the exogenous separation probabilities of good and bad career fits. This section evaluates the relative contributions of each to the model generated unemployment-education gap.

Figure 9(a) presents the unemployment profile by education from the model, the model without productivity differences, and the model with uncertainty channel only. To begin, we shut down productivity differences in good matches by setting $y_0 = y_1$. Doing so causes the U-E gap to slightly close, as the unemployment rate for non-college workers drops from the orange curve with circle markers to the green curve.



(a) Unemployment Rate Profiles

(b) Fraction Explained

Figure 9: Decomposition of the U-E Gap. *Notes*: Panel (a) shows the unemployment rate profiles, by education, after simulating the model economy under the specified parameters. A "C" ("N") in the legend indicates that the line is for college (non-college) workers. The orange shaded region represents the difference in unemployment rates for non-college workers after increasing y_0 to y_1 and lowering δ^b to δ^g. The blue shaded region is the portion of the unemployment-education gap which is attributable to the uncertainty channel. Panel (b) shows the corresponding fraction of the unemployment-education gap that closes after each successive change in the model's parameters.

Next, setting $\delta^b = \delta^g$ shuts down underlying match-specific productivity shocks which give rise to a higher separation probability in bad matches. Doing so results in a large reduction in unemployment for non-college workers and a modest reduction for college workers. Altogether, the portion of the U-E gap that is accounted for by differences in labor productivity and separations by match status is represented by the orange shaded region in Figure 9(a).

The blue shaded region in Figure 9(a) represents the gap in unemployment rates that is due to non-college workers (i) being more likely to end up in a bad match and endogenously separate from it ($N_0 > N_1$) and (ii) learning their best fit at a lower rate ($\phi_0 < \phi_1$). The fraction of the U-E gap at each age bin that is attributed to the uncertainty channel is 27.76%, 30.37%, 17.84%, and 8.62%, respectively.

Figure 9(b) illustrates the fraction explained by each channel by age bin.[35] The purple bars represent the portion of the U-E gap that is attributed to the uncertainty channel (the blue shaded region in Figure 9(a)). The blue bars represent the fraction explained by the match-specific productivity channel. As explained in Section 3.3, we interpret $\delta^b > \delta^g$ as the result of the interaction between underlying match-specific productivity shocks and the uncertainty channel, as workers are more productive at their best fit. Further, $\delta^b > \delta^g$

[35]The decomposition is robust to the order of decomposition, i.e., the results remain unchanged regardless of the sequence in which we break down the U-E gap into different channels. See Appendix C.3.

Table 8: U-E Gap Channels, Wages, and Lifetime Earnings

	Full Model	$y_0 = y_1$	$y_0 = y_1, \delta^b = \delta^g$
Wage, non-college	2.17	2.52	2.57
Wage, college	2.92	2.92	2.92
Lifetime earnings, non-college	543.04	633.30	671.20
Lifetime earnings, college	769.03	769.03	771.69

Notes: Wages and lifetime earnings are the averages across all workers in the simulated economy. Full Model refers to the model under the calibrated parameters listed in Table 6. The second column, $y_0 = y_1$, presents the outcomes when y_0 is set equal to y_1 and all other parameters take their baseline values. The last column displays model outcomes when both $y_0 = y_1$ and $\delta^b = \delta^g$.

contributes to the U-E gap because college workers are more likely to be in their best fit ($p_{i1} > p_{i0}$). Therefore, through our interpretation of $\delta^b > \delta^g$, part of the blue bars in Figure 9(b) are attributable to the uncertainty channel and the sum of the purple and blue bars represents the upper bound of the uncertainty channel's contribution to the U-E gap.

To arrive at an aggregate decomposition, we compute the weighted average of the fraction explained by each channel across age bins, where the weights are the fraction of employment observations at each bin. After eliminating differences in labor productivity at the best career fit, 87.85% of the U-E gap remains. After setting $\delta^b = \delta^g$, 24.26% of the U-E gap remains. Therefore, the uncertainty channel explains between 24.26% and 87.85% of the U-E gap.

Table 8 shows the implications of the U-E gap's channels on wages and lifetime earnings. From the first and second columns, shutting down the labor productivity channel leads to a 16.12% (16.62%) increase in wages (lifetime earnings) for non-college workers. The relative increase in wages and earnings are nearly the same because the labor productivity channel has little impact on the unemployment rate of non-college workers. Transitioning from the second to the third column, where we shut down the match-specific productivity channel, shows that wages (lifetime earnings) increase by 1.98% (5.98%) for non-college workers. Here, the impact on lifetime earnings is higher than wages as the unemployment rate for non-college workers substantially decreases after shutting down the match-specific productivity channel. At this point, the 13.6% (14.9%) gap in wages (lifetime earnings) by education in the last column is driven by the uncertainty channel. The remaining gap in wages is large despite having already increased y_0 to y_1 because non-college workers are still much more likely to be in a bad career fit, which significantly drags down their average labor productivity and wages. Taken together, the decomposition exercise shows that the uncertainty channel can account for a meaningful share of the model-generated education gaps in unemployment, wages, and lifetime earnings.

Table 9: Dissecting the Uncertainty Channel

	Non-college	$\phi_0 = \phi_1$	$N_0 = N_1$	$\phi_0 = \phi_1$, $N_0 = N_1$
Unemployment rate	[9.7, 7.5, 6.0, 5.0]	[11.6, 2.5, 2.4, 2.3]	[5.8, 3.3, 2.8, 2.6]	[4.3, 1.9, 2.0, 2.0]
Separation probability	[3.4, 2.3, 1.7, 1.3]	[3.8, 0.8, 0.7, 0.7]	[2.0, 1.0, 0.8, 0.8]	[1.3, 0.6, 0.6, 0.6]
Labor productivity	2.29	2.56	2.50	2.62
Wage	2.17	2.46	2.44	2.57
Lifetime earnings	543.04	627.73	634.04	675.81
Lifetime utility	504.90	570.95	569.59	604.12

Notes: The "non-college" column presents baseline outcomes for less than college. The second (third) column presents outcomes for less than college after setting ϕ_0 (N_0) equal to ϕ_1 (N_1). The last column shows outcomes for non-college workers with both $\phi_0 = \phi_1$ and $N_0 = N_1$. Unemployment and separations are reported as percentages and by age bin. Lifetime (discounted) utility is the value of a new entrant to the labor market, i.e. $U_{y,0,1}$. All other parameter values are held constant at their calibrated values shown in Table 6. Each reported outcome is the average across workers in the simulated economy.

To further understand the uncertainty channel's role in shaping unemployment, we evaluate the workings of its two components. The second column of Table 9 shows that raising ϕ_0 to ϕ_1 increases the unemployment rate early in the career of non-college workers, as they learn faster and are more likely to be in a bad match. Hence, they experience more separations. However, due to the higher learning speed and experiencing more separations early on, more non-college workers find their best fit by the second age bin and are less likely to be unemployed at the later stages of their career. The third column, $N_0 = N_1$, shows that reducing the number of potential best fits for non-college workers reduces unemployment across the life cycle. Finally, Table 9 demonstrates that endowing non-college workers with the same level of uncertainty as college workers leads to considerable increases in labor productivity, wages, earnings, and lifetime discounted utility.

4.4 Education Policy Implications

To this point, the empirical analysis in Section 2.1 shows that additional years of schooling are associated with lower forecast errors and the quantitative analysis suggests that differences in uncertainty make a meaningful contribution to the gaps in unemployment, wages, and lifetime earnings between non-college and college workers. Together, these findings lead us to ask what are the quantitative implications of implementing different education systems and policies which increase educational attainment. While the model does not speak to the channels through which education impacts uncertainty, it can be used to quantify the differences in uncertainty between specific education groups and

Table 10: Uncertainty Channel Parameters by Detailed Education Group

	Careers (data)	Careers (model)	Separation prob. (data)	Separation prob. (model)	ϕ	N
$<$ AA	2.867	2.869	[3.66, 2.28, 1.80, 1.49]	[3.51, 2.48, 1.83, 1.43]	0.021	9
AA	2.574	2.136	[2.06, 1.66, 1.19, 1.05]	[2.80, 1.71, 1.22, 1.02]	0.017	5
BA	2.276	1.914	[1.36, 0.81, 0.72, 0.72]	[1.45, 0.66, 0.65, 0.63]	0.099	3
$>$ BA	1.674	1.476	[0.99, 0.53, 0.49, 0.49]	[0.88, 0.63, 0.62, 0.61]	0.137	2

Notes: Careers is the number of unique careers and the targeted value is obtained from the NLSY79. Targets for separation probabilities come from the CPS. All other parameters are fixed at the values presented in Table 6. In particular, $<$AA and AA (BA and $>$ BA) have the same value of y_0 (y_1) from Table 6.

how those differences impact unemployment, productivity, and wages.[36]

We first apply the model to evaluate the quantitative implications of implementing a more specialized/vocational education system in the US. To do so, we separate the non-college group into those with and without an associate's degree (but less than a BA). We make this distinction as AA degrees provide what is arguably the closest analog to a specialized, vocational training that is provided in many European economies and that we have the data for that is needed to discipline the model.

We conduct the following exercise. We keep all parameters fixed at the values presented in Table 6 except for the learning speed, ϕ, and career set N. We then assign a learning speed and career set to two sub-groups within the non-college group: $\{\phi_{<AA}, N_{<AA}\}$ for those without an AA and $\{\phi_{AA}, N_{AA}\}$ for those with an AA. The two sets of parameters are chosen to match the number of unique careers and life cycle separation profile for those with and without an AA. The first two rows of Table 10 present the targets, model fit, and estimated learning speed and career sets. Despite being constrained by fixing the remaining parameters to their baseline values, the fit is good to these targets. Further, there is a large gap in the size of the career sets between these two groups.

The second and third columns of Table 11 show the impact of assigning ϕ_{AA} and N_{AA} to the less than associate's group. The largest effects come from reducing $N_{<AA}$ from 9 to 5. Doing so decreases unemployment by nearly 2.5 percentage points and increases labor productivity (wages) by 5.7% (7.4%). Further, through a combined effect of higher wages and a lower unemployment rate, lifetime earnings increase by 10.2%. These findings suggest that there are potentially large benefits to implementing an education system

[36]One example of a specific intervention which could be used to reduce career uncertainty is by exposing students to role models working in specific fields who share about their career path. For example, Breda et al. (2023) find that exposure to female role models working in scientific fields increased the probability of females enrolling in selective STEM programs.

Table 11: Education Policy Analysis (< AA vs. AA holders)

	$< AA$	$\phi_{<AA} = \phi_{AA}$	$N_{<AA} = N_{AA}$	AA
Unemployment rate (%)	8.20	8.28	5.85	6.06
Separation probability (%)	2.62	2.64	1.88	1.95
Labor productivity	2.27	2.24	2.40	2.38
Wage	2.15	2.12	2.31	2.28
Lifetime earnings	535.44	528.15	590.09	580.16
Lifetime utility	499.49	493.41	538.44	531.61

Notes: The "< AA" column presents baseline outcomes for less than AA workers. The second (third) column presents outcomes for less than AA workers after setting $\phi_{<AA}$ ($N_{<AA}$) equal to ϕ_{AA} (N_{AA}). The last column shows outcomes for less than AA workers with both $\phi_{<AA} = \phi_{AA}$ and $N_{<AA} = N_{AA}$. Lifetime (discounted) utility is the value of a new entrant to the labor market, i.e. $U_{y,<AA,1}$. Each reported outcome is the average across workers in the simulated economy.

that provides those with less than an AA the level of specialized training, and narrowed set of careers, akin to those acquired in associate's degree programs.

Lastly, we repeat the analysis two separate times for sub-groups within the college group. The first is a comparison between those with a BA only and those with above a BA. This exercise allows the model to speak to the effects of increasing the attainment of advanced degrees. The second, and is delegated to Appendix C.4, compares those with a STEM and Non-STEM undergraduate degree, which shifts the policy focus towards the effects of encouraging students to enroll in particular majors.[37]

The third and fourth row of Table 10 show that those with above a BA exhibit a higher learning speed and smaller set of potential best fits. Table 12 shows that assigning the above BA uncertainty channel parameters, $\{\phi_{>BA}, N_{>BA}\}$, to BA holders generates a modest improvement in most outcomes. This is because BA holders already have a narrow set of potential best fits with $N_{BA} = 3$, and therefore are likely to find their best fit early in their career. As a whole, Tables 11 and 12 indicate that the largest potential benefits of reducing worker's pre-labor market uncertainty come from narrowing down the set of potential best career fits among workers with less than an associate's degree.

4.5 Match-Specific Productivity

An alternative mechanism to generate the unemployment-education gap is the formation of match-specific productivity. If college workers have, on average, higher skills and labor

[37]The CPS does not report undergraduate major. As such, to do a STEM vs. Non-STEM calibration, we estimate the targeted separation probabilities from the NLSY79. This presents several measurement challenges, which we detail in Appendix C.4, and is why we leave this as a supplementary exercise.

Table 12: Education Policy Analysis (BA vs. > BA holders)

	BA	$\phi_{BA} = \phi_{>BA}$	$N_{BA} = N_{>BA}$	$> \text{BA}$
Unemployment rate (%)	2.89	2.76	2.20	2.12
Separation probability (%)	0.96	0.91	0.75	0.72
Labor productivity	2.96	2.97	2.97	2.98
Wage	2.90	2.91	2.93	2.93
Lifetime earnings	764.82	768.66	778.36	778.86
Lifetime utility	682.82	685.81	692.71	694.30

Notes: The "BA" column presents baseline outcomes for BA workers. The second (third) column presents outcomes for BA workers after setting ϕ_{BA} (N_{BA}) equal to $\phi_{>BA}$ ($N_{>BA}$). The last column shows outcomes for BA workers with both $\phi_{BA} = \phi_{>BA}$ and $N_{BA} = N_{>BA}$. Lifetime (discounted) utility is the value of a new entrant to the labor market, i.e. $U_{y,BA,1}$. Each reported outcome is the average across workers in the simulated economy.

productivity common to all matches, then they can sustain matches with a lower match-specific productivity and experience fewer separations and lower unemployment.[38] This environment could also generate the differences in separation rates, by education, over the life cycle shown in Figure 3(b) as older non-college workers are more likely to have found a match with high productivity and thus, exhibit a lower separation rate.

What, then, distinguishes the uncertainty channel from a mechanism which focuses only on the formation of match-specific productivity? First, following the intuition above, a model of match-specific productivity would predict that the match-specific component of productivity is lower among college graduates. While match-specific productivity is not directly observable, Guvenen et al. (2020) argue that skill mismatch can serve as a proxy for it. We show in Appendix A.3.3 that skill mismatch is, throughout the life cycle, lower among college graduates. This suggests that the average idiosyncratic component of match productivity is higher among college graduates and is contrary to what a standard model of match-specific productivity predicts. Second, environments that rely exclusively on shocks to or learning about match-specific productivity to generate separations predict that the expected duration of a match formed through unemployment is independent of the worker's experience. However, this "resetting" property is counterfactual, as shown in Section 2.3.3 and Appendix A.6. Third, one would expect the separation profile to continuously decline in age as older workers have higher match-specific

[38] This intuition follows from a standard search model with shocks to match-specific productivity (e.g., Mortensen and Pissarides (1994)). A closely related, but alternative theory is that due to college workers having higher skills, firms are more likely to layoff non-college workers following an aggregate productivity shock. However, this theory alone does not address why the unemployment-education gap narrows over the life cycle.

productivity. From Figure 3, the separation profile for college workers is essentially flat from 30 to 59 years old. As displayed in Table 5, our model matches this pattern well by having almost all college workers settled into their best fit after their first ten years of potential experience. Fourth, models of match-specific productivity do not speak to patterns in career mobility, nor do they address the differences in forecast errors by educational attainment we documented in Section 2.1.

As mentioned in Section 3.3, the presence of match-specific productivity shocks can rationalize why bad fits are destroyed at a higher rate. However, this is because workers are less productive outside their true calling, which is tied to the uncertainty channel. Our decomposition suggests that the interaction between the uncertainty and match-specific productivity channels are quantitatively meaningful, as $\delta^b/\delta^g = 4.66$ from Table 6. In this sense, the quantitative results suggest that the two channels should not be viewed in isolation from each other.

5 Conclusion

This paper posits the uncertainty channel as a new explanation for the unemployment-education gap. Using the NLSY79 and CPS, we document a set of facts to support the uncertainty channel: college graduates form more accurate expectations regarding their future occupation, the unemployment-education gap narrows over the life cycle, and separations are, especially for non-college workers, negatively associated with prior work experience and career sampling. To quantify the uncertainty channel, we develop a life cycle search model with uncertainty over one's best career fit, learning, and endogenous separations. The model is parameterized by matching features of the NLSY79 and CPS. Our decomposition reveals that the uncertainty channel accounts for meaningful shares of the model-generated education gaps in unemployment, wages, and lifetime earnings.

Existing research has primarily focused on the level of workers' skills by educational attainment. However, less attention has been given to how certain workers are about their own abilities, and how that certainty—or lack thereof—affects their capability to find their best fit in the labor market. Our empirical and quantitative findings indicate that not only do such differences in uncertainty exist between these two groups of workers, but they also play a significant role in generating differences in labor market outcomes by educational attainment and that reducing uncertainty among the least educated would be an admirable policy objective. With that said, this paper has not addressed the sources of the uncertainty channel. We leave this to future research.

References

Autor, D., D. Dorn, and G. Hanson (2019). When work disappears: Manufacturing decline and the falling marriage market value of young men. *American Economic Review: Insights 1*(2), 161–78.

Autor, D. H. and D. Dorn (2013, August). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review 103*(5), 1553–97.

Baley, I., A. Figueiredo, and R. Ulbricht (2022). Mismatch cycles. *Journal of Political Economy 130*(11), 2943–2984.

Bover, O., M. Arellano, and S. Bentolila (2002). Unemployment duration, benefit duration and the business cycle. *The Economic Journal 112*(479), 223–265.

Breda, T., J. Grenet, M. Monnet, and C. Van Effenterre (2023). How Effective are Female Role Models in Steering Girls Towards STEM? Evidence from French High Schools. *The Economic Journal 133*(653), 1773–1809.

Cairó, I. and T. Cajner (2018). Human capital and unemployment dynamics: Why more educated workers enjoy greater employment stability. *The Economic Journal 128*(609), 652–682.

Cajner, T., İ. Güner, and T. Mukoyama (2025). Gross worker flows over the life cycle. *Journal of Money, Credit and Banking 57*(4), 757–791.

Chéron, A., J.-O. Hairault, and F. Langot (2013). Life-cycle equilibrium unemployment. *Journal of Labor Economics 31*(4), 843–882.

Choi, W., J. Kinsler, A. Orellana, and R. Pavan (2023). College majors and earnings growth. Working paper.

Coskun, S. (2024). Young, educated, unemployed. Working paper.

Créchet, J., E. Lalé, and L. Tarasonis (2024). Life-cycle worker flows and cross-country differences in aggregate employment. Working paper.

Dorn, D. (2009). *Essays on inequality, spatial interaction, and the demand for skills*. Ph. D. thesis, Verlag nicht ermittelbar PhD diss.

Elsby, M. W., B. Hobijn, and A. Sahin (2010). The labor market in the great recession. NBER Working Paper No. 15979.

Elsby, M. W. L., R. Michaels, and G. Solon (2009). The ins and outs of cyclical unemployment. *American Economic Journal: Macroeconomics 1*(1), 84–110.

Esteban-Pretel, J. and J. Fujimoto (2014). Life-cycle labor search with stochastic match quality. *International Economic Review 55*(2), 575–599.

Feng, Y., D. Lagakos, and J. E. Rauch (2024). Unemployment and development. *The Economic Journal 134*(658), 614–647.

Flood, S., M. King, R. Rodgers, S. Ruggles, J. R. Warren, and M. Westberry (2022). Intergrated public use microdata series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS 2022.

Gathmann, C. and U. Schönberg (2010). How general is human capital? A task-based approach. *Journal of Labor Economics 28*(1), 1–49.

Gervais, M., N. Jaimovich, H. E. Siu, and Y. Yedid-Levi (2016). What should I be when I grow up? Occupations and unemployment over the life cycle. *Journal of Monetary Economics 83*, 54–70.

Girsberger, E. M. and R. Meango (2025). The puzzle of educated unemployment in West Africa. Working paper.

Gorry, A. (2016). Experience and worker flows. *Quantitative Economics 7*(1), 225–255.

Gorry, A., D. Gorry, and N. Trachter (2019). Learning and life cycle patterns of occupational transitions. *International Economic Review 60*(2), 905–937.

Gregory, V., G. Menzio, and D. Wiczer (2025). The alpha beta gamma of the labor market. *Journal of Monetary Economics 150*, 103695.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2021). What do data on millions of us workers reveal about lifecycle earnings dynamics? *Econometrica 89*(5), 2303–2339.

Guvenen, F., B. Kuruscu, S. Tanaka, and D. Wiczer (2020). Multidimensional skill mismatch. *American Economic Journal: Macroeconomics 12*(1), 210–44.

Hall, R. E. and M. Kudlyak (2022). Churn and stability: The remarkable heterogeneity of flows among employment, job search, and non-market activities in the US population. Working paper.

Jarosch, G. (2023). Searching for job security and the consequences of job loss. *Econometrica 91*(3), 903–942.

Kambourov, G. and I. Manovskii (2008). Rising occupational and industry mobility in the United States: 1968–97. *International Economic Review 49*(1), 41–79.

Menzio, G. and S. Shi (2010). Block recursive equilibria for stochastic models of search on the job. *Journal of Economic Theory 145*(4), 1453–1494.

Menzio, G. and S. Shi (2011). Efficient search on the job and the business cycle. *Journal of Political Economy 119*(3), 468–510.

Menzio, G., I. A. Telyukova, and L. Visschers (2016). Directed search over the life cycle. *Review of Economic Dynamics 19*, 38–62. Special Issue in Honor of Dale Mortensen.

Mortensen, D. T. and C. A. Pissarides (1994). Job creation and job destruction in the theory of unemployment. *Review of Economic Studies 61*(3), 397–415.

Moscarini, G. and K. Thomsson (2007). Occupational and job mobility in the US. *Scandinavian Journal of Economics 109*(4), 807–836.

Neal, D. (1999). The complexity of job mobility among young men. *Journal of Labor Economics 17*(2), 237–61.

Ozkan, S., J. Song, and F. Karahan (2023). Anatomy of lifetime earnings inequality: Heterogeneity in job-ladder risk versus human capital. *Journal of Political Economy Macroeconomics 1*(3), 506–550.

Papageorgiou, T. (2014). Learning your comparative advantages. *Review of Economic Studies 81*(3 (288)), 1263–1295.

Petrongolo, B. and C. A. Pissarides (2001). Looking into the black box: A survey of the matching function. *Journal of Economic Literature 39*(2), 390–431.

Pissarides, C. A. (2009). The unemployment volatility puzzle: Is wage stickiness the answer? *Econometrica 77*(5), 1339–1369.

Ruggles, S., S. Flood, M. Sobek, D. Backman, G. Cooper, J. A. Rivera Drew, S. Richards, R. Rodgers, J. Schroader, and K. C. Williams (2025). IPUMS USA: Version 16.0 [dataset]. Minneapolis, MN: IPUMS 2025.

Sengul, G. (2017). Learning about match quality: Information flows and labor market outcomes. *Labour Economics 46*, 118–130.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review 95*(1), 25–49.

Shimer, R. (2012). Reassessing the ins and outs of unemployment. *Review of Economic Dynamics 15*(2), 127–148.

Topel, R. H. and M. P. Ward (1992). Job mobility and the careers of young men. *Quarterly Journal of Economics 107*(2), 439–479.

Vardishvili, O. (2024). The macroeconomic cost of college dropout. Working paper.

Wee, S. L. (2013). Born under a bad sign: The cost of entering the job market during a recession. Working paper.

Online Appendix

A Empirical Appendix

A.1 Current Population Survey (CPS)

The Current Population Survey (CPS) is a monthly survey conducted by the U.S. Census Bureau and the Bureau of Labor Statistics, providing information on employment, earnings, and demographic characteristics of the U.S. labor force. The survey follows a rotation pattern in which households are interviewed for four consecutive months, then not interviewed for the next eight months, and finally interviewed again for another four months. We use the individual identifier, CPSIDP, to link individual records across time.

A.1.1 Educational Categories

The measurement of educational attainment was modified in January 1992. Prior to 1992, the CPS recorded the highest grade attended and years of education completed. Since 1992, the CPS has switched to reporting the highest degree obtained. To ensure comparability between them, we harmonize educational categories based on years of education or degree attainment. As shown in Table A1, "Non-College" includes individuals who have completed up to three years of college before 1992 or obtained at most an associate's degree afterward. "BA" encompasses those who completed four years of college in the old question or obtained a bachelor's degree in the new question. We also classify individuals who completed five or more years of college in the old question or obtained a master's degree in the new question as ''Master''. Additionally, ''Professional and Doctorate Degree'' includes individuals with either a professional, or doctorate degree. Overall, ''College'' refers to individuals who completed at least four years of college in the old question or attained at least a bachelor's degree in the new question.

A.1.2 Occupation Distance Measurement

To measure the distance between occupations, we begin by characterizing each occupation by a skill vector, where each element represents the required level of a specific skill to perform that job. In particular, we measure occupational requirement across multiple dimensions: (i) verbal, math, social, and technical skill requirements as in Guvenen et al. (2020); and (ii) abstract, routine, and manual task intensities as in Autor and Dorn

Table A1: Potential Experience by Education

<table>
<tr><th>Category</th><th>Refined Category</th><th>CPS Education</th><th>Potential Exp.</th></tr>
<tr><td>Non-College</td><td>Non-College</td><td>$<$ 4 years of college</td><td>$Age - 18 + 1$</td></tr>
<tr><td rowspan="7">College</td><td rowspan="2">BA</td><td>4 years of college</td><td rowspan="2">$Age - 22 + 1$</td></tr>
<tr><td>Bachelor's degree</td></tr>
<tr><td rowspan="4">Master</td><td>5+ years of college</td><td>$Age - 23 + 1$</td></tr>
<tr><td>5 years of college</td><td>$Age - 23 + 1$</td></tr>
<tr><td>6+ years of college</td><td>$Age - 24 + 1$</td></tr>
<tr><td>Master degree</td><td>$Age - 24 + 1$</td></tr>
<tr><td>Professional and Doctorate Degree</td><td>Professional degree
Doctorate degree</td><td>$Age - 28 + 1$</td></tr>
</table>

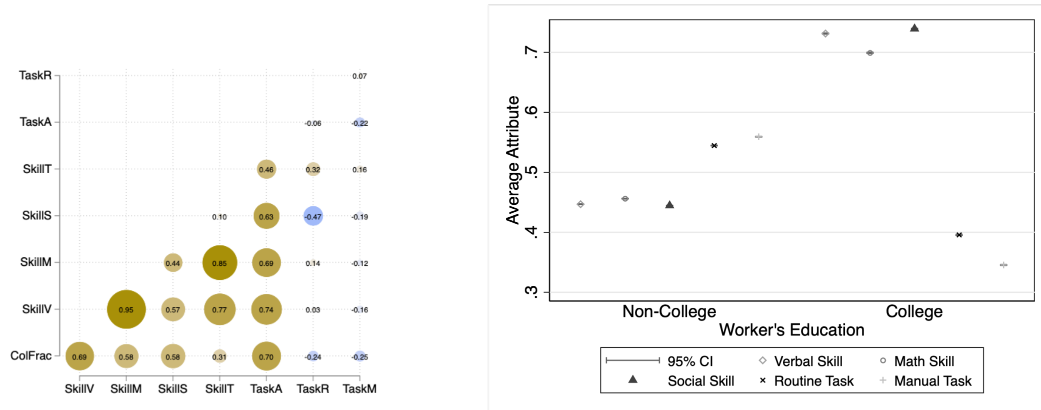
Note: This table shows the mapping between a respondent's educational attainment and presumed years of potential experience.

(2013).[39] Figure A1(a) displays the pairwise correlation between these attributes and the proportion of respondents in the O*NET survey reporting that at least a bachelor's degree is required to perform that job. Jobs with a higher college fraction are positively related to the amount of verbal, math, social, technical skill requirements, as well as the abstract task intensity. Conversely, routine and manual task intensity is negatively correlated with the college fraction. As such, we select verbal, math, and social skills to capture the high-order skills and incorporate the routine and manual task intensity to capture the low-order skills.[40] Furthermore, we examine the average occupational attributes in jobs held by non-college and college workers in the NLSY79 sample of $1,152,280$ employment observations. Figure A1(b) shows that jobs held by college graduates have higher requirements for verbal, math, and social skills, whereas routine and manual task intensities are lower. Overall, these five attributes capture well the lower- and higher-order skills required by occupations.

To give an example of the skill measures and how those map into the angular and Euclidean distance measures, Figure A2 compares eight occupations to dentists. First, Figure A2(a) visually represents the skill mix in each of the occupations. Figure A2(b) plots the angular and Euclidean distances between the fixed occupation of "dentists" and each of the respective occupations.

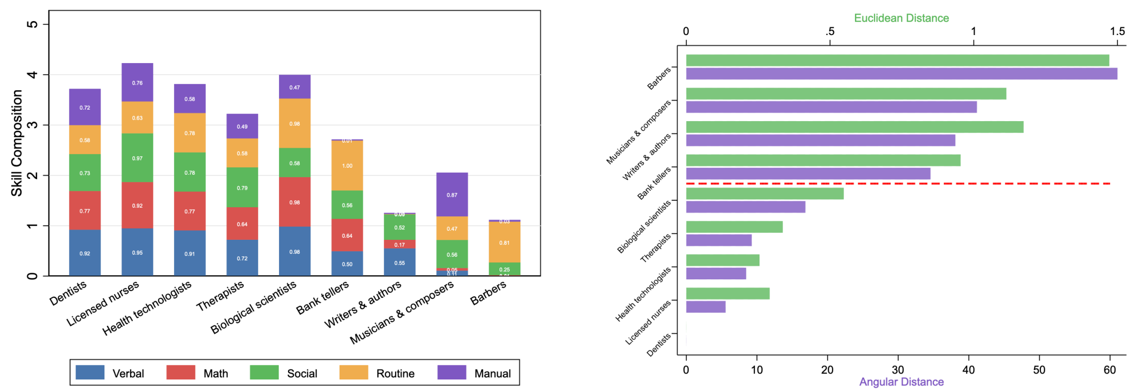
[39] We follow the steps outlined by both Guvenen et al. (2020) and Autor and Dorn (2013) in the measurement of occupational attributes and, for brevity, omit those detailed steps here.

[40] We do not incorporate technical skill requirement or abstract task intensity measure in the skill vector as both are highly correlated with verbal and math skill requirements.



(a) Correlogram of Occupation Attributes (b) Average Attributes

Figure A1: Occupation Attributes. *Note*: Figures are constructed using NLSY79 data.



(a) Skill Mix (b) Distance to Dentists

Figure A2: Comparison of Eight Occupations to Dentists. *Notes*: Panel (a) shows the skill and task requirements for each of the nine occupations used in this example. Panel (b) compares the angular and Euclidean distance between each of the listed occupations to dentists.

A.2 National Longitudinal Survey of Youth (NLSY79)

The National Longitudinal Survey of Youth (NLSY79) is a longitudinal survey that tracks the labor market histories of a youth cohort aged 14 to 22 when first surveyed in 1979. Conducted by the U.S. Bureau of Labor Statistics, it provides comprehensive information on employment, education, training, income, and family status.

A.2.1 Sample Construction

We first construct a weekly panel data from original NLSY79 files, involving three key steps: (i) cleaning the employer history roster and determining employer characteristics,

(ii) identifying necessary demographic variables for each respondent in each survey year, and (iii) identifying the primary job for each week if the worker holds multiple jobs.

Next, to match the time structure of the NLSY79 sample with the CPS, we convert the weekly panel to a monthly panel by identifying the primary labor force status for each month. The primary job for each month is determined as the one with the most working hours. If multiple civilian jobs have the same total working hours, we consider the job with complete occupation and industry records as the primary one. If several jobs have complete records, we retain the one with a known employer ID. If there are still multiple civilian jobs in a particular month, we keep the earliest reported one, indicated by a lower job code in the weekly array.

If the respondent does not hold any job with valid job codes for a given month, we prioritize the remaining labor force statuses in the following order: 3 (employed, but periods not working with an employer are missing) > 4 (unemployed) > 5 (out of the labor force) > 2 (period not working with an employer, unsure if unemployed or out of the labor force) > 7 (military) > 0 (no information). The highest-priority status for the month is assigned as the respondent's primary labor force status. Online Appendix D contains further details on the NLSY79 sample construction.

A.2.2 Sample Selection

We start with monthly employment histories of 12,686 respondents and subsequently restrict the sample to 6,403 males, as female labor force participation exhibited substantial changes throughout the survey period.[41] Next, we filter the observations to include only those from the earliest survey year (1978) until 2018. Table A2 summarizes the sample selection criteria.

We assume that individuals enter the labor market upon completing their highest level of education. For those whose highest education level is recorded as ''None'', we set their employment histories to start in 1978, the earliest year available in our dataset. We drop respondents with unknown graduation dates from our sample, which leads to a sample of 6,386. Subsequently, we exclude individuals who have served in the military, leaving a sample size of 5,361 respondents. Finally, we drop individuals with either incomplete cognitive or non-cognitive scores, resulting in a sample size of 4,823 respondents.

[41] For example, the labor force participation rate of female increases from 50% in 1978 to around 60% starting in 1997.

Table A2: NLSY79 Sample Selection

Criteria	No. Respondents	No. Observations
Restrict to males	6,403	2,317,473
Monthly histories from 1978 to 2018	6,403	2,307,286
Start from the (known) graduation year	6,386	1,805,924
Never served in the military	5,361	1,589,597
Complete ASVAB	5,030	1,511,337
Complete non-cognitive scores	4,823	1,452,307

Note: This table details the steps taken to construct the NLSY79 sample and the corresponding sample size after each sample restriction is implemented.

A.2.3 Measurement of Worker's Aptitudes

To measure a worker's verbal and math skills, we begin with a sample of 4,823 respondents who have complete scores for the word knowledge, paragraph comprehension, arithmetic reasoning, and mathematics knowledge sub-tests of the Armed Services Vocational Aptitude Battery (ASVAB). We normalize the mean and variance of each test score within each age cohort. To identify verbal and math abilities for each individual, we perform Principal Component Analysis (PCA) separately on the first two sub-tests (word knowledge and paragraph comprehension) and the last two sub-tests (arithmetic reasoning and mathematics knowledge). By extracting the first component from each PCA, we obtain measures of verbal and math abilities. Subsequently, we convert these ability indicators into percentile ranks across all individuals.

To measure social skills, we use the Rotter Locus of Control Scale and the Rosenberg Self-Esteem Scale. Following a similar approach as with math and verbal skills, we adjust for the effect of test-taking age and extract the first principal components from the standardized scores of these two tests as the social ability measure.

A.2.4 Measurement of Skill Mismatch

To quantify the mismatch between workers' abilities and occupational requirements, we compute the distance between the percentile ranks of worker abilities and their corresponding occupational attributes. In particular, the mismatch in aptitude j between worker i and occupation o is given by:

$$m_{i,j,o} = \left| q(A_{i,j}) - q(s_{o,j}) \right| , \tag{A.19}$$

where $q(A_{i,j})$ represents the percentile rank of worker i in skill j, and $q(s_{o,j})$ denotes the requirement percentile of occupation o in skill j. The aggregate mismatch is then defined as:

$$m_{i,o} = \sum_j \{\omega_j \left|q(A_{i,j}) - q(s_{o,j})\right|\}, \tag{A.20}$$

where ω_j represents the weights assigned to each skill j, reflecting the relative importance of the difference in that skill to the aggregate skill mismatch. These weights are determined by factor loadings obtained from the normalized first principal component analysis. In particular, the respective weights for verbal, math, and social are $(0.43, 0.42, 0.15)$, which is similar to the weights in Guvenen et al. (2020).

A.3 Additional Motivating Facts

A.3.1 Forecast Error and Separations

From the NLSY79, we observe that 115 (88) out of 1,961 (604) non-college (college) workers accurately predicted their occupation at age 35. This section compares the separation profiles of workers with and without forecast errors within each education group. As shown in Figure A3, within each education group, workers who did not make forecast errors – or, had little uncertainty about their best career – exhibit lower separations than those who did. Furthermore, the difference in separation rates by forecast error is widest early in workers' careers, supporting the notion that individuals with greater uncertainty about their best fit experience higher separation rates in the initial stages of their careers. Over time, as they find their true calling, these separation rates decline.

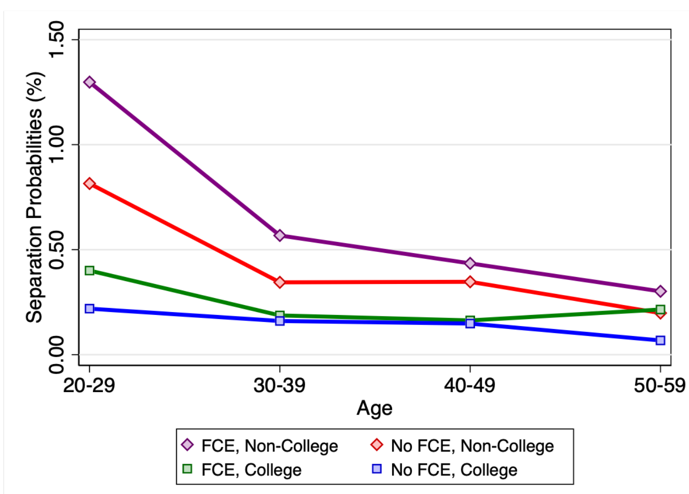
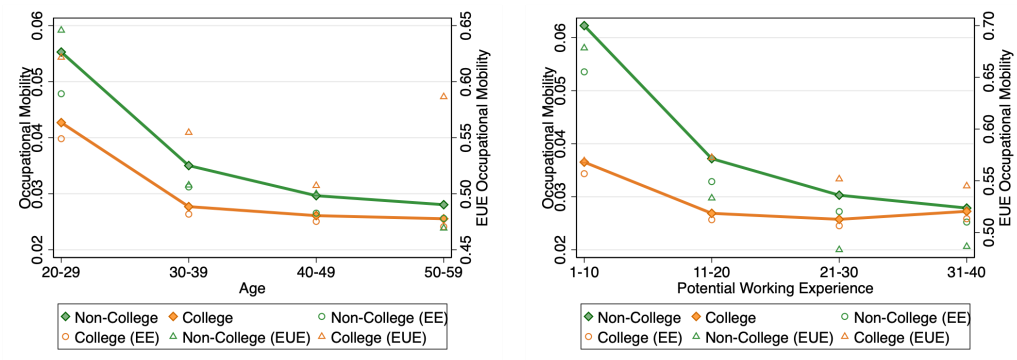


Figure A3: Forecast Error and Separations. *Note*: Constructed using NLSY79 sample.



(a) Age (b) Potential Experience

Figure A4: Occupational Mobility. *Note*: Figures are constructed using CPS data between 1994-2019.

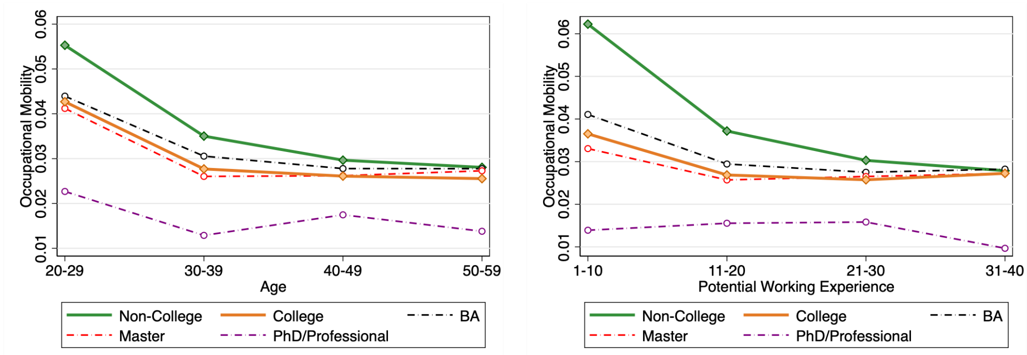
A.3.2 Occupational Mobility

Using CPS data from 1994 to 2019, we compute monthly 3-digit occupational mobility rates by age and education.[42] We do this by separately computing occupational mobility for job-to-job (EE) transitions and transitions from unemployment (EUE). For EE transitions, we restrict to observations with known occupations for two consecutive months. For EUE switches, we track the occupations before and immediately following the unemployment spell. We arrive at the aggregate occupational mobility rate by taking a weighted average across all transitions, incorporating essential correction to address potential measurement error in the mobility rates.

Figure A4 presents occupational mobility rates over ages or potential working years. The diamonds (triangles) represent occupational switches in EE (EUE) transitions, while the solid line is the overall fraction of workers who switch occupations each month. There are three patterns to highlight. First, occupational mobility is decreasing in age (Kambourov and Manovskii, 2008). Second, non-college workers change occupations more frequently. Third, similar to unemployment and separations, the difference in occupational mobility rates across two education groups decreases with age. To further support the notion that highly educated workers experience less occupational mobility given their lower uncertainty, we report the 3-digit occupational mobility rates for detailed educational attainments. Figure A5 shows that holding a Master's, Ph.D., or Professional degree is associated with even lower occupational mobility rates.

[42]Occupational records in survey data are prone to measurement error. To mitigate this concern, we apply the methodology proposed by Moscarini and Thomsson (2007) which leverages the dependent questions introduced in the CPS starting in 1994. The correction process involves three stages: first, flagging transitions susceptible to measurement error in occupational codes; second, subjecting these dubious transitions to the *ANY*3 filter; and finally, passing the remaining suspicious transitions through the *Flag* filter. Detailed procedures are omitted here but are available upon request.

One factor complicating the interpretation of mobility over age is that educational attainment affects the timing of labor market entry. Hence, we also show the occupational mobility rates by presumed years of potential experience, assuming that non-college (college) workers enter the labor market at the age of 18 (22). Figure A4(b) illustrates that while the overall pattern is unchanged, the gap in occupational mobility rates becomes even larger in early career stages than when compared by age.



(a) Age (b) Potential Experience

Figure A5: Occupational Mobility Across Specific College Degrees. *Note*: Figures are constructed using CPS data between 1994-2019.

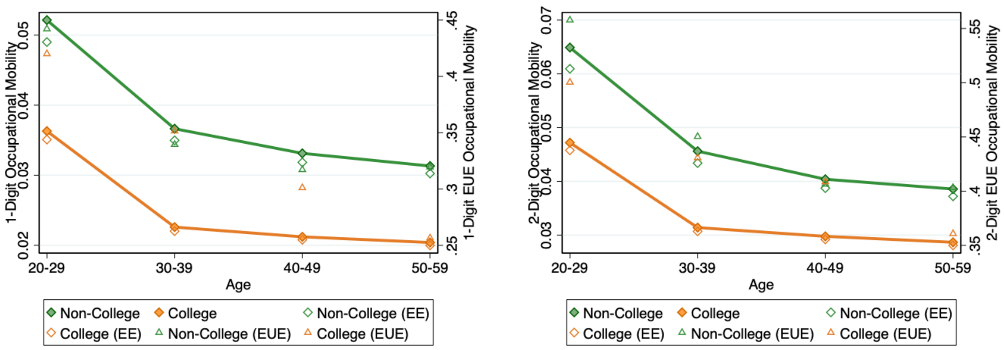
Occupational Mobility in Broader Categories Occupational mobility within broader occupation categories is less susceptible to measurement error because there is less overlap between occupations and, hence, less of a chance that a worker's occupation is misclassified. Figure A6 presents the raw occupational mobility rates by using 1- and 2-digit occupational codes. The patterns are consistent with those shown in Figure A4.

A.3.3 Skill Mismatch

Based on the skill mismatch for each worker-job pair as outlined in Appendix A.2.4, we compute the average skill mismatch disaggregated by age and educational attainment, denoted as $\overline{MM}_{i,j}$:

$$\overline{MM}_{i,j} = \frac{\sum_{k \in i \cap j} MM_k \times \omega_k}{\sum_k \mathbb{1}\{k \in i \cap j\} \times \omega_k}. \tag{A.21}$$

From equation (A.21), $\overline{MM}_{i,j}$ is given by the ratio of the aggregate mismatch among individuals with age i and education j to the number of individuals within that subgroup. We apply the technical weight ω_k to accounts for each respondent's representation in the



(a) 1-digit (b) 2-digit

Figure A6: Occupational Mobility at 1- and 2-digit Occupation Codes. *Note*: Figures are constructed using CPS data between 1994-2019.

U.S. population. Figure A7 shows that aggregate skill mismatch is decreasing with higher educational attainment.[43]

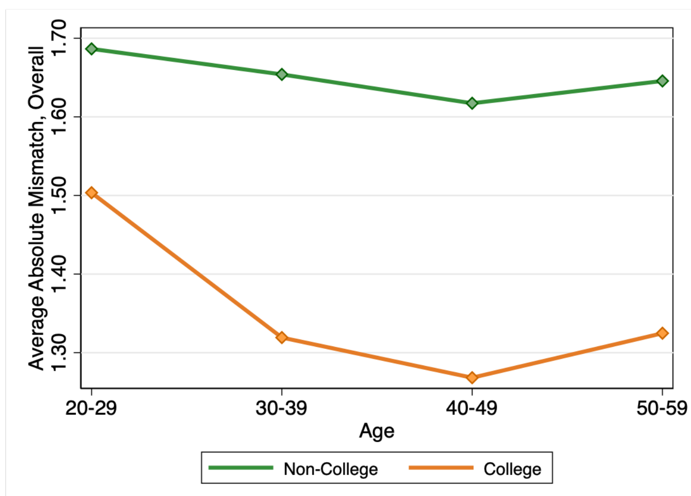


Figure A7: Skill Mismatch by Age and Educational Attainment. *Note*: Constructed using NLSY79 sample.

A.3.4 Dispersion in Skill Requirements

In this section, we compare the variance of occupational skill requirements across age or potential experience and educational attainment. The degree of dispersion is suggestive of workers' uncertainty regarding their comparative advantages. Specifically, workers more certain of their best fit may choose occupations with more imbalanced skill requirements, indicating their assurance in excelling in jobs that emphasize particular skills. We

[43]Similar patterns are observed for each single skill dimension and are available upon request.

measure the degree of skill dispersion using the following metrics:

$$Var_i = \frac{\sum_j (r_{i,j} - \bar{r}_i)^2}{5}, \quad Max - Min_i = max(r_{i,j}) - min(r_{i,j}),$$
$$MeanDev_i = \frac{|\sum_j (r_{i,j} - \bar{r}_i)|}{5}, \quad MedianDev_i = \frac{|\sum_j (r_{i,j} - Median_i)|}{5},$$

where $r_{i,j}$ denotes the skill requirement along skill j by occupation i , and $\bar{r}_i$ ($Median_i$) denotes the mean (median) value of the skill requirement in occupation i.

Table A3 shows that the college workers are employed in occupations with more dispersed skill requirements, lending support to the notion that more educated workers have a higher degree of certainty regarding which kind of job is a best fit for them.

Table A3: Degree of Skill Requirement Imbalance

	Age	Working Experience
Panel A: Variance		
Non-College	[0.16, 0.18, 0.18, 0.18]	[0.17, 0.18, 0.18, 0.18]
College	[0.21, 0.23, 0.23, 0.23]	[0.23, 0.23, 0.23, 0.23]
Panel B: Max-Min Differences		
Non-College	[0.60, 0.61, 0.60, 0.61]	[0.60, 0.61, 0.60, 0.61]
College	[0.65, 0.66, 0.66, 0.66]	[0.66, 0.66, 0.66, 0.67]
Panel C: Mean Absolute Deviation		
Non-College	[0.35, 0.37, 0.37, 0.37]	[0.35, 0.37, 0.37, 0.37]
College	[0.39, 0.40, 0.41, 0.41]	[0.40, 0.41, 0.40, 0.40]
Panel D: Median Absolute Deviation		
Non-College	[0.18, 0.18, 0.18, 0.18]	[0.18, 0.18, 0.18, 0.18]
College	[0.20, 0.21, 0.21, 0.21]	[0.20, 0.21, 0.21, 0.21]

Notes: Working experience refers to years of potential experience. Data from NLSY79, 1979:1-2018:12.

A.3.5 Number of employer, occupation, and career changes

To compare career stability across educational attainment, we examine the average number of cumulative transitions experienced by age and educational attainment. This involves a two-step process. First, we calculate the average number of employer, occupational, and career switches within each subgroup. Second, we compute the cumulative average transitions by aggregating these averages across all preceding age bins.

Table A4: Cumulative Transitions by Age

	$20-29$	≤ 39	≤ 49	≤ 59
Panel A: Employer Transitions				
Non-College	4.43	6.90	8.35	9.08
College	1.91	3.57	4.73	5.57
Panel B: Unique Employers				
Non-College	5.01	6.56	7.33	7.71
College	2.76	3.82	4.37	4.78
Panel C: Occupation Transitions				
Non-College	4.97	7.80	9.13	9.76
College	3.02	5.40	6.61	7.38
Panel D: Career Transitions				
Non-College	3.00	4.65	5.41	5.78
College	1.44	2.44	2.85	3.17

Note: Data from NLSY79, 1979:1-2018:12.

Table A4 shows that individuals tend to accumulate transitions as they age. More importantly, individuals with higher educational attainment tend to experience fewer switches across all transition types at any career stage. Notably, workers without a college degree experience nearly twice as many career transitions over their life cycle.

A.3.6 Unemployment Rate by College Major

We use the American Community Survey (ACS) to compute the unemployment rate by college major, as the CPS does not record college major and the sample size of any individual college major in the NLSY79 is small. To do so, we download the ACS data covering 2009-2019 from IPUMS Ruggles et al. (2025). We restrict our sample to white males between the ages of 16-59 who are not in school, not veterans, and have a valid undergraduate major (for those who attended college). After these steps, we arrive at a sample of 1.6 million observations. We then place individuals who obtained at least a bachelor's degree and have a valid undergraduate major into a broader major category: Engineering and Computer Science, Business, Life and Physical Sciences, Social Sciences, and Other following Choi et al. (2023). We then compute the unemployment rate by college major category and among those who obtained less than a bachelor's degree. Figure A8 presents the results.

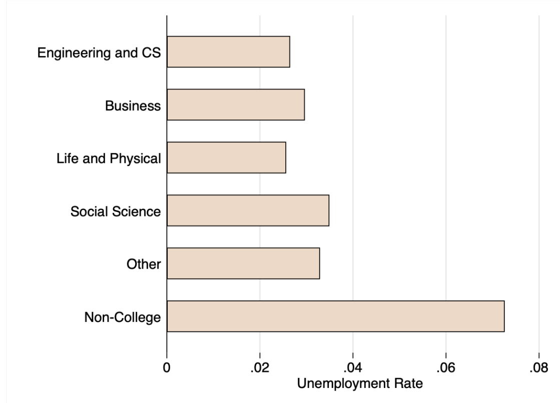


Figure A8: Unemployment Rate by College Major. *Notes:* The categorization of majors follows Choi et al. (2023). Non-college includes all respondents who acquired less than a bachelor's degree. Data are from the American Community Survey between 2009-2019.

A.4 Robustness Checks

A.4.1 Transition Probabilities

Aggregate Employment Profile Table A5 shows that, in the aggregate, college graduates are less likely to be unemployed and have a lower separation risk. College graduates also exhibit a lower job finding probability and rate than those without a college degree.

Table A5: Aggregate Employment Profile, by Education, PP

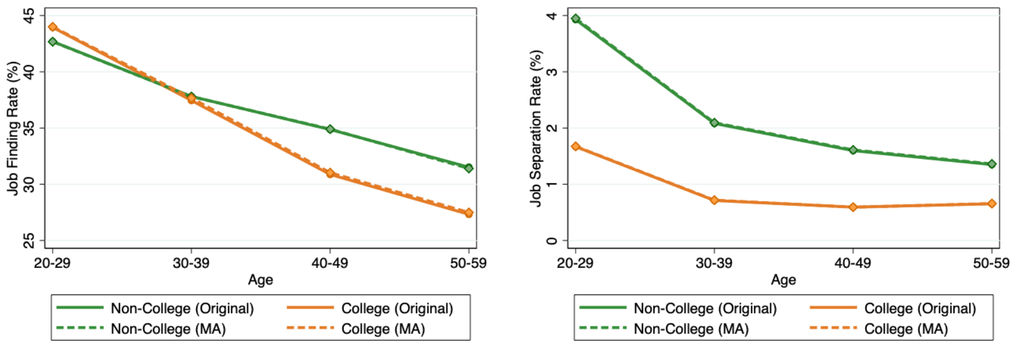
	Urate	JFP	JSP	JFR	JSR
Non-College	6.88	27.92	1.87	37.81	2.21
College	2.74	27.19	0.63	32.37	0.79

Note: The first three columns are computed from CPS, 1976:1 - 2019:12, while the last two are computed from CPS: 1994:1-2019:11.

Job Finding and Separation Rates Following Shimer (2005) and Elsby et al. (2009), the unemployment outflow (f_t) and inflow rates (s_t) for each cohort of age i and education j can be derived starting with the law of motion for unemployment:

$$u_{t+1} = (1 - F_t)u_t + u^s_{t+1} \quad \Rightarrow \quad F_t = 1 - \frac{u_{t+1} - u^s_{t+1}}{u_t}, \tag{A.22}$$

where F_t is the monthly outflow probability. Equation (A.22) states that the number of unemployed workers at month $t+1$, u_{t+1}, is equal to the number of unemployed workers at month t who did not find a job with probability $(1 - F_t)$, plus the number of short-term



(a) Job Finding Rate (b) Separation Rate

Figure A9: Original and 12-month Moving Average Transition Rate. *Note*: Figures are constructed using CPS data between 1994-2019.

unemployed workers who are unemployed at month $t+1$, but employed at month t, denoted by u^s_{t+1}. Therefore, the outflow rate f_t can be derived from $f_t = -\log(1 - F_t)$.

To compute s_t, we start from the law of motion for unemployment:

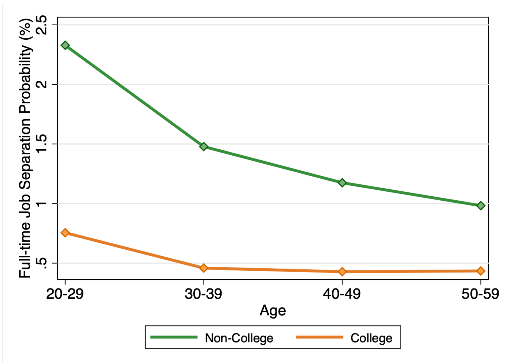
$$\dot{u} = \overbrace{s_t(l_t - u_t)}^{inflow} - \overbrace{u_t f_t}^{outflow} = -(s_t + f_t)(u_t - u^*), \tag{A.23}$$

where u^* is the steady state unemployment and l_t is the size of the labor force. The second equality comes from the labor market equilibrium condition $s_t e^*_t = u^* f_t$. By solving (A.23) and assuming s_t, f_t and l_t are constant between surveys, we can infer s_t from

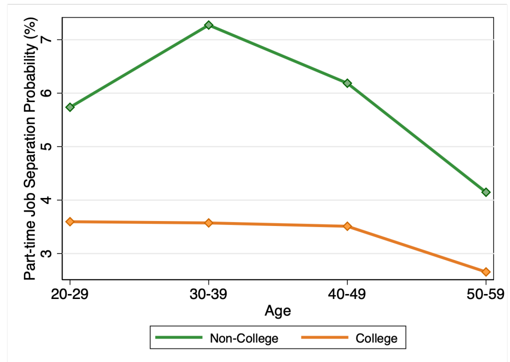
$$u_{t+1} = \frac{(1 - e^{(-s_t - f_t)})s_{t+1}}{f_{t+1} + s_{t+1}} l_t + u_t e^{(-s_t - f_t)}. \tag{A.24}$$

To compute the inflow and outflow rates, we first compute the unemployment rate for each subgroup defined by age i and education j. In the same manner, we calculate the short-term unemployment rate for each subgroup, where short-term unemployment (denoted $u^{ij}_{t,s}$) is defined as a duration of less than 5 weeks. From here, we can readily infer the hazard rates from equations (A.22) and (A.24).[44] Finally, we take a 12-month moving average. Figure A9 shows that the age profile patterns of the transition rates closely resemble those of the transition probabilities shown in Figure 3.

[44] Observations before 1994 were discarded because the unemployment duration variable is only available in IPUMS-CPS data starting from 1994.



(a) Full Time

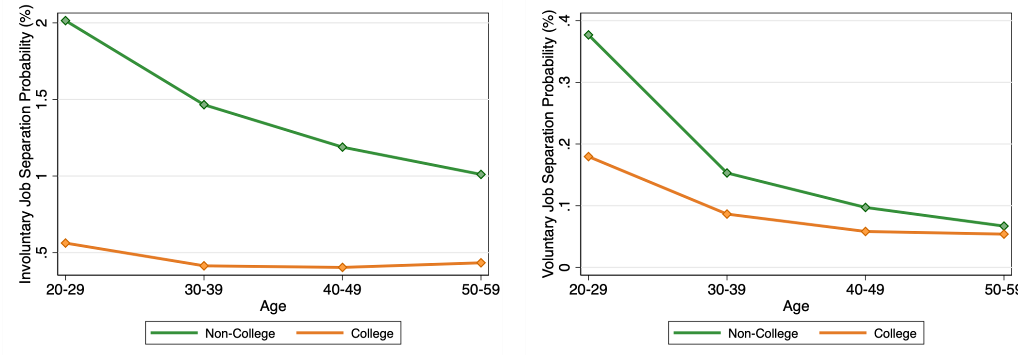


(b) Part Time

Figure A10: Separation Probability by Working Status. *Note*: Figures are constructed using CPS data between 1976-2019.

Separation Probability by Working Status Workers without a college degree are more likely to hold part-time jobs, which might lead to more separations unrelated to their certainty of comparative advantage, such as seasonal employment. To demonstrate that part-time employment is not a key driver of the U-E gap, we provide several pieces of evidence. First, there is no systematic compositional difference across education-age groups in terms of working status. For example, the fraction of full-time employment for non-college workers is about 85%/95%/95%/94% at each age bin, which is close to 91%/96%/96%/95% for college workers. Second, Figure A10 shows that, even among full-time workers, those with less education have higher separation rates. Finally, as shown in Section A.4.3, the observed patterns persist after controlling for month fixed effects.

Involuntary and Voluntary Separations To examine voluntary and involuntary separations in the data, we leverage the reason for unemployment in the CPS. Respondents listing "job loser – on layoff", "other job loser", or "temporary job ended" as their reason for being unemployed are classified as involuntarily unemployed, while those listing "job leaver" are classified as voluntarily unemployed. Figure A11 shows that both types of separations occur at a higher rate among non-college workers. Further, the gap in each type of separation rate narrows over the life cycle. Section 4.2 discusses how our quantitative model is consistent with the fact that involuntary separations occur at a higher rate than voluntary separations for both education groups and that, for both separation types, non-college workers separate at a higher rate.



(a) Involuntary Separations

(b) Quits

Figure A11: Involuntary Separations and Quits. *Note*: Figures are constructed using CPS data between 1976-2019.

A.4.2 U-E Gap Decompositions

We employ the method by Pissarides (2009) to decompose the U-E gap at each age bin into differences in the job finding and separation probabilities (rates). Denoting s_{ij} and f_{ij} as the job separation and finding probabilities (rates) for age group i with educational attainment j, the steady-state unemployment rate for subgroup ij is given by:

$$u_{ij} = \frac{s_{ij}}{s_{ij} + f_{ij}}. \tag{A.25}$$

Taking first differences of (A.25) between education levels j and j' gives

$$1 = \underbrace{\frac{(1 - u_{ij})u_{ij'}\frac{(s_{ij} - s_{ij'})}{s_{ij'}}}{\Delta u_i}}_{\text{Fraction explained by SP}} + \underbrace{\frac{-u_{ij}(1 - u_{ij'})\frac{(f_{ij} - f_{ij'})}{f_{ij'}}}{\Delta u_i}}_{\text{Fraction explained by JFP}}. \tag{A.26}$$

Table A6 presents the fraction of the U-E gap at each age bin i that is attributable to difference in the job finding and separation probabilities (rates). Each decomposition indicates that the U-E gap is primarily driven by differences in separation probability/rate.

Table A6: Decomposition of the U-E Gap by Age Bin

	20-29	30-39	40-49	50-59
Panel A: Job Finding/Separation Probability				
Separation Probability	0.85	1.04	1.29	1.66
Job Finding Probability	0.15	-0.04	-0.29	-0.66
Panel B: Job Finding/Separation Rate				
Separation Rate	0.95	1.01	1.22	1.35
Job Finding Rate	0.05	-0.01	-0.22	-0.35
Panel C: Moving Average Job Finding/Separation Rate				
MA Separation Rate	0.95	1.01	1.21	1.32
MA Job Finding Rate	0.05	-0.01	-0.21	-0.32

Notes: Panel A is constructed using CPS data from 1976-2019. Panels B and C are constructed using CPS data from 1994-2019.

A.4.3 Regression Results

To assess the robustness of the patterns presented in graphs throughout the paper after controlling for standard observables, we estimate:

$$\begin{aligned} Y_{it} = {} & \beta_0 College_i + \beta_1 Potexp_{it} + \beta_2 Potexp_{it}^2 + \beta_3 College_i * Potexp_{it} + Race_i + \\ & MarStatus_{it} + Child_{it} + FamInc_{it} + \Phi_{Occ2} + \Phi_{Ind2} + \Phi_{State} + \Phi_{Year} + \Phi_{Month} + \epsilon_{it}. \end{aligned} \tag{A.27}$$

Our outcomes of interest, $Y_{i,t}$, include indicators for worker i in period t: (i) unemployed or not; (ii) transitions from unemployment to employment; (iii) transitions from employment to unemployment; (iv) transitions to a different occupation; (v) transitions to a different career; and (vi) the magnitude of skill distance in occupational transitions. Our primary variable of interest is $College_i$, which is an indicator for whether individual i has a college degree. The coefficient β_0 captures the association between a college degree and the outcome of interest, while β_3 indicates how this association varies over years of potential experience.

As shown in equation (A.27), we control for a quadratic in years of potential experience, race, marital status, whether the respondent has a child or not, and family income. In addition, we control for job characteristics by including 2-digit occupation and industry fixed effects. Finally, we incorporate year, month, and state fixed effects.

Table A7 indicates that college graduates have statistically significant lower probabilities of unemployment, job separation, occupational switching, career switching, while notable higher job finding probabilities relative to non-college counterparts. Furthermore, conditional on changing occupations, college graduates switch to occupations similar to their prior ones. Besides that, the education gap in each outcome dissipates with potential experience. Overall, these results align with the descriptive patterns.

Table A7: Regression Results in the CPS

	(1)	(2)	(3)	(4)
Panel A: Unemployed Indicator				
College	-0.03750***	-0.03762***	-0.03672***	-0.02638***
College × PotExp	0.00130***	0.00131***	0.00131***	0.00116***
Observations	16,531,741	16,531,741	16,531,741	13,097,696
R^2	0.034	0.036	0.042	0.064
Panel B: Job Finding Indicator				
College	0.01749***	0.02223***	0.02304***	0.01415***
College × PotExp	-0.00176***	-0.00194***	-0.00207***	-0.00198***
Observations	501,664	501,664	501,664	409,425
R^2	0.018	0.022	0.040	0.045
Panel C: Job Separation Indicator				
College	-0.01617***	-0.01608***	-0.01571***	-0.01392***
College × PotExp	0.00058***	0.00058***	0.00059***	0.00057***
Observations	10,083,104	10,083,104	10,083,104	8,145,221
R^2	0.013	0.014	0.015	0.017
Panel D: Occupational Mobility Indicator				
College	-0.01985***	-0.01981***	-0.02034***	-0.01927***
College × PotExp	0.00079***	0.00079***	0.00080***	0.00080***
Observations	852,249	852,249	852,249	801,775
R^2	0.007	0.008	0.008	0.010
Panel E: Career Mobility Indicator				
College	-0.01499***	-0.01505***	-0.01582***	-0.01501***
College × PotExp	0.00050***	0.00050***	0.00051***	0.00052***
Observations	827,086	827,086	827,086	778,243
R^2	0.007	0.007	0.008	0.009
Panel F: Angular Distance in Occupation Switches				
College	-2.88222***	-2.81939***	-2.80947***	-2.60987***
College × PotExp	-0.04694**	-0.04996**	-0.05127***	-0.04174**
Observations	28,940	28,940	28,940	26,537

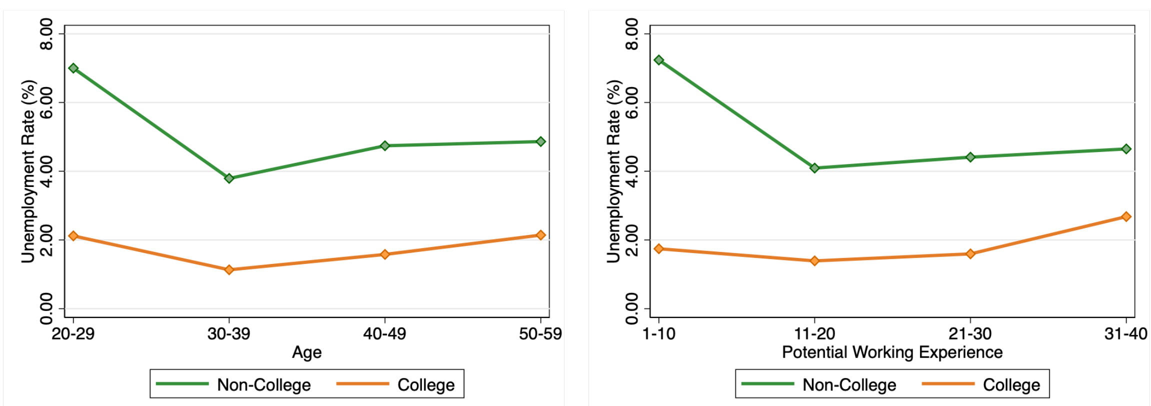
R^2	0.078	0.080	0.082	0.084
State FE		✓	✓	✓
Year FE			✓	✓
Month FE			✓	✓

Notes: All specifications control for industry and occupation fixed effects, where industry and occupational codes for the unemployed are classified according to their last job. Standard errors are robust to heteroskedasticity. The last column additionally controls for family income. Levels of statistical significance are indicated by *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

A.5 NLSY79 Patterns

A.5.1 Unemployment-Education Gap

Figure A12 displays the unemployment rate by age/potential experience and educational attainment in the NLSY79 sample. The overall patterns are consistent with the CPS, with the U-E gap narrowing as individuals age or gain potential experience. Notably, there is an increase in the unemployment rate in later career stages, which is reasonable given that around 85% (15%) of respondents were 40-49 (50-59) years old in 2008, at the onset of the Great Recession.



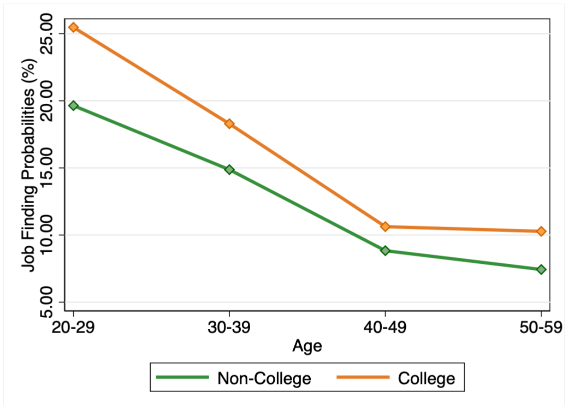
(a) Age (b) Potential Experience

Figure A12: Unemployment-Education Gap in the NLSY79. *Note*: Constructed using NLSY79 sample.

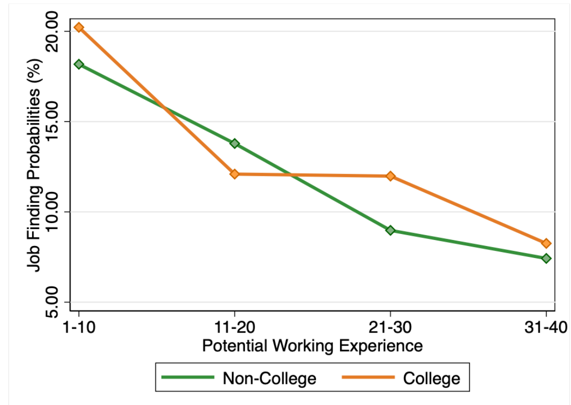
A.5.2 Job Finding and Separation Probabilities

Figure A13 presents the job finding and separation probabilities by age/potential experience and educational attainment. Concerning the job finding probabilities, there is no systematic difference among education groups, particularly over potential experience.

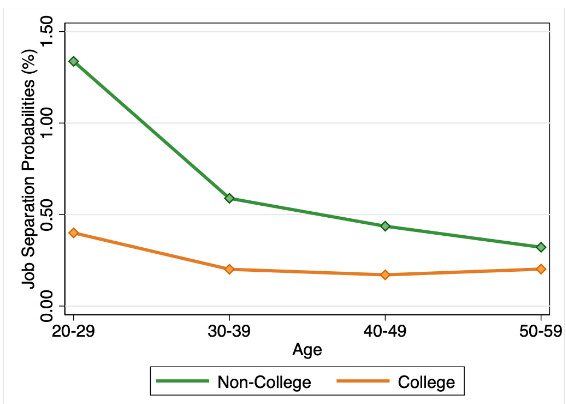
However, consistent with CPS patterns, college workers have systematically lower separation probabilities, with the gap widest in the early career stages and gradually narrowing with age or work experience.



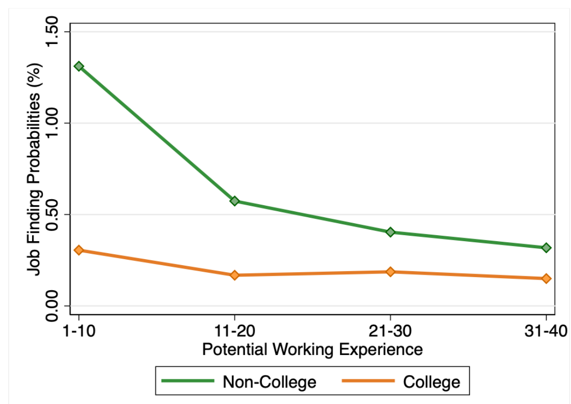
(a) JFP: Age



(b) JFP: Potential Experience



(c) SP: Age

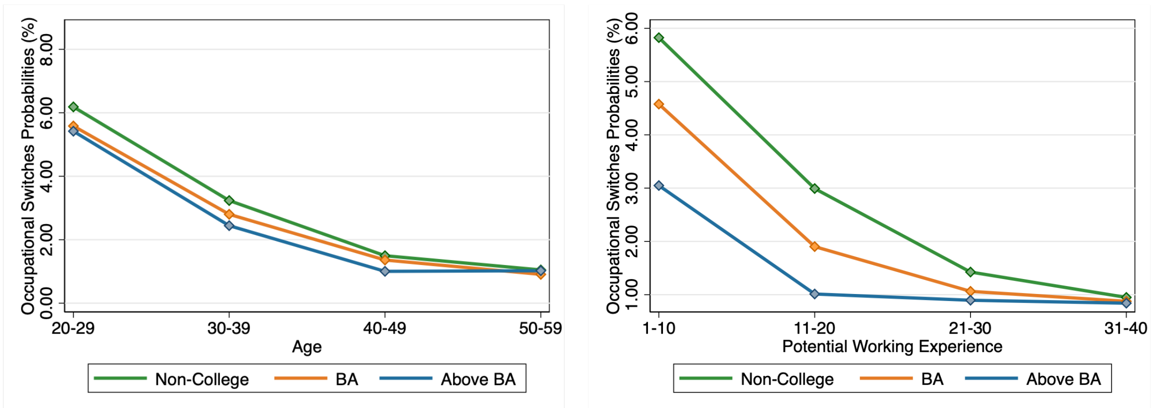


(d) SP: Potential Experience

Figure A13: NLSY79 Job Finding and Separation Probabilities. *Note*: Constructed using NLSY79 sample.

A.5.3 Occupation Mobility

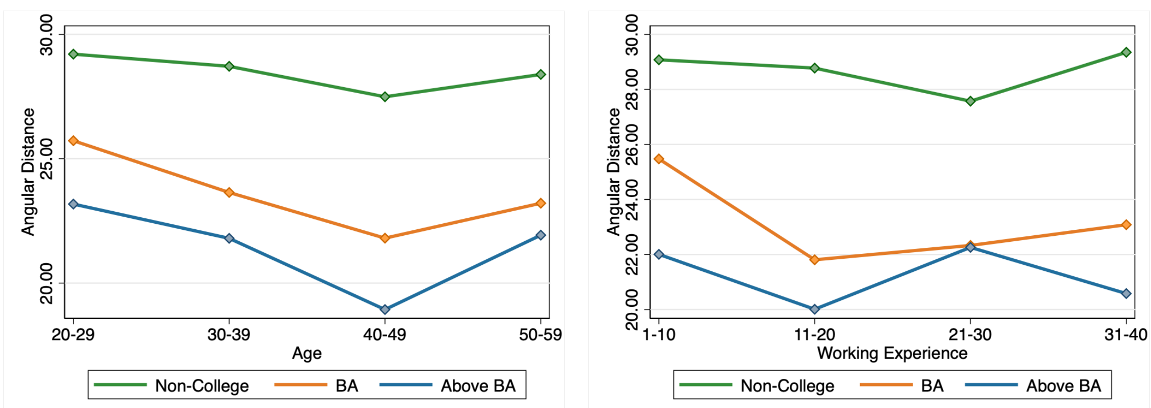
To measure occupational mobility in the NLSY79, we compute the fraction of workers within each age/potential experience and education subgroup who switch occupations between months $t-1$ and t, weighting each observation by the *PANELWEIGHT*. We restrict the sample to pairs of months with valid occupational codes. If the worker was non-employed in the previous month, we identify the occupation preceding the non-employment period. Figure A14 shows that occupational mobility patterns in the NLSY79 align with those in the CPS, with occupational mobility is decreasing with age/potential experience and educational attainment.



(a) Age (b) Potential Experience

Figure A14: Occupational Transitions. *Note*: Constructed using NLSY79 sample.

Figure A15 shows the average angular distance in occupation switches. Consistent with the trends observed in the CPS, higher educational attainment is associated with a lower angular distance at each career stage.



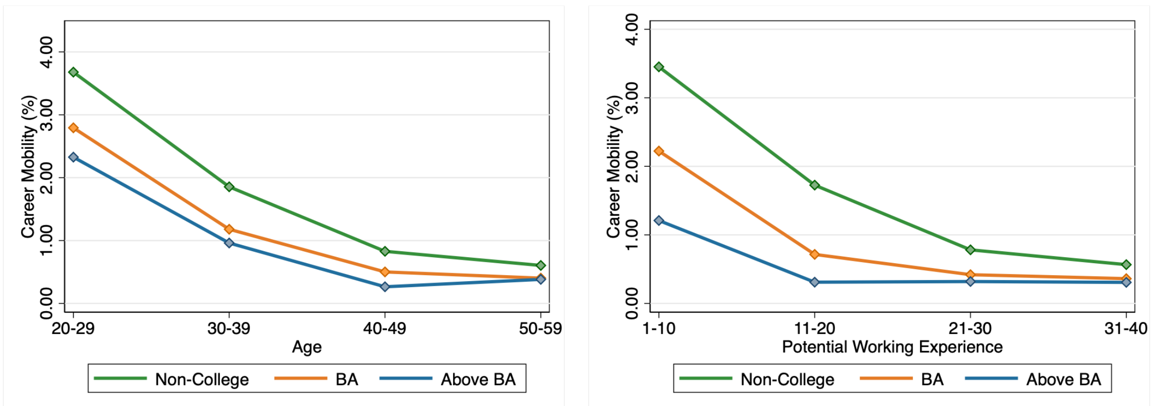
(a) Age (b) Potential Experience

Figure A15: Angular Distance in Occupational Transitions. *Note*: Constructed using NLSY79 sample.

A.5.4 Career Mobility

To measure career mobility in the NLSY79, we first identify a threshold, $\bar{\phi}$, for career transitions. That is determined by examining $37,084$ occupational transitions, where both skill requirements and task intensities are available for both the current and previous occupations. Next, we find that setting $\bar{\phi} = 23.08$ yields an unweighted average correlation of aptitudes $k \in \{verbal, math, social, manual, routine\}$ of approximately 0.00005. As such, a career switch is defined as an occupational transition where the angular distance exceeds the threshold, i.e., $\phi \geq 23.08$.

Figure A16 shows that, similar to occupational mobility, that career mobility in the NLSY79 decreases with both age/potential experience and educational attainment. Moreover, the gap in career mobility rates across education levels narrows over the life cycle.



(a) Age (b) Potential Experience

Figure A16: Career Switches. *Note*: Constructed using NLSY79 sample.

A.5.5 Robustness of NLSY79 Results

This section examines the robustness of the NLSY79 patterns after controlling for standard observables. The regression specification is the same as equation (A.27). Given that occupational inheritance may impact employment stability through parental networking ties, we additionally control for parental occupation, denoted by *ParentOcc*, which is measured in two ways: at the individual level, it equals one if the worker has ever held a job similar to a parent's, and at the observation level, it equals one if the worker's current job matches a parent's. The outcomes of interest remain the same as in the CPS regressions, with an additional focus on skill mismatch, as detailed in Section A.3.3.

Table A8 presents the estimated coefficients for college and the interaction term between college and potential experience. We can see that after controlling for observables, having a college degree is still associated with significantly lower probabilities of being unemployed, separating from employment, switching occupations or careers. Moreover, college workers have less skill mismatch and, conditional on switching occupations, have a lower angular distance in the switch. The interaction terms between college and potential experience suggest that, in general, the education gap in our outcomes of interest tend to narrow with work experience.

Table A8: NLSY79 Regression Results

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Unemployed Indicator						
College	-0.01986***	-0.02005***	-0.01296***	-0.00279***	-0.01298***	-0.01294***
College × PotExp	0.00062***	0.00065***	0.00012***	-0.00011**	0.00012***	0.00011**
Observations	1,197,087	1,187,574	1,187,574	1,003,608	1,187,574	1,183,448
R^2	0.051	0.051	0.057	0.118	0.057	0.056
Panel B: Job Finding Indicator						
College	0.01004	0.00742	0.02036	0.03187*	0.02179	0.02339
College × PotExp	-0.00033	-0.00018	-0.00028	-0.00134	-0.00030	-0.00039
Observations	63,714	63,042	63,042	47,568	63,042	61,634
R^2	0.043	0.044	0.055	0.070	0.055	0.054
Panel C: Job Separation Indicator						
College	-0.00927***	-0.00931***	-0.00585***	-0.00433***	-0.00585***	-0.00567***
College × PotExp	0.00044***	0.00044***	0.00027***	0.00021***	0.00027***	0.00026***
Observations	1,129,938	1,121,132	1,121,132	953,230	1,121,132	1,118,434
R^2	0.009	0.009	0.010	0.013	0.010	0.010
Panel D: Occupational Mobility Indicator						
College	-0.01854***	-0.01840***	-0.00763***	-0.00561***	-0.00760***	-0.00763***
College × Potexp	0.00070***	0.00070***	0.00036***	0.00029***	0.00036***	0.00036***
Observations	1,120,216	1,111,460	1,111,460	945,932	1,111,460	1,111,460
R^2	0.017	0.017	0.024	0.026	0.024	0.024
Panel E: Career Mobility Indicator						
College	-0.01562***	-0.01563***	-0.00951***	-0.00815***	-0.00949***	-0.00952***
College × PotExp	0.00059***	0.00059***	0.00039***	0.00034***	0.00039***	0.00039***
Observations	1,120,210	1,111,454	1,111,454	945,928	1,111,454	1,111,454
R^2	0.014	0.014	0.017	0.018	0.017	0.017
Panel F: Angular Distance in Occupational Switches						
College	-2.89147***	-3.05840***	-3.16422***	-3.04611***	-3.16346***	-3.16341***
College × Potexp	-0.05265**	-0.04957**	-0.03879	-0.02232	-0.03864	-0.03860
Observations	36,687	36,444	36,444	29,754	36,444	36,444
R^2	0.196	0.196	0.199	0.203	0.199	0.199
Panel G: Skill Mismatch						
College	-0.01562***	-0.01563***	-0.00951***	-0.00815***	-0.19913***	-0.19922***
College × PotExp	0.00059***	0.00059***	0.00039***	0.00034***	-0.00002	0.00006
Observations	1,120,210	1,111,454	1,111,454	945,928	1,121,392	1,121,392

R^2	0.014	0.014	0.017	0.018	0.163	0.164
State FE		✓	✓	✓	✓	✓
Year FE			✓	✓	✓	✓
Month FE			✓	✓	✓	✓

Notes: All specifications control for industry and occupation fixed effects, where industry and occupation codes for the unemployed are classified according to their last job. Standard errors are robust to heteroskedasticity. The fourth column controls for family income, while the last two columns additionally control for parents' occupation. Levels of statistical significance are indicated by *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

A.6 Experience and Match Survival

Following Bover et al. (2002), we estimate the association between prior experience and match survival by estimating:

$$\begin{aligned}\text{Survival}_{it} = \sum_{n=2}^{120} \mathbb{1}(\text{Dur}_{it} = n) + \beta_1 \text{Exp}_{it} + \beta_2 \log(\text{Dur}_{it}) * \text{Exp}_{it} + \beta_3 \text{Exp}_{it} * \text{College}_i \\ + \beta_4 \log(\text{Dur}_{it}) * \text{College}_i + \beta_5 \log(\text{Dur}_{it}) * \text{White}_i + \text{College}_i + \text{Age}_{it} \\ + \text{White}_i + \Phi_{Year} + \Phi_{Season} + \Phi_{Ind} + \epsilon_{it},\end{aligned}$$

where Survival_{it} is an indicator for whether the match survives into the subsequent period. We flexibly capture the duration dependence in the survival probability by introducing an additive dummy variable corresponding to each monthly duration. The primary explanatory variables include the amount of experience the worker had accumulated at the formation of the match, Exp_{it}, and its interaction with education attainment, $\text{Exp}_{it} * \text{College}_i$.[45]

Table A9 shows that prior experience is associated with a higher survival probability, and that this effect dissipates with tenure. In addition, $\beta_3 < 0$ suggests that the association between experience and the survival probability is weaker for college workers.

A.7 Unemployable Workers

The higher unemployment rate and separation probability for non-college workers might be driven by a group of "unemployable" workers, i.e. workers who experience an unusually high number of separations. We define unemployable workers as those with at least four EU transitions within the first ten years of their career, representing the 90^{th} percentile of EU transitions among non-college workers during this period.

[45] Exp_{it} is either a binary variable indicating if the prior working experience is longer than 76 months (the median prior working experience among 1,108,438 employment observations) or prior working experience in months.

Table A9: Prior Experience and Match Survival

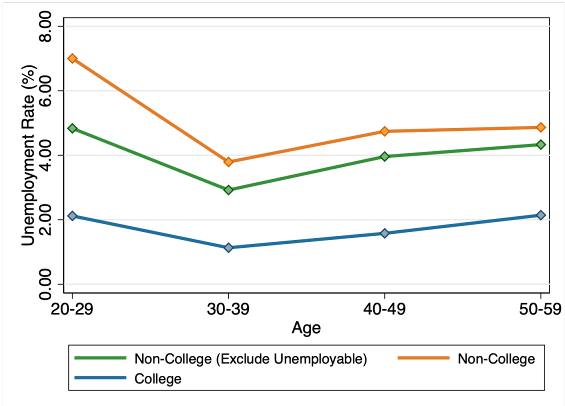
	(1)	(2)	(3)	(4)
Panel A: Experience > 76 Months Indicator				
Exp	0.02037***	0.00731***	0.01157***	0.00790***
Log(Dur) × Exp	-0.00295***	-0.00232***	-0.00282***	-0.00248***
Exp × College		-0.00498***	-0.00427***	-0.00755***
Log(Dur) × College		-0.00721***	-0.00561***	-0.00602***
Observations	1,105,229	1,105,229	1,055,676	484,382
R^2	0.019	0.024	0.019	0.022
Panel B: Months of Prior Experience				
Exp	0.00012***	0.00006***	0.00008***	0.00005***
Log(Dur) × Exp	-0.00002***	-0.00002***	-0.00002***	-0.00001***
Exp × College		-0.00004***	-0.00003***	-0.00004***
Log(Dur) × College		-0.00745***	-0.00574***	-0.00620***
Observations	1,105,229	1,105,229	1,055,676	484,382
R^2	0.019	0.024	0.019	0.022
Year FE		✓	✓	✓
Season FE		✓	✓	✓
1990dd Industry FE			✓	✓

Notes: The second and third specifications include the interaction between Log(Dur) and College, and White. 1990dd are industry fixed effects according to the industrial classification scheme compiled by Autor et al. (2019). Column (4) excludes matches formed through job-to-job transitions. Standard errors are robust to heteroskedasticity. Levels of statistical significance are indicated by *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

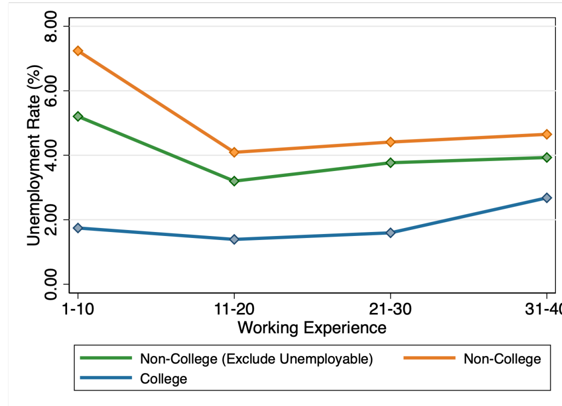
Figure A17 shows that removing the unemployable workers shifts down the unemployment and separation probabilities for non-college workers, while making little difference in the job finding probability. Overall, the gaps in unemployment and separations persist and narrow with age after excluding unemployable workers.

A.8 Sampled Jobs and Match Survival

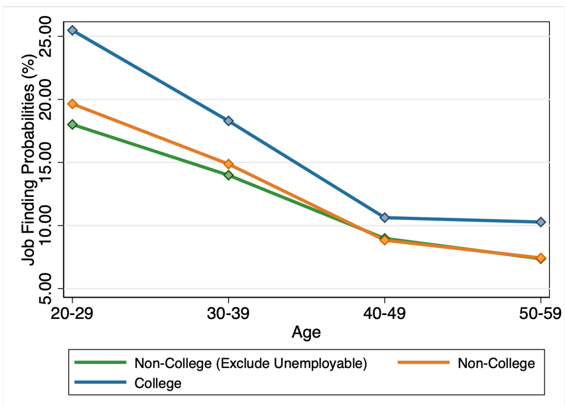
This section presents the complete estimation results for the association between sampled jobs and match survival, following specification (3). As shown in Tables A10 and A11, learning from prior working experience, whether through sampled occupations or careers, is always associated with a higher survival probability for the current match.



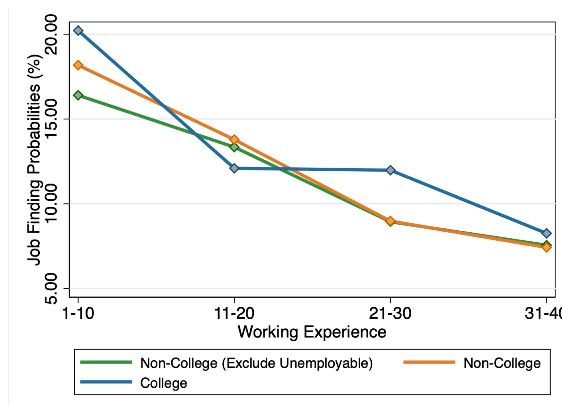
(a) Unemployment Rate



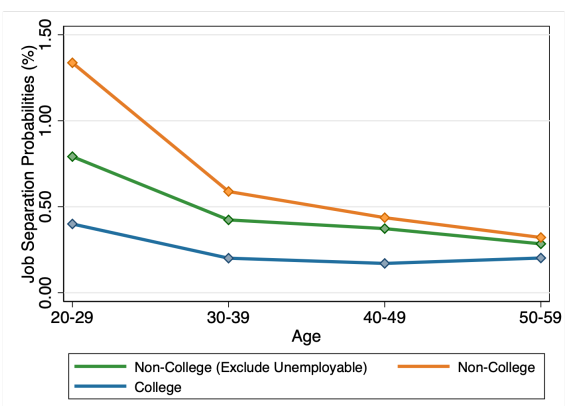
(b) Unemployment: Potential Experience



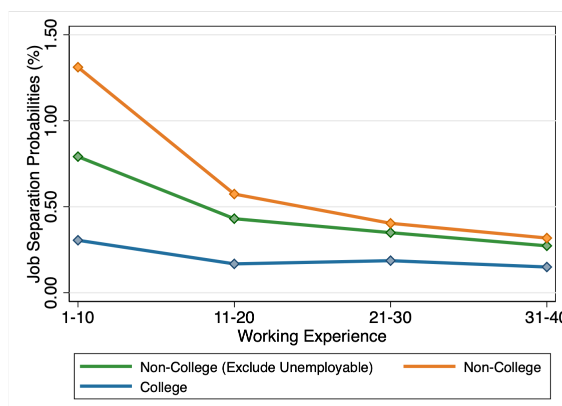
(c) Job Finding Probability



(d) Job Finding Probability: Potential Experience



(e) Separation Probability



(f) Separation Probability: Potential Experience

Figure A17: Life Cycle Patterns Excluding Unemployable Workers. *Notes:* An unemployable non-college worker are those who experience at least four transitions from employment to unemployment in their first ten years of potential experience. Constructed using NLSY79 sample.

Notably, this effect is more pronounced for non-college workers.

Table A10: Sampled Jobs and Survival Probability of Employer Matches

	Sampled Occupations			Sampled Careers	
	Non-College	College		Non-College	College
Sampled Occ.=1	0.00090	0.00105	Sampled Career=1	0.00524***	0.00243***
Sampled Occ.=2	0.00493***	0.00228***	Sampled Career=2	0.00861***	0.00420***
Sampled Occ.=3	0.00770***	0.00393***	Sampled Career=3	0.01199***	0.00533***
Sampled Occ.=4	0.01123***	0.00507***	Sampled Career=4	0.01366***	0.00564***
Sampled Occ.=5	0.01304***	0.00537***	Sampled Career=5	0.01546***	0.00666***
Sampled Occ.=6	0.01692***	0.00577***	Sampled Career=6	0.01717***	0.00795***
Sampled Occ.=7	0.01868***	0.00601***	Sampled Career=7	0.01820***	0.00707***
Sampled Occ.=8	0.02180***	0.00742***	Sampled Career=8	0.01998***	0.00835***
Sampled Occ.=9	0.02261***	0.00838***	Sampled Career=9	0.02130***	0.00920***
Sampled Occ.=10	0.02422***	0.00963***	Sampled Career=10	0.02180***	0.00810***
Sampled Occ.=11	0.02677***	0.01002***			
Sampled Occ.=12	0.02880***	0.01135***			
Sampled Occ.=13	0.03153***	0.01243***			
Sampled Occ.=14	0.03155***	0.01049***			
Sampled Occ.=15	0.03448***	0.01049***			
Tenure	0.00002***	0.00001***	Tenure	0.00001***	0.00001***
Exp	-0.00004***	-0.00002***	Exp	-0.00001***	-0.00001***
Observations	747,187	252,383	Observations	820,893	252,662
R^2	0.034	0.013	R^2	0.023	0.011

Notes: The dependent variable in both panels is an indicator for whether the worker remains employed at the same employer or becomes unemployed. The tenure variable is the worker's tenure with their current employer. All specifications control for individual fixed effects. Standard errors are robust to heteroskedasticity. Levels of statistical significance are indicated by *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Table A11: Sampled Jobs and Survival Probability of Occupation/Career Matches

	Sampled Occupations			Sampled Careers	
	Non-College	College		Non-College	College
Sampled Occ.=1	0.00390***	0.00139*	Sampled Career=1	0.00651***	0.00342***
Sampled Occ.=2	0.00569***	0.00253***	Sampled Career=2	0.01054***	0.00536***
Sampled Occ.=3	0.00896***	0.00434***	Sampled Career=3	0.01479***	0.00707***
Sampled Occ.=4	0.01121***	0.00512***	Sampled Career=4	0.01703***	0.00768***
Sampled Occ.=5	0.01327***	0.00568***	Sampled Career=5	0.01953***	0.00909***
Sampled Occ.=6	0.01660***	0.00589***	Sampled Career=6	0.02182***	0.01085***
Sampled Occ.=7	0.01769***	0.00601***	Sampled Career=7	0.02341***	0.01020***
Sampled Occ.=8	0.02061***	0.00723***	Sampled Career=8	0.02586***	0.01173***
Sampled Occ.=9	0.02146***	0.00827***	Sampled Career=9	0.02782***	0.01296***
Sampled Occ.=10	0.02283***	0.00940***	Sampled Career=10	0.02883***	0.01197***
Sampled Occ.=11	0.02538***	0.00986***			
Sampled Occ.=12	0.02699***	0.01104***			
Sampled Occ.=13	0.02943***	0.01213***			
Sampled Occ.=14	0.02939***	0.00994***			
Sampled Occ.=15	0.03147***	0.01125***			
Tenure	0.00003***	0.00001***	Tenure	0.00004***	0.00001***
Exp	-0.00004***	-0.00002***	Exp	-0.00003***	-0.00002***
Observations	713,170	243,672	Observations	817,210	252,781
R^2	0.029	0.012	R^2	0.023	0.012

Notes: The dependent variable in the left (right) panel is an indicator for whether the worker remains employed in the same occupation (career) or becomes unemployed. The tenure variable in the left (right) panel is the worker's cumulative experience in their current occupation (career). All specifications control for individual fixed effects. Standard errors are robust to heteroskedasticity. Levels of statistical significance are indicated by *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

Table A12: Observable Characteristics by FCE Validity at 35 years old and College Status

	Valid, C	Non-Valid, C	Valid, NC	Non-Valid, NC
Race (%)	11, 18, 71	8, 17, 75	21, 31, 49	15, 24, 61
% with at least one child	77.81	57.56	77.41	57.30
Ability: verbal	0.76	0.75	0.44	0.41
Ability: math	0.79	0.76	0.43	0.41
Ability: social	0.64	0.63	0.47	0.45
Skill required: verbal	0.73	0.72	0.44	0.40
Skill required: math	0.70	0.68	0.45	0.41
Skill required: social	0.74	0.73	0.44	0.41
Skill required: routine	0.40	0.40	0.54	0.55
Skill required: manual	0.34	0.38	0.56	0.56
Mismatch	1.57	1.55	1.69	1.68
Hourly wage	31.65	29.56	13.60	6.26
Unemployed probability	0.02	0.03	0.05	0.08

Notes: Those with a valid (non-valid) FCE are respondents who both listed an expected occupation at age 35 and had at least one observed occupation at age 35. "C" ("NC") stands for college (non-college). Race is the percentages who are Hispanic, Black, and non-Hispanic and non-Black, respectively. Ability, skill requirements, and mismatch measures are detailed in Appendix A.3.3.

A.9 More Details on Anticipated Occupations

As noted in the main text, a relatively small share of our original NLSY79 sample have valid forecast error (FCE) measures where we can compare their expected occupation to their realized occupation at 35 years old or 5 years after their initial interview. This is primarily because a smaller share of workers have no realized occupation at 35 years old or in 5 years from the initial interview, as nearly 86% of our original sample of 4,823 NLSY79 respondents do report their expected occupations. To help address concerns that the results we report in the main text are driven by selection issues, we compare respondents with valid and non-valid FCE measures along several observable characteristics in Tables A12 and A13. While there are some differences across those with a valid and non-valid FCE, these differences are relatively small and show that these groups are relatively similar. Moreover, as mentioned in Section 2.1, we can increase the number of respondents with a valid FCE measure by instead comparing their expected occupation at 35 years old to the weighted average of skill requirements across all occupations worked at between 30 and 40 years old from 2,565 to 3,260. Table 2 shows that our main results are robust to this alternative measure of forecast errors.

A concern regarding the differences in forecast errors across education attainments is that, irrespective of their understanding of their own comparative advantage, workers

Table A13: Observable Characteristics by FCE Validity in 5 years and College Status

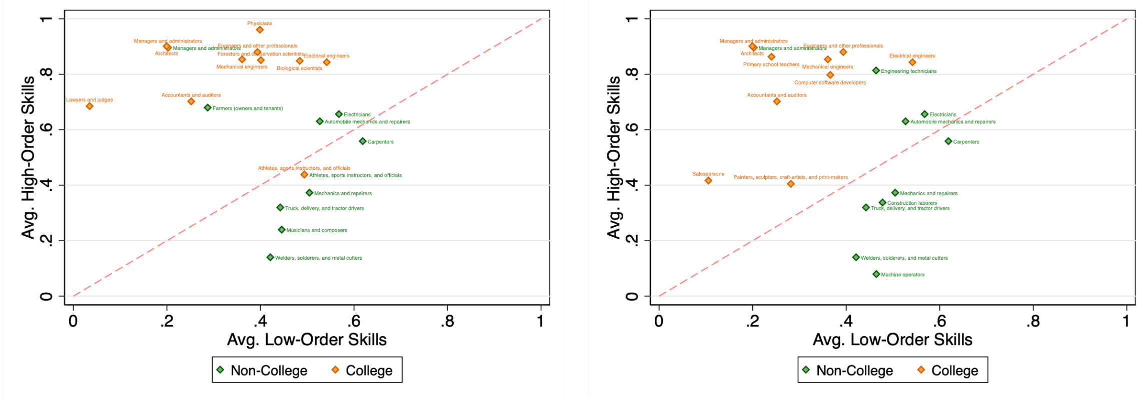
	Valid, C	Non-Valid, C	Valid, NC	Non-Valid, NC
Race (%)	7, 12, 81	10, 18, 72	18, 24, 58	19, 30, 52
% with at least one child	67	69	71	66
Ability: verbal	0.72	0.76	0.42	0.43
Ability: math	0.79	0.78	0.41	0.42
Ability: social	0.65	0.64	0.46	0.46
Skill required: verbal	0.69	0.73	0.41	0.44
Skill required: math	0.67	0.69	0.42	0.44
Skill required: social	0.71	0.74	0.40	0.44
Skill required: routine	0.42	0.40	0.56	0.54
Skill required: manual	0.33	0.36	0.58	0.55
Mismatch	1.65	1.54	1.67	1.70
Hourly wage	45.26	28.76	9.20	10.72
Unemployed probability	0.01	0.02	0.08	0.06

Notes: Those with a valid (non-valid) FCE are respondents who both listed an expected occupation in 5 years and had at least one observed occupation 5 years after their initial interview. "C" ("NC") stands for college (non-college). Race is the percentages who are Hispanic, Black, and non-Hispanic and non-Black, respectively. Ability, skill requirements, and mismatch measures are detailed in Appendix A.3.3.

may optimistically aspire to land in prestigious, well-regarded occupations. As a result, forecast errors tend to be larger for non-college workers, as they are less likely to secure jobs that typically require higher educational qualifications.

Figure A18 displays the most frequently anticipated occupations at age 35 or in 5 years, categorized by educational attainment. The x-axis shows the average low-order (routine and manual) skill requirements of these anticipated occupations, while the y-axis reflects the average high-order (verbal, math, and social) skill requirements. A clear distinction emerges between the expectations of college and non-college workers. College workers, for example, tend to anticipate working in high-skill occupations by the time they are 35, such as lawyers, judges, physicians, electrical engineers, and biological scientists. In contrast, non-college workers are more likely to expect employment in occupations that emphasize low-order skills, such as automobile mechanics, repairers, truck drivers, and carpenters. This pattern holds when looking at their anticipated occupations in 5 years as well. These observations indicate that differences in forecast errors by education are not driven by common occupational aspirations.

Beyond that, college workers tend to anticipate occupations with more dispersion in skill requirements. As shown in Table A14, the degree of skill dispersion in expected occupations for college workers is consistently higher than that for non-college workers, regardless of whether the expectations are short-term or long-term. This aligns with the



(a) At Age 35

(b) In 5 Years

Figure A18: Most Frequent Anticipated Occupations. *Note*: Constructed using NLSY79 sample.

Table A14: Skill Dispersion in Expected Occupation

	Variance	Max-Min	Mean Deviation	Median Deviation
Panel A: Expected Occupation at Age 35				
Non-College	0.0620	0.6361	0.2112	0.1878
College	0.0701	0.6549	0.2297	0.2062
Panel B: Expected Occupation in 5 Years				
Non-College	0.0569	0.6148	0.1996	0.1774
College	0.0725	0.6804	0.2327	0.2068

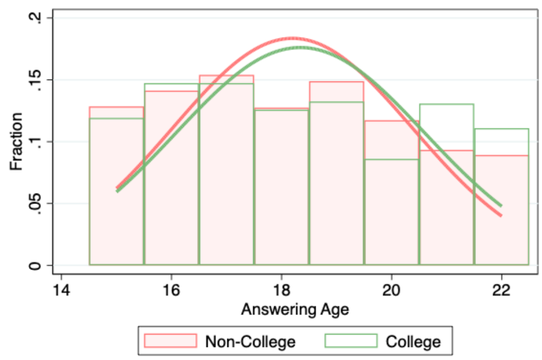
Notes: Data from NLSY79, including 1,961 (604) non-college (college) workers. Let Diff represent the difference in skill dispersion between non-college and college respondents. The p values for a t-test of the null hypothesis H_0 : Diff $=0$ versus the alternative hypothesis H_a : Diff <0 are all less than 0.01. Formulas for each measure of dispersion are provided in Appendix A.3.4.

intuition that workers would anticipate an occupation that has balanced skill requirements if they are not sure of their own skills. On the other hand, an individual who knows that they have, e.g., high math skills may anticipate working in an occupation with relatively high math skill requirements.

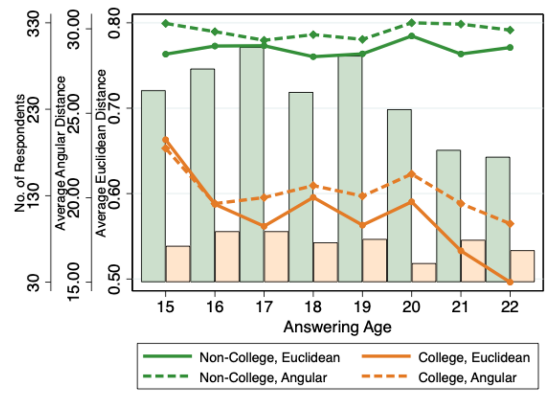
A.9.1 Decomposition of the Euclidean Distance

Let ψ denote the Euclidean distance between two vectors, $\mathbf{s}_i$ and $\hat{\mathbf{s}}_i$. From the Law of cosines,

$$\|\mathbf{s}_i\|^2 + \|\hat{\mathbf{s}}_i\|^2 - 2\|\mathbf{s}_i\|\|\hat{\mathbf{s}}_i\|\cos(\phi) = \psi^2. \tag{A.28}$$



(a) Answering Age



(b) FCE by Education and Answering Age

Figure A19: Forecast Error and Answering Age. *Note*: Constructed using NLSY79 sample.

Adding and subtracting $2\|\mathbf{s}_i\|\|\hat{\mathbf{s}}_i\|$ to the left-hand side and dividing by ψ^2 gives:

$$\underbrace{\frac{(\|\mathbf{s}_i\| - \|\hat{\mathbf{s}}_i\|)^2}{\psi^2}}_{\text{Diff. in Skill Magnitude}} + \underbrace{\frac{2\|\mathbf{s}_i\|\|\hat{\mathbf{s}}_i\|(1-\cos(\phi))}{\psi^2}}_{\text{Diff. in Cosine Similarity}} = 1. \tag{A.29}$$

From (A.29), the first term is the contribution of the difference in the norms of the two vectors to the Euclidean distance, while the second is driven by the angular distance.

A.9.2 Forecast Error by Age and Educational Attainment

As the occupational expectation questions are asked upon respondents' entry into the survey, the difference in forecast errors by educational attainment may be biased if respondents who eventually obtained a college degree were, on average, older when they recorded their expected occupation. Figure A19(a) shows there is no systematic difference in the age at which expectations were recorded across the different education levels. Moreover, Figure A19(b) shows that the gap in forecast errors is present at each age. These findings suggest that the difference in forecast errors by education is not driven by differences in the ages at which occupational expectations were recorded.

A.10 Associate's Degrees and College Dropouts

Associate's degree provides specialized technical or vocational training aimed at equipping individuals with a specific skill set or preparing them for particular careers, similar to four-year college degree in offering workers greater certainty about their comparative

Table A15: Associate's Degrees and College Dropouts

	Unemployment Rate (%)	Separation Prob. (%)
Panel A: Associate's Degree, CPS		
Non-College	[10.13, 6.51, 5.56, 5.18]	[2.72, 1.78, 1.45, 1.21]
Associate's Degree	[5.57, 3.93, 3.72, 4.08]	[1.48, 0.98, 0.92, 0.92]
College	[4.34, 2.33, 2.42, 2.88]	[0.99, 0.56, 0.53, 0.59]
Panel B: College Dropouts, NLSY79		
College Dropouts	[4.80, 3.03, 3.82, 3.55]	[0.96, 0.48, 0.40, 0.31]
Less-Educated Dropouts	[6.03, 4.23, 4.51, 3.59]	[1.20, 0.57, 0.46, 0.28]
More-Educated Dropouts	[3.70, 2.30, 3.45, 3.54]	[0.75, 0.43, 0.37, 0.32]
College Graduates	[2.12, 1.13, 1.58, 2.14]	[0.40, 0.20, 0.17, 0.20]

Notes: "Prob." refers to probability. Panel A is constructed using the CPS sample between 1992-2019. Panel B is constructed using the NLSY79 sample.

advantage, albeit with a shorter exploration period.

In Panel A in Table A15, we compare employment stability across three groups: non-college workers without an associate's degree, associate's degree holders, and college workers in the CPS (the patterns are very similar in the NLSY79). Notably, the separation probability among AA graduates is significantly lower than those without an AA while slightly higher than four year degree holders.

We also examine the employment profile for college dropouts by comparing "less" and "more" educated dropouts to college graduates, where the latter group of dropouts have completed at least two years of college and account for nearly 60% of the 810 college dropouts. Using the NLSY79 sample, we define dropouts as those who enrolled full-time in college but did not attain a Bachelor's degree or higher, yielding a dropout rate of 57.65%, closely aligned with the 54% reported by Vardishvili (2024). Furthermore, we exclude 15 respondents who report "lack of ability or poor grades" and 4 respondents who report being "expelled or suspended" as their reasons for dropping out.

Panel B in Table A15 presents the unemployment rate and separation probability over the life cycle for college graduates and dropouts. College dropouts are more likely to be unemployed than graduates, and within the group of dropouts, more years of completed schooling is associated with a lower unemployment rate. Similarly, the job separation probability is, at each age bin, decreasing in years of college completed.

Table A16: Occ1990dd Occupation List

First-Level Code	First-Level Occupation Title	Second-Level Code	Second-Level Occupation Title
A.1	Executive, Administrative, and Managerial Occupations	A	Managerial and Professional Specialty Occupations
A.2	Management Related Occupations		
A.3	Professional Specialty Occupations		
B.1	Technicians and Related Support Occupations	B	Technical, Sales, and Administrative Support Occupations
B.2	Sales Occupations		
B.3	Administrative Support Occupations		
C.1	Housekeeping and Cleaning Occupations	C	Service Occupations
C.2	Protective Service Occupations		
C.3	Other Service Occupations		
D.1	Farm Operators and Managers	D	Farming, Forestry, and Fishing Occupations
D.2	Other Agricultural and Related Occupations		
E.1	Mechanics and Repairers	E	Precision Production, Craft, and Repair Occupations
E.2	Construction Trades		
E.3	Extractive Occupations		
E.4	Precision Production Occupations		
F.1	Machine Operators, Assemblers, and Inspectors	F	Operators, Fabricators, and Laborers
F.2	Transportation and Material Moving Occupations		

A.11 Occupation Codes

Table A16 displays the first- and second-level occupation categories following the occupation scheme developed by Autor and Dorn (2013) and is originally presented as part of Appendix Table 2 in Dorn (2009).

B Theory Appendix

B.1 Laws of Motion

Let $u_{a,e,i}$ denote the measure of unemployed workers of age a, education e, and history i who are unemployed at the beginning of the learning stage and are searching in a submarket for a career with which they have an unknown or good fit. Further, $\overline{u}_{a,e,i}$ denotes the measure of unemployed workers with a bad fit, $n_{e,a,i}$ the measure of workers employed in a career of unknown or good fit, and $\overline{n}_{e,a,i}$ the measure employed in matches that are a bad fit. A "+" superscript denotes the measures in the next time period.

The law of motion for young, unemployed workers in a career with an unsure or good fit is

$$u^+_{y,e,i} = \begin{cases} \mu\pi_e + (1-\lambda_o)\left[(1-f^*)u_{y,e,1} + (1-\phi_e)\delta^{un}_{1e} n_{y,e,1}\right] & \text{for} \quad i = 1, \\ (1-\lambda_o)\left[(1-f^*)(u_{y,e,i} + l^*\overline{u}_{y,e,i-1}) + (1-\phi_e)\delta^{un}_{ie} n_{y,e,i}\right] & \text{for} \quad i = 2, \dots, N_e - 1, \\ (1-\lambda_o)\left[\delta^g\left(\phi_e \sum_{i=1}^{N_e-1} p_{ie} n_{y,e,i} + n_{y,e,N_e}\right) + (1-f^*)(u_{y,e,N_e} + l^*\overline{u}_{y,e,N_e-1})\right] & \text{for} \quad i = N_e, \end{cases} \tag{B.1}$$

where f^* and l^* represent the job finding probability and decision to leave a career. For brevity, we suppress the subscript a, e, i on the policy functions. Starting with the first line of (B.1), the first term represents new entrants to the labor market, the second term are unemployed workers who do not find a job or become old, and the third term is employed workers who do not learn their fit, lose their job, and are not hit with an aging shock. As for the second line, the first term is unemployed workers with an unsure fit, including those who switched from a bad fit, who do not find a job. The second term is employed workers who do not learn their fit and lose their job. Each measure is multiplied by $1 - \lambda_o$, as these are the young workers who are not hit with an aging shock. Finally, in the third line, the first term represents all young workers who exited the previous period's learning stage knowing their best fit and were hit with a separation shock. The second term represents unemployed workers who do not find a job.

The law of motion for young, employed workers in a career with an unsure or good fit is

$$n^+_{y,e,i} = \begin{cases} (1-\lambda_o)\left[f^* u_{y,e,1} + (1-\phi_e)(1-\delta^{un}_{1e}) n_{y,e,1}\right] & \text{for} \quad i = 1, \\ (1-\lambda_o)\left[f^*(u_{y,e,i} + l^*\overline{u}_{y,e,i-1}) + (1-\phi_e)(1-\delta^{un}_{ie}) n_{y,e,i}\right] & \text{for} \quad i = 2, \dots, N_e - 1, \\ (1-\lambda_o)\left[(1-\delta^g)\left(\phi_e \sum_{i=1}^{N_e-1} p_{ie} n_{y,e,i} + n_{y,e,N_e}\right) + f^*(u_{y,e,N_e} + l^*\overline{u}_{y,e,N_e-1})\right] & \text{for} \quad i = N_e, \end{cases} \tag{B.2}$$

Equation (B.2) has a similar interpretation as (B.1), except that the measure of employed workers consists of unemployed workers who find a job and employed workers who do not lose their job.

Next, the laws of motion for young workers in a bad fit are given by:

$$\overline{u}^+_{y,e,i} = (1-\lambda_o)\left[\delta^b \bar{n}_{y,e,i} + (1-l^*)(1-\overline{f}^*)\overline{u}_{y,e,i} + \phi_e(1-p_{ie})d^* n_{y,e,i}\right], \tag{B.3}$$

$$\overline{n}^+_{y,e,i} = (1-\lambda_o)\left[(1-\delta^b)\bar{n}_{y,e,i} + (1-l^*)\overline{f}^*\overline{u}_{y,e,i} + \phi_e(1-p_{ie})(1-d^*)n_{y,e,i}\right], \tag{B.4}$$

for $i = 1, 2, \ldots, N_e - 1$ and where $\overline{f}^*$ is the job finding probability of the workers in submarkets for bad matches and d^* is the separation probability upon learning the worker is not in their true calling. The first term of (B.3) represents employed workers in a bad fit who lose their job. The second term is workers who are unemployed in a bit fit, do not leave their current career, and do not find a job. The last term captures workers who were employed in an unsure fit, learn that they are in a bad fit, and separate from the match. Equation (B.4) follows a similar intuition.

We now proceed to the laws of motion for old workers, and begin with those who are in a career with an unknown or a good fit:

$$u^+_{o,e,i} = \begin{cases} \sum_a \chi_a\left[(1-f^*)u_{a,e,1} + (1-\phi_e)\delta^{un}_{ie}n_{a,e,1}\right] & \text{for} \quad i = 1, \\ \sum_a \chi_a\left[(1-f^*)(u_{a,e,i} + l^*\overline{u}_{a,e,i-1}) + (1-\phi_e)\delta^{un}_{ie}n_{a,e,i}\right] & \text{for} \quad i = 2, \ldots, N_e - 1, \\ \sum_a \chi_a\left[\delta^g\left(\phi_e \sum_{i=1}^{N_e-1} p_{ie}n_{a,e,i} + n_{a,e,N_e}\right) + (1-f^*)(u_{a,e,N_e} + l^*\overline{u}_{a,e,N_e-1})\right] & \text{for} \quad i = N_e, \end{cases} \tag{B.5}$$

$$n^+_{o,e,i} = \begin{cases} \sum_a \chi_a\left[f^* u_{a,e,1} + (1-\phi_e)(1-\delta^{un}_{ie})n_{a,e,1}\right] & \text{for} \quad i = 1, \\ \sum_a \chi_a\left[f^*(u_{a,e,i} + l^*\overline{u}_{a,e,i-1}) + (1-\phi_e)(1-\delta^{un}_{ie})n_{a,e,i}\right] & \text{for} \quad i = 2, \ldots, N_e - 1, \\ \sum_a \chi_a\left[(1-\delta^g)\left(\phi_e \sum_{i=1}^{N_e-1} p_{ie}n_{a,e,i} + n_{a,e,N_e}\right) + f^*(u_{a,e,N_e} + l^*\overline{u}_{a,e,N_e-1})\right] & \text{for} \quad i = N_e, \end{cases} \tag{B.6}$$

where $a \in \{y, o\}$, $\chi_a = \lambda_o$ if $a = y$, and $\chi_a = 1 - \lambda_d$ if $a = o$. The components of (B.5)-(B.6) are very similar to (B.1)-(B.2), except that there are additional flows into the stocks of old workers from young workers who are hit with an aging shock.

Finally, the law of motion for old workers in a bad fit is

$$\overline{u}^+_{o,e,i} = \sum_a \chi_a\left[\delta^b \bar{n}_{a,e,i} + (1-l^*)(1-\overline{f}^*)\overline{u}_{a,e,i} + \phi_e(1-p_{ie})d^* n_{a,e,i}\right], \tag{B.7}$$

$$\overline{n}^+_{o,e,i} = \sum_a \chi_a\left[(1-\delta^b)\bar{n}_{a,e,i} + (1-l^*)\overline{f}^*\overline{u}_{a,e,i} + \phi_e(1-p_{ie})(1-d^*)n_{a,e,i}\right], \tag{B.8}$$

for $i = 1, 2, \ldots, N_e - 1$.

B.2 Wages

This section details how we compute the model's wages via Nash bargaining and constant renegotiation that are used in the quantitative analysis.

Consider a worker with an unknown fit, age $a = o$, education e, and history i. In the main text, firms post bilaterally efficient contracts that deliver a worker the value of employment, $W_{o,e,i} = x$. Let θ be the value of tightness in this submarket. The value of searching in this submarket is:

$$U_{o,e,i} = z + \beta(1 - \lambda_d)\{U_{o,e,i} + f(\theta)[x - U_{o,e,i}]\}. \tag{B.9}$$

The Bellman equation for a firm posting in this market satisfies the free entry condition:

$$-k_a + q(\theta)[V_{o,e,i} - x] = 0. \tag{B.10}$$

Rewriting the entry condition gives

$$x = V_{o,e,i} - \frac{k_a}{q(\theta)}. \tag{B.11}$$

Substituting (B.11) into (B.9) gives

$$U_{o,e,i} = z + \beta(1 - \lambda_d)\left\{U_{o,e,i} + \max_{\theta} f(\theta)\left[V_{o,e,i} - U_{o,e,i} - \frac{k_a}{q(\theta)}\right]\right\}, \tag{B.12}$$

where the first order condition with respect to θ is

$$f'(\theta)[V_{o,e,i} - U_{o,e,i}] = k_a. \tag{B.13}$$

Substituting (B.13) into (B.11) gives

$$x = U_{o,e,i} + \varepsilon(\theta)[V_{o,e,i} - U_{o,e,i}], \tag{B.14}$$

where $\varepsilon(\theta) = 1 - f'(\theta)\frac{\theta}{f(\theta)}$. From (B.14), workers receive the value of unemployment plus a share, $\varepsilon(\theta)$, of the match surplus.

Let us now denote $w_{o,e,i}$ as the worker's wage and the value employment as $W_{o,e,i}$. The Bellman equation for an employed worker is

$$\begin{aligned} W_{o,e,i} = w_{o,e,i} + \beta(1 - \lambda_d)\{\phi_e[p_{ie}(\delta^g U_{o,e,N_e} + (1 - \delta^g)W_{o,e,N_e}) + \\ (1 - p_{ie})(d^*_{o,e,i}\overline{U}_{o,e,i} + (1 - d^*_{o,e,i})\overline{W}_{o,e,i})] + (1 - \phi_e)[\delta^{un}_{ie}U_{o,e,i} + (1 - \delta^{un}_{ie})W_{o,e,i}]\}. \end{aligned} \tag{B.15}$$

The value of a filled job to the firm, $J_{o,e,i}$, satisfies

$$J_{o,e,i} = p_{ie}y_e + (1-p_{ie})(y_e-\alpha) - w_{o,e,i} + \beta(1-\lambda_d)\{\phi_e[p_{ie}(\delta^g * 0 + (1-\delta^g)J_{o,e,N_e})+ \\ (1-p_{ie})(d^*_{o,e,i} * 0 + (1-d^*_{o,e,i})\overline{J}_{o,e,i})] + (1-\phi_e)[\delta^{un}_{ie} * 0 + (1-\delta^{un}_{ie})J_{o,e,i}]\}, \quad \text{(B.16)}$$

where we have incorporated that the value of a vacancy is equal to zero through free entry.

Bilateral efficiency implies that $J_{o,e,i} > 0 \iff W_{o,e,i} > U_{o,e,i} \iff V_{o,e,i} > U_{o,e,i}$ where $V_{o,e,i} = W_{o,e,i} + J_{o,e,i}$. From Nash bargaining, we have the surplus sharing rule:

$$W_{o,e,i} - U_{o,e,i} = \varepsilon(\theta)[V_{o,e,i} - U_{o,e,i}]. \quad \text{(B.17)}$$

Substituting the surplus sharing rule into (B.15) and solving for $w_{o,e,i}$ gives

$$w_{o,e,i} = (1-\varepsilon(\theta))U_{o,e,i} + \varepsilon(\theta)[p_{ie}y_e + (1-p_{ie})(y_e-\alpha)] - \\ \beta(1-\lambda_d)(1-\varepsilon(\theta))\{\phi_e[p_{ie}U_{o,e,N_e} + (1-p_{ie})\overline{U}_{o,e,i}] + (1-\phi_e)U_{o,e,i}\}. \quad \text{(B.18)}$$

We can apply the same process to derive the following wages for old workers in a bad match:

$$\overline{w}_{o,e,i} = (1-\varepsilon(\theta))\overline{U}_{o,e,i} + \varepsilon(\theta)(y_e-\alpha) - \beta(1-\lambda_d)(1-\varepsilon(\theta))\overline{U}_{o,e,i}, \quad \text{(B.19)}$$

for young workers of education e and history i in an unsure/good match:

$$w_{y,e,i} = (1-\varepsilon(\theta))U_{y,e,i} + \varepsilon(\theta)[p_{ie}y_e + (1-p_{ie})(y_e-\alpha)] - \beta(1-\varepsilon(\theta))\sum_a \chi_a\{ \\ \phi_e[p_{ie}U_{a,e,N_e} + (1-p_{ie})\overline{U}_{a,e,i}] + (1-\phi_e)U_{a,e,i}\}, \quad \text{(B.20)}$$

where $\chi_a = 1-\lambda_o$ if $a = y$ and λ_o for $a = o$. And, finally, for young workers of education e and history i in a bad match:

$$\overline{w}_{y,e,i} = (1-\varepsilon(\theta))\overline{U}_{y,e,i} + \varepsilon(\theta)(y_e-\alpha) - \beta(1-\varepsilon(\theta))\{(1-\lambda_o)\overline{U}_{y,e,i} + \lambda_o\overline{U}_{o,e,i}\}. \quad \text{(B.21)}$$

Table C1: College Wage Premium

	(1)	(2)	(3)	(4)
College	0.37904***	0.29166***	0.37364***	0.27211***
1990dd Occupation FE		✓		✓
1990dd Industry FE			✓	✓
Observations	1111851	1085490	1069617	1068529
R^2	0.244	0.298	0.285	0.324

Notes: All specifications include the vector of individual controls, X, listed below equation (C.1). 1990dd occupation fixed effects are the second-level occupation codes constructed by Dorn (2009). 1990dd industry fixed effects are second-level industry codes according to the industrial classification scheme compiled by Autor et al. (2019). Levels of statistical significance are denoted by ***$(p < 0.01)$. Data are from the NLSY79.

C Quantitative Appendix

C.1 Counting Unique Careers

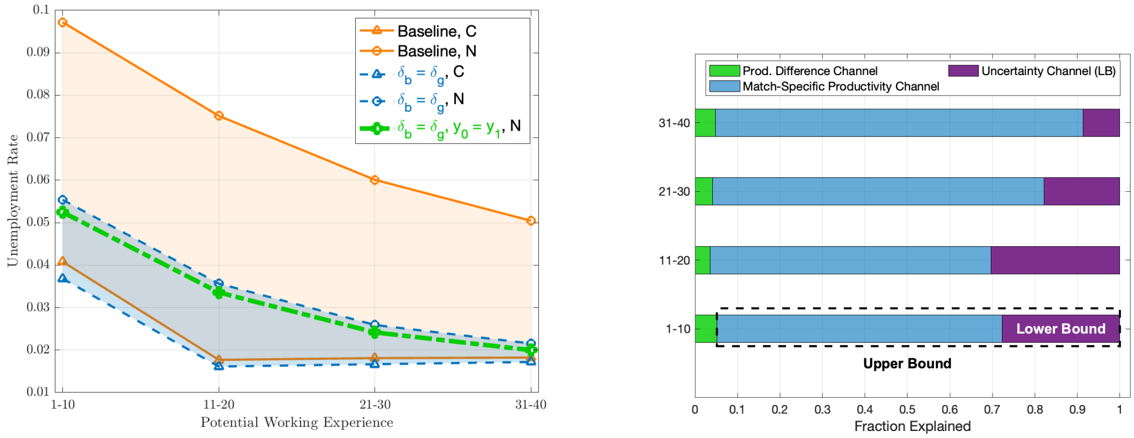
To count the number of unique careers each worker held over their lifetime using the NLSY79 data, we start with individuals with complete occupational information, including occupation codes, skill requirements, and task intensity. A unique career is defined as one where the angular distance between that career and all previous careers is greater than or equal to the threshold $\bar{\phi} = 23.08$. That is, career i is considered unique if its angular distance relative to any previously held career j, denoted by ϕ_{ij}, is greater than $\bar{\phi}$.

C.2 College Wage Premium

This section details the estimated college wage premium in our NLSY79 sample that appear in Table 7. We first start with our monthly NLSY79 panel and restrict to observations between ages 20 and 59. Further, we deflate an individual's reported hourly wage using the PCE2000 index. We then code real wage observations as missing if they fall below (above) 3 (200) dollars per hour. With the hourly wages in hand, we estimate variations of the following specification:

$$w_{it} = \beta_0 + \beta_1 \text{College}_i + \Gamma X_i + \beta_2 \text{Urate}_t + \Phi_{Occ2} + \Phi_{Ind2} + \epsilon_{it}, \tag{C.1}$$

where w_{it} is individual i's log real hourly wage in month t, College is an indicator for whether individual i obtained a BA or above, X is a vector of fixed individual characteristics (race and AFQT score), Urate is the yearly national unemployment rate, Φ_{Occ2} is an occupation fixed effect, and Φ_{Ind2} is an industry fixed effect. Table C1 presents the



(a) Graphical Illustration of Decomposition

(b) Decomposition Fraction

Figure C1: Alternative Decomposition. *Notes*: Panel (a) shows the unemployment rate profiles, by education, after simulating the model economy under the specified parameters. Panel (b) shows the corresponding fraction of the unemployment-education gap that closes after each successive change in the model's parameters.

estimates β_1 across different versions of (C.1).

C.3 Alternative Decomposition

In the decomposition presented in Section 4.3, we first shut down the differences in labor productivity and then eliminated the difference in separation probabilities between good and bad matches. In this section, we reverse the order by first equating δ^b with δ^g and then set y_0 equal to y_1. Proceeding in this way delivers a decomposition where the differences in productivity account for approximately 4.49% of the U-E gap, while the uncertainty channel explains between 24.26% and 95.51%, which is consistent with the contribution of each channel shown in the main text. This decomposition is presented in Figure C1.

C.4 STEM vs. Non-STEM

This appendix details the classification of college majors into STEM and Non-STEM categories, the computation of moments for each group used in our model calibration, and the subsequent calibration results.

C.4.1 Classification

The NLSY79 provides information on respondents' major field of study for their most recent college attended from survey year 1979 until 2018, with the exceptions of 1987 and

Table C2: College Majors by STEM and Non-STEM Classification

Category	Majors
STEM	Computer and Information Sciences; Mechanical Engineering; Electrical, Electronics, Communications Engineering; Biology; Chemical Engineering; Engineering; Civil, Construction & Transportation Engineering; Architecture; Chemistry; Mathematics; Agricultural Business; Engineering Technologies; Aerospace, Aeronautical, Astronautical Engineering; Industrial and Management Engineering; Information Sciences and Systems; Horticulture; Health Professions; Biochemistry; Geology; Natural Resources Management; Data Processing; Agriculture; Physics; Agricultural Economics; Rehabilitation; Architectural Engineering; Animal Science; Medical Specialties; Podiatry or Podiatric Medicine; Zoology; Agronomy; Dental Specialties; Pharmacy; Behavioral Science; Materials Engineering; Systems Analysis; Poultry Science; Pre-med; Environmental and Sanitary Engineering; Veterinary Medicine Specialties; Pre-dentistry; Aerospace Science (Air Force); Petroleum Engineering; Medical Laboratory Technologies; Neurosciences; Earth Sciences; Geological Engineering; Agriculture and Forestry Technologies; Biophysics; Biological Sciences; Forestry; Engineering Physics; Computer Programming; Mining and Mineral Engineering; Ecology; Mathematics; Microbiology
Non-STEM	Business Management and Administration; Accounting; Business and Commerce; Physical Education; Law Enforcement and Corrections and Criminology and Criminal Justice; Political Science and Government; Banking and Finance; Marketing and Purchasing; Communications; Economics; Psychology; History; Music (Performing, Composition, Theory); Applied Design and Graphic Design and Fashion Design; Social Work and Helping Services; Radio - Television; Law; English; Journalism; Hotel and Restaurant Management; Education; Theological Professions; Dramatic Arts; Music (Liberal Arts Program); Nursing; Sociology; Public Administration; Music Education; Personnel Management; Geography; Physical Therapy; Public Relations; Criminology; Religious Studies; Landscape Architecture; Fine Arts; Photography; Anthropology; Industrial Arts, Vocational & Technical Education; Elementary Education; Art; Labor and Industrial Relations; Special Education; General Liberal Arts and Sciences; Religious Education; Spanish; Commercial Art; Parks and Recreation Management; Hospital and Health Care Administration; Educational Administration; Cinematography; Literature, English; Advertising; Creative Writing; Chinese; Transportation and Public Utilities; Foreign Languages Education; International Relations; Business and Management; Pre-law; Merchant Marine; Education of the Deaf; Recreation, Outdoor Recreation; Classics; Comparative Literature; Optometry; Chiropractic; Organizational Behavior; Art History and Appreciation; Social Foundations; Speech Pathology and Audiology; Clinical Psychology; Philosophy; Secondary Education; Psychology for Counseling; English Education, Other

Note: The classification of STEM and Non-STEM follows the Department of Homeland Security's list of STEM designated degree programs.

1991. In our restricted sample, we observe 1,061 college workers with at least a bachelor's degree who reported 140 distinct second-level college majors and 22 distinct first-level college majors.[46] The primary college major for each graduate is identified as the one

[46] A comprehensive list of these majors can be found in the NLSY79 attachment 4: https://nlsinfo.org/content/cohorts/nlsy79/other-documentation/codebook-supplement/nlsy79-attachment-4-fields-study.

Table C3: Uncertainty Channel Parameters for STEM and Non-STEM

	Careers (data)	Careers (model)	Separation prob. (data)	Separation prob. (model)	ϕ	N
Non-STEM	2.14	1.92	[1.36, 0.72, 0.72, 0.72]	[1.41, 0.66, 0.64, 0.63]	0.1104	3
STEM	1.81	1.42	[1.04, 0.60, 0.60, 0.60]	[0.99, 0.64, 0.63, 0.62]	0.0791	2

Notes: Careers is the number of unique careers and the targeted value is obtained from the NLSY79. Targets for separation probabilities come from the CPS. All other parameters are fixed at the values presented in Table 6. In particular, STEM and Non-STEM have the same value of y_1 from Table 6.

reported most frequently across these survey years.

Based on the Department of Homeland Security STEM Designated Degree Program List, we have re-categorized these majors into 60 STEM and 80 Non-STEM categories. Among the 1,061 college male workers in our sample, 642 are classified as having a Non-STEM college major, while 419 have a STEM college major. The complete list of college major titles, categorized by STEM and Non-STEM, is provided in Table C2.

C.4.2 Computation of STEM and Non-STEM Moments

The classification of majors into STEM and Non-STEM categories allows us to compute specific moments for each group, which are crucial for our model calibration.

Following the calculation approach for the baseline target moments described in Section 4.1, we compute the job separation profiles over four potential work experience bins for both STEM and Non-STEM groups. Given that our model does not explicitly account for prompt surges in unemployment rates or job separation probabilities for college workers observed in the NLSY79 at the onset of the Great Recession, we fit the calculated separation rates for each career stage for each education group. Subsequently, we scale up the fitted job separation profile by a factor of 4. This scaling factor was determined by comparing the job separation rates of college workers across age bins in the NLSY79 and CPS samples, ensuring our fitted profiles align more closely with the average job separation profiles observed in the CPS sample, as our other non-uncertainty parameters are determined by fitting CPS data moments. The resulting target fitted separation profiles are listed in Table C3.

For the number of sampled careers, the calculation methodology remains consistent with our baseline calibration. On average, college graduates with a STEM college major sampled 1.81 unique careers throughout their observed life, while those with a Non-STEM college major sampled 2.14 unique careers.

Table C4: Education Policy Analysis (Non-STEM vs. STEM)

	Non-STEM	$\phi_{NS} = \phi_S$	$N_{NS} = N_S$	STEM
Unemployment rate (%)	2.84	2.98	2.17	2.25
Separation probability (%)	0.95	1.00	0.74	0.77
Labor productivity	2.96	2.95	2.98	2.97
Wages	2.91	2.90	2.94	2.93
Lifetime earnings	766.04	761.56	778.81	775.91
Lifetime discounted utility	683.92	680.27	693.30	691.34

Notes: The "Non-STEM" column presents baseline outcomes for Non-STEM workers. The second (third) column presents outcomes for Non-STEM workers after setting ϕ_{NS} (N_{NS}) equal to ϕ_S (N_S). The last column shows outcomes for BA workers with both $\phi_{NS} = \phi_S$ and $N_{NS} = N_S$. Life discounted utility is the value of a new entrant to the labor market, i.e. $U_{y,NS,1}$. Each reported outcome is the average across workers in the simulated economy.

C.4.3 Quantitative Policy Analysis

This section carries out a partial calibration of the model to match the number of unique careers sampled and separation profiles for Non-STEM and STEM majors in the NLSY79. Table C3 presents the targeted moments, model fit, and estimated learning speed and career set size for each group. As STEM majors have lower number of sampled careers, we find $N_S = 2 < 3 = N_{NS}$. Further, as Non-STEM majors exhibit a steeper decline in separation rates, we estimate a higher learning speed of $\phi_{NS} = 0.1104 > 0.0791 = \phi_S$.

Table C4 demonstrates the effect of assigning each of the STEM group's uncertainty channel parameter values, $\{\phi_S, N_S\}$, to the Non-STEM group. Just as in Section 4.4 where we compared the BA to above BA group, reducing the uncertainty of the Non-STEM group by narrowing their career set from 3 to 2 causes a modest reduction in unemployment and has a marginal impact on wages, earnings, and lifetime discounted utility.

D NLSY79 Panel Construction

Weekly Panel This section details the construction of the weekly panel from the NLSY79. The process involves three steps: (i) cleaning the employer history and determining employer characteristics; (ii) identifying demographic variables in each survey year; and (iii) identifying the primary job for each week if employed by multiple employers.

We start by processing the employer history roster, which involves two steps. The first is to standardize the occupation and industry codes across various census classification schemes to the 1990dd scheme developed by David Dorn.[47] This scheme consolidates US Census codes into a balanced panel of occupations or industries for the 2000 and 2002 Census, and also an unbalanced panel for the 1970 Census. When occupation and industry codes do not have corresponding 1990dd codes in the crosswalk file, we review the classification files and manually assign the closest equivalent within the 1990dd classification scheme.

In particular, for occupation codes (for civilian jobs, CPS jobs and the job at last employer) spanning survey year 1979 (round 1) to 2000 (round 19), we convert the original 1970 census occupational codes to the 1990dd classification scheme. For employer characteristics in the survey year 2002 (round 20), we convert the original 3-digit 2000 census occupation codes to the 1990dd classification scheme. For occupation codes from survey year 2004 onwards, we convert the original 4-digit 2002 occupation codes into 3-digit 2000 census codes by taking the first three digits, and then covert them to the 1990dd classification scheme. The crosswalk process for industry codes is similar to that for occupation codes, with one key difference: for industry codes reported from survey years 1979 (round 1) to 2000 (round 19), we first convert IND70 codes to IND80 codes before mapping them to the 1990dd industry classification scheme.

We then identify the employer characteristics for each job in every survey year by referring to the original employer history roster (EHR). When the EHR lacks occupational and industry codes, we supplement this with the corresponding codes from Current Population Survey (CPS) jobs. While the CPS employer is typically the first employer, this is not always the case during the survey years 1980 to 1992. To address job order discrepancies, we refer to the question: ''IS JOB # SAME AS CURRENT JOB? '' If the answer is affirmative, we use the CPS job information to fill in the missing data. We also use industry and occupation codes from the last employer to complete any remaining gaps.

Now, turning to the weekly employment histories with primary job codes, these codes follow the format $\textit{Survey Round} * 100 + \textit{Job Number}$. We first determine the survey round for each reported job, which corresponds to the first one or two digits of the job code. Using the unique respondent ID, survey round, and job number, we merge the weekly employment history with the employer history roster to obtain the employer characteristics for the reported job. Next, through cross-referencing *EMP_NUM_ARRAY* with the job number in the work history array, we can ascertain the current employer is the *x-th* employer the worker has worked for.

Next, we process the demographic variables. Since we already have the demographic characteristics for each respondent in each survey year, we need to align these with the

[47] See https://www.ddorn.net/data.htm for more details.

weekly employment history. To do so, we need to determine the survey year associated with each weekly observation using the available survey dates. For surveys conducted up to 1994, only the survey month is reported, so we need to impute the survey year. The identification process is as follows: we first determine the continuous week corresponding to each survey date. Then, for each weekly observation, we check if its week number falls within the range between the survey date of the most recent preceding survey round (not inclusive) and the current survey round (inclusive).[48] If it does, we assign the survey year of the current round to the observation. Once the survey year is identified, we can pinpoint various demographic characteristics such as race, gender, birth year (or age), marital status, childbearing, residential region, highest grade completed, (imputed) graduation year, enrollment status, ASVAB scores, and non-cognitive test scores (including the Rotter Locus of Control Score and Rosenberg Self-Esteem Scale).

Finally, we identify the primary job for each week, which is defined as the job with the most working hours during that week.[49] If multiple jobs have the same working hours per week, we keep the job reported in the main array.

Monthly Panel Next, we outline the process for converting a weekly panel to a monthly panel. To begin, we determine the calendar year and calendar month for each continuous week using the time crosswalk file. We then determine the primary labor force status of each respondent in each month. If the respondent is employed during a particular month, the primary job is determined as the one with the most working hours within that month. If multiple civilian jobs with the same total working hours, we consider the job with complete occupation and industry records as the primary job. If there are several jobs with complete records, we retain the one with known employer ID as the primary job. If there still exist multiple civilian jobs, we keep the earliest reported one, indicated by a lower job code in the weekly array.

If the respondent does not hold any job with assigned job codes for a given month, we prioritize the remaining labor force statuses following the precedence order adopted by the NLSY79: 3 (employed, but periods not working with an employer are missing) >

[48] An important characteristic of the NLSY surveys is that, with a few exceptions, each respondent in a survey round may have a distinct reference period. Specifically, the reference period is defined as the time between the date of the last interview and the date of the current interview. If a respondent participates in consecutive rounds, they report on events since their last interview date. Even if a respondent misses one or more interviews, they are still asked to report events since their last interview. This approach ensures that the entire time between a respondent's most recent and current interviews is recorded.

[49] In the case where a respondent simultaneously holds multiple jobs, the job number assigned to the main array is determined based on the start date of the job with the lowest job number. This selection is not influenced by any specific attributes of the job, such as the number of hours worked.

4 (unemployed) > 5 (out of the labor force) > 2 (period not working with an employer, unsure if unemployed or out of the labor force) > 7 (military) > 0 (no information). The status with higher precedence is regarded as the primary labor force status for that month.