WORK HISTORIES AND LIFETIME UNEMPLOYMENT[*]

BY IACOPO MORCHIO[1]

University of Vienna, Austria

A long-standing question in economics is how important unobserved differences across workers are for explaining unemployment. I revisit this topic using variation in lifetime unemployment across workers in U.S. data. A comparison of workers often unemployed with the rest shows that although differences in job-finding rates increase over the course of a career, differences in job-separation rates are large right from the start. I develop a directed search model with symmetric unobserved heterogeneity, in which agents learn workers' types from their labor market histories, to rationalize these findings. The model cannot match the data if unobserved heterogeneity is neglected.

1. INTRODUCTION

Economists have long sought to understand the determinants of unemployment. In particular, the degree to which unemployment is due more to adverse shocks or to unobserved differences in individual characteristics is crucial in the design of optimal labor market policy. For instance, if some group of individuals is always more prone to be unemployed than others, then changes in unemployment insurance policy will primarily affect that group. One way to understand the importance of unobserved heterogeneity and shocks is to study who the most frequently unemployed workers are over a long time span, which naturally maps into the study of lifetime unemployment. However, heterogeneity in lifetime unemployment has received little attention in the literature. In contrast, the study of lifetime income has shed considerable light on similar questions.[2]

The contribution of this article is to examine variation in lifetime unemployment and to provide a model to explain such variation. Using data from the 1979 National Longitudinal Survey of Youth (hereafter NLSY/79), I first extend standard measures of income dispersion to the study of lifetime unemployment. I then describe the concentration of lifetime unemployment in certain groups, and how these groups differ in the probabilities of finding and losing a job over the life cycle. Finally, I develop a directed search model that rationalizes these patterns.

I document three novel facts on lifetime unemployment. First, the experience of unemployment is concentrated in a relatively small group of workers. Two-thirds of prime-age unemployment in the NLSY/79 cohort are accounted for by 10% of workers. This also

---

---

holds within education/gender subgroups. Such concentration is mainly due to the higher job-separation rate of the most unemployed workers. Second, I find that time spent in unemployment when a worker is young is a powerful predictor of time spent in unemployment during prime age. Moreover, it has more predictive power than education, occupation, or health. Third, I uncover specific patterns in the probability of finding and losing a job over the life cycle across prime-age unemployment groups. With respect to job-finding rates, differences appear only in the long run. In other words, workers in different prime-age unemployment groups start their careers with similar job-finding rates. Subsequently, the job-finding rate of those who will be the most unemployed in prime-age declines, whereas that of other workers remains relatively constant. With respect to job-separation rates, differences are large right from the start. Those who will belong to the 10% most unemployed workers in prime age have more than twice the likelihood of separating from a job to unemployment at age 20 relative to other workers. Finally, differences in job-separation rates remain high over time.

Why are separation rates so heterogeneous and persistent over the life cycle? Why do differences in the job-finding rate start small and increase over time? And why do the same workers experience both low finding rates and higher separation rates? As I will discuss later, explanations based on standard human capital models (such as Ljungqvist and Sargent, 1998) are at odds with these facts.

I propose a directed search model of learning that is consistent with all the above-mentioned microobservations. In the model, workers can be of two types: high or low. A worker's type is initially unobserved by all agents in the market, who are allowed to learn workers' types from his labor market history. This feature makes the model consistent with the observation that, although differences in separation rates are large from the start, differences in job-finding rates are small at first and then increase over the span of a career. This is because workers in the model who experience frequent separations when young progressively find fewer jobs, because their prospective employers observe a poor labor market history. Thus, the model is able to answer the questions posed above.

Information in the model is symmetric. Thus, at the start of a worker's career, no agent in the market (including the worker himself) knows his type. Search is competitive in the sense that workers decide to search for a job that delivers a certain lifetime value. Upon matching with a firm, workers draw match quality from a type-specific distribution, which is constant for the duration of the match. Firms write complete contracts and matches are destroyed whenever their surplus is negative. Match quality is an experience good, as in Jovanovic (1979). The output of a match is unobserved until a shock is realized, at which point it becomes known to the firm–worker match. The match is then either continued or destroyed, leaving the worker unemployed in the latter case. Past realizations of match quality for each worker are observed by the market, and are used to update the probability that the worker is of high type accordingly. The probability of being high type formalizes the notion of a "résumé" based on the worker's labor market history. Low values for match quality will lead the market to believe that the worker is more likely to be of low type, and vice versa. Thus, a worker's type is gradually learned from his labor market history and workers with different résumés apply for jobs in different submarkets.

To the best of my knowledge, this is the first competitive search model in which job-finding rates, job-separation rates, and the speed of learning are all endogenously and simultaneously determined.[3] This is accomplished by allowing workers to choose the submarket in which they search, understanding that each lifetime value entails a different job-finding rate, and by having the worker's outside option evolve dynamically with his history, thus determining both his future desired lifetime value and his probability of separating from a job.

---

[3] The model in Doppelt (2016) is the closest to the one studied here, but the two papers differ along several dimensions. Doppelt uses a directed search model in which wages are set through bargaining to examine the job-finding rate/unemployment duration relationship. I develop a competitive search framework to investigate heterogeneity in lifetime unemployment instead.

Heterogeneity in lifetime unemployment originates from three sources in the model. The first is heterogeneity across workers, since one type of workers is always more likely to draw low match quality values, which will ultimately lead to separations. The second is bad luck, since any given type can draw low match qualities and thus separate to unemployment. The third is information frictions, where low-type workers wrongly infer from lucky draws that they are likely to be high types. As a result, they might leave their current jobs to sample a better one, only to experience a bad draw later on and become unemployed again. Conversely, high-type workers might experience unlucky draws, so that their résumé will look worse the next time they look for a job, and thus, they will experience longer unemployment duration.

The model is estimated using data from the NLSY/79. The model is successful in reproducing the observed concentration and persistence of unemployment, as well as the patterns of job-finding rates and job-separation rates over the life cycle. The model delivers concentration of unemployment since low-type workers have a higher probability of drawing low-quality matches than high-type workers and they have lower expected productivity. Thus, such workers face a higher separation rate and a lower job-finding rate at every age. The model delivers persistence because low-type workers tend to experience frequent separations both when young and when of prime age, and job-finding rates that decline with age as the market recognizes them as low-type workers. Information frictions are crucial to match the life cycle patterns of job-finding and job-separation rates by unemployment group. I argue that a model based on human capital, or on other forms of bad luck that accumulate over the lifetime, instead of information frictions, would be inconsistent with these patterns since it would have the counterfactual implication that differences in separation rates start small and increase over the life cycle.

By progressively shutting down features of the baseline model, I find that neglecting heterogeneity across workers makes it impossible to match the concentration and persistence of unemployment observed in the data. Although uncertainty in match quality draws helps in matching the life cycle profile of job-separation rates and the concentration of prime-age unemployment, heterogeneity across workers is crucial in matching the documented persistent differences in job-finding and job-separation rates across workers. Furthermore, uncertainty in match quality draws is important because it slows down learning: If there is no uncertainty and workers only differ in mean match quality, learning is too fast and it is impossible to match the progressive decrease in job-finding and job-separation rates by prime-age unemployment groups.

Are information frictions important in determining labor market outcomes? In a quantitative exercise, I shut down information frictions and show that they are responsible for the entire decline in monthly job-finding rates among the top 10% of prime-age unemployed (from 21% at age 20 to 10% at age 35). This is because 89% of the top 10% unemployed are low types and, although their type is initially unknown, it is slowly revealed by their labor market histories. This translates into progressively lower job-finding rates for these workers. Information frictions also explain part of the decline in the separation rates of the most unemployed workers. In short, such frictions are important for explaining the labor market outcomes of young workers. However, the role of information frictions later in life is negligible. By age 30, types have effectively been learned for most of the population, so that most of the concentration and persistence of unemployment after this age is due to heterogeneity across workers.

This article mainly contributes to two strands of the literature. First, it relates to the large empirical literature that investigates the composition of the unemployment pool and heterogeneity in job-finding rates (Clark et al., 1979; Addison and Portugal, 1989; Lockwood, 1991). I add to this literature by showing that most of the prime-age unemployment pool is composed of a relatively small group of workers, who exit employment at a higher rate and stay unemployed for lengthier periods. Other papers have found that lifetime unemployment is relatively concentrated (Michelacci et al., 2012, for the United States; Schmillen and Moller, 2012, for Germany; and Brooks, 2005, for Canada), but have not focused on prime-age unemployment. I further add to this literature by decomposing concentration into job-finding and job-separation rates, documenting young to prime-age persistence and showing the life cycle patterns of job-finding and job-separation rates by prime-age unemployment groups. My results for the speed of

employer learning are similar to those of Lange (2007), who finds that employers learn relatively quickly and expectation errors on productivity decline by 50% during the first three years of employment.[4]

Second, I contribute to the theoretical literature on job search and learning by developing a competitive search model of unemployment and learning from labor market histories, in which job-finding rates, job-separation rates, and the speed of learning are all jointly determined in equilibrium. This is also the first model to study lifetime unemployment data and use it in estimation. Other models of job search have proposed the combination of unobserved heterogeneity and learning as a candidate explanation for the scars of unemployment (Michaud, 2018) or heterogeneity in job-finding rates (Shimer, 2008; Gonzalez and Shi, 2010; Fernández-Blanco and Preugschat, 2018). This article combines a directed-search environment with a mechanism similar to Michaud (2018) regarding separations, but also adds résumés, learning from labor market histories and heterogeneity in the shape of match quality distributions across types. Doppelt (2016) contemporaneously developed a model similar to the one in this article, to investigate the negative relationship between job-finding rates and unemployment duration found in the data. I share with Hagedorn and Manovskii (2013), Mustre-del-Río (2012), and Jung and Kuhn (2019) the idea that longer employment duration signals higher match quality. Finally, my model is, in spirit, a life cycle model of search and matching which environment is similar to that of Menzio et al. (2016). I build on their model of homogeneous workers and allow for unobserved heterogeneity and learning.

The article is organized as follows: Section 2 describes the data on lifetime unemployment and the patterns of job-finding and job-separation rates over the life cycle I uncover. Section 3 sets up the model and Section 4 discusses its identification. Section 5 outlines the results of model simulations and quantitative exercises. Section 6 discusses possible competing mechanisms for explaining the data and the role of some model assumptions. Section 7 discusses the model's policy implications and concludes. All online appendixes contain details on data construction and robustness checks for the main data results, except for online Appendix A.4 that contains four graphs of the model-simulated data and online Appendix A.5 that contains all proofs not included in the article.

2. THE DATA

I use weekly job histories taken from NLSY/79 data to compute lifetime unemployment statistics. The NLSY is one of the best-known panel data sets available for the United States. It follows a cohort of more than 10,000 individuals starting from 1979, when they were aged 14–22. The data were gathered annually until 1994, and biannually since then.

I use only the cross-sectional representative sample of the NLSY and restrict my attention to males who have completed only a high school education by age 30.[5] Furthermore, I exclude any worker who has less than 100 weeks of reported employment/unemployment from age 20–30, or less than 100 weeks from age 35–50.[6] This provides a final sample of 1,083 workers, who are observed for about 1,300 weeks on average. Nonetheless, the results are robust to more inclusive definitions of the sample.[7] Finally, following Chetty (2008), I consider an individual

---

[4] Other empirical work focuses on employer learning as a source of increase in wage heterogeneity over the career (see, e.g., Kahn and Lange, 2014).

[5] This means that I include in the sample only individuals who have completed no more and no less than high school at age 30. I do this to attain as homogeneous a sample as possible. High school graduate males are the largest education/gender subgroup in the NLSY/79. Moreover, Menzio et al. (2016) show that, in terms of labor market outcomes, this subgroup is a good representation of the behavior of U.S. labor market aggregates over the life cycle. In online Appendix A.2.2, I show that the findings are robust for other education/gender subgroups.

[6] This is to address measurement error issues when computing lifetime unemployment statistics. I examine the extent of measurement error in online Appendix A.2.3.

[7] For results using the entire sample, see online Appendix A.2.2.

TABLE 1
LIFETIME UNEMPLOYMENT: DATA VS UNIFORM MATCHING MODEL

	(1) NLSY/79	(2) Unif. Match
Avg. % time in unemployment	4.7	4.9
Avg. % time in U, excluding top 10%	2.0	3.1
Avg. % time in U, excluding top 20%	0.8	2.0
% never unemployed:	58	40

NOTES: Column (1): Lifetime unemployment statistics computed on NLSY/79, individuals aged 35-55. Sample includes only high school educated, male individuals with more than 100 observations of weekly job histories in their prime age, ending 2010. Column (2): averages computed by simulating sequences of 500 job-finding–job-separation events using flow equations of Mortensen-Pissarides model, calibrated to the average job-finding and job-separation probabilities in NLSY/79 sample.

unemployed in a given week if that week is part of a nonemployment spell during which the individual reports being unemployed and searching for a job for at least one week.[8]

2.1. *Prime-Age Unemployment Is Concentrated.* I first document that prime-age unemployment is concentrated among relatively few individuals. I start by defining young-age unemployment as the proportion of an individual's work history spent in unemployment out of total weeks employed or unemployed[9] from age 20 to 30:

$$(1) \qquad \bar{u}_i^y = \frac{\sum_{t=1}^{T_i^y} u_{i,t}^y}{T_i^y},$$

where $u_{i,t}^y$ is a variable that takes value 1 for weeks in which individual i was unemployed, and 0 for weeks in which he was employed, and T_i^y is the number of weeks that individual i was either employed or unemployed between the ages of 20 and 30. Similarly, I define prime-age unemployment as the proportion of a work history spent in unemployment from age 35 to 50. Since I will show that there are important connections between young and prime-age unemployment, the five-year gap is necessary to ensure that the correlations are not in part due to the aftermath of a recession, or to long unemployment spells that span between the two periods.

As shown in Table 1, there are large differences in unemployment outcomes across workers. The first finding is that prime-age unemployment is concentrated among relatively few workers. After ranking individuals by the fraction of time spent in unemployment, I compute the fraction of weeks spent in unemployment by the bottom 90% of the sample:[10]

$$(2) \qquad \bar{u}_{u^p < q_{90}}^p = \frac{\sum_{i=1}^{N} 1(\bar{u}_i^p < q_{90}(u^p)) \sum_{t=1}^{T_i^p} u_{i,t}^p}{\sum_{i=1}^{N} 1(\bar{u}_i^p < q_{90}(u^p)) T_i^p},$$

where $1(u_i^p < q_{90}(u^p))$ is an indicator function that takes value 1 if the prime-age unemployment of individual i is below the 90th percentile of the prime-age unemployment distribution, and

---

[8] This is also to address a specificity in the design of unemployment and nonemployment spells in the NSLY/79. I provide more details in online Appendix A.1. Results are robust to alternative definitions of unemployment, as described in online Appendix A.3.

[9] My definitions are similar to those used by Schmillen and Möller (2012).

[10] These measures are commonly used in the literature on income inequality (see, e.g., Atkinson, 1970). Their application to lifetime unemployment is relatively uncommon, with the exception of Schmillen and Möller (2012) and Brooks (2005).

0 otherwise, whereas T_i^p is the number of weeks in which individual i was either employed or unemployed while of prime age.[11]

The 10% most unemployed individuals account for about 2/3 of the prime-age unemployment observed in the data. Moreover, about half of the sample has never been unemployed in the reference period. Notice that the fact that prime-age unemployment is concentrated among relatively few workers is very different from the phenomenon noted (among others) by Clark et al. (1979), who show that most of the unemployment pool is accounted for by workers staying unemployed, instead of workers entering and exiting unemployment. The phenomenon documented by Clark et al. is mainly driven by heterogeneity in unemployment duration, whereas I will show below that the frequency of unemployment spells is even more important in explaining the concentration of prime-age unemployment.

As in the discussion of income inequality, measures of concentration may not be meaningful unless they are compared to what a standard framework would imply for the distribution of unemployment. If only one person out of 10,000 were unemployed, the fact that unemployment is concentrated would not be very interesting. Moreover, it is important to stress that these numbers do not accurately represent differences in the "underlying" job-finding and job-separation rates for groups of workers. My estimates of job-finding and job-separation probabilities are likely to be biased estimates of the underlying probabilities faced by different groups of workers, because by creating groups based on the amount of unemployment experienced in prime age, I am selecting those individuals who experienced exceptionally high amounts of unemployment, who might be the most "unlucky" among a specific group. To understand the magnitude of these results, I compare the concentration of unemployment observed in the data to what a standard search and matching framework à-la Mortensen and Pissarides (1994) would imply.[12]

The simulations show that the standard model, when estimated to reproduce the average job-finding and job-separation rates in the sample, has trouble replicating the observed concentration in prime-age unemployment, since it predicts too many transitions in and out of unemployment for the majority of workers. This fact is important, because it suggests that heterogeneity across workers is likely to be crucial in making sense of labor market outcomes, and of the ins and outs of unemployment during prime age.

I now proceed to compute weekly average job-finding/job-separation rates for workers in their prime. Job-separations always refer to separations to unemployment throughout the article. As can be seen from Table 2, the concentration of unemployment is due both to a lower job-finding rate[13] (by about four times) and a higher separation rate (by about eight times) for the top 10% (see Table 2). This group of workers therefore appears to have both longer unemployment duration and shorter employment duration. Since unemployment is a nonlinear function of both finding and separation rates, failure to account for both simultaneously means not getting the distribution of lifetime unemployment right. Interestingly, the difference in separation rates accounts for a larger proportion of the heterogeneity in unemployment outcomes than the difference in finding rates.

---

[11] Clearly, this is not the only way to compute this average. Another possibility is to compute:

$$(3) \qquad \bar{u}_{u^p < q_{90}}^p = \frac{\sum_{i=1}^{N} \left[1(\bar{u}_i^p < q_{90}(u^p)) u_i^p \right]}{\sum_{i=1}^{N} 1(\bar{u}_i^p < q_{90}(u^p))},$$

which is the average of each individual's prime-age unemployment. The two averages differ since T_i^p varies across individuals (because some individuals are observed for more weeks than others). In particular, they can differ significantly if $\mathbb{COV}(T^p, u^p) \neq 0$, for instance, if those often unemployed also tend to be more often out of the labor force. In fact, this is indeed the case (see online Appendix A.2.1). I find that there is relatively little difference between the two methods of computing the average, and that they deliver very similar results. Indeed, the concentration of unemployment is even larger (about 70% accounted for by the top 10%) according to the second methodology (see online Appendix A.2.1).

[12] Simulations of 500 weeks of transitions are used because in my NLSY/79 sample, prime-age workers are observed for about 600 weeks on average. Increasing the number of simulated weeks worsens the performance of the standard model further.

[13] Online Appendix A.1.1 describes the calculation of job-finding and job-separation rates.

TABLE 2

SUMMARY STATISTICS BY PARTS OF THE PRIME-AGE UNEMPLOYMENT DISTRIBUTION

	Top 10%	Rest of Sample	Ratio Top 10/Rest
Avg. % time in unemployment	35.5	2.0	18.1
f: Prob. of $U \to E$ (weekly%)	1.61	6.15	0.26
δ: Prob. of $E \to U$ (weekly%)	0.96	0.13	7.6
Predicted % time in U of top 10%,	9.7		
f alone ($\delta/(\delta + f_{\text{top10}})$):			
Predicted % time in U of top 10%,	22.3		
δ alone ($\delta_{\text{top10}}/(\delta_{\text{top10}} + f)$):			
Avg. Log Wage (2000)	$\simeq -41\%$		

SOURCE: Author's calculations on NLSY/79. Male, high-school educated individuals aged 35–55. Predicted % time in U calculated using the formula $u = \delta/(\delta + f)$.

TABLE 3

MARKOV TRANSITION MATRIX, FROM DISTRIBUTION OF YOUNG (AGES 20–30) UNEMPLOYMENT TO PRIME-AGE (35–55) UNEMPLOYMENT. OVERALL SAMPLE (TOP PANEL) AND ONLY HIGH SCHOOL MALES (BOTTOM PANEL)

	Rest	Top 10% When (35–55)
All Sample		
Rest	92.89	7.11
Top 10% when (20–30)	64.26	35.74
High School Workers		
Rest	93.23	6.77
Top 10% when (20–30)	61.11	38.89

SOURCE: Author's calculations on NLSY/79.

To summarize, I find that prime-age unemployment is concentrated in relatively few workers, that the standard search-and-matching framework cannot replicate this finding, and that most of the concentration of unemployment in prime age is accounted for by heterogeneity in job-separation rates, instead of job-finding rates.

2.2. *Unemployment Is Persistent over the Life Cycle.* I now document that young and prime-age unemployment are strongly correlated. Table 3 shows that workers who were in the top 10% of the young-age unemployment distribution are five times more likely to be in that same part of the distribution when of prime age. In short, young and prime-age unemployment are connected and, for a wide range of observables in the NLSY/79, young unemployment is the best predictor of prime-age unemployment. Accordingly, regression analysis (see Table 9 in the online Appendix) confirms that young unemployment is a strong predictor of prime-age unemployment, and that this is not due to observables such as education, marital status, and IQ (IQ as measured by the Army Force Qualification Test score).

Finally, notice that such persistence is not explained by other kinds of observable heterogeneity, such as differences across occupations (i.e., the choice of an "unlucky occupation" when young, as in Schmillen and Möller, 2012) or in health status (poor health that adversely affects labor market outcomes). I perform a battery of regressions (see Table 4) that shows that ethnic origin, education, prior and current occupation, ex post health, and IQ do not substantially explain the amount of persistence in unemployment. This result is particularly strong because current occupations and ex post health are endogenous to prior labor market experience, and as such should capture part of the persistence in unemployment. For instance, a worker who has frequently been unemployed when young will typically work in less stable occupations in prime age, which should capture some of the young-prime-age correlation. Similar considerations can be made for ex post health.

TABLE 4
REGRESSION OF % OF UNEMPLOYMENT IN PRIME AGE (35–50) ON % OF YOUNG UNEMPLOYMENT (20–30)

	(1) No Young U	(2) Base	(3) Occ	(4) Occ 2	(5) Occ 2 + Hlt
% Time U. when 20–30		0.288*** (0.018)	0.285*** (0.018)	0.283*** (0.018)	0.277*** (0.018)
Demographics					
Married	−0.0428***	−0.0251***	−0.0244***	−0.0251***	−0.0242***
Education and AFQT					
Some college	0.0102	0.00950	0.00463	0.00425	0.00210
High school	0.0254***	0.0226**	0.0175*	0.0156	0.0102
<High school	0.0746***	0.0584***	0.0544***	0.0514***	0.0429***
AFQT Quartile = 2	−0.0138	−0.00165	−0.00201	−0.00221	0.000873
AFQT Quartile = 3	−0.0247**	−0.00662	−0.00607	−0.00657	−0.00172
AFQT Quartile = 4	−0.0259**	−0.00707	−0.00549	−0.00528	−0.000916
Standard controls	X	X	X	X	X
Prevalent occupation (one digit)			X	X	X
Occupation in 1990 (1d)				X	X
Health status (2000)					X
R^2	0.119	0.212	0.217	0.220	0.244
Observations	2,188	2,188	2,188	2,188	2188

NOTES: Column (1) reports results from a regression on all workers, including only controls and no young unemployment; column (2) includes young unemployment; column (3) adds dummies for the occupation in which the individual spent most time; column (4) adds dummies for occupation in 1990; and column (5) adds self-reported health at age 40. All regressions include controls for education at age 30, ethnic group, age in 1979, marital status at age 30, AFQT test score quartile.
SOURCE: Author's calculations on males from NLSY/79. Other controls' coefficients are not reported for reading convenience. More complete results can be found in Table 14 in the online Appendix. Standard errors in parentheses.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Moreover, comparing the first two columns of Table 4 shows that several observables, particularly IQ measures and marital status, lose importance when young unemployment is taken into account. This suggests that a same fixed characteristic might be correlated with all three measures, but that young unemployment may be measuring it more precisely.[14]

2.3. *Job-Finding and Job-Separation over the Life Cycle.*   As a final piece of evidence, I compute job-finding and job-separation probabilities depending on age, for ages 20–35, by groups of prime-age unemployment (Figure 1). My intention is to show that those who have experienced more prime-age unemployment had different labor market outcomes during their early careers as well. I compute average rates from probit regressions of job-finding rates and job-separation rates on a third-degree polynomial in age by prime-age unemployment groups, controlling for observables and year-specific fixed effects to clean the effect of recessions.[15] Between the ages of 20 and 30, the job-separation rate of the top 10% of prime-age unemployed is more than twice that of the rest of the sample, and this gap increases over the years, both in absolute and relative terms. On the other hand, the two groups experience virtually identical job-finding

---

[14] Including noncognitive skills information from the NLSY/79 in the regressions does not have any impact on the estimates, possibly because of measurement error. The fact that a better observable is hard to find in the NLSY/79 does not mean that one does not exist in general. For instance, Lindqvist and Vestman (2011) find that, in Sweden, high-quality measures of noncognitive ability based on personal interviews conducted by psychologists are strongly associated with the risk of unemployment.

[15] The results are essentially identical, though more noisy, if the averages are computed with one-year age intervals instead of restricting to a functional form. I chose the polynomial shape for presentation purposes. Results under the age-group specification appear in Figure 10 of Supporting Information and are used as an alternative identification scheme for the model in a robustness check (available upon request).
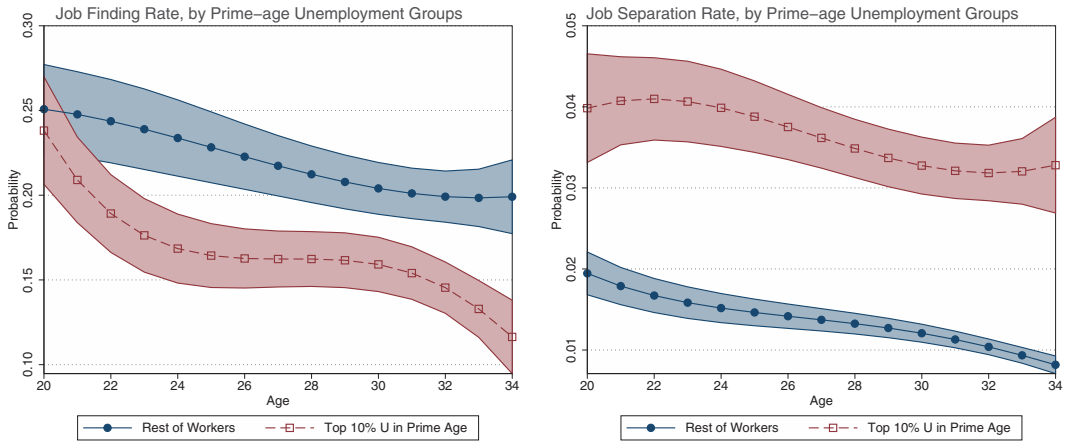
NOTES: Sample of male, high-school educated workers. Average at means of other covariates; probit regressions include controls for further education achieved after 30, region of residence, ethnic group, year, and marital status.
SOURCE: Author's calculations using NLSY/79. Shaded areas are 95% confidence bands.

FIGURE 1

MONTHLY JOB-FINDING (LEFT PANEL) AND JOB-SEPARATION (RIGHT PANEL) PROBABILITIES, BY GROUP OF PRIME-AGE UNEMPLOYED [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

rates at age 20, but a gap appears and widens over time as workers age, particularly due to the decline in the job-finding rate of the top 10% of prime-age unemployed.

This suggests that, in the eyes of potential employers, the two groups of workers did not look substantially different at the beginning of their working careers, since they were hired with similar probabilities, but became increasingly distinct over time. Nonetheless, the high separation rates experienced by the top 10% of prime-age unemployed during their 20s suggest that such workers were recognized to be different *during* an employment relationship.[16] In other words, at the start of their career, young workers who later experienced substantially different labor market outcomes appeared similar; however, as they accumulated jobs and separations, job-finding rates diverged, suggesting that information on them had gradually become available.

I further decompose the separation rate using the beta-matched employer–employee variables available in the NLSY/79 (see Table 15 in the online Appendix), which provides some evidence supporting this interpretation. Workers in the top 10% of prime-age unemployment are fired three times more often than other workers when young, although they are also substantially more likely to experience an involuntary separation (due to a layoff or an establishment closure) and to quit to look for another job.[17] In prime age, the various reasons for losing a job seem to be of equal importance. Involuntary separation and quitting account for most of job separations observed in the data at all stages of the life cycle.

The wages of the top 10% group progressively decline over the life cycle relative to the rest of the workers (see Figure 8 in the online Appendix), confirming that differences across workers become more pronounced over the course of their careers. This suggests that, after sampling a number of jobs, they are recognized as less productive, or they sort into different jobs to avoid frequent separations in the future, or they fail to accumulate skills that lead to higher wages.

These facts motivate the need for a theory of unemployment that is capable of replicating the concentration of unemployment among relatively few workers and the persistence of unemployment over the life cycle, phenomena that have important implications for the design of labor market policy. For instance, the concentration of unemployment suggests that a relatively

---

[16] Section 6 discusses why explanations based on human capital are insufficient to explain such patterns.

[17] In the model section, I choose not to take a stand on the interpretation of different labelings of job separations. My modeling strategy is based on the surplus of a firm–worker match and is thus silent on the interpretation of specific job separations.

small group will receive the bulk of unemployment insurance benefits, and will be the most affected by changes in the level of such benefits. However, I argue that the relatively low ex ante difference in job-finding rates and the large ex post differences in both job-finding rates and wages suggest that important information frictions are at work in the first years of a worker's career, and that information on workers slowly accumulates over their careers. The model presented here will feature this mechanism, which has important implications for understanding the concentration of unemployment, the connection between young and prime-age unemployment, and the effects of labor market policies.

3. MODEL

In this section, I set up a directed search model of heterogeneity in labor market outcomes, roughly based on Moen (1997), Menzio and Shi (2011), and Gonzalez and Shi (2010). The ingredients of the model are motivated by the evidence presented in the previous section.

To obtain believable life cycle profiles of separation rates, I add heterogeneity in match quality draws, as in Menzio et al. (2016), which also embeds—in a reduced-form manner—the idea of "sorting into the right job" as a determinant of heterogeneity in career paths. Heterogeneity across workers, information frictions, and the notion of a worker's "résumé" is added to capture the fact that the group of most unemployed workers experiences higher separation rates at the start of their career, and that such separation rate diminishes later. This is because such workers are being separated often (as in Gibbons and Katz, 1991) and are learning that they have low productivity. Thus, their outside option declines, which reduces their probability of separating from a job. Moreover, such workers progressively find fewer jobs and earn lower wages, since their résumés worsens over time, thus reducing their expected productivity in the eyes of potential employers. Finally, heterogeneity across workers can explain the relatively high levels of persistence in unemployment found in the data.

3.1. *Environment.* The economy is populated by a measure of firms $M > 1$ and a measure one of workers, who are either employed or unemployed. In every period, a fraction λ of workers die, and are replaced by newly born, unemployed workers. Each worker is born either as type H or type L, High and Low, respectively, with type being unknown both to firms and workers. Low types occur with probability l and high types with probability $1 - l$. All agents are risk-neutral and discount the future at the factor $1/(1 + r)$.

Let p be the probability of a worker being high-type. There exists a continuum of submarkets indexed by $\{v, p\}$, the expected lifetime value v offered by firms to workers in that submarket and the prior p of workers applying to that submarket.[18] Matches are endogenously destroyed when their surplus is negative. Some matches end randomly with probability δ.

3.2. *Search and Matching.* Firms can post vacancies in any submarket at cost κ. Search is directed, in the sense that workers with prior $\bar{p}$ have to choose the submarket $\{v, \bar{p}\}$ in which to search. Thus, the ratio of vacancies to searching workers in each submarket is denoted by $\theta(v, p)$, the tightness in that submarket. The number of matches in each submarket is determined by the matching function $m = g(\theta)$, such that the job-finding probability is $f(\theta) = m/u$, which satisfies $f' > 0$, $f'' < 0$, $f(0) = 0$, and $\lim_{\theta \to \infty} f(\theta) = 1$, and the job-filling probability is $q(\theta) = m/v = f(\theta)/\theta$. Unemployed workers can search for a job, whereas employed workers cannot.[19] When unemployed, workers get flow utility b.

---

[18] This is to make fewer assumptions on the distributions of match quality that follow. In principle, the model can be rewritten to feature submarkets indexed only by $\{v\}$, provided that further assumptions are made on the match quality distributions so that workers with different p will apply to different submarkets. In the current version, I allow workers with different p to apply to contracts delivering the same lifetime value, although in equilibrium, for every value $p \in [0, 1]$, only one submarket $\{v(p), p\}$ will be active.

[19] Adding on-the-job-search might alter the quantitative predictions of the model, but its main mechanism would remain intact.

3.3. *Information and Learning.* High- and low-type workers each draw the productivity of a job from different distributions. Denote by $H(x)$ and $L(x)$ the cumulative distribution functions of match quality for high and low types, respectively, with support $X \subseteq [0, \bar{x}]$, such that $H(x)$ strictly first-order stochastically dominates $L(x)$, that is, $H(x) < L(x) \, \forall \, x < \bar{x}$. Once a worker is matched, a match-specific quality shock is drawn from the worker's type distribution. Match quality is constant over the entire duration of the match.[20]

At the beginning of a firm-worker match, its output is unobserved. Match quality is an experience good, as in Jovanovic (1979). Thus, once a worker with prior p has been matched, the match produces the expected payoff $\mathbb{E}(x \mid p)$ in each period, until the worker/firm pair has observed the match's output or a random separation occurs.[21] The output of the match is observed with probability π in each period.

The history of a worker's observed match quality draws is observable by all agents in the market.[22] Thus, upon observing output, agents gain information on a worker's type.[23]

3.4. *Contracts.* I assume that employment contracts are complete, in the sense that they specify a wage w paid by the firm to the worker and a probability of separation d at every point in time, as a function of the promised expected lifetime utility, the history of the worker, and the history of the firm–worker match.[24] As shown by Menzio and Shi (2011), since contracts are complete and utility is transferable, it is optimal for firms to offer contracts that are bilaterally efficient, so that they maximize the sum of the firm's lifetime profits and the worker's lifetime utility. Thus, the firm finds it optimal to offer a probability of separation d delivering bilateral efficiency (i.e., matches are kept as long as their lifetime value is higher than the outside option of the worker[25]) and a wage w such that the lifetime utility v is delivered in expectation to a worker in a $\{v, p\}$ match.[26] However, there are many different sequences of w that deliver

---

[20] Menzio et al. (2016) find that, in a similar model, the probability that a match changes quality during an employment relation is around 1%, thus making the constant match assumption a reasonable simplifying approximation.

[21] From the perspective of writing the surplus of a match, this is equivalent to assuming that both parties obtain zero value until productivity is observed and then obtain the sum of flow utilities for all previous periods.

[22] Although this seems like a strong assumption, it is possible to show that a model in which match quality is constant over the duration of a match and observed only by the firm produces similar dynamics if wage renegotiation is allowed every period. The intuition is that, over the duration of a match, the worker and the rest of the market learn that the worker is more likely to be a high-type, since a separation did not occur. Thus, the worker's outside option increases and he negotiates a higher wage. Such a model would be much harder to solve and its mechanism would be less transparent, since it would need to introduce another state variable (the duration of a match) and solve a dynamic asymmetric information problem over a match's duration. I adopt this simplifying assumption to focus on the discussion of lifetime unemployment inequality.

[23] Although unobserved heterogeneity is often studied as an adverse-selection problem, since it is assumed that workers know more about their type than prospective employers, information in the model diffuses symmetrically on both sides of the market. The results of Guerrieri et al. (2010) motivate this assumption by showing that, in a directed search environment with adverse selection, it is always possible to write separating contracts. This makes the symmetric information assumption somewhat less important in directed search models, although the quantitative implications of a model featuring adverse selection might change.

[24] This assumption reflects the idea that matches can be kept as long as they are profitable to both parties, so that the relation between labor market histories and learning is not strongly dependent on the contract environment, but rather on the features of match quality distributions and on the evolution of the outside option of workers. The idea that types are being learned over workers' careers does not hinge necessarily on the particular contract space assumed here, but its quantitative implications might be affected. For instance, one could think of an environment in which match quality is the firm's private information and workers learn from whether they are kept on or fired by the firm.

[25] This mechanism is similar to that of Gibbons and Katz (1991): When agents find out that a worker–firm match has low productivity, it is interrupted. Although they model firing, I do not take a stand on whether a separation should be labeled a firing, a layoff, or a quit.

[26] This does not mean that wages will only depend on p: Wages can depend on match quality x too, and in general will be state-contingent. For instance, when match quality is close to the value of unemployment, maintaining the match is bilaterally efficient, but only a low enough wage will sustain the match (if a wage is to be paid every period). The full state-contingent contract specifies wages for any combination of x and p such that, ex ante, lifetime utility v is delivered to the worker in expectation.

the same lifetime utility to the worker. In the present model, it is not necessary to resolve this indeterminacy since the focus is on the patterns of job-finding and job-separation rates.[27]

The intuition is that jobs will be endogenously destroyed whenever the value of the match is lower than the outside opportunity of the worker. Thus, low-type workers will face more frequent separations if they typically draw match qualities that are below the value of unemployment. Thus, π and the properties of the match quality distributions $H(x)$ and $L(x)$ measure the *informational content of job duration*. If $\pi = 0$, match quality is never observed and job duration is not informative of the worker's type.

Since the history of past match qualities drawn by a worker is observable by potential employers, it follows that p is a sufficient statistic for the entire history of a worker's match quality, and can be considered as the worker's "résumé."

Timing is as follows:

1. Workers die w.p. λ and are replaced by unemployed workers with belief $1 - l$.
2. Firms and workers observe the output of a match with probability π.
3. Workers revise their beliefs: $p' = p$ if match quality is still unobserved and no shocks occur; $p' = P(x, p)$ if match quality has just been observed.
4. Production occurs and wages are paid.
5. Unemployed workers with belief p choose the submarket $\{v', p\}$ in which to search.
6. Unemployed workers match w.p. $f(\theta(v', p))$. Separations (both exogenous and endogenous) occur.
7. Newly matched workers draw match quality from $H(x)$ or $L(x)$ depending on their type.

Bayes' rule implies that the beliefs of employed workers who observe the realization of match quality evolve according to

$$(4) \qquad p' = P(x, p) = \frac{p\, h(x)}{p\, h(x) + (1 - p)\, l(x)},$$

where $h(x)$ and $l(x)$ are the density functions of match quality draws for high and low types, respectively.

Heterogeneity in job duration across workers and the updating of p are closely related: When a worker's productivity is observed and the match is maintained, it must mean that match quality was high enough to support the match, an event that is more likely for high types because $H(x)$ first-order stochastically dominates $L(x)$. Conversely, when a worker's productivity is observed and the match is destroyed, it must mean that match quality was too low, an event that is more likely for low types. Clearly, this depends on how high the flow utility of unemployment is relative to the typical match quality draws of low types. However, depending on the features of the match quality distributions, matches might also be destroyed because workers prefer to have the opportunity to sample a new job with higher productivity.

3.5. *Bellman Equations.* The value function of an unemployed worker with prior p can be written as

$$(5) \qquad U(p) = b + \beta \Big[\max_{v} [f(v, p)(v - U(p))] + U(p) \Big],$$

where $\beta = \frac{1-\lambda}{1+r}$.

The joint firm–worker value of a match for which output is known can be written as

$$(6) \qquad S(x, p) = x + \beta \Big\{ \max_{d \in [\delta, 1]} [(1 - d)\, S(x, p) + d\, U(p)] \Big\}.$$

---

[27] When discussing wages in the model and in the data, I will often resort to the specific case of Nash bargaining.

Thus, the joint value of a match for which output is unknown and the worker has prior p can be written as

(7)
$$S_u(p) = (1 - \pi)[\mathbb{E}(x \mid p) + \beta((1 - \delta) S_u(p) + \delta U(p))] +$$
$$\pi\left(p \int S(x, p') \, dH(x) + (1 - p) \int S(x, p') \, dL(x)\right),$$

where $p' = P(x, p)$ denotes the next period's prior depending on the realization of match quality (while suppressing the dependence on p and x for readability).

The value of a firm posting a vacancy in submarket $\{v, p\}$ is

(8)
$$V(v, p) = -\kappa + q(\theta(v, p))\beta(S_u(p) - v),$$

and the tightness function must satisfy

(9)
$$\kappa \geq q(\theta(v, p))\beta(S_u(v, p) - v) \quad \forall v, p,$$

which makes θ consistent with the firm's optimal vacancy creation. Equation (9) holds with equality if $\theta > 0$. Essentially, condition (9) implies that if $\theta = 0$, such tightness is consistent with the firm's optimal choice only if the benefit from creating a vacancy is smaller than the cost.

3.6. *Equilibrium.* DEFINITION 1. A Bayesian Markov perfect Block Recursive Equilibrium (BRE) for this economy consists of a value function for the unemployed worker $U(p)$, a policy function for the unemployed worker $v^*(p)$, value functions for the joint value of a match $S_u(p)$ and $S(x, p)$, a separation policy $d(x, p)$, a tightness function $\theta(v, p)$, and a law of motion for beliefs $p'(x, p)$ such that:

1. $U(p)$, $v^*(p)$, $S(x, p)$, $S_u(p)$, $d(x, p)$, and $\theta(v, p)$ are independent of the aggregate state ψ.
2. $\theta(v, p)$ satisfies (9) $\forall v, p$ and $\theta(v, p) \geq 0$ with complementary slackness.
3. $U(p)$ and $v^*(p)$ satisfy (5) $\forall p$.
4. $S(x, p)$ and $d(x, p)$ satisfy (6) $\forall w, x$.
5. $S_u(x, p)$ satisfies (7) $\forall w, x$.
6. $p'(x, p)$ satisfies (4).

The BRE is much easier to solve than a recursive equilibrium because, as shown in Menzio and Shi (2011), the value functions and policy functions of agents depend only on the individual states $\{p, x\}$ and not on aggregate states. Aggregate statistics can be computed, after solving the individual problem, from the aggregation of individual choices. Moreover, computing transitions out of the steady state is straightforward since all policy functions and laws of motion are independent of the aggregate state.

3.6.1. *Characterization of Equilibrium.* LEMMA 1. Given $p \in [0, 1]$, $S(x, p)$ is increasing in $x \in [0, \bar{x}]$. Also, $d(x, p)$ is a step function that takes value 1 if $x < U(p)(1 - \beta)$ and δ otherwise.

PROOF. See online Appendix A.5. The intuition is straightforward: Matches are maintained as long as the discounted value of the stream of match quality is higher than the value of unemployment. Clearly, distributions of match quality that have more probability mass on values that are lower than b will lead to more frequent separations.

COROLLARY 1. *The ex ante probability of drawing match quality lower than b, and as a con-sequence experiencing a separation to unemployment, is greater or equal for low types than for high types.*

The intuition is that the lower bound of $U(p)$ is equal to $b/(1 - \beta)$; therefore, any match quality draw below b will certainly lead to a separation. Because of first order stochastic domi-nance (FOSD), $L(b) > H(b)$: low types are always more at risk of drawing lower values of match quality because they draw from a worse distribution.

REMARK. In the BRE of the economy, the unique solution to equilibrium condition (9) is

$$
(10) \qquad \theta(v, p) = \begin{cases} q^{-1}(k/(\beta(S_u(p) - v)) & \text{if} \quad \beta(S_u(p) - v) \geq k \\ 0 & \text{otherwise.} \end{cases}
$$

□

Since the function $S_u(p)$ is continuous in p for $p \in [0, 1]$, the market tightness function θ is continuous in p. Furthermore, since $V(v, p)$ is a decreasing function of v, $\theta(v, p)$ is a decreasing function of v. Intuitively, since firms have to pay higher wages to deliver the promised level of lifetime utility, their expected profits decline as v increases, so that a higher job-filling probability is required to pay for the cost of creating a vacancy, thus implying lower tightness in that submarket.

REMARK. $P(x, p)$ is continuous in x and p, $\forall x < \bar{x}$. Moreover, $\partial P / \partial p > 0$.

Continuity of the posterior belief P and the fact that it is increasing in p stems trivially from the functional form. However, first-order stochastic dominance is not sufficient to ensure that beliefs are monotone in x, which would require further restrictions on the shape of the distributions.

PROPOSITION 1. *The optimal choice of a worker can be written as the choice of a tightness level $\theta^*(p)$, $\forall p \in [0, 1]$, and is unique given p.*

PROOF. Notice that, since $f(\theta)/q(\theta) = \theta$, Equation (5) can be rewritten as:

$$
(11) \qquad U(p) = b + \max_{\theta} \{f(\theta)\beta(S_u(p) - U(p)) - \kappa\theta\} + \beta U(p),
$$

by using Equation (9) for submarkets with positive tightness.

Given p, this is a well-defined concave problem in θ because, by the assumptions on the matching function, $f(\theta)$ is an increasing, concave, and twice-differentiable function. Thus, the derivative of the value function with respect to θ is

$$
\frac{\partial U}{\partial \theta} = f'(\theta)\beta(S_u(p) - U(p)) - \kappa,
$$

which admits a zero for $\theta > 0$ if $\beta(S_u(p) - U(p)) > 0$.   □

LEMMA 2. *Let $TU(p)$ denote the right-hand side of Equation (11). T satisfies Blackwell's suf-ficient conditions, and thus, it is a contraction admitting a unique fixed point.*

PROOF. See online Appendix A.5.   □

PROPOSITION 2. *Under the assumption that $H(x) < L(x)$ $\forall x \in [0, \bar{x})$, U is strictly increasing.*

PROOF. See online Appendix A.5. The intuition is that a higher résumé p translates into higher expected productivity, which, in turn, leads to a higher expected job-finding rate for every promised lifetime utility, thus to a higher value of unemployment U. It also means that workers will trade off higher job-finding rates for lower promised lifetime utility when choosing the submarket in which to search.   □

PROPOSITION 3. $\theta^*(p)$ *is strictly increasing in p, and therefore, job-finding rates $f(\theta^*(p))$ are increasing in p.*

PROOF. See online Appendix A.5. The intuition is that, as unemployment entails the risk of not finding a job, the difference between the values of employment and unemployment is increasing in the résumé of the worker, and therefore workers with better résumés can find jobs faster.   □

COROLLARY 2. *The ex ante probability of separation to unemployment after discovering match quality $\mathbb{E}(d \mid p, type)$ is increasing in p for every type.*

This result follows from the fact that U is increasing in p. Using Lemma 1, the ex ante separation probability (without considering exogenous separations) can be written as

$$\mathbb{E}(d|p, \text{type}) = \begin{cases} \int \mathbb{I}[x < (1-\beta)U(p'(p,x))]dL(x) & \text{if type} = L \\ \int \mathbb{I}[x < (1-\beta)U(p'(p,x))]dH(x) & \text{if type} = H \end{cases},$$

where $\mathbb{I}[x < (1-\beta)U(p'(p,x))]$ is an indicator function taking value 1 if x is lower than $(1-\beta)U(p'(p,x))$ and 0 otherwise. Because posterior beliefs $p'(p,x)$ are an increasing function of p and $U(p)$ is an increasing function of p, it follows that, for every type, a better résumé will translate to a higher probability of separation, ceteris paribus. This is because, with a higher résumé, the outside option will be higher or equal for any draw of x, leading to a separation with a higher probability. Thus, workers with a better résumé anticipate that they can find better jobs on average, so they will separate more often when they make a bad draw. This seems to suggest, counterintuitively, that high-type workers might separate more often from their jobs. However, whether low types have higher job-separation rates than high types also depends on how dispersed the distributions of match quality are for each type, and on how often each type makes draws that are below the value of unemployment, as discussed in Corollary 1.

It is not possible to sign separation rates depending on worker types unequivocally, because separation rates are determined by two opposing forces: the probability of drawing low values of match quality, and the option value of making new, better draws in future jobs. Nonetheless, since job-finding rates are unequivocally increasing in résumés instead, I argue that these two separate channels driving heterogeneity in separation rates give more flexibility to the model, allowing types to experience separation rates that depend both on their match quality draws and on their labor market histories.

4. QUANTITATIVE MODEL AND IDENTIFICATION

The model is identified by estimating parameters to replicate features of the job-finding rates and job-separation rates observed in the NLSY/79. The idea behind identification is that the concentration and persistence of unemployment, and the differences in job-separation and job-finding rates between the different parts of the prime-age unemployment distribution, are informative of the quantity of low-type workers present in the economy and the differences between the match quality distributions of each type. This strategy is partly inspired by Menzio and Shi (2011) and Menzio et al. (2016), who use the life cycle patterns of job-finding rates,
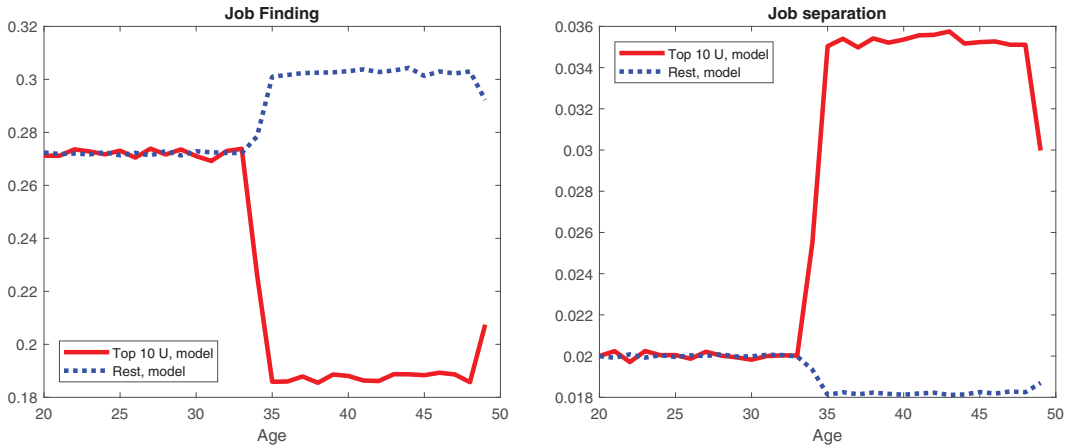
NOTE: The figure illustrates that all rates by prime-age unemployment groups exhibit discontinuities, despite average rates being continuous in age.

FIGURE 2

JOB-FINDING RATES (LEFT PANEL) AND JOB-SEPARATION RATES (RIGHT PANEL), OF THE TOP 10% PRIME-AGE UNEMPLOYED (RED, CONTINUOUS) VERSUS THE REST (BLUE, DOTTED) IN A HOMOGENEOUS WORKER MODEL WITHOUT MATCH QUALITY RISK
[COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

job-separation rates, and employment-to-employment transitions to identify the parameters of the match quality distribution and the probability of observing productivity during a match. My model works similarly to the aforementioned models during a match's duration; therefore, I apply the same strategy but I distinguish between the job-finding/job-separation rates experienced by the top 10% of the prime-age unemployed and the rest of the population.

It is important to notice the job-finding and job-separation rates of the top 10% prime-age unemployed and the rest exhibit sharp discontinuities around age 35. This is a mechanical consequence of looking at the outcomes of the top 10% after age 35 versus the rest of the population: Since the top 10% are exactly those who experienced more unemployment, they have higher separation rates and/or lower job-finding rates by construction starting at age 35. Intuitively, the most unemployed workers must, by definition, experience either a lower job-finding rate or a higher job-separation rate on average. In practice, they will experience a mixture of both. Similarly, due to selecting away the top 10% unemployed, the other workers will either have higher job-finding rates or lower job-separation rates on average. Thus, a mechanical divergence in at least one rate emerges between the two groups around age 35. This does not occur before to the same extent because, although the top 10% unemployed in prime age are exactly the most "unlucky" of workers after that age, that is not necessarily true before, as long as prime-age unemployment is not perfectly predictive of worker heterogeneity.

Figure 2 shows that, even in a standard random search model with no heterogeneity and no uncertainty about match quality, job-finding and job-separation rates by prime-age unemployment groups exhibit a discontinuity around age 35.[28] It is easy to see that, in a model with homogeneous workers, differences by prime-age unemployment groups do not imply any difference in young age, because they are not symptomatic of anything other than "having been unlucky after age 35." All rates by prime-age unemployment groups after age 35 are clearly biased estimates of the true, underlying job-finding and job-separation probability of workers. Comparing the model-generated statistics with the data allows to correct for this bias during estimation.

---

[28] Results are robust to changes in average rates. Differences across prime-age unemployment groups naturally become smaller, but are always present, as the probability of finding or losing a job increases, because this reduces inequality in the risk of unemployment.

To see how the match quality distribution affects job-separation rates, consider Corollary 1. Given a résumé p, the probability that a match is destroyed is at least $H(b)$ for high types and $L(b)$ for low types. Therefore, the way in which the probability mass is distributed over match qualities determines separation rates for the different types, which are also heterogeneous in their labor market histories because of the effect of the outside option (see Corollary 2). In general, all types will experience job-separation rates that decline with age because of a luck effect; the more jobs they draw, the higher the likelihood that workers find a job where they are productive.

Turning to job-finding rates, Proposition 3 establishes that job-finding rates are increasing in p. Since Proposition 3 also establishes that $U(p)$ always grows slower than $S_u(p)$, the larger the difference in expected productivity between types, the larger the difference in job-finding rates between types will be.

In summary, the features of match quality distributions will determine jointly job-finding rates and job-separation rates of types. Distributions that feature a high mass on low values of match quality, but a long right tail, can deliver high separation rates and high job-finding rates. On the other hand, distributions featuring a high mass on low values of match quality and a short right tail will deliver high separation rates and low job-finding rates. Finally, concentrated distributions, such that uncertainty about match quality is low, will deliver low separation rates.

The concentration and persistence of unemployment is informative of the quantity of low-type workers and the match quality distributions. Consider a case in which workers have the same starting résumé p (the population prior), and the match quality distribution of low types has more probability mass on realizations below b than the match quality distribution of high types. In other words, young low-type workers at the start of their career may experience a higher-than-average number of separations during their youth, depending on how their match quality distribution compares to the flow value of unemployment and how likely they are to draw a good match. As information on their type accumulates, their outside option declines, thus reducing their probability of separation, ceteris paribus. However, they may still experience higher separation rates due to their poor match quality distribution. Moreover, they will experience lower job-finding rates because their expected productivity is lower.

The model also features a role for luck in determining labor market histories. If the distribution of high types has enough common support with low types, high types who are unlucky enough to draw several low-quality matches will experience frequent separations as a consequence of their bad draws, and will be considered "low type" with a high probability. As a result, they will experience lower job-finding rates later on.

Young-age separation rates depend on how fast output is observed (as captured by the parameter π), whereas the speed of learning also depends on how far apart the two match quality distributions are. Thus, if types draw match qualities from very similar distributions, learning will be slow, job-finding rates will be similar across types, and there will be less concentration of unemployment. If the two types draw from very different distributions, learning will be fast and unemployment will be concentrated among a small number of workers. The scale and shape of the match quality distributions thus influence the life cycle profile of job-finding rates and job-separation rates. Notice that it is possible to obtain concentration of unemployment even with only one type of worker, simply by altering the features of the match quality distribution. However, this would be inconsistent with the fact that unemployment is persistent over the life cycle and with the life cycle patterns of job-finding rates and job-separation rates by unemployment group.

Persistence of unemployment depends on how far apart the two match quality distributions are, how risky they are, and how large the measure of low-types is. If low types always have a high risk of being unemployed (i.e., of drawing low match quality values) while high types are almost never unemployed, persistence will be high and will be determined almost uniquely by the measure of low types. To see this, suppose that low types constitute 10% of workers and high types are very seldom unemployed. In that case, persistence as measured by the probability of ending up in the top 10% of prime-age unemployment, given that the worker was in the top

10% of young unemployment, will be close to 100%. This is because the unemployment pool comprises mostly low types, both at young and old ages. However, looking at persistence among the top 20% would yield only about 50%, since the rest of the population is seldom unemployed and is unlikely to be among the most unemployed both when young and in prime age.

Conversely, if the two distributions of match quality are similar, persistence of unemployment will also depend on the speed of learning and on the role of luck in determining unemployment for both types. In all cases, however, the persistence of unemployment over the life cycle can be used to pin down the measure of low-type workers in the economy.

4.1. *Estimation.* In this section, I simulate the lives of a large sample of workers to compute lifetime statistics and estimate the model to replicate as closely as possible the observed patterns of transition rates in the NLSY/79. Estimation is performed by applying the simulated method of moments. I minimize the loss function

$$(12) \qquad\qquad L(\omega) = m(\omega)' W m(\omega),$$

where ω is the vector of parameters, $m(\omega)$ is a column vector of the differences between the model-generated moments and the data moments, and W is a weighting matrix.[29]

I set the model period to be one month. I assume that workers are born at age 20 (the starting age of the data) and I choose the death probability λ so as to match an average working life of 40 years. I choose the interest rate r as to give a compounded annual interest rate of 4%.

In line with numerous other models of directed search (such as Shimer, 2005; Mortensen and Nagypal, 2007; Menzio and Shi, 2011; Menzio et al., 2016), I restrict the matching probability to be of the form $f(\theta) = \min\{\theta^{0.5}, 1\}$.

The flow value of unemployment b is considered to include both the value of leisure and of unemployment benefits, and is chosen so as to match a ratio of 0.7 between b and average wages,[30] which is in line with the estimates of Hall and Milgrom (2008).

The two match quality distributions H and L are assumed to be Weibull distributions[31] with scale parameters σ_H, σ_L and shape parameters ϕ_H, ϕ_L. The shape and scale of match quality distributions, the probability π of observing a worker's output, the random separation probability δ, and the measure of low-type workers l are identified by the observed patterns of job-finding rates, job-separation rates over the life cycle for the top 10% unemployed and the rest of workers, and the observed concentration and persistence between young and prime-age unemployment for the top 10% and top 20%, as presented in the tables in Section 2. The estimation algorithm targets piecewise polynomially smoothed[32] patterns of job-finding and job-separation rates over the life cycle, but results are essentially identical using raw age averages instead. Since the model is unit-free, one of the scales has to be set exogenously. Therefore, I normalize $\sigma_H = 1$.

---

[29] Computation of the variance–covariance matrix of moments and standard errors is not trivial, since one of the moment restrictions is based on the estimates of Hall and Milgrom (2008) and its covariance with the remaining moments cannot be computed. Currently, W is set in such a way that all moments are scaled to their data average. In other words, I minimize the sum of the square differences $\frac{m(\omega)}{\hat{m}}$, where $\hat{m}$ are the data moments. In this way, I minimize the sum of relative distances from data averages. Also, singleton moments that are not part of life cycle patterns are weighted 10 times more to partly compensate for the fact that they are singleton moments instead of yearly patterns. Results do not vary significantly when this relative weight is altered.

[30] Computing the ratio between b and wages requires the computation of wages implied by the choices of workers in terms of lifetime utility v. For estimation purposes, I assume that wages are determined by Nash bargaining under the assumption that the Hosios condition holds. Although this might be an important assumption, the model is identified even without any restriction on b because the algorithm fits job-finding and job-separation rates over the life cycle for two different groups of workers. This restriction is meant to provide more discipline for the value of unemployment: The results are substantially identical if it is removed.

[31] The Weibull distribution is a common choice in this regard (see, e.g., Menzio et al., 2016).

[32] The polynomials are allowed to have different parameterizations before and after age 35 to account for the post-35 selection bias induced by dividing groups based on prime-age unemployment.

TABLE 5
BASELINE CALIBRATION RESULTS: TARGETS CALCULATED ON NLSY/79

Parameter	Symbol	Value	Target	Data	Model
Scale param. high type	σ_H	1.0000	Normalization		
Scale param. low type	σ_L	0.6271	JS rate profiles,		
Shape param. high type	ϕ_H	3.4261	JF rate profiles,		
Shape param. low type	ϕ_L	2.2427	% prime-age U acc. by top 10	0.59	0.53
Prob. of obs. output	π	0.0395	% prime-age U acc. by top 20	0.83	0.75
Flow value of Unemp.	b	0.7225	Ratio b/avg. wage	0.71	0.70
Vacancy cost	κ	3.2004	JF rate, bottom 90, age 20–25	0.25	0.28
Random sep. rate	δ	0.0022	JS rate, bottom 90, age 40–50	0.006	0.006
Measure of low types	l	0.2593	Persistence of U, top 10	0.39	0.32
			Persistence of U, top 20	0.43	0.38

The vacancy creation cost κ is chosen to match the job-finding rate of bottom 90% of the prime-age unemployment distribution at ages 20–25.

The estimation table reports only singleton targets. The patterns of job-finding and job-separation rates, which are vectors, are shown later in graph form for readability. Overall, the estimation algorithm fits eight parameters with 125 restrictions, of which 120 are the job-finding and job-separation rate patterns over the life cycle and the remaining restrictions are shown in Table 5.

5. RESULTS

5.1. *Estimation Results.* Despite being estimated with overidentifying restrictions, the model does a very good job in replicating the main features of the data. As can be seen in Table 5, the model is capable of delivering realistic amounts of concentration and persistence of unemployment. The model falls only slightly short in delivering concentration of unemployment: the top 10% account for 53% of prime-age unemployment in the model as compared to 59% in the data, whereas the top 20% account for 75% of prime-age unemployment in the model and more than 80% in the data. This is a drastic improvement over the standard Mortensen–Pissarides model, which only predicts half of observed concentration for the top 10%. I discuss why the model cannot get closer to the concentration observed in the data in a later paragraph.

The model gets also close to the levels of persistence of unemployment observed in the data as measured by the Markov transition matrix between young unemployment and prime-age unemployment. The probability of being in the top 10% of the prime-age unemployment distribution, after having been in the top 10% of the young unemployment distribution, is 0.32 in the model and 0.39 in the data. For the top 20%, the model yields 0.38 against 0.43 in the data. As I will show in a later paragraph, models without unobserved heterogeneity can only deliver a little above 10% and 25%, respectively.

The match quality distributions of high- and low-type workers are substantially different: that of low types has more mass close to zero and a lower mean than that of high types (Figure 3).

The estimated value of the probability of a firm observing the worker's output ($\pi = 0.0395$) implies that the average duration of a "bad match" is about 25 months. Quantitatively, this is a lower probability of separation than what is observed in other data sets such as the Current Population Survey,[33] although this is partly because gross flows observed in the NLSY/79 are smaller.

The estimated measure of low-type workers in the economy is around 25%. The model retains a role for luck as a determinant of labor market outcomes; indeed, the job-finding and job-separation rates among the top 10% are essentially those of the most unlucky among low types, as I will show below.

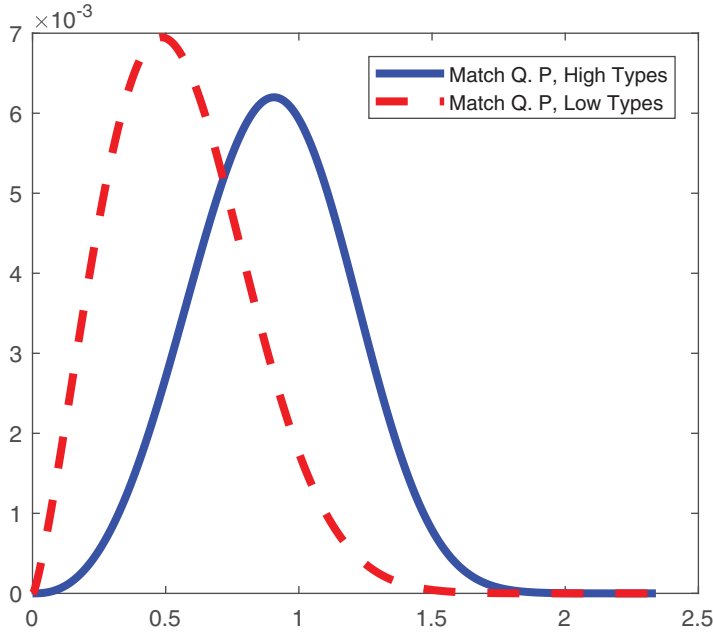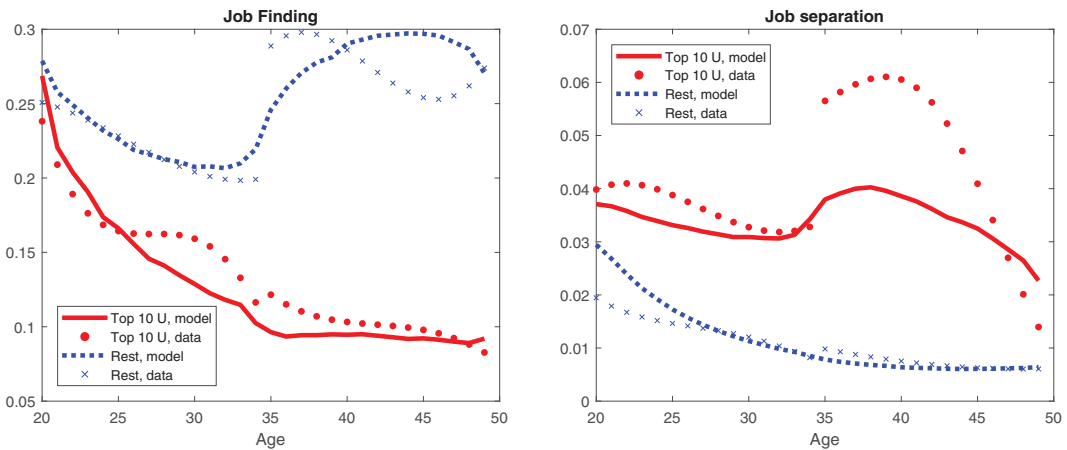[33] See, for instance, Shimer (2012).

FIGURE 3

NOTE: Results under baseline estimation.

FIGURE 4

MODEL VERSUS DATA: JOB-FINDING RATES (LEFT) AND JOB-SEPARATION RATES (RIGHT) OF THE TOP 10% PRIME-AGE UNEMPLOYED (RED) VERSUS THE REST (BLUE) [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

Figure 4 shows that the job-finding rate among the top 10% of prime-age unemployed declines over the life cycle as in the data, whereas the job-finding rate of other workers rises during prime age. As anticipated in the identification section, focusing on the labor market outcomes of the top 10% unemployed workers after age 35 generates a sharp discontinuity at that age in job-finding and job-separation rates by prime-age unemployment groups, both in the model and in the data. This is due to selection bias: mechanically, the top 10% are those who have either found fewer jobs or lost more jobs after age 35, but this is not necessarily the case before age
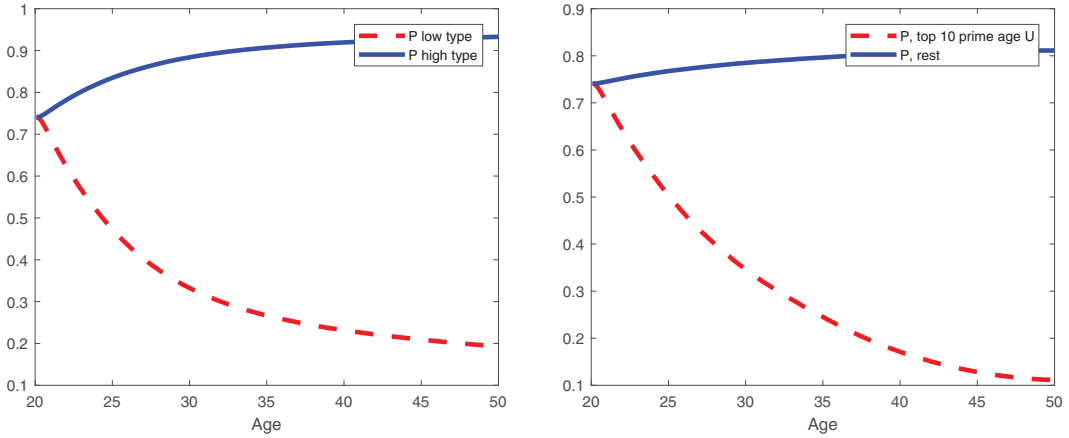
NOTE: Model results under baseline estimation.

FIGURE 5

PROBABILITY OF BEING HIGH-TYPE: BY HIGH/LOW TYPE (LEFT) AND BY TOP 10% OF PRIME-AGE UNEMPLOYMENT (RIGHT)
[COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]
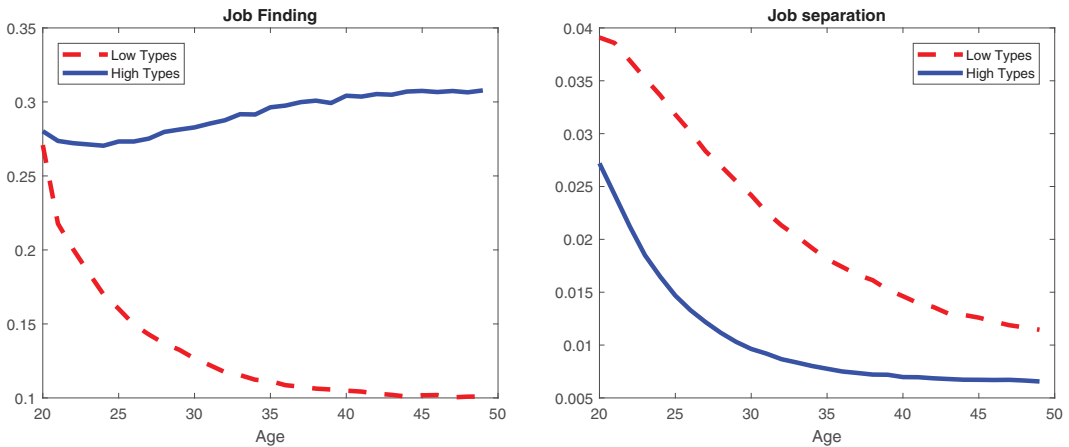
NOTE: Model results under baseline estimation.

FIGURE 6

JOB-FINDING RATES (LEFT) AND JOB-SEPARATION RATES (RIGHT) OF HIGH TYPES (BLUE, CONTINUOUS) VERSUS LOW TYPES (RED, BROKEN) [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

35, prior to the selection criterion. As I will show later, the model generates smooth job-finding and job-separation rates for every worker type.

The model is particularly successful at fitting the patterns of job-finding rates by prime-age unemployment groups, both for the most unemployed and the rest, and performs well in fitting job-separation rates, although it undershoots the job-separation rate of the most unemployed workers in prime age. It may be that the model cannot replicate the full amount of concentration observed in the data because it lacks on-the-job search. Thus, since workers are allowed to quit in the model, a higher probability π of observing match quality would imply a counterfactual pattern of job-separation rates for workers not in the top 10% unemployed. This is because high-type workers often quit, especially at the beginning of their career, to sample a better job (see also Figure 6).

The estimated match quality distributions also have implications for wages, although this is not part of the estimation strategy. In the model's environment, it is not possible to uniquely pin down wages without first specifying a contract space. Here, I analyze how the model performs under Nash bargaining (see Figure 15 in the online Appendix), which I show to be compatible with the choices of workers in terms of lifetime utility in online Appendix A.6.

Under Nash bargaining, the model gets remarkably close to the wage differentials between the top 10% of prime-age unemployed and other workers observed in the data. Differences in wages between the top 10% unemployed and other workers increase over the life cycle because the résumés of the former are worsening, whereas those of the latter are improving. Therefore, the gap in expected productivity widens. Moreover, when match quality is observed, low types typically draw lower values than high types, even conditional on remaining employed.

Figure 5 plots the average probability that a worker is of high type (the résumé) depending on her age, for low and high types and whether or not they are in the top 10% of the prime-age unemployment distribution. The graphs depict what I call "learning over the life cycle": As workers draw new values of productivity, the market gradually learns which workers are high-type and which are low-type. The patterns of job-finding and job-separation rates are a consequence of this mechanism. Notice that although average résumés converge relatively quickly over the life cycle, there is substantial variation in résumés across individuals even within types, as shown in Figure 14 in the online Appendix.

Consider first the job-finding rate: the gap between the 10% most unemployed workers and the rest widens over the life cycle because of two forces: First, as the market learns who are low- and who are high-type workers, the gap in job-finding rates between workers of different types widens. This can be seen by comparing job-finding rates and job-separation rates between high- and low-type workers in the model (Figure 6). Second, the low types' share in the unemployment pool increases with the age of workers (Figure 12 in the online Appendix). Thus, the job-finding rate of the top 10% unemployed is essentially the job-finding rate of the unluckiest of the low types: Indeed, the model predicts that 89% of the top 10% unemployed in prime age are low-type workers. Again, notice that the model does not feature any inherent discontinuity at age 35: All workers' labor market outcomes are continuous in age.

As in the case of job-finding rates, the job-separation rate of low-type workers is affected by learning over the life cycle. At ages 20–30, the job-separation rate of low-type workers is substantially higher than that of high types, since the former typically draw low values of match quality and face frequent separations.[34] However, both workers and the market learn during this stage, so that the low-type workers' outside option declines and they become progressively less likely to separate from a job. Moreover, a luck effect exists: sooner or later, every worker can find a job in which he is productive and remains there. The subsequent rise in separation rates observed for the top 10% of prime-age unemployed is due to selection bias: This empirical strategy is selecting the most unemployed individuals, who tend to be the unluckiest of the low-type workers.

The model is also capable of reproducing a duration-dependence relation in job-finding rates (see Figure 13 in the online Appendix), similar to that documented by Hornstein (2012) and Wiczer (2015). This relation arises due to a composition mechanism akin to Gonzalez and Shi (2010): Workers with higher market priors find jobs first, followed by workers with lower market priors.

5.2. *Counterfactual Simulations.*   In this section, I simulate alternative scenarios, by removing model features one by one to study their relative importance in fitting the data. Results are summarized in Table 6. All models have been reestimated on the same loss function as the baseline model.

---

[34] Carrillo-Tudela and Kaas (2015) show that similar dynamics occur in a model with adverse selection and on-the-job search.

TABLE 6
BASELINE CALIBRATION RESULTS VERSUS COUNTERFACTUALS

	(1) One Type $\mathbb{VAR}(x) > 0$ HC	(2) Two Types $\mathbb{VAR}_l = 0$ $\mathbb{VAR}_h = 0$	(3) Three Types $\mathbb{E}(x|l) =$ $\mathbb{E}(x|h)$	Baseline	Data
Separation rates, top 10					
age 20–25	2.83	0.90	3.16	3.60	4.04
age 25–30	2.80	0.71	2.86	3.23	3.62
age 40–45	3.91	1.57	3.40	3.66	5.50
Separation rates, rest					
age 20–25	2.38	0.74	2.48	2.41	1.70
age 25–30	1.67	0.68	1.64	1.46	1.37
age 40–45	0.49	0.59	0.61	0.62	0.70
Finding rates, top 10					
age 20–25	15.61	22.00	21.28	21.21	19.62
age 25–30	15.45	13.13	15.31	14.74	16.27
age 40–45	12.25	9.47	9.15	9.39	10.14
Finding rates, rest					
age 20–25	16.08	34.20	24.90	25.22	24.30
age 25–30	16.07	28.21	22.28	21.76	21.77
age 40–45	20.45	46.13	30.36	29.44	27.15
Concentration—Prime age					
U acc. by top 10	43.74	51.06	52.10	52.55	58.56
U acc. by top 20	72.83	70.37	76.43	74.52	82.66
Persistence					
Top 10 young—Top 10 primes	16.34	23.70	26.32	31.90	38.89
Top 20 young—Top 20 primes	29.61	41.54	34.93	37.81	42.59

NOTES: Column (1) is a model with no unobserved heterogeneity, and no uncertainty in match quality and fixed observable skills. Column (2) adds uncertainty in match quality and accumulation/depreciation of skills. Column (3) is a model with heterogeneity in average productivities, but no match quality uncertainty. Column (4) has heterogeneity in match quality uncertainty, but not in average productivities. All numbers are percentage points. All models have been recalibrated on the same loss function.

First, I estimate a version of the model featuring heterogeneity in match quality, on-the-job human capital accumulation, and stochastic human capital depreciation when unemployed, but no unobserved heterogeneity across workers in terms of a fixed type (column (1)). I find that such a model cannot deliver realistic amounts of concentration and persistence of unemployment. In particular, the predicted persistence of unemployment in the top 10% is little above 15% as opposed to 39% in the data. This is because human capital accumulation and depreciation introduce "reshuffling" in the skill level of workers, acting as mean-reverting forces. Thus, instead of having fixed differences in productivity, every worker can now become unskilled if he stays unemployed long enough or skilled if he manages to obtain a sufficiently high level of match quality. This negative impact of "reshuffling" of skills on persistence mirrors similar economic intuitions for the earnings losses from displacement in Michaud (2018) and Jung and Kuhn (2019). Furthermore, model 1 cannot replicate the pattern of differences in separation rates at young ages between unemployment groups, because it does not feature enough heterogeneity in match quality draws: The bottom 90% of workers have about the same separation rate as the 10% most unemployed at ages 20–30, as opposed to the 2.3% difference in the data. Finally, such a model cannot replicate the fact that the job-finding rate of the most unemployed falls in the early part of their career.

I now estimate a model without human capital nor uncertainty about match quality, but with heterogeneity in mean productivity across types (column (2)). In other words, I force the distribution of match quality to be degenerate. Such a model cannot predict the higher separation rate at ages 25–30 for one group of workers, because as soon as workers draw their first job, they immediately learn their type and give up looking for a better job forever. This

occurs because, without uncertainty about match quality, a worker will learn his type with certainty at the first observation of output. In short, learning is too fast and bad luck plays almost no role. Moreover, since the distribution of match quality is degenerate, there is no other mechanism that can deliver heterogeneity in separation rates. Such a model also predicts a decrease in job-finding rates that is too sudden and too large. This is another consequence of excessively fast learning: Job-finding rates immediately respond to the sharp change in the prior of workers. Moreover, this model can predict the amounts of concentration of unemployment observed in the data at the cost of producing counterfactual job-finding rates for most workers. This is due to the struggle of the model in delivering concentration by means of heterogeneity in job-finding rates only. Finally, this model falls short in delivering persistence of unemployment among the top 10% (23% as opposed to 39% in the data).

In the final experiment, I estimate a model with heterogeneity in the variance of the match quality distribution, but no differences in mean productivity (column (3)).[35] Such a model comes closer to the results of the baseline model, suggesting that variance in match quality draws is even more important than differences in the mean of draws. Nonetheless, it still predicts less persistence of unemployment and less heterogeneity in job-separation rates among the young than the baseline model.

These quantitative exercises confirm that all the ingredients are important for explaining the patterns observed in the data. Heterogeneity in the mean of match quality draws is important for explaining differences in job-finding rates. Heterogeneity in the variance of match quality is important for explaining heterogeneity in job-separation rates, for obtaining concentration of unemployment, and for slowing down learning at the start of the career. Slower learning is important since it translates into a more realistic decline in job-finding rates and job-separation rates among the most unemployed workers.

5.3. *Decomposing Learning over the Life Cycle.*   In this section, I maintain the baseline estimation but shut down information frictions by making types known right from the start. This exercise answers two questions. First, how important are information frictions in explaining the data? Second, how much would workers, and the economy in general, benefit from knowing their types early on? Results are presented in Table 7.

When types are known from the start, the separation rate of the most unemployed at ages 20–30 is reduced by about half a percentage point (one-fifth). This occurs because these workers are already aware that they are low-type. Therefore, their outside option is already low and, as a result, they do not separate from jobs after lucky draws. The concentration of prime-age unemployment is hardly affected, since it is mainly due to heterogeneity across workers and bad luck in drawing match quality values.

Information frictions account for the entire decline in job-finding rates from age 20 to 40 for the most unemployed workers: If types were already known, firms would anticipate their low average productivity and these workers would find jobs with lower probability right from the start. Finally, the persistence of unemployment increases if types are known from the start. This result is because the most unemployed workers in prime-age experience only slightly lower separation rates, but substantially lower (-56%) job-finding rates when young, making them more likely to be unemployed during the initial years of their career.

Overall, the economy benefits from information frictions being removed: welfare, as measured by total output produced from the entry of workers in the labor market until age 50, minus vacancy-posting costs, increases by 1.22%. Most of this change is accounted for by the fact that high-type workers are now offered more jobs, so that they match faster and therefore increase

---

[35] In practice, this is done by letting the shape parameters ϕ_i of the Weibull distributions be estimated freely by the algorithm, whereas σ_l solves the nonlinear equation

$$\mu_h = \mu_l,$$

where $\mu_i = \sigma_i \Gamma(1 + \frac{1}{\phi_i})$ is the mean of the Weibull distribution and $\Gamma(x)$ is the Gamma function.

TABLE 7
BASELINE RESULTS VERSUS COUNTERFACTUALS

	Types Known	Baseline	Difference
Separation rates, top 10			
age 20–25	2.90	3.60	−19.49
age 25–30	2.40	3.23	−25.70
age 30–35	2.47	3.17	−21.97
age 40–45	3.34	3.66	−8.78
Separation rates, rest			
age 20–25	2.39	2.41	−0.96
age 25–30	1.45	1.46	−0.38
age 30–35	0.99	0.99	0.14
age 40–45	0.61	0.62	−2.28
Finding rates, top 10			
age 20–25	9.19	21.21	−56.68
age 25–30	8.94	14.74	−39.40
age 30–35	8.84	11.77	−24.86
age 40–45	8.85	9.39	−5.69
Finding rates, rest			
age 20–25	21.57	25.22	−14.48
age 25–30	21.33	21.76	−1.96
age 30–35	22.54	21.17	6.43
age 40–45	31.48	29.44	6.95
Concentration—Prime age			
U acc. by top 10	54.38	52.55	3.48
U acc. by top 20	75.54	74.52	1.36
Persistence			
Top 10 young—Top 10 primes	39.76	31.90	24.65
Top 20 young—Top 20 primes	44.00	37.81	16.38
Welfare			
Value at $t = 0$, low types	76.68	100.00	−23.32
Value at $t = 0$, high types	109.90	100.00	9.90
Welfare	101.22	100.00	1.22

NOTE: Column (1): Types are known from the beginning. Column (2): Baseline model. Column (3): Percentage difference between columns (1) and (2). All numbers are percentage points.

employment and output. Although the majority of workers benefit from types being known, a fraction suffers. High-type workers increase their ex ante lifetime utility by around 10%, although low-type workers see a decline of about 23%. This is because, while high-type workers benefit from types being known by finding jobs faster, exactly the opposite happens to low-type workers, who forgo the initial part of their lives in which the fact that their type is not known plays in their favor, both in terms of higher job-finding rates and in terms of better expectations about their future productivity.

5.4. *Résumés in the Model and in the Data.* If it were possible to measure résumés p in the data, the model predicts that in period t, one would find a positive relationship between résumés and job-finding rates, and a negative relationship between résumés and separation rates. The model also predicts that there exist measures that are strongly correlated with the résumé p of a worker. In this subsection, I investigate whether some particular measures have similar predictive power both in the model and in the data.[36]

Two such measures inspired by the theory are the empirical average probability of separation experienced by each individual worker up to period $t - 1$ and the duration of the last unemployment spell. The two measures can jointly account for 22% of the variation in résumés observed in the model. The first measure is informative because low types are more likely to experience

[36] Notice that, in the model, the résumé cannot be measured by looking only at the employment/nonemployment history of the worker: The résumé includes information on match quality that is unobservable in the NLSY/79 data.

TABLE 8
REGRESSIONS OF JOB FINDING (COLUMNS (1) AND (2)) AND JOB SEPARATION (COLUMNS (3) AND (4)) EVENTS ON THE
PREVIOUSLY OBSERVED PROBABILITY OF SEPARATION AND THE DURATION OF THE LAST UNEMPLOYMENT SPELL, ON NLSY/79
DATA (COLUMNS (1) AND (3)) AND MODEL-GENERATED DATA (COLUMNS (2) AND (4))

	(1) JF (Data)	(2) JF (Model)	(3) JS (Data)	(4) JS (Model)
Empirical individual probability of separation	−0.0545***	−0.117***	0.114***	0.148***
	(0.003)	(0.008)	(0.002)	(0.003)
Last unemployment duration	−0.000738***	−0.00460***	0.0000128*	0.000151***
	(0.000)	(0.000)	(0.000)	(0.000)
Sample average	0.161	0.206	0.012	0.012
R^2	0.028	0.012	0.029	0.004
Observations	86,790	446,840	953,955	6,961,398

Standard errors in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

separations and thus to have poorer résumés (lower p). The second is informative because high types are more likely to have better résumés (higher p) and thus find jobs faster. The last unemployment duration is more informative about the current résumé than the previous ones, because it incorporates all information up to the unemployment spell.

I test the informational content of these measures by regressing monthly job-finding and job-separation events at period t against the ex post average separation rate at the individual level and the duration of the last unemployment spell (Table 8), both in the model and in the data. In these regressions, a third-degree polynomial in age and yearly dummies are added as controls. The age polynomial is particularly important because job-separation rates are higher for young workers both in the model and in the data, and this can create an artificial correlation between past job-separation rates at the individual level and the probability of separation today.

The results of this exercise confirm that measures of past job-separation events and past unemployment duration are strongly correlated with the current probability of finding a job and separating from one, both in the model and in the data. All coefficients have the same sign both in the model and in the data, although their magnitudes differ. As expected, those who experienced a higher likelihood of transitioning from employment to unemployment in the past experience lower job-finding rates and higher job-separation rates, both in the model and in the data.

Let me discuss past separation rates first. For every 1 percentage point (p.p. from now on) higher past separation rate, the probability of finding a job declines by approximately 0.05 p.p. in the data and 0.1 p.p. in the model. The difference in magnitude can be partly explained by the fact that the model fails to deliver as high a separation rate for the top 10% prime-age unemployed as that observed in the data, resulting in a difference in scale between the data and the model. However, the impact of past separation events on the probability of finding a job appears to be small, relative to the average job-finding probability of about 16 percentage points. Even a 10 p.p. higher past job-separation rate translates into only a 1 p.p. lower job-finding rate. However, the predictive power of past separation rates for present separation rates is relatively stronger: a 1 p.p. higher past separation rate predicts a 0.1 p.p. higher separation rate at time t, both in the model and in the data, as compared to an average separation rate in the data of about 1.3 percentage points.

With regard to unemployment duration, its power to predict job-finding rates is similar: One more week of unemployment duration predicts a 0.07 p.p. lower job-finding rate in the data and 0.4 p.p. in the model. Again, this discrepancy may be due to the fact that although unemployment duration is strongly tied to résumés in the model, there might be other sources of variation in the data (e.g., measurement error, human capital, or preferences) that are not accounted for in the model. Interestingly, the relationship between past unemployment duration and job-separation rates is extremely weak both in the model and in the data.

The predictive power of these variables seems quantitatively small, but the purpose of the exercise is not to strongly predict period-by-period outcomes by means of résumé measures, but rather to show that such plausible measures have a similar correlation with labor market outcomes both in the model and in the data. In fact, all such measures contain significant measurement error even in the model, and as such their coefficients are downward-biased even for predicting résumés. For instance, observing several separations might not be a signal that a worker is low-type. In fact, he could be a high-type worker who found jobs that indeed correctly signal his type, but which were not productive enough to be to his liking. Moreover, résumés themselves are error-contaminated measures of a worker's type. Thus, although the model ties them strongly to job-finding rates, which is not the case for job-separation rates. Finally, although the quantitative relevance of the considered variables in predicting labor market outcomes seems small in terms of R^2, it is important to consider that all period-by-period events still retain a strong random component, and as such exhibit more variation than what can be explained with observables, even in the model. For instance, although the type of the worker is the main determinant of job-separation rates in the model, it can account for less than 1% (in terms of pseudo-R^2) of the variance in a probit regression of job-separation events on worker type.

6. DISCUSSION

I have shown that a theory of information frictions and heterogeneity is capable of jointly explaining the patterns of job-finding rates and job-separation rates by unemployment group over the life cycle. An alternative explanation might be that workers who are often unemployed tend to lose, or fail to accumulate, human capital because they lack on-the-job training and experience human capital depreciation (as in Ljungqvist and Sargent, 1998). To the extent that human capital is observable, if workers started with some level of human capital, depreciation would lead the most unemployed workers to experience lower job-finding rates, possibly explaining one of the facts. However, even if lower human capital yielded higher separation rates, accumulation and depreciation would imply that heterogeneity in job-separation rates *rises* over the career, because the most unemployed would lose human capital (or fail to accumulate it) and would possibly face higher separation rates, whereas the rest of the workers would experience fewer separations. In that case, we should observe similar separation rates by prime-age unemployment group at the start of the career, and divergence subsequently, whereas in the data, differences in separation rates are large right from the beginning.

Column (1) of Table 6 partially tests for these implications by estimating a version of the model that features homogeneous workers, retains risk in match quality draws, and adds human capital accumulation/depreciation as a source of persistence in unemployment. Figure 11 in the online Appendix shows the behavior of the baseline model with heterogeneity (column (4) of Table 6), compared to the model without heterogeneity and human capital accumulation (column (1)). The latter delivers almost no young to prime-age persistence of unemployment, and predicts patterns of job-finding rates and job-separation rates by prime-age unemployment groups that are inconsistent with the data.

One way to make the human–capital model consistent with the data is to give a special importance to the first job. A model in which a separation in the first job adversely affects the rest of the career for some workers would produce patterns of job-finding and job-separation rates that are very similar to those found in the data. However, it seems hard to motivate such a role for the first job without having some difference in fundamentals to drive the results.

7. CONCLUSIONS

Using NLSY/79 data, I show that unemployment during prime age is concentrated among relatively few workers, who experience both long spells of unemployment and frequent job separations. Moreover, unemployment is persistent in the sense that those who were often

unemployed when young tend to be often unemployed during their prime. I build a model that delivers both high concentration of unemployment during prime age and persistence of unemployment over the life cycle and is consistent with the patterns of job-finding rates and job-separation rates by prime-age unemployment groups. This is accomplished by means of a combination of incomplete information and heterogeneity across workers. I find that information frictions are important for explaining workers' labor market outcomes at the beginning of their career. In particular, a model without information frictions cannot match the heterogeneity in separation rates among young workers observed in the data, nor the persistence of unemployment over the life cycle.

The findings have a number of implications for labor market policy. First, the estimation results suggest that a significant fraction of workers experience more job instability than the rest over the entire life cycle, and therefore are likely to be the most affected by changes in unemployment insurance and employment protection policies. For instance, increases in the degree of employment protection would asymmetrically decrease the job opportunities of the two groups. Although the most stable group might be almost unaffected by the change, job creation for workers with poor résumés would be more strongly affected, since employers would anticipate that a match with these workers would likely terminate. Second, changes in labor market policies might change the speed of learning in the market in several ways, the most direct of which being their effect on job-finding rates. For instance, increasing unemployment insurance might increase unemployment duration, and thus, the number of jobs a worker is able to sample and add to his résumé. However, depending on how contracts and learning work in the market, it is also possible that matches that would otherwise have survived would be destroyed, or vice versa, thus affecting the amount of information a résumé conveys. For instance, consider a case in which employers infer a worker's type from his past employment durations. If there are large firing taxes, employers might decide to keep relatively unproductive workers, thus conveying a different message to future prospective employers. On the other hand, in this same environment, a separation would convey a stronger negative signal, because it would mean that a firm preferred to pay the firing tax instead of maintain the match. Although the model developed in this article is not suited to answering this particular question (due to the assumption of complete contracts), an extension of the model that allows for alternative contract spaces might be able to shed light on this and other related questions.   □

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table 9: Regression of % of unemployment in prime age (35–50) on % of young unemployment (20–30): column (1) reports results from a regression on all workers, including only controls and no young unemployment; (2) includes young unemployment; and (3) is the same as (2) but only for high-school educated workers
Table 10: Left column: equally weighted averages
Table 11: Participation-adjusted averages computed on NLSY/79, individuals aged 35–50
Table 12: Summary statistics by parts of the prime-age (35–50) unemployment distribution and by education subgroups
Table 13: Accounting for possible measurement error: concentration and persistence of unemployment according to alternative definitions of the sample
Table 14: Regression of % of unemployment in prime age (35–50) on % of young unemployment (20–30): column (1) reports results from a regression on all workers, including only controls and no young unemployment; (2) includes young unemployment; and (3) adds dummies for the occupation in which the individual spent most time
Table 15: Weekly probability of job termination, by reason and group of prime-age unemployment

Figure 7: Labor force participation rate by prime-age unemployment groups and by five-years age groups.
Figure 8: Log difference in hourly wage, between top 10% prime-age unemployment group and the rest; obtained as coefficients from a regression that controls for education, region of residence, ethnic group, year, and marital status
Figure 9: Job-finding (left panel) and job-separation (right panel) probabilities, by group of prime-age nonemployed, using alternative definition of nonemployment
Figure 10: Job-finding (top panels) and job-separation (bottom panels) probabilities, by group of prime-age unemployed
Figure 11: Comparison between baseline model and human capital model without heterogeneity; job-finding (left panels) and job-separation rates (right panels), by prime-age unemployment groups
Figure 12: Share of unemployed workers who are low types, by age, under baseline estimation
Figure 13: Model-generated data: duration-dependence relation in job-finding rates, at 1, 3, 6, and 12 months of unemployment duration
Figure 14: Model-generated data: standard deviation of résumés, by age, under baseline estimation
Figure 15: Difference in wages between top 10 % prime-age unemployed and rest; data (dashed) versus model (continuous) under baseline estimation and Nash Bargaining wage determination
Figure 16: Model-generated data: wages by type and age, under baseline estimation
Figure 17: Model-generated data: standard deviation of wages by résumé, under baseline estimation

REFERENCES

ADDISON, J. T., AND P. PORTUGAL, "Job Displacement, Relative Wage Changes, and Duration of Unemployment," *Journal of Labor Economics* 7 (1989), 281–302.
ATKINSON, A. B., "On the Measurement of Inequality," *Journal of Economic Theory* 2 (1970), 244–63.
BROOKS, B., *Chronic Unemployment: A Statistical Profile* (Ottawa, Ontario: Income Statistics Division and Socio-Economic Analysis and Modeling Division, 2005).
CARRILLO-TUDELA, C., AND L. KAAS, "Worker Mobility in a Search Model with Adverse Selection," *Journal of Economic Theory* 160 (2015), 340–86.
CHETTY, R., "Moral Hazard versus Liquidity and Optimal Unemployment Insurance," *Journal of political Economy* 116 (2008), 173–234.
CLARK, K. B., L. H. SUMMERS, C. C. HOLT, R. E. HALL, M. N. BAILY, AND K. B. CLARK, "Labor Market Dynamics and Unemployment: A Reconsideration," *Brookings Papers on Economic Activity* 1979 (1979), 13–72.
DOPPELT, R., "The Hazards of Unemployment: A Macroeconomic Model of Job Search and Résumé Dynamics," Technical Report, Penn State University, 2016.
FERNANDEZ-BLANCO, J., AND E. PREUGSCHAT, "On the Effects of Ranking by Unemployment Duration," *European Economic Review* 104 (2018), 92–110.
GIBBONS, R., AND L. F. KATZ, "Layoffs and Lemons," *Journal of Labor Economics* 9 (1991), 351–80.
GONZALEZ, F. M., AND S. SHI, "An Equilibrium Theory of Learning, Search, and Wages," *Econometrica* 78 (2010), 509–37.
GUERRIERI, V., R. SHIMER, AND R. WRIGHT, "Adverse Selection in Competitive Search Equilibrium," *Econometrica* 78 (2010), 1823–62.
HAGEDORN, M., AND I. MANOVSKII, "Job Selection and Wages over the Business Cycle," *American Economic Review* 103 (2013), 771–803.
HALL, R. E., AND P. R. MILGROM, "The Limited Influence of Unemployment on the Wage Bargain," *American economic review* 98 (2008), 1653–74.
HORNSTEIN, A., "Accounting for Unemployment: The Long and Short of It," Federal Reserve Bank of Richmond Working Paper, 2012.
HUGGETT, M., G. VENTURA, AND A. YARON, "Sources of Lifetime Inequality," *American Economic Review* 101 (2011), 2923–54.
JOVANOVIC, B., "Job Matching and the Theory of Turnover," *Journal of Political Economy* 87 (1979), 972–90.
JUNG, P., AND M. KUHN, "Earnings Losses and Labor Mobility over the Life Cycle," *Journal of the European Economic Association* 17 (2019), 678–724.

KAHN, L. B., AND F. LANGE, "Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures," *The Review of Economic Studies* 81 (2014), 1575–613.

KURUSCU, B., "Training and Lifetime Income," *American Economic Review* 96 (2006), 832–46.

LANGE, F., "The Speed of Employer Learning," *Journal of Labor Economics* 25 (2007), 1–35.

LINDQVIST, E., AND R. VESTMAN, "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment," *American Economic Journal: Applied Economics* 3 (2011), 101–28.

LJUNGQVIST, L., AND T. J. SARGENT, "The European Unemployment Dilemma," *Journal of political Economy* 106 (1998), 514–50.

LOCKWOOD, B., "Information Externalities in the Labour Market and the Duration of Unemployment," *The Review of Economic Studies* 58 (1991), 733–53.

MENZIO, G., AND S. SHI, "Efficient Search on the Job and the Business Cycle," *Journal of Political Economy* 119 (2011), 468–510.

———, I. A. TELYUKOVA, AND L. VISSCHERS, "Directed Search over the Life Cycle," *Review of Economic Dynamics* 19 (2016), 38–62.

MICHAUD, A. M., "A Quantitative Theory of Information, Worker Flows, and Wage Dispersion," *American Economic Journal: Macroeconomics* 10 (2018), 154–83.

MICHELACCI, C., J. PIJOAN-MAS, AND H. RUFFO, "Inequality in Unemployment Risk and in Wages," 2012 Meeting Papers 794, Society for Economic Dynamics, 2012.

MOEN, E. R., "Competitive Search Equilibrium," *Journal of political Economy* 105 (1997), 385–411.

MORTENSEN, D. T., AND E. NAGYPAL, "More on Unemployment and Vacancy Fluctuations," *Review of Economic dynamics* 10 (2007), 327–47.

———, AND C. A. PISSARIDES, "Job Creation and Job Destruction in the Theory of Unemployment," *The Review of Economic Studies* 61 (1994), 397–415.

MUSTRE-DEL RÍO, J., "Job Duration, Wages, and the Cleansing and Sullying Effects of Recessions," Working Paper, Federal Reserve Bank of Kansas City, 2012.

SCHMILLEN, A., AND J. MÖLLER, "Distribution and Determinants of Lifetime Unemployment," *Labour Economics* 19 (2012), 33–47.

SHIMER, R., "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review* 95 (2005), 25–49.

———, "The Probability of Finding a Job," *American Economic Review* 98 (2008), 268–73.

———, "Reassessing the Ins and Outs of Unemployment," *Review of Economic Dynamics* 15 (2012), 127–48.

WICZER, D. G., "Long-Term Unemployment: Attached and Mismatched?," Federal Reserve Bank of St. Louis Working Paper No. 2015-42, 2015.