Information Frictions, Match Quality and Lifetime Unemployment

Iacopo Morchio∗†

Draft: September 9th, 2014

Please see www.sites.google.com/site/iacopomorchio for the updated version

Abstract

We show that, in the US, unemployment during prime-age is concentrated in a small number of workers, that this is due both to longer unemployment duration and more frequent separations, and that young unemployment is a strong predictor of prime-age unemployment. The standard search model is at odds with these facts. We introduce a model which aim is to reproduce the concentration of unemployment and the observed heterogeneity in finding and separation rates across workers, and use a novel identification strategy based on lifetime statistics. The model allows to quantify the relative importance of information frictions and match quality for explaining wage patterns and the concentration of unemployment; results show that information frictions are responsible for about one-fifth of separations occurring during the first year of career, and for the relatively low wage differentials at the start of the career between workers with very different lifetime outcomes. Information frictions have little influence on the concentration and persistence of unemployment over the life cycle, which is mainly driven by strong heterogeneity across workers. Observable skills, distinct from unobservable heterogeneity, have a substantially lower impact on wage differentials and lifetime inequality in unemployment. The model also allows the evaluation of labor market policies.

Keywords: Concentration, Inequality, Learning, Sorting, Unemployment.

JEL Classification Numbers: J24, J64.

∗Universidad Carlos III de Madrid
†I am very grateful to Andrés Erosa and Javier Fernández-Blanco for their support and comments. Special thanks to my friend and colleague Salvatore Lo Bello for the many discussions on the topic. Thanks to Maia Guell, Fatih Guvenen, Moritz Kuhn, Ellen McGrattan, Sevi Mora, Fabrizio Perri, Omar Rachedi, José-Victor Rios-Rull for useful comments and suggestions.
1 Introduction

In the US economy, most of prime-age unemployment between 1985 and 2010 is accounted for by a relatively small fraction of workers: 10% of workers account for about two-thirds of observed unemployment between ages 35 to 50. This is both due to lower finding rates and higher separation rates for that top 10%. Moreover, time spent in unemployment when young is a powerful predictor of unemployment in prime age. The standard search-and-matching framework is at odds with these facts, because it features too many transitions in and out of the labor force for the majority of workers, and no source of persistence of unemployment outcomes over the life cycle.

This paper documents patterns of inequality in unemployment outcomes over the lifetime and introduces a model to make a sense of such patterns. The model is also a way to take care of the selection bias of computing lifetime statistics by parts of the unemployment distribution: while part of the differences in unemployment outcomes can be attributed to heterogeneity/human capital loss or other reasons, another part is due to the fact that this empirical strategy is selecting those individuals who experienced exceptionally high amounts of unemployment.

Our model is a theory of job finding rates and job separation rates based on two main ingredients: heterogeneity across workers and incomplete information about workers’ types. There are two types of workers, high and low, who extract match quality from two different distributions, and their type is initially unknown by firms. All firms in the market slowly learn workers’ types from workers’ labor market outcomes. We then estimate the model using data on job finding and job separation rates by parts of the prime-age unemployment distribution, and on the concentration and persistence of unemployment.

The mechanism works as follows. Search is directed in the sense that workers decide to search for a job with a certain wage: workers who are believed to be high types with higher probability will apply to higher wages. Firms have limited commitment to the matches they form, and can destroy a match at any point in time. Output of a match is unobserved, until an observation shock is realized: then, the firm decides to keep the worker or to destroy the match, leaving the worker unemployed. The occurrence of this shock is observed by the market, which updates the probability that the worker is of high type depending on the outcome of the observation event. Separations will lead the market to believe that the worker is more likely to be of low type, while continuations will have the opposite effect. As
a result, workers’ types will be slowly learned along their careers.

Estimation results show that the combination of progressive learning and heterogeneity delivers both concentration of unemployment in prime age and persistence of unemployment outcomes. It delivers concentration because low-type workers have a higher probability of extracting low quality matches than high-type workers, and have a lower expected productivity; thus such workers face a higher separation rate and a lower job finding rate at every age. It delivers persistence because low-type workers tend to experience frequent separations both when young and when prime-age, and job finding rates that decline with age as the market recognizes them as low-type workers.

We find that neglecting heterogeneity across workers makes it impossible to match the concentration and persistence of unemployment observed in the data, even in the presence of heterogeneous match quality. Information frictions play an important role in the first part of workers’ lives: the data show that wage differentials between workers with very different lifetime unemployment are relatively small at the beginning of their career, and our simulations show that information frictions are responsible for the initial small difference. Progressive learning, rather than human capital accumulation (or the lack of it) leads the wage gap between the always-unemployed and the rest to widen over the workers’ careers. We also estimate that about one fifth of observed separations in the first year of employment are due to incomplete information on workers’ types. Basically, since young low type workers are still not sorted from high type ones, such workers apply to excessively high wages that yield to more frequent separations once their match quality is discovered. However, the role of information frictions later in life is negligible, and most of the concentration and persistence of unemployment are due to heterogeneity across workers.

The model can also be used as a laboratory for investigating duration dependence in labor market outcomes. Low-type workers tend to find jobs with lower probability, thus giving rise to a duration dependence relation in job finding rates due to a sorting mechanism.

The paper is organized as follows. Section 2 discusses the related literature; section 3 presents data on the concentration and persistence of unemployment, and decompositions between finding and separation rates. Section 4 presents the model of heterogeneity and learning over the life cycle. Section 5 discusses the sources of variation used to identify the model. Section 6 presents results. Section 7 concludes.
A large literature investigates the composition of the unemployment pool and the reasons behind heterogeneity in job finding rates; Clark et al. (1979) were the first to show that most of unemployment is due to workers experiencing rather long spells of unemployment, rather than to workers going frequently in and out of unemployment. In this paper we make a different point and argue that most of the prime-age unemployment pool is composed by a relatively small group of workers continuously going out of employment, and staying unemployed for a long time.

The fact that the job finding rate declines with unemployment duration has been widely studied; see for instance Shimer (2008). Different explanations for duration dependence relations have been proposed by several authors. Gonzalez and Shi (2010) propose a model in which workers are heterogeneous in their ability to find jobs; Fernández-Blanco and Preugschat (2014) postulate that workers have different probabilities of forming productive matches, so that firms in equilibrium rank candidates by unemployment duration before screening workers. Wiczer (2014) develops a model of switches across occupations, in which skills are occupation specific, so that workers do not want to change occupation in case they remain unemployed. Our model features a duration dependence relation in a way similar to Gonzalez and Shi (2010) because workers who have a higher probability of being high types tend to find jobs faster.

The model we develop is, in spirit, a life-cycle model of search and matching. Menzio, Telyukova and Visschers (2012) is the closest model to the one presented in this paper, having in common directed search and job-specific match quality. However, while they want to provide a life cycle theory of the transitions in and out of unemployment and employment over the life cycle, we want to understand the sources of heterogeneity in lifetime unemployment, and study the role played by information frictions in determining lifetime outcomes. We find that models without heterogeneity and sorting (like Menzio, Telyukova and Visschers (2012) or Chèron, Hairault and Langot (2013)), despite featuring potential sources of persistence such as human capital accumulation, cannot replicate the amounts of concentration and persistence of unemployment we document.

This paper is also related to the literature studying adverse selection and search. In part the model works with a mechanism similar to Gibbons and Katz (1991): when firms find out that a worker has low productivity, they fire her. Wright, Guerrieri and Shimer (2010) show that, in directed search environments with in-
complete information, it is always possible to construct contracts that yield to separating equilibria. Their results do not apply here because in our model information is incomplete on both sides of the market, in the sense that neither workers nor firms know what the worker’s type is and have the same information about it.

To the best of our knowledge, macroeconomic models have dedicated relatively little space to information frictions as a source of heterogeneity in labor market outcomes across workers. A notable exception is Michaud (2014), who proposes a theory of information frictions as a candidate explanation for the wage scars of unemployment, and whose model has some features in common with the one we present.

3 The Data

We use weekly job histories from NLSY/79 data to compute lifetime unemployment statistics. The NLSY is one of the best-known panel data available for the US, following a cohort of more than ten thousand individuals from 1979 onwards. Those who are being followed in the NLSY79 ranged ages 14 to 22 in 1979; information has been gathered annually until 1994, and biennially since then.

We first document that prime-age unemployment is concentrated in relatively few individuals. We start by defining young-age unemployment as the fraction of the work history an individual spent in unemployment, over total weeks employed or unemployed\(^1\) from age 20 to 30:

\[
\bar{u}_y^i = \frac{\sum_{t=1}^{T_y^i} u_{i,t}}{T_y^i}
\]

where \(u_{i,t}^y\) is a variable taking value 1 in weeks in which individual \(i\) was unemployed, and 0 if individual \(i\) was employed, and \(T_y^i\) is the number of weeks that individual \(i\) was either employed or unemployed between ages 20 and 30. Similarly, we define prime-age unemployment as the fraction of work history spent in unemployment from age 35 to 55. Since we will show that there are important connections between young and prime-age unemployment, the five-years gap is necessary in order to avoid that part of the correlations are not artificially due to the aftermath of a recession, or to long unemployment spells that connect between subsequent years.

\(^1\text{Our definitions are similar to Schmillen and Möller (2012).}\)
We use only the cross-sectional representative sample of the NLSY, and exclude every worker who has less than 100 weeks of reported employment/unemployment from age 20 to 30, and 100 weeks from age 35 to 55; this gives us a sample of 5422 workers. Further, we restrict attention to the relatively narrower sample of males who are high-school educated at age 30. This leaves us with a total of 1027 individuals followed for 30 years. However, results are robust to more inclusive definitions of the sample.

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

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

where \(1(\bar{u}_i^p < q_{90}(u^p))\) is an indicator function taking value 1 if prime-age unemployment of individual \(i\) was below the 90th quantile of the prime-age unemployment distribution, and 0 otherwise, while \(T_i^p\) is the number of weeks in which individual \(i\) was either employed or unemployed during prime-age. The 10% most unemployed individuals account for about 2/3 of prime-age unemployment observed in the data. Moreover, about half of these individuals have never been unemployed in the reference period. Notice that the fact that prime-age unemployment is concentrated in relatively few workers is a very different point from the one raised for instance by Clark et al. (1979), who show that most the pool of the unemployed is accounted for by workers staying unemployed, rather than workers going in and out of unemployment.

We then proceed to compute monthly average job finding/job separation rates

---

2This is to have as homogeneous a sample as possible. High-school males are the biggest sex-education subgroup in the NLSY/79. Moreover, Menzio, Telyukova and Visschers (2012) show that, in terms of labor market outcomes, this subgroup is a good representation of the behavior of US labor market aggregates over the life cycle.

3Clearly this is not the only way of computing this average. Another possibility is to compute instead

\[
\tilde{u}^{p}_{u_p < q_{90}} = \frac{\sum_{i=1}^{N} 1(\bar{u}_i^p < q_{90}(u^p)) u_i^p}{N}
\]  

that is, the average of each individual’s prime-age unemployment. The two averages are different since \(T_i^p\) differs across individuals, because some are observed for more weeks than others; in particular, there can be a significant difference if \(\text{COV}(T_i^p, u_i^p) \neq 0\), for instance if those often unemployed tend to be more often out of the labor force. We find that there is little difference between the two ways of computing the average, and that this does not matter for results on the concentration of unemployment, which is even larger (about 70% accounted for by top 30%) if using this second methodology.
Table 1: Left column: averages 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. Right column: averages computed by simulating sequences of job finding - job separation events using flow equations of Mortensen-Pissarides model, calibrated to average job finding and job separation probabilities in NLSY/79 sample.

<table>
<thead>
<tr>
<th></th>
<th>NLSY/79</th>
<th>Unif. Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. % time in unemployment</td>
<td>3.6</td>
<td>(target) 3.6</td>
</tr>
<tr>
<td>Avg. % time in U, excluding top 10 %</td>
<td>1.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Avg. % time in U, excluding top 20 %</td>
<td>0.6</td>
<td>1.9</td>
</tr>
<tr>
<td>% never unemployed:</td>
<td>56</td>
<td>29</td>
</tr>
</tbody>
</table>

for workers in their prime[4]. We find that the concentration of unemployment is both due to a (≃ 3 times) lower finding rate[5] and a (≃ 9 times) higher separation rate for that top 10 % (see table[4] in the appendix); this group of workers 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 at the same time means not getting the distribution of unemployment right. Interestingly, the difference in separation rates accounts for a larger fraction of the heterogeneity in unemployment outcomes than the difference in finding rates.

Similarly to what happens when discussing income inequality, measures of concentration might not be meaningful if they are not compared with what a standard framework would imply for the distribution of unemployment. Moreover, since unemployment is a random variable, lifetime statistics by parts of the unemployment distribution suffer from selection bias, because the empirical strategy is selecting individuals who experienced exceptionally high amounts of unemployment, thus the underlying job finding and job separation rates might be substantially different. In order 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.

Simulations show that the standard, calibrated to reproduce the job finding and job separation rates of the sample, has trouble replicating the observed concentration in prime-age unemployment: the standard search model features too many transitions in and out of unemployment for the majority of workers. This fact is im-

---

[4] Since we will calibrate the model to monthly probabilities, we do not adjust for short-term unemployment as in Shimer (2012).

[5] We describe how we compute job finding and separation rates in appendix A.
important, because it suggests that heterogeneity across workers is likely to be crucial to make sense of labor market outcomes, and of the ins and outs of unemployment, during prime age.

Second, we document that young and prime-age unemployment are strongly correlated. Workers who were in the top 10% of the young-age distribution are five times more likely to be in the same part of the distribution when prime-age. In short, young and prime-age unemployment are connected and, among a wide range of observables available in the NLSY/79, young unemployment is the best predictor of prime-age unemployment. Noticeably, regression analysis (see table 5 in the Appendix) confirms that young unemployment is a very strong predictor of prime-age unemployment, and that this is not due to observables such as education or IQ.

<table>
<thead>
<tr>
<th>All sample</th>
<th>Rest</th>
<th>Top 10 % when (35-55)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rest</td>
<td>92.68</td>
<td>7.32</td>
</tr>
<tr>
<td>Top 10 % when (20-30)</td>
<td>65.94</td>
<td>34.06</td>
</tr>
<tr>
<td>High-School Workers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rest</td>
<td>93.53</td>
<td>6.47</td>
</tr>
<tr>
<td>Top 10 % when (20-30)</td>
<td>58.82</td>
<td>41.18</td>
</tr>
</tbody>
</table>

Table 2: 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). Source: own calculations on NLSY/79.

Little additional information can be obtained by decomposing further the separation rate: using the matched employer-employee dataset available along the NLSY/79, we show that workers who were in the top 10% most unemployed in prime-age had about twice the likelihood of separating from their employers for any reason than the rest of the sample, without one single particular reason being more important than others (see table 6 in Appendix).

As a final piece of evidence, we compute job finding and job separation probabilities depending on age. We use 5-years long age groups because the selection of the top 10% prime-age unemployed can generate heavy nonlinearities, which
a polynomial in age might not capture\textsuperscript{6}. We can see that, at ages 20-30, the job separation rate of the top 10% of prime-age unemployed is more than double the job separation rate of the rest of the sample. Instead, between the two groups there is only a 7 percentage points (about 1/3 of the sample average) difference in job finding rates at ages 20-30, but this difference becomes much more pronounced during prime age.

Figure 1: Job finding (left panel) and Job Separation (right panel) probabilities, by group of prime-age unemployed. Sample of male, high-school educated workers. Source: own calculations on NLSY/79. Shaded areas are 95% confidence bands.

This suggests that, in the eyes of potential employers, the two groups of workers were not substantially different at the beginning of their working careers, but that such differences became more pronounced later; however, 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. That is, before an employment relationship had been established, young workers who came to experience substantially different careers looked similar; however, as they accumulated jobs and separations, workers experienced increasingly different job finding rates, suggesting that information on them had slowly become available.

The wages of the top 10% unemployed progressively fall over the life cycle, relatively to those of the rest of the population (see figure \textsuperscript{7}), confirming that differences across workers become larger over workers’ careers. This suggests that, after many separation events, such workers may sort into different jobs in order to avoid frequent future separations, or that they may 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 in relatively few workers and the

\textsuperscript{6}We use 5 years as the length of an age group due to the relatively small sample.
persistence of unemployment over the life cycle; such concentration and persistence can have important consequences for the design of labor market policy. For instance, the concentration of unemployment suggests that relatively few people are likely to obtain the bulk of unemployment insurance, and will be the most affected by it. However, we 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 workers’ careers, and that workers are being slowly sorted by employers over their careers. Our model will feature this mechanism, which has important implications for understanding the concentration of unemployment, the connection of young and prime-age unemployment, and the effects of labor market policies.

4 Model

The model is roughly based on Delacroix and Shi [2006] and Gonzalez and Shi [2010]. We add heterogeneity across workers, information frictions and a notion of ‘resumè’ of the worker in order to capture sorting in different jobs over the life cycle and the fact that a group of workers experiences higher separation rates. The quantitative version of the model will feature more ingredients in order to allow to disentangle more mechanisms, but we first present a simpler version with the key ingredients.

4.1 Environment

The economy is populated by a measure of firms $M > 1$ and a measure one of workers, who are either employed or unemployed. Every period, a fraction $\lambda$ of workers die, and are replaced by newly born, unemployed workers. Each worker is born of type $i = \{H, L\}$, High and Low respectively, unknown both to firms and workers; low types occur with probability $l$, high types with probability $1 - l$. All agents are risk neutral and discount the future at rate $1/(1 + r)$.

Let $p$ be the market prior probability over a worker being of high type. The $p$ of each worker is public in the economy, and evolves along the history of the worker, as explained below. There exists a continuum of submarkets indexed by $\{w, p\}$, the wage $w$ earned in that submarket and the prior $p$ of workers applying to that submarket. Wages are perfectly rigid in each submarket; however, firms have limited commitment and can destroy a match at the beginning of every period.
4.2 Search and Matching

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

4.3 Information and Learning

Denote by $H(x)$ and $L(x)$ the cumulative distribution functions of match quality, for high and low types respectively, with support $X \subseteq \mathbb{R}$, such that $H(x)$ first order stochastically dominates $L(x)$; that is, $H(x) \leq L(x)$ $\forall x \in X$. Once a match with a worker has been established, a match-specific quality shock is drawn from the workers’ type distribution. Match quality is constant over the whole duration of the match.

During a match with a worker, with probability $1 - \pi$, firms do not observe the worker's output. After a match with a worker with belief $p$ has been established, in each period the firm pays the wage $w$ to the worker, and gets expected payoff $\mathbb{E}(x | p)$, until a separation occurs. With probability $\pi$, the firm observes the output of the match; by assumption, the occurrence of this event is known to the market if the match continues but not if the match is destroyed. Although the occurrence of $\pi$ in case of continuation is observed by the market, the output produced by the worker is not. Denote by $d(w, x)$ the choice of a firm to destroy the match; after observing the worker’s output, the firm can decide whether to keep the match ($d = 0$) or destroy it ($d = 1$). Moreover, matches can end randomly with probability

---

[1] Menzio, Telyukova and Visschers [2012] find that the probability that a match changes quality during an employment relation is around 1%, thus making the constant match assumption a reasonable simplifying approximation.

[8] Since match quality is persistent, removing observability of $\pi$ would imply the addition of another state variable, the job’s duration, because the informational content of not experiencing a separation declines along the match's duration. To see why, consider the history of a worker who does not observe the occurrence of $\pi$: at the end of the first period, he has been observed with probability $\pi$ or not observed w.p. $1 - \pi$. In the second period however, he has already been observed with probability $\pi + (1 - \pi)\pi$ and has not been observed with probability $(1 - \pi)^2$; in the third period, he has already been observed with probability $\pi + (1 - \pi)\pi + (1 - \pi)^2\pi$ and so on. By induction, the probability that the worker has not been observed yet is $\sum_{t=0}^{D} (1 - \pi)^t$ where $D$ is the duration of the match, which then becomes necessary to compute the informational content of job continuation. We plan to extend the model to add this ingredient in the future, but preliminary simulations show that the current version is a good approximation of a model with duration as another state variable.

[9] If one postulates that there is a small cost for revealing information, the firm will never choose to reveal the information it observed because it does not profit by revealing it in any way.
The market can gain information on a worker’s type during the match’s duration. When the observation shock realizes, continuation is good news: output has to be higher than the wage in order to ensure continuation, which is more likely for high-type workers. Instead, a separation is bad news: either a random separation event occurred, or output was lower than the wage. Thus, \( \pi \) and the properties of the match quality distribution \( H(x) \) and \( L(x) \) measure the informational content of job duration. If \( \pi = 0 \), firms never observe the type of the worker and job duration is not informative of the worker’s type.

To see how the properties of the distributions convey information on a worker’s type, consider the simple case in which

\[
H(x) = \begin{cases} 
1 & \text{if } x \geq y_H \\
0 & \text{otherwise} 
\end{cases} \quad (4)
\]

\[
L(x) = \begin{cases} 
1 & \text{if } x \geq y_L, \quad y_L < y_H \\
0 & \text{otherwise} 
\end{cases} \quad (5)
\]

that is, the distribution of match quality is degenerate and output of high types is always higher than the output of low types. If the worker had applied to a wage \( y_H \geq w > y_L \) and the random separation rate \( \delta \) was zero, a separation would immediately signal that output was lower than the wage, thus revealing with certainty that the worker is of low type. Conversely, if \( w < y_L \), the market would not learn anything from a separation because such event will occur only for random reasons. Similarly, consider the case in which \( H(x) = L(x) \). In this case, neither continuation nor separations give any information on the worker’s type, because both events will be triggered with the same probability for both high and low types; in fact, there is only one type of worker.

In fact, \( p \) is a sufficient statistic for the number of times a worker was observed by a firm and not separated, and the number of separations he experienced, and can be considered the worker’s ‘resumé’.

Timing is as follows:

1. Workers die w.p. \( \lambda \), replaced by unemployed workers with belief \( 1 - l \).

\footnote{While it is reasonable to assume that the number of past jobs and their duration is observable, the fact that wages are is somewhat more controversial. Notice, however, that while the model features heterogeneous rates of learning for different wages, it is not necessary for the market to know anything else than a worker’s resumé \( p \) and her employment status to figure out the evolution of \( p' \), because unemployed workers with resumé \( p \) will apply to a wage \( w(p) \) and this can be rationally anticipated. To the best of my knowledge, there are few studies that formalize the notion of resumé, and those who do make the somewhat extreme assumption that match quality is being observed too (Doppelt 2014).}
2. Exogenous separations occur; w.p. \( \pi \), firms observe workers’ output \( \Rightarrow \) Separations occur.

3. Workers revise beliefs: \( p' = p \) if worker is still unobserved and no shocks occur, \( p' = C(w,p) \) if worker is observed and match continues, \( p' = F(w,p) \) if worker is observed and worker is separated, \( p' = p \) if worker has been observed in the past and is separated.

4. Production occurs, wages are paid.

5. Unemployed workers search for a job. They can choose to search in submarket \( \{w', p'\} \).

6. Workers match w.p. \( f(\theta(w',p)) \) and extract match quality from \( H(x) \) or \( L(x) \) depending on their type.

Bayes’ rule implies that beliefs of employed workers, who are observed and whose match continues, evolve according to

\[
p' = C(w,p) = \frac{p \left[ (1 - H(w)) \right]}{p \left[ (1 - H(w)) \right] + (1 - p) \left[ (1 - L(w)) \right]} \tag{6}
\]

while beliefs of employed workers, who had not been observed yet and separated, evolve according to

\[
p' = F(w,p) = \frac{p \left[ \delta + \pi H(w) \right]}{p \left[ \delta + \pi H(w) \right] + (1 - p) \left[ \delta + \pi L(w) \right]} \tag{7}
\]

The intuition is that, when a worker is observed and the match continues, it must mean that match quality was high enough to support the current wage, an event that is more likely for high types. Viceversa, when a worker is observed and separates, it must mean that match quality was not high enough, an event that is more likely for low types.

4.4 Bellman Equations

The value function of an unemployed worker with prior \( p \) can be written as

\[
U(p) = b + \beta \left[ \max_{w'} \left[ f(w',p)(W(w',p) - U(p)) \right] + U(p) \right] \tag{8}
\]

where \( \beta = \frac{1-\lambda}{1+\gamma} \).

It is useful to define the continuation probability of a worker who has been observed, while working at wage \( w \) and prior \( p \), as
\( \chi(w, p) = (1 - \delta) \left[ \left(p \int (1 - d(w, x)) dH(x) + (1 - p) \int (1 - d(w, x)) dL(x) \right) \right] \) \( (9) \)

Thus, the value function of an employed worker, who has still not been observed (match quality unknown by the firm, at wage \( w \) and with prior \( p \) can be written as

\[
W_u(w, p) = w + \beta \left[ (1 - \pi) (1 - \delta) W_u(w, p) + \delta U(p') \right] + \pi \left[ \chi(w, p) W_k(w, p') + (1 - \chi(w, p)) U(p') \right] \]

\( (10) \)

where we dDnote by \( p'_C = C(w, p) \) the next period’s prior in case of continuation, and by \( p'_F = F(w, p) \) the next period’s prior in case of firing, suppressing the belief’s dependence on \( w \) and \( p \) in the notation for convenience.

while the value of an employed worker who has already been observed and kept her job (match quality known) can be written as

\[
W_k(w, p) = w + \beta \left[ (1 - \delta) W_k(w, p) + \delta U(p) \right] \]

\( (11) \)

The value of a firm for which output is known can be written as

\[
J_k(w, x) = \max_{d \in \{0, 1\}} \left[ (1 - d)(x - w + \beta(1 - \delta) J_k(w, x)) \right] \]

\( (12) \)

while the value of a firm matched with a worker at prior \( p \) and wage \( w \), for which output is unknown, can be written as

\[
J_u(w, p) = \mathbb{E}(x | p) - w + \beta(1 - \delta) \left[ \pi(p \int J_k(w, x) dH(x) + (1 - \pi) J_u(w, p) \right] \]

\( (13) \)

The value of posting a vacancy in submarket \((w, p)\) is

\[
V(w, p) = -\kappa + q(\theta(w, p))\beta J_u(w, p) \]

\( (14) \)

and the tightness function must satisfy

\[
\kappa \geq q(\theta(w, p))\beta J_u(w, p) \quad \forall w, p \]

\( (15) \)

which makes \( \theta \) consistent with the firm’s optimal vacancy creation; \( (15) \) holds with
equality if $\theta > 0$. Basically, condition $[15]$ 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.

### 4.5 Equilibrium

**Definition 1.** A RBE (Recursive Block Equilibrium) for this economy consists of:

- A value function for the unemployed worker $U(p)$,
- A policy function for the unemployed worker $w'(p)$,
- A value function for the employed worker $W(w,p)$,
- A value function for the informed firm $J_k(w,x)$,
- A separation policy for the informed firm $d(w,x)$,
- A value function for the uninformed firm $J_u(w,p)$,
- A tightness function $\theta(w,p)$ and laws of motion for beliefs $C(w,p)$ and $F(w,p)$ such that

1. $U(p), w'(p), W(w,p), J_k(w,x), d(w,x), J_u(w,p), \theta(w,p)$ are independent of the aggregate state $\psi$
2. $\theta(w,p)$ satisfies $[15] \forall w, p$
3. $U(p)$ and $w'(p)$ satisfy $[8] \forall p$
4. $J_k(w,x)$ and $d(w,x)$ satisfy $[12] \forall w, x$
5. $J_u(w,p)$ satisfies $[13] \forall w, x$
7. $C(w,p)$ satisfies $[6]$
8. $F(w,p)$ satisfies $[7]$

**Proposition 1.** The unique recursive equilibrium for this economy is a RBE.

**Proof:** work in progress. The argument is the same as in Menzio and Shi (2009) and Menzio, Telyukova and Visschers (2012). The idea is that all value functions and policy functions are independent of the aggregate state, and so is the market tightness function.

The RBE is much easier to solve than a Recursive Equilibrium, because value functions and policy functions of agents depend only on the states $w, p, x$ and not on aggregate states. Aggregate statistics can be computed, after solving the RBE, from the aggregation of individual choices. Moreover, computing transitions out of the steady state is easy because all policy functions and laws of motion are independent of the aggregate state.
4.6 Characterization of Equilibrium

It is easy to see that \( d(w, x) \) is a step function that takes value 1 when \( w > x \) and 0 otherwise, that is, matches that produce more than what they cost to be maintained are not destroyed.

**Lemma 1.** Given \( p \in [0, 1] \), \( J_u(w, p) \) admits a unique fixed point that is continuous in \( w \), \( J_u \in [0, J_u] \), and for \( J_u \in (0, J_u) \), \( \partial J_u / \partial w < 0 \), \( \partial J_u / \partial p > 0 \).

**Proof.** Given a match quality value \( x \) and a wage \( w \in [0, x] \), \( J_k(w, x) \) can be rewritten as

\[
J_k(w, x) = \begin{cases} 
\frac{x-w}{1-\beta(1-\delta)} & \text{if } w \leq x \\
0 & \text{otherwise}
\end{cases}
\]  

(16)

Substituting \( J_k(w, x) \) into \( J_u(w, p) \) yields

\[
J_u(w, x) = \frac{\mathbb{E}(x | p) - w + \beta(1-\delta) \pi \left[ p \int_{w}^{\infty} \frac{x-w}{1-\beta(1-\delta)} dH(x) + (1-p) \int_{w}^{\infty} \frac{x-w}{1-\beta(1-\delta)} dL(x) \right]}{1 - \beta(1-\delta)(1-\pi)}
\]  

(17)

The right-hand side of expression (17) is decreasing in \( w \) because current flow profits are decreasing in \( x \), the future expectation of flow profits is computed on fewer match qualities (there are fewer match qualities that support wage \( w \)) and, for every match quality \( x \), profits are lower.

Moreover, \( J_u(w, x) \) is increasing in \( p \). To see this, first remember that

\[
\mathbb{E}(x | p) = p \int_{0}^{\infty} x dH(x) + (1-p) \int_{0}^{\infty} x dL(x)
\]  

(18)

Since \( H(x) \) first-order stochastically dominates \( L(x) \), \( \int_{0}^{\infty} x dH(x) > \int_{0}^{\infty} x dL(x) \), thus current flow profits are increasing in \( p \). As for the future value of the firm, first-order stochastic dominance implies that \( \forall w > 0, \int_{w}^{\infty} x - w dH(x) > \int_{w}^{\infty} x - w dL(x) \), thus showing that \( \partial J_u / \partial p > 0 \).

To see that \( J_u(w, p) \) has an upper bound, consider the case in which the wage is equal to zero. In this case, \( J_u(w, p) = \frac{\mathbb{E}(x)p}{1-\beta(1-\delta)} \), completing the proof. □

**Lemma 2.** In the RBE of the economy, the unique solution to equilibrium condition (15) is
\[ \theta(w, p) = \begin{cases} 
q^{-1}(k/\beta J_u(w, p)) & \text{if } \beta J_u(w, p) \geq k \\
0 & \text{otherwise} 
\end{cases} \] (19)

Since the function \( J_u(w, p) \) is continuous in \( w \) for \( w \in [0, \mathbb{E}(x | p)] \), the market tightness function \( \theta \) is continuous in \( w \). Furthermore, since \( J_u \) is a decreasing function of \( w \), \( \theta(w, p) \) is a decreasing function of \( w \). The intuition is that, as firms have to pay higher wages, their expected profits are lower, so that a higher job filling probability is required to pay for the cost of creating a vacancy, thus implying a lower tightness in that submarket. Finally, as \( J_u \) is an increasing function of \( p \), \( \theta(w, p) \) is an increasing function of \( p \). When a worker has a higher probability of being a high type, her expected productivity is lower, thus for any given wage, the equilibrium tightness function will be higher because more firms will post vacancies in that submarket.

**Corollary 1.** The matching probability \( f(\theta(w, p)) \) is continuous in \( \{w, p\} \), decreasing in \( w \) and increasing in \( p \).

The corollary follows trivially from the fact that \( \theta(w, p) \) is a continuous function of \( J_u(w, p) \), which is a continuous function of \( w, p \), and that \( \theta(w, p) \) is decreasing in \( w \) and increasing in \( p \), and that \( f(\theta) \) is continuous and increasing in \( \theta \). \( \square \)

In the future I will proceed to establish results on desired wages of workers as a function of \( p \), and on the associated matching probability in submarkets \( \{w, p\} \). The intuition is that, as workers with higher \( p \) have a higher expected productivity and a lower separation probability for every given wage \( w \), they will demand higher wages and possibly face higher job finding rates. The last point will depend on the trade-off, at given \( p \), between higher wages and lower job finding probability.

### 5 Quantitative Model and Identification

In the quantitative version of the model, we add another dimension of heterogeneity: some workers have higher skills and thus produce more than one unit per unit of match productivity when employed. Such skills are observable by firms, contrary to the type of the worker. The probability of being skilled is a function of the type of the worker: high and low type workers are born skilled with probability \( \alpha_h \) and \( \alpha_l \), respectively. While the productivity of unskilled workers is equal to the productivity of the match \( x \), the productivity of a skilled worker is equal to \( x s \), where \( s \geq 1 \) is the skill multiplier. Since high and low types are born skilled with
different probabilities, the fact that they are skilled or unskilled conveys information on their type at the start of their careers.

The model is identified by calibrating parameters in order to replicate features of job finding rates, job separation rates and wage patterns observed in the NLSY79. The idea behind the calibration is that the concentration and persistence of unemployment, and the differences in job separation and job finding rates by parts of the prime-age unemployment distribution, are informative of the amount of low-type workers present in the economy and of the differences between the match quality distributions of types, while the life-cycle profile of wages are informative of differences both in the match quality distributions and in skill multipliers. This strategy is partly inspired by Menzio and Shi (2011) and Menzio, Telyukova and Visschers (2012), who use the life-cycle patterns of job finding rates, job separation rates and employment-to-employment transitions in order to identify the parameters of the match quality distribution and the probability of observing productivity during a match. Our model works similarly during a match’s duration, so we apply the same strategy but we distinguish between the job finding/separation rates experienced by the top 10% prime-age unemployed and the rest of the population.

To see why the match quality distribution affects separation rates, consider the separation policy of the firm $d(w,x)$. Given a wage $w$, the probability that a firm destroys a match upon discovering match quality is $1 - H(w)$ for high types and $1 - L(w)$ for low types, which means that the way in which the probability mass is distributed over match qualities determines separation rates at each wage for different types.

To see why the match quality distribution affects job finding rates, consider equation 15 which states that in equilibrium the tightness of submarket \{w,p\} depends on the expected profits of the firm for a worker with prior $p$. In turn, expected profits depend on $\mathbb{E}(x \mid p)$ and on $\mathbb{E}(J_k(w) \mid p)$, that is on current expected productivity and on future productivity if the match will not be destroyed. Basically, job finding rates depend on expected match quality given the prior, and on the expected match quality for the part of the distribution above the separation cutoff.

Summing everything up, a distribution featuring high mass on low values of match quality, but a long right tail, will deliver high separation rates and high job finding rates. On the other hand, a distribution featuring high mass on low values of match quality and a short right tail will deliver high separation rates and low job finding rates. Finally, a concentrated distribution, such that uncertainty about
match quality is low, will deliver low separation rates.

To see why the concentration and persistence of unemployment is informative on the amount 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 of low types has more probability mass on low realizations than the match quality distribution of high types. This means that young, low type workers who are starting their careers will typically experience a larger-than-average amount of separations during their youth. As information on their type accumulates, these workers will slowly sort into lower-wage jobs, but will still experience higher separation rates because of the worse match quality distribution, and will experience lower job-finding rates because their expected productivity will be lower. The mechanism does not necessarily apply only to low types: high types who have been unlucky, and extracted many low-quality matches, will experience frequent separations and will be considered "low-types" with a high probability, thus experiencing lower job-finding rates when older.

Young-age separation rates depend on how fast output is observed (parameter \( \pi \)), while the speed of learning depends on how far apart the two distributions of match quality are: if types extract match qualities from very similar distributions, learning will be slow, desired wages will be similar and the concentration of unemployment will be low too. If the two types extract from very different distributions, learning will be fast and unemployment will be concentrated in few workers. The scale and shape of the match distributions will thus influence the life-cycle profile of job finding rates, job separation rates and wages. Notice that it is possible to obtain concentration of unemployment even with only one type of workers, just by changing features of the match quality distribution. However, this would be inconsistent with the fact that unemployment is persistent along the life cycle, and with the patterns of job finding rates and job separation rates by unemployment groups.

Persistence of unemployment relies on a form of incomplete learning and on the variance of the match quality distributions: if match quality was not uncertain, after high young unemployment low-type workers would be sorted into different wages, and would never experience unemployment if not for random reasons (provided that their match quality distribution supports wages higher than the flow value of unemployment). In this extreme case, persistence of unemployment would be caused only by not-yet-sorted low-type workers, who would be still working at a higher wage than that which can be sustained by their match quality, so that
persistence would be most affected by the measure of low-type workers present in the economy, because a certain level of persistence could be observed only in the presence of a large enough measure of low-type workers. Intuitively, if low-type workers are less than 10% of the population for instance, and they all sort into low-wage jobs before age 35, young unemployment will not be informative of prime-age unemployment. At the other extreme, if match quality was very uncertain for both types, learning over the life cycle would be very slow, and separations would give little information on a worker’s type. Thus, persistence in unemployment would be mainly a consequence of the speed of learning, and have a lower connection with the measure of low types.

5.1 Calibration

We then proceed to simulate the lives of a large sample of workers in order to compute lifetime statistics, and calibrate the model to replicate as closely as possible the observed patterns of wages and transition rates in the NLSY79. Calibration is performed by applying the Simulated Method of Moments: we minimize the loss function

\[ L(\theta) = m(\theta)' W m(\theta) \]  

(20)

where \( \theta \) is the vector of parameters of the model, \( m(\theta) \) is a column vector of the differences between the model-generated moments and the data moments, and \( W \) is a weighting matrix\(^{11}\).

We calibrate the model at the monthly frequency. We assume workers are born at age 20, the starting age of our data, and choose the death probability \( \lambda \) in order to match an average working life of 40 years. We choose the interest rate \( r \) as to give a compounded annual interest rate of 4%.

In line with many other models of directed search (Shimer (2005); Mortensen and Nagypal (2007); Menzio and Shi (2011); Menzio, Telyukova and Visschers (2012)), we restrict the matching probability to be of the form \( f(\theta) = \min\{\theta^{0.5}, 1\} \).

The flow value of unemployment \( b \) is chosen as to match a ratio between \( b \) and average wages of 0.71, in line with the estimates of Hall and Milgrom (2007).

The two match quality distributions \( H \) and \( L \) are assumed to be Weibull distributions\(^{12}\) with scale parameters \( \sigma_H, \sigma_L \) and shape parameters \( \phi_H, \phi_L \). Shape and scale of match quality distributions, the probability \( \pi \) of observing a worker’s out-

---

\(^{11}\)Computation of variance-covariance matrix of moments and of standard errors is work in progress.

\(^{12}\)The Weibull distribution is a common choice in this regard. See for instance Menzio, Telyukova and Visschers (2012).
put, the random separation probability $\delta$ and the measure of low-type workers $l$, are calibrated as to match the observed patterns of job finding rates, job separation rates over the life cycle by rest of population and top 10% unemployed, and the observed concentration and persistence between young and prime-age unemployment of top 10% and of top 20%, as in the tables presented in section 3.

The skill multiplier $s$, the probabilities $\alpha_h$ and $\alpha_l$ of being skilled are calibrated as to match the initial observed difference in job finding rates and the observed difference between the wages of the top 10% unemployed and the rest of the population over the life cycle. The intuition is that the multiplier $s$ matters for wages, as skilled workers will demand higher wages. At the same time, the probabilities of being skilled will influence the strength of the initial signal given by the presence or absence of skills, and the gap between the wages of high type workers and low type workers: as a consequence, such gap translates in a wage differential between the top 10% unemployed and the rest.

The vacancy creation cost $\kappa$ is calibrated as to match the job finding rate of bottom 90% of the prime age unemployment distribution at ages 20-25.

<p>| Table 3: Baseline calibration results. Targets calculated on NLSY/79. |</p>
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th>Target</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest rate</td>
<td>$r$</td>
<td>0.0034</td>
<td>Annual interest rate 4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Death Probability</td>
<td>$\lambda$</td>
<td>0.0021</td>
<td>Avg. work life 40 years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scale parameter, high type</td>
<td>$\sigma_H$</td>
<td>1.7499</td>
<td>Wage differential top 10% - rest,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scale parameter, low type</td>
<td>$\sigma_L$</td>
<td>17.5938</td>
<td>JS rate profiles,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shape parameter, high type</td>
<td>$\phi_H$</td>
<td>0.5338</td>
<td>JF rate profiles,</td>
<td>0.60</td>
<td>0.60</td>
</tr>
<tr>
<td>Shape parameter, low type</td>
<td>$\phi_L$</td>
<td>0.7635</td>
<td>% prime-age U acc. by top 10</td>
<td>0.83</td>
<td>0.82</td>
</tr>
<tr>
<td>Prob. of observing output</td>
<td>$\pi$</td>
<td>0.0097</td>
<td>% prime-age U acc. by top 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prob. of low skilled</td>
<td>$\alpha_l$</td>
<td>0.8212</td>
<td>Wage differential top 10% - rest,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prob. of high skilled</td>
<td>$\alpha_h$</td>
<td>0.9662</td>
<td>JF rate profiles,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skill Multiplier</td>
<td>$s$</td>
<td>1.1422</td>
<td>Initial W diff, top 10% - rest,</td>
<td>-0.21</td>
<td>-0.19</td>
</tr>
<tr>
<td>Flow value of Unemployment</td>
<td>$b$</td>
<td>0.8101</td>
<td>Ratio $b/avg.$ wage</td>
<td>0.70</td>
<td>0.64</td>
</tr>
<tr>
<td>Vacancy cost</td>
<td>$\kappa$</td>
<td>3.2368</td>
<td>Avg. Job finding rate, bottom 90, age 20-25</td>
<td>0.29</td>
<td>0.30</td>
</tr>
<tr>
<td>Random separation rate</td>
<td>$\delta$</td>
<td>0.0038</td>
<td>JS rate, bottom 90, age 40-50</td>
<td>0.007</td>
<td>0.008</td>
</tr>
<tr>
<td>Measure of low types</td>
<td>$l$</td>
<td>0.4224</td>
<td>Persistence of U, top 10</td>
<td>0.41</td>
<td>0.30</td>
</tr>
<tr>
<td>Persistence of U, top 20</td>
<td></td>
<td></td>
<td></td>
<td>0.45</td>
<td>0.48</td>
</tr>
</tbody>
</table>

The calibration table reports only singleton targets: patterns of job finding/separation rates and wages are vectors and are shown later in graphs for readability. Overall, the estimation algorithm fits 12 parameters with 36 restrictions.
6 Results

6.1 Calibration results

Despite being calibrated with over-identifying restrictions, the model does quite a good job in replicating the main features of the data. As can be seen in table 3, the model is quite capable of delivering realistic amounts of concentration and persistence of unemployment; the model delivers three quarters of the persistence as measured by the Markov transition matrix between being unemployed when young and when prime-age, and matches almost perfectly the observed concentration both at the top 10% and 20% of the distribution of prime-age unemployment, where the standard Mortensen-Pissarides model only predicts one-third.

The match quality distributions of high and low type are substantially different: at the calibrated values, the match quality distribution of low types has a large mass close to zero, and a long right tail, while the match quality distribution of high types is narrower and more concentrated on higher match qualities (figure 2).

![Figure 2: Match quality distribution of high types (red) and low types (blue), under baseline calibration](image)

The calibrated value of the probability of a firm observing the worker’s output \( \pi = 0.07 \) implies that the average duration of a “bad match” is about 14 months.

The calibrated measure of low-type workers in the economy is around 42%, a relatively large number. As I will show in the discussion section, this number has important implications for the composition of the unemployment pool and for the
concentration and persistence of unemployment over the life cycle.

Skilled workers benefit from a 14% higher productivity; the probability that a high and low type workers are skilled are, respectively, 97% and 82%, making unskilled workers a minority among both low and high type workers. These probabilities imply that the signal of being skilled is substantially uninformative of a worker’s type at the beginning of their career, but being unskilled is a strong signal that the worker might be of low type\textsuperscript{13}.

Figure 3: Model (continuous line) versus data (dashed); Job finding rates (left) and job separation rates (right) of top 10% prime-age unemployed (red) vs rest (blue). Results under baseline calibration.

Figure 3 shows that the job finding rate of the top 10% of prime-age unemployed declines over the life cycle as in the data, while the job finding rate of the rest of workers rises during prime-age. The model does also a very good job in explaining wage differentials between the top 10% and the rest until age 40, after which it explains only two-thirds (figure 8 in Appendix).

Figure 4 plots the probability that a worker is of high type depending on her age, by low and high types and by part of the prime-age unemployment distribution. The figure depicts what we term “learning over the life cycle”: as separations and continuations occur, the market slowly sorts out who are high-type and who are low-type workers. The patterns of job finding and job separation rates are a consequence of this mechanism.

Let us look first at the job finding rate: as the market learns who are low and

\[ P_{\text{skilled}} = \frac{\alpha_h (1-l) + \alpha_l (1-h)}{\alpha_h (1-l) + \alpha_l (1-h) + (1-h)}, \]

\[ P_{\text{unskilled}} = \frac{(1-\alpha_h) (1-l)}{(1-h) + (1-h) + (1-\alpha_l)} \]

Thus a skilled worker starts with prior 0.62 and an unskilled worker with prior 0.21.

\textsuperscript{13} By Bayes’ rule
who are high type workers, the gap in job finding rates between workers widens. This can be seen by comparing job finding rates and job separation rates of high and low type workers in the model (figure 5). The result follows from this mechanism, and from the fact that more than 85% of the unemployment pool is made of low type workers (figure 9 in Appendix). Thus, the job finding rate of the top 10% unemployed is essentially the job finding rate of the most unlucky of low types: the model predicts that 99% of the top 10% unemployed in prime age are low type workers.

Job separation rates are substantially higher for the top 10% unemployed, both when young and when prime-age; if anything, the model undershoots the job separation rate of the top 10% when young, and overshoots the job finding rate of the rest of workers after age 40. One reason for the failure of the model in correctly predicting the descent in the job finding rate could be that workers accumulate assets
over the life cycle and this increases their outside option value, making them demand higher wages and lowering their job finding rate as they age (see for instance Michelacci and Ruffo (2013)), while this model does not feature assets accumulation. Similarly to 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 declines because these workers initially apply to too high wages, extract low values of match quality and face frequent separations. However, both workers and the market learn from these separations, so that workers apply to progressively lower wages, thus facing lower separation rates.

With the calibrated version of the model, we now proceed to simulate a number of alternative scenarios in order to quantify the relative importance of each mechanism.

6.2 The importance of Learning over the Life Cycle

I now simulate what would happen if firms were to know with certainty the type of a worker from the beginning. Results are summarized in the second column of table 6.4.

We estimate that incomplete information about workers’ types is responsible of 22% of separations during the first year of career, of 13% of separations before age 25, and of 10% of job finding before age 25, but that it has almost no influence on later job finding and separation rates. Moreover, a model without learning from workers’ careers would imply a wage differential of -41% between the top 10% of prime age unemployment and the rest at age 20, as opposed to the true differential of -21%. This is because, as types are already known by firms, workers start applying to different wages from the very beginning of their careers, so that wage differentials are constant over the life cycle while in the data such gap widens as workers age. That information frictions explain 20% higher wages at the start of the career, for workers who will have exceptionally high unemployment and low wages in their lives, is perhaps one of the most important findings of the paper.

Concentration of unemployment during prime age is substantially unaffected by the absence of learning: this is because, by age 35, workers have accumulated enough labor market history and their types are already revealed, so the fact that firms already know the types of workers has no influence. However, persistence of unemployment is positively affected by information on types: this is because in the baseline model low type workers tend to find jobs faster at the beginning of their careers, at a speed more similar to that of high type workers, thus they initially
experience lower unemployment than what they do in the model in which types are known.

6.3 The role of heterogeneity and skills

Since there are two types of workers, heterogeneity can be removed either by setting the probability that a worker is a low type $l$ equal to 1 or to 0; we perform simulations under the two scenarios.

Results are summarized in columns 4 and 5 of table 6.4; the model without heterogeneity fails in delivering persistence of unemployment over the life cycle. The reason is that young unemployment is not carrying any information in a model without heterogeneity, because all workers are equally likely to face separations and to find jobs; there are still small differences between skilled and unskilled workers, but since this characteristic is observable, this mechanism passes only by the job finding rate and is not sufficient to deliver a plausible persistence of unemployment.

A level of concentration similar to the one in the data is achieved by the model with $l = 0$, but this is also because such model delivers an unemployment rate substantially lower than the one in the data, for which obtaining high concentration is relatively easy even for the Mortensen-Pissarides model. Quite obviously, the models without heterogeneity have nothing to say on wage differentials.

6.4 Observable skills vs unobservable skills

In the last counterfactual simulation, we remove the possibility that some workers have higher observable skills; we find that such skills have an important impact on the job finding rate, which are lower by 4-5 points in the model without observable skills, but only a relatively small one on job separation rates, and that their absence does not affect neither concentration nor persistence of unemployment. Skills account for 2 percentage points in wage differentials between the top 10% prime-age unemployed and the rest.

6.5 Duration dependence

The model is also capable of reproducing a duration dependence relation in job finding rates (figure 6), similar to the one documented by Hornstein (2012) and Wiczer (2014). The relation arises because of a composition mechanism similar to Gonzalez and Shi (2010): workers with higher market prior find jobs first, followed...
Separation rates

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>No Learning</th>
<th>No Learning</th>
<th>No Learning</th>
<th>All types</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>No heterogeneity</td>
<td>All high types</td>
<td></td>
</tr>
<tr>
<td>age 20</td>
<td>2.56</td>
<td>2.00</td>
<td>0.56</td>
<td>5.27</td>
<td>2.75</td>
</tr>
<tr>
<td>age 20-25</td>
<td>1.92</td>
<td>1.69</td>
<td>0.46</td>
<td>3.90</td>
<td>2.06</td>
</tr>
<tr>
<td>age 25-30</td>
<td>1.11</td>
<td>1.08</td>
<td>0.40</td>
<td>2.12</td>
<td>1.29</td>
</tr>
<tr>
<td>age 40-50</td>
<td>0.78</td>
<td>0.77</td>
<td>0.39</td>
<td>1.31</td>
<td>0.89</td>
</tr>
</tbody>
</table>

Finding rates

<table>
<thead>
<tr>
<th></th>
<th>age 20</th>
<th>age 20-25</th>
<th>age 25-30</th>
<th>age 40-50</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>32.2</td>
<td>23.7</td>
<td>22.0</td>
<td>23.7</td>
</tr>
<tr>
<td></td>
<td>25.0</td>
<td>21.5</td>
<td>21.6</td>
<td>23.4</td>
</tr>
<tr>
<td></td>
<td>59.4</td>
<td>58.3</td>
<td>58.9</td>
<td>59.2</td>
</tr>
<tr>
<td></td>
<td>18.8</td>
<td>18.6</td>
<td>18.5</td>
<td>18.4</td>
</tr>
<tr>
<td></td>
<td>29.0</td>
<td>18.7</td>
<td>15.7</td>
<td>17.1</td>
</tr>
</tbody>
</table>

Concentration

Prime-age U accounted by top 10:
60
Prime-age U accounted by top 20:
82

Persistence

Persistence top 10 young - top 10 primes:
30
Persistence top 20 young - top 10 primes:
48

Wage diff top 10 - rest, age 20-25:
-19

7 Conclusions

Using NLSY/79 data, we show that unemployment during prime age is concentrated in relatively few workers, who experience both long spells of unemployment and frequent separations from their jobs. Moreover, unemployment is persistent in the sense that those who were often unemployed when young tend to be often unemployed during their primes. The standard search model is at odds with these facts. I build a model that delivers both high concentration of unemployment during prime age and persistence of unemployment over the life cycle; the model delivers such result by a combination of incomplete information and heterogeneity across workers. We 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 delivers a too high wage gap between different workers at the start of their work life, and a lower separation rate than the one observed in
the data for young workers. Finally, we find that differences in unobserved skills, rather than in observed skills, are responsible for the bulk of our results. In the future, we will simulate the effect of increases in unemployment subsidies and other labor market policies in our environment.

References


A Appendix

A.1 Construction of job finding and job separation probabilities

Following Clark and Summers (1979) and Wiczer (2014), we consider workers who exit the labor force as if they were not in the at-risk population; for each group of workers $N_j$, which can be the whole sample ($N_j = N$), or the top 10% of the unemployment distribution and its complement, we use the formulas

$$F_j = \frac{\sum_{i \in N_j} \sum_{t=1}^{U^p_i} f_{i,t}}{|N_j| \sum_{t=1}^{U^p_i}}$$

$$S_j = \frac{\sum_{i \in N_j} \sum_{t=1}^{E^p_i} s_{i,t}}{|N_j| \sum_{t=1}^{E^p_i}}$$

where $|N_j|$ stands for the number of elements (individuals) in group $N_j$; $f_{i,t}$ is a variable defined only in weeks spent in unemployment, which were followed by weeks spent in either unemployment or employment, and takes value 1 if the following week the worker was employed, and 0 otherwise; $s_{i,t}$ is defined only in weeks spent in employment, followed by weeks spent in either employment or unemployment, and takes value 1 if the following week the worker was unemployed, and 0 otherwise; $U^p_i$ is the number of weeks worker $i$ was employed during prime age; and $E^p_i$ is the number of weeks worker $i$ was employed during prime age.
Table 4: Summary statistics by parts of the prime-age unemployment distribution. Source: own 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>
<thead>
<tr>
<th></th>
<th>Top 10 %</th>
<th>Rest of Sample</th>
<th>Ratio Top 10 / Rest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. % time in unemployement</td>
<td>29</td>
<td>1.5</td>
<td>19.1</td>
</tr>
<tr>
<td>$\delta$: Prob. of U $\rightarrow$ E (monthly %)</td>
<td>8</td>
<td>26</td>
<td>0.3</td>
</tr>
<tr>
<td>$f$: Prob. of E $\rightarrow$ U (monthly %)</td>
<td>3.5</td>
<td>0.4</td>
<td>8.75</td>
</tr>
<tr>
<td>Predicted % time in U of top 10 %, $\delta$ alone:</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted % time in U of top 10 %, $f$ alone:</td>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Avg. Log Wage (2000) $\simeq$ -40 %

Table 5: Source: own calculations on NLSY/79. Regression of % of prime-age unemployment on % of young unemployment only for high-school educated males, for all workers + controls (2), for all workers with controls only and no young unemployment (3). Controls include sex, education, ethnic group, age in 2010, marital status, AFQT test score quartile.

<table>
<thead>
<tr>
<th></th>
<th>(1) Only HS, Males</th>
<th>(2) All, with controls</th>
<th>(3) All, No Young U</th>
</tr>
</thead>
<tbody>
<tr>
<td>% U when young (20-30)</td>
<td>0.3635***</td>
<td>0.239***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>1029</td>
<td>3127</td>
<td>3127</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.186</td>
<td>0.218</td>
<td>0.127</td>
</tr>
</tbody>
</table>

*p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Weekly probability of job termination, by reason and group of prime-age unemployment. Third column gives ratio of probability between top 10 and rest. Source: own calculations on matched employer-employee data of NLSY/79. ‘Involuntary’ category merges layoffs, establishment closures and temporary jobs ended.

<table>
<thead>
<tr>
<th></th>
<th>Rest</th>
<th>Top 10 % (35-55)</th>
<th>Ratio Top 10 / Rest</th>
</tr>
</thead>
<tbody>
<tr>
<td>When young (20-30):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fired</td>
<td>0.06</td>
<td>0.14</td>
<td>2.33</td>
</tr>
<tr>
<td>Involuntary</td>
<td>0.43</td>
<td>0.78</td>
<td>1.8</td>
</tr>
<tr>
<td>Quit to Look</td>
<td>0.4</td>
<td>0.6</td>
<td>1.5</td>
</tr>
<tr>
<td>In Primes (35-55):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fired</td>
<td>0.02</td>
<td>0.1</td>
<td>5</td>
</tr>
<tr>
<td>Involuntary</td>
<td>0.09</td>
<td>0.4</td>
<td>4.44</td>
</tr>
<tr>
<td>Quit to Look</td>
<td>0.02</td>
<td>0.09</td>
<td>4.5</td>
</tr>
</tbody>
</table>
Figure 7: Log difference in Hourly Wage, between Top 10% Prime-Age Unemployment Group and Rest. Sample of male, high-school educated workers. Source: own calculations on NLSY/79. Shaded area is 95% confidence bands.

Figure 8: Difference in wages between top 10% prime-age unemployed and rest; data (dashed) versus model (continuous) under baseline calibration.
Figure 9: Share of workers who are low types, by age; under baseline calibration.