Information Frictions, Match Quality and Lifetime Unemployment

Iacopo Morchio*†

JOB MARKET PAPER
Draft: January 15th, 2015

Abstract

Using panel data from the US, I document three new stylized facts on unemployment. First, 10% of workers account for two-thirds of unemployment in prime-age. Second, young unemployment predicts prime-age unemployment. Third, differences in job-finding rates between the most unemployed and the rest increase over the life cycle, while differences in separation rates shrink. I show that a model of heterogeneity across workers and information frictions, in which agents learn workers' types from their labor market history, is quantitatively consistent with all these facts. I find information frictions to be responsible for the whole decrease in job-finding rates of the most unemployed workers over the life cycle. The concentration and persistence of prime-age unemployment are mainly explained by heterogeneity across workers, while information frictions have a negligible role. The model has novel implications for labor market policy: I find that severance payments have asymmetric effects, affect mainly the most unemployed, and reduce the speed of learning later in life.

Keywords: Concentration, Inequality, Learning, Sorting.

*Universidad Carlos III de Madrid. Email: imorchio@eco.uc3m.es Web: http://sites.google.com/site/iacopomorchio
†I am very grateful to Andrès Erosa and Javièr Fernàndez-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, Ángel Hernando Veciana, Timothy Kehoe, Matthias Kredler, Moritz Kuhn, Igor Livshits, Matilde Machado, Ellen McGrattan, Sevi Mora, Fabrizio Perri, Omar Rachedi, Josè-Victor Rios-Rull, Hernan Seoane, Jan Stühler and all participants to the UC3M Ideas Workshop and the UMN Labor Workshop for useful comments and suggestions. Financial support from the Spanish Ministry of Education is gratefully acknowledged.

1 Introduction

Using data from the 1979 National Longitudinal Survey of Youth, I document three novel facts on lifetime unemployment. First, two-thirds of observed prime-age unemployment between 1985 and 2010 is accounted for by 10% of workers[1]. Such concentration is due to both lower job-finding rates and higher job-separation rates of the most unemployed workers. Bad luck alone cannot explain why unemployment is so concentrated: the standard search-and-matching framework is at odds with this fact, because it features too many transitions in and out of employment for the majority of workers. Second, time spent in unemployment when young is a powerful predictor of time spent in unemployment during prime-age. By means of regression analysis, I show that this is not due to observable heterogeneity such as education, occupation, or health. Third, I show that the 10% most unemployed workers and the rest start their careers with similar job-finding rates, and that the job-finding rate of the most unemployed declines over the years while the one of the rest of workers stays relatively constant. Instead, differences in monthly job-separation rates[2] shrink: they start as large as 4 percentage points at age 20 and descend to two percentage points at age 35.

Why are separation rates so heterogeneous and persistent over the life cycle? Why does the job-finding rate of the most frequently separated workers decline? And why do the same workers experience both low finding rates and higher separation rates? The fact that those with a low job-finding rate tend to have a high separation rate is crucial to account for heterogeneity in lifetime unemployment: since unemployment is a nonlinear function of both, theories that account only for one or the other cannot reproduce the concentration of unemployment observed in the data.

The challenge is to build a theory that is consistent with all the micro facts presented above. I propose a directed-search model which succeeds in this regard. In the model workers can be of two types, high and low. A worker's type is initially unobserved by all agents in the market, who are allowed to learn workers' types from labor market histories. This feature allows the model to be consistent with the fact that, while differences in job-finding rates increase over the career, differences in separation rates become smaller. In the model, this

[1] This is true even within education-sex subgroups.

[2] Conditionally on facing a separation, I find that the likelihood of experiencing a firing/layoff/temporary job ended/quit does not vary between lifetime unemployment groups. Thus, this is not because one group was more frequently fired than the other, for instance.

is because workers who face frequent separations when young progressively find fewer jobs, and sort into different jobs to reduce their separation probability.

Information is symmetric: at the start of a worker's career, no agent in the market (including the worker herself) knows her type. Search is directed in the sense that workers decide to search for a job with a certain wage. Upon matching with a firm, workers draw match quality from type-specific distributions, which is constant for the whole duration of the match. Firms write fixed wages contracts and are free to destroy a match at will. Match quality is an experience good as in Jovanovic (1979); output of a match is unobserved until a shock is realized, upon which output becomes the firm's private information. Then, the firm keeps the worker or destroys the match, leaving the worker unemployed. The occurrence of a continuation or separation is observed by the market, which updates the probability that the worker is of high type accordingly. The probability of being high-type formalizes a notion of "résumé" based on the worker's labor market history: 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. Thus workers' types are slowly learned from labor market histories and workers with different résumés apply to different wages. To the best of my knowledge, this is the first model in which wages, job-finding rates, job-separation rates and the speed of learning are all endogenously determined at the same time, because workers are allowed to choose the wages they search for: since each wage entails a job-finding rate, an expected duration of a match, and different updates of the résumé, workers are effectively optimizing over all these trade-offs at once.

Heterogeneity in lifetime unemployment comes from three sources in the model. First, it can be the result of bad luck, because any given type might draw low match qualities, which will ultimately lead to separations. Second, it can be the result of information frictions, that is low-type workers apply to wages that are too high to sustain their match qualities. Third, workers with different labor market histories find jobs at different rates.

I estimate the model using data from the NLSY/79. The model is very successful in reproducing the observed concentration and persistence of unemployment, as well as the patterns of job-finding rates, job-separation rates and wages over the life cycle. The model delivers concentration of unemployment because low-type workers have a higher probability of drawing 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. Information frictions are crucial to match the life-cycle patterns of job-finding and job-separation rates by unemployment groups. I argue that a model based on human capital, rather than information frictions, would be inconsistent with these patterns because it would have the counterfactual implication that differences in separation rates increase over the life-cycle.

I calibrate an array of competing models and find that neglecting heterogeneity across workers makes it impossible to match the concentration and persistence of unemployment observed in the data. While uncertainty in match quality draws helps in matching the life-cycle profile of job-separation rates and the concentration of unemployment, heterogeneity across workers is crucial to match the persistent differences in job-finding and job-separation rates across workers I document. 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.

Information frictions play an important role in the first part of workers' lives. In a quantitative exercise, I shut down information frictions and show that they are responsible for the entire decline in monthly job-finding rates of the top 10% of prime-age unemployed (from 22 % at age 20 to 15 % at age 35). This is because 99% of the top 10% unemployed are low types: while their type is initially unknown, it is slowly revealed by their labor market histories. This translates in progressively lower job-finding rates for these workers. Moreover, wage differentials between workers with very different lifetime unemployment are relatively small at the beginning of the career, and my simulations show that this too can be explained by information frictions. Progressive learning makes gaps in wages between the always-unemployed and the rest widen over the workers' careers, in a similar way as a model with human capital accumulation would. Information frictions also explain a portion of the decline in the separation rates of the most unemployed workers. However, the role of information frictions later in life is negligible: by age 30 types have effectively been learned and most of the concentration and persistence of unemployment after this age are due to heterogeneity across workers.

Finally, my model has new policy implications. In particular, I show that severance payments have nontrivial effects on the speed of learning and unemployment: by increasing unemployment duration, especially for low types, severance payments slow down learning because they give workers fewer opportunities to update their résumé. Moreover, lower separation rates also imply that résumès of low-type workers are being updated less frequently with bad news. Overall, unemployment in prime-age decreases: although the job-finding rate declines, the reduction in job-separation rates and the fact that workers demand lower wages to find jobs faster more than offset this decrease.

This paper 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) were the first to show that most of unemployment is accounted for by workers experiencing long spells of unemployment, rather than by workers going in and out of unemployment. In this paper, I 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, during all their lives. The literature on lifetime unemployment is relatively scarce, possibly due to the limited availability of long panel data. My results on the concentration of unemployment in the US are mirrored in the empirical work of Schmillen and Moller (2012), who use long time series from administrative data from Germany, and in Brooks (2005), who looks at workers in Canada in the years 1993-2001. Neither of these studies compares concentration to what is implied by standard models of unemployment. My approach at studying inequality in unemployment risk is similar to the one used in Michelacci, Pijoan-Mas and Ruffo 2011: using NLSY/79 data, the authors show that unemployment over the lifetime is more unequally distributed than what the standard search and matching framework implies. However, none of the studies above looks at the concentration in prime-age unemployment, nor documents young-prime-age persistence, nor decomposes concentration into job-finding and job-separation rates. A vast literature studies heterogeneity in job-finding rates and unemployment duration, both empirically (Addison and Portugal 1989) and theoretically (Lockwood 1991; Shimer 2008; Gonzalez and Shi 2010; Fernàndez-Blanco and Preugschat 2014; Wiczer 2014).

Second, I develop a model of unemployment and learning from job histories in which wages, job-finding rates, job-separation rates and the speed of learning are all jointly determined in equilibrium; this is also the first model to be estimated on (and to study) lifetime unemployment data. Other models of job search have proposed learning as a candidate explanation for the scars of unemployment (Michaud 2014) and duration dependence in job-finding rates (Gonzalez and Shi 2010). The model I develop shares a mechanism similar to Michaud (2014) regarding separations, but adds résumés, learning from labor market history and heterogeneity in the shape of match quality distributions across types: I find that all these ingredients are important to match heterogeneity in lifetime unemployment. My model's environment is similar to Gonzalez and Shi (2010), but in their model workers are heterogeneous in their ability to find jobs, and learn about their ability by finding jobs or not. Instead, in my model workers are heterogeneous in productivity and learn from their employment history, which maps into differences in job finding rates and job-separation rates. Similarly to Gonzalez and Shi (2010), my model also features a duration dependence relation because workers who have a higher probability of being high types tend to find jobs faster. My results on 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% in the first 3 years of employment. Other empirical work focuses on employer learning as a source of increase in wage heterogeneity over the career: see for instance Kahn and Lange (2013).

This paper also relates to the empirical literature that looks at the effect of unemployment on subsequent earnings and labor market outcomes. Since the pioneering study by Heckman et al. (1980), many papers have addressed whether unemployment leaves "scars" on subsequent wages and increases chances of future unemployment; see for instance Von Wachter, Manchester and Song (2009), Von Wachter and Bender (2006), Barnett and Michaud (2012). See also Couch and Placzek (2010) for a review of the studies on the effects of job displacement on earnings. Other recent studies (Kahn 2010; Oreopoulos et al. 2012) look at individuals who graduated from college during a recession, and find that this has negative, persistent effects on the earnings of otherwise identical workers. My model generates ex-post heterogeneity in labor market outcomes by allowing the market to separate workers using their job history, and in principle could be extended to allow for other additional channels discussed in the literature.

Finally my model 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, I want to understand the sources of heterogeneity in lifetime unemployment, and study the role played by information frictions in determining lifetime outcomes. I find that models without heterogeneity and learning (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 I document.

2 The Data

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

I 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 me a sample of 5422 workers[3]. Further, I restrict attention to the relatively narrower sample of males who are only high-school educated at age 30[4]. This leaves us with a total of 1029 individuals followed for 30 years. However, results are robust to more inclusive definitions of the sample[5].

[3]This is to address measurement error issues when computing lifetime unemployment statistics. I study the extent of measurement error in Appendix B.3.

[4]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 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. In appendix B.2 I show that findings are robust to other education-sex subgroups.

[5]For results on the whole sample, see Appendix B.2.

2.1 Prime-age unemployment is concentrated

I first document that prime-age unemployment is concentrated in relatively few individuals. I start by defining young-age unemployment as the fraction of the work history an individual spent in unemployment, over total weeks employed or unemployed[6], from age 20 to 30:

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

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_i^y is the number of weeks that individual i was either employed or unemployed between ages 20 and 30. Similarly, I define prime-age unemployment as the fraction of work history spent in unemployment from age 35 to 55. Since I 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.

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, I compute the fraction of weeks spent in unemployment by the bottom 90% of the sample[7]:

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

where $1(u_i^p < q_{90}(u^p))$ 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[8].

The 10% most unemployed individuals account for about 2/3 of prime-age unemployment

[6] My definitions are similar to Schmillen and Möller (2012).

[7] These measures are common in the literature on income inequality; see for instance Atkinson (1970). Their application to lifetime unemployment is relatively uncommon, with the exception of Schmillen and Möller (2012) and Brooks (2005).

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

	NLSY/79	Unif. Match	
		300 wks	500 wks
Avg. % time in unemployment	3.6	(target) 3.6	(target) 3.6
Avg. % time in U, excluding top 10%	1.5	2.5	3.1
Avg. % time in U, excluding top 20%	0.6	1.9	2.6
% never unemployed:	56	29	5

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 columns: averages computed by simulating sequences of 300 (column 2) and 500 (3) 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.

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 of the unemployment pool is accounted for by workers staying unemployed, rather than workers going in and out of unemployment.

I then proceed to compute monthly average job-finding/job-separation rates for workers in their primes[9]. I find that the concentration of unemployment is both due to a ($\simeq$ 3 times) lower finding rate[10] and a ($\simeq$ 9 times) higher separation rate for that top 10% (see table 2); 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

$$\tilde{u}^p_{u^p < q_{90}} = \frac{\sum_{i=1}^{N} \left[1(\bar{u}^p_i < q_{90}(u^p)) u^p_i \right]}{\sum_{i=1}^{N} 1(\bar{u}^p_i < q_{90}(u^p))} \tag{3}$$

that is, the average of each individual's prime-age unemployment. The two averages are different since T^p_i differs across individuals, because some are observed for more weeks than others; in particular, there can be a significant difference if $\mathbb{COV}(T^p, u^p) \neq 0$, for instance if those often unemployed tend to be more often out of the labor force. In fact, this is indeed the case (see Appendix B.1). I find that there is relatively 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 10%) if using this second methodology (see Appendix B.1).

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

[10]I describe how I compute job-finding and separation rates in appendix A.

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.

	Top 10%	Rest of Sample	Ratio Top 10 / Rest
Avg. % time in unemployment	29	1.5	19.1
δ: Prob. of U $\rightarrow$ E (monthly%)	8	26	0.3
f: Prob. of E $\rightarrow$ U (monthly%)	3.5	0.4	8.75
Predicted % time in U of top 10%, δ alone:	12		
Predicted % time in U of top 10%, f alone:	5		
Avg. Log Wage (2000)	$\simeq$ - 40%		

Table 2: 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 $u = \delta/(\delta + f)$.

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. If only one person out of 10000 was unemployed, the fact that unemployment is concentrated would not be very interesting. Moreover, it is important to stress that these numbers do not represent accurately 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, because by creating groups based on the amounts 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. 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. I produce simulations of 300 and 500 weeks of transitions because in my NLSY/79 sample I observe prime-age workers for about 700 weeks on average; 95% of workers are observed for more than 470 weeks, and less than 1 % of workers is observed for less than 250 weeks. This is for robustness: increasing the number of simulated weeks leads to worse performance of the standard model, so I construct at least one case that is favorable to it.

Simulations show that the standard model, 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 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.

2.2 Unemployment is persistent over the life-cycle

I now 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 8 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, marital status or IQ.

Little additional information can be obtained by decomposing further the separation rate: using the matched employer-employee dataset available along the NLSY/79, I 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 14 in Appendix).

Finally, notice that such persistence is not due to observable heterogeneity: one might think for instance that such differentials could be explained by differences across occupations (the choice of an "unlucky occupation" when young, as in Schmillen and Möller (2012)) or in health status (worse health means worse labor market outcomes). I perform several batteries of regressions (see Appendix B.4) including ethnic origin, education, prior occupations and current occupations, ex-post health and IQ and find that none of these variables substantially reduces the amount of persistence I observe in the data. This result is particularly

All sample	Rest	Top 10% when (35-55)
Rest	92.68	7.32
Top 10% when (20-30)	65.94	34.06

High-School Workers	Rest	Top 10% when (35-55)
Rest	93.53	6.47
Top 10% when (20-30)	58.82	41.18

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). Source: own calculations on NLSY/79.

strong because current occupations and ex-post health are endogenous to prior labor market experience, and as such are bound to capture part of the persistence of unemployment. For instance, a worker that has been unemployed often when young will tipically work in more unstable occupations in prime-age, and this should capture part of the young-prime-age correlation I find. Similar considerations are valid for ex-post health.

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, from age 20 to age 35, by groups of prime-age unemployment. I want to show that those who have experienced large amounts of prime-age unemployment had different labor market outcomes during the first years of their career too. I compute marginal effects from linear regressions of job-finding rates and job-separation rates on a 4-th degree polynomial on age, controlling for year-specific fixed effects in order to clean the effect of recessions[11]. I can see that, at ages 20-30, the job-separation rate of the top 10% of prime-age unemployed is 4 percentage points higher than the job-separation rate of the rest of the sample (higher than

[11]Results are substantially identical if I compute the averages using 5-years long age groups instead of restricting to a functional form. I choose the polynomial shape for presentation purposes; results under the age-group specification will be used to identify the model and will be reported in figure 4.

the sample average), and this difference declines to 2 p.p at age 35. Instead, between the two groups there is only a 4 percentage points difference (about 1/7 of the sample average) in job-finding rates at age 20, but this difference becomes more pronounced as workers age, particularly because of the decline in the job-finding rate of the top 10% of prime-age unemployed.

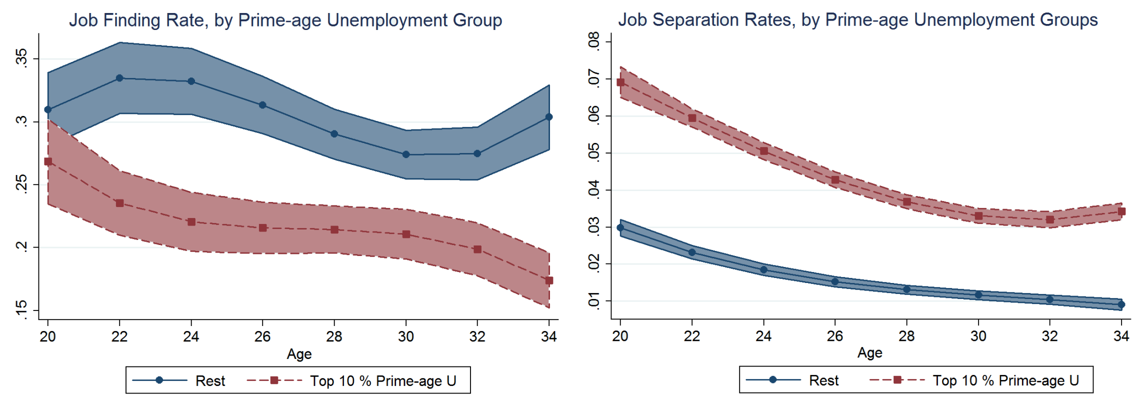

Figure 1: job-finding (left panel) and job-separation (right panel) probabilities, by group of prime-age unemployed. Sample of male, high-school educated workers. Source: own calculations on NLSY/79. Shaded areas are 95% confidence bands.

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, because they were hired with similar probabilities, but such differences became more pronounced later[12]; 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 2), 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 repli-

[12] I address why human capital-based explanations are insufficient to explain such patterns in section 6.

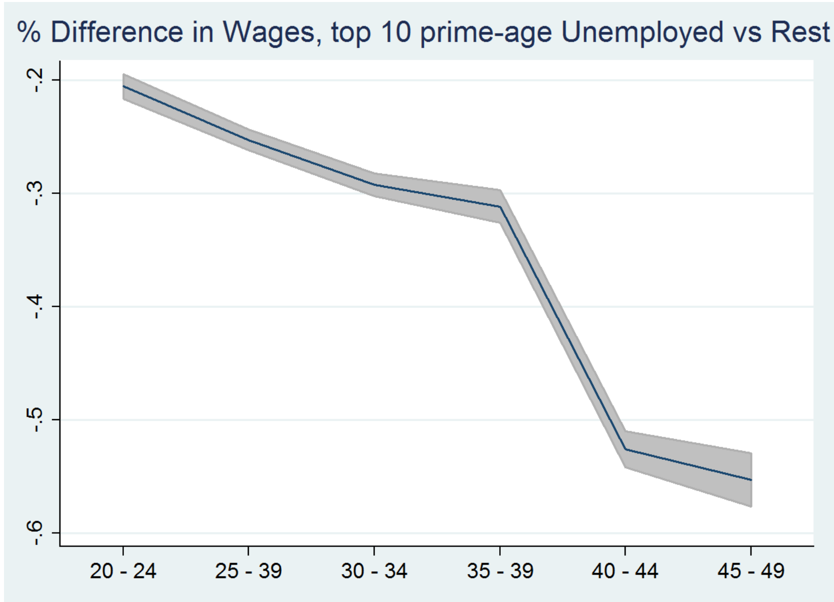

Figure 2: Log difference in Hourly Wage, between Top 10% Prime-Age Unemployment Group and Rest. Sample of male, high-school educated workers. Source: own calculations on NLSY/79. Shaded area is 95% confidence bands.

cating the concentration of unemployment in relatively few workers and the 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, 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 workers' careers, and that workers are being slowly sorted by employers over their careers. My 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.

3 Model

I now proceed to set up a model of heterogeneity in labor market outcomes, roughly based on Delacroix and Shi (2006) and Gonzalez and Shi (2010); the ingredients of such model are inspired by the evidence presented in the previous section.

In order to obtain believable life-cycle profiles of separation rates, I add heterogeneity in match quality draws as in Menzio, Telyukova and Visschers (2012). Heterogeneity across workers, information frictions, and a notion of 'résumé' of the worker are added in order to capture the fact that a group of workers experiences higher separation rates at the start of the career, and that such separation rate diminishes later. This can be because such workers are being separated often (similarly to Gibbons and Katz 1991) and are learning that they have low productivity, thus they sort into lower-paying jobs as to reduce their separation rate. Moreover, such workers progressively find less jobs and earn lower wages: this can be rationalized by the fact that their résumé gets worse with every separation, thus reducing their expected productivity in the eyes of potential employers. Moreover, heterogeneity across workers can rationalize the high levels of persistence of unemployment found in the data.

The quantitative version of the model will feature more ingredients in order to allow to disentangle more mechanisms, but I first present a simpler version with the key ingredients.

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. Every period, a fraction λ 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 probability of a worker being high-type. 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[13]. Wages are perfectly rigid in each submarket; however, firms can destroy a match at will at the beginning of every period. Some matches end randomly with probability δ.

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

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 $\bar{p}$ 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.

3.3 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 $X \subseteq [0, \bar{x}]$, such that $H(x)$ strictly first order stochastically dominates $L(x)$; that is, $H(x) < L(x)\ \forall x < \bar{x}$. Once a 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[14].

At the beginning of a firm-worker match, output of the match is unobserved. Match quality is an experience good as in Jovanovic (1979): 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 $\mathbb{E}(x \mid p)$, until the firm gets to observe the worker's output or a random separation occurs. With probability $1 - \pi$, firms do not observe the worker's output. With probability π, the firm observes the output of the match; by assumption, the occurrence of this event is known to the market if the match continues[15]. Although the occurrence of π in case of continuation is observed by the market, the output produced by the worker is not[16].

[14] 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.

[15] Since match quality is persistent, removing observability of π 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 π: at the end of the first period, he has been observed with probability π 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 $(1 - \pi)^D$ where D is the duration of the match, which then becomes necessary to compute the informational content of job continuation. I 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.

[16] 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.

Denote by $d(w, x)$ the choice of a firm to destroy the match; after observing the worker's output, the firm will either keep the match ($d = 0$) or destroy it ($d = 1$).

Agents gain information on a worker's type during the match's duration. When the firm observes output, 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, π 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

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

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

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 δ 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.

It follows that 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 'résumé'.[17].

[17]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 employment history to calculate

Timing is as follows:

1. Workers die w.p. λ, replaced by unemployed workers with belief $1-l$.
2. W.p. π, firms observe workers' output.
3. Separations (exogenous and endogenous) occur.
4. Workers revise beliefs: $p' = p$ if worker is still unobserved and no shocks occur, $p' = C(w,p)$ if worker has not been observed in the past, is observed today and match continues, $p' = F(w,p)$ if worker has not been observed in the past and worker is separated, $p' = p$ if worker has been observed in the past and is separated.
5. Unemployed workers search for a job. They choose to search in submarket $\{w', p'\}$.
6. Workers match w.p. $f(\theta(w', p))$.
7. Newly matched workers draw match quality from $H(x)$ or $L(x)$ depending on their type.
8. Production occurs, wages are paid.

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

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

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

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

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

p, because unemployed workers with a certain résumé 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 résumé, and those who do make the somewhat extreme assumption that match quality is being observed too (Doppelt (2014)).

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.

3.4 Bellman Equations

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

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

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

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

$$W_k(w,p) = w + \beta\Big[(1-\delta)\,W_k(w,p) + \delta\,U(p)\Big] \tag{9}$$

Define the continuation probability of a worker who has been observed, while working at wage w and prior p, as

$$\chi(w,p) = (1-\delta)\Bigg[\bigg(p\int(1-d(w,x))\,dH(x) + (1-p)\int(1-d(w,x))\,dL(x)\bigg)\Bigg] \tag{10}$$

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

$$\begin{aligned} W_u(w,p) = \;\; & w + \beta\Bigg[(1-\pi)\bigg((1-\delta)\,W_u(w,p) + \delta\,U(p'_F)\bigg) + \\ & \pi\bigg(\chi(w,p)\,W_k(w,p'_C) + (1-\chi(w,p))\,U(p'_F)\bigg)\Bigg] \end{aligned} \tag{11}$$

where I denote 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.

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

$$J_k(w,x) = \max_{d \in \{0,1\}} \Big[(1-d)(x - w + \beta(1-\delta) J_k(w,x)) \Big] \tag{12}$$

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

$$\begin{aligned} J_u(w,p) = \;& \mathbb{E}(x \mid p) - w + \beta(1-\delta)\Big[\pi\big(p \int J_k(w,x)\, dH(x) + \\ & (1-p)\int J_k(w,x)\, dL(x)\big) + (1-\pi)\, J_u(w,p)\Big] \end{aligned} \tag{13}$$

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

$$V(w,p) = -\kappa + q(\theta(w,p))\beta J_u(w,p) \tag{14}$$

and the tightness function must satisfy

$$\kappa \geq q(\theta(w,p))\beta J_u(w,p) \quad \forall w, p \tag{15}$$

which makes θ 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.

3.5 Equilibrium

Definition 1. *a Markov Perfect BRE (Block Recursive 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 ψ
2. $\theta(w,p)$ satisfies 15 $\forall\ w,p$ and $\theta(w,p) \geq 0$ with complementary slackness.

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

6. $W_u(w,p)$ satisfies 11 and $W_k(w,p)$ satisfies 9.

7. $C(w,p)$ satisfies 6

8. $F(w,p)$ satisfies 7

I look only at the Markov Perfect equilibrium to restrict possible off-equilibrium paths on the agents' choices[18]. The BRE 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 BRE, 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.

3.6 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 $p \in [0,1]$, $J_u(w,p)$ is continuous in w, $J_u \in [0, \bar{J}_u]$, and for $J_u \in (0, \bar{J}_u)$, $\partial J_u/\partial w < 0$, $\partial J_u/\partial p > 0$.*

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

$$J_k(w,x) = \begin{cases} \frac{x-w}{1-\beta(1-\delta)} & \text{if } \ w \leq x \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

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

[18] Agents can infer that a worker has applied to wage $w'(p)$ in equilibrium. However, if the equilibrium is not Markov perfect, workers might have an incentive to deviate to other wages to reduce their probability of being separated and look better in the eyes of outside firms in case the match is destroyed. This would introduce asymmetric information (workers know more than firms) and break block recursivity (firms have to form expectations on the mixed strategies of workers). An alternative approach would be to assume that the wage earned in each submarket is observable on the résumé, thus making this equilibrium refinement unnecessary.

$$J_u(w,x) = \frac{\mathbb{E}(x \mid p) - w + \beta(1-\delta)\pi \left[p \int\limits_w^\infty \frac{x-w}{1-\beta(1-\delta)}\, dH(x) + (1-p) \int\limits_w^\infty \frac{x-w}{1-\beta(1-\delta)}\, dL(x) \right]}{1-\beta(1-\delta)(1-\pi)} \tag{17}$$

The right-hand side of expression 17 is decreasing in w because current flow profits are decreasing in w, 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

$$\mathbb{E}(x \mid p) = p \int\limits_0^\infty x\, dH(x) + (1-p) \int\limits_0^\infty x\, dL(x) \tag{18}$$

Since $H(x)$ first-order stochastically dominates $L(x)$, $\int\limits_0^\infty x\, dH(x) > \int\limits_0^\infty x\, dL(x)$, thus current flow profits are increasing in p. As for the future value of the firm, first-order stochastic dominance implies that $\forall w > 0$, $\int\limits_w^\infty x - w\, dH(x) > \int\limits_w^\infty x - w\, dL(x)$, 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 BRE of the economy, the unique solution to equilibrium condition 15 is*

$$\theta(w,p) = \begin{cases} q^{-1}(k/(\beta\, J_u(w,p)) & \textit{if } \; \beta\, J_u(w,p) \geq k \\ 0 & \textit{otherwise} \end{cases} \tag{19}$$

Since the function $J_u(w,p)$ is continuous in w for $w \in [0, \mathbb{E}(x \mid p)]$, the market tightness function θ 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 θ. $\square$

Remark: $C(w,p)$ is continuous in w and $p \;\; \forall w < \bar{x}$, and $\partial C/\partial p > 0$. Moreover, $\partial C/\partial w \leq 0$ whenever $l(w)/(1 - L(w) \leq h(w)/(1 - H(w)$.

Continuity of posterior belief C stems trivially from the functional form and from the fact that the conditional distribution function is always < 1 for $w < \bar{x}$. The fact that the belief function is increasing in p is because of Bayes' rule. Finally, that the belief in case of continuation is increasing in the wage whenever $l(w)/(1 - L(w)) \geq h(w)/(1 - H(w))$ can be easily proved by differentiating the posterior belief function w.r.t. w. Intuitively, if increasing the wage yields a greater *relative* reduction in feasible match qualities for low types than for high types (because $l(w)$ is the density function of match qualities at w for low types), asking for higher wages will yield higher posterior beliefs in case of continuation of the match. Similar results can be obtained for the belief F, thus they are omitted here.

Proposition 1. *The Markov-perfect Block Recursive Equilibrium exists.*

Proof: Conditions 2, 4, 5, 7, 8 of the definition of equilibrium are satisfied by the functions provided above. Looking at the worker's choice, the right-hand side of equation 8 is a contraction mapping on U. The function 9 is continuous and increasing in w; thus the value of employment 11 is continuous because it is the sum of continuous functions and compositions of continuous functions. By the properties of the tightness function discussed in lemma 3.6, the fact that the job-finding probability is continuous and concave in θ, the fact that the value of employment 11 is continuous in w, using standard arguments it can be proved that a unique value function U exists, which is positive, bounded and continuous on $p \in [0,1]$. Furthermore, the set of maximizers $W^*(p)$ is nonempty, closed and upper hemi-continuous. Notice that, differently from Gonzalez and Shi (2010), I do not need to verify that matches will always be accepted to prove the existence of an equilibrium, because workers are learning nothing from job search, thus they apply only to wages that they would accept. This shows that U, W_u and W_k satisfy conditions 3 and 6.

Finally, notice that all functions derived until now are completely independent of the aggregate state ψ, satisfying condition 1. □

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 depends on the trade-off, at given p, between higher wages, lower job-finding probability and higher separation probability.

4 Quantitative Model and Identification

In the quantitative version of the model, I 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 α_h and α_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 xs, 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 estimating parameters in order to replicate features of job-finding rates, job-separation rates and wage patterns observed in the NLSY79. The idea behind identification 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. My model works similarly during a match's duration, so I apply the same strategy but I 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 descovering match quality is $H(w)$ for high types and $L(w)$ 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.

Turning to how 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.

I now explain 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 drew 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 π), while the speed of learning depends on how far apart the two distributions of match quality are: if types draw 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 draw 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 over the life cycle, and with the life-cycle patterns of job-finding rates and job-separation rates by unemployment groups.

Persistence of unemployment depends on how far apart the two distributions of match quality are, how risky they are and how large the measure of low-types is. If low-types have always high risk of being unemployed (that is, 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 why, suppose that low-types are 10% of workers: in that case, persistence as measured by the probability of ending up being in the top 10% of prime-age unemployment, given that one has been in the top 10% of young unemployment, will be 100%. However, there will be no additional persistence at the top 20% because the rest of the population is never unemployed both when young and in prime-age.

Instead, if the two distributions of match quality are very close, persistence of unemployment will depend also on how fast learning is, 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-types present in the economy.

4.1 Calibration

I now 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. Estimation is performed by applying the Simulated Method of Moments: I minimize the loss function

$$L(\omega) = m(\omega)' W \, m(\omega) \tag{20}$$

where ω is the vector of parameters of the model, $m(\omega)$ is a column vector of the differences between the model-generated moments and the data moments, and W is a weighting matrix[19].

I set the model period to be one month. I assume workers are born at age 20, the starting age of my data, and choose the death probability λ in order 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 many other models of directed search (Shimer (2005); Mortensen and Nagypal (2007); Menzio and Shi (2011); Menzio, Telyukova and Visschers (2012)), 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 as including both the value of leisure and unemployment benefits, and 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[20] with scale parameters σ_H, σ_L and shape parameters ϕ_H, ϕ_L. 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 calibrated 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

[19] Computation of variance-covariance matrix of moments and of standard errors is not trivial, because one moment restriction comes from the estimates of Hall and Milgrom (2007) and its covariance with the remaining moments cannot be computed. At the moment, W is set in such a way that moments are scaled to their data average, that is 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.

[20] The Weibull distribution is a common choice in this regard. See for instance Menzio, Telyukova and Visschers (2012).

2. Notice however that the model is unit-free, so one of the scales has to be set exogenously. I normalize $\sigma_H = 1$.

The skill multiplier s, the probabilities α_h and α_l of being skilled are calibrated 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 κ is calibrated as to match the job-finding rate of bottom 90% of the prime-age unemployment distribution at ages 20-25.

Parameter	Symbol	Value	Target	Data	Model
Scale parameter, high type	σ_H	1.0000	Wage profile diff, top 10 % - rest,		
Shape parameter, high type	ϕ_H	41.7871	JS rate profiles,		
Scale parameter, low type	σ_L	0.5070	JF rate profiles,		
Shape parameter, low type	ϕ_L	2.2090	% prime-age U acc. by top 10	0.60	0.62
Prob. of observing output	π	0.0636	% prime-age U acc. by top 20	0.83	0.84
Prob. of low skilled	α_l	0.7313	Initial W diff, top 10 % - rest,	-0.21	-0.18
Prob. of high skilled	α_h	0.9681	JF rate profiles		
Skill Multiplier	s	1.0911	Wage profile diff, top 10 % - rest		
Flow value of Unemployment	b	0.5652	Ratio b/avg. wage	0.70	0.61
Vacancy cost	κ	2.8425	job-finding rate, bottom 90, age 20-25	0.29	0.29
Random separation rate	δ	0.0027	JS rate, bottom 90, age 40-50	0.007	0.007
Measure of low types	l	0.3447	Persistence of U, top 10	0.41	0.36
			Persistence of U, top 20	0.45	0.48

Table 4: Baseline calibration results. Targets calculated on NLSY/79.

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 11 parameters with 36 restrictions.

5 Results

5.1 Calibration results

Despite being calibrated with over-identifying restrictions, the model does a very good job in replicating the main features of the data. As can be seen in table 4, the model is quite capable of delivering realistic amounts of concentration and persistence of unemployment. The model fits very well the persistence as measured by the Markov transition matrix between being unemployed when young and when prime-age: the probability of being in the top 10% of the unemployment distribution when prime-age, after having been in the top 10% of the unemployment distribution when young, is 0.36 in the model and 0.41 in the data. At the top 20%, the same statistic is 0.48 in the model and 0.45 in the data.

The model matches almost perfectly the observed concentration of the distribution of prime-age unemployment: the top 10% accounts for 60% of prime-age unemployment both in the model and in the data, while the top 20% accounts for more than 80% of prime-age unemployment in the model and in the data. The standard Mortensen-Pissarides model only predicts one-third of observed concentration at the top 10%.

The match quality distributions of high- and low-type workers are substantially different: at the calibrated values, the match quality distribution of low types has more 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 3).

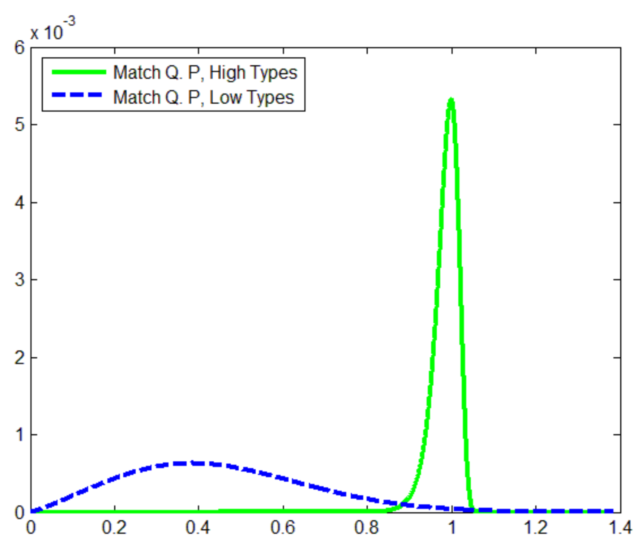

Figure 3: Match quality distribution of high types (red) and low types (blue), under baseline calibration

The calibrated value of the probability of a firm observing the worker's output $\pi = 0.0636$ implies that the average duration of a "bad match" is about 16 months.

The calibrated measure of low-type workers in the economy is around 34%, 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 9% higher productivity; the probability that a high- and low-type workers are skilled are, respectively, 97% and 73%, making unskilled workers a minority among both low and high types. 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 a low-type[21].

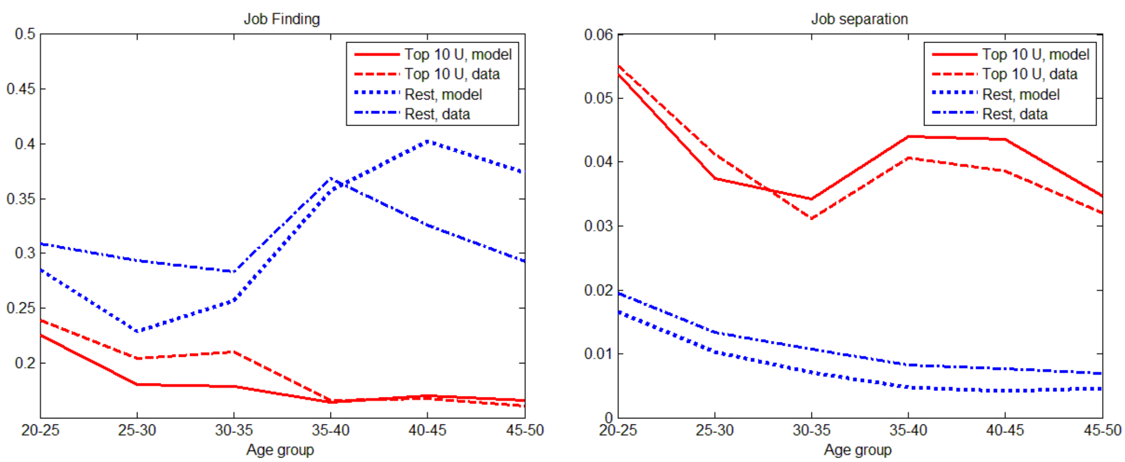

Figure 4: Model versus data; job-finding rates (left) and job-separation rates (right) of top 10% prime-age unemployed (red) vs rest (blue). Results under baseline calibration.

Figure 4 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 is very successful in fitting the patterns of job-separation and job-finding rates by prime-age unemployment groups, both for the most unemployed and the rest. The model does also a very good job in explaining wage differentials between the top

[21]By Bayes' rule

$$\begin{aligned} p_{\text{skilled}} &= \frac{\alpha_h(1-l)}{\alpha_h(1-l)+\alpha_l l} \\ p_{\text{unskilled}} & \frac{(1-\alpha_h)(1-l)}{(1-\alpha_h)(1-l)+(1-\alpha_l)l} \end{aligned}$$

Thus a skilled worker starts with prior 0.7156 and an unskilled worker with prior 0.1841 .

10% and the rest until age 40, after which it explains only two-thirds (figure 9 in Appendix).

Figure 5 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 I term "learning over the life cycle": as separations and continuations occur, the market slowly learns 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.

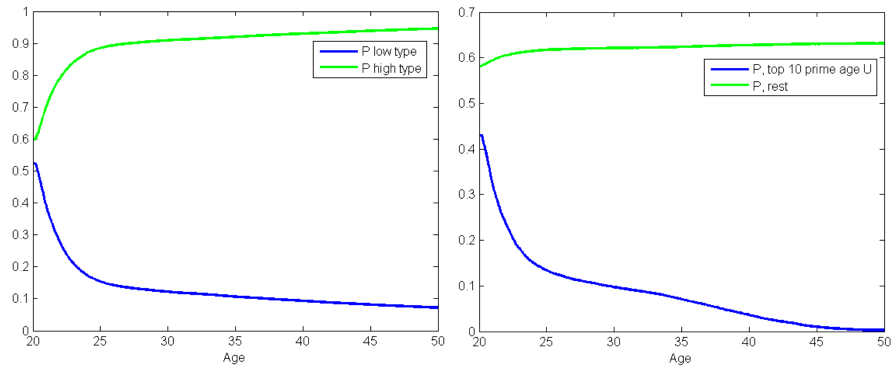

Figure 5: Probability of being a high type: by high/low type (left) and by top 10% of prime-age unemployment (right). Model results under baseline calibration.

Let us look first at the job-finding rate: as the market learns who are low- and 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 6). The result follows from this mechanism, and from the fact that more than 80% of the unemployment pool is made of low types (figure 10 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-finding rate of the bottom 90% when young, and overshoots their job-finding rate 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

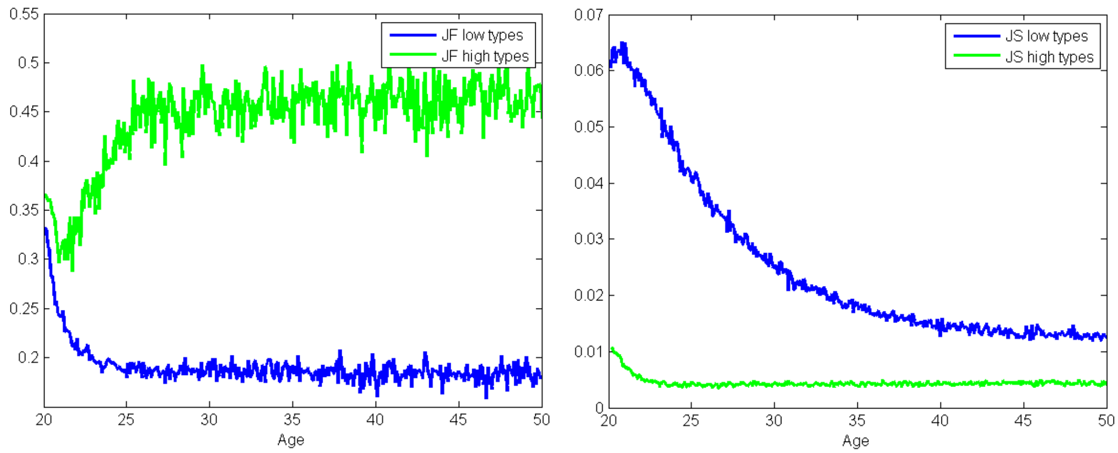

Figure 6: job-finding rates (left) and job-separation rates (right) of low types (blue) vs high types (green). Model results under baseline calibration.

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. 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 most unlucky of low-type workers.

5.2 Counterfactual Simulations

I now simulate what would happen in alternative scenarios, removing model features one by one to study their relative importance in fitting the data. Results are summarized in table 5. All models have been recalibrated on the same loss function of the baseline model.

First, I calibrate a version of the model featuring only some observable differences in productivity, no unobserved heterogeneity in productivity, nor any uncertainty on match quality (column 1): the model is completely incapable of replicating the concentration of unemployment observed in the data (at the top 10%, 27% against 60%), although it picks some persistence thanks to observable skills, which translate in persistent differences in job-finding rates. Separation rates are way off those observed in the data, both for the most unemployed (1.26% against 5.51% at age 20-25) and the rest (1.26% against 1.99% at age 20-25): this is because there is no uncertainty in match quality. This model is akin to a

	Baseline	Data	(1) 1 type $\mathbb{VAR}(x) = 0$	(2) 1 type $\mathbb{VAR}(x) > 0$ HC	(3) 2 types $\mathbb{VAR}_l = 0$ $\mathbb{VAR}_h = 0$	(4) 2 types $\mathbb{E}(x\|l) =$ $\mathbb{E}(x\|h)$
Separation rates, top 10						
age 20-25	5.37	5.51	1.26	2.95	4.48	4.29
age 25-30	3.76	4.12	1.26	2.16	2.35	3.54
age 40-45	4.37	3.86	2.30	4.15	2.54	4.60
Separation rates, rest						
age 20-25	1.99	1.95	1.26	2.71	2.61	3.66
age 25-30	1.28	1.34	1.26	1.58	1.58	2.07
age 40-45	0.75	0.77	1.27	0.79	1.29	0.62
Finding rates, top 10						
age 20-25	22.46	23.88	16.59	21.13	20.30	23.37
age 25-30	17.99	20.41	16.59	22.49	13.10	21.63
age 40-45	16.98	16.73	13.54	17.83	12.43	18.15
Finding rates, rest						
age 20-25	28.57	30.90	29.71	21.27	27.95	24.24
age 25-30	22.89	29.33	29.79	22.73	27.19	22.39
age 40-45	40.11	32.62	35.27	28.72	33.86	29.42
Concentration - Prime-age						
U acc. by top 10	62.10	60.00	27.17	43.67	31.27	58.04
U acc. by top 20	83.69	83.00	41.73	73.43	43.89	89.11
Persistence						
Top 10 young - Top 10 primes	36.49	41.00	29.72	12.79	37.04	16.47
Top 20 young - Top 20 primes	48.49	45.00	33.33	26.03	40.12	29.57
Δ Wage top 10 - rest						
age 20-25	-17.77	-20.57	-25.81	-0.09	-11.20	-1.06
age 40-45	-32.03	-52.50	-27.78	-1.72	-23.45	-2.59

Table 5: Baseline calibration results vs counterfactuals. Column 1 is a model with no unobserved heterogeneity, 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.

standard search-and-matching model with observable differences in productivity, and as such can generate heterogeneity in lifetime unemployment only through differences in job-finding rates.

I then augment the model with uncertainty about match quality, on-the-job human capital accumulation and stochastic human capital depreciation when unemployed (column 2): such a model improves significantly over the first one, in particular because uncertainty in match quality draws allows to get closer to the data in terms of concentration of prime-age unemployment (at the top 10%, 44% against 60% in the data and 27% in model 1). Moreover, uncertainty about match quality allows to replicate some of the life-cycle profile of separation rates: the top 10% prime-age unemployed start with a job-separation rate of 2.95 at age 20-25, against 5.51 in the data, and the separation rate at age 40 for these workers is almost matched (4.15 against 3.86). However, such a model fails completely in delivering sufficiently large persistence of unemployment (at the top 10%, 13% against 41%). This is because human capital accumulation and depreciation introduce "reshuffling" in the skill level of workers: 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 get a sufficiently high level of match quality: this reduces the persistence of unemployment over the life cycle with respect to model 1. Moreover, model 2 cannot replicate the patterns of wage differentials and differences in separation rates at young ages by unemployment groups, because it does not feature enough heterogeneity in match quality draws: the bottom 90% of workers have about the same separation rate of the most unemployed at ages 20-30, differently from the 3.6% difference existing in the data. Finally, such model cannot replicate the fact that the job-finding rate of the most unemployed falls at the beginning of their career.

I now calibrate a model with no human capital, heterogeneity in mean productivity across types, but no uncertainty about match quality (column 3): that is, I estimate a model forcing the distribution of match quality to be degenerate. Such model can predict a higher separation rate at age 20-25 for one group of workers, because some workers are initially applying to wages that are too high to sustain their match quality (top 10% separation rate is 4.48 in this model against 5.51 in the data). However, the decrease is too sudden: at age 25-30, the separation rate of the most unemployed is already almost identical to the one of the rest of workers. This is because learning is too fast and bad luck plays little role: when there is no uncertainty about match quality, a worker who asks for a wage above the productivity of low types will learn his type with very high precision[22] at the first separation or continuation. Moreover, since the distribution of match quality is degenerate, there is no

[22]In the case of separations, there is still slight uncertainty about the type because of exogenous separations. Continuations instead immediately result in a probability 1 of being high-type.

other mechanism that delivers heterogeneity in separation rates. Such model also predicts a too sudden and too large decrease in job-finding rates, which become as low as 13% at age 25 against 20% in the data. This is another consequence of excessively fast learning, and of the struggle of the model in delivering concentration of unemployment by having to rely solely on heterogeneity in job-finding rates. Finally, this model can improve over the model without heterogeneity (model 1) in delivering persistence of unemployment (37% against 36% in the data at the top 10%), but fails dramatically in delivering concentration, underperforming even the model without heterogeneity (model 2) in this regard; this is because there is no uncertainty in match quality.

In the last experiment, I calibrate a model with heterogeneity in the variance of the match quality distribution, but no differences in mean productivities[23] (column 4). Such model delivers separation rates that are closer to the data for the top 10% of prime-age unemployed (4.29 against 5.51 in the data), but predicts too high separation rates for the rest of workers when young (3.66 against 1.95 at ages 20-25). Moreover, such model fails in delivering sufficient young-prime-age persistence of unemployment (at the top 10%, 17% against 36% in the data), as well as wage differentials that are consistent with the data. This is because differences in the distributions of match quality translate into relatively large differences in separation risk, but also in relatively small differences in job-finding rates, and in even smaller wage differentials.

These quantitative exercises confirm that all 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 and wages. 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 translates into a more realistic descent of job-finding rates and job-separation rates for the most unemployed workers. Notice that even a model with homogeneous types, but uncertainty in match quality draws (model 2), is capable of delivering concentration of unemployment: this is because such concentration can be obtained if there is sufficient

[23] In practice, this is done by letting the shape parameters ϕ_i of the Weibull distributions be estimated freely by the algorithm, while 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.

heterogeneity in separation rates across workers, which can be the consequence of bad luck in match quality draws.

5.3 Decomposing Learning over the Life Cycle

In this section I keep the baseline calibration but shut down model features pertaining to the learning structure to understand their importance in explaining the data. Results are presented in table 6

First, I shut down entirely learning over the career, by making types already known at the beginning. This reduces the separation rate of the most unemployed at age 20-30 by 1.2 percentage points (one-fifth), because these workers are already aware that they are low types and apply to lower wages as to avoid frequent separations. However, it increases the separation rate of the rest of workers by one-tenth, because most workers know they are high types and, having a high finding rate, have an incentive to "gamble" for higher levels of match quality by asking for higher wages: they have nothing to lose from separations because their type is already known. The persistence and concentration of unemployment are barely affected, as these are mainly due to heterogeneity across workers and bad luck in drawing match quality values. Finally, information frictions account for the whole decline in job-finding rates from age 20 to 40 for the most unemployed workers, and for the increase in wage differentials over the life-cycle between the most unemployed and the rest: if types were already known, low types would ask for lower wages and find jobs with lower probability right from the start. That information frictions are responsible for these two facts at a time is one of the most important results of the paper.

In the second experiment, I keep information frictions but shut down the initial signals given by skills, rendering them uninformative[24]. I find that initial signals have a relatively small influence on separation rates and job-finding rates, which would both be higher in the absence of such signals for the 10% most unemployed workers at ages 20-25. This is because low-type workers are more likely to be unskilled: removing the initial information of skills, such workers are thought to be high types with a higher probability, thus find jobs faster, but also get separated more frequently because they apply to higher wages. Removing signals

[24]That is, I keep differences in productivity implied by skills but set the priors of both skilled and unskilled workers to $1 - l$, the population mean.

	(1) Types known	(2) No Info Priors	(3) No skill diff	Baseline
Separation rates, top 10				
age 20-25	4.31	5.51	5.51	5.37
age 25-30	3.14	3.88	3.96	3.76
age 30-35	3.07	3.52	3.67	3.44
age 40-45	4.22	4.48	4.60	4.37
Separation rates, rest				
age 20-25	2.19	1.98	1.98	1.99
age 25-30	1.17	1.30	1.27	1.28
age 30-35	0.89	0.97	0.96	0.96
age 40-45	0.73	0.75	0.74	0.75
Finding rates, top 10				
age 20-25	17.77	23.60	19.56	22.46
age 25-30	17.47	17.86	14.26	17.99
age 30-35	17.33	17.74	14.16	17.80
age 40-45	16.90	16.82	13.30	16.98
Finding rates, rest				
age 20-25	30.33	29.01	24.78	28.57
age 25-30	24.63	22.67	18.22	22.89
age 30-35	26.84	25.20	20.69	25.57
age 40-45	41.74	39.42	33.59	40.11
Concentration - Prime-age				
U acc. by top 10	62.76	60.67	61.62	62.10
U acc. by top 20	83.07	83.76	84.89	83.69
Persistence				
Top 10 young - Top 10 primes	35.67	36.40	35.28	36.49
Top 20 young - Top 20 primes	49.62	48.78	50.00	48.49
Δ Wage top 10 - rest				
age 20-25	-34.58	-15.66	-14.55	-17.77
age 40-45	-34.22	-31.97	-30.02	-32.03

Table 6: Baseline results vs counterfactuals. Column 1: types are known from the beginning. Column 2: types are unknown and skills are uninformative. Column 3: types are unknown, skills are informative but do not give productivity differentials. All numbers are percentage points.

have almost no effect on the remaining 90% of workers. Concentration and persistence of unemployment are barely affected by the removal of signals. Finally, wage differentials would be lower at the start of the career, because low-type workers are thought to be high types with a higher probability, thus apply to higher wages.

In the last experiment, I keep the informative content of skills but cancel the productivity gain they imply, by setting $s = 1$. This reduces the job-finding rate of all skilled workers, because it reduces their expected productivity. The effect on the job-separation rate of the 10% most unemployed workers is even more pronounced than in the case of signals, because now even skilled low-type workers apply to wages that are close to their reservation value, thus increasing their expected separation rate at all points of the life cycle with respect to the baseline case. Again, concentration and persistence of unemployment would be barely affected, while wage differentials at the start of the career would be lower because of a pure composition effect: since high types were more likely to be skilled, they applied to higher wages in the baseline case.

In conclusion, I find information frictions to be important for explaining separation rates at age 20-30 and the patterns of job-finding rates, both for the most unemployed workers and the rest. Moreover, information frictions are responsible for the whole increase in the wage differential between the most unemployed workers at the rest until age 40, and for two-thirds of such differential later on. Observable heterogeneity in productivity have a secondary role for explaining the patterns of job-finding rates, job-separation rates and wages. However, information frictions and observable skills have a negligible role for explaining the concentration and persistence of unemployment over the life cycle, which are mainly due to unobserved heterogeneity across workers and bad luck within the group of low-type workers.

5.4 Duration dependence

The model is also capable of reproducing a duration dependence relation in job-finding rates (figure 7), 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 by workers with lower market priors. I plan to expand this section in the future by decomposing the duration dependence relation in effects of learning vs observable skills.

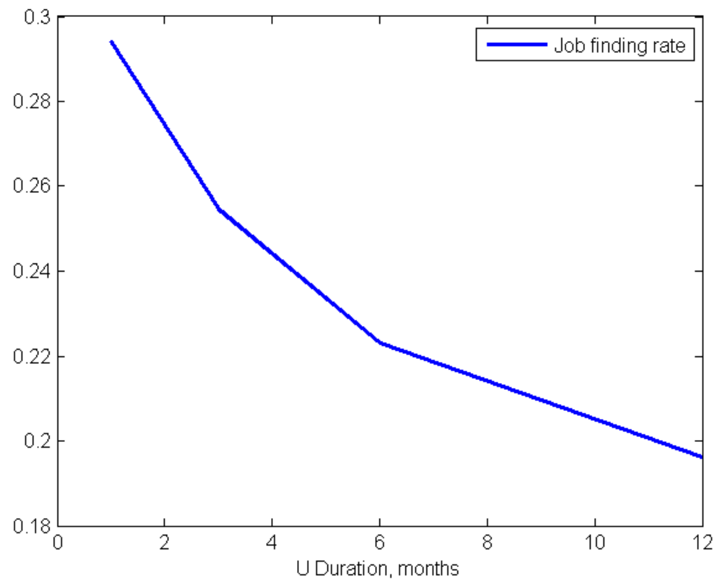

Figure 7: Model-generated data: duration dependence relation in job-finding rates, at 1, 3, 6 and 12 months of unemployment duration.

6 Discussion

6.1 Heterogeneity or Human Capital?

I have shown that a theory of information frictions and heterogeneity is capable of explaining at the same time the patterns of job-finding rates, job-separation rates and (part of) the patterns of wages by unemployment groups 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 face human capital depreciation (as in Ljunqvist 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, depreciation would imply that heterogeneity in job-separation rates *rises* along the career, because the most unemployed would lose human capital and possibly face higher separation rates, while the rest of workers would experience fewer separations. As a result, we should observe a divergence in separation rates by unemployment groups, and not a convergence such as the one I document.

Column 2 of table 5 partially tests for these implications by calibrating a version of the model featuring no ex-ante heterogeneity across workers, match quality draws, fixed wages and human capital accumulation/depreciation as the only source of persistence in unemployment. Such model delivers only a fraction of the concentration of prime-age unemployment observed in the data, almost no young-prime-age persistence, and patterns of job-finding rates and job-separation rates that are inconsistent with the data.

6.2 Policy implications

Preliminary simulations show that the model has novel policy implications, particularly in regard to severance payments. The introduction of severance payments, even of a relatively mild size (three months of average wages), has several effects. It decreases firms' expected profits, particularly when matching with not-yet-sorted workers and low types, because high types are being separated so rarely that they are scarcely affected by the introduction of severance payments. Thus, unemployment duration for low-type workers increases (their job finding rate lowers by about 1 percentage point). Moreover, the endogenous response of workers to lower job-finding rates is to ask for lower wages (by about 3.6%).

As a result, the speed of learning later in life decreases for three reasons. First, workers stay unemployed longer, thus having less opportunities to update their résumé. Second, firms destroy matches less frequently, so that low-type workers receive bad news less frequently. Third, those who are likely to be low types ask for lower wages, thus increasing the probability that more match qualities support that wage.

Finally, severance payments effects are strongly asymmetric: high types are scarcely affected by the policy, because they are very unlikely to be separated in the first place. The policy has its strongest effects among not-yet-sorted workers and low types. It lowers dramatically their job-finding rate (almost halved in the case of a severance payment equal to 5 months of wages), because such workers typically draw low values of match quality, thus a firm that hires them has a high probability to realize negative profits. Almost all workers benefit from a large reduction in separation rates, so much so that unemployment decreases, on average. However, output drops significantly, particularly for low-type workers (-11% when severance payments equal five months of average wages).

	3 months SP	5 months SP	Baseline
Output, Steady State			
Low types	0.63	0.59	0.66
High types	1.02	1.02	1.02
Wages, Steady State			
Low types	0.79	0.78	0.82
High types	0.99	0.98	1.00
Priors - Low types			
Age 25	10.51	11.32	9.34
Age 30	7.88	8.89	5.80
Age 35	6.84	7.72	5.11
Priors - High types			
Age 25	94.6	94.27	95.20
Age 30	96.0	95.52	97.10
job-separation rate - Low types			
Steady State	1.03	0.75	1.39
Age 20-25	4.82	4.43	5.38
Age 25-30	2.31	1.76	3.20
Age 40-50	0.94	0.71	1.34
job-separation rate - High types			
Steady State	0.31	0.27	0.37
Age 20-25	0.35	0.30	0.55
Age 25-50	0.31	0.28	0.43
job-finding rate - Low types			
Steady State	17.31	13.41	17.61
Age 20-25	20.50	17.88	21.83
Age 25-30	17.30	14.14	18.65
Age 40-50	17.04	14.70	18.30
job-finding rate - High types			
Steady State	33.13	31.42	
Age 20-25	34.21	33.93	36.36
Age 25-50	39.46	39.24	46.17
Unemployment rate - Low types			
Steady State	6.67	6.61	8.32
Age 20-25	22.87	23.88	23.52
Age 25-30	12.35	11.99	15.09
Age 40-50	5.23	4.63	6.86
Unemployment rate - High types			
Steady State	1.52	1.46	1.59
Age 20-25	5.54	5.45	6.03
Age 25-50	0.79	0.74	0.93
Persistence of Unemp.			
Top 10 young - Top 10 prime-age	33	30	36
Top 20 young - Top 20 prime-age	43	36	48
Δ Wage top 10 - rest	-20	-20	-18

Table 7: The impact of introducing severance payments into the model. Column 1: 3 months severance payment. Column 2: 12 months severance payment. Results are averages of simulated data under baseline calibration. All rates are percentage points. High types age categories collapsed into 25-50 because of no differences across age categories after age 25.

7 Conclusions

Using NLSY/79 data, I 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. I build a model that delivers both high concentration of unemployment during prime-age and persistence of unemployment over the life-cycle, and that is consistent with the patterns of job-finding rates and job-separation rates by prime-age unemployment groups. The model delivers such result by 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 delivers a too high wage gap between different workers at the start of their work life, a higher separation rate than the one observed in the data for young workers, and a flat job-finding rate for the most unemployed workers in prime-age. Finally, I find that unobserved heterogeneity, rather than differences in observed skills, is responsible for the bulk of my results.

Preliminary simulations also show that the model has novel policy implications. Severance payments have asymmetric effects: they affect mostly the most unemployed workers, and decrease the speed of learning later in life.

References

Addison, John T., & Portugal, Pedro. 1989. Job Displacement, Relative Wage Changes and the Duration of Unemployment. *Journal of Labor Economics*, **7**(3).

Atkinson, Anthony B. 1970. On the Measurement of Inequality. *Journal of Economic Theory*, **2**.

Barnett, Justin, & Michaud, Amanda. 2012. Wage Scars from Job Loss. *Working Paper.*

Brooks, Bradley. 2005. Chronic Unemployment: a Statistical Profile. *Statistics Canade, Analysis in Brief*, **31**.

Chèron, Arnaud, Hairault, Jean-Olivier, & Langot, FranÃ§oise. 2013. Life-Cycle Equilibrium Unemployment. *Journal of Labor Economics*, **31**(4).

Clark, Kim B., Summers, Lawrence H., Holt, Charles C., Hall, Robert E., Baily, Martin Neil, & Clark, Kim B. 1979. Labor Market Dynamics and Unemployment: A Reconsideration. *Brookings Papers on Economic Activity*, **1979**(1).

Couch, Kenneth A., & Placzek, Dana W. 2010. Earnings Losses of Displaced Workers Revisited. *American Economic Review*, **101**(1).

Delacroix, Alain, & Shi, Shouyong. 2006. Directed Search On the Job and the Wage Ladder. *International Economic Review*, **47**(2).

Doppelt, Ross. 2014. The Hazards of Unemployment. *Working Paper*.

Fernández-Blanco, Javiér, & Preugschat, Edgar. 2014. Unemployment Duration, Exit Rates and Re-employment Wage. *Working Paper*.

Gibbons, Robert, & Katz, Lawrence F. 1991. Layoffs and Lemons. *Journal of Labor Economics*, **9**(4).

Gonzalez, Francisco M., & Shi, Shouyong. 2010. An Equilibrium Theory of Learning, Search and Wages. *Econometrica*, **78**(2).

Hall, Robert E., & Milgrom, Paul R. 2007. The Limited Influence of Unemployment on the Wage Bargain. *American Economic Review*, **98**(4).

Heckman, James J., & Borjas, George J. 1980. Does unemployment cause future unemployment? Definitions, questions and answers from a continuous time model of heterogeneity and state dependence. *Economica*, **47**(187).

Hornstein, Andreas. 2012. Accounting for Unemployment: the Long and Short of It. *Working Paper*.

Jovanovic, Boyan. 1979. Job Matching and the Theory of Turnover. *Journal of Political Economy*, **87**(5).

Kahn, Lisa B. 2010. The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics*, **17**(2).

Kahn, Lisa B., & Lange, Fabian. 2013. Employer Learning, Productivity and the Earnings Distribution: Evidence from Performance Measures. *Review of Economic Studies, Forthcoming.*

Lange, Fabian. 2007. The Speed of Employer Learning. *Journal of Labor Economics*, **25**(1).

Ljunqvist, Lars, & Sargent, Thomas J. 1998. The European Unemployment Dilemma. *Journal of Political Economy*, **106**(3).

Lockwood, Ben. 1991. Information Externalities in the Labour Market and the Duration of Unemployment. *The Review of Economic Studies*, **58**(4).

Menzio, Guido, & Shi, Shouyoung. 2011. Efficient Search on the Job and the Business Cycle. *Journal of Political Economy*, **119**(3).

Menzio, Guido, Telyukova, Irina, & Visschers, Ludo. 2012. Directed Search over the Life Cycle. *Working Paper.*

Michaud, Amanda. 2014. An Information Theory of Worker Flows and Wage Dispersion. *Working Paper.*

Michelacci, Claudio, & Ruffo, Hernan. 2013. Optimal Life-Cycle Unemployment Insurance. *Working Paper.*

Michelacci, Claudio, Pijoan-Mas, Josep, & Ruffo, Hernan. 2011. Inequality in Unemployment Risk and in Wages. *Working Paper.*

Mortensen, Dale T., & Nagypal, Eva. 2007. More on Unemployment and Vacancy Fluctuations. *Review of Economic Dynamics*, **10**(3).

Mortensen, Dale T., & Pissarides, Christopher A. 1994. Job Creation and Job Destruction in the Theory of Unemployment. *Review of Economic Studies*, **61**(3).

Oreopoulos, Philip, von Wachter, Till, & Heisz, Andrew. 2012. The Short- and Long-Term Career Effects of Graduating in a Recession. *American Economic Journal: Applied Economics*, **4**(1).

Schmillen, Achim, & Möller, Joachim. 2012. Distribution and determinants of lifetime unemployment. *Labour Economics*, **19**(1).

Shimer, Robert. 2005. The Cyclical Behavior of Unemployment and Vacancies. *American Economic Review*, **95**(1).

Shimer, Robert. 2008. The Probability of Finding a Job. *American Economic Review P&P*, **98**(2).

Shimer, Robert. 2012. Reassessing the Ins and Outs of Unemployment. *Review of Economic Dynamics*, **15**.

von Wachter, Till, & Bender, Stefan. 2006. In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers' Careers. *American Economic Review*.

von Wachter, Till, Song, Jae, & Manchester, Joyce. 2009. Long-Term Earnings Losses Due to Mass Layoffs During the 1982 Recession: An Analysis Using U.S. Administrative Data from 1974 to 2004. *Mimeo*.

Wiczer, David. 2014. Long Term Unemployment: Attached and Mismatched? *Working Paper*.

A Appendix

A.1 Construction of job-finding and job-separation probabilities

Following Clark and Summers (1979) and Wiczer (2014), I 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, I use the formulas

$$F_j = \frac{\sum_{i \in N_j} \sum_{t=1}^{U_i^p} f_{i,t}}{\sum_{i \in N_j} U_i^p} \tag{21}$$

$$S_j = \frac{\sum_{i \in N_j} \sum_{t=1}^{E_i^p} s_{i,t}}{\sum_{i \in N_j} E_i^p} \tag{22}$$

where $f_{i,t}$ 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_i^p is the number of weeks worker i was unemployed during prime-age; and E_i^p is the number of weeks worker i was employed during prime-age.

Dep. Variable: % U when 35-55

	(1) Only HS, Males	(2) All, with controls	(3) All, No Young U
% U when young (20-30)	0.3635***	0.239***	
	(0.020)	(0.02)	
N	1029	3127	3127
R^2	0.186	0.218	0.127

p-values in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Source: own calculations on NLSY/79. Regression of % of prime-age unemployment on % of young unemployment: only for high-school educated males (1), 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.

B Supplementary Data Analysis

B.1 Impact of Labor Force Participation

In this subsection I investigate whether the labor force participation margin is relevant for the results I present on the concentration of unemployment. One possibility is that the most unemployed individuals get discouraged about their possibilities of finding jobs; thus, they might tend to permanently drop out of the labor force more frequently than the rest of the sample. To study whether this is the case, I compute the average participation rate of individuals by unemployment groups. It is easy to see that the top 10% of prime-age unemployed tends to participate less often to the labor force. The two groups follow a substantially parallel trend until age 40, after which the top 10% do tend to drop out of the labor force more frequently. However, when we look at individuals who did not participate for a full year, this difference reduces dramatically, suggesting that although the top 10% tends to spend more time out of the labor force, this does not mean that they always drop out completely.

It is unlikely that changes in sample composition are driving most results on the concentration of unemployment; however, I address this concern by studying how much the participation margin matters for computing lifetime statistics and the concentration of unemployment. As explained in section 2, another possible way to compute the average by unemployment groups is

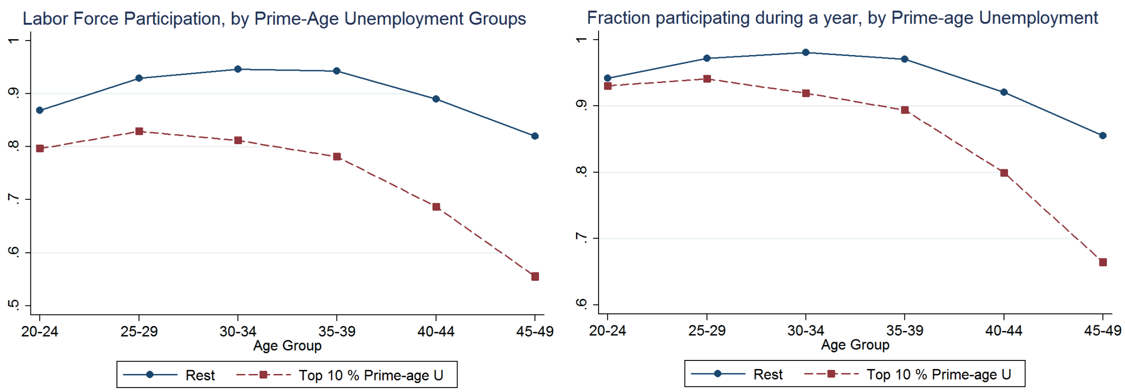

Figure 8: Labor Force Participation rate; by prime-age unemployment groups and by 5-years age groups. **Left**: fraction of weeks spent either in employment or unemployment by group. **Right**: fraction of individuals in group who worked at least a week during a year.

$$\tilde{u^p}_{u^p<q_{90}} = \frac{\sum_{i=1}^{N}\left[1(\bar{u}_i^p < q_{90}(u^p))u_i^p\right]}{N} \tag{23}$$

so that every individual has the same weight in the computation of the average, regardless of the number of periods he has been employed or unemployed. I will refer to this as the *equally-weighted* average, and to the average presented in the paper as the *participation-weighted* average.

In principle, it is not clear which of the two averages should be used. Since the top 10% of prime-age unemployed tends to be out of the labor force more often, these individuals have a lower weight in the participation-weighted average than in the equally-weighted average. Thus, the latter represents the concentration of unemployment if we were to observe the top 10% in the labor force as often as the rest of the sample. With respect to this logic, the participation-weighted average is likely to bias downward my estimates of aggregate prime-age unemployment, and of the concentration of unemployment. As shown in table 9, the equally-weighted formula indeed imply substantially identical averages when excluding the most unemployed, but a higher average of overall prime-age unemployment. Thus, this implies a higher concentration of unemployment, compared to the results with participation-weighted averages; the performance of the standard model is even worse in delivering concentration of unemployment when using equally-weighted averages for comparison.

	NLSY/79	Unif. Match
Avg. % time in unemployment	4.6	(target) 4.6
Avg. % time in U, excluding top 10%	1.6	3.1
Avg. % time in U, excluding top 20%	0.6	1.9

Table 9: Left column: **equally-weighted** 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.

B.2 Sample selection

In this subsection I show that sample selection plays little role in computing both the concentration and persistence of unemployment. I compare the statistics computed in section 2 with the same statistics[25] computed using the whole sample of workers by education and gender.

	Only HS	Whole Sample
Prime-age unemployment Avg. % time in U	3.6	3.4
Avg. % time in U, exclud. top 10%	1.6	1.4
Avg. % time in U, exclud. top 20%	0.6	0.6
Persistence Prob. top 10 prime-age given top 10 young	41.18	34.06
Prob. top 10 prime-age given Rest young	6.47	7.32
Prob. top 20 prime-age given top 20 young	44.88	43.07
Prob. top 20 prime-age given Rest young	13.71	14.23

Table 10: **Participation-adjusted** averages computed on NLSY/79, individuals aged 35-55. Left column: only high school educated, male individuals with more than 100 observations of weekly job histories in their prime-age, ending 2010. Right column: whole cross-sectional sample of NLSY/79, satisfying the same restriction on weekly job histories.

I also investigate whether concentration and persistence of unemployment, as well as dif-

[25]In this case, the average time in unemployment during prime-age is the *participation-adjusted* average.

ferentials in job-finding rates and job-separation rates I document, vary significantly across education subgroups. I keep only males and divide the NLSY/79 into high-school dropouts, high-school educated and some-college and above (those who took some college courses but did not complete college, and college-educated). Results are summarized in table 11. For all subgroups, all facts stand. Unemployment is more concentrated than what the Mortensen-Pissarides standard model implies; high young unemployment predicts high prime-age unemployment; and inequality is more due to heterogeneity in job-separation rates than job-finding rates.

	Dropouts	High-School	$\geq$ Some College
% U Accounted for by top 10	47	59	70
Predicted by MP	23	24	44
% U Accounted for by top 20	69	83	92
Predicted by MP	38	39	70
Persistence: prob. of top 10 prime-age			
from top 10 young	26	41	22
from rest young	7	6	7
Avg. % time in unemployment			
top 10% prime-age	53	29	15
Rest	4.2	1.5	0.5
δ: Prob. of U $\rightarrow$ E (monthly%)			
top 10% prime-age	4	8	11
Rest	18	26	35
f: Prob. of E $\rightarrow$ U (monthly%)			
top 10% prime-age	5.0	3.5	2.0
Rest	0.8	0.4	0.2
Predicted % time in U of top 10%, δ alone:	21	12	5
Predicted % time in U of top 10%, f alone:	16	5	2

Table 11: Summary statistics by parts of the prime-age (35-55) unemployment distribution and by education subgroups. Source: own calculations on NLSY/79. Predicted % time in U calculated using the formula $u = \delta/(\delta + f)$.

B.3 Measurement Error

When computing lifetime unemployment statistics, it is crucial to have enough observations per individual. If an individual had been observed only for few weeks, and was always unemployed, taking the average over those weeks would incorrectly attribute a lifetime unemployment of 100% to that individual. To address the extent of measurement error, I compute the concentration of unemployment in top 10 %, top 20 %, and the persistence of unemployment for different values of the lower bound of weeks of reported employment/unemployment, both when 20-30 and when 35-55. Results are reported only for the high-school subsample. Although measures of persistence and unemployment tend to fall, because the most unemployed also tend to stay out of the labor force more often, results on concentration are substantially unchanged, and persistence remains high: in the worst-case scenario in which the sample is required to have at least 500 weeks of reported employment/unemployment both when young and when prime-age (totaling 1000 weeks over 1560 maximum weeks available), the top 10% most unemployed when young still have between 4 and 5 times the likelihood of being the most unemployed when prime-age, and the most prime-age unemployed still account for two-thirds of unemployment.

	100 weeks (baseline)	300 weeks	500 weeks
Prime-age U	3.6	3.3	2.6
Prime-age U, without top 10%	1.5	1.3	0.9
Prime-age U, without top 20%	0.6	0.5	0.3
Persistence (top 10 - top 10)	41%	35%	29%
(rest - top 10)	6%	7%	8%
Persistence (top 20 - top 20)	45%	41%	35%
(rest - top 20)	14%	14%	16%
N. Individuals	1029	918	633

Table 12: Accounting for possible measurement error: concentration and persistence of unemployment according to alternative definitions of the sample. High school males with at least 100 weeks (column 1), 300 weeks (2), 500 weeks (3) of reported employment/unemployment. Source: own calculations on NLSY/79.

B.4 The role of Occupations and Health

One might think that differences across occupations are behind the strong young-prime-age correlations found in the data. For instance, the choice of a "bad occupation" when young might lead a worker to experience high unemployment both when young and in the future. I show that occupations explain relatively little of the observed young-old persistence by augmenting previous regressions with occupational controls (table 13). I use the CENSUS 1970 classification at the major category level, and I control both for the most prevalent occupation between 1979 and 2001 (that is, the occupation in which the individual worked the most during those years) and for the occupation in 1990. Sample size diminishes because occupation codes are not always available for workers in the NLSY/79; however, the strong predictive power of young unemployment remains substantially unchanged. Occupation in 1990 appears to be the most important correlate variable, diminishing the young-prime-age persistence of unemployment by 0.06. However, one must consider that occupational choice in 1990 is not independent on past labor market history, and its relevance is likely to be upward-biased because of reverse causality.

I finally consider whether the deterioration of health correlates with prime-age unemployment. I use ex-post health (in 2006) as a control as to construct a worst-case scenario: since health in 2006 can be the result of past unemployment, it will partly correlate with young unemployment and prime-age unemployment, thus in principle lowering the estimate of the impact of young unemployment. I find that, although health correlates with prime-age unemployment, it has a negligible influence on the predictive power of young unemployment.

	(1)	(2)	(3)	(4)	(5)
	Base	Educ	Educ+Occ	Educ+Occ 2	Educ+Occ 2+Health
% U when Young (20-30)	0.299***	0.256***	0.256***	0.254***	0.247***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age	-0.000262	-0.000233	-0.000144	-0.000159	-0.000195
Female	0.00128	0.00211	0.000769	0.000910	-0.000363
Ethnic Group					
Black	0.0204*	0.0170*	0.0160*	0.0165*	0.0145*
Hispanic	-0.000845	0.00827	0.00744	0.00760	0.00623
Marital Status					
Married		-0.0340***	-0.0334***	-0.0332***	-0.0319***
Separated		-0.0139	-0.0130	-0.0132	-0.0141
Divorced		-0.0119*	-0.0111*	-0.0114*	-0.0119*
Widowed		-0.00665	-0.00637	-0.00631	-0.00508
Education, age 30					
Some College		0.00770	0.00477	0.00505	0.00429
High School		0.00929*	0.00578	0.00564	0.00289
Dropout		0.0295***	0.0268***	0.0265***	0.0230**
AFQT Quartile					
2		-0.00499	-0.00562	-0.00518	-0.00323
3		-0.0106*	-0.0108*	-0.00999	-0.00688
4		-0.0128*	-0.0122*	-0.0113	-0.00896
Health 2000					
Very Good					-0.00167
Good					0.00367
Fair					0.0196**
Poor					0.0685***
Constant	0.0403	0.0440	0.0369	0.0371	0.0373
Standard Controls	X	X	X	X	X
Education, AFQT and MaStat		X	X	X	X
Prevalent Occ. (1 digit)			X	X	X
Occupation in 1990 (1d)				X	X
Health Status (2000)					X
N	3896	3896	3896	3896	3896
R^2	0.151	0.179	0.183	0.185	0.195

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: Source: own calculations on NLSY/79. Complementary regressions of % of prime-age unemployment on % of young unemployment for all workers. Sample restricted to individuals for which all controls are available for all models. Controls always include sex, ethnic group and age in 2010. (2) adds AFQT test score quartile, education and marital status, (3) adds prevalent occupation during working life dummies, (4) adds occupation in 1990 dummies, (5) adds health status in 2000 dummies. Omitted categories: male, white, never married, college-educated, 1st quartile AFQT, Technical/professional occupations, Excellent Health. Occupation coefficients and standard errors (with the exception of young unemployment) are not reported for reading convenience.

	Rest	Top 10% (35-55)	Ratio Top 10 / Rest
When young (20-30):			
Fired	0.06	0.14	2.33
Involuntary	0.43	0.78	1.8
Quit to Look	0.4	0.6	1.5
In Primes (35-55):			
Fired	0.02	0.1	5
Involuntary	0.09	0.4	4.44
Quit to Look	0.02	0.09	4.5

Table 14: 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.

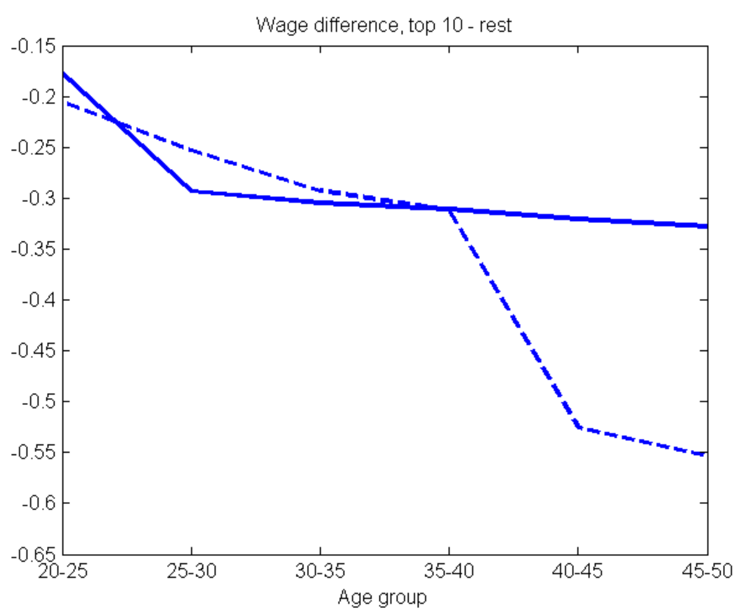

Figure 9: Difference in wages between top 10 % prime-age unemployed and rest; data (dashed) versus model (continuous) under baseline calibration.

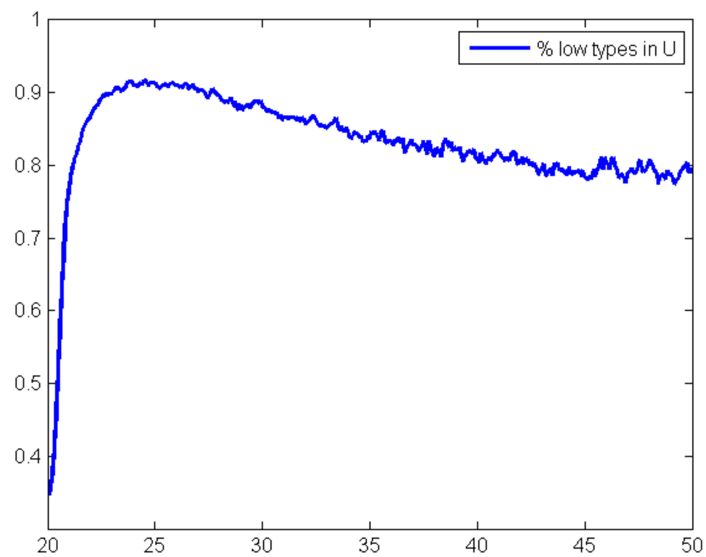

Figure 10: Share of workers who are low types, by age; under baseline calibration.